A modular click ligand-directed approach to label endogenous aminergic GPCRs in live cells

Xavier Gómez-Santacana[a,b]*, Marin Boutonnet[a], Carles Martínez-Juvés[b], Juan Lorenzo Catena[b], Enora Moutin[a], Thomas Roux[c], Eric Trinquet[c], Laurent Lamarque[c], Julie Perroy[a], Laurent Prézeau[a], Jurriaan M. Zwier[c†], Jean-Philippe Pin[a]*, Amadeu Llebaria[b]*.

[a] Institut de Génomique Fonctionnelle, Université de Montpellier, UMR 5203 CNRS and U 1191 INSERM, France. [b] MCS, Institute for Advanced Chemistry of Catalonia – CSIC, Barcelona, Spain. [c] PerkinElmer, Codolet, France.

KEYWORDS. ligand-directed labeling, protein labeling, G protein-coupled receptors, GPCR, dopamine receptors, D_1 receptor, fluorescence, endogenous proteins, native proteins.

ABSTRACT: New technologies based on luminescence have been essential to monitor the organization, signaling, trafficking or ligand binding of G Protein-Coupled Receptors (GPCRs), but they rely on the overexpression of genetically modified receptors. As more and more studies indicate the importance of studying native receptors in their natural environment, it is essential to develop approaches allowing the specific labeling of native receptors. Here we report an innovative ligand directed approach to specifically label residues of native GPCRs upon ligand binding. To this end, we developed a ligand-directed toolbox based on a novel approach that uses molecular modules to build fluorescent ligand-directed probes that can label an archetypical aminergic GPCR (D_1R). Our probes can be readily prepared before the labeling reaction from two molecular modules: an activated electrophilic linker which includes a fluorescent dye and a GPCR ligand that may include nucleophilic groups. Thanks to a fast and specific click reaction, the nucleophilic ligand can barely react with the activated linker before it is bound to the native target GPCR and the labeling reaction occurs. Subsequently, the ligand unbinds the GPCR pocket, leaving the receptor fluorescently labeled and fully functional. This novel labeling approach allowed us to label both D_1 receptor in transfected cells and native receptors in neuronal cell lines. This approach will pave the way to develop new reagents and assays to monitor endogenous GPCRs distribution, trafficking, activity or binding properties in their native environment.

INTRODUCTION

G-protein coupled receptors (GPCR) are a complex superfamily of membrane proteins that have a key role in regulating many key physiological processes. More than 700 marketed drugs target GPCRs, representing around 35% of the approved drugs.[1] Due to their biological relevance, there is a high interest in both academy and industry in unveiling the signal transduction and developing new ligands and compound tools that could be ultimately used in classical medicine, but also in innovative theranostic fields.

Traditionally, technologies to monitor the organization, trafficking, signaling or ligand binding of GPCRs were based on radioactive approaches but, in the last two decades, new technologies based on luminescence have been developed. These technologies rely on the overexpression of genetically modified (and/or fluorescently tagged) receptors or proteins of interest and include popular assay technologies such as BRET,[2] FRET [3] or time resolved FRET (TR-FRET) using lanthanide complexes. [4,5] One successful example is the Tag Lite® technology developed by Cisbio Bioassays (now PerkinElmer), which is based on expressing a genetically modified receptor or protein of interest fused to a suicide enzyme tag, such as the SNAP-tag® (ST), CLIP-tag® (CT) or HALO-tag [6] (Figure 1A) in recombinant living cells. These tagged receptors can be covalently labeled with any fluorophores linked to a benzylguanine or benzylcytosine able to form a covalent bond with the SNAP-tag® or CLIP-tag®, respectively (Figure 1A). The use of lanthanide cryptates (such as Lumi4-Tb®) allows time-resolved fluorescent or TR-FRET measurements, that can be used in internalization, oligomerization or binding assays.[4,7] This allows to perform saturation or competition final point binding assays or association, dissociation or association-competition kinetic binding assays.[8]

As stated above, these approaches rely on the use of artificial systems, involving genetically manipulated proteins and the corresponding overexpression in heterologous cells. However, the properties of a modified receptor overexpressed in heterologous cells are unlikely similar to those of the native receptor in its natural environment. Moreover, labeling native receptors may be useful for analyzing human samples for clinical studies. Therefore, an attractive objective in chemical biology is to extend the high performance assays and protein labeling selectivity of the SNAP-tag®-based technologies to unmodified receptors in their native environments such as in primary cell cultures or tissues.[6,9]

