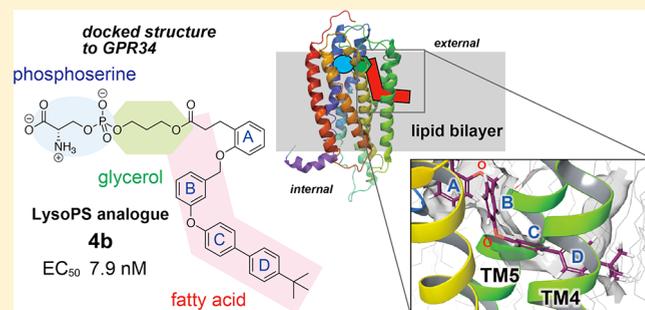


Probing the Hydrophobic Binding Pocket of G-Protein-Coupled Lysophosphatidylserine Receptor GPR34/LPS₁ by Docking-Aided Structure–Activity AnalysisMisa Sayama,^{†,∇} Asuka Inoue,^{‡,⊥,∇} Sho Nakamura,[†] Sejin Jung,[†] Masaya Ikubo,[†] Yuko Otani,[†] Akiharu Uwamizu,[‡] Takayuki Kishi,[‡] Kumiko Makide,^{‡,⊥} Junken Aoki,^{*,‡,‡,‡} Takatsugu Hirokawa,^{*,§,||} and Tomohiko Ohwada^{*,†,Ⓜ}[†]Laboratory of Organic and Medicinal Chemistry, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan[‡]Laboratory of Molecular and Cellular Biochemistry, Graduate School of Pharmaceutical Sciences, Tohoku University, 6-3, Aoba, Aramaki, Aoba-ku, Sendai, Miyagi 980-8578, Japan[§]Molecular Profiling Research Center of Drug Discovery (molprof), National Institute of Advanced Industrial Science and Technology (AIST), 2-3-26 Aomi, Koto-ku, Tokyo 135-0064, Japan^{||}Division of Biomedical Science, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki 305-8575, Japan[⊥]PRESTO, Japan Science and Technology Agency (JST), 4-1-8, Honcho, Kawaguchi-shi, Saitama 332-0012, Japan[#]AMED-CREST, Japan Agency for Medical Research and Development, 1-7-1 Otemachi, Chiyoda-ku, Tokyo 100-0004, Japan

Supporting Information

ABSTRACT: The ligands of certain G-protein-coupled receptors (GPCRs) have been identified as endogenous lipids, such as lysophosphatidylserine (LysoPS). Here, we analyzed the molecular basis of the structure–activity relationship of ligands of GPR34, one of the LysoPS receptor subtypes, focusing on recognition of the long-chain fatty acid moiety by the hydrophobic pocket. By introducing benzene ring(s) into the fatty acid moiety of 2-*deoxy*-LysoPS, we explored the binding site's preference for the hydrophobic shape. A tribenzene-containing fatty acid surrogate with modifications of the terminal aromatic moiety showed potent agonistic activity toward GPR34. Computational docking of these derivatives with a homology modeling/molecular dynamics-based virtual binding site of GPR34 indicated that a kink in the benzene-based lipid surrogates matches the L-shaped hydrophobic pocket of GPR34. A tetrabenzene-based lipid analogue bearing a bulky *tert*-butyl group at the 4-position of the terminal benzene ring exhibited potent GPR34 agonistic activity, validating the present hydrophobic binding pocket model.



INTRODUCTION

G-protein-coupled receptors (GPCRs) are membrane receptors containing seven transmembrane helix (TM) domains and are considered major therapeutic targets. During the past decade, a number of GPCRs have been orphaned, and their ligands have been identified as endogenous lipids. Among them, GPR34/LPS₁^{1,2} is specifically activated by lysophosphatidylserine (LysoPS), a lysophospholipid generated by enzymatic hydrolysis of a component of membrane bilayers, phosphatidylserine (PS).³ GPR34 is highly expressed in immune cells, such as mast cells, macrophages, and microglia.¹ GPR34-deficient mice show a defective immune response after pathogen challenge⁴ and exhibit significant changes in microglial morphology and phagocytosis.⁵ GPR34 also plays a role in the progression of some tumors, such as human gastric

adenocarcinoma,^{6,7} mucosa-associated lymphoid tissue (MALT) lymphoma,⁸ and colorectal cancer.⁹

However, although GPR34 is related to immune pathology, the underlying mechanisms remain unknown, and chemical tools, such as potent agonists, antagonists, or inverse agonists of GPR34, are needed to investigate these issues. We have designed and synthesized several LysoPS analogues that exhibit agonist activity toward GPR34.^{10–13} LysoPS is an amphipathic molecule that has a phosphoserine as a polar headgroup linked via a glycerol moiety to a fatty acid (e.g., oleic acid (18:1)) as the hydrophobic tail. Our previous structure–activity relationship (SAR) study of LysoPS analogues suggested that the phosphoserine headgroup is necessary for GPR34-activating

Received: May 10, 2017

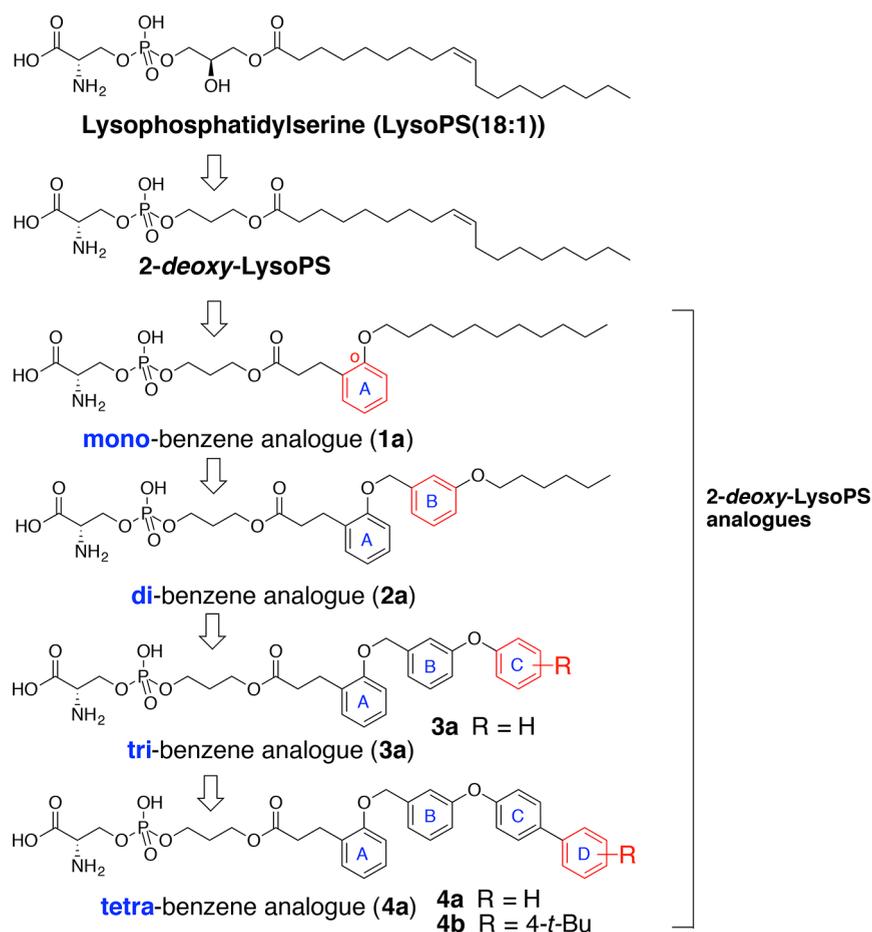


Figure 1. Structural development of lysophosphatidylserine and its 2-deoxy derivatives.

potency.¹² On the other hand, the length and the shape (i.e., geometry and position of the olefin) of the fatty acid moiety also seemed to contribute to receptor activation.¹² However, from the viewpoint of ligand–receptor recognition, the interaction of such long-chain fatty acids with the hydrophobic pocket of GPCRs, including GPR34, has been little studied.

Therefore, in this work, we set out to investigate the molecular basis of the dependency of GPR34 activity on the structure of the fatty acid moiety of LysoPS analogues. We focused on the GPR34 agonist activity of 2-deoxy-LysoPS analogues, which lack the *sn*-2 hydroxyl group of the glycerol moiety, and explored the preferred shape of the hydrophobic fatty acid moiety by introducing benzene ring(s) to restrict the molecular conformation. Our findings indicated that GPR34-activating potency depends strongly on the configuration and substitution of aromatic rings in these fatty acid surrogates. On the basis of the obtained SAR, we probed the hydrophobic binding pocket of GPR34 by means of computational docking of these synthesized compounds with a homology modeling/molecular dynamics-based virtual binding site of GPR34. We found that the matching/mismatching of ligand analogues with a putative L-shaped binding pocket between TM4 and TM5 was consistent with the observed SAR. The *cis*-olefin moiety of the endogenous oleoyl group (18:1) also fit well to the L-shaped binding pocket. To test the validity of this binding model, we used it to design a tetrabenzene-based lipid analogue bearing a bulky *tert*-butyl group at the 4-position of the terminal benzene ring. We synthesized this compound and confirmed that it showed potent GPR34 agonistic activity.

RESULTS AND DISCUSSION

Tactics for Designing LysoPS Analogues. In a previous study, we had found that analogues of the fatty acid moiety with an ortho-substituted benzene ring (ring A) had greater activity than meta- or para-substituted analogues (Figure 1).¹² Accordingly, in this study, we aimed to elaborate ortho-substituted analogues by expanding the hydrophobic fatty acid moiety of 2-deoxy-LysoPS analogues (1a → 2a → 3a → 4a) through introduction of additional benzene ring(s) (Figure 1). SAR analysis of the synthesized compounds indicated that GPR34-activating potency depends on the configuration and substitution of aromatic rings in these fatty acid surrogates. In particular, the tribenzene-containing fatty acid surrogate (3a) showed moderate agonistic activity toward GPR34 (Table 1). We further explored the SAR, focusing on the terminal moiety of the fatty acid surrogate, by means of computational docking of these compounds with a homology modeling/molecular dynamics-based virtual binding site of GPR34. To test the validity of the binding models, we used the virtual binding site to design a tetrabenzene compound bearing a bulky *tert*-butyl group at the 4-position of the terminal benzene ring D (4b) as a candidate agonist, which we synthesized and tested (Figure 1). The chemical structures of all the compounds synthesized in this work are compiled in Chart S1.

Synthesis of LysoPS Analogues. We previously established synthetic routes to 2-deoxy-LysoPS analogues by means of benzyl protection–hydrogenative debenzylation¹⁰ and Boc protection–acid-catalyzed deprotection procedures.¹² Because

Table 1. SAR of 2-Deoxy-LysoPS Analogues Containing Fatty Acid Surrogates^a

R group	Mouse GPR34	R group	Mouse GPR34
	EC50 (LogEC50) [Emax] <RIA> n =		EC50 (LogEC50) [Emax] <RIA> n =
Positive Control	180 nM (-6.74 ± 0.08) [21.3 ± 2.1%] <1> 5	18:1-LysoPS	420 nM (-6.38 ± 0.09) [14.0 ± 1.0%] <0.26 ± 0.05> 7
2-deoxy-LysoPS	> 3 μM (> -5.5) [NA] <NA> 3	3a	220 nM (-6.66 ± 0.15) [19.4 ± 4.2%] <0.48 ± 0.25> 3
1a	1.6 μM (-5.79 ± 0.09) [14.6 ± 1.7%] <0.040 ± 0.019> 3	3b	> 3 μM (> -5.5) [NA] <NA> 3
1b	> 3 μM (> -5.5) [NA] <NA> 3	3c	> 3 μM (> -5.5) [NA] <NA> 3
2a	> 3 μM (> -5.5) [NA] <NA> 3	3d	1.3 μM (-5.89 ± 0.17) [14.5 ± 5.4%] <0.077 ± 0.038> 3
2b	1.7 μM (-5.76 ± 0.34) [6.1 ± 2.4%] <0.033 ± 0.027> 3	3e	800 nM (-6.10 ± 0.05) [18.8 ± 2.5%] <0.11 ± 0.05> 3
2c	1.2 μM (-5.91 ± 0.07) [20.7 ± 4.0%] <0.071 ± 0.035> 3		

^aActivities are represented in terms of EC₅₀. LogEC₅₀ (M) and E_{max} (% AP-TGFα release) values are calculated from a sigmoidal concentration–response curve and shown as mean ± SEM of indicated numbers of independent experiments (*n*). EC₅₀ values are calculated from a mean value of LogEC₅₀. RIA (relative intrinsic activity, a dimensionless parameter; ref 49) is an estimation of agonist activity and represented as an E_{max}/EC₅₀ value relative to that of a positive control compound (compound 6 in the previous work, ref 13). By definition, the RIA value of a positive control compound is equal to 1. “NA” means “not available” because of very low activity.

the target fatty acid surrogates in the present work contain a benzyl position sensitive to hydrogenation, we chose the Boc-protection–acid-catalyzed deprotection procedure. The syntheses of 2-deoxy-LysoPS analogues with fatty acid surrogates (4a–4f) are summarized in Scheme 1.

