Structural basis for receptor selectivity and inverse agonism in S1P5 receptors

The bioactive lysophospholipid sphingosine-1-phosphate (SIP) acts via five different subtypes of SIP receptors (SIPRs) - SIP1-5. SIP5 is predominantly expressed in nervous and immune systems, regulating the egress of natural killer cells from lymph nodes and playing a role in immune and neurodegenerative disorders, as well as carcinogenesis. Several SIPR therapeutic drugs have been developed to treat these diseases; however, they lack receptor subtype selectivity, which leads to side effects. In this article, we describe a 2.2 Å resolution room temperature crystal structure of the human S1P5 receptor in complex with a selective inverse agonist determined by serial femtosecond crystallography (SFX) at the Pohang Accelerator Laboratory X-Ray Free Electron Laser (PAL-XFEL) and analyze its structure-activity relationship data. The structure demonstrates a unique ligand-binding mode, involving an allosteric sub-pocket, which clarifies the receptor subtype selectivity and provides a template for structure-based drug design. Together with previously published SIPR structures in complex with antagonists and agonists, our structure with S1P5-inverse agonist sheds light on the activation mechanism and reveals structural determinants of the inverse agonism in the S1PR family.
Sphingosine-1-phosphate (S1P) is a lysosphingolipid bio-regulator produced from ceramide in activated platelets, injured cells, and cells stimulated by protein growth factors\(^1,2\). S1P is released in the blood\(^3\), where it regulates angiogenesis\(^4\), cell proliferation, migration, and mitosis\(^5\) by activating five subtypes of the S1P G-protein-coupled receptors—S1P\(_1\)–S1P\(_5\). S1P\(_1\) couples only to Gi protein, S1P\(_4\), and S1P\(_5\) signal through Gi and G\(_{12/13}\)\(^6\), and both S1P\(_2\) and S1P\(_3\) couple to Gi, G\(_{12/13}\), and G\(_{q}\). S1P receptors (S1PRs) have different expression profiles—S1P\(_1\)–S1P\(_3\) is expressed in all organs throughout the body, while S1P\(_4\) expression is limited to the immune system, and S1P\(_5\) is predominantly expressed in the nervous (oligodendrocytes) and immune (NK cells) systems\(^8\). S1P\(_5\) also inhibits PAR-1-mediated platelet activation\(^9\). This receptor plays an important role in autoimmune\(^10\) and neurodegenerative disorders\(^10,11\) as well as carcinogenesis\(^12\). For example, S1P\(_5\) agonists elicit neuroprotective effects in Alzheimer’s and Huntington’s diseases\(^10\), while S1P\(_5\) inhibition leads to apoptosis of cancerous NK cells in large granular leukemia (LGL)\(^12\). Non-selective modulators such as fingolimod\(^13\), as well as dual S1P/SIP ligands siponimod\(^14\) and ozanimod\(^13,16\), have been approved for the treatment of multiple sclerosis\(^7\), Crohn’s disease\(^13\), and other autoimmune disorders. However, the exact pharmacological role of S1P\(_5\) remains unclear, mostly due to the lack of well-characterized potent and highly selective S1P\(_5\) ligands with in vivo activity. While inhibition of S1P\(_5\) is considered a prospective treatment for LGL\(^12\), antagonism of S1P\(_1\) leads to serious adverse effects such as lung capillary leakage, renal reperfusion injury, and cancer angiogenesis\(^19\). Therefore, high-resolution structures of S1P\(_5\) in complex with highly selective ligands would shed light on receptor selectivity and provide templates for structure-based design of selective therapeutic drugs with more focused function and fewer side effects.

The first crystal structure of an S1PR was published in 2012\(^20\), revealing the inactive state conformation of the human S1P\(_1\) in complex with a selective antagonist sphingolipid mimetic ML056.
Recently, a crystal structure of S1P3 bound to its endogenous agonist, as well as cryo-EM structures of S1P1, S1P2, and S1P5 in complex with G proteins, and S1P3 in complex with G and β-arrestin, provided insights in the activation mechanism for the S1PR family. However, no structures of this family members in complex with inverse agonists have been reported to date.

In this article, we present the crystal structure of S1P5 in complex with a selective inverse agonist ONO-5430608 determined by serial femtosecond crystallography (SFX) and analyze it alongside structure-activity relationship data from cell-based functional assays using extensive mutagenesis, molecular docking, molecular dynamics, and AlphaFold simulations.

**Results**

**Structure determination using an X-ray free-electron laser (XFEL)**

Human S1P3 receptor was engineered for crystallization by fusing a thermostabilized apocytochrome b562RIL into the third intracellular loop (ICL3) and adding a haemagglutinin signal peptide, FLAG-tag, and a linker on the N-terminus as well as a PreScission Protease site and decahistidine tag on the C-terminus (Supplementary Fig. 1). Crystals of S1P3 bound to an inverse agonist ONO-5430608 were obtained by lipidic cubic phase (LCP) crystallization reaching a maximum size of 30 μm. Our initial attempts at solving the structure using synchrotron data were unsuccessful. Crystals of S1P3 bound to ONO-5430608 were then optimized to grow at a high crystal density with an average size of -5–10 μm and used for room temperature SFX data collection at PAL-XFEL (Supplementary Fig. 2). The crystal structure was solved at a 2.2 Å resolution in the P2_12_12 space group (Supplementary Table 1). A high systematic background scattering from the direct XFEL beam (Supplementary Fig. 3) combined with pseudotranslation led to high structure refinement R-factors, although it did not affect the excellent quality of electron density maps (see Methods and Supplementary Fig. 4). The receptor crystallized with two monomers per asymmetric unit, forming an antiparallel dimer through the TM4-TM4 interface (Supplementary Fig. 5).
Inactive conformation of S1P5 in complex with ONO-5430608

The S1P5 structure in complex with ONO-5430608 shares the classical architecture with other class A α-branched lipid receptors, including a heptahelical transmembrane bundle (7TM), two pairs of disulfide bonds stabilizing extracellular loops 2 and 3 (ECL2 and ECL3), an amphipathic C-terminal helix 8 running parallel to the membrane on the intracellular side, and an N-terminal helix capturing the ligand-binding site. As expected, the receptor is captured in the inactive conformation (Fig. 1a, b) based on its overall alignment with the inactive state S1P1 (PDB ID 3V2Y, Ca RMSD = 0.84/0.78 Å on 90% of residues for chains A/B of our S1P5 structure) and the active state S1P2 (PDB ID 7EW1, Ca RMSD = 1.40/1.40 Å on 90% of residues for chains A/B of our S1P5 structure) as well as on the conformation of conserved activation-related motifs described below.

A dual toggle switch L(F)3.36-W6.48 (superscripts refer to the Ballesteros-Weinstein residue numbering scheme in class A GPCRs) together with L3.36-W6.48.48 motif have been characterized as the classical microswitches in class A GPCRs that transmit activation-related conformational changes from the ligand-binding pocket towards an outward movement of TMs 5 and 6 and inward displacement of TM7 on the intracellular side. In all S1PRs, the dual toggle switch is conserved as L3.36-W6.48; however, the P-I-F motif deviates from the consensus, especially in S1P5, it is represented as I5.50-V3.40-F6.44 (Fig. 1e). Nevertheless, the I-F-V motif in S1P5 apparently serves a similar role as the classical P-I-F motif in other receptors, as the side chains of V4.40 and W6.48 and the shift of W264.48 are accompanied by a rotamer switch of L1193.36 (Fig. 1c). Similar dual (also known as “twin”) toggle switch L(F)3.36-W6.48 has been shown to play a key role in the activation of several other receptors, such as CB1, AT1, and MC4.

