Published in final edited form as:

Angew Chem Int Ed Engl. 2015 December 7; 54(50): 15246–15249. doi:10.1002/anie.201508506.

In-membrane chemical modification (IMCM) for site-specific chromophore labeling of GPCRs

Lukas Sušac^[a], Dr. Casey O'Connor^{[a],[b]}, Prof. Dr. Raymond C. Stevens^[b], and Prof. Dr. Kurt Wüthrich^{[a],[c]}

Kurt Wüthrich: wuthrich@scripps.edu

^[a]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)

^[b]Departments of Biological Sciences and Chemistry, Bridge Institute, University of Southern California, 3430 South Vermont Avenue, Los Angeles, CA 90089 (USA)

^[c]Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)

Abstract

We present “in-membrane chemical modification” (IMCM) for obtaining selective chromophore labeling of intracellular surface cysteines in G protein-coupled receptors (GPCRs) with minimal mutagenesis. Practical use of IMCM is illustrated with site-specific introduction of chromophores for NMR and fluorescence spectroscopy in the human κ -opioid receptor (KOR) and the human A_{2A} adenosine receptor (A_{2A}AR). IMCM is applicable for a wide range of *in-vitro* studies of GPCRs, including single molecule spectroscopy, and is a promising platform for in-cell spectroscopy experiments.

Graphical Abstract

Keywords

G protein-coupled receptors; site-specific labeling; NMR spectroscopy; fluorescence spectroscopy; membrane proteins

In-membrane chemical modification (IMCM) is a novel method, which enables site-specific introduction of chromophores through chemical modification of natural amino acids with minimal use of mutagenesis. Since low expression yields tend to be a limiting factor for biophysical studies, which can be exacerbated by the use of mutations,^[1] the ability to attach chemical probes in a site-specific manner to wild type proteins is a significant advantage. Here, IMCM is tailored for chemical modification of cysteines in G protein-coupled receptors (GPCR), which are represented by more than 800 different proteins in the human body^[2] and are targets for more than 30% of all prescription drugs in human medicine.^[3] GPCRs function by transmitting signals across the plasma membrane to the cell interior. Spectroscopic studies based on observation of chromophores located on the intracellular receptor surface are used to complement information from crystal structures, by monitoring dynamic processes that underlie GPCR function.^[1,4] Cysteine residues at the intracellular surface of GPCRs are a preferred target for chemical modification with chromophore-carrying reagents (see below). Conventionally, such chemical reactions have been performed with GPCRs in detergent micelles, with selectivity obtained primarily through extensive mutagenesis, so that screening of numerous mutant GPCRs for high expression and functionality has typically been required.^[1,4] In contrast, the new “in-membrane chemical modification” (IMCM) approach makes use of natural protection of most cysteines by the membrane environment, and thus enables selective cysteine labeling on the intracellular receptor surface with minimal or no mutagenesis. IMCM-labeling is an attractive technique for a wide range of GPCRs (Figure 1), as illustrated here with ¹⁹F-NMR and fluorescence spectroscopy of the human κ -opioid receptor (KOR) and the human A_{2A} adenosine receptor (A_{2A}AR).

Analysis of 32 currently available unique GPCR crystal structures (Table S1) yielded a comprehensive overview of the spatial distribution of cysteine residues as potential attachment sites for spectroscopic probes (Figure 1), which showed that most of these GPCRs are amenable to selective labeling by IMCM, since intracellular cysteines represent 15% of all GPCR cysteines and are 15-fold more abundant than extracellular cysteines (Figure 1b). 27 of the 32 GPCRs are devoid of extracellular cysteines, so that with efficient protection of the transmembrane (TM) cysteines (Figure 1b), selective labeling of cysteine residues on the intracellular surface can be achieved without mutagenesis. To this end, IMCM uses crude membrane preparations, with the GPCRs still embedded in a biological membrane, so that protection of the TM cysteines is afforded by the natural membrane environment.

Based on the absence of fluorine atoms in most natural biological materials, the use of ^{19}F -NMR labels has the advantage that there is no spectral background when the labeling is performed in a site-specific manner.^[5] Methods for site-specific introduction of ^{19}F -labeled non-natural amino acids are available,^[6] but these are rarely practical in systems with low expression yields. Therefore, post-translational modification of natural amino acids, e.g., cysteine or lysine, has been widely used^[1,4] and is also the strategy followed by IMCM. When applying IMCM for ^{19}F -NMR studies, we used a small ^{19}F -NMR probe, 2,2,2-trifluoroethanethiol (TET), for chemical modification of accessible cysteines, and conventional TET-labeling in detergent micelles served as a reference.

