

Supporting Online Material

Crystal Structure of a Lipid G-protein Coupled Receptor

Michael A. Hanson^{1*}, Christopher B. Roth¹, Euijung Jo², Mark T. Griffith¹, Fiona L. Scott¹, Greg Reinhart¹, Hans Desale¹, Bryan Clemons¹, Stuart M. Cahalan², Stephan C. Schuerer⁴, M. Germana Sanna², Gye Won Han⁴, Peter Kuhn³, Hugh Rosen^{2,5}††, Raymond C. Stevens^{4**}††

This PDF file includes:

Materials and Methods

Table S1 to S3

Figs. S1 to S6

References

Materials and Methods

Molecular biology for generation of *Spodoptera frugiperda*

Synthetic complimentary DNA encoding the full length S1P₁ receptor was ordered from GenScript, Inc., where it was synthesized using overlap extension PCR techniques. Upon delivery, a second construct was created by employing a two-step cloning strategy. The first step used splicing by overlap extension (SOE) PCR (1) to insert a cysteine-free bacteriophage T4 lysozyme (T4L C54T, C97A) (2) within the intracellular loop 3 (ICL3) of S1P₁-WT receptor (replacing Ser232-Lys244). The second step utilized standard PCR techniques to amplify the resulting S1P₁-T4L fusion, followed by subcloning into four distinct expression cassettes designed to explore tag placement and promoter strength in the context of baculovirus-directed protein expression in *Sf9* insect cells. Optimal expression was driven by the polyhedrin promoter with a construct design containing a cleavable HA signal sequence on the N-terminus of the receptor and a cleavable purification tag consisting of a 10x poly histidine tag upstream of a FLAGTM epitope tag.

Generation of the crystallization construct for the S1P₁-T4L receptor was accomplished by systematically truncating the C-terminus using standard site directed mutagenesis protocols. Each truncated species was tested for expression and the shortest variant that both expressed well and was stable was carried forward into crystallization trials.

Expression and purification of S1P₁-T4L for crystallization

High-titer recombinant baculovirus (>10⁸ viral particles per mL) was obtained following transposition and transfection protocols from Invitrogen (<http://www.invitrogen.com>). Viral titers were determined by the flow cytometric method after staining cells with gp64-PE

(Expression Systems) (3). For expression of recombinant protein, *Sf9* cells at a density of $2-3 \times 10^6$ cells/mL were infected with the high-titer viral stock at an multiplicity of infection (MOI) of 3. Cells were harvested by centrifugation at 48 hr post infection and stored at $-80\text{ }^\circ\text{C}$ until use.

Insect cell membranes were disrupted by thawing frozen cell pellets in a hypotonic buffer containing 10 mM HEPES (pH 7.5), 10 mM MgCl_2 , 20 mM KCl and protease inhibitor cocktail (Roche). Extensive washing of the isolated raw membranes was performed by repeated centrifugation in the same hypotonic buffer (~2-3 times), and then in a high osmotic buffer containing 1.0 M NaCl, 10 mM HEPES (pH 7.5), 10 mM MgCl_2 , 20 mM KCl (~3-4 times), followed by Dounce homogenization to resuspend the membranes in fresh wash buffer, thereby removing soluble and membrane associated proteins from the suspension of membranes. Purified membranes were resuspended in 10 mM HEPES (pH 7.5), 10 mM MgCl_2 , 20 mM KCl, and 40% glycerol then flash-frozen with liquid nitrogen and stored at $-80\text{ }^\circ\text{C}$ until further use.

Prior to solubilization, purified membranes were thawed at room temperature in the presence of 250 μM ML056 (Avanti Polar Lipids), 2.0 mg/mL iodoacetamide (Sigma) and protease inhibitor cocktail (Roche). After incubation for 30 min at $20\text{ }^\circ\text{C}$, membranes were chilled to $4\text{ }^\circ\text{C}$ and incubated for an additional 30 min. After ligand binding and preparative steps were complete, the membranes were solubilized by incubation in the presence of 0.5% (w/v) *n*-Dodecyl β -*D*-maltoside (DDM; Anatrace) and 0.1% (w/v) cholesteryl hemisuccinate (CHS; Sigma) for ~2.5-4 hr at $4\text{ }^\circ\text{C}$. The unsolubilized material was removed by centrifugation at $150,000 \times g$ for 45 min. The supernatant was incubated with TALON IMAC resin (Clontech) in buffer containing 50 mM HEPES (pH 7.5), 800 mM NaCl, 0.5% (w/v) DDM, 0.1% (w/v) CHS, and 20 mM imidazole. After overnight binding, the resin was washed with ten column volumes of 50 mM HEPES (pH 7.5), 800 mM NaCl, 10% (v/v) glycerol, 25 mM imidazole, 0.1% (w/v) DDM,

0.02% (w/v) CHS, 10 mM MgCl₂, 8 mM ATP (Sigma) and 250 μM ML056, followed by four column volumes of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 50 mM imidazole, 0.05% (w/v) DDM, 0.01% (w/v) CHS and 250 μM ML056. The receptor was eluted with 25 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 200 mM imidazole, 0.05% (w/v) DDM, 0.001% (w/v) CHS and 250 μM ML056 in a minimal volume. Purified receptor in the presence of ML056 was concentrated from ~1 mg/mL to ~100 mg/mL with a 100 kDa molecular weight cut-off Vivaspin concentrator (GE Healthcare). Receptor purity, monodispersity and concentration were estimated using SDS-PAGE and analytical size-exclusion chromatography.

In meso crystallization of S1P₁ with ML056

Protein samples of S1P₁ in complex with ML056 were reconstituted into lipidic cubic phase (LCP) by mixing with molten lipid in a mechanical syringe mixer. The protein-LCP mixture contained 40% (w/w) protein solution, 50% (w/w) monoolein (Sigma) and 10% (w/w) cholesterol (Avanti Polar Lipids). LCP crystallization trials were performed using an *in meso* crystallization robot as previously described (4). 96-well glass sandwich plates (Paul Marienfeld GmbH, Germany) were filled with 25 nL protein-laden LCP boluses overlaid by 0.8 μL of precipitant solution in each well and sealed with a glass coverslip. Crystallization set-ups were performed at room temperature (~20 °C). Plates were incubated and imaged at 20 °C using an automated incubator/imager (RockImager 1000, Formulatrix). Data-collection quality crystals were obtained in precipitant condition containing 0.1 M Tricine (pH 7.75), 34 - 36% (v/v) PEG400, 80 mM sodium citrate and 4% (v/v) glycerol. Optimal crystal size was achieved by allowing the crystals to nucleate for 1 day at 14 °C after which the experiment was transferred to

a 17 °C incubator were the crystals grew to full size (~50 μm x 10 μm x 5 μm) over a period of 2 days (Figure S1A). Crystals were harvested directly from LCP matrix using MiTeGen micromounts and flash frozen in liquid nitrogen.