An allosteric sodium-binding site located in the middle of the 7TM bundle near D2.50 is highly conserved in class A GPCRs. Binding residues, such as D822.50, S1223.39, and N2987.45, we could not observe inside the ligand-binding pockets of both receptor structures, and this was also the case for ONO-5430608. Thus, mutations of any of the three amino acids in S1P5 significantly (by an order of magnitude or more) decrease the response for both ONO-5430608 and S1P. Although the locations of residues, known to interact with the phosphate group of S1P, the core tetrahydrobenzazepine rings fill in space in the middle of the pocket, while the naphthyl-ethoxy group unexpectedly swings over and extends into a previously unidentified allosteric sub-pocket. The subpocket is surrounded by non-conserved residues from TM1, TM2, and TM7 and opened in our structure due to a rotamer switch of Y99.47 compared to structures of other S1P receptors (Fig. 2a, b). The distinct amino acid composition of this allosteric subpocket suggests that it can serve as a selectivity determinant for S1P-specific ligands and makes the hallmark of the structure described in this work.

Functional characterization of the ligand binding hotspots in S1P5

To validate the observed ligand binding pose and further expand our knowledge about the ligand selectivity and relative importance of specific residues, we tested 25 structure-inspired ligand-binding pocket mutants of S1P5 by a BRET-based CAMP production assay using the endogenous agonist S1P and the co-crystallized inverse agonist ONO-5430608 (Fig. 2c, Supplementary Table 2, and Supplementary Fig. 6). In line with the binding pocket structure description given above, we consequently characterize important interactions in each part.

In S1P5, the polar upper part of the binding pocket is highly conserved among the whole S1PR family (Fig. 2b). It consists of residues Y19.52, K24.47, N92.60, R111.36, and E123.29 and accommodates the phosphate and primary amine groups of S1P. The receptor’s potential for multiple polar interactions in this region is utilized in anchoring zwitterionic groups of synthetic ligands of S1PRs. Thus, in our S1P5 structure, the carboxyl group of ONO-5430608 is stabilized by polar interactions with Y19.52, K24.47, and R111.36, while the protonated tertiary amine group makes a salt bridge with E123.29, similar to interactions of the phosphate and secondary amine groups of ML056 in S1P5 structure. The zwitterionic headgroup of the endogenous S1P ligand bound to S1P5 is shifted towards TM1 while retaining the same interactions except for the N-terminal K27.

The mutations disrupting polar interactions with zwitterionic ligand head groups: Y19.52/E123.29/A/F, K24.47/ R111.36/A/Q, N92.60/A/C, and E123.29/A/Q either fully abolish or significantly (by an order of magnitude or more) decrease the response for both ONO-5430608 and S1P (Fig. 2c and Supplementary Table 2). Notably, some mutations have different effects on S1P and ONO-5430608. While mutations of N92.60, R111.36, and E123.29 completely eliminate response to the inverse agonist, they only decrease the potency for S1P. A similar effect of mutations of homologous amino acids on S1P potency was previously observed for S1P3. In this case, each of the three amino acids independently interacts with the amine group of S1P (see PDB ID 7C4S). On the other hand, in our S1P5-ONO-5430608 structure, these three amino acids are interconnected and form a stable cluster that further interacts with the tertiary amine and the carboxyl group of ONO-5430608. Thus, mutations of any of the three amino acids in S1P5 would only partially perturb S1P complex, while they would disrupt the cluster and completely eliminate the binding of ONO-5430608.

Although the locations of residues, known to interact with the phosphate group of S1P from either functional or structural data, are largely conserved between S1P receptors, the effects of their mutations on S1P potency are different. Namely, mutations of N-terminal Y29.19 and K34.24 to alanine render S1P/S1P5, respectively, non-responsive to

---

Nature Communications | (2022) 13:4736
https://doi.org/10.1038/s41467-022-32447-1
SIP, while corresponding mutations preserve the interaction with SIP. These data suggest a different orientation of the phosphate headgroup of SIP within the binding pocket in different receptors.

The hydrophobic part of the orthosteric binding pocket in S1PRs accommodates the lipidic tail of the endogenous ligand or its synthetic analogs such as ML056 and ONO-5430608. The residues on its bottom are conserved among S1PRs (Fig. 2b) and well characterized. The top part of the hydrophobic subpocket in SIP, which in our structure accommodates the tetrahydrobenzazepine double-ring system of ONO-5430608, consists of residues V115, L292, and Y89 as less characterized, although they play an important role in ligand binding. In particular, mutations L292A/V decrease the selectivity of ligands targeting it. Interestingly, this pocket is present in S1P receptors utilizing an enhanced molecular dynamics simulation technique, originally developed by Laio and Parrinello and known as metadynamics (metaMD), as well as by targeted mutagenesis.

Selectivity of ligands targeting it. Interestingly, this pocket is present in S1P receptors utilizing an enhanced molecular dynamics simulation technique, originally developed by Laio and Parrinello and known as metadynamics (metaMD), as well as by targeted mutagenesis.

Fig. 3 | Conformational flexibility of Y2.57 and its effect on inverse agonism in SIP receptors. (a) Two distinct upward and downward conformations of Y2.57 as observed in crystal structures of S1P-ML056 (PDB ID 3YX7) and S1P-ONO-5430608 (this work, PDB ID 7YXA), respectively. (b) The downward orientation of Y2.57 is incompatible with the active state of the dual toggle switch L3.36-W6.48 because of a steric clash. S1P-ONO-5430608 (this work, PDB ID 7YXA, inactive state) is shown in pink, and SIP-siponimod (PDB ID 7EW1, active state) is shown in purple. C free energy profiles of the Y2.57 side chain torsion angle χ1 as calculated by metaMD for S1P, S1P, and S1P. Dotted lines correspond to Y2.57 conformations in corresponding experimental structures. C Free energy profiles of the L1.36 side chain torsion angle in S1P with two alternative orientations of Y2.57 as calculated by metaMD. Dotted lines correspond to L1.36 conformations in corresponding experimental structures.
Insights from molecular docking

To assess the importance of the Y892.57 conformation in ligand binding, we performed molecular docking of ONO-5430608 ligand series \cite{25} (Fig. 4a) into two S1P5 models: the crystal structure (Y892.57 in the downward conformation) and a metaMD snapshot (Y892.57 in the upward conformation). As expected, the docking scores correlate well with the ligand affinity \cite{23}, only in case of the crystal structure (Fig. 4b, c): the most potent group ‘A’ ligands (IC_{50} between 1 and 100 nM) have docking scores of \(-37 \pm 5 \text{ kJ/mol}\), whereas the least potent group ‘C’ ligands (IC_{50} between 1 and 3 \mu M) have scores \(-23 \pm 4 \text{ kJ/mol}\), and for the intermediate group ‘B’ (IC_{50} between 100 nM and 1 \mu M) scores are \(-32 \pm 5 \text{ kJ/mol}\). For the metaMD snapshot, scores show no correlation with the ligand affinity (Fig. 4d, e). Accordingly, ligand docking poses also confirm that Y892.57 needs to be in a downward conformation for the ONO-5430608-like compounds to adopt conformations similar between each other. Namely, all of group ‘A’ compounds closely resemble the co-crystallized ligand pose (Fig. 4b, c). They retain interactions of the negatively-charged headgroup with Y19/K24, as well as the interaction of the positively charged amino group with E1123.29, and the position of the double-ring system is preserved. For the upward confirmation of Y892.57, the docking of the group “A” ligands show no consistency between each other and the obtained data from the mutation screening (Fig. 4d, e).

Notably, the SAR data for the ONO-5430608 ligand series (Fig. 4a) suggest a role of the substituent position on the core double-ring system in the ligand binding. Namely, most of the lower affinity ligands (group ‘C’) have a tetrahydroisoquinoline or tetrahydronaphthalene scaffold instead of the tetrahydrobenzazepine, which is more common among group ‘A’ and ‘B’ ligands. Likely, the affinity drop occurs due to the overall ligand shape, rather than the ring size. Namely, most ligands with both substituents placed on the same side of the middle...
plane across the double-ring system display higher affinity, while all ligands with two substituents placed on different sides have a low affinity (Fig. 4a), with the only exception, Example 18-2, which however has an amine group placed within the isoquinoline system, compared to other same-side substituents ligands. This notion also suggests a common framework for designing S1P5-selective ligands.

