Electronic Supplementary Material (ESI) for Chemical Communications

Glucose Neopentyl Glycol (GNG) Amphiphiles for Membrane Protein Solubilization, Stabilization and Crystallization

Pil Seok Chae*, Rohini Rana, Kamil Gotfryd, Søren G. F. Rasmussen, Andrew C. Kruse, Kyung Ho Cho, Stefano Capaldi, Emil Carlsson, Brian Kobilka, Claus J. Loland, Ulrik Gether, Surajit Banerjee, Bernadette Byrne*, John K. Lee* & Samuel H. Gellman*

Supporting figures

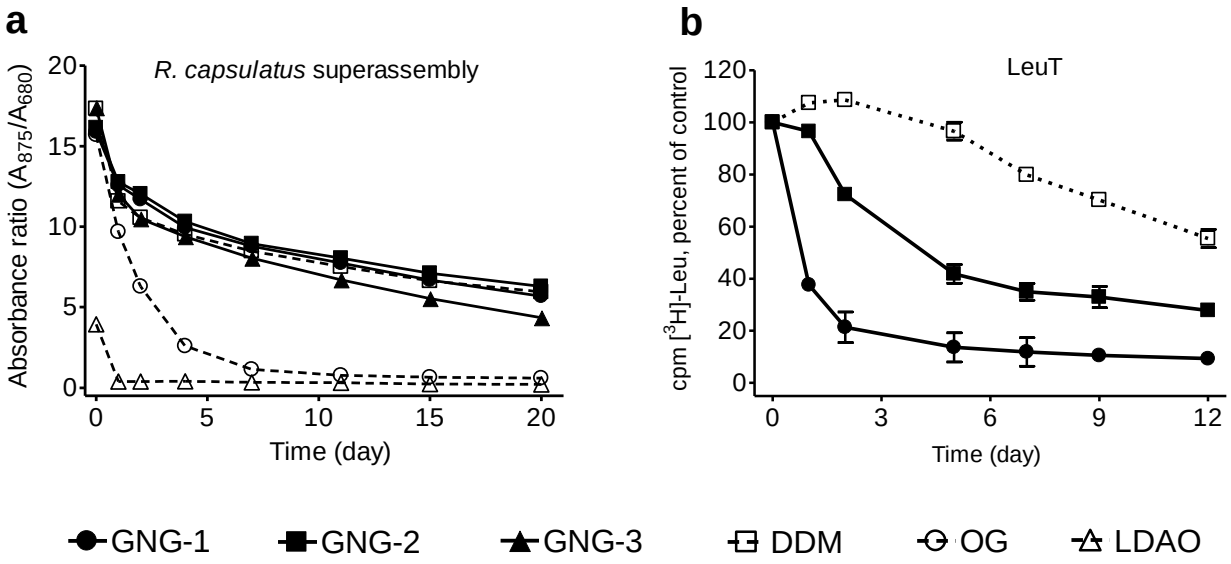


Figure S1. Stability time course of (a) LHI-RC complexes and (b) LeuT in GNG amphiphiles (GNG-1, GNG-2, and GNG-3) and conventional detergents (DDM, LDAO and OG). Agents were tested at CMC + 0.2 wt % for both proteins stored at room temperature. For the stability assessment of LHI-RC complexes, the absorption ratios (A_{875}/A_{680}) were measured and calculated at regular intervals (n = 1). LeuT activity ([^{3}H]-Leu binding) was evaluated using a scintillation proximity assay (SPA). Results are expressed as % activity relative to the day 0 measurements (mean ± SEM, n = 2).

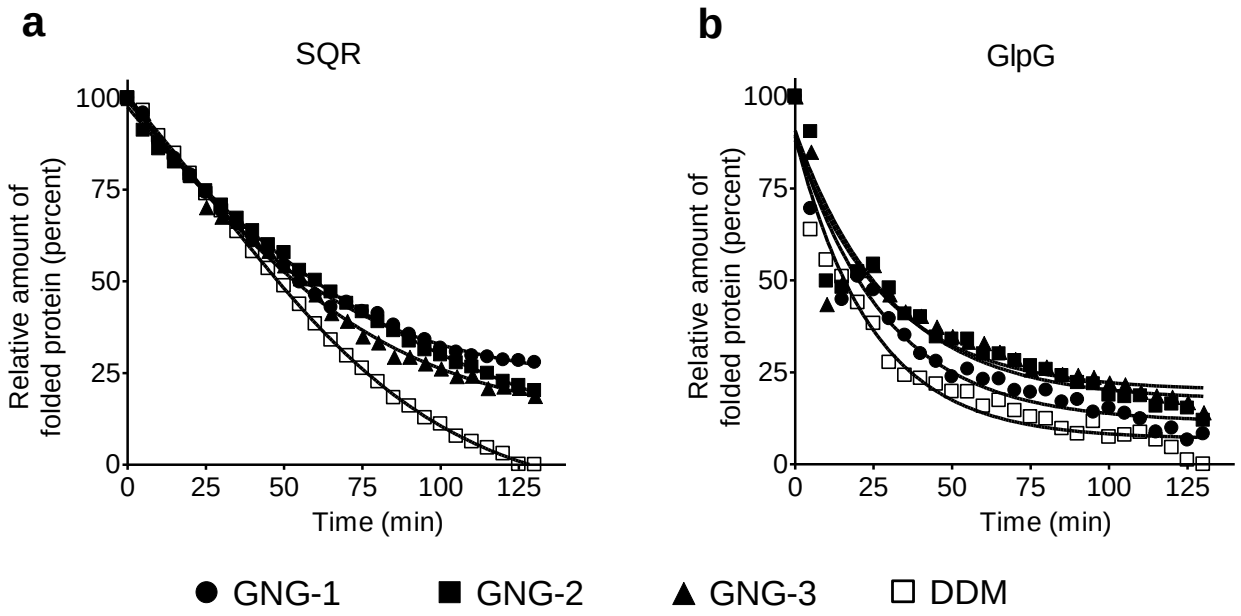
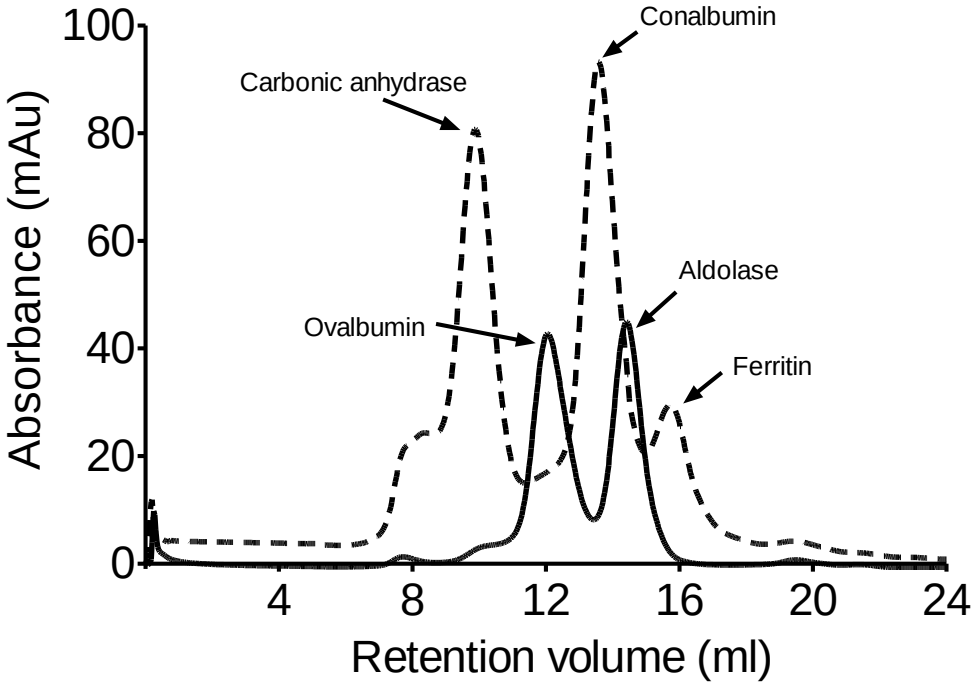


