Protein Labeling

International Edition: DOI: 10.1002/anie.201508506
German Edition: DOI: 10.1002/ange.201508506

# In-Membrane Chemical Modification (IMCM) for Site-Specific Chromophore Labeling of GPCRs

*Lukas Sušac, Casey O'Connor, Raymond C. Stevens, and Kurt Wüthrich**

***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. This method takes advantage of the natural protection of most cysteines by the membrane environment. Practical use of IMCM is illustrated with the 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 to a wide range of in vitro studies of GPCRs, including single-molecule spectroscopy, and is a promising platform for in-cell spectroscopy experiments.*

In-membrane chemical modification (IMCM) is a novel method that enables the 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. IMCM is tailored for the chemical modification of cysteine residues in G-protein-coupled receptors (GPCRs), 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 the 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 the 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 by $^{19}$F-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 in the Supporting Information) yielded a comprehensive overview of the spatial distribution of cysteine residues as potential attachment sites for spectroscopic probes (Figure 1). The results show 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 is applied to 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 the 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, for example, 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 the small $^{19}$F-NMR probe 2,2,2-trifluoroethanethiol (TET) for the 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, the 1D $^{19}$F-NMR spectra for KOR contained an intense, broad resonance at $\delta \approx 9.5$ ppm (Figure 2a). With the variant

[*] L. Sušac, Dr. C. O'Connor, Prof. Dr. K. Wüthrich
Department of Integrative Structural and Computational Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
E-mail: wuthrich@scripps.edu
Dr. C. O'Connor, Prof. Dr. R. C. Stevens
Departments of Biological Sciences and Chemistry, Bridge Institute
University of Southern California
3430 South Vermont Avenue, Los Angeles, CA 90089 (USA)
Prof. Dr. K. Wüthrich
Skaggs Institute of Chemical Biology, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
and
Institute of Molecular Biology and Biophysics, ETH Zurich
Otto-Stern-Weg 5, 8093 Zurich (Switzerland)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201508506.

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

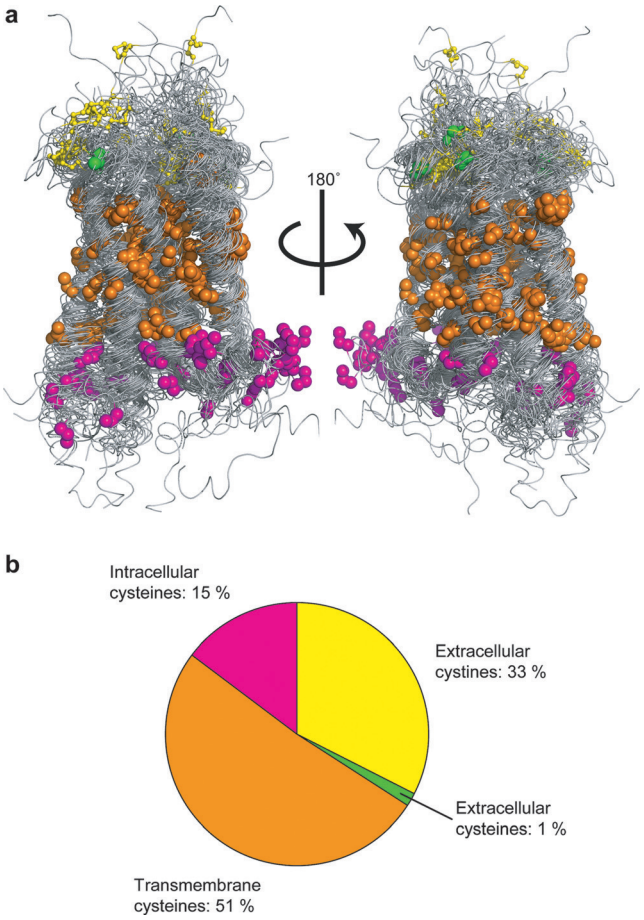


***Figure 1.*** Cysteine locations in GPCR crystal structures. a) Superposition of 32 unique GPCR crystal structures listed in Table S1. Cysteine residues are shown as spheres and classified into four categories based on their locations and oxidation states: extracellular cystines (yellow), extracellular cysteines (green), transmembrane cysteines (orange), and intracellular cysteines (magenta). b) Distribution of the cysteine residues among the four groups introduced in (a).