Data collection and processing

X-ray diffraction data were collected on the 23ID-D/B beamline (GM/CA-CAT) at the Advanced Photon Source (Argonne, IL) using a 10 μm diameter minibeam with a MarMosaic 300 CCD detector (Table S1). Crystals embedded in LCP became invisible in the opaque mesophase upon flash-freezing into liquid nitrogen, and a similar rastering and data-collection strategy was followed as previously described (5) for about 400 crystal samples. Unattenuated exposures of 5 seconds per degree were required to observe diffraction to ~3 Å resolution (Figure S1B). Due to the rapid onset of radiation damage, data collection was limited to a maximum of 6° oscillation per crystal or position using 1° oscillations, of which only the first 1 – 2 frames diffracted to the maximum resolution. Data were integrated and scaled using *XDS* (6) and merged using *SCALA* (Figure S2).

Due to the difficulties inherent in scaling disparate wedges of data from multiple crystals, a medium-resolution initial dataset was first collected at low exposure for use as a scaling reference. This ensures that each small segment of data collected with maximum exposure has sufficient common reflections to allow robust scaling and merging with the full dataset. Data from each crystal was collected, indexed and integrated using *XDS* and scaled to the growing dataset using *XSCALE* and the medium resolution data as a scaling reference. We processed the data using the traditional method of inclusion based on isotropic resolution cutoffs applied to the data after indexing and integration. Using the resolution cutoff methodology where entire

resolution shells of data were eliminated based on poor merging statistics, we were able to obtain a 90.1% complete 3.35 Å dataset (Table S1).

Microdiffraction Data Assembly Method

Given the challenges in working with microcrystals, 10 μm minibeam, rapid radiation decay, and datasets from a large number of crystal samples, we examined different methods to assemble the data with the goal of improving the quality of the resulting electron density map and subsequent model. After a list of reflection observations were obtained from the integration step, the wedge of data was divided into reflection observations from a single image consisting of 1° of rotation about phi. Each of these individual sets of observations was iteratively combined with the main dataset by standard scaling procedures using the medium-resolution dataset as a reference standard as implemented in *XSCALE*. Most often the scaling procedure would fail, resulting in very high agreement factors presumably due to a large number of poorly measured reflections resulting from radiation damage to the crystal. This prompted a more granular data assembly method consisting of each individual reflection observation included based on a correlation threshold to the reflection peak fitting profiles.

Beginning at zero, the correlation threshold was increased in increments of 5% where all of the observations with a peak profile correlation greater than the threshold were subsequently scaled to the main dataset, and the observations with a peak profile correlation less than or equal to the threshold were filtered out of the reflection observation list. If the observations could be scaled to the main dataset with a total R_{merge} less than 16%, they were included and the algorithm moved on to the next set of integrated observations. If, however, the R_{merge} remained high indicating a poor fit to the data, the peak profile correlation threshold was increased 5% from the

previous round and the observations were rescaled with the higher acceptance criteria. The process was repeated for each set of reflection observations corresponding to 1° rotation about phi for a total of 6° on average for each crystal. If the data collected from a crystal were not useable, the process would eventually eliminate all of the observations and proceed to the next set of data. If the data collected from the crystal were useable, the final correlation threshold would typically vary between 0 and 40% depending on the degree of radiation damage and overall diffraction quality. Data for all of the crystals were analyzed and combined as appropriate in this fashion, resulting in a final scaled set of reflection observations that were converted to a multi-batch MTZ file using the CCP4 programs Pointless and scaled using *SCALA* (7). Using this procedure we were able to combine enough data to generate a 97.2% complete data set overall at 2.8 Å resolution with > 90 % completeness in the high resolution bin.

Since the peak profile filtering methodology could potentially risk an introduction of bias into the data via a positive intensity selection, especially at the higher resolution limits, we compared electron density maps from data processed with the microdiffraction data assembly method (maps calculated at 2.8 Å) to maps from data processed with the resolution cutoff filter method (maps calculated at 3.35 Å). Phases for the maps were calculated from protein models refined separately into each respective dataset so that phases used to calculate the 3.35 Å maps were not biased by refinement from the 2.8 Å data. Examination of the resulting maps indicated that disconnected electron density for the ligand could be seen in the difference $|F_o| - |F_c|$ maps as well as the $2|F_o| - |F_c|$ maps calculated from the 3.35 Å data processed by the resolution cutoff filter method (Fig. S3A). However, the electron density maps calculated from the 2.8 Å data processed by the microdiffraction data assembly method were of significantly higher quality (Fig. S3B). In this case most of the side chains could be placed in the electron density and there

was clear difference density $|F_o|-|F_c|$ for the ligand in the binding pocket. It is unlikely the electron density map improvements in resolution are due to artifacts in the resolution increase and filling in of missing reflections since the data set is 97% complete. We can conclude that in this case the risk of increasing the noise of the overall dataset is offset by incorporation of useable data through this methodology. This method may be generally applicable to situations where data collection is limited by severe radiation damage and estimates of uncertainty associated with each observation are not robust compared to data collected from a single crystal. Our positive selection method overcomes this hurdle, allowing inclusion of the best measured estimates of diffraction intensities that would otherwise be inaccessible ultimately improving the quality of the crystallographic maps.

Analysis of the 2.8 Å maps revealed improved details about the binding pocket and the N-terminal capping helix not seen unambiguously in the previous 3.35 Å data set processed by truncation of the resolution bins. The use of site specific mutagenesis was used to confirm the additional map details that were observed. Statistics for both methods are presented in the crystallographic summary table (Table S1) and both sets of PDB coordinates and structure factors have been deposited (PDB ID codes 3V2W and 3V3Y) for future evaluation and methods development in this area of research.

Structure determination

Using the 3.35 Å and 2.8 Å data sets, initial structures were obtained by molecular replacement (MR) with the program *PHASER* (8) using two independent search models of the β_2 -adrenergic receptor (β_2 AR; without ligand) and T4L from the β_2 AR/carazolol structure (PDB ID 2RH1). All refinements were performed with the *REFMAC5* (9) , *PHENIX* (10) and

autoBUSTER (11) followed by manual examination and rebuilding of the refined coordinates in the program *COOT* (12) using both sigma-A weighted $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ maps, as well as omit maps. Given the challenges in working at these resolutions with ligand placement and partial occupancy, the ligand was modeled in using the available electron density along with the knowledge/confirmation of ligand placement from site specific mutagenesis data. Statistics for both methods are presented in the crystallographic summary table (Table S1). The PDB id codes are 3V2W for the 3.35 Å structure refined using conventional resolution cut off method and 3V3Y for the 2.80 Å data using the novel microdiffraction data assembly method.