The human κ -opioid receptor contains five TM cysteines and three cysteines (Cys161, Cys340, Cys345) exposed on the intracellular surface.^[7] After TET-labeling in micelles, 1D ^{19}F -NMR spectra of KOR contained an intense, broad resonance at ~9.5 ppm (Figure 2a). In the variant protein KOR[C161M, C340A, C345S] (Figure 2b), the ^{19}F -NMR spectrum contained a weak line at the same chemical shift, showing that the signal in Figure 2a was due mostly to TET-labeling of the three intracellular cysteines and that TET-labeled TM cysteines accounted for a low-intensity background signal. When KOR[C340A, C345S] was labeled in micelles, a signal at ~9.5 ppm was observed (Figure 2c), which corresponds to the superposition of the signal from the TET-label on Cys161 and the background intensity in Figure 2b. Repeating the same experiments with TET-labeling by IMCM (Figure 2d–f), we observed again a ^{19}F -NMR signal at ~9.5 ppm for the wild type KOR (Figure 2d). In KOR[C161M, C340A, C345S], the background signal from TM cysteines was completely removed (Figure 2e), so that the signal in Figure 2f must originate entirely from the TET-labeled Cys161 in KOR[C340A, C345S]. Overall, without introducing any mutations in the TM region, the indigenous intracellular Cys161 was thus shown to be a reliable reporter site in the variant protein KOR[C340A, C345S], provided that TET-labeling was performed with IMCM.

The human $\text{A}_{2\text{A}}$ AR has six TM cysteines and no intra- or extracellular cysteines.^[8] 1D ^{19}F -NMR data for $\text{A}_{2\text{A}}$ AR labeled with TET in micelles exhibited a broad ^{19}F -NMR signal due to labeling of indigenous TM cysteines (Figure 3a). In contrast, IMCM of $\text{A}_{2\text{A}}$ AR with TET resulted in an empty NMR spectrum, demonstrating that labeling of indigenous TM cysteines was completely suppressed (Figure 3b). We then engineered the reporter cysteine A289C at the tip of TM helix VII to monitor the intracellular surface of $\text{A}_{2\text{A}}$ AR. In the

variant protein A_{2A}AR[A289C], TET-labeling in micelles yielded a ¹⁹F-NMR spectrum with two signals of closely similar intensity at ~9.5 and ~11.5 ppm (Figure 3c). TET-labeling by IMCM resulted in a related spectrum where, however, the signal at ~9.5 ppm had greatly reduced intensity (Figure 3d). These two peaks in the spectrum of Figure 3d must both originate from the TET-label attached to the engineered A289C and therefore manifest a conformational polymorphism. This information is complementary to previous reports that TM helix VII undergoes activation-related conformational changes.^[8b,8c]

Validation of the IMCM approach for ¹⁹F-NMR studies was obtained from two lines of experiments. Firstly, IMCM was applied with the previously extensively studied β₂-adrenergic receptor (β₂AR), which contains four TM cysteines and three intracellular cysteines. Previous work^[1b] had established, by extensive mutational experiments, that the TM cysteines were all protected from TET labeling even when the chemical reaction was performed with micelle-reconstituted β₂AR. The resulting ¹⁹F-NMR spectrum contained three lines which were, by additional mutational studies, individually assigned to the three intracellular cysteines in positions 265, 327 and 341 (Figure S1a). Using IMCM for TET-labeling, the same spectral features were obtained (Figure S1b), showing that the IMCM preparation coincides with the one from the “proven” conventional approach. Clearly, the previously reported results on β₂AR could have been obtained more efficiently with the use of IMCM. Secondly, intact protein electrospray ionization mass spectrometry (ESI-MS) of A_{2A}AR labeled in micelles with TET showed a heterogeneous receptor preparation containing unlabeled A_{2A}AR and four species resulting from partial labeling of four different TM cysteines (Figure S2a). In contrast, IMCM-labeling of A_{2A}AR produced a homogenous sample of unlabeled receptor (Figure S2b), due to complete suppression of cysteine-labeling in the TM helices. ESI-MS data of engineered A_{2A}AR[A289C] labeled by IMCM showed a homogenous preparation of singly TET-labeled receptor (Figure S2d), confirming selective labeling of the engineered intracellular cysteine site. This selectivity was lost when TET-labeling of A_{2A}AR[A289C] was carried out in micelles, where ESI-MS showed a mixture of six different species (Figure S2c).