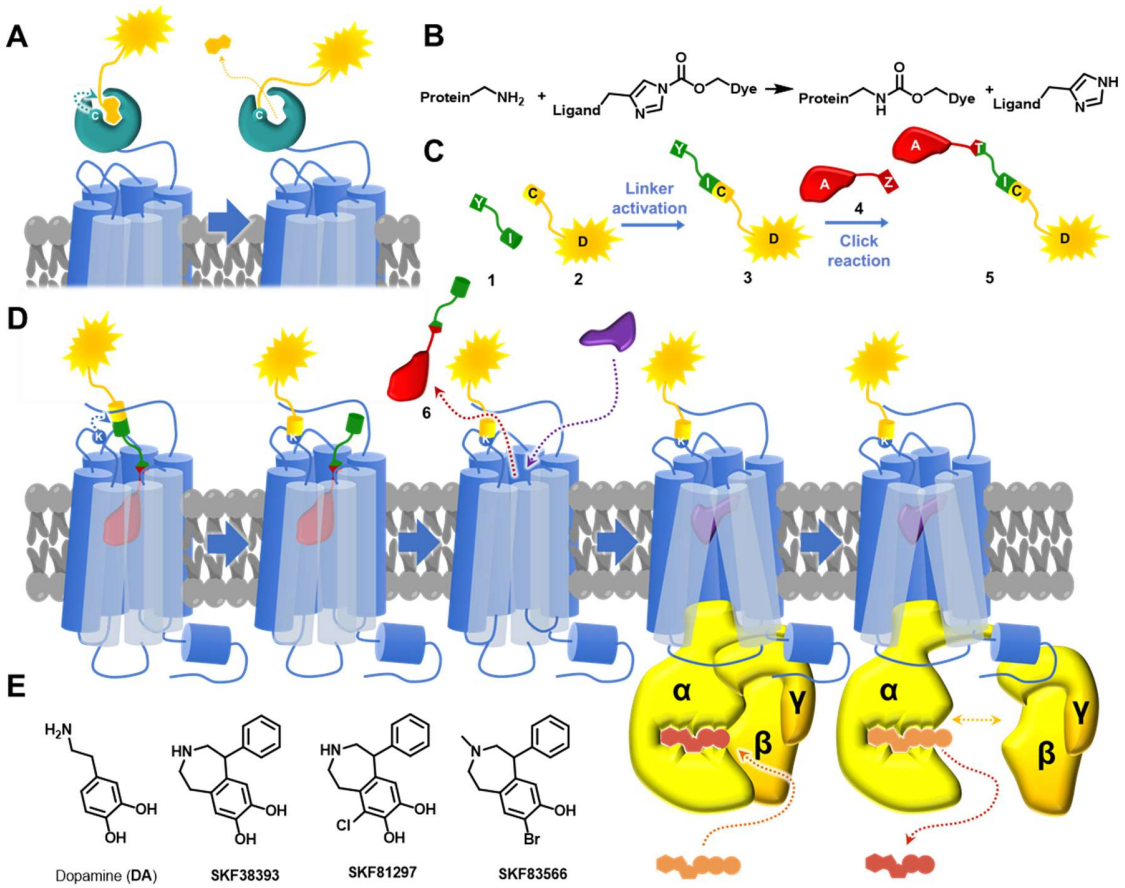


Figure 1: GPCR labeling approaches design. (A) GPCRs can be labeled using a SNAP-tag fused in the *N*-terminus. Fluorescent dyes covalently conjugated to a benzyl guanylyl (BG) moiety that bind to the SNAP-tag® and a cysteine residue reacts through a SN2 reaction to give the resulting SNAP-tag® labeling. (B) Ligand directed labeling approaches based on acyl imidazole reactive moiety (LD-AI) consist of label or tag covalently conjugated to an acyl imidazole reactive moiety, and this is conjugated to a specific ligand (i.e. affinity unit). Upon ligand binding, a native lysine can react with the activated acyl to form covalent labeling with the tag through a stable carbamate bond. (C) Click ligand-directed approaches (CLD) can be used with nucleophilic affinity units and are prepared in situ immediately before the labeling step. To prepare the ligands, an acyl group (**C**) linked to a dye (**D**) (**1**) is activated with an imidazole (I) linker equipped an alkynyl terminal (**Y**) (**2**), to form the activated clickable dye (**3**), which will be reacted with the ligand (**L**) linked to an azide (**Z**) terminus (**4**) through a copper-catalyzed click reaction, to give the final labeling compound (**5**). (D) Click ligand-directed (CLD) labeling is based on the binding of compound 5 to the target native GPCR and the subsequent reaction of a lysine residue in the vicinity with the imidazole-activated carboxylate (**IC**) of the linker. A covalent carbamate bond with the dye is formed and the resulting ligand 6 can be released and washed-out, leaving a native GPCR labeled and fully functional. Other agonists or antagonists will be able to bind the labeled receptor after the procedure.

A critical challenge of such an approach is to define a suitable chemical reaction to selectively label the native protein of interest. Several approaches to conjugate chemical labels to native amino acid residues have been reported to date.[10] However, the low reactivity and abundance of certain amino acids residues add difficulties to the specificity of the conjugation reaction to the desired residue and this difficulty is growing as the complexity of the system is increasing: from isolated proteins to cell cultures, tissues or even organisms. Ligand directed approaches, extensively developed in Hamachi's laboratory, may offer a solution to this problem. They combine in a single molecule an affinity unit for a protein of interest (i.e. a ligand), a fluorescent label and a reactive moiety (i.e. electrophilic group) with a relative geometry and distance that, upon ligand binding, directs the labeling reaction to an amino acid in the vicinity of a binding pocket. These reactive moieties include tosylates to conjugate mainly cysteines[11] or acyl imidazole (Figure 1B), dibromophenyl benzoates or *N*-acyl-*N*-alkyl sulfonamides for lysine labeling.[12–15] However, this approach requires the use of an affinity unit which is chemically compatible with the reactive moiety. For example, LD probes that include affinity units with nucleophilic functional groups (e.g. amine, thiol, phenol, etc) may be chemically unstable due to the possible intra- and intermolecular reaction with the electrophilic groups (reactive moiety). To overcome this limitations, affinity-guided catalyzed labeling with DMAP or oxime linked ligands have been proposed[16,17] but also the present click ligand-directed labeling (CLD) can be advantageous. In CLD, the labeling probe is assembled by a fast click reaction of the chemically incompatible affinity unit and the

reactive linker prior to the application on the native system to be labeled (Figure 1C).

In the present work, we report the development and validation of an innovative CLD approach to specifically and covalently attach fluorescent dyes to endogenous GPCRs maintaining the full functionality of the target receptor allowing its analysis in its native environment. In particular, our approach allows the labeling of dopamine receptor D_1, belonging to the class A aminergic GPCRs, which are commonly bound by ligands that include nucleophilic groups, such as amines or phenols. We have used a novel modular approach, which makes use of different molecular pieces combined by click chemistry to assemble a toolbox of ligand-directed labeling probes with red fluorescent dyes to label an endogenous GPCR, in this present case the D_1 receptor. This novel labeling approach allowed us to label native D_1 receptor both in transfected cells and endogenously expressed in neuronal cell lines.

RESULTS

Approach design and synthesis of the units. We used dopamine D_1 as an archetypical aminergic class A GPCR, which is mainly expressed in the central nervous system (CNS), especially in the dorsal and ventral striatum. Its endogenous agonist is the neurotransmitter dopamine (DA), however many synthetic agonists and antagonists have been reported to date, such as the agonists SKF38393 and SKF81297 and the potent antagonist SKF83566 (Figure 1E).

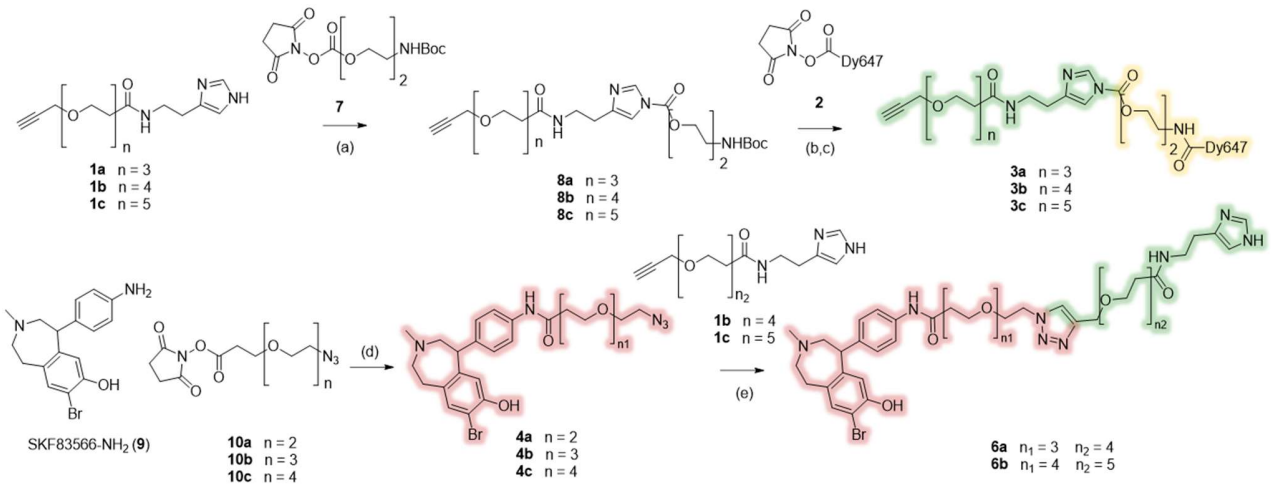
To design a molecular probe to label native D_1 receptor, we were inspired by the work of Hamachi group in the development of labeling approaches and we selected the acyl imidazole ligand-directed (LDAI) approach (Figure 1B). However, the nature of D_1 ligands is incompatible with LDAI for two reasons: (i) they all include a secondary or tertiary amine that can react with the acyl imidazole moiety or catalyze its hydrolysis and (ii) they include a phenol or a catechol that can be deprotonated and react with the acyl imidazole.