Mutagenesis and Assays

Site-directed mutagenesis along the putative aromatic pocket

Plasmid encoding N-terminal, triple HA-tagged S1P₁ fusion protein in pcDNA3.1 was purchased from cDNA.org. S1P₁ receptor mutants were generated by overlapping PCR mutagenesis and sequences were verified prior to use in transient transfections for screening and the subsequent generation of stable single cell clones.

CHO-K1 cells were purchased from the American Type Culture Collection (Manassas, VA). For transient transfection, cells were plated on 10-cm dishes at 80% confluence and transfected with 12 μ g of each plasmid using Fugene HD (Roche Applied Science, Indianapolis, IN) in Opti-MEM (Invitrogen). After 16 hr, each pool-transfected dish was split into 6-well plates and the cells were incubated for additional 24 hr. The cells were next incubated for 4 hr in serum-free Dulbecco's modified Eagle's medium before addition of the ligands.

Signaling in stable single cell clones

In order to improve the consistency of the data on transiently transfected S1P₁ CHO cells, we subsequently made stable single cell clones of triple HA-tagged WT and W269L receptors. For developing stable single cell clones, the cells from each pool-transfected dish were split into 6-well plates at 20% confluence and antibiotic-resistant cells were selected after Geneticin treatment for 2 wk. The Geneticin-selected cells were diluted by limiting dilution and were plated into 96-well plates to give a single cell per 10 wells. Each single cell clone was grown and examined for its receptor expression by Western blotting using anti-HA antibody. Single cell clones were also examined for their ERK phosphorylation upon S1P stimulation by ELISA. Selected single cell clones were examined for the receptor expression by Western blotting using anti-HA antibody. Clones were also screened for the downstream ERK phosphorylation upon 1 μ M S1P stimulation. The positive clones from the Western blotting were consistently positive for signaling through ERK phosphorylation. The positive single clones were examined for the time course for ERK phosphorylation upon stimulation with S1P and CYM-5442. The phosphorylation time course was consistent with that previously described, 5-10 min being the time points for the maximal ERK phosphorylation (13).

Ligand-mediated ERK phosphorylation was measured using PathScan Phospho-p44/42

MAPK (Thr202/Tyr204) Sandwich ELISA kit (Cell Signaling Technology, Danvers, MA). Cells expressing WT or mutant S1P₁ were serum-starved for 4 hr. In the antagonist experiments, cells were incubated with ML056 (Avanti Polar Lipids, Alabaster, AL) at 10 μ M for 1 hr before agonist treatment. Cells were then stimulated for 5 min (determined to give maximal ERK phosphorylation for all agonists) with increasing concentrations of S1P, FTY-720P or CYM-5442 and phosphorylation of p44/42 MAPK was assayed according to manufacturer's protocol. The dose response curves for agonist-mediated p44/42 MAPK phosphorylation were analyzed and EC₅₀ was determined using Prism (Graphpad Software, San Diego, CA).

Based on the previously published homology modeling (*14, 15*), a series of minimal point mutations were made along the putative aromatic cluster on TMVI to see if any of the mutants had defective binding or signaling in response to S1P or CYM-5442 (synthesized in Dr. Edward Roberts' lab as described in Gonzalez-Cabrera 2008, Roberts et al US patent application 2010). The mutations were carefully made using conservative substitutions, such as hydrophobic to hydrophobic substitution, rather than alanine substitution or hydrophobic to polar substitution that could change the helical conformation and property drastically. When we compared transient transfections of F210L, F265L, and W269L mutants, all three mutants showed decreased or abolished CYM-5442-induced ERK phosphorylation. However, S1P-induced ERK phosphorylation was minimally affected in W269L (Table S3) in contrast to the previously published mutations of F210 and F265 (*14, 16*), indicating that S1P binding is not highly dependent on W269, while F210 and F265 impact globally on the structure and function of the pocket. We therefore concentrated on analyzing the contribution of W269, a residue universally conserved in Class A GPCRs. Stable cell lines expressing W269L mutant were compared with WT receptor for ERK phosphorylation in response to stimulation by S1P, FTY-720-P and CYM-

5442. The EC50 shift was not evident for FTY-720P (showing that global receptor structure and function were well preserved), was modest for S1P and was >60-fold for CYM-5442. In radio-ligand cold-competition whole cell binding studies using S1-³³P (Figure S6), the IC50s for cold-competition by S1P were 34.8 and 11.5nM respectively for the WT and W269L mutation. CYM-5442 competed for S1P binding on the WT but not on the W269L receptor, showing that the binding pockets have some overlap but that CYM-5442 has a requirement for W269 for both binding and receptor activation. We then examined whether the location of tryptophan at 269 is critical for CYM-5442 binding or if the presence of tryptophan somewhere within the ligand binding pocket is sufficient. F265 is located one helical-turn down from W269 and F273 is one helical-turn up along helix VI. F210 in helix V was proposed to be facing W269 in the helical bundle (15). We made a series of double mutant S1P₁ receptors: F210W/W269F, F265W/W269F, and F273W/W269F. The replacement of phenylalanine with tryptophan at various locations in the pocket did not rescue the W269F mutant from loss of CYM-5442-binding (data not shown). To conclude, tryptophan could be neither moved up (F273W) nor down (F265W) along the helix VI and could not be moved to the facing helix V (F210W) to restore CYM-5442 binding or function.

Both the planar bulk and the electronics of W269 within the cavity of the binding site might stabilize and optimize CYM-5442 binding. Alternatively, the isomerization of tryptophan and the π - π interaction between the CYM-5442 phenyl moiety and the W269 indole ring might play a critical role in CYM-5442 binding and subsequent receptor activation.

Both WT and W269F mutant S1P₁ cells were examined by fluorescence microscopy using anti-HA antibody. They all showed strong staining of the whole cells, indicating that the receptors were normally expressed and localized on the cell surface in the absence of ligands

(Figure S5). Upon stimulation with S1P or CYM-5442, both WT and W269F receptors were internalized (Figure S5), indicating that the mutant S1P₁ receptor was synthesized, modified, recycled, and degraded in a normal manner.

