Platform Reagents Enable Synthesis of Ligand-Directed Covalent Probes: Study of Cannabinoid Receptor 2 in Live Cells


[a] Laboratorium für Organische Chemie, Eidgenössische Technische Hochschule Zürich, Vladimir-Prelog-Weg 3, 8093 Zürich, Switzerland
[b] Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham NG7 2UH, UK; Centre of Membrane Proteins and Receptors (COMPARE), University of Birmingham and University of Nottingham, Midlands, UK
[c] Department of Chemical Physiology & Biochemistry, Oregon Health & Science University, Portland, Oregon 97239-3098, United States
[d] Institute of Biomolecular Chemistry, National Research Council (ICB-CNR), Via Campi Flegrei 34, 80078 Pozzuoli, Italy
[e] School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK
[f] Department of Chemical Biology, Max Planck Institute for Medical Research, Jahnstr. 29, 69120 Heidelberg, Germany
[g] Roche Pharma Research & Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., 4070 Basel, Switzerland
[h] Vollum Institute, Oregon Health & Science University, Portland, Oregon 97239-3098, United States
Supporting information

Table of Contents

SUPPLEMENTARY FIGURES........................................................................................................... 3
SUPPLEMENTARY TABLES ............................................................................................................. 28
MOLECULAR MODELING ............................................................................................................... 29
MOLECULAR DYNAMICS: COMPUTATIONAL METHODS ............................................................ 29
REFERENCES ................................................................................................................................. 30
Figure S1. Root-mean-square deviation (rmsd) plot over 1 µs of MD1 and MD2 trajectories calculated for 2b and 2z, after best fitting of the protein backbone atoms (except the flexible N-terminus and ICL3 loop). The rmsd plot was smoothed with a five-point window running average.
Figure S2. Distance plot between the electrophilic carbon atom (C*) of 2b and $N_{\varepsilon}$ of either Lys23 or Lys103 monitored over the two MD1 and MD2 trajectories in the 2b-CB$_2$R complex.
Figure S3. Distance plot calculated over MD2 trajectory between electrophilic carbon atom ($C^*$) of 2b and $N\delta/\varepsilon$ of His98 side chain in the 2b-CB2R complex.
Figure S4. TR-FRET based characterization of covalent NBD transfer. SNAP-Lumi4-Tb labelled CB₂R membranes were treated with O-NBD probe 2a, 2c, 2d or 2e (100 nM). Following incubation for 120 min. at 25 °C SR-144,528 (1 µM) was added. N = 3.
Figure S5. TR-FRET based characterization of covalent fluorophore transfer. SNAP-Lumi4-Tb labelled CB2R membranes were treated with ligand-directed covalent (LDC) probe 11a, 11b, 11c, 11d or 11e (100 nM). Following incubation for 120 min at 25 °C SR-144,528 (1 µM) was added. N = 3.
Figure S6. TR-FRET based characterization of covalent fluorophore transfer. SNAP-Lumi4-Tb labelled CB2R membranes were treated with LDC probe 11a, 11b, 11c, 11d or 11e (100 nM). Following incubation for 20 h at 25 °C SR-144,528 (1 µM) was added. N = 3.
Figure S7. TR-FRET based characterization of covalent fluorophore transfer. SNAP-Lumi4-Tb labelled CB2R membranes were treated with ligand-directed covalent (LDC) probe 11e (100 nM) or 15 (200 nM). Following incubation for 120 min at 37 °C SR-144,528 (1 µM for 11e and 10 µM for 15) was added. N = 3.
Figure S8. Solution spectra of probes 11e (A, 3 µM) and 15 (B, 10 µM) in DMSO.
Figure S9. Saturation binding profile following a 2 h incubation at 37 °C with 11e using a membrane preparation of HEK293 cells expressing either CB₁R or CB₂R, previously labelled with SNAP-Lumi4-Tb.
Figure S10. Root-mean-square deviation (rmsd) plot over MD1 and MD2 trajectories calculated for 15z and 2z after best fitting of protein backbone atoms (except the flexible N-terminus and ICL3 loop). The rmsd plot was smoothed with a five-point window running average.
Figure S11. Distance monitored over MD1 and MD2 trajectories between the electrophilic carbon atom (C*) of 15z and Nε of either Lys23 or Lys103 in 15z-CBzR complex.
Figure S12. Representative frames from MD1 of 15z-CB2R complex featuring the shortest distances between the electrophilic carbon (C*) of 15z (in ball representation) and Lys103 Nε atom. A stick representation is used for heavy atoms of the ligand (colored in olive drab) and for protein sidechains (colored in light gray according to the ribbon) within 5 Å of the ligand. Water molecules within 10 Å from the ligand involved in H-bonds are shown in stick representation. Hydrogen, nitrogen, oxygen, and sulfur atoms are painted white, blue, red, and yellow, respectively. Half-transparency is employed for the ribbon representation of protein regions overlying the ligand in the selected view. A green and pink wire representation is adopted for H-bonds and (C*)-Nε distance, respectively.