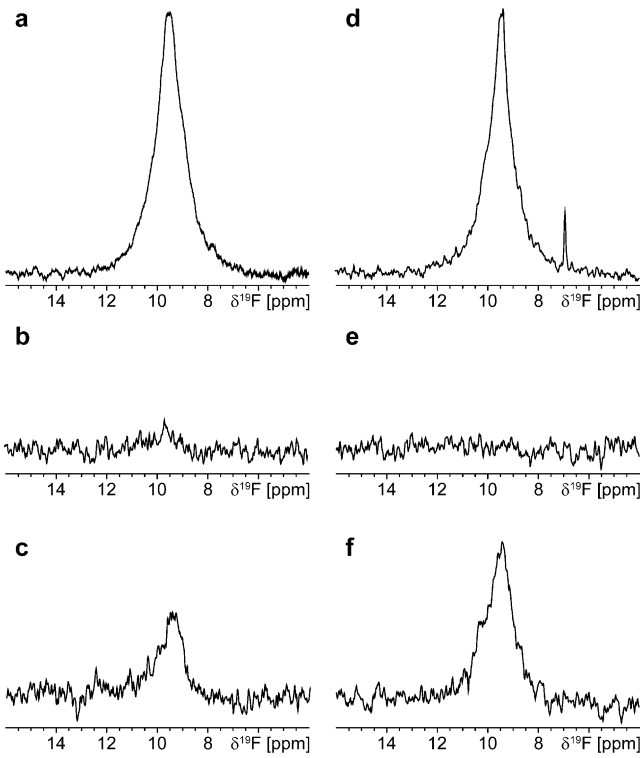


***Figure 2.*** Labeling of human KOR for $^{19}$F-NMR. a–c) 1D $^{19}$F-NMR spectra after TET labeling in micelles of KOR (a), KOR[C161M, C340A, C345S] (b), and KOR[C340A, C345S] (c). d–f) 1D $^{19}$F-NMR spectra after TET labeling by IMCM of KOR (d), KOR[C161M, C340A, C345S] (e), and KOR[C340A, C345S] (f). Note that the protein concentration used in experiments (b) and (c) was about two-fold lower than in (a) and (d–f). The sharp peak at $\delta \approx 7.0$ ppm corresponds to a small concentration of free TET.

protein KOR[C161M, C340A, C345S] (Figure 2b), the $^{19}$F-NMR spectrum contained a weak line at the same chemical shift, thus 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 $\delta \approx 9.5$ 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 again observed a $^{19}$F-NMR signal at $\delta \approx 9.5$ ppm for the wild-type KOR (Figure 2d). With KOR[C161M, C340A, C345S], the background signal from TM cysteines was completely removed (Figure 2e), so 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 shown to be a reliable reporter site in the variant protein KOR[C340A, C345S], provided that TET labeling was performed with IMCM.

The human $A_{2A}AR$ has six TM cysteines and no intra- or extracellular cysteines.[8] 1D $^{19}$F-NMR data for $A_{2A}AR$ labeled with TET in micelles exhibited a broad $^{19}$F-NMR signal owing to labeling of indigenous TM cysteines (Figure 3a). By contrast, IMCM of $A_{2A}AR$ with TET resulted in an empty NMR spectrum, thus demonstrating that labeling of the 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 $A_{2A}AR$. In the variant protein $A_{2A}AR$[A289C], TET labeling in micelles yielded a $^{19}$F-NMR spectrum with two signals of closely similar intensity at $\delta \approx 9.5$ and 11.5 ppm (Figure 3c). TET labeling by IMCM resulted in a related spectrum, however, the signal at $\delta \approx 9.5$ ppm showed 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,c]

Validation of the IMCM approach for $^{19}$F-NMR studies was obtained from two lines of experiments. Firstly, IMCM was applied with the previously extensively studied $\beta_2$-adrenergic receptor ($\beta_2AR$), which contains four TM cysteines and three intracellular cysteines. Previous work[1b] had