Structural insights into inverse agonism

It has been shown that S1P5 exhibits a relatively high level of basal activity\(^4\), while our functional assay revealed that ONO-5430608 acts as an inverse agonist for the G\(_i\)-protein-mediated signaling pathway, reliably decreasing the basal activity level detected by the BRET-based cAMP sensor (Supplementary Fig. 6).

Our S1P5 structure in complex with the inverse agonist ONO-5430608 along with previously reported agonist and antagonist-bound structures of S1P1,3,5 shed light on the mechanism of inverse agonism. Specifically, the above-mentioned conformational flexibility of Y892.57 may provide a structural background for the basal activity of S1P5. We used metaMD to estimate free energy profiles along the reaction coordinate corresponding to the torsion rotation of the L1193.36 side chain for S1P5 with Y892.57 restrained in the upward and downward orientations. The upward orientation of Y892.57 is compatible with both active and inactive conformations of the dual toggle switch 1.36-W6.48, while the downward orientation of Y892.57 selects the inactive conformation (Fig. 3d). The dual toggle switch is found in the previously reported active state structures of S1P3-siponimod as well as in S1P1,23 and S1P3,21 agonist-bound complexes. It induces activation of the P-I-F motif and an outward movement of the intracellular part of TM6 resulting in G-protein signaling cascade. On the other hand, the dual toggle switch is observed in the inactive conformation in our S1P5-ONO-5430608 structure and in the previously published antagonist-bound S1P1,20. The inverse agonist ONO-5430608 induces the downward conformation of Y892.57 that opens the allosteric subpocket and suppresses the switching of L1193.36 locking the dual toggle switch in the inactive state (Fig. 3b). Therefore, the conformational flexibility of Y892.57 in S1P5 provides a structural basis for both receptor subtype selectivity and inverse agonism.

Naturally occurring mutations in S1P5

In order to characterize additional functionally important residues in S1P5, we performed mapping of known point mutations from genomic databases onto the crystal structure (Fig. 2d). Multiple databases carry information about S1P5 point mutations including gnomAD (229 SNVs)\(^4\), which contains genomic information from unrelated individuals, and COSMIC (124 point mutations)\(^4\), which accumulates somatic mutations in cancer. The most frequent gnomAD mutation points correspond to Y2.57 conformations in AlphaFold models. g, h Performance of existing experimental structures versus the best AlphaFold structure in virtual screening of compounds from three benchmark tests (ONO, selective, and non-selective), as judged by their AUC and enrichment (top 10%), respectively. Bar heights represent mean ± 95% CI for \(n = 3\) independent docking trials with effort = 1.

---

Fig. 5 | AlphaFold prediction of S1PR structures. a–e Conformations of Y2.57 and the dual toggle switch L1.36-W6.48 in 50 AlphaFold models of each S1P1–5 subtype, respectively. Three distinct conformations of S1P5 are shown in dark violet, light violet, and black. Free energy profiles of the Y2.57 side chain torsion angle \(\chi_1\) in S1P1 (yellow line), S1P3 (blue line) and S1P5 (pink line) calculated by metaMD. Dotted lines correspond to Y2.57 conformations in experimental structures. g, h Performance of existing experimental structures versus the best AlphaFold structure in virtual screening of compounds from three benchmark tests (ONO, selective, and non-selective), as judged by their AUC and enrichment (top 10%), respectively. Bar heights represent mean ± 95% CI for \(n = 3\) independent docking trials with effort = 1.
L318S/Q in helix 8 (3% of the population) was shown to impair G12
signaling; however, according to our functional data it only slightly
decreases the potency of SIP in G1-mediated signaling (Fig. 2c). It was
previously proposed43 that a possible cause of this mutation on the
signaling impairment is the prevention of palmitoylation of the
downstream C322A or C323K.Fr A concomitant cause might be a
shift in the helix 8 position due to the loss of a hydrophobic contact
between the mutated residue L318S/Q and the membrane.

Several individuals have missense mutations in the ligand binding
pocket; for example, 3 out of 235,080 samples43 contain R111L-F
mutation possibly affecting the contacts with the zwitterionic ligand
headgroup (Fig. 2). Mutation of another headgroup recognition resi-
due, E1123.G, is less frequent (1 of 234,568). As shown in our functional
data, mutations of both of these residues to neutral ones disrupt
response to ligands (Fig. 2c). Additionally, two mutations are located
side-by-side in the putative ligand entrance gateway, C43-F and,
M296-V, are present in the population43 at 10
times frequencies. While
C43-F shows little effect in our functional tests, M296-V disrupts G1
signaling response for both SIP and ONO-5430608 (Fig. 2c). Another
conserved in SIP receptors, except for SIP2, residue A295-G54 has a
hydrophobic contact with the ligand (Fig. 2a), which becomes altered
in case of the A295S mutation. Mutation of A295S may also directly influence the state of the toggle switch (R1113.K243W64A in
SIP3) and may interfere with protein activation, as it was shown for
several other receptors, e.g., β2-adrenergic receptor46 and CCR5.47 One
of the key residues in the sodium-binding site, N2984, has several
variations in population: S, D, or K. While the effects of N2984-D and
N2984-S are unclear, N2984-K would mimic sodium-binding, stabi-
zizing the inactive state of the receptor.48

Somatic mutations appearing in COSMIC and not found in the
population may be linked to severe cancer impairments. For example, S125R disrupts the conserved hydrogen-bond network involving
S772 and W1594, destabilizing contacts between TMs 2, 3, and 449 and,
likely, disturbing the 7TM fold due to the introduction of a charged
residue in a mostly hydrophobic environment.

Comparison with AlphaFold predictions
Recently, a redesigned artificial intelligence-based protein
structure-predicting system AlphaFold v.2 achieved a notable
breakthrough in approaching the accuracy in protein structure
modeling, previously available only from experimental methods.
AlphaFold-based approaches started to find multiple applications in
structural biology5, however, their full capacity and limitations
remain to be uncovered. Here, we evaluated the ability of AlphaFold
to predict structural features responsible for receptor selectivity
and inverse agonism in the SIP family. For that, we generated 50 de
novo AlphaFold models for each of the five SIPs without using
existing structures as templates. Overall, the models demonstrated
reasonable correspondence to the available experimental struc-
tures; for example, Ca RMSDs in the 7TM region between the SIP3
models and the inactive state crystal structure (SIP3-ONO-5430608)
is 1.3 ± 0.2 Å and the active state structure (PDB ID 7EW1, SIP3-
Siponimod) is 3.0 ± 0.2 Å.

The conformational heterogeneity of Y(F)2.L7 observed in experi-
mental SIPR structures and metaMD simulations were also well cap-
tured by AlphaFold predictions (Fig. 5a–f). In all SIP, SIP3, and SIP4
models, Y2.L has an upward conformation, except for a single SIP
model, in which this residue adopts a downward orientation similar
to that previously observed in all atom MD simulations5. Furthermore, 19 out of 50 SIP3 models display a downward Y2.L orientation, while all the others have an upward Y2.L orientation. Notably, SIP2 is the only receptor, in which Y2.L is replaced with F2.57 which adopts a downward conformation in all generated models. The downward orientation of F2.57 in SIP2, similar to that of Y2.L in SIP3, opens the allosteric sub-
pocket, which may be targeted to achieve ligand selectivity.