Biarylethers (5a and 5b) were constructed by means of Chan–Evans–Lam coupling.^{14,15} Reduction of the formyl group, followed by condensation with the 3-phenylpropionic acid methyl ester (7) by means of Mitsunobu reaction afforded the methyl ester intermediates (8a and 8b) of aromatic fatty acid surrogates. Suzuki–Miyaura coupling afforded compounds containing an additional benzene ring (9a–9f). Basic hydrolysis of the methyl ester (9a–9f) furnished the desired carboxylic acids (10a–10f). Various fatty acid surrogates were synthesized and coupled to the polar moiety intermediate 11, which was synthesized through phosphoramidite coupling of the protected phosphoserine part and 1,3-propanediol as previously described¹² to furnish protected 2-deoxy-LysoPS analogues

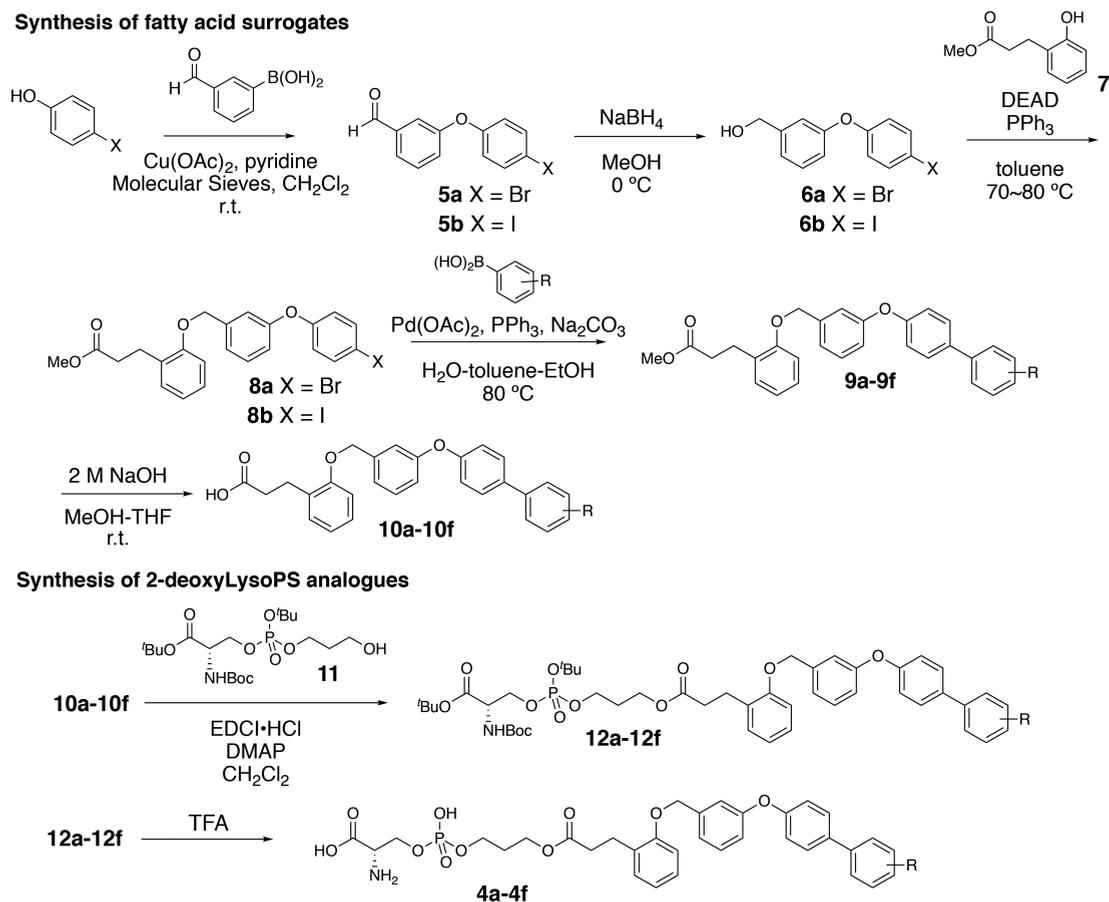
(12a–12f). Global deprotection with TFA furnished the 2-deoxy-LysoPS analogues as the TFA salts (4a–4f).

Details of the synthesis of the acyl moiety of other compounds (13–34) are given in the Supporting Information.

Activity of Fatty Acid Surrogate-Containing LysoPS Analogues toward GPR34. The activities of the synthesized compounds toward GPR34 were determined by means of TGFα shedding assay,¹⁶ as described in the Experimental Section. We principally used mouse receptors in the cell-based assay because of scalability to in vivo studies in mice in the future (Tables 1–3). For selected analogues (3a, 3f, 3g, 3h, 3v, 4a, 4c, and 4e), we also examined the response to the corresponding human GPR34, and we confirmed that the response of LysoPS analogues to human GPR34 matched well to that of mouse GPR34 (Table 4). We consider the SAR for mouse GPR34 activity in the following discussion.

When we introduced another benzene ring (ring B) near ring A of our previously synthesized compound 1a, the resulting compound 2a (Figure 1) exhibited very weak potency toward

Scheme 1. General Synthetic Routes to 2-Deoxy-LysoPS Analogues



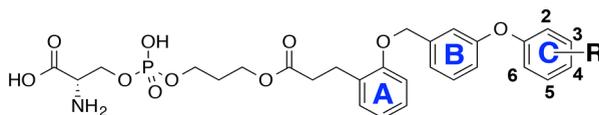
GPR34 ($EC_{50} \geq 3 \mu\text{M}$; Table 1 and Figures S1a and S2b). When the second benzene moiety of 2a was replaced with a triple bond (1b), the agonist activity was also very weak ($EC_{50} \geq 3 \mu\text{M}$; for structures, see Table 1, Chart S1, and Figures S1a and S2a). The second benzene ring inserted into the terminal moiety of 1a did not greatly increase the GPR34-activating potency (2b, $EC_{50} = 1.7 \mu\text{M}$; 2c, $EC_{50} = 1.2 \mu\text{M}$; Table 1 and Figure S2b). When we introduced a third benzene moiety (ring C) into the remaining flexible alkyl chain moiety of 2a, affording analogues 3a–3e (for structures, see Table 1 and Chart S1), compound 3a was as active as LysoPS ($EC_{50} = 220 \text{ nM}$; Table 1 and Figures S1b and S2c). However, a slight change in the configuration of ring C abrogated the GPR34 activity (3b, $EC_{50} \geq 3 \mu\text{M}$; 3c, $EC_{50} \geq 3 \mu\text{M}$; 3d, $EC_{50} = 1.3 \mu\text{M}$; Table 1 and Figures S1b and S2c). Therefore, an appropriate configuration of ring C is important for GPR34 activation. The activity of saturated cyclohexyl-substituted analogue 3e was higher than those of compounds 3b–3d (3e: $EC_{50} = 800 \text{ nM}$; Table 1 and Figures S1b and S2c), suggesting that the presence of a hydrophobic moiety at the terminal position is important for GPR34 activation and also that π interaction is involved in 3a.

We also synthesized 2-deoxy-LysoPS analogues linked at the meta- and para-positions of ring A of 1a. These meta-substituted analogues (1c, 2d–2f, 3w, and 3x) and para-substituted analogue (3y) lacked activity toward GPR34 ($EC_{50} \geq 3 \mu\text{M}$; Table S1 and Figure S2a–c). These results are similar to the SARs found previously for LysoPS derivatives.^{12,13}

Substituent Effects at the Terminal Benzene Ring on GPR34-Activating Activity. Because 3a showed more potent

agonistic activity than previously found 1a toward GPR34, we focused on derivatization of 3a to further characterize the preferred environment around ring C. Various substituents were introduced into ring C (3f–3v; for structures and activities, see Table 2 and Figure 2a and b).

Introduction of a methyl group at the 2- (3f), 3- (3g), or 4-position (3h) changed the GPR34-activating potency in different ways: the 4-Me group markedly increased the activity (3h, $EC_{50} = 39 \text{ nM}$), whereas the 2-Me or 3-Me groups decreased the activity (3f, $EC_{50} = 650 \text{ nM}$; 3g, $EC_{50} = 550 \text{ nM}$) compared with unsubstituted 3a ($EC_{50} = 220 \text{ nM}$) (Table 2, Figure 2a, and Figure S2d). Introduction of a chloro group at the 2-, 3-, or 4-position of ring C showed a similar trend: the 3-Cl or 4-Cl group showed a similar or slightly increased activity compared with that of 3a (3j, $EC_{50} = 170 \text{ nM}$; 3k, $EC_{50} = 120 \text{ nM}$), whereas a 2-Cl group diminished the activity (3i, $EC_{50} = 1.2 \mu\text{M}$). 3,4-Dichloro substitution caused synergistic activation of GPR34 (3l, $EC_{50} = 110 \text{ nM}$) (Table 2 and Figures S1c and S2d). The results for these small probes suggest that ring C is accommodated in the receptor pocket and that substitution at the 4-position can be accommodated but that substitution at the 2-position results in steric conflict. Other substituents at the 4-position, such as fluoro (3m, $EC_{50} = 120 \text{ nM}$), bromo (3n, $EC_{50} = 82 \text{ nM}$), iodo (3o, $EC_{50} = 100 \text{ nM}$), ethyl (3p, $EC_{50} = 180 \text{ nM}$), and methoxy (3q, $EC_{50} = 260 \text{ nM}$) resulted in similar or greater GPR34-activating potency to the unsubstituted parent compound (3a, $EC_{50} = 220 \text{ nM}$), but they were less effective than 4-methyl substitution (3h, $EC_{50} = 39 \text{ nM}$) (Table 2 and Figures S1d and S2d). A hydrophilic cyano or amino group at the 4-position diminished the activity toward GPR34

Table 2. SAR of 3a Derivatives with Three Benzene Rings^a

R group	Mouse GPR34	R group	Mouse GPR34	R group	Mouse GPR34
	EC50 (LogEC50) [E _{max}] <RIA> n =		EC50 (LogEC50) [E _{max}] <RIA> n =		EC50 (LogEC50) [E _{max}] <RIA> n =
3a	H 220 nM (-6.66 ± 0.15) [19.4 ± 4.2%] <0.48 ± 0.25> 3	3l	3,4-diCl 110 nM (-6.98 ± 0.03) [18.0 ± 1.4%] <0.71 ± 0.37> 3	3r	4-CN 2.0 μM (-5.71 ± 0.13) [18.3 ± 2.5%] <0.041 ± 0.021> 3
3f	2-Me 650 nM (-6.19 ± 0.08) [20.3 ± 1.1%] <0.28 ± 0.07> 5	3m	4-F 120 nM (-6.93 ± 0.19) [18.7 ± 2.2%] <0.66 ± 0.38> 3	3s	4-NH ₂ 2.7 μM (-5.57 ± 0.03) [21.3 ± 1.8%] <0.089 ± 0.010> 4
3g	3-Me 550 nM (-6.26 ± 0.08) [21.3 ± 1.8%] <0.29 ± 0.06> 6	3n	4-Br 82 nM (-7.09 ± 0.01) [17.7 ± 2.1%] <0.97 ± 0.56> 3	3t	2- <i>t</i> -Bu > 3 μM (>-5.5) [NA] <NA> 3
3h	4-Me 39 nM (-7.41 ± 0.08) [18.2 ± 1.5%] <3.7 ± 0.9> 6	3o	4-I 100 nM (-6.99 ± 0.06) [19.1 ± 2.6%] <0.82 ± 0.41> 3	3u	3- <i>t</i> -Bu 210 nM (-6.67 ± 0.11) [19.9 ± 2.8%] <0.39 ± 0.2> 3
3i	2-Cl 1.2 μM (-5.91 ± 0.04) [19.8 ± 2.4%] <0.069 ± 0.038> 3	3p	4-Et 180 nM (-6.75 ± 0.11) [19.3 ± 3.9%] <0.42 ± 0.21> 3	3v	4- <i>t</i> -Bu 900 nM (-6.04 ± 0.16) [16.9 ± 3.2%] <0.079 ± 0.039> 3
3j	3-Cl 170 nM (-6.77 ± 0.12) [17.1 ± 2.0%] <0.53 ± 0.26> 3	3q	4-OMe 260 nM (-6.59 ± 0.04) [16.6 ± 2.2%] <0.28 ± 0.15> 3	4a	4-Ph 45 nM (-7.34 ± 0.06) [17.8 ± 3.9%] <2.2 ± 1.5> 3
3k	4-Cl 120 nM (-6.94 ± 0.11) [17.8 ± 1.5%] <0.76 ± 0.38> 3				

^aActivities are represented in terms of EC₅₀. LogEC₅₀ (M) and E_{max} (% AP-TGFα release) values are calculated from a sigmoidal concentration–response curve and shown as mean ± SEM of indicated numbers of independent experiments (n). EC₅₀ values are calculated from a mean value of LogEC₅₀. RIA (relative intrinsic activity, a dimensionless parameter; ref 49) is an estimation of agonist activity and represented as an E_{max}/EC₅₀ value relative to that of a positive control compound (previous work). By definition, the RIA value of a positive control compound is equal to 1. “NA” means “not available” because of very low activity.

(**3r** (4-CN), EC₅₀ = 2.0 μM; **3s** (4-NH₂), EC₅₀ = 2.7 μM) (Table 2 and Figure S2d). Therefore, the region of the binding pocket adjacent to the 4-position of ring C may be hydrophobic.