**Figure S13.** Representative frames from MD1 of 15z-CB2R complex featuring the shortest distances between the electrophilic carbon (C*) of 15z (in ball representation) and Lys23 Ne atom. A stick representation is used for heavy atoms of the ligand (colored in olive drab) and for protein sidechains (colored in light gray according to the ribbon) within 5 Å of the ligand. Water molecules within 10 Å from the ligand involved in H-bonds are shown in stick representation. Hydrogen, nitrogen, oxygen, and sulfur atoms are painted white, blue, red, and yellow, respectively. Half-transparency is employed for the ribbon representation of protein regions overlying the ligand in the selected view. A green and pink wire representation is adopted for H-bonds and (C*)-Ne distance, respectively.
Figure S14. Representative frames from MD1 of 15z-CB2R complex featuring the shortest distances between the electrophilic carbon (C*) of 15z (in ball representation) and Lys33 Nε atom. A stick representation is used for heavy atoms of the ligand (colored in olive drab) and for protein sidechains (colored in light gray according to the ribbon) within 5 Å of the ligand. Water molecules within 10 Å from the ligand involved in H-bonds are shown in stick representation. Hydrogen, nitrogen, oxygen, and sulfur atoms are painted white, blue, red, and yellow, respectively. Half-transparency is employed for the ribbon representation of protein regions overlying the ligand in the selected view. A green and pink wire representation is adopted for H-bonds and (C*)-Nε distance, respectively.
Figure S15. Distance plot calculated over MD1 and MD2 trajectories between the electrophilic carbon atom (C*) of 15z and either Nε of Lys33 or Nδ/ε of His98 side chain in the 15z-CB2R complex.
Figure S16. Conventional flow cytometry of HEK293 T-Rex hCB₂R expressing cells (A) and HEK293 wild-type cells (B) upon incubation with 2b (6.25 – 200 nM). Non-specific binding was determined in cells pre-incubated with SR-144,528 (10 µM). The mean fluorescence intensity (MFI) is shown with subtracted background.
**Figure S17.** Conventional flow cytometry of HEK293 T-Rex hCB2R expressing cells (A) and HEK293 wild-type cells (B) upon incubation with 11e (6.25 – 200 nM). Non-specific binding was determined in cells pre-incubated with SR-144,528 (10 µM). The mean fluorescence intensity (MFI) is shown with subtracted background.
**Figure S18.** Exemplary gallery images of imaging flow cytometry with 2b: (A) 2b, MIQ = –0.5 - 0.5. (B) 2b, MIQ > 2. (C) 2b + SR-144,528, MIQ = –0.5 - 0.5. (D) 2b + SR-144,528, MIQ > 2.
**Figure S19.** Exemplary gallery images of imaging flow cytometry with 11e; (A) 11e, MIQ < −1. (B) 11e, MIQ = −0.5 - 0.5. (C) 11e, MIQ > 2. (D) 11e + SR-144,528, MIQ < −1. (E) 11e + SR-144,528, MIQ = −0.5 - 0.5. (F) 11e + SR-144,528, MIQ > 2.
Figure S20. Exemplary gallery images of imaging flow cytometry with 15: (A) 15, MIQ < –1. (B) 15, MIQ = –0.5 - 0.5. (C) 15, MIQ > 2. (D) 15 + SR-144,528, MIQ < –1. (E) 15 + SR-144,528, MIQ = –0.5 - 0.5. (F) 15 + SR-144,528, MIQ > 2.
Figure S21. Exemplary gallery images of imaging flow cytometry with control cells (no probe); (A) used for 2b and (B) used for 11e and 15.
Figure S22. Imaging flow cytometry data of 2b; (A) Histogram. (B) Positive of 2b. (C) Positive of 2b + SR-144,528. (D) Negative control. (E) Median internalization quotient (MIQ) shown as a comparison of internal to whole cell fluorescence intensity.
Figure S23. Imaging flow cytometry data of 11e; (A) Histogram. (B) Positive of 11e. (C) Positive of 11e + SR-144,528. (D) Negative control. (E) Median internalization quotient (MIQ) shown as a comparison of internal to whole cell fluorescence intensity.
Figure S24. Imaging flow cytometry data of 15: (A) Histogram. (B) Positive of 15. (C) Positive of 15 + SR-144,528. (D) Negative control. (E) Median internalization quotient (MIQ) shown as a comparison of internal to whole cell fluorescence intensity.
**Figure S25.** 60× fluorescent labeling of CB₂R using 15 in (A) AtT-20(CB₂R), (B) AtT-20-wild-type cells, and (C) AtT-20(CB₂R) in the presence of CB₂R antagonist SR-144,528. Nuclei were stained with Hoechst 33342.
SUPPLEMENTARY TABLES

**Table S1.** Evaluation of chemical stability of selected probes in the Aqueous stability assay (ASTA). Data shown is the difference of peak area at time points 0 and 2 h as determined by HPLC.