In all available experimental SIPR structures, the conserved dual
toggle switch L1.3.K-W6.48 displays either active or inactive conformation. AlphaFold predicted both of these conformations for all receptors
except for SIP1, in which only the active conformation was present
in all models (Fig. 5a–e). However, AlphaFold models did not fully reflect the mutual relationship between conformations of Y957-L1193.36
and observed by metaMD in S1P3. Thus, all AlphaFold-predicted SIP3
models cluster into three groups (Fig. 5e), including the energetically
unfavorable conformation with Y957-L1193.36 in downward-upward
orientations while missing the energetically favorable conformation
with Y957-L1193.36 in upward-downward positions. Consequently, we
conclude that the current version of AlphaFold could not consistently
generate an SIPR structure in a specific signaling state, sometimes
mixing the features of different conformations in a single model. These
findings are corroborated by a recent study of several other GPCRs.53

One of the most intriguing AlphaFold-related questions is how
useful the predicted models are for structure-based drug design44. To
test it in application to SIP3 targets, we constructed three bench-
marks, mimicking virtual ligand screening campaigns, and compare
the available experimental structures and AlphaFold models by their
ability to distinguish high-affinity ligands from low-affinity binders
decays. Our results demonstrated that crystal structures outperform
AlphaFold-generated models in several scenarios (Fig. 5g, h and Sup-
plementary Fig. 7). Namely, our SIP3 crystal structure showed sub-
stantially better overall ranking and top-10% enrichment among both
ONO-5430608-like inverse agonists (“ONO” benchmark) and SIP3-
selective ligands (“Selective” benchmark). In the case of the non-
selective ligand benchmark (mostly SIP3 agonists), the best per-
formance was achieved for several experimental SIP3 structures deter-
mixed in complex with non-selective ligands, e.g., SIP3-Siponimod
complex (Fig. 5g, h), while our SIP3 structure fared on par with
AlphaFold models.

Discussion
Here, we present the 2.2 Å crystal structure of the human SIP3 receptor
in complex with its selective inverse agonist. The structure was
obtained by room temperature SFX data collection at PAL-XFEL using
sub-10 μm crystals. In combination with site-directed mutagenesis,
functional assays, metaMD simulations, and docking studies, this
structure revealed molecular determinants of ligand binding and
selectivity as well as shed light on the mechanism of inverse agonism in
the SIPR family. The obtained structure also allowed us to map loca-
tions of known missense SNVs from gnomAD and COSMIC genome
databases and annotate their potential functional roles providing
future insights into personalized medicine approaches.

We found that the inverse agonist ONO-5430608 binds to the
receptor’s orthosteric site, suppressing SIP3 basal activity. Highly
conserved residues Y957(GP), K243(K5), R1113(K5), and E1123.G5 play
an essential role in the recognition of both ONO-5430608 and its native
ligand SIP. The naphthyl group of ONO-5430608 occupies an allosteric
subpocket that was not previously observed in any other SIPR
structure. While the orthosteric site is highly conserved in the SIPR family,
the allosteric subpocket is composed of unique residues and is present
in our SIP3 structure due to the conformational switch of a single residue V5.27 Functionally important residues were revealed by
structure-guided site-directed mutagenesis and G1 signaling assays.
We further used metaMD simulations to explore the conformational
flexibility of V5.27 in SIP3 and established its role in receptor subtype
selectivity and inverse agonism. The role of V5.27 in the binding of
selective ligands was also confirmed by comparative molecular docking
simulations. Furthermore, taking advantage of the availability of
several experimental structures of SIPRs in different functional states,
we tested the ability of AlphaFold to predict de novo specific con-
formational states for SIP3 and to provide reliable templates for
structure-based virtual ligand screening. While the AlphaFold-
generated models showed a close similarity to experimental structures and captured conformational diversity of conserved structural motifs, the models did not provide a full description of specific signaling states and showed subpar performance in virtual ligand screening compared to experimental structures.

Our structure along with our functional and computer modeling data may facilitate the rational design of ligands that could further serve as lead or tool compounds for detailed elucidation of biological function of S1P₅, and therapeutic developments. S1P₅ is emerging as a promising drug target. Inhibiting S1P₅ by an inverse agonist could create new therapeutic strategies against neuroinflammation and degeneration where the high ligand selectivity would diminish the off-target effects. While S1P₅ has a broad expression profile, S1P₅ is expressed predominantly in brain tissues; thus, a highly selective compound would afford more localized control over associated CNS disorders not affecting peripheral processes in the body.

**Methods**

**Protein engineering for structural studies** The human wild-type gene S1PR5 (UniProt ID Q9H22S) was codon-optimized for Gibson for insect cell expression and modified by adding a hemagglutinin signal peptide (HA; KTIIALSYIFCLVFA), a FLAG-tag for expression detection, and an Ala-Gly-Arg-Ala linker at the N-terminus. An apoyctochrome b562RIIL (BRIL) was inserted in the third intracellular loop between A223 and R241 to stabilize the receptor and facilitate crystallization. The C-terminus was truncated after Val321, and a PreScission cleavage site was added after it to enable the removal of the following 10 His tag used for IMAC purification (Supplementary Fig. 1). The resulting construct was cloned into a pFastBac1 (Invitrogen) plasmid. The full DNA sequence of the S1P₅ crystallization construct is provided in Supplementary Table 3.

**Protein expression** Using the Bac-to-Bac system (Invitrogen), a high titer (10⁸ particles per ml) virus encoding the crystallization construct was obtained. Sf9 (Novagen, cat. 71104) cells were infected at a density (2–3) × 10⁵ cells per ml and a multiplicity of infection (MOI) 4–8, incubated at 28 °C, 120 rpm for 50–52 h, harvested by centrifugation at 2,000 g and stored at −80 °C until further use.

**Protein purification** Cells were thawed and lysed by repetitive washes (Dounce homogenization on ice, and centrifugation at 128,600 g for 30 min at 4 °C) in hypotonic buffer (10 mM HEPES pH 7.5, 20 mM KCl, and 10 mM MgCl₂) and high osmotic buffer (10 mM HEPES pH 7.5, 20 mM KCl, 10 mM MgCl₂, and 1 M NaCl) with an addition of protease inhibitor cocktail [PIC; 500 μM 4-(2-aminomethyl)benzenesulfonyl fluoride hydrochloride (Gold Biotechnology), 1 μM E-64 (Cayman Chemical), 1 μM leupeptin (Cayman Chemical), 150 μM aprotinin (A.G. Scientific)] with the ratio of 50 μl per 100 ml of lysis buffer. Membranes were then resuspended in 10 mM HEPES pH 7.5, 20 mM KCl, 10 mM MgCl₂, 2 mM iodoacetamide, PIC (100 μl per 50 ml of resuspension buffer), and 25 μM ONO-5430608 (4-{6-[2-(1-Naphthyl)ethoxy]-4-methylcoumarin, Invitrogen}) dissolved in DMF at 10 mM. This CPM stock solution was diluted to 1 mM in DMSO and then added to working buffer at 10 μM. 1 μg of the target protein was added to 50 μl of working buffer (25 mM HEPES, 250 mM NaCl, 10 mM imidazole, 0.05/0.01%v/v DDM/CHS) with CPM, and the melting curve was recorded on a Rotor-Gene Q real-time PCR cycler (Qiagen) using a temperature ramp from 28 to 98 °C with 2°C/min rate. The fluorescence signal was measured in the Blue channel (excitation 365 nm, emission 460 nm), and the melting temperature was calculated as the maximum of the fluorescence signal derivative with respect to temperature.

**LCP crystallization** Purified and concentrated S1P₅ was reconstituted in LCP, made of monooolein (Nu-Chek Prep) supplemented with 10%w/v cholesterol (Affymetrix), in 2:3 (v/v) protein:lipid ratio using a syringe lipid mixer. The obtained transparent LCP mixture was dispensed onto 96-wells glass sandwich plates (Marienfeld) in 40 nl drops and covered with 900 nl precipitant using an NTS-LCP robot (Formulatrix) to grow crystals for synchrotron data collection. To prepare crystals for XFEL data collection, the protein-laden LCP mixture was injected into 100 μl Hamilton gas-tight syringes filled with precipitant as previously described. All LCP manipulations were performed at room temperature (20–23 °C), while plates and syringes were incubated at 22 °C. Crystals of S1P₅ grew to their full size of ~30 μm (in plates) or <10 μm (in syringes) within 3 days in precipitant conditions containing 100–300 mM KH₂PO₄ monobasic, 28–32%v/v PEG400, and 100 mM HEPES pH 7.0.