When a bulkier hydrophobic substituent such as a *tert*-butyl group was introduced at the 2- or 3-position, GPR34-activating potency changed in a similar manner to the case of a small

methyl or chloro group: the 2-*tert*-butyl group decreased the activity (**3t**, EC₅₀ ≥ 3 μM) and the 3-*tert*-butyl-substituted compound had almost the same activity as the unsubstituted **3a** (**3u**, EC₅₀ = 210 nM (Table 2)). On the other hand, when a *tert*-butyl group was introduced at the 4-position, GPR34-activating potency was decreased significantly as compared with that of **3a** (**3v**, EC₅₀ = 900 nM) (Table 2, Figure 2b, and

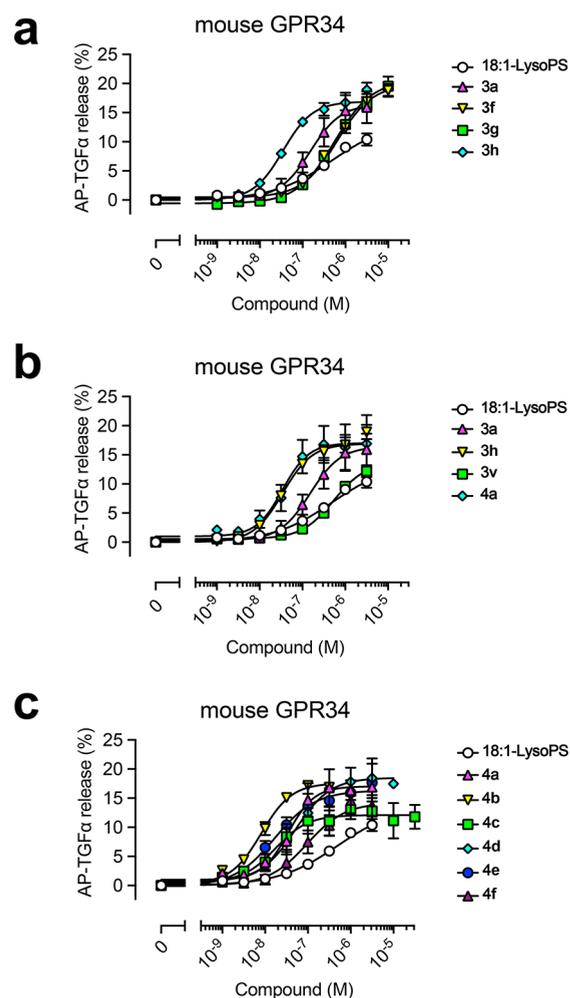


Figure 2. SAR of 2-Deoxy-LysoPS Derivatives for Mouse GPR34. HEK293 cells transiently transfected with mouse-GPR34-encoding expression vector or an empty vector were treated with compounds. Receptor-specific AP-TGF α release responses were determined by subtracting background responses in the empty vector-transfected cells. Data are mean and SEM (standard error of the mean) for 3–6 independent experiments.

Figures S1e and S2d). In contrast, when an anisotropic phenyl group was introduced, the analogue (**4a** (4-Ph), EC_{50} = 45 nM) was as active as the 4-methyl derivative (**3h** (4-Me), EC_{50} = 39 nM) (Figure 1, Table 2, Figure 2b, and Figure S2e). These observations are all consistent with the idea that ring C is recognized by GPR34 and that the binding site contains a hydrophobic pocket that can accommodate nonbulky substituents, especially at the 4-position of ring C.

Homology Modeling of GPR34 and Computational Docking with Active Compound 3a. The present SAR data showed that GPR34 strictly recognizes the fatty acid moiety of **3a**. To understand this finding in terms of receptor–ligand interactions, we aimed to explore the structure of the active site of GPR34 by means of computational docking of ligands to the protein. Because the crystal structures of LysoPS receptors have not been reported so far, we adopted a homology modeling/molecular dynamics approach to obtain the human GPR34 structure. Thus, our approach to docking compounds to the LysoPS binding sites of human GPR34 involved the following three main steps: (1) homology modeling of GPR34, (2) computational docking of compound **3a**, (3) optimization of

binding models by molecular dynamics. For prediction of the ligand–protein binding mode, our SAR information (Table 2) should be especially useful to define the hydrophobic binding pocket of the acyl moiety. Although the SAR information in this work is based on mouse GPR34, mouse and human GPR34 share high sequence identity (90%), and we confirmed that the rank-orders of activation potency of selected compounds toward mouse and human GPR34s were similar (Table 4 and Figure 3).

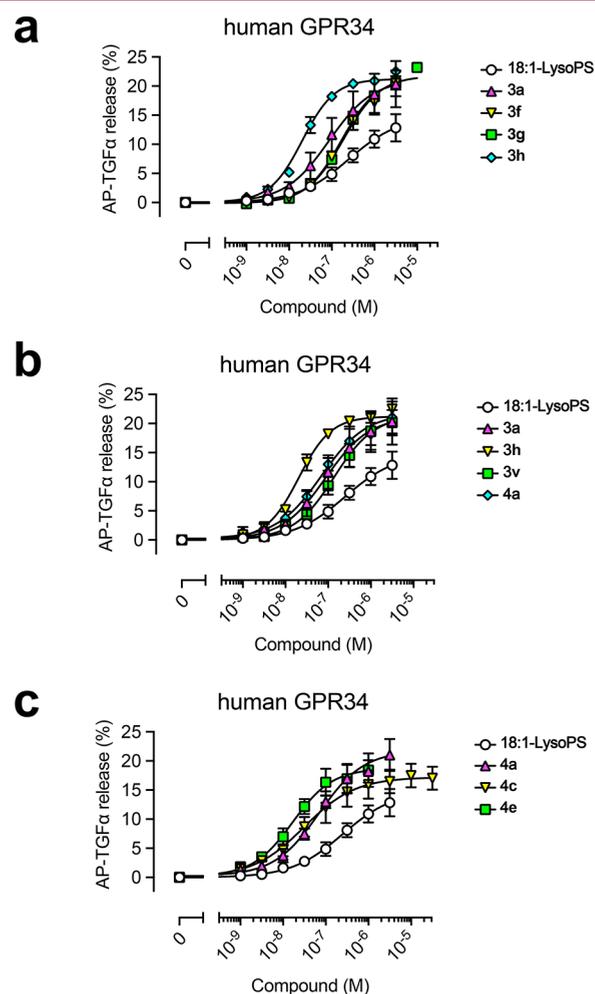


Figure 3. SAR of 2-deoxy-LysoPS derivatives for human GPR34. HEK293 cells transiently transfected with human-GPR34-encoding expression vector or an empty vector were treated with compounds. Receptor-specific AP-TGF α release responses were determined by subtracting background responses in the empty vector-transfected cells. Data are mean and SEM (standard error of the mean) for 3 or 4 independent experiments.

Homology Modeling of GPR34. GPR34 belongs to the P2Y family, which in turn belongs to the δ group of class A GPCRs.¹⁷ So far, crystal structures of four members of the δ group GPCRs, PAR1,¹⁸ P2Y1,¹⁹ P2Y12,^{20,21} and GPR40,^{22,23} have been solved. Among them, P2Y12 has the highest identity with human GPR34 (26% identity in terms of the full-length protein; 30% identity in terms of the modeling region after truncation of the N- and C-terminals; Figure S3). Thus, we chose P2Y12 (PDB ID: 4PXZ) in the agonist-bound state as a template for GPR34 homology modeling.

Computational Docking of Compound 3a. We obtained 19 binding models in total by docking ionized 3a to three homology models of human GPR34 with the induced fit docking protocol^{24,25} followed by a 10 ns molecular dynamics (MD) simulation, as described in the [Experimental Section](#). In all 19 models, the polar headgroup of the serine-phosphate moiety of ligand 3a was located close to the exterior membrane surface, and the hydrophobic fatty acid tail was located deep inside the receptor pocket. This orientation is similar to that of the original ligand in the P2Y12 crystal structure, 2-methylthioadenosine-5'-diphosphate (2MeSADP; for the chemical structure, see [Chart S2](#)), which binds with its hydrophilic phosphate group directed to the extracellular region and its hydrophobic 2-methylthioadenine moiety embedded inside the receptor.²¹ The geometrical orientation of 3a was also similar to that of the complex of another lysophospholipid receptor, S1P₁, with its antagonist.²⁶

Optimization of Binding Models by Molecular Dynamics. On the basis of the ligand RMSD values, the 19 binding poses of 3a were clustered into five groups. Two of them represent major configurations, consisting of 10 (cluster 1) and 6 poses (cluster 2), respectively ([Table S2](#)). Representative binding models (top 30% of cluster members) having the best MM-GBSA ΔG_{bind} values were chosen from clusters 1 and 2 (designated as models 1-1, 1-2, and 1-3 (cluster 1) and models 2-1 and 2-2 (cluster 2)) (see [Figure 4](#)). These models were subjected to a 100 ns MD simulation in an explicit membrane environment to confirm their stability.²⁷ [Figure 4](#) shows the results for the five models after this calculation. The ligands did not change their position drastically in the binding pocket during the 100 ns MD simulation ([Figure S4](#) and [Table S2](#)). In models 1-1, 1-2, and 1-3 ([Figure 4a–c](#), left), the hydrophobic tail of 3a (blue balls) is surrounded by TM3, TM4, and TM5, whereas in models 2-1 and 2-2 ([Figures 4d](#) and [e](#), left), the hydrophobic tail of 3a (blue balls) is surrounded by TM3, TM5, and TM6. In particular, the orientation of ring C of 3a is different between models 1-1, 1-2, and 1-3 and models 2-1 and 2-2. On the other hand, the polar headgroup of 3a (orange balls, except ring A) occupies a similar position in all models. An interaction diagram of model 1-1 is shown in [Figure S5](#) as a representative example.

Selection and Evaluation of 3a Bound GPR34 Models Based on the SAR Data. We examined the compatibility of five models with the experimental SAR data for analogues of 3a.

1. Comparison of the Binding Pocket Environment around Ring C of 3a. We characterized the shape and physicochemical features of the receptor-binding pocket at which the acyl moiety of 3a interacts by means of SiteMap analysis.²⁸ In all of the representative models, there is a hydrophobic pocket around ring C, as shown in [Figure 4](#) (right). Ring C can interact hydrophobically with aromatic amino-acid side chain(s) in models 1-1 (Tyr139^{3,37} and Phe219^{5,39}), 1-2 (Tyr139^{3,37} and Phe219^{5,39}), and 2-1 (Tyr135^{3,33}) (Ballesteros–Weinstein nomenclature is also shown).²⁹ In particular, in models 1-1 and 1-2, the L-shaped hydrophobic pocket fits well to the ligand kink created by ether oxygen between rings B and C. Furthermore, this L-shape seems important for strict recognition of ring C because derivative 3b, which is closely related to 3a, lacked activity toward GPR34; because rings B and C of 3b are directly connected, this molecule would not fit the L-shaped pocket of the receptor, impairing the hydrophobic interactions between ring C and Tyr and Phe residues. These considerations are

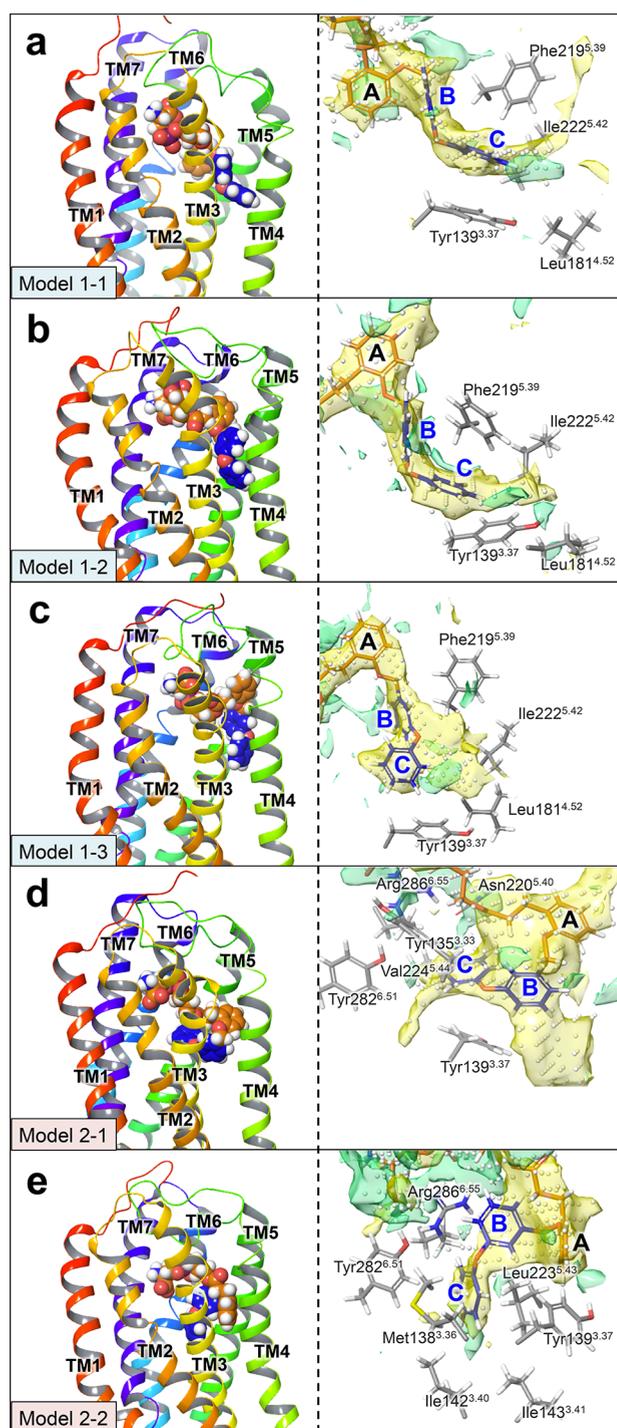


Figure 4. Five representative binding models from clusters 1 and 2 (a, model 1-1; b, model 1-2; c, model 1-3; d, model 2-1; e, model 2-2). Each left figure presents the side view of the binding model of ligand 3a in complex with the human GPR34 model (ribbon). The extracellular half of the receptor is shown. Each right figure presents a close-up view of the acyl moiety of 3a (stick) in the receptor binding site, calculated using SiteMap, and the receptor side chains near the ligand terminal benzene ring (ring C). Rings B and C are colored blue, and the other part of 3a is colored orange. The hydrophobic pocket is shown as a yellow surface, and the hydrophilic pocket is shown as a green surface. White dots represent the grid where placement of ligand atoms is allowed in the SiteMap analysis. Leu181 and Ile222 in human GPR34 are mutated to Val in mouse GPR34. As starting points for molecular modeling of GPR34, the crystal structure of P2Y12 as the agonist bound state (PDB ID: 4PXZ) was used.