To avoid these problems, we decided to use a novel CLD labeling strategy (Figure 1CD). We used a clickable linker **1** equipped with an alkynyl terminus (**Y**) and an imidazole (**I**) to activate a carboxyl or a carbonyl group (**C**) of the molecular module **2** containing a red fluorescent dye (**D**, Dy647) (Figure 1C). After this carbonyl activation, we performed a click copper-catalyzed azide–alkyne cycloaddition (CuAAC) with the alkynyl moiety of the resulting activated dye **3** with the molecular module **4**, which consist of an affinity unit (i.e. D_1 antagonist, **A**) equipped with a linker with an azide terminus (**Z**). This fast click CuAAC reaction leads to a triazole (**T**) and yields the final labeling probes **5** (Figure 1C). Upon binding of the affinity unit (**A**), a receptor surface nucleophilic lysine (**K**) will attack the imidazole-activated carbonyl (**IC**) to form a covalent stable bond and thus label native D_1 receptor (Figure 1D). The labeling byproduct **6** will be washed away, leaving the D_1 orthosteric pocket unmodified and the receptor fully functional. Therefore, other ligands such as agonists, antagonist or even fluorescent antagonists that allows FRET with the labeled tag will be able to bind the receptor after the labeling process (Figure 1D).

Before synthesizing the different molecular modules of the CLD labeling probe, we first analyzed the amino acid sequence of D_1 receptor and its structure (obtained from GPCRdb[18,19]). We identified the lysine K165 or K167 as the potential labeling residues. These amino acids are located in the extracellular loop 2 (ECL2) and their solvent exposed amine was anticipated to react with an activated acyl-imidazole to form a permanent carbamate bond (Figure S1). We examined the distance from the binding pocket to the target lysines in the model and we observed that the linker should be accommodated through a channel between loops ECL2 and ECL3, which are rigidified by disulfur bridges, and then bend over the ECL2 to reach the target lysine residues (Figure S1). This significant number of variables makes difficult to define an optimal linker geometry and length that may give a proper labeling probe **5**. Therefore, we opted to explore flexible PEG linkers with different lengths, through a combination of different lengths of the subunits **1** and **3.** (Figure 1C, Scheme 1).

The key steps of the synthesis of the different CLD molecular modules are depicted in Scheme1. Nonetheless, the detailed synthesis of the different subunits is described in the supporting information. We selected the red dye Dy647-NHS (**2**) as a fluorescent partner that was required to be transformed into an acylimidazole reactive group. Thus, we synthesized three clickable imidazole PEG linkers with di-tri- or tetra-ethylene glycol, respectively, bearing an *O*-propargyl and a histamine carboxamide in each terminus respectively (**1a-c**, Scheme 1, supporting information). As activated imidazole alkylcarbonyl compounds are reported to be unstable,[20] we opted to add an extra linker to synthesize an activated imidazole carbonyloxy linkers (Figure 1B) as amine reactive groups, which would modify the protein lysine residues by forming a carbamate group. This involves three step synthesis, starting from the preparation of the acylimidazoles **8a-c** group from imidazole-containing clickable linkers **1a-c** and the NHS-activated carbonate with a BOC-protected amine **7** (Scheme 1, step a). Next deprotection of the Boc group in acidic conditions gave the amines, which reacted preferably with the fluorescent NHS-ester (**2**), o give the advanced molecular modules **3a-c** (Scheme 1, steps b-c).

As affinity unit, we selected the antagonist SKF83566 (Figure 1E) due to its high affinity and selectivity against D_1 receptor, which were considered likely to be retained after being modified with three different PEG-linkers (**4a-c**, Scheme1). This assumption was supported by related high affinity D_1 ligands based in the modification of SKF83566 structure.[21] First, the *para*-aniline analog of SKF83566 (SKF83566-NH_2, **9**) was synthesized (supporting information)[21] and acylated with an activated carboxylic acid with di- tri- or tetra-ethylene glycol linkers, respectively, having an terminal azido group (**10a-c**) to give clickable PEG-ligands **4a-c** (Scheme 1, step d). Additionally, two imidazole ligands **6a-b** were synthesized through a copper-catalyzed click (CuAAC) reaction from the corresponding azido ligand and alkynyl linker (Scheme 1). These compounds **6a-b** correspond to the by-products resulting from the D_1-receptor chemical labeling (Figure 1D)..



Scheme 1: Key synthetic steps for the synthesis of molecular modules 4, 6 and 3. Reagents and conditions: (a), MeCN, 60 ºC, 20-48h, 21-60%; Cu_2O 0.25 eq, BzOH 4.0 eq, THF/H2O 1:4, rt, 0.5-1.5 h, 25-79%, (c) Pyr 3.0 eq, DMF, 25 ºC, 3h, 27-79%; (d) TFA, 10K eq, 30 min, (f) DIPEA 5.0 eq, DMSO, 25ºC, 1 h, 15-33% (two steps).

Pharmacological characterization of the ligand units. The pharmacological activity of compounds **4a-c** and **6a-b** on human D_1 receptor involved both functional and binding assays and was compared with the activity of the parent D_1 antagonist SKF83566 (Figure 1E).

First, the antagonist behavior of compounds **4a-c** and **6a-b** was assessed using a functional HTRF® cAMP Gs Dynamic assay (PerkinElmer) in HEK293 cells transfected with human F-ST-D_1 receptor. Increasing concentrations of compounds SKF83566, **4a-c** and **6a-b** were able to inhibit the activation of D_1 induced by the agonist SKF38493 (300 nM). Despite all compounds display inhibitory potencies in the nanomolar range, the ligands equipped with the shorter PEG-linkers showed a higher potency, being SKF83566 the compound showing a higher pK_i and compound **9b** the less potency antagonist (table 1, Figure 2B).

The binding properties of compounds SKF83566**, 4a-c** and **6a-b** were assessed with Tag-Lite® assays (PerkinElmer) with HEK293 cells, previously transfected with human F-ST-D_1 receptor tagged with SNAP-Lumi4-Tb®. We first performed saturation binding experiments with green D_1 antagonist (GDA, PerkinElmer) and, afterwards, kinetic association-competition binding assays with increasing concentrations of the D_1 antagonists and a constant concentration of GDA 120 nM. We read the HTRF ratio every minute during a minimum of 60 minutes at 37 ºC. From the obtained binding curves, the kinetic and equilibrium binding properties of each ligand could be calculated, such as the on-rate constant (k_{on}), off-rate constant (k_{off}), residence time (RT) or dissociation constant (K_d) (Table 1, Figure S2). All ligands tested showed binding affinities in the nanomolar range with an inverse correlation of the linker length and binding affinity. The residence time (RT) observed by the different ligands are in the same range from 5 to 8 min, being the longer ligands **6a** and **6b** the ones with a longer residence time altogether with SKF83566. These results indicate that, once the labeling probe is prepared and after receptor labeling, short times of perfusion are needed to wash-out the generated labeling byproduct **6** and leave the receptor binding site available. Additionally, the last 6 points of each binding curve were selected and plotted to obtain a competition dose response curve of each ligand. The inhibition constant (K_i) values obtained from these curves were in line with the K_d values obtained from the association-competition curves (Table 1, Figure 2B). Interestingly, there is a shift of 1.2 to 1.5 log units between the binding affinity values (K_d, K_i) and the functional IC_{50} K_i, probably due to the different nature of competition between agonist and antagonists.

Table 1: Pharmacological properties of compounds 4a-c and 6a-b. The inhibition of D_1 functional activity by compounds **4a-c** and **6a-b** was assessed with and HTRF® cAMP Gs Dynamic assay with a constant 300 nM concentration of the agonist SKF38393. Tag-Lite® association-competition binding assays were performed with the to obtain the kinetic binding properties of compounds **4a-c** and **6a-b** with F-ST-D_1. With the same assay the corresponding K_i were also obtained. The values are means of a minimum of 3 independent experiments with the corresponding standard error of the mean (SEM).