<table>
<thead>
<tr>
<th>Buffer pH</th>
<th>2c</th>
<th>2d</th>
<th>2e</th>
<th>11a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105</td>
<td>97</td>
<td>102</td>
<td>103</td>
</tr>
<tr>
<td>4</td>
<td>94</td>
<td>103</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>99</td>
<td>105</td>
<td>98</td>
<td>104</td>
</tr>
<tr>
<td>8</td>
<td>102</td>
<td>103</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>_†</td>
<td>_†</td>
<td>_†</td>
<td>_†</td>
</tr>
</tbody>
</table>

 Degradation peak already observed at t = 0

**Table S2.** Median fluorescence intensity (MFI) and percentage of cells with fluorescence intensity greater than threshold (%positive) of imaging flow cytometry experiments with 2b. The corresponding relative MFI (rMFI) is calculated as the ratio of MFI of labelled cells to MFI of unlabeled cells (control).

<table>
<thead>
<tr>
<th>Compound</th>
<th>MFI</th>
<th>rMFI</th>
<th>%positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>15052</td>
<td>1.00</td>
<td>3.29</td>
</tr>
<tr>
<td>2b</td>
<td>28589</td>
<td>1.90</td>
<td>16.7</td>
</tr>
<tr>
<td>2b + SR-144,528</td>
<td>23916</td>
<td>1.59</td>
<td>10.9</td>
</tr>
</tbody>
</table>

**Table S3.** Median fluorescence intensity (MFI) and percentage of cells with fluorescence intensity greater than threshold (%positive) of imaging flow cytometry experiments with 11e or 15. The corresponding relative MFI (rMFI) is calculated as the ratio of MFI of labelled cells to MFI of unlabeled cells (control).

<table>
<thead>
<tr>
<th>Compound</th>
<th>MFI</th>
<th>rMFI</th>
<th>%positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>6003</td>
<td>1.00</td>
<td>3.86</td>
</tr>
<tr>
<td>11e</td>
<td>56143</td>
<td>9.35</td>
<td>88.2</td>
</tr>
<tr>
<td>11e + SR-144,528</td>
<td>34997</td>
<td>5.83</td>
<td>82.3</td>
</tr>
<tr>
<td>15</td>
<td>50892</td>
<td>8.48</td>
<td>86.6</td>
</tr>
<tr>
<td>15 + SR-144,528</td>
<td>24551</td>
<td>4.09</td>
<td>62.3</td>
</tr>
</tbody>
</table>
MOLECULAR MODELING

The Cryo-EM structure of CB$_2$R complexed with the agonist WIN 55,212-2 (6PT0) was used as a template to dock CB$_2$R ligands. The docking experiments were performed with the software GOLD (CCDC) with default settings. The best 10 docking poses for each compound were energy-minimized within the binding pocket using MOE (CCG, Montreal) and examined visually to select the most reasonable docking pose with respect to molecular interactions and internal conformational strain. The final selection was based on checking consistency with the available structure-activity relationship information.

MOLECULAR DYNAMICS: COMPUTATIONAL METHODS

The inactivating mutations present in PDB:5ZTY were back-mutated to wild type residues and the missing intracellular loop 3 (ie3) was modeled using Modeller program. Both 2b and 15 probes were optimized using the GAMESS program at the Hartree–Fock level with the STO-3G basis set, followed by a single-point HF energy evaluation at the 6-31G* level to derive partial atomic charges for the ligand by the RESP procedure. The energy-minimized complexes were embedded in a pre-equilibrated palmitoyl-oleoyl-phosphatidyl-choline (POPC) lipid bilayer and solvated in an aqueous medium using the charmngui web-interface (http://www.charmmgui.org). Potassium and chloride ions were added to ensure electric neutrality and 0.15 M ionic strength. MD simulations were carried out with the pmemd.cuda module of the Amber20 package, using lipid 14 (lipids), ff14SB force (proteins), and gaff (ligands) force field parametrization. The system underwent 10,000 steps of energy minimization keeping the solute atoms harmonically restrained to their starting positions ($K_r = 10$ kcal mol$^{-1}$ Å$^{-1}$). The system was gradually heated to 100 K for 500 ps keeping fixed solute and lipid atoms using a Langevin thermostat, followed by a subsequent heating up to the final temperature of 310 K for 1 ns, keeping fixed the solute and lipid atoms, using an anisotropic Berendsen weak-coupling barostat to equilibrate the pressure. The system then underwent equilibration with positional restraints on the solute: 25 ns with restraints on all solute atoms ($K_r = 5$ kcal mol$^{-1}$ Å$^{-1}$), followed by 5 ns with restraints on Cα protein atoms alone ($K_r = 5$ kcal mol$^{-1}$ Å$^{-1}$). The production run was carried out for 1 µs. The ligand cluster analysis was done with the hierarchical agglomerative method after best fit of protein backbone. The Cpptraj module and the UCSF Chimera 1.17 program were used to perform MD analysis and to draw the figures, respectively.
REFERENCES