To test applications of IMCM for site-specific fluorescent labeling of GPCRs, the variant receptor A_{2A}AR[A289C] was IMCM-labeled with fluorescein-5-maleimide. Indeed, a fluorescent protein sample was obtained (Figure 3e) and ESI-MS analysis showed selective attachment of a single fluorescein marker to A_{2A}AR[A289C] (Figure S3). The control IMCM experiment with A_{2A}AR did not result in a fluorescent receptor preparation (Figure 3f), demonstrating the absence of labeling of TM cysteines, as was confirmed by ESI-MS, which showed only unlabeled A_{2A}AR (Figure S3). Maleimide-based probes are the most diverse class of commercially available cysteine-reactive compounds, which has thus been shown to be suitable for IMCM-labeling. This underscores the broad applicability of IMCM for site-specific chromophore labeling for spectroscopic studies, and suggests IMCM as a general strategy for selective covalent labeling of GPCRs with thiol- and maleimide-based reagents.

In summary, we successfully applied IMCM for selective covalent attachment of chromophores for biophysical studies to indigenous or engineered intracellular cysteines in human A_{2A}AR and KOR. We demonstrated that IMCM can facilitate spectroscopic studies

of GPCRs with ^{19}F -NMR and fluorescence probes by reducing or eliminating the need for mutagenesis to achieve site-specific labeling on the intracellular surface. Inspection of the cysteine distribution in GPCR crystal structures (Figure 1) showed that a wide range of GPCRs can be selectively labeled with the IMCM approach. This new experiment may result in further increased importance of cysteine residues as sites for chemical modification in biophysical studies of GPCRs, since other residues with selective side chain chemical reactivity do not show comparably favorable distributions in the three-dimensional GPCR structures. Here, this is illustrated with the distribution of lysine residues (Figure S4), which have also been a frequent target for chemical modification of proteins for spectroscopic studies.^[9] The favorable situation for cysteines can be rationalized by the impact of the oxidative extracellular environment, which results in formation of intramolecular disulfide bonds in GPCRs (Figure 1). Finally, in addition to the impressive potential for *in vitro* studies of GPCRs, it would appear that along the lines of “cell surface engineering” described by Saxon and Bertozzi,^[10] IMCM could be a starting point for in-cell labeling experiments designed to monitor GPCRs in their natural environment by spectroscopic techniques.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by PSI:BiologY grant U54 GM094618. K.W. is the Cecil H. and Ida M. Green Professor of Structural Biology at The Scripps Research Institute. L.S. was supported by a Boehringer Ingelheim Fonds PhD fellowship and by the German Academic Exchange Service (DAAD). We thank M. Tran from The Scripps Center for Mass Spectrometry for help with ESI-MS analysis; J. Velasquez for help with molecular biology; K. Villers, C. Hanson and M. Chu for help with baculovirus expression.

References

1. a) Gether U, Lin S, Ghanouni P, Ballesteros JA, Weinstein H, Kobilka BK. EMBO J. 1997; 16:6737–6747. [PubMed: 9362488] b) Liu JJ, Horst R, Katritch V, Stevens RC, Wüthrich K. Science. 2012; 335:1106–1110. [PubMed: 22267580]
2. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. Mol Pharmacol. 2003; 63:1256–1272. [PubMed: 12761335]
3. Hopkins AL, Groom CR. Nat Rev Drug Discovery. 2002; 1:727–730. [PubMed: 12209152]
4. a) Loewen MC, Klein-Seetharaman J, Getmanova EV, Reeves PJ, Schwalbe H, Khorana HG. Proc Natl Acad Sci U S A. 2001; 98:4888–4892. [PubMed: 11320239] b) Chung KY, Kim TH, Manglik A, Alvares R, Kobilka BK, Prosser RS. J Biol Chem. 2012; 287:36305–36311. [PubMed: 22893704] c) Fay JF, Farrens DL. J Biol Chem. 2012; 287:33873–33882. [PubMed: 22846992] d) Rahmeh R, Damian M, Cottet M, Orcel H, Mendre C, Durroux T, Sharma KS, Durand G, Pucci B, Trinquet E, Zwier JM, Deupi X, Bron P, Baneres JL, Mouillac B, Granier S. Proc Natl Acad Sci U S A. 2012; 109:6733–6738. [PubMed: 22493271] e) Kim TH, Chung KY, Manglik A, Hansen AL, Dror RO, Mildorf TJ, Shaw DE, Kobilka BK, Prosser RS. J Am Chem Soc. 2013; 135:9465–9474. [PubMed: 23721409] f) Horst R, Liu JJ, Stevens RC, Wüthrich K. Angew Chem Int Ed. 2013; 41:10762–10765. g) Mathiasen S, Christensen SM, Fung JJ, Rasmussen SG, Fay JF, Jorgensen SK, Veshaguri S, Farrens DL, Kiskowski M, Kobilka B, Stamou D. Nat Methods. 2014; 11:931–934. [PubMed: 25086504] h) Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, Lerch MT, Kobilka TS, Thian FS, Hubbell WL, Prosser RS, Kobilka BK. Cell. 2015; 161:1101–1111. [PubMed: 25981665]