Figure S2. CPM assays for (a) SQR and (b) GlpG after solubilization with individual detergents. SQR was extracted from the membrane using 2 wt % C12E9, purified in buffer containing 0.05 wt % C12E9 and exchanged into buffer containing 0.2 wt % decyl-β-D-maltoside (DM). GlpG was initially extracted from the *E. coli* membrane with 1 wt % DDM and isolated in buffer containing 0.03 wt % DDM. CMP-Sia was initially extracted from *S. cerevisiae* membranes with 1% DDM and purified in buffer containing 0.03% DDM. Each purified proteins was diluted 1:150 into solutions containing CMC + 0.04 wt % of individual amphiphiles. The CPM analysis was performed over 130 min at 30°C using a microplate spectrofluorometer set at an excitation wavelength of 387 nm and an emission wavelength of 463 nm. Measurements were taken every 5 min after automatic agitation of the plate. Note that the vertical axes in these graphs have no absolute meaning. The "Relative amount of folded protein" in each case is defined as follows: 100% corresponds to the fluorescence emission intensity at time = 0 min while 0% corresponds to the lowest value measured among the samples treated with amphiphiles during the 130 min assay period. The fluorescence of CPM-containing solution with no protein was not measurably dependent on the detergent micelles used in the experiments.

a



b

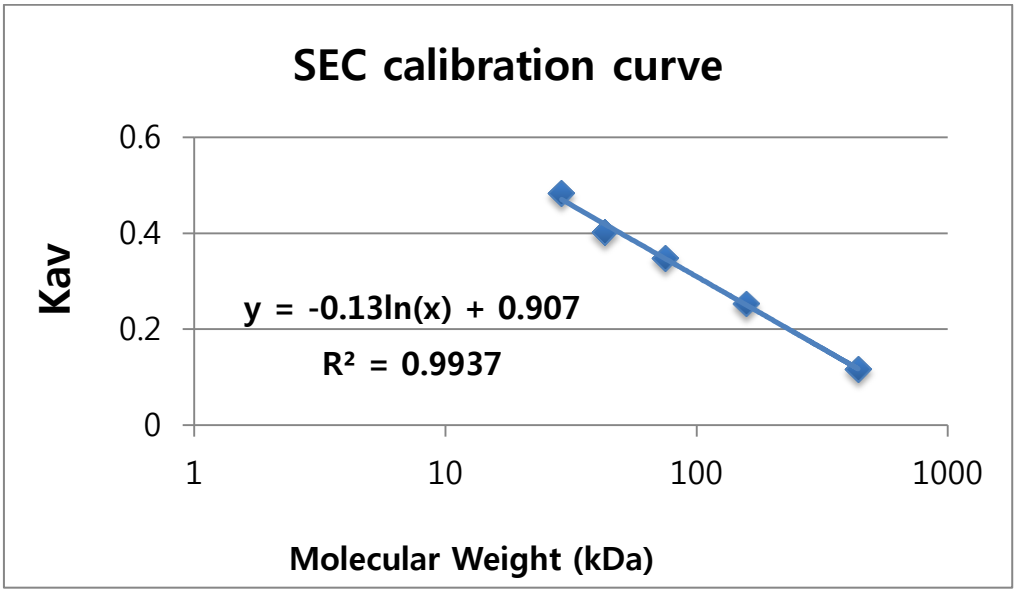


Figure S3. Size exclusion chromatography (SEC) analysis with standard protein makers (a) and the resulting linear calibration curve (b) for the molecular weight (MW) determination of PDCs. Five standard proteins (Carbonic anhydrase (29 kDa), Ovalbumin (43 kDa), Conalbumin (75 kDa), Aldolase (158 kDa) and Ferritin (440 kDa)) were subjected to SEC experiments using same elution buffer condition and column employed for SEC analysis of GlpG protein in the main text. No measurable change was observed by the addition of DDM into the elution buffer. The linear calibration curve was obtained using a least squares regression. Based on the calibration curve, the MW of DDM-solubilzed and GNG-2-solubilized GlpG were estimated to be 91 kDa and 56 kDa, respectively.