	Functional	**Binding** (association-competition)		**Binding** (competition)
Compound	**pK_i ± SEM**	**RT ± SEM (min)**	**pK_d ± SEM**	**pK_i±SEM**
SKF83566	7.52 ± 0.18	5.6 ± 0.9	8.62 ± 0.14	8.71 ± 0.07
4a	7.03 ± 0.08	8.0 ± 1.2	8.10 ± 0.12	8.12 ± 0.11
4b	6.87 ± 0.12	5.0 ± 0.6	8.01 ± 0.08	8.13 ± 0.11
4c	6.85 ± 0.09	5.0 ± 1.0	7.94 ± 0.09	7.98 ± 0.07
6a	6.91 ± 0.10	7.5 ± 1.0	8.20 ± 0.10	8.20 ± 0.11
6b	6.02 ± 0.09	5.6 ± 1.2	7.26 ± 0.10	7.28 ± 0.12

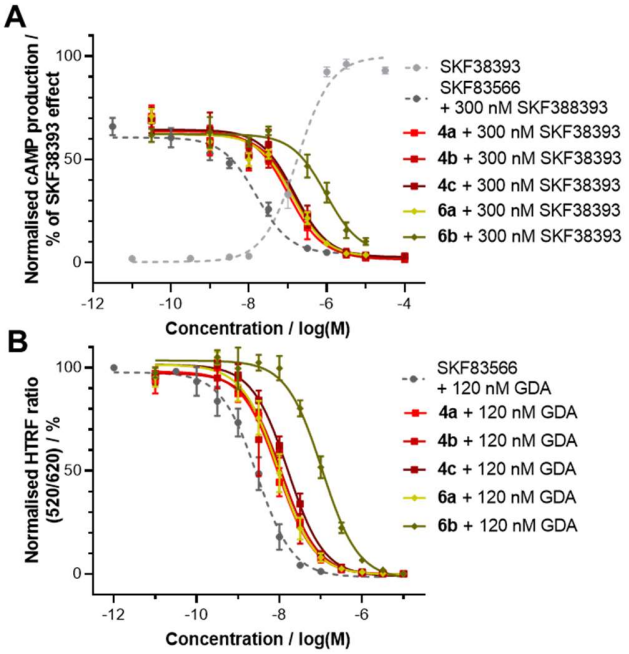


Figure 2: Pharmacological profile of units 1, 4 and 6: (A) Functional dose response curves of **4a-c** and **6a-b** with a constant 300 nM concentration of the agonist SKF38393 using a cAMP® assay with F-ST-D_1. (B) Competition binding curves of **4a-c** and **6a-b** with a constant 120 nM GDA extracted from TagLite® association-competition binding assays. Data points correspond to the means of a minimum of 3 independent experiments with the SEM as error bars.

Click reaction and receptor labeling. The final labeling probes **5a-i** were assembled just immediately before the application to the cells expressing human D_1 receptor through CuAAC reaction between the fluorescent alkyne linkers **3a-c** and the azido-ligands **4a-c** (Figure 3A). The click reaction to prepare **5a-i** was performed at 10-12 nmol scale in water with equimolecular quantities of the alkynyl-dye **3** and the azido ligand **4** (200 µM), 4 mol-eq of benzoic acid and Cu_2O (0.5-1.0 mg) to obtain the nine different possible compounds **5a-f**. As expected, the attempts to isolate the compounds **5** were unsuccessful, due to chemical degradation of the probes. Therefore, we opted to use directly the resulting solution, diluting it with buffer at pH=7.4 after removing the Cu_2O by decantation and treating it with a copper-absorbing resin to prevent cytotoxicity[22].

We expected that CLD probes are bound to the receptor and select a reactive lysine at distance to attack the acyl-imidazole group and form the *N*-acyl bond. However, this affinity directed labeling competes with the reaction of the CLD probes with other reactive molecules present in the cell system leading to non-specific reactions. To compare between the D_1 receptor selective affinity labeling and the non-specific labeling, probes **5** were also incubated with an excess of 50 µM of SKF83566. After a washing step the specific labeling of the receptor was calculated by the subtraction of the total one and the non-specific fluorescence readouts. However, to discard that the signal obtained corresponds to unreacted probe compound **5** bound to the receptor, the cells were later incubated for 0.5h with 50 µM of SKF83566 and washed. Thus, we were able to displace and remove any unreacted compound **5** and the formed reaction by-product **6** (Figure 1D). Next, the measured fluorescence corresponded only to the covalently labeled dye molecules (specific plus non-specific) allowing us to measure the extent of the chemical reaction with the CLD probes.

As a first step, we aimed to monitor the incubation time to optimize D_1 receptor labeling. Thus, we tested the labeling of F-ST-D_1 and HA-D_1 with **5e** as an intermediate-size CLD probe at five different concentrations with incubations of 0.5, 1, 2, 4 and 6 hours. The SNAP-tag form F-ST-D_1 was co-labeled with Lumi4-Tb to read TR-FRET measurements and obtain a cleaner signal for the optimization steps.[5] Since the Flag and the SNAP-tag® of F-ST-D_1 contain lysine residues that could be susceptible of labeling, the labeling experiments were performed in parallel with cells expressing HA-D_1, as HA-tag contains no lysine. The HTRF ratios obtained with F-ST-D_1 signal and the fluorescence of Dy647 obtained from HA-D_1 were compared in all the experiments. The results suggested that the incubations with micromolar concentrations of the reactive probe gave the maximum percentage of specific labeling after two hours (Figure S3). Thus, the labeling properties of acyl imidazole CLD probes **5a-i** were tested using the same protocol at 2.5 hours of incubation. Compound **5h** emerged as the probe giving the best results, which includes a significant separation between the non-specific labeling and the total labeling and a saturating behavior at high concentrations in HA-D_1 expressing cells.

To further evaluate the labeling properties of compound **5h** and, especially determining the labeling residue, we prepared two constructs of HA-D_1 with a single point mutation of the solvent-exposed lysines with arginines (K165R, K167R) and one construct with the double mutation (K165R-K167R). We selected arginine as a replacing amino acid of lysine since it maintains the positive charge and thus the structural properties of the lysine but it has a higher pK_a, involving a decreased nucleophilicity towards acyl transfer reactions at neutral pH. Thus, the **5h** would not hypothetically react with the introduced arginines, but the similar physical properties would maintain the structural and functional properties of the receptor. Indeed, cAMP assays revealed that the three mutants and wild type (WT) HA-D_1 are functionally equivalent (Table 1, S1; Figure S4).

We next performed the labeling experiments using the CLD labeling probe **5h** with the four HA-D_1 constructs. In such experiments, we found convenient to add a filtration step through a Durapore® PVDF filter (0.22 µm) after the click reaction and before the cell incubation. This was helping to remove possible compound aggregates in the cell mixture that could increase the non-specific fluorescence. In parallel, the F-ST-D_1 was labeled with BG-Dy647 to estimate of the fluorescence corresponding to the totality of receptors expressed at the cell surface. The results confirmed significant specific labeling of the wild type HA-D_1 receptor, resulting in labeling of 20±3% of the receptors labeled at 6 µM concentration and B_{max} of 27±3% (Figure 3B) after fitting the different experimental points with a Langmuir isotherm function. Interestingly, the mutated HA-D_1 K165R showed