Jump-In Stable Cell Lines with Homogeneous Surface Expression for Quantitative Studies

Although stable transfection provided us with more consistent data than transient transfection, there have been concerns regarding traditional limitations to stable transfection, such as gene silencing and cell-to-cell variability. Therefore, we further established Jump-In cell lines with either WT or mutant S1P₁. Jump-In integration is irreversible and only a single copy of the gene of interest is integrated at the same site in every cell.

Jump-InTM TITM CHO-K parental cells, Gateway cloning vectors (pDONR 221, pJTI R4 DEST, and pJTI R4 Int), and enzymes (BP clonase II and LR clonase) were purchased from Invitrogen Corp. First, triple HA-tagged WT, W269L, and W269F S1P₁ were cloned into entry clones using the BP recombination reaction. They were retargeted into an appropriate pJTI R4 DEST vector to yield a pJTI R4 EXP retargeting expression vector. Jump-In TI CHO-K cells were transfected with both pJTI R4 EXP S1P₁ and pJTI R4 Int vector for retargeting. After expanding retargeted cells, retargeted Jump-In TI cells were selected using blasticidin at 10 µg/mL for 4 wk.

Jump-In cell lines were examined for the equivalent receptor expression by Western blotting using anti-HA antibody. In addition to WT and W269L S1P₁, the W269F mutant was made to examine whether the tryptophan is replaceable by phenylalanine for ligand-induced ERK phosphorylation. All three Jump-In stable cell lines were stimulated with increasing concentrations of either S1P or CYM-5442 and were examined for their ERK phosphorylation. EC₅₀ for the WT, W269L, and W269F were 0.097 nM, 0.54 nM, and 0.56 nM S1P, respectively,

giving 6-fold higher EC_{50} for W269L and W269F compared to WT. On the other hand, CYM-5442-induced ERK phosphorylation gave 69-fold higher EC_{50} for W269L and only 7-fold higher EC_{50} for W269F. These data confirmed that W269 in the S1P₁ receptor plays a more critical role in CYM-5442-induced receptor activation, compared to S1P-induced receptor activation. Further, these experiments revealed that the presence of the aromatic ring structure is key to delivering the optimal signal from CYM-5442 binding to the bitopic binding pocket of the S1P₁ receptor.

During the course of the structural refinement, it became apparent that there were additional important polar interactions contributing to ligand binding and interacting with the sphingosine head group. The residues N101, Y29 and K34 were tested for their effect on FTY720-P induced signaling in transiently transfected CHO cells (Table S3). When the mutant N101A was tested for its ability to signal in response to FTY720-P, it was found to have a significantly right shifted EC_{50} confirming the finding that this residue is involved in sphingolipid binding. Importantly, this mutant was also deficient in CYM-5442 induced signaling indicating that this ligand shared a common interaction in the sphingolipid head group binding region (Table S3).

The individual point mutants Y29A and K34A were created and the receptor was tested for its ability to signal in the presence of FTY720-P (Table S3). Neither single mutant affected the EC_{50} . However, when the two mutations were combined, the double mutant elicited a significant effect on the pERK1/2 (Table S3) signaling.

Radioligand binding and competition

In order to determine the upstream cause of the defective ERK phosphorylation of W269L upon CYM-5442 stimulation, we examined the initial step of GPCR activation - ligand

binding. CHO cells (5×10^5 per well) stably expressing WT S1P₁ or W269L mutant were serum-starved for 4 hrs. They were then incubated at 4 °C for 30 min in the binding buffer containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 15 mM NaF, 0.5 mM EDTA, 1 mM Na₃VO₄, 0.5% fatty acid-free bovine serum albumin, and protease inhibitor mixture (Roche) with 0.1 nM [³³P]S1P (American Radiolabeled Chemicals, Inc., St. Louis, MO) and increasing concentrations of S1P, ML056, SEW2871, or CYM-5442. Cells were washed three times with cold binding buffer. Cell-bound radioactivity was measured by lysing the cells with 0.5% SDS followed by liquid scintillation counting. The raw data was normalized so that the level of [³³P]S1P bound to each cell line (WT or W269L) in the absence of competing ligand was referenced as 100% for its own cell line. The level of [³³P]S1P bound to WT was significantly higher (up to 4.7-fold) than that bound to W269L in the absence of competing ligands.

With increasing concentrations of cold S1P, [³³P]S1P binding was competitively reversed in a dose dependent manner in both WT and W269L as expected (Figure S7A). The S1P₁-specific antagonist ML056 also reversed [³³P]S1P binding competitively in a dose dependent manner in both WT and W269L mutant receptors (Figure S7B). On the other hand, CYM-5442 was unable to compete for [³³P]S1P binding in the W269L mutant whereas it was able to compete for [³³P]S1P binding in WT receptor (Figure S7C). This suggests that the S1P binding pocket and CYM-5442 binding pocket are partially overlapping but not identical. The results indicate that W269 plays a more critical role in binding CYM-5442 than binding S1P.

Agonist Modeling

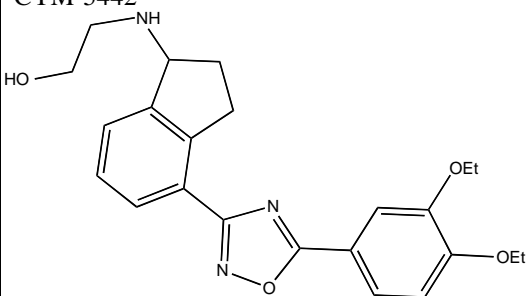
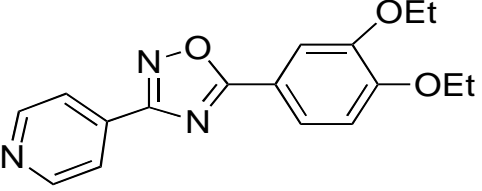
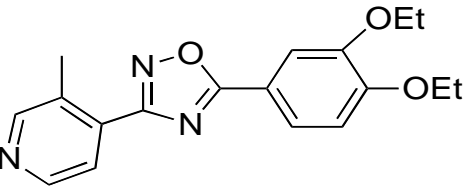
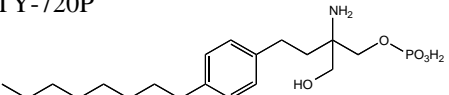
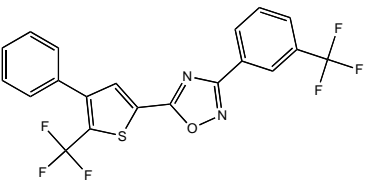
In developing a working model of agonist binding, SAR evidence from the ML056-like series of compounds provided insight into chemical changes that result in the conversion from antagonism to agonism. First, it was found that increasing the acyl chain length of the meta-