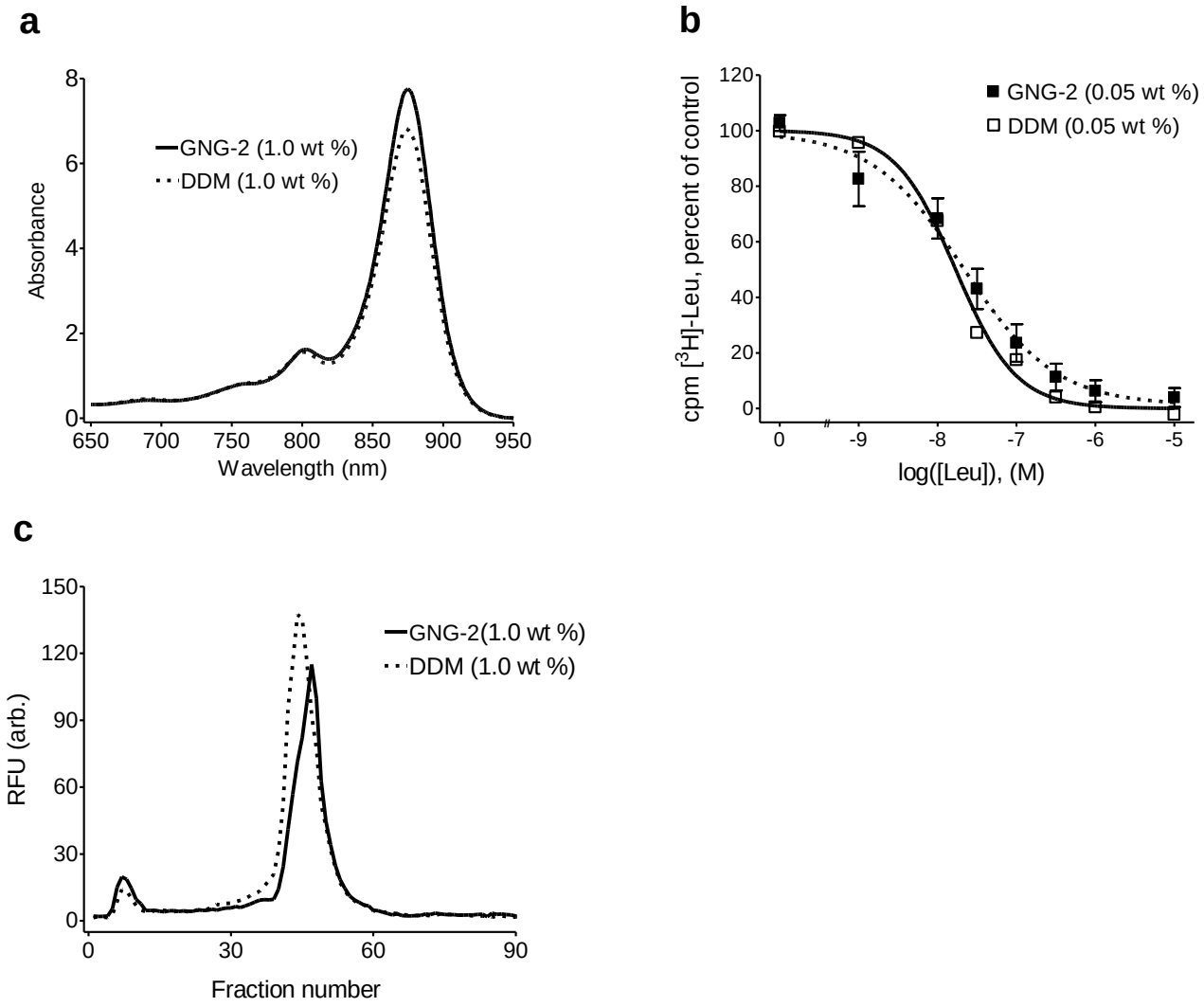


Figure S4. Comparison of proteins extracted from the membrane with either GNG-2 or DDM. (a) LHI-RC superassembly; quaternary structure integrity was assessed via absorbance. (b) LeuT activity was assessed by competition ligand binding ([^{3}H]-Leu/Leu) using a scintillation proximity assay (mean ± SEM, n = 2). Measured ligand K_d values were 22 ± 6 and 16 ± 1 nM for GNG-2 and DDM, respectively. (c) CMP-Sia-GFP fusion protein; integrity was assessed by size exclusion chromatography, which was monitored via fluorescence.

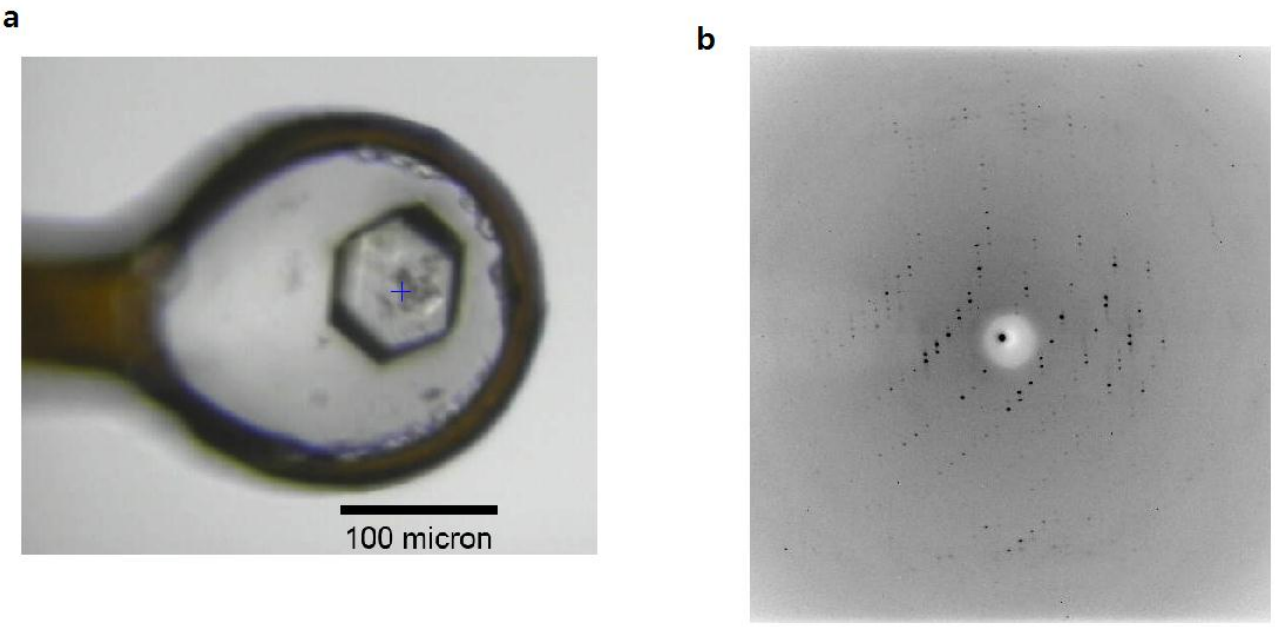


Figure S5. (a) Images and (b) X-ray diffraction patterns for the crystals of acetate transporter obtained in GNG-3 using the hanging drop technique. Diffraction limit was 4.1 Å. The diffraction experiment was conducted at beamlines ALS 4.2.2.