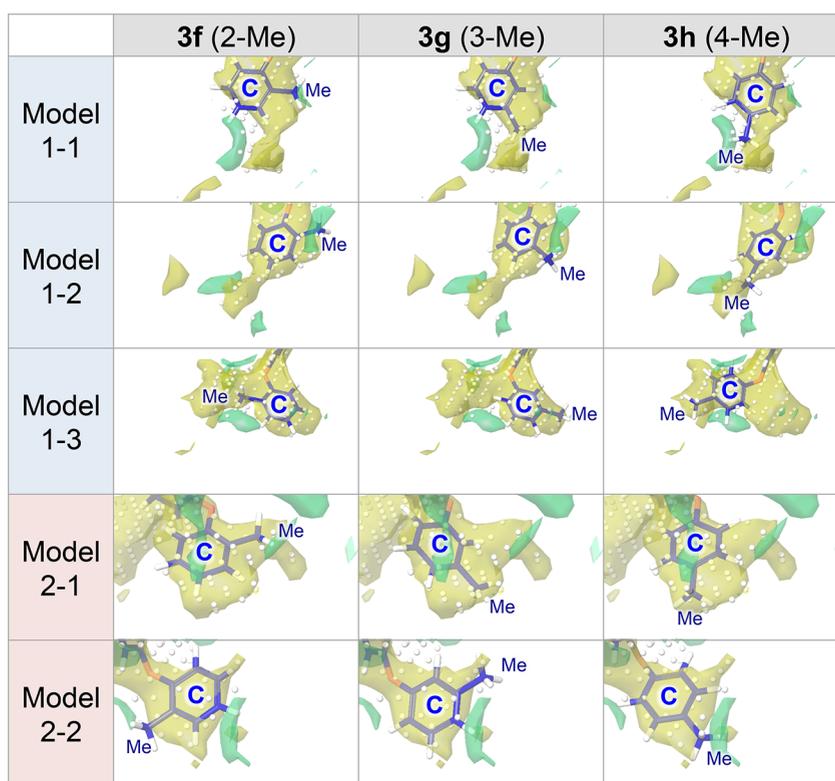


Figure 5. SiteMap analysis of methyl substitution of ring C in the five representative models. Only the regions accommodating ring C of ligands 3f–3h are shown. Ligands 3f–3h are shown as blue sticks. The hydrophobic pocket is shown as a yellow surface, and the hydrophilic pocket is shown as a green surface. White dots represent the grid where placement of ligand atoms is favorable in SiteMap analysis.

consistent with the SAR-based assessment that ring C of 3a is strictly recognized by the receptor (Table 1).

2. Influence of Methyl Substitution Position on Ring C. The observed substituent effects on ring C of 3a suggested that substitution of a methyl or a chloro group at the 4-position on ring C is favored over that at the 2- or 3-position for GPR34 activation (3f (2-Me), $EC_{50} = 180$ nM; 3g (3-Me), $EC_{50} = 230$ nM; 3h (4-Me), $EC_{50} = 27$ nM for human GPR34; Table 4, Figure 3a, and Figure S2f). Methyl-substituted compounds 3f–3h were docked to the five representative receptor models shown in Figure 4 using the Glide docking program^{30,31} on the assumption that the location of the hydrophilic serine-phosphate moiety in the receptor pocket is the same in each case (see Experimental Section) (Figure 5). In models 1-1 and 1-2, the 4-methyl compound (3h) occupied the hydrophobic pocket of the receptor more effectively than the 2- (3f) or 3-methyl (3g) compound in accordance with the relative GPR34-activating potencies. However, in models 1-3, 2-1, and 2-2, the methyl group in all compounds 3f–3h interacted similarly with the hydrophobic pocket of the receptor, and thus, these models cannot account for the observed preference for 4-substitution.

3. Influence of Bulkiness of the 4-Substituent on Ring C. A *tert*-butyl group at the 4-position on ring C (3v) markedly decreased the GPR34 activity as compared with a 4-methyl group (3h), whereas a phenyl group at the 4-position of ring C (4a) was as effective as the 4-methyl group (3h) for GPR34 activation (3h (4-Me), $EC_{50} = 27$ nM; 3v (4-*t*-Bu), $EC_{50} = 300$ nM; 4a (4-Ph), $EC_{50} = 78$ nM for human GPR34; Table 4, Figure 3b, and Figure S2f). These results suggested that linear and planar substitution at this position is compatible with GPR34 activity.

To examine whether each of the five models could account for the difference in potency between 3v and 4a, we examined the binding poses obtained using Glide docking software (Figure 6).^{30,31}

In models 1-1 and 1-2, 3v was successfully docked to the receptor (Figure 6a), but the *tert*-butyl group of 3v was not accommodated well in the hydrophobic pocket region (yellow) predicted by SiteMap.²⁸ This might explain why 3v showed decreased activity. In contrast, the phenyl group of docked 4a was accommodated more effectively in the hydrophobic pocket of the receptor in models 1-1 and 1-2 in accordance with the fact that 4a showed high activity. Detailed scrutiny of models 1-1 and 1-2 revealed a narrowing of the hydrophobic pocket in which the phenyl group could still be accommodated but the *tert*-butyl group would clash (Figure 6b). On the other hand, in model 1-3, the *tert*-butyl group of 3v and the phenyl group of 4a were both located outside the hydrophobic pocket region, and thus, this model cannot explain the observed difference of GPR34 activity between 3v and 4a. In model 2-1, the *tert*-butyl group of 3v is fully accommodated in the hydrophobic pocket, and thus this model cannot account for the decreased activity of 3v. Model 2-1 gave a binding pose for 4a, but it was completely different from the original binding mode of 3a in complex with receptor model 2-1 as shown in Figure 4d (3a is also shown as orange sticks in Figure 6a). Instead, the binding pose of 4a in model 2-1 was similar to those in models 1-1 and 1-2: ring C was placed between TM4 and TM5 (Figure 6a, shown as purple sticks). On the other hand, model 2-2 did not give binding poses for 3v and 4a.

These results indicate that models 1-1 and 1-2 are more compatible with the SAR data than are models 1-3, 2-1, and 2-2. Thus, we consider that an L-shaped binding pocket occupied

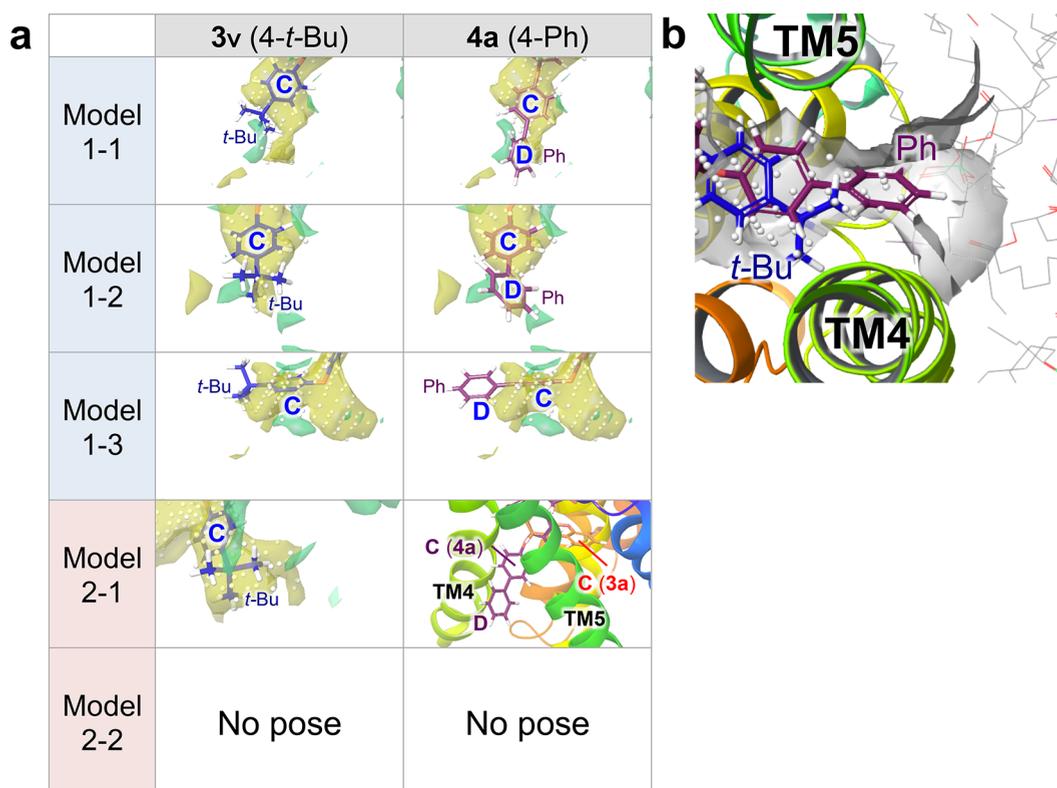


Figure 6. SiteMap analysis of the effects of 4-*tert*-butyl and 4-phenyl substitution in five representative receptor models. Only the regions accommodating ring C of ligands **3v** and **4a** are shown. Ligands **3v** and **4a** are shown in stick form (**3v**, blue; **4a**, purple). (a) The hydrophobic pocket, analyzed by SiteMap, is shown as a yellow surface, and the hydrophilic pocket is shown as a green surface. White dots represent the grid where ligand placement is favorable in SiteMap analysis. For **4a** in complex with model 2-1, the hydrophobic moiety was placed in a different pocket from the pocket for the originally bound **3a** (orange stick model) as shown in the model 2-1 column. (b) Top view of the narrowing between TM4 and TM5 in model 1-1. 1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) molecules (gray lines) of the membrane bilayer are extracted from the MD trajectory (after 100 ns) and superimposed onto the docking model. The binding pocket surface calculated using SiteMap is shown in gray. White dots represent the grid where ligand atom placement is favorable in SiteMap analysis.

by the acyl moiety of analogues of **3a** is likely to exist between TM4 and TM5.

4. Confirmation of Selected Active-Site Models. It is important to note that protein structure is flexible, and therefore, we carried out ensemble docking of **4a** (4-Ph) with multiple structures of GPR34 to confirm that the hydrophobic pocket is located between TM4 and TM5 (as in models 1-1 and 1-2) but not around TM3, TM5, and TM6 (as in models 2-1 and 2-2). Specifically, **4a** (4-Ph) was docked to ten structures of GPR34 obtained from 100 ns MD simulations of five models (ensemble docking, Table S3). Even taking protein dynamics (i.e., the flexibility of side chains of the receptor) into consideration, binding of the acyl moiety around TM3, TM5, and TM6 seems highly unlikely. These results provide further support for the location of the hydrophobic binding pocket between TM4 and TM5 as in models 1-1 and 1-2.

Design of Agonists Based on the Selected GPR34–Compound Binding Models. Finally, to confirm the validity of models 1-1 and 1-2, we used them to design a LysoPS derivative as a candidate agonist. In models 1-1 and 1-2, an L-shaped aperture exists between TM4 and TM5 of the receptor. As a result, in the docking pose of **4a**, the 4-phenyl group (ring D) is placed out of the receptor to reach the membrane region (Figure 6b). Thus, we considered that the activity would be retained even if a bulky substituent (such as a *tert*-butyl group) was attached to the terminal ring D of **4a** because the substituent would not disturb the ligand binding mode. To

examine this idea, we docked **4b** (Figure 1), which has a 4-*tert*-butyl group on ring D, into model 1-1. In the resulting binding model, aromatic ring D extends outside the receptor, and thus **4b** should be GPR34 active (Figure 7). The same rationale is applicable to **4c** (4-Ph), **4d** (4-*O*-*n*-C₆H₁₃), and **4e** (3-*t*-Bu).

Therefore, we next synthesized and evaluated derivatives (**4b**–**4e**) bearing various substituents on ring D of the

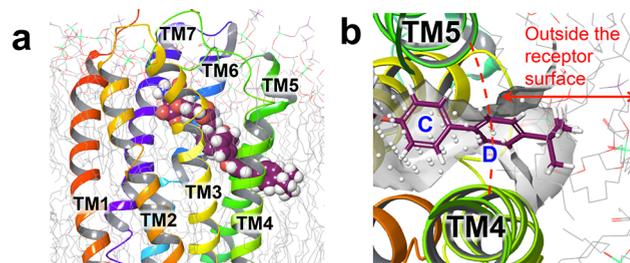
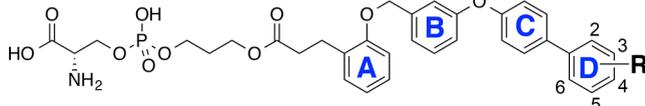


Figure 7. Docking model of tetra benzene analogue **4b** bearing a 4-*tert*-butyl group on ring D. (a) Docking model of **4b** (purple) with GPR34 model 1-1 (ribbon). (b) Top view showing rings C and D. POPC molecules (gray lines) of the membrane bilayer are extracted from the MD trajectory (after 100 ns) and superimposed onto the docking model. The binding pocket surface calculated using SiteMap is shown in gray. White dots represent the grid where ligand atom placement is favorable in SiteMap analysis.