substituted analogs of ML056 eventually resulted in conversion of the series from antagonists to agonists. This implies that the volume occupied by the hydrophobic portion of the ligand plays a key role in the agonist triggering event (17). The second finding related to the substitution pattern around the phenyl ring, whereas the meta-substituted acyl chain is capable of antagonist binding, the para substituted acyl chain is not. This implies that a volume change in the vicinity of the phenyl ring is a key point for determining agonism. Through docking studies, it was found that the ML056 antagonist binding pocket of the S1P₁ receptor could indeed accommodate acyl chain lengths up to 8 carbons while maintaining optimal sphingosine head group interactions. This is consistent with the antagonistic properties of these compounds on the S1P₁ and S1P₃ receptors. However, when the acyl chain was increased to 10 carbons, the binding pocket could not accommodate the increase in hydrophobic volume without disrupting the polar and ionic interactions of the sphingosine headgroup. In order to induce side-chain displacements conducive to accommodating the increased bulk the sphingolipid agonists, we utilized a constrained induced fit docking protocol. We found that the increased volume requirements of the meta substituted agonists could be satisfied by allowing subtle changes to occur within the binding pocket in response to an induced fit docking protocol. These subtle changes are within the scope of binding pocket conformational changes expected upon transitioning to an active state receptor (18-21). The optimal model was created by allowing the side-chains of the residues lining the hydrophobic portion of the binding pocket to move in response to a para-substituted regio-isomer of ML056 while holding the sphingosine head group interactions fixed.

Table S1. Data collection and refinement statistics

Processing Method	S1P ₁ -ML056	
	Traditional resolution cutoff method	Microdiffraction data assembly method
Data collection		
Number of crystals	33	77
Space group	<i>P2₁2₁2</i>	<i>P2₁2₁2</i>
Cell dimensions a, b, c (Å)	107.9, 69.7, 81.9	107.9, 69.7, 81.9
Number of reflections measured	22,196	92,094
Number of unique reflections	8,293	15,297
Resolution (Å)	20 – 3.35 (3.53 – 3.35)	20 – 2.80 (2.95 – 2.80)
R _{merge} [*]	0.18 (0.78)	0.16 (0.42)
Mean I/sd(I)	3.5 (1.1)	6.1 (1.9)
Completeness (%)	90.1 (80.8)	97.2 (94.4)
Redundancy	2.7 (2.0)	6.0 (3.1)
Refinement		
Resolution (Å)	20 – 3.35 (3.53 – 3.35)	20 – 2.80 (3.0 – 2.80)
Number of reflections (test set)	8,286 (594)	15,275 (742)
R _{work} / R _{free}	0.22 / 0.28	0.23 / 0.27
Number of atoms		
All	3,423	3,523
Proteins	3,386	3,475
Ligand (ML056)	23	23
Others (Lipid, sugar and waters)	14	25
Average B value (Å²)		
All	77	77
S1P ₁	72	71
T4 lysozyme	86	83
Ligand (ML056)	81	59
Others (Lipid, sugar and waters)	100	69
RMSD		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	0.9	0.7
Ramachandran statistics		
Favored regions (%)	96	96
Allowed regions (%)	4	4
Outliers (%)	0	0

Table S2. Selectivity of CYM-5442 and its structural analogues for the S1P_x receptors

Structures of CYM-5442 & Analogues	EC ₅₀ (nM)				
	S1P ₁	S1P ₂	S1P ₃	S1P ₄	S1P ₅
<p>CYM-5442</p> 	1.35 (100%) ± 0.25	NA	NA	NA	20% at 10µM
<p>CYM-5181</p> 	5.6 ± 1.4	ND	661	NA	ND
<p>CYM-5178</p> 	0.15	ND	397	NA	ND
<p>FTY-720P</p> 	0.2	>1,000	5	5.9	0.6
<p>SEW-2871</p> 	13	>10,000	>10,000	>10,000	>10,000

NA, no activity; ND, not determined

CYM-5442 and its structural analogues were tested in 12-point (up to 10 µM) concentration response curves.

SEW-2871 from (22)

FTY720-P from (23)

Table S3. EC₅₀ (nM) of the dose-response curves of ligand-induced ERK phosphorylation in WT vs. various mutants.

Stable single cell clones or stable mixed populations of WT, W269L, N101A, Y29A, K34A or Y29A / K34A mutant S1P₁ receptors were stimulated with increasing concentrations of S1P, FTY720-P or CYM-5442. Data are reported as individual replicates in parenthesis where the deviation is very large.

Ligand	S1P			FTY720-P			CYM5442		
	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev	N
WT (stable)	0.29	0.14	6	0.40	0.36	3	7.42	1.40	8
W269L (stable)	6.15	2.02	5	0.51	0.81	3	435	120	6
N101A (stable)	n/a			23.80	14.00	3	340 (822, 106, 91)	418	3
Y29A (stable mixed)	n/a			0.18	0.02	2	n/a		
K34A (stable mixed)	n/a			0.28	0.11	2	n/a		
Y29A / K34A (stable mixed)	n/a			255 (453, 56)	281	2	n/a		