Table S1. Critical micelle concentrations (CMC) and hydrodynamic radii (R_h) of the micelles (mean ± SD, n = 3) for GNGs and representative conventional detergents.

	MW[a]	CMC (mM)	CMC (wt %)	R_h (nm)[b]
GNG-1	628.8	~1.6	~0.10	2.64 ± 0.04
GNG-2	656.8	~0.10	~0.0092	10.57 ± 0.03
GNG-3	540.6	~1.0	~0.058	3.07 ± 0.02
DDM	510.1	~0.17	~0.0087	3.47 ± 0.04
OG	292.4	~25[c]	~0.73	1.5~2.3[c]

[a]Molecular weight. [b]Hydrodynamic radius of micelles as determined by dynamic light scattering. [c]These values were obtained from the literature. [11]

Protein stability evaluation

Solubilization and stability assay for *R. capsulatus* superassembly

The solubilization and stability of the *R. capsulatus* superassembly were assessed according to the published protocol.[1] Briefly, we obtained specialized photosynthetic membranes from an engineered strain of *Rhodobacter (R.) capsulatus*, U43[pUHTM86Bgl], lacking the LHII light-harvesting complex. We start solubilization of the LHI-RC superassembly by thawing and homogenizing frozen aliquots of *R. capsulatus* membranes at room temperature and incubating them with mild agitation at 32˚C for 30 min. Subsequently, 30-min incubation was performed after adding individual detergents/amphiphiles (DDM, OG, LDAO, or GNG amphiphiles) at 1.0 wt % in the solid form. The solution was then subjected to ultracentrifugation at 315,000 *g* at 4℃ for 30 min to remove membrane debris. UV-Vis spectra of the solubilized protein solutions were measured in a range of 650 ~ 950 nm to assess solubilization efficiency.

For the stability assay, DDM-solubilized sample was transferred into a new 1.7 mL microcentrifuge tube containing Ni-NTA resin (Qiagen, Inc.; Valencia, CA; pre-equilibriated and stored in an equal volume of buffer containing 10 mM Tris, pH 7.8, and 100 mM NaCl). Following a 1 h incubation at 4℃ for protein binding, the resins were washed twice with 0.5 mL of binding buffer (a pH 7.8 Tris solution containing DDM at 1xCMC) and eluted three times with 0.20 mL elution buffer solutions containing 1 M imidazole (otherwise, this buffer was identical to binding buffer; the pH of each solution was readjusted to pH = 7.8). The DDM-purified solutions were collected and diluted with 0.4 mL of the binding buffer. Then small aliquots (0.05 mL) of the DDM-purified protein solutions were mixed with 0.95 mL solutions containing individual detergents/amphiphiles (DDM, OG, LDAO, or GNG amphiphiles) at concentrations CMC+ 0.04 wt % or CMC+ 0.2 wt %. UV-Vis spectra of these solutions were taken to monitor the stability of the protein at room temperature over 20 days. Protein denaturation was assessed by measuring the 875 nm/680 nm absorbance ratio.

Membrane Solubilization and Protein Purification (GlpG and SQR)

GlpG was expressed as fusion proteins with a C-terminal GFP-His tag in *Escherichia coli*. All steps were carried out at 4 °C. Membranes containing GlpG were resuspended in PBS, 10 mM Imidazole pH 8.0, 150 mM NaCl, 10% glycerol and solubilized in 1% DDM for 1 hr with mild agitation. Supernatant containing DDM-solubilised protein was harvested after ultracentrifugation at 100,000 *g* for 45 min. The GFP-His fusions, GlpG was bound to Ni^{2+}-NTA resin (1 ml per 1 mg of GFP fusion) pre-equilibrated with Buffer A (PBS, 10 mM Imidazole pH 8.0, 150 mM NaCl, 10% glycerol, 0.03% DDM) using stirred mixing for 2-3 hr. The resin was washed with 10 CV of Buffer A, then 35 CV of Buffer A supplemented 30 mM Imidazole, followed by elution using 2-3 CV of Buffer A supplemented with 250 mM Imidazole. Equal amounts of His-tagged TEV protease was added to the GFP-His fusions in the eluate, and the samples dialysed overnight against Buffer B (20 mM Tris (pH 7.5), 150 mM NaCl, 0.03% DDM). Cleaved GlpG was isolated in the flowthrough fractions using reverse Ni^{2+}-NTA binding. Samples were concentrated to a 0.5 ml volume using centrifugal concentrators, and submitted to a final polishing gel filtration step using a Superdex 200 10/300 column pre-equilibrated with Buffer B. GlpG was concentrated to 5mg/ml using centrifugal concentrators.

SQR was expressed in *E. coli* as a untagged construct. Membranes (~400 mg) containing SQR were resuspended in 20 mM potassium phosphate buffer (pH 7.4) containing 0.2 M EDTA and solubilized in 2% $C_{12}E_9$ for 15 min. Supernatant containing detergent-solubilized protein was harvested following ultracentrifugation at 100,000*g* for 45 min, and filtered through a 0.2 μm filter. SQR was bound to pre-equilibrated Q-sepharose Fast Flow resin in an

XK26/20 column (~ 24 ml). The column was washed with 2 CV of Buffer C (20 mM potassium phosphate (pH 7.4), 0.2 M EDTA, 0.05% C12E9), 2 CV of Buffer C supplemented with 100 mM NaCl, followed by elution using a (100-350) mM NaCl gradient. Fractions containing SQR were concentrated using an Amicon stirred cell concentrator, and filtered. The SQR was then applied onto a Phoros 50 HQ resin using an XK16/20 column (~20 ml) pre-equilibrated with Buffer C, followed by a Sephacryl 300 26/60 pre-equilibrated with buffer D (20 mM potassium phosphate (pH 7.4), 0.05% $C_{12}E_9$). The final buffer exchange was performed on a Superdex 200 10/300 gel filtration column pre-equilibrated with 20 mM Tris (pH 7.6), 0.2% decyl-β-D-maltoside (DM). SQR was concentrated to 12 mg/ml using centrifugal concentrators.