**Diffraction data collection and structure determination** XFEL data for S1P₅-ONO-5430608 crystals were collected at the NCI (Nanocrystallography and Coherent Imaging) beamline of the Pohang Accelerator Laboratory X-ray Free Electron Laser (PAL-XFEL), Pohang, South Korea. The PAL-XFEL was operated in SASE mode at the wavelength of 1.278 Å (9.7 keV) and 0.2% bandwidth, delivering individual X-ray pulses of 25-fs duration focused into a spot size of 2 × 3 μm using a pair of Kirkpatrick-Baez mirrors. LCP laden with dense suspension of protein microcrystals was injected at room temperature inside a sample chamber filled with helium (23 °C, 1 atm) into the beam focus region using an LCP injector with a 50-μm-diameter capillary at a flow rate of 0.15 μl/min. Microcrystals ranged in size from 5 to 10 μm. Diffraction data were collected at a pulse repetition range of 30 Hz with a Rayonix MX225-HS detector, operating in a 4 × 4 binning mode (1440 × 1440 pixels, 30 fps readout rate). The beam was not attenuated and delivered full intensity (5 × 10⁸ photons per pulse). A total number of 490,000 detector images were collected. Due to a high systematic background, Cheetah v. 2019-1 was initially used only to apply dark
current calibration, and all images were used for further processing. The overall time of data collection from a sample with a total volume of about 36 μl was approximately 4 h and yielded 6918 indexed frames with 7492 crystal lattices.

During the XFEL data collection, a high systematic background scattering from upstream to the interaction point occurred due to a high-intensity X-ray lasing conditions (Supplementary Fig. 4). Matplotlib v.3.3.2 was used for (a) radial average of the scattered intensity), which prevented from establishing suitable Cheetah hit finding parameters during the beamtime and complicated the overall data processing. All data processing was performed using CrystFEL v. 0.8.0. Here we describe steps that we took to improve data quality as much as possible starting from the available data with a high background level. For all CrystFEL runs (Supplementary Table 4), peak search was limited with max-res = 340, min-res = 50 to search for peaks in the region between the beamstop and the LCP ring, and the frames were limited to a 12,000 subset of all frames, selected with minimum 5 in the region between the beamstop and the LCP ring, and the frames search was limited with max-res = 340, min-res = 50 to search for peaks ground level. For all CrystFEL runs (Supplementary Table 4), peak determination statistics are shown in Supplementary Methods 1.8.6 in Ref. 51 for details) without any further modifications (Supplementary Data file 1).

Spot integration parameters had the biggest impact on the merged data quality. First, changing the spot integration model from rings-nograd model, which assumes flat background around a spot, to rings-grad, which performs 2D-fitting of each spot background profile, decreased overall Rmerge from 29.7% to 19.4% (Supplementary Table 4 column D) and increased the highest resolution shell CC* from 0.618 to 0.666. Second, increasing local-bg-radius from 3 to 5, and using int-radius = 3.5, instead of default 4.5, 8 further improved data quality with the highest resolution shell CC* equal to 0.716 (Supplementary Table 4, columns E-F). Following reviewer’s suggestions, we attempted to improve overall data resolution via applying partiality modeling (column C), less aggressive push-res option with or without-overpredict option (columns H and I, respectively). None of these strategies yielded better results than the initial processing (column G). The final merging was performed with partialiatior, iterations = 2, push-res = 3.0, and model = gppm (Supplementary Table 1).

The structure was initially solved by molecular replacement using phenix.phaser60 with two independent search models of the polyalanine SIP, 7TM domain (PDB ID 3V2Y) and BRIL from the high-resolution A2aAR structure (PDB ID 4EHI). Model building was performed by cycling between manual inspection and building with Coot v. 0.9.6 using both 2mFo-DFc and mFo-DFc maps and automatic refinement with phenix.refine v. 1.19.2 using automatic torsion angle NCS restraints and 2 TLS groups. Ligand restraints were generated using the web server GRADE v. 1.2.19 (http://grade.globalphasing.org). The SIP2 structures from two molecules A and B in the asymmetric unit show very high similarity (Ca RMSD 1.0 Å within 7TM; 1.3 Å all-atom RMSD). The main difference includes flexible ECL1 and conformations of several side chains exposed to the lipid bilayer and solvent. The final data collection and refinement statistics are shown in Supplementary Table 1. The relatively high Rfree of the structure can be partially explained by the high systematic background scattering and modulations of the diffraction intensities. The modulations are produced by two factors: (1) the NCS operator (x, y, z) → (1/8 + x, y, -z) seen as a Patterson peak at (3/8, 1/2, 0) with a 0.3 of the origin peak height, and (2) the lattice-translocation defect (LTD) seen as a Patterson peak at (1/4, 0, 0) with a 0.1 of the origin peak height. We corrected our data partially for LTD as described previously, which resulted in a Rfree drop by 0.6% during the refinement. The final resolution cutoff was determined by paired refinement.

AlphaFold predictions
Prediction runs were executed using AlphaFold v. 2.11 + 110948 with a non-docker setup (https://github.com/kalininlab/alphafold_non_docker, git commit 7ccdb7) and an updated run alphafold.sh wrapper with added -random-seed parameter. The use of structural templates was disabled by setting “max_template_date” to 1900-01-01, thus, no SIRP structures were used for prediction, and all AlphaFold models analyzed in this work were constructed based on multiple sequence alignment alone. 50 AF2-models (ranked...,pdb models) were generated for each of 5 human SIRPs with protein sequences obtained from UniProt. For each receptor, 10 prediction runs with different seeds (random-seed = <run number>) were executed; each run generated five models. Structures were used as provided by the AlphaFold’s pipeline with Amber relaxation (see Supplementary Methods 1.8.6 in Ref. 51 for details) without any further modifications (Supplementary Data file 1).

MD simulations
Molecular dynamics simulations were conducted for the wild-type homologs SIP1, SIP3, and SIP5 receptors based on the X-ray structures 3V2Y (residues V16-K300), 7CAS (G14-R311), and the structure reported in the present study (S12-C323), respectively. All engineered mutations were reverted back to the WT amino acids, and all missing fragments were filled using Modeller v. 9.24. Receptors were embedded into lipid bilayers consisting of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (POPC) lipids and solvated with TIP3P waters and Na+/Cl− ions (to guarantee the electroneutrality of the systems and the ionic strength of 0.15 M) by means of the CHARMM-GUI web-service. The obtained in this way starting models (with 61,666/61,763/56,303 atoms including 119/117/123 POPC molecules in the SIP/SIP/SIP, systems, respectively) were subject to standard CHARMM-GUI minimization and equilibration protocol, i.e., the steepest descent minimization (5000 steps) was followed by a series of short equilibration simulations in the NPT ensemble using Berendsen thermostat and barostat with the restraints on protein and lipids gradually released.

We employed a metadynamics (metaMD) approach to estimate free energy profiles along the rotation of the Kh torsion angle in the side chain of Y2.57 in SIP1, SIP5, and SIP5 as well as free energy profiles along the rotation of the Kh torsion angle in the side chain of L1.14 in SIP1, with two alternative orientations of Y2.57. This method is based on the addition of biasing repulsive potentials (‘hills’, typically Gaussians) to the total potential of the system to enhance the sampling of the configurational space along the chosen reaction coordinates. The deposition rate for hills in metaMD simulations was 1 ps; the width and height of deposited hills were equal to 0.1 rad (-5.7°) and 0.5 kJ/mol, respectively. The metaMD simulations were run for 10 ns each. Two conformations corresponding to the free energy minima along the rotation of the Kh torsion of Y2.57 in SIP1 were selected for the subsequent metaMD simulations of L1.14, in which the orientation of Y2.57 was harmonically restrained in the upward or downward positions. To test for convergence of the metaMD simulations, we applied the following method: the free energy difference between two regions of the obtained free energy profiles, corresponding to the crystallographic orientations of Y2.57 (Supplementary Fig. 8a-c) or to the orientations of L1.14 in the active and inactive SIP3 structures (Supplementary Fig. 8d, e) as a function of simulation time was plotted. In case of convergence, this difference should not change with the progress of simulations as the systems diffuse freely along the reaction coordinate.