Figures

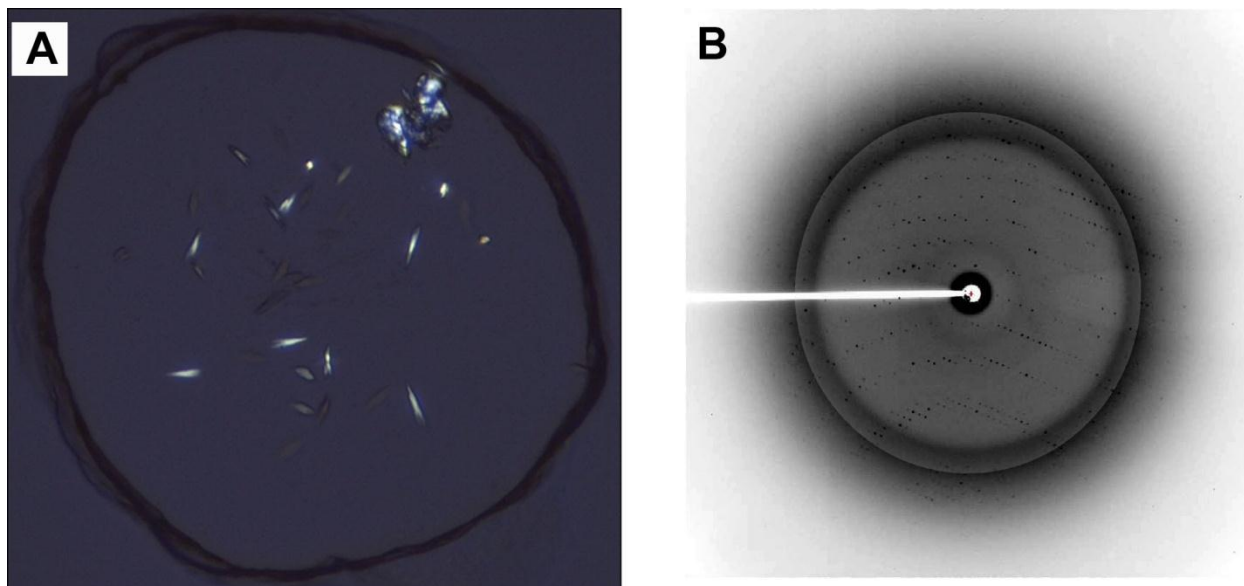


Fig S1. Representative crystals with cross-polarization under a microscope (A) and by X-ray diffraction (B). The diffraction image was divided into two zones for adjustment of brightness and contrast to view the low and high-resolution spots simultaneously.

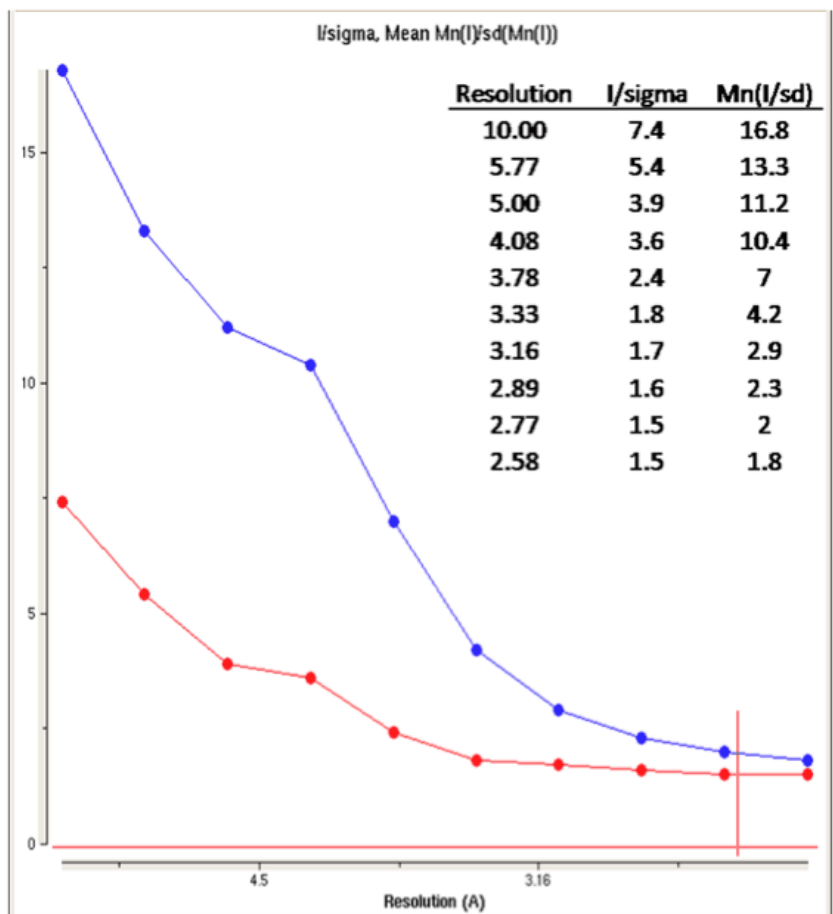


Fig. S2. After scaling the final integrated data in *XDS*, the unmerged statistics were imported into the program *SCALA* for merging. The data were merged in *SCALA* while optimizing the error model to give an estimate of the Intensity over uncertainty estimate (*I/sigma*: red trace), and average Intensity over standard deviation of intensity (*Mn(I/sd)*: blue trace). Because of the multiplicity of the data, the standard deviation of the integrated intensities can be used as an estimate in the uncertainty of the measurements. Data integrated and scaled with *XDS* but merged with *SCALA* constitute the final dataset to which the model was fit. The resolution cutoff was determined based on the *Mn(I/sd)* and examination of the incremental change in the mean intensity as a function of resolution.

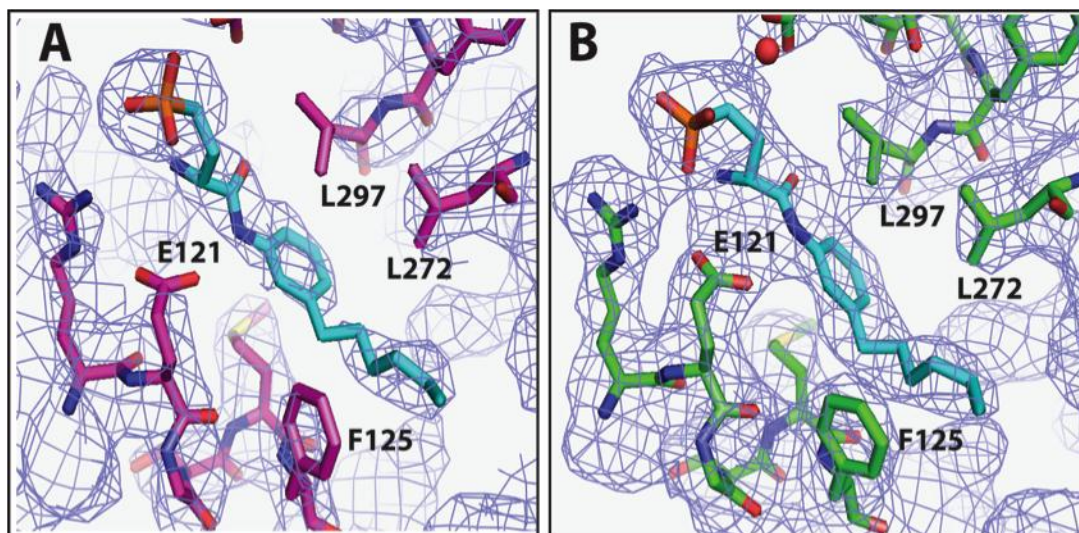
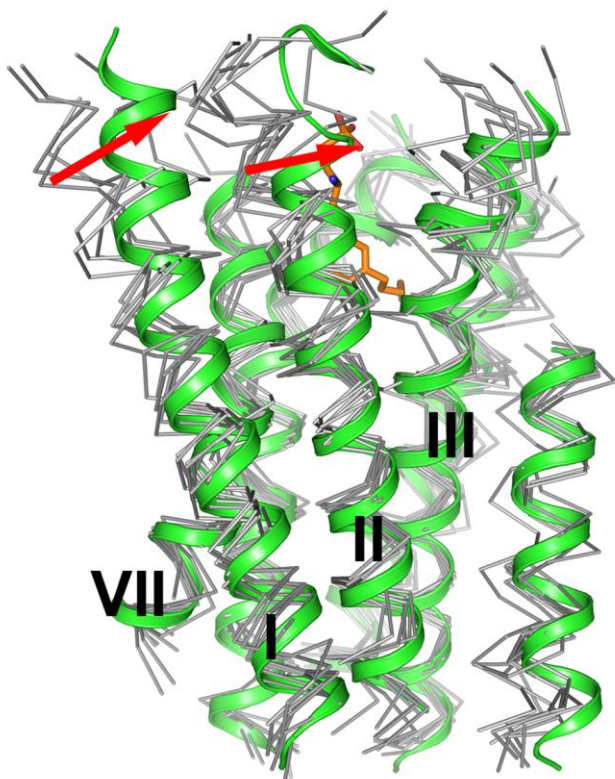