Samples for CPM Assay and Gel Filtration Analysis

CPM dye (Invitrogen), stored in DMSO (Sigma), was diluted (1:100) in Buffer B supplemented with 5 mM EDTA. Amphiphile or DDM were used at CMC+0.04 wt% concentrations in 20 mM Tris (pH 7.5), 150 mM NaCl as test buffers. 1 µl of the purified protein; GlpG (5 mg/ml) and SQR (12 mg/ml) was individually added to test buffers (150 µl) in Greiner 96-well plates, and left for equilibration at RT for 5 min, before adding 3 µl diluted CPM dye. The fluoresence of each well was monitored at 30 °C for 2 hr.

For gel filtration analysis, 10 µl of purified GlpG (5 mg/ml) was diluted in 1000 µl test buffer. GNG -2 or DDM were used at CMC+0.04 wt% concentrations in 20 mM Tris (pH 7.5), 150 mM NaCl as test buffers. 500 µl aliquots of the diluted protein were applied onto a Superdex 200 30/100 gel filtration column, before and after incubation at 30 °C for 2 hr. The column was pre-equilibrated with the respective test buffer prior to sample loading.

LeuT expression, purification and functional assay

The wild type of the leucine transporter from *A. aeolicus* was overexpressed *E. coli* C41(DE3) harbouring pET16b encoding LeuT-His_8, essentially as previously described.[2] Plasmid was kindly provided by E. Gouaux (Vollum Institute, Portland, Oregon, USA). Briefly, after isolation of bacterial membranes followed by solubilization in 1 % DDM, the LeuT was purified by Ni^{2+}-NTA affinity chromatography in buffer composed of 20 mM Tris-HCl (pH 8.0), 1 mM NaCl, 199 mM KCl and containing 0.05 wt % DDM. Two fractions with highest protein concentrations were pooled to achieve ca. 1.5 mg/ml stock, which was subsequently diluted 1:10 in same buffer without DDM, but containing GNG-1 or GNG-2 in final concentrations of CMC + 0.04 wt % or CMC + 0.2 wt %, respectively. Control samples included DDM used at the above-mentioned final concentrations. Following protein storage at RT, at the indicated time points, samples were centrifuged and the protein concentration was measured. Concomitantly, for the corresponding time points, protein activity was assessed as the ability of [^{3}H]-Leu binding monitored using scintillation proximity assay (SPA).[3] Briefly, SPA was assessed in 200 mM NaCl in the presence of tested compounds at the above-mentioned concentrations for reaction mixtures consisting of 5 µL of the respective protein samples, 20 nM [^{3}H]-leucine and copper chelate (His-Tag) YSi beads (both from PerkinElmer, USA). SPA was monitored using MicroBeta liquid scintillation counter (PerkinElmer). In separate series of experiments, the LeuT was purified in 0.05 wt % DDM or GNG-2, and the activity of the respective samples was assessed by competition ligand binding ([^{3}H]-Leu/Leu) using SPA protocol.

CMP-Sia solubilization

CMP-Sia was expressed as a fusion protein with a C-terminal GFP-His tag in *Saccharomyces cerevisiae* cells, as described above. All steps were carried out at 4 °C. Membranes containing CMP-Sia were resuspended in PBS, 10 mM Imidazole pH 8.0, 150 mM NaCl, 10% glycerol and and the protein concentration measured. The membranes

were adjusted to a concentration of 1 mg/ml and 1 ml aliquots were solubilized individually in DDM and GNG-2 at final detergent concentrations of 1.0 wt % for 1 hr at 4 ℃ with mild agitation. 100 µl aliquots were removed from each tube, and a fluorescence reading was taken for each sample before and after ultracentrifugation at 150,000 *g* for 1 hr to remove insoluble material. The solubilization efficiency (%) is the fluorescence reading of the soluble supernatant divided by the fluorescence reading of the total sample times 100. The remaining soluble fraction for each condition was submitted to fluorescent SEC (FSEC) using a Superose 6 column (GE Healthcare) equilibrated with buffer containing the appropriate agent (DDM or GNG-2).

Amphiphile Synthesis

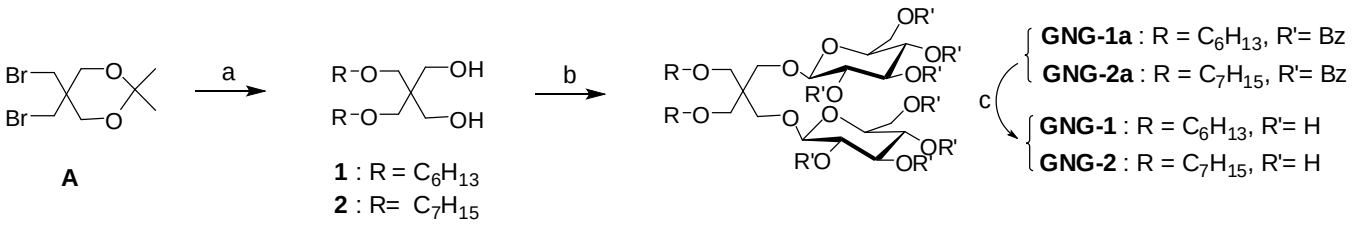
General procedure for glycosylation reactions

This reaction was performed according to a literature method[6] with slight modification. A mixture of alcohol derivative, AgOTf (2.4 equiv.) and 2,4,6-collidine (1.8 equiv.) in anhydrous CH_2Cl_2 (40 mL) was stirred at -45°C. A solution of perbenzoylated glucosylbromide (2.4 equiv.) in CH_2Cl_2 (40 mL) was added dropwise over 0.5 h to this suspension. Stirring was continued for 0.5 h at -45°C, and then the reaction mixture was allowed to warm to 0 °C and left stirring for 1.5 h. After completion of the reaction (as detected by TLC), pyridine was added to the reaction mixture, and it was diluted with CH_2Cl_2 (40 mL) before being filtered over celite. The filtrate was washed successively with a 1 M aqueous $Na_2S_2O_3$ solution (40 mL), a 0.1 M aqueous HCl solution (40 mL), and brine (2 x 40 mL). Then the organic layer was dried with anhydrous Na_2SO_4, and the solvent was removed by rotary evaporation. The residue was purified by silica gel column chromatography (EtOAc/hexane), which provided the desired product as a glassy solid.