Table 3. SAR of 4a Derivatives with Four Benzene Rings^a


R group	Mouse GPR34	R group	Mouse GPR34
	EC50		EC50
	(LogEC50)		(LogEC50)
	[Emax]		[Emax]
	<RIA>		<RIA>
	n =		n =
4a	H	4d	4-OC6H13
	45 nM		49 nM
	(-7.34 ± 0.06)		(-7.31 ± 0.06)
	[17.8 ± 3.9%]		[17.4 ± 0.5%]
	<2.2 ± 1.5>		<3.4 ± 0.2>
	3		4
4b	4- <i>t</i> -Bu	4e	3- <i>t</i> -Bu
	7.9 nM		14 nM
	(-8.10 ± 0.07)		(-7.84 ± 0.05)
	[16.5 ± 0.7%]		[15.9 ± 2.5%]
	<21 ± 4>		<7.7 ± 0.9>
	4		3
4c	4-Ph	4f	2-Me
	18 nM		78 nM
	(-7.75 ± 0.18)		(-7.11 ± 0.18)
	[11.6 ± 2.2%]		[13.4 ± 1.6%]
	<4.3 ± 2.6>		<0.94 ± 0.61>
	3		3

^aActivities are represented in terms of EC₅₀. LogEC₅₀ (M) and E_{max} (% AP-TGF α release) values are calculated from a sigmoidal concentration–response curve and shown as mean \pm SEM of indicated numbers of independent experiments (*n*). EC₅₀ values are calculated from a mean value of LogEC₅₀. RIA (relative intrinsic activity, a dimensionless parameter; ref 49) is an estimation of agonist activity and represented as an E_{max}/EC₅₀ value relative to that of a positive control compound (previous work). By definition, the RIA value of a positive control compound is equal to 1. “NA” means “not available” because of very low activity.

tetrabenzene fatty acid surrogate (for structures, see Table 3 and Chart S1).

As expected, **4b–4e** exhibited activity toward GPR34 (**4b**, EC₅₀ = 7.9 nM; **4c**, EC₅₀ = 18 nM; **4d**, EC₅₀ = 49 nM; **4e**, EC₅₀ = 14 nM for mouse GPR34, **4c**, EC₅₀ = 41 nM; **4e**, EC₅₀ = 15 nM for human GPR34; Tables 3 and 4, Figures 2c and 3c, and Figure S2e and f). This is in sharp contrast to the finding that the activity toward GPR34 was decreased when a *tert*-butyl group was attached directly to ring C of **3a** (**3v**, EC₅₀ = 900 nM). In the case of 2-methyl substitution on ring D (**4f** (2-Me)), the GPR34 activity was not as high as those of **4b–4e** (**4f**, EC₅₀ = 78 nM; Table 3, Figure 2c, and Figure S2e), but **4f** retained considerable activity. The reduction of activity of **4f** can be interpreted in terms of increased steric hindrance around the 4-position of ring C. Compound **4d**, which bears an *n*-hexyloxy group at the 4-position of ring D, showed high potency toward GPR34 (**4d**, EC₅₀ = 49 nM). Replacement of the terminal methyl group with an amino group diminished the activity (data not shown), supporting the idea that the hydrophobic environment of the terminal substituent R of **4a** derivatives is exposed to the membrane bilayer. These results further support the view that models 1-1 and 1-2 are plausible binding models for GPR34 activation.

In the recently solved crystal structure of GPR40 complexed with 2-[(3*S*)-6-[[3-[2,6-dimethyl-4-(3-methylsulfonylpropoxy)phenyl]phenyl]methoxy]-2,3-dihydro-1-benzofuran-3-yl]acetic acid (TAK-875 (**35**); for the chemical structure, see Chart S2), the ligand protruded between TM3 and TM4.²² In the case of

derivatives of **35**, introduction of additional groups on the benzene ring located outside the receptor has been reported to have little disrupting effect on receptor binding.^{32,33} Our result is in good agreement with this finding in the case of GPR40. Furthermore, in the crystal structure of GPR40 in complex with (5*aR*,6*S*,6*aS*)-3-({2',6'-dimethyl-4'-[3-(methylsulfonyl)propoxy][1,1'-biphenyl]-3-yl-methoxy}-5,5*a*,6,6*a*-tetrahydrocyclopropa[4,5]cyclopenta[1,2-*c*]pyridine-6-carboxylic acid (MK-8666 (**36**)) and (2*S*,3*R*)-3-cyclopropyl-3-[(2*R*)-2-(1-((1*S*)-1-[5-fluoro-2-(trifluoromethoxy)phenyl]-ethyl)-piperidin-4-yl)-3,4-dihydro-2*H*-1-benzopyran-7-yl]-2-methylpropanoic acid (AgoPAM AP8 (**37**); for the chemical structures of **36** and **37**, see Chart S2),²³ **36** also protruded between TM3 and TM4 as in the case of **35** and allosteric agonist, and **37** was bound in the lipid-facing binding pocket between TM4 and TM5. During the binding process of a lipid-like ligand, the pathway of the ligand entry from the membrane bilayer was suggested in the sphingosine 1-phosphate receptor subtype 1 (S1P₁)^{26,34} and the cannabinoid receptor subtype 2 (CB₂).³⁵ These reports support the viewpoint that the hydrophobic part of the ligand can interact with the membrane bilayer or the receptor membrane-embedded surface upon docking as shown in our present models (Figure 7b and also see Figure 8).

Finally, we docked the endogenous LysoPS (18:1) into preferred model 1-1 using the Glide docking program.^{30,31} The obtained binding mode suggested the *cis*-olefin moiety of the (18:1) oleoyl group fitted well to the relevant L-shaped binding

Table 4. SAR of 2-Deoxy-LysoPS Analogues for Human GPR34^a

	R group	Human GPR34		R group	Human GPR34
		EC50 (LogEC50) [E _{max}] <RIA> n =			EC50 (LogEC50) [E _{max}] <RIA> n =
Positive Control		150 nM (-6.82 ± 0.14) [25.6 ± 2.5%] <1> 3	18:1-LysoPS		250 nM (-6.61 ± 0.31) [15.5 ± 4.2%] <0.64 ± 0.35> 4
3a		90 nM (-7.02 ± 0.22) [21.7 ± 4.3%] <1.7 ± 0.6> 3	3v		300 nM (-6.53 ± 0.09) [28.2 ± 4.6%] <0.70 ± 0.21> 3
3f		180 nM (-6.74 ± 0.06) [22 ± 2.4%] <1.2 ± 0.1> 5	4a		78 nM (-7.11 ± 0.06) [23.0 ± 2.8%] <2.1 ± 0.4> 3
3g		230 nM (-6.64 ± 0.03) [24.2 ± 1.6%] <1.2 ± 0.2> 3	4c		41 nM (-7.39 ± 0.11) [17.0 ± 1.7%] <3.2 ± 1.2> 3
3h		27 nM (-7.57 ± 0.03) [22.1 ± 0.4%] <9.8 ± 2.1> 3	4e		15 nM (-7.83 ± 0.10) [18.8 ± 3.0%] <8.6 ± 0.2> 3

^aActivities are represented in terms of EC₅₀. LogEC₅₀ (M) and E_{max} (% AP-TGF α release) values are calculated from a sigmoidal concentration–response curve and shown as mean \pm SEM of indicated numbers of independent experiments (*n*). EC₅₀ values are calculated from a mean value of LogEC₅₀. RIA (relative intrinsic activity, a dimensionless parameter; ref 49) is an estimation of agonist activity and represented as an E_{max}/EC₅₀ value relative to that of a positive control compound (compound 6 in the previous work, ref 13). By definition, the RIA value of a positive control compound is equal to 1. “NA” means “not available” because of very low activity.

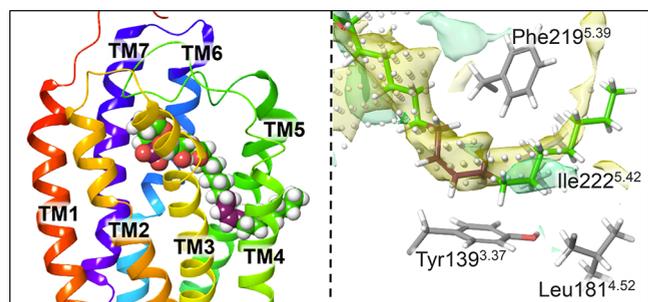


Figure 8. Binding model of (18:1) Oleoyl-LysoPS. The left figure presents the side view of the binding model of LysoPS (18:1) in complex with GPR34 model 1-1 (ribbon). The extracellular half of the receptor is shown. The right figure presents a close-up view of the oleic acid of LysoPS in the receptor binding site, calculated using SiteMap, and the receptor side chains near the ligand terminus. The *cis*-olefin in oleic acid is colored purple to highlight the kink in the ligand shape. The hydrophobic pocket is shown as a yellow surface, and the hydrophilic pocket is shown as a green surface. White dots represent the grid where placement of ligand atoms is allowed in the SiteMap analysis. Leu181 and Ile222 in human GPR34 are mutated to Val in mouse GPR34.

pocket (Figure 8). This is consistent with the fact that LysoPS (18:0), which bears a saturated fatty acid chain, favoring trans conformations, shows significantly decreased GPR34-activation activity. Thus, the L-shaped pocket can explain the origin of the

fatty acid selectivity of GPR34, which favors unsaturated fatty acids over saturated ones.

CONCLUSIONS

Our present focused SAR study revealed that GPR34-activating potency strongly depends on the configuration and substitution of aromatic rings in a series of fatty acid surrogates. In this study, we explored the preference of the receptor's binding site for the shape of the hydrophobic tail of 2-deoxy-LysoPS analogues bearing aromatic ring(s) in the fatty acid moiety. Tribenzene-containing compound 3a showed increased agonistic activity toward GPR34. Our SAR study of 3a and its analogues revealed strong dependence of GPR34-activating ability upon substituents at the terminal aromatic substituent (ring C). In particular, substitution at the 4-position of ring C afforded potent GPR34 agonists.

To understand the hydrophobic interactions between the ligand and the receptor, we carried out computational docking to binding-site models obtained by homology modeling/molecular dynamics calculations based on the crystal structure of agonist-bound human P2Y12. The preferred model of the GPR34 receptor contains an L-shaped hydrophobic pocket, which is complementary to the kink in the molecular shape of the fatty acid surrogates 3a. Docking calculations with 3v, the analogue having a *tert*-butyl group at the 4-position of terminal ring C in the tribenzene system, indicated that the group would be located in a narrow hydrophobic passage between TM4 and TM5, where it would sterically interfere with the binding

pocket. Indeed, **3v** (4-*t*-Bu) showed decreased activity compared with that of **3a**. On the other hand, docking calculations predicted that introduction of a bulky *tert*-butyl group at the 3- or 4-position of terminal ring D of **4a** should have little disturbing effect on the interaction with the active site. Accordingly, we synthesized **4b–4e** and found that these analogues indeed retained high agonistic activity toward GPR34. Therefore, the strong dependency of GPR34 agonist activity on the shape of the fatty acid moiety of LysoPS analogues found in this work can be explained in terms of interactions between the fatty acid surrogates of the LysoPS analogue and the hydrophobic binding pocket of the receptor. In other words, matching/mismatching of the kink in the molecular shape of the benzene-based lipid surrogate with the L-shaped hydrophobic binding pocket of GPR34 is a key determinant of the activity. It is noteworthy that the position of the kink in the bioactive lipid surrogates corresponds well to the location of the *cis*-olefin moiety of (18:1) oleoyl LysoPS, an endogenous ligand and activator of GPR34 receptor. We believe this feature of the hydrophobic pocket of GPR34 will be helpful in the design of potent and specific modulators for use in detailed studies of the pathophysiological roles of this receptor.

EXPERIMENTAL SECTION

General Synthetic Procedures. Melting points were determined with a Yanaco micro-melting point apparatus without correction. ^1H (400 MHz) and ^{13}C (100 MHz) NMR spectra were recorded on a Bruker Avance400. Chemical shifts were calibrated with tetramethylsilane as an internal standard or with the solvent peak and are shown in ppm (δ) values; coupling constants are shown in hertz (Hz). The following abbreviations are used: s = singlet, d = doublet, t = triplet, q = quartet, dd = double doublet, dt = double triplet, dq = double quartet, h = hexet, m = multiplet, and brs = broad singlet. An electron spray ionization time-of-flight (ESI-TOF) mass spectrometer (Bruker micrOTOF-05) was used to obtain high-resolution mass spectra (HRMS). All commercially available compounds and solvents were used as received. Combustion analyses were carried out in the Microanalytical Laboratory of the Graduate School of Pharmaceutical Sciences, the University of Tokyo. All tested compounds showed $\geq 95\%$ purity on the basis of combustion analysis.

Details of the synthesis of compounds **1–3** and **13–34** are given in the [Supporting Information](#). The synthetic procedure for **4a** (**5a–12a**) is described below.

3-(4-Bromophenoxy)benzaldehyde (5a). To a 100 mL round-bottomed flask containing copper acetate (1.1055 g, 6.0864 mmol) in anhydrous CH_2Cl_2 (60 mL) were added pyridine (960 μL , 11.92 mmol), 4 Å molecular sieves, *p*-bromophenol (1.0568 g, 6.1084 mmol), and 3-formylphenylboronic acid (1.3488 g, 8.9956 mmol). The resulting mixture was stirred for 47 h at room temperature. Polyvinylpyridine (5.0550 g) was added, and stirring was continued for 10 min. The mixture was filtered on Celite, and the solvent was evaporated to give a green solid. The residue was column-chromatographed on an open column with silica-gel (*n*-hexane:Et₂O = 10:1) to afford **5a** as a colorless oil (644.7 mg, 2.326 mmol, 38%).

^1H NMR (400 MHz, CDCl_3): 9.964 (1H, s), 7.638 (1H, ddd, J = 7.6 Hz, 1.2 Hz, 1.2 Hz), 7.515 (1H, dd, J = 7.9 Hz, 7.9 Hz), 7.476 (2H, d, J = 9.0 Hz), 7.452 (1H, dd, J = 2.4 Hz, 1.4 Hz), 7.277 (1H, ddd, J = 8.1 Hz, 2.6 Hz, 1.0 Hz), 6.919 (2H, d, J = 8.9 Hz). ^{13}C NMR (100 MHz, CDCl_3): 191.38, 157.85, 155.43, 138.12, 133.00, 130.57, 125.20, 124.63, 121.06, 118.07, 116.73.