Fig. S3. Electron density maps (blue) contoured at 1σ demonstrating differences in quality of the data processing methods. **A.** Data processed by the traditional resolution assembly method with a final resolution of 3.35 \AA (magenta stick) yields $2|F_o|-|F_c|$ maps missing density in key areas of binding pocket, such as side chain density for Glu121, Phe125 and Leu297. Ligand density is incomplete as well. **B.** Data processed by the microdiffraction data assembly method with a final resolution of 2.8 \AA (green stick), yields $2|F_o|-|F_c|$ maps with clear density for most of the side chains in the receptor structure and all of the side chains that define the binding pocket. Red dot top of the figure represents a water. The ligand density is complete and continuous in the $2|F_o|-|F_c|$ map and is clearly visible in a $|F_o|-|F_c|$ difference map contoured at 3σ (not shown). Maps were calculated using the clipper utilities distributed with CCP4; missing F_{obs} were not estimated using calculated structure factors to improve the map quality, a feature common to maps calculated after refinement in *REFMAC*.



PDBID	Receptor	Chain	No. Atoms	RMSD core	RMSD helices
3EML	a2a	A	104	0.9	2.7
2RH1	b2AR	A	104	1.0	2.4
3PBL	D3	A	104	1.0	2.0
3ODU	CXCR4	A	97	1.9	3.6
1U19	rhod	A	104	1.6	5.8
2VT4	b1AR	C	104	1.0	2.3
3RZE	H1	A	104	1.0	2.3

Fig S4. Overlay of all published antagonist bound structures (grey C α trace) with S1P₁ (green ribbon). The alignment was performed using the core residues (Ballesteros-Weinstein: 1.50-2.53; 3.30-3.54; 5.49-5.62; 6.43-6.61; 7.40-7.53) and the RMSD reported as “RMSD core” in the table. In the case of CXCR4 the number of core residues was reduced to produce an overlay with an RMSD < 2.0 Å. After alignment the RMSD was calculated for all of the C α atoms in the transmembrane region (RMSD helices). The position of helices I and II are repositioned (red arrows) closer to helix III in S1P₁ relative to other receptors possibly to open access to the binding pocket.

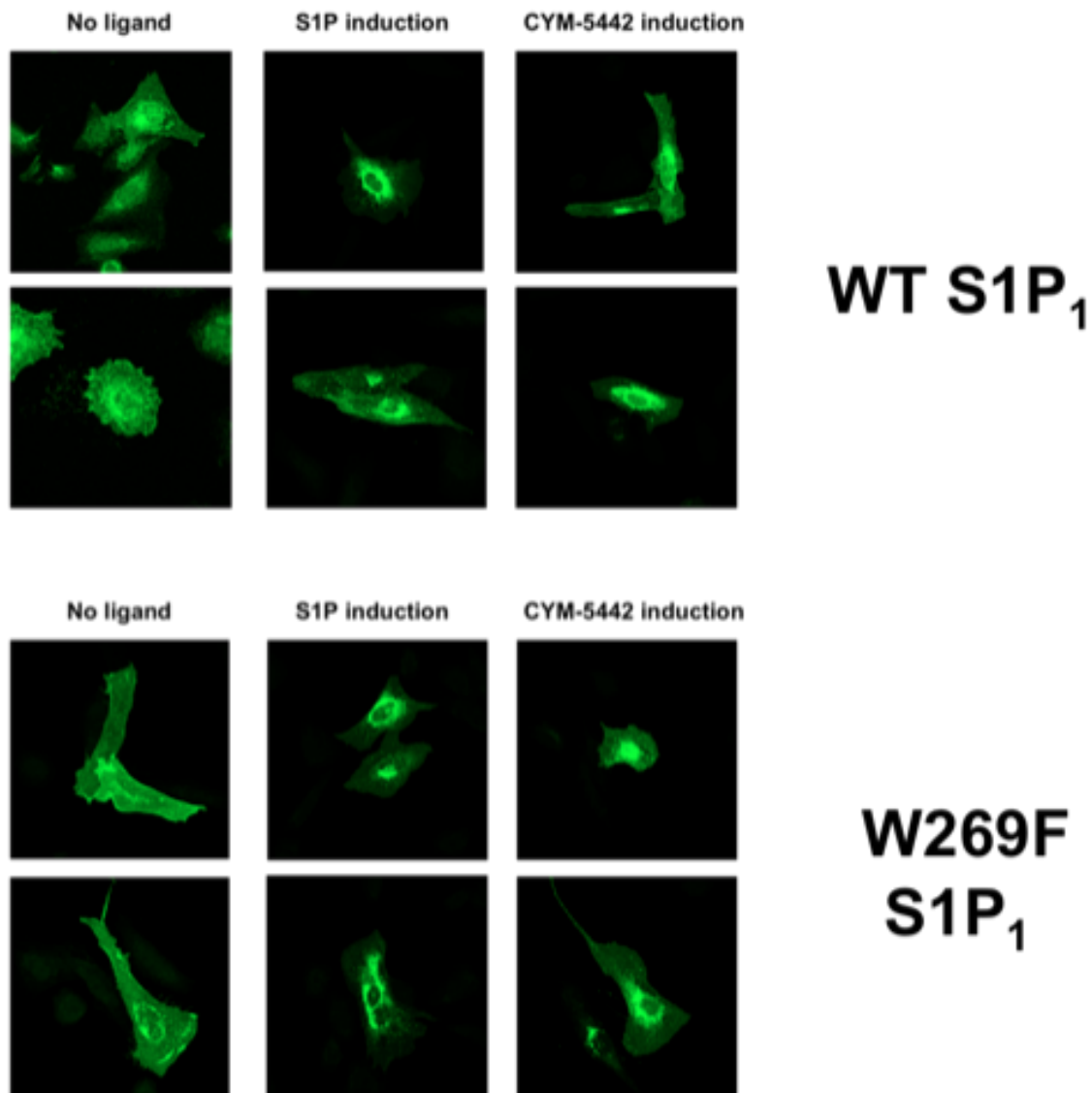
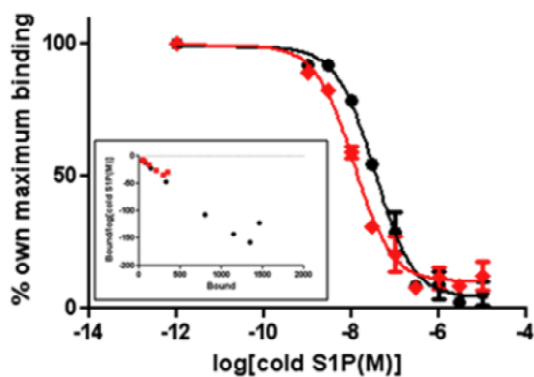
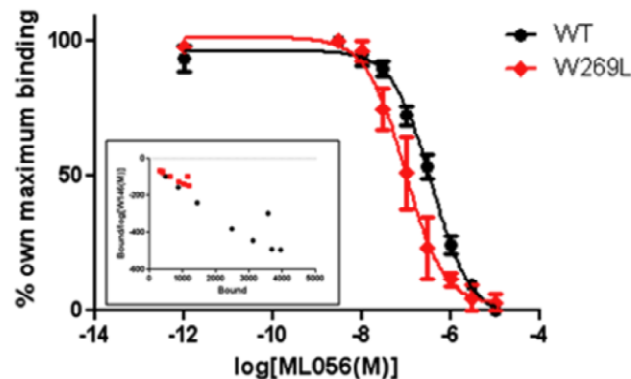