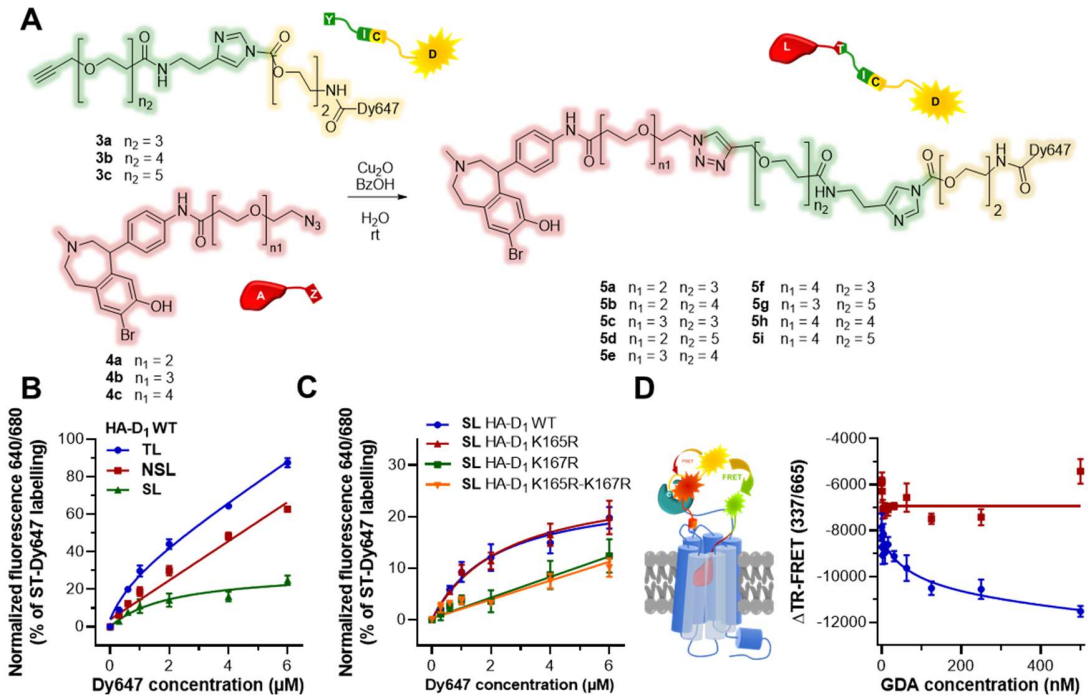


Figure 3: Click reaction to build the CLD labeling probe and labeling of D_1 receptor expressed in HEK293. (A) The click reaction was performed combining each compounds 3a-c with each 4a-c to give the 9 different combinations 5a-i. The concentration of compounds **3** and **4** was 200 µM concentration in presence of 4 eq of BzOH and Cu_2O catalysis. (B,C) The crude reaction of **5h** was diluted and employed to label HA-D_1 with increasing concentrations from 0 to 6 µM (dye concentration) alone (total labeling, blue dots) or in presence of 50 µM of the D_1 antagonist SKF83566 to measure the non-specific labeling (red dots and line). Subtraction of total labeling and non-specific labeling gave the specific labeling (green dots and line). Compound 5h labeled specifically HA-D_1 WT (B), while did not label specifically HA-D_1 K167R (C), indicating that K167 is the target residue for labeling. Graphs correspond to the mean of four replicates performed in triplicate with the SEM as error bars (D) Application of increasing concentrations of Green D_1 antagonist (GDA) in HEK293 cells expressing F-ST- D_1 labeled with both with 5h and BG-Lumi4-Tb decreased the TR-FRET produced between Lumi4-Tb and the red dye labeled on K167 for cells non-treated with the antagonist SKF83566 (blue dots and line). This indicates that D_1 binding pocket GDA was bound to an unoccupied binding site. No decrease of TR-FRET was observed if the same cells were treated with the antagonist SKF83566 (i. e. with the binding site occupied) (red dots and line). The graph is a representative example of two replicates performed in quintuplet and error bars correspond to the SEM.

very similar labeling properties: 20±4% of the receptors labeled at 6 µM and a B_{max} 29±8% (Figure 3C, S5). This indicates that the involvement of Lys165 in the D_1 labeling is limited. In contrast, HA-D_1 mutants K167R and K165R-K167R showed a reduced labeling compared to the wild type and K165R versions and the different concentration points did not fit well with a Langmuir function, revealing an active role of Lys167 in the labeling reaction with the CLD probe **5h** and pointing this amino acid as a likely site of *N*-acylation reaction.

The LD strategies that we are developing are intended to leave the receptor fully functional, since the affinity unit conjugated to the triazole and imidazole (**6**) can be washed-out and leave the binding site unoccupied. We checked this point by labeling HEK293 cells transiently expressing F-ST-D_1 with both the CLD probe **5h** and BG-Lumi4Tb. Thus, upon excitation of the terbium cryptate labeled on the SNAP-tag®, part of the energy would be transferred to the red FRET acceptor (i.e. Dy647) labeled on Lys167 to produce fluorescence emission at 665 nm (i.e. TR-FRET).[5] Next, the cells were treated with increasing concentrations of green D_1 antagonist (GDA, PerkinElmer) to interfere in the TR-FRET transfer between Lumi4Tb and Dy647 upon binding. Definitely, a decrease of the 665 nm emission was observed in the presence of GDA and such decrease was fitted to a binding function (Figure 3D). The K_d obtained (77 ± 53 nM) was compatible with the K_d of GDA obtained in Tag-Lite® assays (184 ± 9 nM). In contrast, the double-labeled cells, which were additionally treated with an excess of 50 µM of the antagonist SKF83566, showed no decrease of red fluorescence upon GDA application (Figure 3D). This demonstrates an almost complete binding of GDA to the binding pocket of D_1 receptors previously labeled with the CLD probe **5h.**

Visualization of labeling of D_1 receptor. We visualized the Dy647 labeling of D_1 receptor by the probe **5h** in live HEK293 cells using widefield microscopy. First, we performed the preparation of the CLD probe **5h** through a click reaction (*vide supra*) and we applied a solution 2.5 µM to living HEK293 cells transiently expressing Myc-D_1-Venus. After the incubation and washing steps, the fluorescence of Dy647 was observed on the plasma membrane (Figure 4A) in a similar manner than F-ST-D_1 with the SNAP-tag® labeled with BG-Dy647 (Figure S6) and with a high degree of

co-localization (R^2=0.8462) between Dy647 and the intracellular Venus fluorescence (Figure 4AB). Additionally, we did not observe specific labeling with the CLD probe **5h** in HEK293 cells non-expressing D_1 receptor and the click reaction mixture excluding the clickable fluorescent module **3b** did not affect the Venus fluorescence in Myc-D_1-Venus transfected cells, nor induced fluorescence in the red channel (Figure S6). These data show a specific labeling of D_1 receptor in living cells

As Myc-tag includes one lysine residue, we repeated the labeling procedure with HEK293 cells transiently expressing HA-D_1 and we next fixed the cells with PFA. Again, the fluorescence of Dy647 was observed on the plasma membrane and with a high degree of co-localization with the fluorescence originated from immunostaining (Figure 4C), showing that the dye covalently bound to D_1 lysine residues with the CLD technology is sufficient to visualize D_1 expression. We next aimed to visualize the organization and dynamics of D_1 receptor, such as monitoring the agonist-induced receptor internalization. To achieve this, **5h** was applied to label Myc-D_1-Venus receptor expressed in the plasma membrane of HEK293 cells and, after observing the fluorescence in the plasma membrane, we applied an excess of an orthosteric agonist (SKF81297 1 µM, Figure 1E). After 30 min, we could observe most of the fluorescence in endosomes in the cytosol of the cells, indicating an internalization of HA-D_1 receptors (Figure4E). Hence, the stability of CLD labeling enables tracking of native D_1 receptor trafficking. This experiment further confirms that the binding site of CLD-labeled native D_1 remains fully functional, as the agonist efficiently induced internalization.