For the metaMD simulations, Nose–Hoover thermostat and Parrinello–Rahman barostat were used. The temperature and pressure were set to 323.15 K and 1 bar with temperature and pressure coupling...
time constants of 1.0 ps⁻¹ and 0.5 ps⁻¹, respectively. All MD simulations were performed with GROMACS⁷⁹ v. 2020.2 using PLUMED plugin⁸⁰ to enable metaMD. The time step of 2 fs was used for all production simulations. The CHARMM36 force field⁸¹ was used for the proteins, lipids, and ions.

**SAR and molecular docking**

For SIP₃ docking studies, we used chain B from our SIP₃-ONO-5430608 crystal structure and a metaMD snapshot with the upward conformation of Y892.57. Chain B was selected based on the quality of 2mFo-DFc maps around the ligand and surrounding residues. Molecular docking was performed using ICM Pro v. 3.9-1b (Molsoft, San Diego). We removed ligands and converted the receptor models into an ICM format using default settings, which includes building missing side chains, adding hydrogens, energy-based Gln/Asn/His conformation optimization, and removal of all water molecules. The same docking box was selected for both models, aligned by their 7TM domains, to encompass both orthosteric and allosteric binding pockets. For each ligand we repeated docking runs 5 times with the effort parameter (ligand sampling depth) set at 16, each time saving three best conformations. Ligand structures and their affinities (IC₅₀ values from radioligand binding assays) at SIP₃ receptors were taken from the published patent⁸⁰.

In the AlphaFold models analysis, 50 SIP₃ models predicted by the AlphaFold algorithm were compared with both chains of our SIP₃ crystal structure and other available S1PR crystal structures. All structures were prepared as described above. SIP₃ ligands from ChEMBL⁷² v. 29 were accessed via the web-interface (https://www.ebi.ac.uk/chembl/) using the SIP₃’s ChEMBL target ID. Ligands were converted to 3D and charged at pH 7.0 using Molsof ICM. For each model, ligand screening was performed three times with docking effort 1. Three ligand benchmarks (Supplementary Fig. 7) were used: 1. “ONO” series: active molecules from ref. 25 (group A, 1 nM < IC₅₀ < 100 nM), inactive molecules from ref. 25 (group C, 1 μM < IC₅₀ < 3 μM) and decoys; 2. “Selective” series: active molecules from refs. 25, 27, 37 (group A or IC₅₀ < 100 nM), correspondingly, inactive molecules from refs. 25, 77 (group C or IC₅₀ > 3 μM, correspondingly), and decoys; 3. “Non-selective” series: active molecules from ChEMBL (pChembly < 7.0, mostly SIP₃ agonists), inactive molecules from ChEMBL (pChembly < 5.0), and decoys. Decay molecules were selected from the Enamine REAL library [https://enamine.net/compound-collections/real-compounds/real-database], matching the distribution of active molecules by charge and weight. The benchmarks have the following ratios of active/inactive decoy molecules: 6:5:60 for “ONO”, 12:10:120 for “Selective”, and 15:39:1207 for “Non-selective”, with the imbalance parameter (ratio of the total library size to the number of active molecules in it) of 11.8, 11.8, and 8.8, respectively. Docking scores and ligand structures are provided in Supplementary Data file 2. For estimation of the virtual screening quality, metrics enrichment at 10% and receiver operating characteristic (ROC)–area under the ROC curve (AUC) were used, as implemented in RDKIT⁷⁴ v. 2021-03-4. Data were calculated using the three-parameter dose–response curve fit in GraphPad Prism v. 9.3. Three independent experiments were performed in triplicate.

**Functional assays with BRET-based cAMP sensor**

G protein-mediated signaling responses to endogenously agonist SIP and inverse agonist ONO-5430608 were assayed for human WT and mutant SIP₃ receptors using Bioluminescence Resonance Energy Transfer (BRET) based EPAC biosensor⁸². Briefly, transfections were carried out by Lipofectamine 3000 according to standard protocol using HEK293T cells seeded in a 100 mm cell culture plate. HeLa cell cultures were transfected separately by 3 μg of each expression plasmids based on pcDNA3.1(+) vector using common Lipofectamine 3000 protocol. After 12–18 h incubation in a CO₂ incubator at 37 °C for receptor expression, the cell culture plates were placed on ice, the media was aspirated completely, and the cells were washed once with ice-cold TBS to remove any residual media. Then the cells were fixed using 400 μl of 4% w/v paraformaldehyde, followed by three 400–500 μl washes with TBS. After surface blocking with 2% w/v protease-free BSA (A3059, Sigma) solution in TBS, HRP-conjugated anti-HA high affinity antibody (3F10) (Roche) at a dilution of 1:2000 in TBS + 1% w/v protease-free BSA and TMB ready-to-use substrate (T0565, Sigma) were used for ELISA procedure. The ELISA results were normalized by Janus Green staining. Cells transfected with empty vectors (pcDNA3.1(+)) were used to determine background.

**Plasmids for functional assays**

The human wild-type S1PR5 gene (UniProt ID Q9H228) with an N-terminal 3× HA epitope (YPYDVPDYA) tag was cloned into pcDNA3.1(+) (Invitrogen) at KpnI/SalI and XhoI/SalI. Point mutations were introduced by overlapping PCR. All DNA sequences were verified by Sanger sequencing (Evrogen ISC). Sequences of all primers used in this work are listed in Supplementary Table 5.

**Cell surface expression determined by ELISA**

Cell surface expression of SIP₃ receptor variants was determined by whole-cell ELISA⁸³. Briefly, HEK293T cells were seeded in 24-well cell culture plates (0.2 × 10⁶ cells in 0.5 ml of medium per well) and transfected separately by 3 μg of each expression plasmids based on pcDNA3.1(+) vector using common Lipofectamine 3000 protocol. After 12–18 h incubation in a CO₂ incubator at 37 °C for receptor expression, the cell culture plates were placed on ice, the media was aspirated completely, and the cells were washed once with ice-cold TBS to remove any residual media. Then the cells were fixed using 400 μl of 4% w/v paraformaldehyde, followed by three 400–500 μl washes with TBS. After surface blocking with 2% w/v protease-free BSA (A3059, Sigma) solution in TBS, HRP-conjugated anti-HA high affinity antibody (3F10) (Roche) at a dilution of 1:2000 in TBS + 1% w/v protease-free BSA and TMB ready-to-use substrate (T0565, Sigma) were used for ELISA procedure. The ELISA results were normalized by Janus Green staining. Cells transfected with empty vectors (pcDNA3.1(+)) were used to determine background.

**Data availability**

Coordinates and structure factors for the S1P₃-ONO-5430608 structure have been deposited in the Protein Data Bank (PDB) under the accession code 7YXA. Raw SFX diffraction data have been deposited to CXIDB database under accession number 196. Publicly available amino acid sequences for S1PRs used in this study were obtained from the UniProt database under accession numbers: P21453, Q95136, Q99300, Q95977, Q9H228. Publicly available structures used in this study can be found in the Protein Data Bank under accession codes: 3V2W, 3V2Y, 4EIY, 7C45, 7EYV, 7EW1, 7EW2, 7EW4. SNV data for SIP₃ used in this work are available from public databases gnomAD [https://gnomad.broadinstitute.org/?gene=ENS00000180739?dataset=gnomad_r2.1] and COSMIC [https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=S1PR5]. AlphaFold structures, sequences, and scripts used to generate them are provided in Supplementary Data file 1. Structures of the compounds used for docking to experimental and AlphaFold
structures and their docking scores are provided in Supplementary Data file 2. Source Data are provided in this paper.