Fig. S5. Normal cell surface expression and internalization of the mutant S1P₁ receptor. Both WT and W269F mutant S1P₁ CHO cells were examined by fluorescence microscopy using anti-HA antibody. They all showed strong staining of the whole cells, indicating that the receptors were normally expressed and localized on the cell surface in the absence of ligands. Upon stimulation with S1P or CYM-5442, both WT and W269F receptors were internalized, indicating that the mutant S1P₁ receptor was synthesized, modified, recycled, or degraded normally.

(A) S1P Competition



(B) ML056 Competition



(C) CYM-5442 Competition

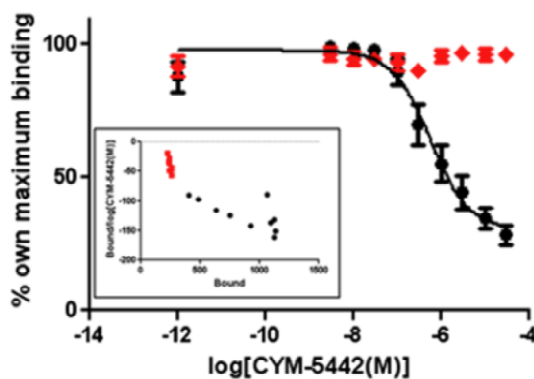


Fig. S6. Radioligand binding and competition with the S1P₁ receptor. Stable single cell clones of WT and W269L mutant S1P₁ receptors were incubated with 0.1 nM [³³P]S1P in the presence of increasing concentrations of (A) S1P, (B) ML056, (C) CYM-5442. [³³P]S1P binding was competitively reversed with S1P and ML056 in both WT and W269L single cell clones whereas CYM-5442 was unable to compete for [³³P]S1P binding in the mutant W265L receptor. The insets show the Scatchard plot of the competition curves. Means and SEM (n=3).

(24)

References:

1. K. L. Heckman, L. R. Pease, *Nature protocols* **2**, 924 (2007).
2. D. M. Rosenbaum *et al.*, *Science* **318**, 1266 (2007).
3. T. Mulvania, B. Hayes, D. Hedin, *Bioprocess J*, 47 (2004).
4. V. Cherezov, A. Peddi, L. Muthusubramaniam, Y. F. Zheng, M. Caffrey, *Acta crystallogr* **60**, 1795 (2004).
5. V. Cherezov *et al.*, *J Royal Soc Interface* **6 Suppl 5**, S587 (2009).
6. W. Kabsch, *Acta crystallogr* **66**, 125 (2010).
7. N. Collaborative Computational Project, *Acta crystallogr* **D50**, 760 (1994).
8. A. J. McCoy *et al.*, *J Appl Crystallogr* **40**, 658 (2007).
9. G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta crystallogr* **53**, 240 (1997).
10. P. D. Adams *et al.*, *Acta crystallogr* **66**, 213 (2010).
11. G. Bricogne *et al.*, (2011). BUSTER version 2.8.0 Cambridge, United Kingdom: Global Phasing Ltd.
12. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta crystallogr* **66**, 486 (2010).
13. E. Jo *et al.*, *Chem Biol* **12**, 703 (2005).
14. Y. Fujiwara *et al.*, *J Biol Chem* **282**, 2374 (2007).
15. S. C. Schurer *et al.*, *ACS Chem Biol* **3**, 486 (2008).
16. W. J. Valentine *et al.*, *J Biol Chem* **286**, 30513 (2011).
17. M. D. Davis, J. J. Clemens, T. L. Macdonald, K. R. Lynch, *J Biol Chem* **280**, 9833 (2005).
18. F. Xu *et al.*, *Science* **332**, 322 (2011).
19. S. G. Rasmussen *et al.*, *Nature* **477**, 549 (2011).
20. S. G. Rasmussen *et al.*, *Nature* **469**, 175 (2011).
21. G. Lebon *et al.*, *Nature* **474**, 521 (2011).
22. M. G. Sanna *et al.*, *J Biol Chem* **279**, 13839 (2004).
23. S. Mandala *et al.*, *Science* **296**, 346 (2002).
24. We appreciate the feedback and discussions with Professors Jack Johnson (TSRI), Andrew Leslie (MRC LMB), Wolfgang Kabsch (MPI) and Tassos Perrakis (NKI-AVL) on the experimental microdiffraction data assembly method to improve the quality of electron density maps.