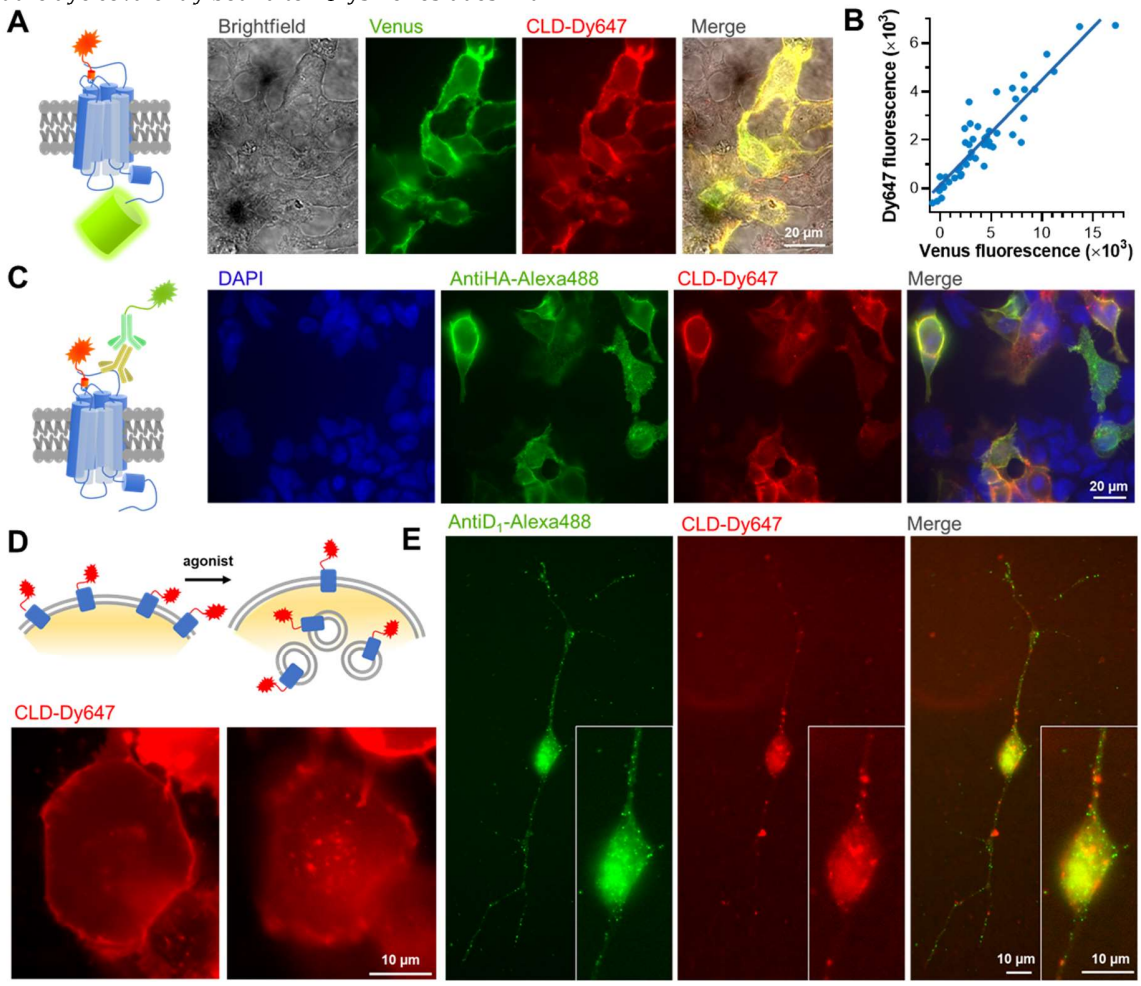


Figure 4: Visualization of D_1 receptors labeling with CLD labeling probe 5h. (A) Live HEK293 cells transiently expressing Myc-D_1-Venus were treated with CLD probe **5h** (2.5 µM), washed 4 times and imaged. Venus (upper right) and Dy647 fluorescence (lower left) originated from labeling with **5h** were located mostly at the cell membrane with a high-degree of co-localization (lower right, R^2=0.8462). (D) Live HEK293 cells transiently expressing HA-D_1 were labeled with **5h** using the same protocol, fixed with PFA, immunostained and imaged. Immunostaining fluorescence (upper right) and CLD fluorescence (lower left) were located mostly at the cell membrane with a high-degree of co-localization (lower right). (C) Living HEK293 cells transiently expressing HA-D_1 labeled with 5h display fluorescence at the cell membrane, but after 30 min of SKF81297 (1 µM) application, the fluorescence is translated in the in the cytosol, indicating receptor internalization. (E) SH-SY5Y cells differentiated to a neuron-like cells were labeled with 5h using the same protocol, fixed with PFA, immunostained and imaged. Immunostaining fluorescence (left) and Dy647 fluorescence (center) were located mostly at the cell membrane with a moderate degree of co-localization (right).

We then demonstrated that our approach allows the efficient labeling of endogenous D_1 receptors in their native environment. We used SH-SY5Y human neuroblastoma cells differentiated into viable neurons.[23] Therefore, we used our CLD molecular probe **5h** to label human D_1 receptor endogenously expressed by these cells using the labeling protocol described above. Upon microscope visualization, we observed CLD labeling co-localization with the fluorescence originated from immunostaining, confirming that **5h** label endogenous D_1 receptor, albeit with small compound aggregates (Figure 4E).

DISCUSSION

We developed a new modular ligand-directed technology (CLD) to chemically label native GPCRs with fluorescent dyes based on affinity reactive probes that can be prepared prior to the labeling reaction through click reactions (CLD) labeling. We have adapted an heterogenous copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) chemistry, employing Cu_2O as catalyst, which was developed for native tissue photopharmacology[24] and is compatible with in vivo testing.[25] In the CLD labeling, the use of copper-metal-chelating resins was required to eliminate traces of copper that may be toxic for cells and tissues.[22]

On the other hand, this modular chemical approach allows us to employ affinity units with relatively nucleophilic character (such aminergic GPCR ligands), since the final labeling probe is built just before the labeling. Similar approaches have been used in the field of tethered photopharamacology, enabling the covalent conjugation of glutamate derivatives directed to membrane receptors, even in vivo.[24,25] Additionally, the modular nature of our technology give us the opportunity to easily extend the application of the molecular unit **3a-c** to label other receptors or proteins of interest. This only requires the development of an affinity unit module **4** specific for the new protein of interest.

Other ligand directed approaches have been reported to label GPCR for bradykinin B_2[12], Adenosine A_{2A} recetors[26] and opioid receptors.[27] However, we managed to label an endogenous archetypical aminergic GPCR (i.e. D_1 receptor) with CLD technology using am affinity unit with relatively high nucleophilic properties. We have obtained a moderate labelling yield if we compare it with well-established technologies on engineered-modified receptors, such as SNAP-tag® labeling but this work represents a proof of concept of how to overcome the pitfall of labeling endogenous aminergic GPCRs with no need of genetic engineering. To that end, we used an established electrophilic group in ligand directed labeling (i.e. acyl imidazole), which is reported to primary target lysine residues.[12,13] Indeed, D_1 receptor only incorporates two extracellular solvent-exposed lysine residues in the loop ECL2 (K165, K167). For that reason, we performed the labeling experiments with three mutant variants, which revealed that Lys167 has a major role on the labeling reactivity of **5h**, but other residues cannot be discarded to have a minor role as mutants K167R and K165R-K167R still have residual specific labeling by **5h**. Certainly, LDAI labeling approaches are described to also target serines, threonines and tyrosines [10,20] and those residues are also highly present in the extracellular surface of D_1 receptor (supporting information). However, the low extent of specific labeling and the non-saturating labeling curves indicate a minor role of those residues compared to Lys167.