3-(4-Bromophenoxy)phenylmethanol (6a). To a solution of **5a** (622.5 mg, 2.246 mmol) in MeOH (12 mL) was added portionwise NaBH_4 (169.1 mg, 4.470 mmol) at 0 °C. The reaction mixture was stirred at 0 °C for 30 min and then diluted with CH_2Cl_2 (10 mL), saturated aqueous solution of NaHCO_3 (5 mL), and water (15 mL). The aqueous layer was extracted three times with CH_2Cl_2 (20 mL).

The combined organic layer was washed with brine, dried over Na_2SO_4 , and filtered. The solvent was evaporated to give a colorless oil. The residue was column-chromatographed on an open column with silica-gel (*n*-hexane:AcOEt = 3:1) to afford **6a** as a colorless oil (582.7 mg, 2.088 mmol, 93%).

^1H NMR (400 MHz, CDCl_3): 7.428 (2H, d, J = 9.0 Hz), 7.326 (1H, dd, J = 7.9 Hz, 7.9 Hz), 7.112 (1H, dd, J = 7.6 Hz, 0.3 Hz), 7.008 (1H, brs), 6.916 (1H, dd, J = 7.9 Hz, 1.1 Hz), 6.885 (2H, d, J = 9.0 Hz), 4.671 (2H, s), 1.753 (1H, brs). ^{13}C NMR (100 MHz, CDCl_3): 157.07, 156.33, 143.09, 132.69, 129.99, 121.94, 120.58, 118.01, 117.17, 115.77, 64.81.

Methyl 3-(2-((3-(4-Bromophenoxy)benzyl)oxy)phenyl)propanoate (8a). To a solution of **7** (564.8 mg, 3.134 mmol), **6a** (563.9 mg, 2.020 mmol), and PPh_3 (1.0263 g, 3.9128 mmol) in anhydrous toluene (30 mL) was added dropwise DEAD (2 M in toluene, 2.0 mL) at room temperature. The yellow mixture was heated at 70–80 °C with stirring for 3 h under an Ar atmosphere; the color changed to orange. The solvent was evaporated to give an orange oil containing a white powder. The residue was column-chromatographed on an open column with silica-gel (*n*-hexane:AcOEt = 10:1) to afford **8a** as a pink oil (640.0 mg, 1.450 mmol, 72%).

^1H NMR (400 MHz, CDCl_3): 7.433 (2H, d, J = 8.9 Hz), 7.353 (1H, dd, J = 7.9 Hz, 7.9 Hz), 7.187–7.147 (3H, m), 7.060 (1H, brs), 6.943 (1H, dd, J = 8.4 Hz, 1.0 Hz), 6.919–6.845 (4H, m), 5.066 (2H, s), 3.640 (3H, s), 2.973 (2H, t, J = 7.9 Hz), 2.610 (2H, t, J = 7.8 Hz). ^{13}C NMR (100 MHz, CDCl_3): 173.56, 157.11, 156.27, 156.21, 139.47, 132.72, 130.13, 130.04, 129.13, 127.57, 122.03, 120.90, 120.67, 118.07, 117.25, 115.86, 111.53, 69.19, 51.49, 33.98, 26.13. HRMS (ESI-TOF) $[\text{M} + \text{Na}]^+$ calcd for $\text{C}_{23}\text{H}_{21}\text{BrNaO}_4^+$: 463.0515, 465.0495; found: 463.0498, 465.0483.

Methyl 3-(2-((3-([1,1'-Biphenyl]-4-yloxy)benzyl)oxy)phenyl)propanoate (9a). Compound **8a** (300.4 mg, 0.6807 mmol) was dissolved in EtOH (1.5 mL) and anhydrous toluene (3.0 mL). Phenylboronic acid (133.5 mg, 1.095 mmol), H_2O (3 mL), Na_2CO_3 (317.8 mg, 2.998 mmol), PPh_3 (13.3 mg, 0.0507 mmol), and $\text{Pd}(\text{OAc})_2$ (6.2 mg, 0.028 mmol) were added to the solution. The mixture was degassed by means of two freeze–pump–thaw cycles and stirred at 80–90 °C under an Ar atmosphere for 17 h. H_2O (10 mL) was added, and the aqueous layer was extracted three times with AcOEt (15 mL). The combined organic layer was washed with brine, dried over Na_2SO_4 , and filtered. The solvent was evaporated to leave a colorless oil containing black droplets. The residue was column-chromatographed on an open column with silica-gel (*n*-hexane:AcOEt = 10:1) to afford **9a** as a colorless oil (256.9 mg, 0.5858 mmol, 86%).

^1H NMR (400 MHz, CDCl_3): 7.578–7.538 (4H, m), 7.447–7.404 (2H, m), 7.358 (1H, dd, J = 7.9 Hz, 7.9 Hz), 7.347–7.305 (1H, m), 7.183–7.143 (3H, m), 7.119 (1H, brs), 7.084 (2H, d, J = 8.7 Hz), 7.001 (1H, dd, J = 8.1 Hz, 0.9 Hz), 6.890 (1H, ddd, J = 7.1 Hz, 7.1 Hz, 1.0 Hz), 6.863 (1H, d, J = 7.7 Hz), 5.070 (2H, s), 3.607 (3H, s), 2.977 (2H, t, J = 7.7 Hz), 2.610 (2H, t, J = 7.7 Hz). ^{13}C NMR (100 MHz, CDCl_3): 173.67, 157.53, 156.46, 156.33, 140.48, 139.34, 136.50, 130.12, 129.94, 129.14, 128.76, 128.45, 127.56, 127.04, 126.88, 121.70, 120.84, 119.27, 118.05, 117.22, 111.55, 69.28, 51.44, 33.99, 26.15. HRMS (ESI-TOF) $[\text{M} + \text{Na}]^+$ calcd for $\text{C}_{29}\text{H}_{26}\text{NaO}_4^+$: 461.1723; found: 461.1733.

3-(2-((3-([1,1'-Biphenyl]-4-yloxy)benzyl)oxy)phenyl)propanoic Acid (10a). Compound **9a** (232.9 mg, 0.5311 mmol) was dissolved in MeOH (1.0 mL) and THF (1.0 mL), and to the solution was added 2 M aqueous solution of NaOH (1.0 mL). The solution was stirred at room temperature for 3.5 h, and then 2 M aqueous solution of HCl was added to acidify the mixture to pH 2. The aqueous layer was extracted three times with AcOEt (10 mL). The combined organic layer was washed with brine, dried over Na_2SO_4 , and filtered. The solvent was evaporated to leave a colorless oil. The residue was column-chromatographed on an open column with silica-gel twice (first *n*-hexane:AcOEt = 2:1, second *n*-hexane:AcOEt = 6:1–2:1) to afford **10a** as a white solid (189.5 mg, 0.4464 mmol, 84%).

^1H NMR (400 MHz, CDCl_3): 7.566–7.526 (4H, m), 7.429–7.391 (2H, m), 7.346 (1H, dd, J = 7.8 Hz, 7.8 Hz), 7.326–7.287 (1H, m), 7.186–7.147 (3H, m), 7.104–7.057 (3H, m), 6.994 (1H, dd, J = 8.1

H_z, 0.9 Hz), 6.903–6.849 (2H, m), 5.056 (2H, s), 2.965 (2H, t, *J* = 7.7 Hz), 2.638 (2H, t, *J* = 7.7 Hz). ¹³C NMR (100 MHz, CDCl₃): 179.28, 157.56, 156.44, 156.33, 140.46, 139.25, 136.49, 130.12, 129.98, 128.80, 128.74, 128.44, 127.66, 127.02, 126.87, 121.69, 120.86, 119.32, 118.07, 117.19, 111.55, 69.29, 33.86, 25.84. HRMS (ESI-TOF) [M – H][–] calcd for C₂₈H₂₄O₄: 423.1602; found: 423.1631. Anal. Calcd for C₂₈H₂₄O₄·0.1H₂O: C, 78.89; H, 5.72. Found: C, 78.80; H, 5.99. Mp 111.8–113.5 °C (recrystallized from *n*-hexane/CH₂Cl₂, colorless needles).

tert-Butyl *O*-((3-((3-(2-((3-([1,1'-Biphenyl]-4-yloxy)benzyl)oxy)phenyl)propanoyl)-oxy)propoxy)(*tert*-butoxy)phosphoryl)-*N*-(*tert*-butoxycarbonyl)-*L*-serinate (**12a**). Compounds **10a** (112.8 mg, 0.2657 mmol) and **11** (95.7 mg, 0.210 mmol) were dissolved in anhydrous CH₂Cl₂ (1.5 mL). DMAP (7.5 mg, 0.061 mmol) was added, and the flask was cooled to 0 °C in an ice–water bath. EDCI-HCl (79.0 mg, 0.412 mmol) was added at 0 °C, and the whole was stirred at room temperature under an Ar atmosphere for 20 h. Additional EDCI-HCl (39.8 mg) and MeOH (1.0 mL) were added to the reaction mixture, and stirring was continued for an additional 24.3 h. H₂O (10 mL) was added, and the aqueous layer was extracted three times with CH₂Cl₂ (15 mL). The combined organic layer was washed with brine, dried over Na₂SO₄, and filtered. The solvent was evaporated to give a colorless oil. The residue was column-chromatographed on an open column with silica-gel (*n*-hexane:AcOEt = 2:1 to 1:1) to yield **12a** as a colorless sticky oil (115.4 mg, 0.1339 mmol, 64%, a mixture of diastereomers).

¹H NMR (400 MHz, CDCl₃): 7.581–7.542 (4H, m), 7.451–7.413 (2H, m), 7.383–7.309 (2H, m), 7.194–7.145 (3H, m), 7.112 (1H, brs), 7.084 (2H, d, *J* = 8.7 Hz), 6.997 (1H, dd, *J* = 8.1 Hz, 0.8 Hz), 6.909–6.857 (2H, m), 5.499 (1H, d, *J* = 7.8 Hz), 5.084 (2H, s), 4.366–4.301 (2H, m), 4.231–4.176 (1H, m), 4.147–4.092 (2H, m), 4.027–3.978 (2H, m), 2.979 (2H, t, *J* = 7.6 Hz), 2.624 (2H, t, *J* = 7.7 Hz), 1.991–1.881 (2H, m), 1.473–1.432 (27H, m). ¹³C NMR (100 MHz, CDCl₃): 173.05, 168.32, 157.49, 156.43, 156.31, 155.21, 140.43, 139.34, 136.48, 130.09, 129.95, 129.01, 128.75, 128.43, 127.57, 127.03, 126.85, 121.64, 120.81, 119.25, 117.99, 117.20, 111.57, 83.68, 83.61, 82.64, 82.61, 79.90, 69.27, 67.38, 64.03, 63.98, 60.39, 54.43, 54.35, 34.00, 29.75, 29.74, 29.71, 29.51, 29.44, 28.26, 27.90, 26.08. ³¹P NMR (161 MHz, CDCl₃): –5.56, –5.70. HRMS (ESI-TOF) [M + Na]⁺ calcd for C₄₇H₆₀NNaO₁₂P⁺: 884.3745; found: 884.3743.

O-((3-((3-(2-((3-([1,1'-Biphenyl]-4-yloxy)benzyl)oxy)phenyl)propanoyl)oxy)propoxy)(hydroxy)phosphoryl)-*L*-serine (**4a**). Compound **12a** (98.4 mg, 0.114 mmol) was dissolved in TFA (1.0 mL) at 0 °C, and the solution was stirred at room temperature for 1.5 h, diluted with CH₂Cl₂ (6 mL), and evaporated. The residue was subjected to flash column chromatography (CHCl₃:MeOH:AcOH = 9:1:0 to 8:1:1). The resulting colorless oil was dissolved in CH₂Cl₂ and TFA at 0 °C, and the solution was evaporated at 40 °C overnight to afford **4a** as the TFA salt (white powder, 41.8 mg, 0.0644 mmol, 56%).

¹H NMR (400 MHz, CDCl₃/TFA-*d*): 7.578–7.550 (4H, m), 7.452–7.434 (2H, m), 7.380–7.315 (2H, m), 7.213–7.162 (2H, m), 7.116–7.056 (4H, m), 7.011 (1H, dd, *J* = 8.1, 1.0 Hz), 6.917–6.883 (2H, m), 5.095 (2H, s), 4.600–4.587 (2H, m), 4.474 (1H, brs), 4.169 (2H, t, *J* = 6.3 Hz), 4.018 (2H, brs), 2.968 (2H, t, *J* = 7.4 Hz), 2.716 (2H, t, *J* = 7.5 Hz), 1.943–1.914 (2H, m). ³¹P NMR (CDCl₃/TFA-*d*): –1.36. HRMS (ESI-TOF) [M – H][–] calcd for C₃₄H₃₅NO₁₀P[–]: 648.2004; found: 648.2014. Anal. Calcd for C₃₄H₃₆NO₁₀P·1.5CF₃CO₂H: C, 54.15; H, 4.61; N, 1.71. Found: C, 54.21; H, 4.86; N, 1.78. Mp 112–144 °C (colorless cubes).

TGF α Shedding Assay. TGF α shedding assay of the synthesized compounds was performed using mouse GPR34 LysoPS receptors, as described previously.¹⁶ This assay detects activation of GPCRs by measuring ectodomain shedding of a membrane-bound proform of alkaline phosphatase-tagged TGF α (AP-TGF α) into conditioned media. Because AP-TGF α shedding response occurs almost exclusively downstream of G $\alpha_{12/13}$ and G α_q signaling, we additionally transfected a chimeric G $\alpha_{q/i1}$ subunit to detect activation of G γ -coupled GPR34.