General Procedure for the de-*O*-benzoylations under Zemplén's conditions[6]

The *O*-benzoylated compounds were dissolved in MeOH and then treated with the required amount of a methanolic solution of 0.5 M NaOMe such that the final concentration of NaOMe was 0.05 M. The reaction mixture was stirred for 6 h at room temperature, and then neutralized with Amberlite IR-120 (H^+ form) resin. The resin was removed by filtration and washed with MeOH, and solvent was removed from the combined filtrate *in vacuo*. The residue was purified by silica gel column chromatography ($MeOH/CH_2Cl_2$). Further purification, by recrystallization using CH_2Cl_2/MeOH/diethyl ether, afforded fully de-*O*-benzoylated product as a white solid.

Supplementary scheme 1



(a) 1-hexanol or 1-heptanol, NaH, DMF, 120°C; *p*-TSA, MeOH, room temperature, 92% (in two steps); (b) perbenzoylated glucosylbromide (2.4 equiv.), AgOTf, CH_2Cl_2, -45°C → room temperature, 94%; (c) NaOMe, MeOH, room temperature, 97%.

2,2-Bis-hexyloxymethyl-propane-1,3-diol (**1**) and 2,2-bis-heptyloxymethyl-propane-1,3-diol (**2**) were synthesized according to the reported procedure[7] by using 1-hexanol or 1-heptanol instead of 1-decanol, as described in **supplementary scheme 1**.

2,2-Bis-hexyloxymethyl-propane-1,3-diol (**1**): **^{1}H NMR** (300 MHz, $CDCl_3$): δ 3.65 (d, *J* = 5.8 Hz, 3H), 3.51 (s, 4H), 3.42 (t, *J* = 6.6 Hz, 4H), 2.87 (t, *J* = 5.8 Hz, 2H), 1.61-1.53 (m, 4H), 1.38-1.25 (m, 12H), 0.89 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 73.4, 72.3, 65.7, 44.7, 31.8, 29.7, 26.0, 22.8, 14.2 ; **MS (MALDI-TOF)**: calcd. for $C_{17}H_{36}O_4$ $[M+Na]^+$ 327.2506, found 327.2509.

2,2-Bis-heptyloxymethyl-propane-1,3-diol (**2**): **^{1}H NMR** (300 MHz, $CDCl_3$): δ δ 3.65 (d, *J* = 5.8 Hz, 3H), 3.51 (s, 4H), 3.42 (t, *J* = 6.6 Hz, 4H), 2.87 (t, *J* = 5.8 Hz, 2H), 1.60-1.53 (m, 4H), 1.36-1.26 (m, 16H), 0.89 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 73.4, 72.3, 65.7, 44.7, 32.0, 29.7, 29.3, 26.3, 22.8, 14.3 ; **MS (MALDI-TOF)**: calcd. for $C_{19}H_{40}O_4$ $[M+Na]^+$ 355.2819, found 355.2809.

GNG-1a and **GNG-2a** were synthesized according to the general procedure for glycosylation.

GNG-1a: **^{1}H NMR** (300 MHz, $CDCl_3$): δ 8.17-8.11 (m, 4H), 8.02-7.97 (m, 4H), 7.94-7.88 (m, 4H), 7.87-7.81 (m, 4H), 7.70-7.60 (m, 6H), 7.58-7.24 (m, 20H), 5.63 (t, *J* = 9.8 Hz, 2H), 5.51 (t, *J* = 9.8 Hz, 2H), 5.40 (dd, *J* = 9.6, 8.4 Hz, 2H), 4.46 (dd, *J* = 12.0, 3.4 Hz, 2H), 4.33 (dd, *J* = 12.0, 4.8 Hz, 2H), 3.91 (d, *J* = 7.8 Hz, 2H), 3.83 (d, *J* = 9.2 Hz, 2H), 3.52-3.12 (m, 12H), 1.48-1.35 (m, 4H), 1.33-1.16 (m, 12H), 0.86 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 166.3, 166.0, 165.3, 165.0, 133.7, 133.6, 133.4, 133.3, 130.2, 130.0, 129.9, 129.8, 129.3, 129.2, 129.1, 128.6, 128.5, 101.7, 72.8, 72.1, 71.9, 71.7, 69.9, 69.1, 68.9, 63.3, 45.2, 31.9, 29.7, 26.0, 22.8, 14.3; **MS (MALDI-TOF)**: calcd. for $C_{85}H_{88}O_{22}$ $[M+Na]^+$ 1483.5660, found 1483.5591.

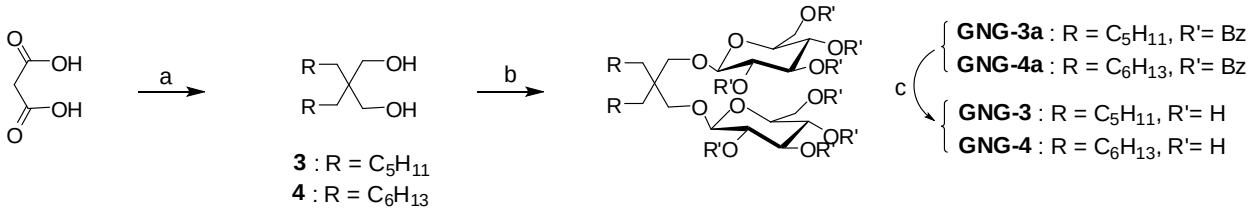
GNG-2a: **^{1}H NMR** (300 MHz, $CDCl_3$): δ 8.17-8.11 (m, 4H), 8.02-7.97 (m, 4H), 7.94-7.88 (m, 4H), 7.87-7.81 (m, 4H), 7.70-7.60 (m, 6H), 7.58-7.24 (m, 20H), 5.63 (t, *J* = 9.8 Hz, 2H), 5.51 (t, *J* = 9.8 Hz, 2H), 5.40 (dd, *J* = 9.6, 8.4 Hz, 2H), 4.46 (dd, *J* = 12.0, 3.4 Hz, 2H), 4.33 (dd, *J* = 12.0, 4.8 Hz, 2H), 3.91 (d, *J* = 7.8 Hz, 2H), 3.83 (d, *J* = 9.2 Hz, 2H), 3.52-3.12 (m, 12H), 1.48-1.35 (m, 4H), 1.33-1.16 (m, 16H), 0.86 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 166.2, 166.0, 165.3, 164.9, 133.7, 133.6, 133.3, 130.2, 130.0, 129.9, 129.8, 129.1, 129.0, 128.6, 128.5, 101.7, 72.8, 72.1, 71.8, 71.7, 69.9, 69.1, 68.9, 63.2, 45.2, 35.0, 29.7, 29.3, 26.3, 22.8, 14.3 ; **MS (MALDI-TOF)**: calcd. for $C_{87}H_{92}O_{22}$ $[M+Na]^+$ 1511.5973, found 1511.5973.