Most probably, the labeling yield and residue specificity could be improved in the future with new electrophilic species that can improve both the specificity and reaction rate. Under the current experimental conditions, a non-specific labeling comes from CLD aggregates dispersed in the culture, which can hardly be confused with receptor labeling, but which requires subtracting the non-specific signal from the total signal in the fluorimeter, or imaging a region without aggregates under the microscope to highlight the specific signal. Additionally, we do not discard in future approaches to employ copper-free click chemistry such as strain-promoted azide-alkyne cycloaddition (SPAAC) or inverse electron-demand Diels–Alder (IEDDA).[28,29] However, these approaches require the use of bulkier functional groups that can compromise the water-solubility of the probe and might cause undesirable interactions with the protein of interest, thus reducing the binding affinity to the receptor. The presented CLD approach allowed us to visualize HA-D_1 receptor expressed in the plasma membrane of HEK293 or endogenous D_1 in neuron-like differentiated cells. This technology can be used as chemical biology tool compounds to develop a great variety of applications in the field of native GPCR research, in part, because the binding site remains unaltered after labeling and can accommodate other orthosteric ligands, as we have demonstrated using TR-FRET experiments. Additionally, we were able to monitor the internalization of D_1 upon agonist application and demonstrating once more that the binding site of the CLD-labelled receptors remains unaltered and the receptor is fully functional after labeling, as agonists bound and induced internalization. This was one an example of possible applications but other multiple choices can be explored such as monitoring receptor dynamics or partnering between proteins. These may include the study of protein-protein interactions in endogenous tissues using FRET between fluorophores covalently conjugated to different native proteins and a possible translation of the Tag-Lite® technology to native receptors and eventually in tissue samples, organs and in vivo.

CONCLUSION

There are increasing indications that native GPCRs likely have properties different from the recombinant ones expressed in heterologous cells. Covalent labeling of native receptors, using approaches such as the one proposed here, will certainly open a multitude of ways to study their properties in their natural environment. These include their internalization and recycling, their binding properties, their association with partners, their precise localization, and many other things. Our study demonstrates that covalent labeling of native receptors is possible without affecting much their properties as the labeled receptors are still able to respond to their cognate ligands. The principle of our approach may also be of interest to any scientist interested in cell surface proteins including other types of receptors,

transporters or channel as long as high affinity ligands exist to specifically target them.

ASSOCIATED CONTENT

Supporting Information. Supporting figures, synthetic strategies, materials and methods and compound characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

* MCS, Institute for Advanced Chemistry of Catalonia
Spanish National Research Council (CSIC)
Jordi Girona 18-16, 08034 Barcelona, Spain
amadeu.llebaria@iqac.csic.es
xavier.gomez@iqac.csic.es
* Institut de Génomique Fonctionnelle
Université de Montpellier, CNRS UMR 5203 and Inserm U1191
141, rue de la Cardonille, 34094 Montpellier cedex 5
jean-philippe.pin@igf.cnrs.fr

Present Addresses

†JMZ is currently affiliated with QPS Netherlands BV, Petrus Campersingel 123, 9713 AG Groningen, The Netherlands.

Author Contributions

The manuscript was written through contributions of all authors and all have given approval to the final version of the manuscript.

Funding Sources

The project on which these results are based has received funding from Agence Nationale de la Recherche (ANR-17-CE11-0046), Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación and ERDF - A way of making Europe (projects CTQ2017-89222-R, PCI2018-093047 and PID2020-120499RB-I00) by the Catalan government (2017 SGR 1604) the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement No. 801342 (TecniospringINDUSTRY, TECSPR19-1-0062) and the Government of Catalonia's Agency for Business Competitiveness (ACCIÓ).

ACKNOWLEDGMENT

We thank the ARPEGE Pharmacology-Screening-Interactome platform facility (UMS Bio-campus, Montpellier, France) for the use of the plate readers, the MRI Montpellier Ressources Imagerie (UMS Bio-campus, Montpellier, France) for images acquisition, Lourdes Muñoz and Carme Serra from SIMChem (IQAC CSIC) for the use of analytical support and helpful discussions, L. Clusan for the generation of the Flag-Snaptag-D_1 receptor, S. Soldevila, E. Bourrier, K. Belkacemi, J. Haubrich, S. Panarello, F. Malhaire, J. Kniazeff, C. Lavergne and X. Rovira for helpful discussions and technical support.

REFERENCES

(1) Sriram, K.; Insel, P. A. GPCRs as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **2018**, *93* (4), 251–258. https://doi.org/10.1124/MOL.117.111062.

(2) Stoddart, L. A.; Johnstone, E. K. M.; Wheal, A. J.; Goulding, J.; Robers, M. B.; MacHleidt, T.; Wood, K. V.; Hill, S. J.; Pfleger, K. D. G. Application of BRET to Monitor Ligand Binding to GPCRs. *Nat. Methods 2015 127* **2015**, *12* (7), 661–663. https://doi.org/10.1038/nmeth.3398.

(3) Kauk, M.; Hoffmann, C. Intramolecular and Intermolecular FRET Sensors for GPCRs – Monitoring Conformational Changes and Beyond. *Trends Pharmacol. Sci.* **2018**, *39* (2), 123–135. https://doi.org/10.1016/J.TIPS.2017.10.011.

(4) Scholler, P.; Zwier, J. M.; Trinquet, E.; Rondard, P.; Pin, J. P.; Prézeau, L.; Kniazeff, J. Time-Resolved Förster Resonance Energy Transfer-Based Technologies to Investigate G Protein-Coupled Receptor Machinery: High-Throughput Screening Assays and Future Development. *Prog. Mol. Biol. Transl. Sci.* **2013**, *113*, 275–312. https://doi.org/10.1016/B978-0-12-386932-6.00007-7.

(5) M. Zwier, J.; Bazin, H.; Lamarque, L.; Mathis, G. Luminescent Lanthanide Cryptates: From the Bench to the Bedside. *Inorg. Chem.* **2014**, *53* (4), 1854–1866. https://doi.org/10.1021/ic402234k.

(6) Lang, K.; Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **2014**, *114* (9), 4764–4806. https://doi.org/10.1021/cr400355w.

(7) Zwier, J. M.; Roux, T.; Cottet, M.; Durroux, T.; Douzon, S.; Bdioui, S.; Gregor, N.; Bourrier, E.; Oueslati, N.; Nicolas, L.; Tinel, N.; Boisseau, C.; Yverneau, P.; Charrier-Savournin, F.; Fink, M.; Trinquet, E. A Fluorescent Ligand-Binding Alternative Using Tag-Lite® Technology. *J. Biomol. Screen.* **2010**, *15* (10), 1248–1259. https://doi.org/10.1177/1087057110384611.

(8) Sykes, D. A.; Charlton, S. J. Single Step Determination of Unlabeled Compound Kinetics Using a Competition Association Binding Method Employing Time-Resolved FRET. *Methods Mol. Biol.* **2018**, *1824*, 177–194. https://doi.org/10.1007/978-1-4939-8630-9_10.

(9) Takaoka, Y.; Ojida, A.; Hamachi, I. Protein Organic Chemistry and Applications for Labeling and Engineering in Live-Cell Systems. *Angew. Chemie - Int. Ed.* **2013**, *52* (15), 4088–4106. https://doi.org/10.1002/anie.201207089.

(10) Tamura, T.; Hamachi, I. Chemistry for Covalent Modification of Endogenous/Native Proteins: From Test Tubes to Complex Biological Systems. *J. Am. Chem. Soc.* **2018**, *141* (7), 2782–2799. https://doi.org/10.1021/jacs.8b11747.