5. Didenko T, Liu JJ, Horst R, Stevens RC, Wüthrich K. *Curr Opin Struct Biol.* 2013; 23:740–747. [PubMed: 23932201]
6. Jones DH, Cellitti SE, Hao X, Zhang Q, Jahnz M, Summerer D, Schultz PG, Uno T, Geierstanger BH. *J Biomol NMR.* 2010; 46:89–100. [PubMed: 19669620]
7. Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson A, Huang X-P, Carroll I, Mascarella W, Westkaemper R, Mosier P, Roth B, Cherezov V, Stevens R. *Nature.* 2012; 485:327–332. [PubMed: 22437504]
8. a) Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, IJzerman AP, Stevens RC. *Science.* 2008; 322:1211–1217. [PubMed: 18832607] b) Lebon G, Warne T, Edwards P, Bennett K, Langmead C, Leslie A, Tate C. *Nature.* 2011; 474:521–525. [PubMed: 21593763] c) Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC. *Science.* 2011; 332:322–327. [PubMed: 21393508] d) Liu W, Chun E, Thompson AA, Chubukov P, Xu F, Katritch V, Han GW, Roth CB, Heitman LH, IJzerman AP, Cherezov V, Stevens RC. *Science.* 2012; 337:232–236. [PubMed: 22798613]
9. Brinkley M. *Bioconjug Chem.* 1992; 3:2–13. [PubMed: 1616945]
10. Saxon E, Bertozzi CR. *Science.* 2000; 287:2007–2010. [PubMed: 10720325]

Figure 1.

Cysteine locations in GPCR crystal structures. a) Superposition of 32 unique GPCR crystal structures listed in Table S1. Cysteine residues are shown as spheres and classified into four categories based on their locations and oxidation states: extracellular cysteines (yellow), extracellular cysteines (green), transmembrane cysteines (orange), and intracellular cysteines (magenta). b) Distribution of the cysteine residues among the four groups introduced in (a).

Figure 2.

Labeling of human KOR for ${}^{19}\text{F}$ -NMR. a–c) 1D ${}^{19}\text{F}$ -NMR spectra after TET-labeling in micelles of KOR (a), KOR[C161M, C340A, C345S] (b), and KOR[C340A, C345S] (c). d–f) 1D ${}^{19}\text{F}$ -NMR spectra after TET-labeling by IMCM of KOR (d), KOR[C161M, C340A, C345S] (e), and KOR[C340A, C345S] (f). Note that the protein concentration used in the experiments (b) and (c) was about two-fold lower than in the measurements of (a) and (d–f). The sharp peak near 7.0 ppm corresponds to a small concentration of free TET.

Figure 3.

Site-specific IMCM labeling of human $\text{A}_{2\text{A}}\text{AR}$ for ^{19}F -NMR and fluorescence spectroscopy. a,b) 1D ^{19}F -NMR spectra of $\text{A}_{2\text{A}}\text{AR}$ labeled with TET either in micelles (a) or using IMCM (b). c,d) 1D ^{19}F -NMR spectra of $\text{A}_{2\text{A}}\text{AR}[\text{A}289\text{C}]$ labeled with TET either in micelles (c) or using IMCM (d). e) $\text{A}_{2\text{A}}\text{AR}[\text{A}289\text{C}]$ IMCM-labeled with fluorescein-5-maleimide is shown under white light and UV light. f) Same presentation as in (e) for wild type $\text{A}_{2\text{A}}\text{AR}$ that was subjected to identical labeling conditions.