GNG-1 and **GNG-2** were synthesized according to the general procedure for de-*O*-benzoylation.

GNG-1: **^{1}H NMR** (300 MHz, $CDCl_3$): δ 4.33 (d, *J* = 7.7 Hz, 2H), 4.00-3.85 (m, 4H), 3.76-3.60 (m, 4H), 3.55-3.28 (m, 14H), 3.23 (t, *J* = 8.4 Hz, 2H), 1.64-1.53 (m, 4H), 1.46-1.30 (m, 12H), 0.95 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 105.3, 78.2, 78.0, 75.4, 72.8, 71.8, 70.6, 70.3, 62.9, 46.7, 33.0, 30.9, 27.2, 23.9, 14.6; **MS (MALDI-TOF)**: calcd. for $C_{29}H_{56}O_{14}$ $[M+Na]^+$ 651.3563, found 651.3530.

GNG-2: **^{1}H NMR** (300 MHz, $CDCl_3$): δ δ 4.33 (d, *J* = 7.7 Hz, 2H), 4.00-3.85 (m, 4H), 3.76-3.60 (m, 4H), 3.55-3.28 (m, 14H), 3.23 (t, *J* = 8.4 Hz, 2H), 1.64-1.53 (m, 4H), 1.46-1.30 (m, 16H), 0.95 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 105.3, 78.2, 78.0, 75.3, 72.8, 71.8, 70.6, 70.3, 62.9, 46.7, 33.2, 30.9, 30.5, 27.5, 23.9, 14.6 ; **MS (MALDI-TOF)**: calcd. for $C_{31}H_{60}O_{14}$ $[M+Na]^+$ 679.3876, found 679.3857.

Supplementary scheme 2



(a) NaH, 1-iodohexane or 1-iodoheptane, THF, room temperature; $LiAlH_4$, THF, room temperature, 93% (two steps); (b) perbenzoylated glucosylbromide (2.4 equiv.), AgOTf, CH_2Cl_2, -45°C → room temperature, 92%; (c) NaOMe, MeOH, room temperature, 95%.

2,2-Bis-hexyl-propane-1,3-diol (**3**) and 2,2-bis-heptyl-propane-1,3-diol (**4**) were synthesized according to the reported procedure[7] by using 1-iodohexane or 1-iodoheptane instead of 1-iododecane as described in **supplementary scheme 2**.

2,2-Bis-hexyl-propane-1,3-diol (**3**): **^{1}H NMR** (300 MHz, $CDCl_3$): δ 3.56 (s, 4H), 2.52 (s, 2H), 1.38-1.12 (m, 20H), 0.89 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 69.7, 41.2, 32.0, 31.0, 30.5, 23.0, 22.9, 14.3 ; **MS (MALDI-TOF)**: calcd. for $C_{13}H_{28}O_2$ $[M+Na]^+$ 239.1982, found 239.1973.

2,2-Bis-heptyl-propane-1,3-diol (**4**): **^{1}H NMR** (300 MHz, $CDCl_3$): δ 3.56 (s, 4H), 2.69 (s, 2H), 1.38-1.12 (m, 24H), 0.89 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 69.6, 41.2, 32.1, 31.0, 30.8, 29.5, 23.1, 22.9, 14.3; **MS (MALDI-TOF)**: calcd. for $C_{15}H_{32}O_2$ $[M+Na]^+$ 267.2295, found 267.2296.

GNG-3a and **GNG-4a** were synthesized according to the general procedure for glycosylation.

GNG-3a: ^{1}H NMR (300 MHz, $CDCl_3$): δ δ 8.21-8.15 (m, 4H), 8.06-7.96 (m, 4H), 7.95-7.89 (m, 4H), 7.89-7.82 (m, 4H), 7.73-7.66 (m, 6H), 7.58-7.24 (m, 20H), 5.58 (t, *J* = 9.6 Hz, 2H), 5.47 (t, *J* = 9.6 Hz, 2H), 5.39 (dd, *J* = 9.6, 7.8 Hz, 2H), 4.44 (dd, *J* = 12.2, 3.2 Hz, 2H), 4.32 (dd, *J* = 12.2, 5.2 Hz, 2H), 3.75 (d, *J* = 7.8 Hz, 2H), 3.49 (d, *J* = 9.1 Hz, 2H), 3.14-3.06 (m, 2H), 2.92 (d, *J* = 8.9 Hz, 2H), 1.30-0.93 (m, 20H), 0.58 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 166.2, 166.0, 165.3, 164.8, 133.8, 133.7, 133.6, 133.4, 133.3, 130.4, 130.3, 130.2, 130.1, 130.0, 129.9, 129.8, 129.7, 129.4, 129.2, 129.1, 129.0, 128.7, 128.6, 128.5, 101.5, 72.7, 72.1, 71.8, 71.7, 69.9, 63.1, 40.6, 31.9, 30.6, 30.2, 22.8, 22.4, 14.3; **MS (MALDI-TOF)**: calcd. for $C_{81}H_{80}O_{20}$ $[M+Na]^+$ 1395.5, found 1395.6

GNG-4a: ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.21-8.15 (m, 4H), 8.06-7.96 (m, 4H), 7.95-7.89 (m, 4H), 7.89-7.82 (m, 4H), 7.73-7.66 (m, 6H), 7.58-7.24 (m, 20H), 5.58 (t, *J* = 9.6 Hz, 2H), 5.47 (t, *J* = 9.6 Hz, 2H), 5.39 (dd, *J* = 9.6, 7.8 Hz, 2H), 4.44 (dd, *J* = 12.2, 3.2 Hz, 2H), 4.32 (dd, *J* = 12.2, 5.2 Hz, 2H), 3.75 (d, *J* = 7.8 Hz, 2H), 3.49 (d, *J* = 9.1 Hz, 2H), 3.14-3.06 (m, 2H), 2.92 (d, *J* = 8.9 Hz 2H), 1.30-0.93 (m, 24H), 0.58 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 166.2, 166.0, 165.3, 164.9, 133.8, 133.7, 133.6, 133.4, 133.3, 130.4, 130.3, 130.2, 130.1, 130.0, 129.9, 129.8, 129.7, 129.4, 129.2, 129.1, 128.7, 128.6, 128.5, 101.5, 72.8, 72.1, 71.9, 71.7, 70.0, 63.2, 40.6, 32.1, 30.6, 30.5, 29.4, 22.9, 22.4, 14.3; **MS (MALDI-TOF)**: calcd. for $C_{83}H_{84}O_{20}$ $[M+Na]^+$ 1423.5, found 1424.0

GNG-3 and **GNG-4** were synthesized according to the general procedure for de-*O*-benzoylation.