(11) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. Ligand-Directed Tosyl Chemistry for Protein Labeling in Vivo. *Nat. Chem. Biol. 2009 55* **2009**, *5* (5), 341–343. https://doi.org/10.1038/nchembio.157.

(12) Miki, T.; Fujishima, S. H.; Komatsu, K.; Kuwata, K.; Kiyonaka, S.; Hamachi, I. LDAI-Based Chemical Labeling of Intact Membrane Proteins and Its Pulse-Chase Analysis under Live Cell Conditions. *Chem. Biol.* **2014**, *21* (8), 1013–1022. https://doi.org/10.1016/J.CHEMBIOL.2014.07.013.

(13) Fujishima, S. H.; Yasui, R.; Miki, T.; Ojida, A.; Hamachi, I. Ligand-Directed Acyl Imidazole Chemistry for Labeling of Membrane-Bound Proteins on Live Cells. *J. Am. Chem. Soc.* **2012**, *134* (9), 3961–3964. https://doi.org/10.1021/JA2108855/SUPPL_FILE/JA210885 5_SI_001.PDF.

(14) Takaoka, Y.; Nishikawa, Y.; Hashimoto, Y.; Sasaki, K.; Hamachi, I. Ligand-Directed Dibromophenyl Benzoate Chemistry for Rapid and Selective Acylation of Intracellular Natural Proteins. *Chem. Sci.* **2015**, *6* (5), 3217–3224. https://doi.org/10.1039/C5SC00190K.

(15) Tamura, T.; Ueda, T.; Goto, T.; Tsukidate, T.; Shapira, Y.; Nishikawa, Y.; Fujisawa, A.; Hamachi, I. Rapid Labelling and Covalent Inhibition of Intracellular Native Proteins Using Ligand-Directed N-Acyl-N-Alkyl Sulfonamide. *Nat. Commun.* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-04343-0.

(16) Wang, H.; Koshi, Y.; Minato, D.; Nonaka, H.; Kiyonaka, S.; Mori, Y.; Tsukiji, S.; Hamachi, I. Chemical Cell-Surface Receptor Engineering Using Affinity-Guided, Multivalent

Organocatalysts. *J. Am. Chem. Soc.* **2011**, *133* (31), 12220–12228. https://doi.org/10.1021/JA204422R/SUPPL_FILE/JA204422R_SI_001.PDF.

(17) Tamura, T.; Song, Z.; Amaike, K.; Lee, S.; Yin, S.; Kiyonaka, S.; Hamachi, I. Affinity-Guided Oxime Chemistry for Selective Protein Acylation in Live Tissue Systems. *J. Am. Chem. Soc.* **2017**, *139* (40), 14181–14191. https://doi.org/10.1021/jacs.7b07339.

(18) Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E. GPCRdb in 2018: Adding GPCR Structure Models and Ligands. *Nucleic Acids Res.* **2018**, *46* (D1), D440–D446. https://doi.org/10.1093/NAR/GKX1109.

(19) Kooistra, A. J.; Mordalski, S.; Pándy-Szekeres, G.; Esguerra, M.; Mamyrbekov, A.; Munk, C.; Keserű, G. M.; Gloriam, D. E. GPCRdb irencesn 2021: Integrating GPCR Sequence, Structure and Function. *Nucleic Acids Res.* **2021**, *49* (D1), D335–D343. https://doi.org/10.1093/NAR/GKAA1080.

(20) Matsuo, K.; Kioi, Y.; Yasui, R.; Takaoka, Y.; Miki, T.; Fujishima, S. H.; Hamachi, I. One-Step Construction of Caged Carbonic Anhydrase I Using a Ligand-Directed Acyl Imidazole-Based Protein Labeling Method. *Chem. Sci.* **2013**, *4* (6), 2573–2580. https://doi.org/10.1039/C3SC50560J.

(21) Bakthavachalam, V.; Baindur, N.; Madras, B. K.; Neumeyer, J. L. Fluorescent Probes for Dopamine Receptors: Synthesis and Characterization of Fluorescein and 7-Nitrobenz-2-Oxa-1,3-Diazol-4-Yl Conjugates of D-1 and D-2 Receptor Ligands. *J. Med. Chem.* **1991**, *34* (11), 3235–3241. https://doi.org/10.1021/jm00115a012.

(22) Kennedy, D. C.; McKay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. Cellular Consequences of Copper Complexes Used to Catalyze Bioorthogonal Click Reactions. *J. Am. Chem. Soc.* **2011**, *133* (44), 17993–18001. https://doi.org/10.1021/JA2083027/SUPPL_FILE/JA2083027_SI_001.PDF.

(23) Shipley, M. M.; Mangold, C. A.; Szpara, M. L. Differentiation of the SH-SY5Y Human Neuroblastoma Cell Line. *JoVE (Journal Vis. Exp.* **2016**, *2016* (108), e53193. https://doi.org/10.3791/53193.

(24) Izquierdo-Serra, M.; Bautista-Barrufet, A.; Trapero, A.; Garrido-Charles, A.; Diaz-Tahoces, A.; Camarero, N.; Pittolo, S.; Valbuena, S.; Perez-Jimenez, A.; Gay, M.; Garcia-Moll, A.; Rodriguez-Escrich, C.; Lerma, J.; De La Villa, P.; Fernandez, E.; Pericas, M. A.; Llebaria, A.; Gorostiza, P. Optical Control of Endogenous Receptors and Cellular Excitability Using Targeted Covalent Photoswitches. *Nat. Commun.* **2016**, *7* (1), 1–12. https://doi.org/10.1038/ncomms12221.

(25) Garrido-charles, A.; Huet, A.; Matera, C.; Thirumalai, A.; Hernando, J.; Llebaria, A.; Moser, T.; Gorostiza, P. Fast Photoswitchable Molecular Prosthetics Control Neuronal Activity in the Cochlea Aida. *J. Am. Chem. Soc.* **2022**. https://doi.org/10.1021/jacs.1c12314.

(26) Stoddart, L. A.; Kindon, N. D.; Otun, O.; Harwood, C. R.; Patera, F.; Veprintsev, D. B.; Woolard, J.; Briddon, S. J.; Franks, H. A.; Hill, S. J.; Kellam, B. Ligand-Directed Covalent Labelling of a GPCR with a Fluorescent Tag in Live Cells. *Commun. Biol.* **2020**, *3* (1), 1–9. https://doi.org/10.1038/s42003-020-01451-w.

(27) Arttamangkul, S.; Plazek, A.; Platt, E. J.; Jin, H.; Murray, T. F.; Birdsong, W. T.; Rice, K. C.; Farrens, D. L.; Williams, J. T. Visualizing Endogenous Opioid Receptors in Living Neurons Using Ligand-Directed Chemistry. *Elife* **2019**, *8*, e49319. https://doi.org/10.7554/ELIFE.49319.

(28) Dommerholt, J.; Van Rooijen, O.; Borrmann, A.; Guerra, C. F.; Bickelhaupt, F. M.; Van Delft, F. L. Highly Accelerated Inverse Electron-Demand Cycloaddition of Electron-Deficient Azides with Aliphatic Cyclooctynes. *Nat. Commun.* **2014**, *5* (1), 1–7. https://doi.org/10.1038/ncomms6378.

(29) Oliveira, B. L.; Guo, Z.; Bernardes, G. J. L. Inverse Electron Demand Diels-Alder Reactions in Chemical Biology. *Chemical Society Reviews*. Royal Society of Chemistry August 21, 2017, pp 4895–4950. https://doi.org/10.1039/c7cs00184c.

Graphical Table of Contents