References

1. An, S., Goetzl, E. J. & Lee, H. Signaling mechanisms and molecular characteristics of G protein-coupled receptors for lysophosphatidic acid and sphingosine 1-phosphate. J. Cell. Biochem. 30–31, 147–157 (1988).

2. Maceyka, M., Harikumar, K. B., Milstien, S. & Spiegel, S. Sphingosine-1-phosphate signaling and its role in disease. Trends Cell. Biol. 20, 50–60 (2012).

3. Inoue, A. et al. Illuminating G-protein-coupling selectivity of GPCRs. Cell. Physiol. Biochem. 34, 158–171 (2014).

4. LaMontagne, K. et al. Antagonism of sphingosine-1-phosphate receptors by FTY720 inhibits angiogenesis and tumor vascularization. Cancer Res. 66, 221–231 (2006).

5. Hla, T. Signaling and biological actions of sphingosine 1-phosphate. Pharmacol. Res. 47, 401–407 (2003).

6. Inoue, A. et al. Illuminating G-protein-coupling selectivity of GPCRs. Cell 177, 1933–1947 e25 (2019).

7. Takakuwa, N., Okamoto, Y., Yoshioka, K. & Takuwa, Y. Sphingosine-1-phosphate signaling and cardiac fibrosis. Inflamm. Regen. 33, 96–108 (2013).

8. Im, D. S. et al. Characterization of a novel sphingosine 1-phosphate receptor, Edg-8. J. Biol. Chem. 275, 14281–14286 (2000).

9. Liu, H. et al. Sphingosine-1-phosphate modulates PAR1-mediated human platelet activation in a concentration-dependent biphasic manner. Sci. Rep. 11, 15308 (2021).

10. Schmidt, K. G. et al. Sphingosine-1-phosphate receptor 5 modulates early-stage processes during fibrogenesis in a mouse model of systemic sclerosis: a pilot study. Front. Immunol. 8, 1242 (2017).

11. Nelson, S. A. et al. Discovery of A-971432, an orally bioavailable selective sphingosine-1-phosphate receptor 5 (S1P5) agonist for the potential treatment of neurodegenerative disorders. J. Med. Chem. 58, 9154–9170 (2015).

12. Shah, M. V. et al. Molecular profiling of LGL leukemia reveals role of sphingolipid signaling in survival of cytotoxic lymphocytes. Blood 112, 770–781 (2008).

13. Bargiela, N. F., Garcia, C. M., Arufe, V. G., Hermida, J. R. V. & Herranz, I. M. Fingolimod in multiple sclerosis: profile of use in habitual practice. Eur. J. Hosp. Pharm. 27, 346–349 (2020).

14. Synnott, P. G., Bloudek, L. M., Sharaf, R., Carlson, J. J. & Pearson, S. D. The effectiveness and value of sipoimid for secondary progressive multiple sclerosis. J. Manag. Care Spec. Pharm. 26, 236–239 (2020).

15. Lamb, Y. N. Ozanimod: first approval. Drugs 80, 841–848 (2020).

16. Scott, F. L. et al. Ozanimod (RPC1063) is a potent sphingosine-1-phosphate receptor-1 (S1P1) and receptor-5 (S1P5) agonist with autoimmune disease-modifying activity. Br. J. Pharmacol. 173, 1778–1792 (2016).

17. Kurata, H. et al. Discovery of a 1-methyl-3,4-dihydropyridinethalene-based sphingosine-1-phosphate (S1P) receptor agonist ceralifimod (ONO-4641). A S1P1 and S1P5 selective agonist for the treatment of autoimmune diseases. J. Med. Chem. 60, 9508–9530 (2017).

18. Pérez-Jeldres, T. et al. Targeting cytokine signaling and lymphocyte traffic via small molecules in inflammatory bowel disease: JAK inhibitors and S1PR agonists. Front. Pharmacol. 10, 212 (2019).

19. Sanna, M. G. et al. Enhancement of capillary leakage and restoration of lymphocyte egress by a chiral S1P1 agonist in vivo. Nat. Chem. Biol. 2, 434–441 (2006).

20. Hanson, M. A. et al. Crystal structure of a lipid G protein-coupled receptor. Science 335, 851–855 (2012).

21. Maeda, S. et al. Endogenous agonist-bound S1PR3 structure reveals determinants of G protein-subtype bias. Sci. Adv. 7, eabf5325 (2021).

22. Zhao, C. et al. Structural insights into sphingosine-1-phosphate recognition and ligand selectivity of S1PR3–Gi signaling complexes. Cell Res. 32, 218–221 (2022).

23. Yuan, Y. et al. Structural and docking analysis of S1PR1 and S1PR5 reveals mechanisms of activation and drug recognition. Cell Res. 31, 1263–1274 (2021).

24. Xu, Z. et al. Structural basis of sphingosine-1-phosphate receptor 1 activation and biased agonism. Nat. Chem. Biol. 18, 281–288 (2022).

25. Watanabe, T., Kusumi, K. & Yuichil Inagaki, I. (2017). Tetra-hydronaphthalene Derivative. United States patent No. US 2019/0031605A1. U.S. Patent and Trademark Office https://patentscope.wipo.int/search/en/detail.jsf?docId=US236791513.

26. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. Structure 20, 967–976 (2012).

27. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. Nat. Protoc. 4, 706–731 (2009).

28. Chrenick, J. E. et al. Crystal structure of antagonist bound human lysophosphatidic acid receptor 1. Cell 161, 1633–1643 (2015).

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. 25, 366–428 (1995).

30. Zhou, Q. et al. Common activation mechanism of class A GPCRs. eLife 8, e50279 (2019).

31. Hilger, D. The role of structural dynamics in GPCR-mediated signaling. FEBS J. 288, 2461–2489 (2021).

32. Schoepp, N. G. et al. Crystal structures of agonist-bound human cannabinoid receptor CB1. Nature 55, 468–471 (2017).

33. Shao, Z. et al. Structure of an allosteric modulator bound to the CB1 cannabinoid receptor. Nat. Chem. Biol. 15, 1199–1205 (2019).

34. Wingler, L. M. et al. Angiotensin and biased analogs induce structurally distinct active conformations within a GPCR. Science 367, 888–892 (2020).

35. Israeli, H. et al. Structure reveals the activation mechanism of the MC4 receptor to initiate satiation signaling. Science 372, 808–814 (2021).

36. Katrich, V. et al. Allosteric sodium in class A GPCR signaling. Trends Biochem. Sci. 39, 233–244 (2014).

37. Zarzycka, B., Zaidi, S. A., Roth, B. L. & Katritch, V. Harnessing ion-bonding sites for GPCR pharmacology. Pharmacol. Rev. 71, 571–595 (2019).

38. Katritch, V., Cherezov, V. & Stevens, R. C. Structure-function of the G protein-coupled receptor superfamily. Annu. Rev. Pharmacol. Toxicol. 53, 531–556 (2013).

39. Nygaard, R., Frimer, T. M., Holst, B., Rosenkilde, M. M. & Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. Trends Pharmacol. Sci. 30, 249–259 (2009).

40. Fujiwara, Y. et al. Identification of the hydrophobic ligand binding pocket of the S1P1 receptor. J. Biol. Chem. 282, 2374–2385 (2007).

41. Laio, A. & Parrinello, M. Escaping free-energy minima. Proc. Natl Acad. Sci. USA 99, 12562–12566 (2002).

42. Niedernberg, A., Blaukat, A., Schöneberg, T. & Kostenis, E. Regulated and constitutive activation of specific signalling pathways by the human S1P5 receptor. Br. J. Pharmac. 138, 481–493 (2003).

43. Karcewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443 (2020).

44. Tate, J. G. et al. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res. 47, D941–D947 (2019).

45. Hornik, J. Analysis of naturally occurring genetic variants in the sphingosine-1-phosphate receptor family. PhD Thesis. The Pennsylvania State University, University Park (2019).
46. Arakawa, M. et al. Structural and functional roles of small group-conserved amino acids present on helix-H7 in the β2-adrenergic receptor. Biochim. Biophys. Acta Biomembr. **1808**, 1170–1178 (2011).

47. Steen, A. et al. Biased and constitutive signaling in the CC-chemokine receptor CCR5 by manipulating the interface between transmembrane helices 6 and 7. *J. Biol. Chem.* **288**, 12511–12521 (2013).

48. Wang, J. et al. The structural study of mutation-induced inactivation of human muscarinic receptor M4. *IUCR J* **7**, 294–305 (2020).

49. Bertalan, É., Le, A., Kollárová, H. & Tar, J. Surface potentials of A-frame DNA: an investigation of the atomic potential. *Nucleic Acids Res.* **52**, 3468–3480 (2022).

50. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

51. Akdel, M. et al. A structural biology community assessment of AlphaFold 2 applications. Preprint at https://www.bionxiv.org/content/10.1101/2021.09.26.461876v1 (2021).

52. Yuan, S., Wu, R., Mao, Y. et al. Rapid conformational analysis for fab classes with CHARMm and its applications. *Front. Comput. Biol.* **6**, 578–590 (2021).

53. Barty, A. et al. Cheetah: Software for high-throughput reduction and analysis of serial femtosecond crystallography data. *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 139–152 (2021).

54. Tong, A. B. et al. Could AlphaFold revolutionize chemical therapeutics? *Nat. Struct. Mol. Biol.* **28**, 771–772 (2021).

55. Alexandrov, A. I., Mileni, M., Chien, E. Y. T., Hanson, M. A. & Stevens, R. C. Microscale fluorescent thermal stability assay for membrane proteins. *Structures* **16**, 351–359 (2008).

56. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat. Commun.* **5**, 3309 (2014).

57. Barty, A. et al. Cheetah: Software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47**, 1118–1131 (2014).

58. Hunter, J. D. Matplotlib: a 2D graphic environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

59. White, T. A. et al. CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45**, 335–341 (2012).

60. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

61. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).

62. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).

63. Lovelace, J. J. & Borgstahl, G. E. O. Characterizing pathological imperfections in macromolecular crystals: lattice disorders and modulations. *Crystallogr. Rev.* **26**, 3–50 (2020).

64. Wang, J., Kamtekar, S., Berman, A. J. & Steitz, T. A. Correction of X-ray intensities from single crystals containing lattice-translocation defects. *Acta Crystallogr. D. Biol. Crystallogr.* **61**, 67–74 (2004).

65. Diederichs, K. & Karplus, P. A. Better models by discarding data? *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1215–1222 (2013).

66. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **54**, 5.6.1–5.6.37 (2016).

67. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

68. Pfendtner, J., Branduardi, D., Parrinello, M., Pollard, T. D. & Voth, G. A. Nucleotide-dependent conformational states of actin. *Proc. Natl Acad. Sci. USA* **106**, 12723–12728 (2009).

69. Abraham, M. J. et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX **1-2**, 19–25 (2015).

70. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLumed 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

71. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2016).

72. Davies, M. et al. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* **43**, W612–W620 (2015).

73. Ma, B. et al. Novel potent selective orally active S1P5 receptor antagonists. *ACS Med. Chem. Lett.* **12**, 351–355 (2021).

74. Landrum, G. RDKit: open-source cheminformatics. https://www.rdkit.org. RDKit v.2021-03-4 https://doi.org/10.5281/ZENODO.5085999 (2021).

75. Waskom, M. L. Seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 321 (2021).

76. Pandey, S., Roy, D. & Shukla, A. K. Measuring surface expression and endocytosis of GPCRs using whole-cell ELISA. *Methods Cell Biol.* **149**, 131–140 (2019).

77. Salahpour, A. et al. BRET biosensors to study GPCR biology, pharmacology, and signal transduction. *Front. Endocrinol.* **3**, 105 (2012).

Acknowledgements

We thank S. Ustinova, A. Podzorov, A. Awadleh, P. Utrobin, and Yu. Semenov for technical assistance, T. Maryum for providing ONO-5430608. We acknowledge the Paul Scherrer Institut, Villigen, Switzerland for the provision of synchrotron radiation beamtime at beamline PX1 of the SLS and would like to thank Dr. V. Olieric for assistance. The authors are grateful to the staff of the accelerator and beamline departments at PAL-XFEL for their technical support. The XFEL experiments were performed at the NCI-PAL-XFEL experimental station under proposal No. 2019-2nd-NCI-012. Protein production and crystallization were supported by the Ministry of Science and Higher Education of the Russian Federation agreement 075-00337-20-03, project FSGM-2020-0003 (V.B., A.M., E.L., A.Gu., M.S., A.L.). XFEL and synchrotron data collection and treatment were supported by the Russian Ministry of Science and Higher Education grant No. 075-15-2021-1354 (V.B., A.M., E.L., A.R.). SXF data collection strategy was developed with the support from the Russian Foundation for Basic Research (RFBR) project 18-02-40020 (V.B., A.M., E.M.). Functional assays were developed and implemented with the support of the Russian Science Foundation project 22-74-10036 (A.M., A.L., E.L.). Molecular docking simulations were supported by the Russian Science Foundation project 22-24-00454 (E.M., M.K.). J.P. and Y.C. were supported by the National Research Foundation of Korea (grant No. NRF-2017M3A9F6029736). U.W. is supported by the National Science Foundation (NSF) BioXFEL Science and Technology Center award 1231306. W.L. is supported by the Advancing Healthier Wisconsin Endowment (AHW) fund. The authors are grateful to the Global Science Experimental Data Hub Center (GSDC) for data computing and the Korea Research Environment Open NETwork (KREONET) for network service provided by the Korea Institute of Science and Technology Information (KISTI) and the Data Processing Center of Moscow Institute of Physics and Technology for high-performance data computing infrastructure and technical support.

Author contributions

E.L. and A.Gu. optimized the constructs, developed the expression and purification procedure, expressed and purified the protein, screened the ligands, and crystallized the protein–ligand complexes. E.M., D.V., A.M., and V.C. collected X-ray diffraction data at PAL-XFEL. J.P. set up the XFEL experiment, beamline, controls, and data acquisition; operated the beamline. H.H., U.W. helped develop and operate the LCP injector. W.L. helped with the XFEL sample preparation. S.P. contributed to the
experimental system installation at the beamline. G.P. contributed to DAQ and data handling. H.J.H. contributed to the Rayonix detector installation and operation. A.Gu., E.M., E.L., K.K, S.B. and A.M. collected synchrotron data at SLS. A.Ge. performed and analyzed cell signaling and cell surface expression assays. E.M., D.V., and V.B. processed diffraction data. G.B. helped with X-ray diffraction data interpretation and analysis. E.M., V.B., V.C. performed structure determination and refinement. V.B., V.C., E.L., A.Gu., E.M., A.M., A.L. performed project data analysis/interpretation. M.E., A.L., G.K., and M.S. helped with construct optimization, protein expression, and purification. P.P. advised on the protein construct design. E.M., M.K. performed molecular docking. P.O., E.M. performed MD simulations. E.M., I.G., V.B. performed AlphaFold simulations and their data analysis. A.L., I.O., P.Kh., A.R., Y.C. helped with experimental work and project organization. E.L., E.M., A.Gu., P.O., A.M., V.B., V.C. wrote the manuscript with the help from other authors. V.C, A.M., V.B., and V.G. initiated the project. A.M. and V.B. organized the project implementation, were responsible for the overall project management and co-supervised the research. V.C. supervised the overall project.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-32447-1.