Briefly, agonistic activities of all synthesized compounds toward LysoPS receptors were measured as follows. HEK293A cells were seeded in 100 mm dishes with 10 mL of cell suspension at a density of

2×10^5 cells/mL in 10% (v/v) fetal bovine serum (FBS)-supplemented Dulbecco's modified Eagle's medium (DMEM) and cultured in a CO₂ incubator. After 24 h, the cells were transfected with a mixture of plasmids encoding AP-TGF α (2.5 μ g per dish), the chimeric G $\alpha_{q/i1}$ subunit (0.5 μ g per dish) and GPR34 (2.5 μ g per dish) using Lipofectamine 2000 reagent (Life Technologies; 12.5 μ L per dish). As a control (mock transfection; for measuring receptor-independent responses), cells were transfected with a mixture of plasmids encoding AP-TGF α (2.5 μ g per dish), the chimeric G $\alpha_{q/i1}$ subunit (0.5 μ g per dish) and empty vector (2.5 μ g per dish). After 24 h, cells were harvested with trypsin-EDTA solution, washed with PBS, and suspended in Hank's balanced salt solution (HBSS; 35 mL per dish) containing 5 mM HEPES (pH 7.4). For GPR34-expressing cells and the corresponding control cells, cell suspension was seeded in a 96-well plate (80 μ L per well) and mixed with the LPA receptor antagonist Ki16425 at a final concentration of 10 μ M (10 μ L per well). After 30 min incubation, cells were treated with compounds diluted in 0.01% (w/v) bovine serum albumin (BSA; essentially fatty acid-free grade; Sigma-Aldrich) containing HBSS (10 μ L per well; total volume of 100 μ L per well), and incubated for 1 h at 37 °C. The plate was centrifuged for 2 min at 190g, and conditioned medium (CM) was transferred to a blank 96-well plate (80 μ L per well). AP solution (10 mM *p*-nitrophenyl phosphate (*p*-NPP), 40 mM Tris-HCl (pH 9.5), 40 mM NaCl, and 10 mM MgCl₂) was added to both the CM plate and the cell plate (80 μ L per well). Absorbance at 405 nm (OD₄₀₅) of the two plates was measured before and after incubation for 1 h at room temperature using a microplate reader. AP-TGF α release was calculated according to the following equations, after subtracting baseline responses in mock-transfected cells (also see [Supporting Information](#) for details of the calculations)

$$\text{AP-TGF}\alpha_{\text{CM}} (\%) = (\text{OD}_{405_{\text{CM}}} / (\text{OD}_{405_{\text{CM}}} + \text{OD}_{405_{\text{cell}}})) \times 1.25 \times 100$$

where factor 1.25 was used to obtain AP-TGF α in total CM (100 μ L) from measured AP-TGF α in transferred CM (80 μ L).

$$\text{AP-TGF}\alpha \text{ release} = \text{AP-TGF}\alpha_{\text{CM}} \text{ under stimulated conditions} (\%) - \text{AP-TGF}\alpha_{\text{CM}} \text{ under vehicle-treated conditions} (\%)$$

Typically, AP-TGF α_{CM} under vehicle-treated conditions ranged from 10 to 25% depending on the transfected plasmids and assay conditions.

Data are shown as receptor-specific responses calculated as

$$\text{receptor-specific AP-TGF}\alpha \text{ release} (\%) = \text{AP-TGF}\alpha \text{ release in receptor-transfected cells} (\%) - \text{AP-TGF}\alpha \text{ release in mock-transfected cells} (\%)$$

LogEC₅₀ (EC₅₀) and E_{max} values were calculated for active compounds with plateau or semiplateau responses by fitting data to a four-parameter sigmoid curve using GraphPad Prism 6 (GraphPad, USA).

Homology Modeling. GPR34 homology models were built using MODELER 9.13³⁶ based on the crystal structure of P2Y12 (PDB ID: 4PXZ) bound to the P2Y12 full agonist 2-methylthio-adenosine-5'-diphosphate (2MeSADP).²¹ The human GPR34 sequence was obtained from UniProt (code: Q9UPC5).³⁷ The N-terminal (from Met1 to Asn42) and C-terminal (from Arg345 to Thr381) regions of the GPR34 sequence were deleted because template structure was not available for these regions. Sequence alignment was conducted with ClustalW³⁸ using Multiple Sequence Viewer (Schrödinger, LLC, NY) as a graphical interface. The alignment was manually modified to match the cysteine residues that form two conserved disulfide bonds in GPR34 and P2Y12 (Cys46-Cys299 and Cys127-Cys204 in GPR34)³⁹ and to fill the gap in the transmembrane helices. The final alignment is shown in [Figure S3](#) with the percentage identity (30%). The region from Ser331 to Phe343 in GPR34 was constrained to α -helix form to create helix 8. Ten models were constructed and three models that had the best MODELER DOPE scores were used in the next docking process.

Docking of GPR34 Agonist 3a. The constructed GPR34 homology models were refined for docking simulations with Protein Preparation Wizard in Maestro (Schrödinger, LLC, NY)⁴⁰ using PROPKA with pH 7.0 in the H-bond assignment step.⁴¹ Suitable 3D coordinates and states of compound **3a** were generated using the OPLS2005 force field and the Epik calculation with pH 7.0 on LigPrep (version 3, Schrödinger, LLC, NY). The most suitable ionization states of the carboxyl, amino, and phosphodiester groups to give the total molecular charge -1 were used.

Compound **3a** was docked to the three GPR34 homology models using the standard induced fit docking (IFD) protocol (Schrödinger, LLC, NY).^{24,25} Box center was defined as the centroid of the P2Y12 agonist implemented in GPR34 homology models. In the IFD protocol, **3a** was first docked to the rigid protein using Glide (version 6, Schrödinger, LLC, NY)^{30,31} with the softened potential of van der Waals radii scaling of 0.5 for the protein and 0.5 for the ligand. The resulting top 20 poses of ligands were used in the next Prime side-chain prediction step. Residues within 5 Å of any ligand pose were subjected to conformational search and minimization with an implicit membrane using Prime (version 3, Schrödinger, LLC, NY).^{42,43} The resulting 20 protein conformers were used for the redocking process in Glide SP mode with the default potential of van der Waals radii scaling of 1.0 for the protein and 0.8 for the ligand. Nineteen complex structures were thus generated from the three homology models.

Molecular Dynamics. The 19 models of GPR34 complexed with **3a** were subjected to 10 ns MD-based energy minimization using Desmond 3.8.5.18.⁴⁴ The OPLS2005 force field was used. Initial ligand–receptor complex models were aligned to the P2Y12 structure obtained from PDBTM^{45–47} and embedded in the POPC membrane placed on the prealigned structure. The system was placed in TIP3P water molecules solvated with 0.15 M NaCl. The system was first subjected to relaxation MD on CPU Desmond and then to production MD on GPU Desmond for 10 ns. Production MD was performed in the NP γ T ensemble at 300 K with 1.01325 bar and 4000 bar·Å as surface tension using Langevin dynamics. Long-range electrostatic interactions were computed using the Smooth Particle Mesh Ewald method.⁴⁸

The GPR34–**3a** complex models after 10 ns MD simulation were extracted and refined using Protein Preparation Wizard into Maestro (Schrödinger, LLC, NY). All 19 complex structures were aligned using the Protein Structure Alignment tool in Maestro (Schrödinger, LLC, NY). Ligands were extracted from the complexes, and ligand poses were clustered according to the ligand heavy atom RMSDs in place in Maestro. The 19 poses were divided into five clusters of which two were major (namely, cluster 1 containing 10 poses and cluster 2 containing 6 poses).

Prime MM-GBSA ΔG_{bind} was calculated with the VSGB solvation model using input ligand partial charges and an implicit membrane. The MM-GBSA ΔG_{bind} scores are listed in Table S2 for the 19 complexes. The top scored three complexes from cluster 1 (models 1-1, 1-2, and 1-3) and two complexes from cluster 2 (models 2-1 and 2-2) were considered for further MD-based energy minimization. Additional 100 ns production MD on GPU Desmond was performed from the last snapshot of the previous 10 ns MD simulation (Figure S4). The conditions were the same as those of the 10 ns simulation.

Ligand Docking and SiteMap Analysis. The structure of the receptor model after the 100 ns MD simulation was extracted from the MD trajectory and refined using Protein Preparation Wizard in Maestro (Schrödinger, LLC, NY). We selected key compounds (**3f**–**3h**, **3v**, **4a**, and **4b**) based on the observed SAR for docking experiments. The selected **3a** derivatives were docked to the refined structures using Glide SP mode. The centroid of the complexed **3a** was defined as the grid center. In ligand docking, the docking position was restricted to the reference position including the serine-phosphate moiety, as shown in Figure S6. Three types of reference core atoms were tried, and the pose with the best docking score was selected except in the case of **3g** docked to model 2-1, where the docking pose in which ring C was placed between TM4 and TM5 gave the best docking score. However, for easy comparison with the docking poses

for **3f** and **3h**, respectively, the docking pose of **3g** in which ring C was placed around TM3, TM5, and TM6 was selected.

In the postdocking minimization stage, strain correction terms were applied. SiteMap (version 3) analysis²⁸ was conducted with each representative model (models 1-1, 1-2, 1-3, 2-1, and 2-2) using default parameters and compared with the binding pose of each compound. The figures were created on the graphical user interface of Maestro or PyMOL Molecular Graphics System (version 1.7, Schrödinger, LLC, NY).

Evaluation of 4a Binding Pose by Ensemble Docking. To further evaluate the docking pose of **4a**, we extracted 9 additional structures from the trajectories of the 100 ns MD simulation, every 5 ns from 55 to 95 ns. Protein structures were prepared, and **4a** was docked to each structure in the same way as mentioned above. The position of the acyl moiety of **4a** in each docking model was visually checked, and the results are summarized in Table S3.

Docking of LysoPS (18:1). LysoPS (18:1) was docked to the refined receptor model after 100 ns MD simulation using Glide SP mode. The centroid of complexed **3a** was defined as the grid center. Postdocking minimization was conducted for 25 poses, and among them, 5 poses were written. The docking pose with the best docking score was selected (Figure 8).

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.7b00693.

Additional graphics, structures of compounds, and full biological data, including details of the syntheses of the compounds (PDF)

Molecular formula strings (CSV)

Compound **3a** in model 1-1 (PDB)

Compound **3a** in model 2-1 (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

*Chemical study (T.O.): phone, +81-3-5841-4730; fax, +81-3-5841-4735; e-mail, ohwada@mol.f.u-tokyo.ac.jp.

*Biological study (J.A.): phone, +81-22-795-6860; fax, +81-22-795-6859; e-mail, jaoki@m.tohoku.ac.jp.

*Computational study (T.H.): phone, +813-3599-8638; fax, +81-3-3599-2064; e-mail, t-hirokawa@aist.go.jp.

ORCID

Yuko Otani: 0000-0002-8104-283X

Tomohiko Ohwada: 0000-0001-5390-0203

Author Contributions

[▽]M.S. and A.I. contributed equally to this work.

Author Contributions

M.S., S.N., S.J., M.I., Y.O., and T.O. (The University of Tokyo) performed the chemical studies including design and synthesis of compounds. A.I., A.U., T.K., K.M., and J.A. (Tohoku University) performed biological study using the TGF α shedding assay. M.S. (The University of Tokyo) and T.H. (AIST and Tsukuba University) performed the computational studies of protein–ligand binding models.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by grants from the Ministry of Education, Science, Sports, and Culture of Japan to J.A., T.O., and T.H. We thank Prof. Dr. Masaji Ishiguro, Niigata

University of Pharmacy and Applied Life Sciences, Niigata, Japan, for his initial guidance of the computational study. This work was supported by the Platform for Drug Discovery, Informatics, and Structural Life Science from the Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was supported by JST, PRESTO Grant Number JPMJPR1331 (to A.I.); the PRIME from Japan Agency for Medical Research and development, AMED (to A.I.); Japan Society for the Promotion of Science (JSPS) Grant Number 16H05085 (J.A.); AMED-CREST, AMED (to J.A.).