GNG-3: ^{1}H NMR (300 MHz, $CDCl_3$): δ 4.36 (d, *J* = 7.7 Hz, 2H), 3.90 (d, 2H), 3.77 (d, 2H), 3.74-3.67 (m, 2H), 3.43-3.29 (m, 8H), 3.23 (t, *J* = 8.4 Hz, 2H), 1.42-1.19 (m, 20H), 0.94 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 105.1, 78.4, 77.9, 75.4, 73.3, 71.9, 63.0, 42.2, 33.1, 32.0, 31.7, 31.5, 23.9, 23.7, 14.6; **MS (MALDI-TOF)**: calcd. for $C_{25}H_{48}O_{12}$ $[M+Na]^+$ 563.3038, found 563.3034.

GNG-4: ^{1}H NMR (300 MHz, $CDCl_3$): δ 4.36 (d, *J* = 7.7 Hz, 2H), 3.90 (d, 2H), 3.77 (d, 2H), 3.74-3.67 (m, 2H), 3.43-3.29 (m, 8H), 3.23 (t, *J* = 8.4 Hz, 2H), 1.42-1.19 (m, 24H), 0.94 (t, 6H); **^{13}C NMR** (75 MHz, $CDCl_3$): δ 105.1, 78.4, 77.9, 75.3, 73.3, 71.9, 63.0, 42.2, 33.2, 31.9, 31.7, 30.5, 23.9, 23.7, 14.6; **MS (MALDI-TOF)**: calcd. for $C_{27}H_{52}O_{12}$ $[M+Na]^+$ 591.3351, found 591.3361

References

1. Bazzacco, P.; Sharma, K. S.; Durand, G.; Giusti, F.; Ebel, C.; Popot, J.-L.; Pucci,B. *Biomacromolecules* **2009**, *10*, 3317-3326.
2. Chae, P. S. et al., *J. Am. Chem. Soc.* **2010**, *132*, 16750-16752.
3. Quick, M.; Javitch, J. A. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3603-3608.
4. Rosenbaum, D. M. et al., *Science* 2007, *318*, 1266-1273.
5. Kobilka, B. K. *Anal Biochem* 1995, *231*, 269-271.
6. Ashton, P. R.; Boyd, S. E.; Brown, C. L.; Jayaraman, N.; Nepogodiev, S. A.; Stoddart, J. F. *Chem. Eur. J.* **1996**, *2*, 1115-1128.
7. Chae, P. S. et al., *Nat. Methods* **2010**, *7*, 1003-1008.

References (in main text)

(8) (l) P. S. Chae, K. Gotfryd, J. Pacyna, L. J. W. Miercke, S. G. F. Rasmussen, R. A. Robbins, R. R. Rana, C. J. Loland, B. Kobilka, R. Stroud, B. Byrne, U. Gether, S. H. Gellman, *J. Am. Chem. Soc.* **2010**, *132*, 16750-16752. (m) Chae, P. S.; Rasmussen, S. G. F.; Rana, R. R.; Gotfryd, K.; Chandra, R.; Goren, M. A.; Kruse, A. C.; Nurva, S.; Loland, C. J.; Pierre, Y.; Drew, D.; Popot, J. –L.; Picot, D.; Fox, B. G.; Guan, L.; Gether, U.; Byrne, B.; Kobilka, B.; Gellman, S. H. *Nat. Methods* **2010**, *7*, 1003-1008. (n) Hovers, J.; Potschies, M.; Polidori, A.; Pucci, B.; Raynal, S.; Bonneté, F.; Serrano-Vega, M. J.; Tate, C. G.; Picot, D.; Pierre, Y.; Popot, J. –L; Nehmé, R.; Bidet, M.; Mus-Veteau,

I.; Busskamp, H.; Jung, K. –H.; Marx, A.; Timmins, P. A.; Welte, W. *Mol. Membr. Biol.* **2011**, *28*, 170-180. (o) Chae, P. S.; Rasmussen, S. G. F.; Rana, R. R.; Gotfryd, K.; Kruse, A. C.; Nurva, S.; Gether, U.; Guan, L.; Loland, C. J.; Byrne, B.; Kobilka, B. K.; Gellman,S. H. *Chem. Eur. J.* **2012**, *18*, 9485-9490

(9) (a) Rasmussen, S. G. F.; Choi, H. -J.; Fung, J. –J; Casarosa, P.; Pardon, E.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. *Nature* **2011**, *469*, 236-240. (b) Rosenbaum, D. M.; Zhang, C.; Lyons, J.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H. -J.; DeVree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shwa, D. E.; Weis, W. I.; Caffrey, M.; Gmeiner, P.; Kobilka, B. K. *Nature* **2011**, *469*, 175-180. (c) Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunnahara, R. K.; Kobilka, B. K. *Nature* **2011**, *477*, 540-541. (d) Kruse, A.C.; Hu, J.; Pan, A. C.; Arlow, D. H.; Rosenbaum, D. M.; Rosemond, E.; Green, H. F.; Liu, T.; Chae, P. S.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Wess,J.; Kobilka, B. K. *Nature* **2012**, *482*, 552-556. (e) Haga, K.; Kruse, A. C.; Asada, H.; Kobayashi, T. Y.; Shiroishi, M.; Zhang, C.; Weis, W. I.; Okada, T.; Kobilka, B. K.; Haga, T.; Kobayashi, T. *Nature* **2012**, *482*, 552-556. (f) Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S. *Nature* **2012,** *485*, 321-326.

(14) G. Deckert, P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, R. V. Swanson. *Nature* **1998**, *392*, 353-358.