ABBREVIATIONS USED

LysoPS, lysophosphatidylserine; PS, phosphatidylserine; LPS, lysophosphatidylserine receptor; SIP, Sphingosine-1-phosphate; TGF α , transforming growth factor- α ; AP-TGF α , alkaline phosphatase-tagged TGF α ; TM, transmembrane helix; EL, extracellular loop; 2MeSADP, 2-methylthio-adenosine-5'-diphosphate; DOPE, discrete optimized protein energy; MD, molecular dynamics; MM-GBSA, molecular mechanics with generalized Born and surface area solvation; SP, standard precision; Mp, melting point; EDCI, *N*-(3-(dimethylamino)propyl)-*N*-ethylcarbodiimide hydrochloride; MS, molecular sieves; HBSS, Hanks' balanced salt solution; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; CM, conditioned media; DMEM, Dulbecco's modified Eagle's medium; DAPI, 4',6-diamidino-2-phenylindole; CI, confidence interval; SEM, standard errors of mean; IFD, induced fit docking; GPU, graphics processing unit; POPC, 1-palmitoyl-2-oleoylphosphatidylcholine; PDBTM, Protein Data Bank of Transmembrane Proteins; NP γ T, constant normal pressure and lateral surface tension of membranes and constant temperature

REFERENCES

- (1) Sugo, T.; Tachimoto, H.; Chikatsu, T.; Murakami, Y.; Kikukawa, Y.; Sato, S.; Kikuchi, K.; Nagi, T.; Harada, M.; Ogi, K.; Ebisawa, M.; Mori, M. Identification of a lysophosphatidylserine receptor on mast cells. *Biochem. Biophys. Res. Commun.* **2006**, *341*, 1078–1087.
- (2) Kitamura, H.; Makide, K.; Shuto, A.; Ikubo, M.; Inoue, A.; Suzuki, K.; Sato, Y.; Nakamura, S.; Otani, Y.; Ohwada, T.; Aoki, J. GPR34 is a receptor for lysophosphatidylserine with a fatty acid at the sn-2 position. *J. Biochem.* **2012**, *151*, 511–518.
- (3) Makide, K.; Kitamura, H.; Sato, Y.; Okutani, M.; Aoki, J. Emerging lysophospholipid mediators, lysophosphatidylserine, lysophosphatidylthreonine, lysophosphatidylethanolamine and lysophosphatidylglycerol. *Prostaglandins Other Lipid Mediators* **2009**, *89*, 135–139.
- (4) Liebscher, I.; Müller, U.; Teupser, D.; Engemaier, E.; Engel, K. M. Y.; Ritscher, L.; Thor, D.; Sangkuhl, K.; Ricken, A.; Wurm, A.; Piehler, D.; Schmutzler, S.; Fuhrmann, H.; Albert, F. W.; Reichenbach, A.; Thiery, J.; Schöneberg, T.; Schulz, A. Altered immune response in mice deficient for the G protein-coupled receptor GPR34. *J. Biol. Chem.* **2011**, *286*, 2101–2110.
- (5) (a) Preissler, J.; Grosche, A.; Lede, V.; Le Duc, D.; Krügel, K.; Matyash, V.; Szulzewsky, F.; Kallendrusch, S.; Immig, K.; Kettenmann, H.; Bechmann, I.; Schöneberg, T.; Schulz, A. Altered microglial phagocytosis in GPR34-deficient mice. *Glia* **2015**, *63*, 206–215. (b) Tokizane, K.; Konishi, H.; Makide, K.; Kawana, H.; Nakamura, S.; Kaibuchi, K.; Ohwada, T.; Aoki, J.; Kiyama, H. Phospholipid localization implies microglial morphology and function via Cdc42 in vitro. *Glia* **2017**, *65*, 740–755.
- (6) Yu, W.; Ma, S.; Wang, L.; Zuo, B.; Li, M.; Qiao, Z.; Pan, X.; Liu, Y.; Wang, J. Upregulation of GPR34 expression affects the progression and prognosis of human gastric adenocarcinoma by PI3K/PDK1/AKT pathway. *Histol. Histopathol.* **2013**, *28*, 1629–1638.
- (7) Jin, Z. T.; Li, K.; Li, M.; Ren, Z. G.; Wang, F. S.; Zhu, J. Y.; Leng, X. S.; Yu, W. D. G-protein coupled receptor 34 knockdown impairs the proliferation and migration of HGC-27 gastric cancer cells in vitro. *Chin. Med. J. (Engl)* **2015**, *128*, 545–549.
- (8) Ansell, S. M.; Akasaka, T.; McPhail, E.; Manske, M.; Braggio, E.; Price-Troska, T.; Ziesmer, S.; Secreto, F.; Fonseca, R.; Gupta, M.; Law, M.; Witzig, T. E.; Dyer, M. J.; Dogan, A.; Cerhan, J. R.; Novak, A. J. t(X;14)(p11;q32) in MALT lymphoma involving GPR34 reveals a role for GPR34 in tumor cell growth. *Blood* **2012**, *120*, 3949–3957.
- (9) Iida, Y.; H Tsuno, N.; Kishikawa, J.; Kaneko, K.; Murono, K.; Kawai, K.; Ikeda, T.; Ishihara, S.; Yamaguchi, H.; Sunami, E.; Kitayama, J.; Yatomi, Y.; Watanabe, T. Lysophosphatidylserine stimulates chemotactic migration of colorectal cancer cells through GPR34 and PI3K/Akt pathway. *Anticancer Res.* **2014**, *34*, 5465–5472.
- (10) Iwashita, M.; Makide, K.; Nonomura, T.; Misumi, Y.; Otani, Y.; Ishida, M.; Taguchi, R.; Tsujimoto, M.; Aoki, J.; Arai, H.; Ohwada, T. Synthesis and evaluation of lysophosphatidylserine analogues as inducers of mast cell degranulation. Potent activities of lysophosphatidylthreonine and its 2-deoxy derivative. *J. Med. Chem.* **2009**, *52*, 5837–5863.
- (11) Uwamizu, A.; Inoue, A.; Suzuki, K.; Okudaira, M.; Shuto, A.; Shinjo, Y.; Ishiguro, J.; Makide, K.; Ikubo, M.; Nakamura, S.; Jung, S.; Sayama, M.; Otani, Y.; Ohwada, T.; Aoki, J. Lysophosphatidylserine analogues differentially activate three LysoPS receptors. *J. Biochem.* **2015**, *157*, 151–160.
- (12) Ikubo, M.; Inoue, A.; Nakamura, S.; Jung, S.; Sayama, M.; Otani, Y.; Uwamizu, A.; Suzuki, K.; Kishi, T.; Shuto, A.; Ishiguro, J.; Okudaira, M.; Kano, K.; Makide, K.; Aoki, J.; Ohwada, T. Structure-activity relationships of lysophosphatidylserine analogs as agonists of G-protein-coupled receptors GPR34, P2Y10, and GPR174. *J. Med. Chem.* **2015**, *58*, 4204–4219.
- (13) Jung, S.; Inoue, A.; Nakamura, S.; Kishi, T.; Uwamizu, A.; Sayama, M.; Ikubo, M.; Otani, Y.; Kano, K.; Makide, K.; Aoki, J.; Ohwada, T. Conformational constraint of the glycerol moiety of lysophosphatidylserine affords compounds with receptor subtype selectivity. *J. Med. Chem.* **2016**, *59*, 3750–3776.
- (14) Chan, D. M.; Monaco, K. L.; Wang, R.-P.; Winters, M. P. New N- and O-arylations with phenylboronic acids and cupric acetate. *Tetrahedron Lett.* **1998**, *39*, 2933–2936.
- (15) Evans, D. A.; Katz, J. L.; West, T. R. Synthesis of diaryl ethers through the copper-promoted arylation of phenols with arylboronic acids. An expedient synthesis of thyroxine. *Tetrahedron Lett.* **1998**, *39*, 2937–2940.
- (16) Inoue, A.; Ishiguro, J.; Kitamura, H.; Arima, N.; Okutani, M.; Shuto, A.; Higashiyama, S.; Ohwada, T.; Arai, H.; Makide, K. TGF α shedding assay: an accurate and versatile method for detecting GPCR activation. *Nat. Methods* **2012**, *9*, 1021–1029.
- (17) Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63*, 1256–1272.
- (18) Zhang, C.; Srinivasan, Y.; Arlow, D. H.; Fung, J. J.; Palmer, D.; Zheng, Y.; Green, H. F.; Pandey, A.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Coughlin, S. R.; Kobilka, B. K. High-resolution crystal structure of human protease-activated receptor 1. *Nature* **2012**, *492*, 387–392.
- (19) Zhang, D.; Gao, Z.-G.; Zhang, K.; Kiselev, E.; Crane, S.; Wang, J.; Paoletta, S.; Yi, C.; Ma, L.; Zhang, W.; et al. Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **2015**, *520*, 317–321.
- (20) Zhang, K.; Zhang, J.; Gao, Z.-G.; Zhang, D.; Zhu, L.; Han, G. W.; Moss, S. M.; Paoletta, S.; Kiselev, E.; Lu, W.; Fenalti, G.; Zhang, W.; Muller, C. E.; Yang, H.; Jiang, H.; Cherezov, V.; Katritch, V.; Jacobson, K. A.; Stevens, R. C.; Wu, B.; Zhao, Q. Structure of the human P2Y₁₂ receptor in complex with an antithrombotic drug. *Nature* **2014**, *509*, 115–118.
- (21) Zhang, J.; Zhang, K.; Gao, Z.-G.; Paoletta, S.; Zhang, D.; Han, G. W.; Li, T.; Ma, L.; Zhang, W.; Müller, C. E.; Yang, H.; Jiang, H.; Cherezov, V.; Katritch, V.; Jacobson, K. A.; Stevens, R. C.; Wu, B.;

- Zhao, Q. Agonist-bound structure of the human P2Y₁₂ receptor. *Nature* **2014**, *509*, 119–122.
- (22) Srivastava, A.; Yano, J.; Hirozane, Y.; Kefala, G.; Gruswitz, F.; Snell, G.; Lane, W.; Ivetac, A.; Aertgeerts, K.; Nguyen, J.; Jennings, A.; Okada, K. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. *Nature* **2014**, *513*, 124–127.
- (23) Lu, J.; Byrne, N.; Wang, J.; Bricogne, G.; Brown, F. K.; Chobanian, H. R.; Colletti, S. L.; Di Salvo, J.; Thomas-Fowlkes, B.; Guo, Y.; Hall, D. L.; Hadix, J.; Hastings, N. B.; Hermes, J. D.; Ho, T.; Howard, A. D.; Josien, H.; Kornienko, M.; Lumb, K. J.; Miller, M. W.; Patel, S. B.; Pio, B.; Plummer, C. W.; Sherborne, B. S.; Sheth, P.; Souza, S.; Tummala, S.; Vonrhein, C.; Webb, M.; Allen, S. J.; Johnston, J. M.; Weinglass, A. B.; Sharma, S.; Soisson, S. M. Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. *Nat. Struct. Mol. Biol.* **2017**, Pub ASAP.
- (24) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534–553.
- (25) Sherman, W.; Beard, H. S.; Farid, R. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* **2006**, *67*, 83–84.
- (26) Hanson, M. A.; Roth, C. B.; Jo, E.; Griffith, M. T.; Scott, F. L.; Reinhart, G.; Desale, H.; Clemons, B.; Cahalan, S. M.; Schuerer, S. C.; Sanna, M. G.; Han, G. W.; Kuhn, P.; Rosen, H.; Stevens, R. C. Crystal structure of a lipid G protein-coupled receptor. *Science* **2012**, *335*, 851–855.
- (27) Sabbadin, D.; Ciancetta, A.; Moro, S. Bridging molecular docking to membrane molecular dynamics to investigate GPCR–ligand recognition: The human A_{2A} adenosine receptor as a key study. *J. Chem. Inf. Model.* **2014**, *54*, 169–183.
- (28) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146–148.
- (29) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.
- (30) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (31) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (32) Negoro, N.; Sasaki, S.; Mikami, S.; Ito, M.; Suzuki, M.; Tsujihata, Y.; Ito, R.; Harada, A.; Takeuchi, K.; Suzuki, N.; Miyazaki, J.; Santou, T.; Odani, T.; Kanzaki, N.; Funami, M.; Tanaka, T.; Kogame, A.; Matsunaga, S.; Yasuma, T.; Momose, Y. Discovery of TAK-875: A Potent, Selective, and Orally Bioavailable GPR40 Agonist. *ACS Med. Chem. Lett.* **2010**, *1*, 290–294.
- (33) Negoro, N.; Sasaki, S.; Mikami, S.; Ito, M.; Tsujihata, Y.; Ito, R.; Suzuki, M.; Takeuchi, K.; Suzuki, N.; Miyazaki, J.; Santou, T.; Odani, T.; Kanzaki, N.; Funami, M.; Morohashi, A.; Nonaka, M.; Matsunaga, S.; Yasuma, T.; Momose, Y. Optimization of (2,3-dihydro-1-benzofuran-3-yl)acetic acids: discovery of a non-free fatty acid-like, highly bioavailable G protein-coupled receptor 40/free fatty acid receptor 1 agonist as a glucose-dependent insulinotropic agent. *J. Med. Chem.* **2012**, *55*, 3960–3974.
- (34) Stanley, N.; Pardo, L.; Fabritiis, G. D. The pathway of ligand entry from the membrane bilayer to a lipid G protein-coupled receptor. *Sci. Rep.* **2016**, *6*, 22639.
- (35) Hurst, D. P.; Grossfield, A.; Lynch, D. L.; Feller, S.; Romo, T. D.; Gawrisch, K.; Pitman, M. C.; Reggio, P. H. A Lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *J. Biol. Chem.* **2010**, *285*, 17954–17964.
- (36) Šali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (37) UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **2015**, *43* (D1), D204–D212.
- (38) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
- (39) Ritscher, L.; Engemaier, E.; Stäubert, C.; Liebscher, I.; Schmidt, P.; Hermsdorf, T.; Römpler, H.; Schulz, A.; Schöneberg, T. The ligand specificity of the G-protein-coupled receptor GPR34. *Biochem. J.* **2012**, *443*, 841–850.
- (40) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221–234.
- (41) Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (42) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. A Hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 351–367.
- (43) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the role of crystal packing forces in determining protein sidechain conformations. *J. Mol. Biol.* **2002**, *320*, 597–608.
- (44) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Yibing, S.; Shaw, D. E. Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*; Tampa, Florida, November 11–17, 2006.
- (45) Tusnády, G. E.; Dosztányi, Z.; Simon, I. Transmembrane proteins in the Protein Data Bank: identification and classification. *Bioinformatics* **2004**, *20*, 2964–2972.
- (46) Tusnády, G. E.; Dosztányi, Z.; Simon, I. PDB_TM: selection and membrane localization of transmembrane proteins in the protein data bank. *Nucleic Acids Res.* **2005**, *33*, D275–D278.
- (47) Kozma, D.; Simon, I.; Tusnády, G. E. PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Res.* **2013**, *41*, D524–D529.
- (48) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (49) Ehlert, F. J.; Griffin, M. T.; Sawyer, G. W.; Bailon, R. A simple method for estimation of agonist activity at receptor subtypes: comparison of native and cloned M3 muscarinic receptors in guinea pig ileum and transfected cells. *J. Pharmacol. Exp. Ther.* **1999**, *289*, 981–992.