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

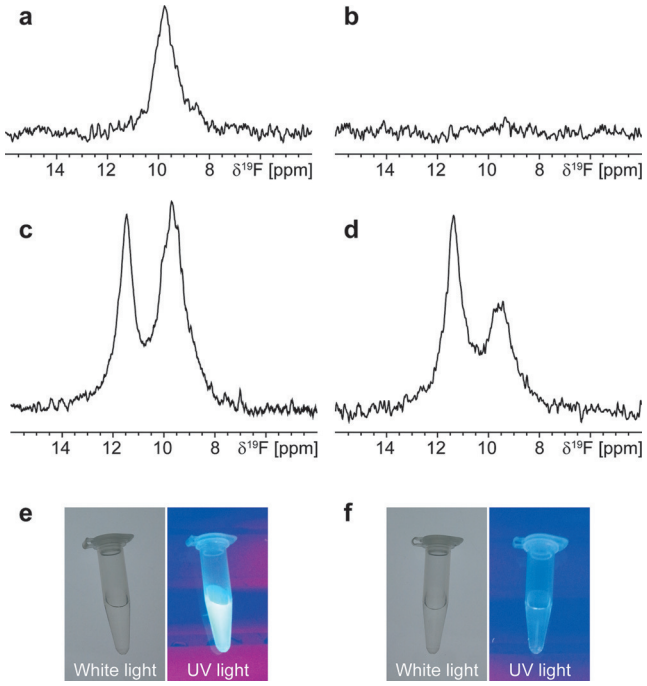


***Figure 3.*** Site-specific IMCM labeling of human $A_{2A}AR$ for $^{19}F$-NMR and fluorescence spectroscopy. a,b) 1D $^{19}F$-NMR spectra of $A_{2A}AR$ labeled with TET either in micelles (a) or by using IMCM (b). c,d) 1D $^{19}F$-NMR spectra of $A_{2A}AR$[A289C] labeled with TET either in micelles (c) or by using IMCM (d). e) $A_{2A}AR$[A289C] IMCM-labeled with fluorescein-5-maleimide is shown under white light and UV light. f) The same presentation as in (e) is shown for wild-type $A_{2A}AR$ that was subjected to identical labeling conditions.

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 $\beta_2AR$. The resulting $^{19}F$-NMR spectrum contained three lines which, through additional mutational studies, were individually assigned to the three intracellular cysteines in positions 265, 327 and 341 (Figure S1a in the Supporting Information). By using IMCM for TET labeling, the same spectral features were obtained (Figure S1b), thus showing that the IMCM preparation coincides with the one from the "proven" conventional approach. Clearly, the previously reported results on $\beta_2AR$ could have been obtained more efficiently through 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). By contrast, IMCM labeling of $A_{2A}AR$ produced a homogenous sample of unlabeled receptor (Figure S2b), owing to complete suppression of cysteine labeling in the TM helices. ESI-MS data for engineered $A_{2A}AR$[A289C] labeled by IMCM showed a homogenous preparation of singly TET-labeled receptor (Figure S2d), thus 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 the site-specific fluorescent labeling of GPCRs, the variant receptor $A_{2A}AR$-[A289C] was IMCM-labeled with fluorescein-5-maleimide. 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), thus 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 and these results show that they are 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 the selective covalent labeling of GPCRs with thiol- and maleimide-based reagents.

In summary, we successfully applied IMCM for the selective covalent attachment of chromophores for biophysical studies to indigenous or engineered intracellular cysteine residues in human $A_{2A}AR$ and KOR. We demonstrated that IMCM can facilitate spectroscopic studies of GPCRs with $^{19}F$-NMR and fluorescent 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 finding may result in a further increase in the 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. This is illustrated by the broader 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 the formation of intramolecular disulfide bonds in GPCRs (Figure 1). Finally, in addition to its impressive potential for in vitro studies of GPCRs, IMCM could be a starting point for in-cell labeling experiments designed to monitor GPCRs in their natural environment by spectroscopic techniques, along the lines of the "cell-surface engineering" described by Saxon and Bertozzi.[10]

## Acknowledgements

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 Metabolomics and Mass Spectrometry for help with ESI-MS analysis; J. Velasquez for help with molecular biology; and K. Villers, C. Hanson, and M. Chu for help with baculovirus expression.

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

**Keywords:** fluorescence spectroscopy · G-protein-coupled receptors · membrane proteins · NMR spectroscopy · site-specific labeling

**How to cite:** *Angew. Chem. Int. Ed.* **2015**, *54*, 15246–15249
*Angew. Chem.* **2015**, *127*, 15461–15464

[1] a) U. Gether, S. Lin, P. Ghanouni, J. A. Ballesteros, H. Weinstein, B. K. Kobilka, *EMBO J.* **1997**, *16*, 6737–6747; b) J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, *Science* **2012**, *335*, 1106–1110.

[2] R. Fredriksson, M. C. Lagerstrom, L. G. Lundin, H. B. Schioth, *Mol. Pharmacol.* **2003**, *63*, 1256–1272.

[3] A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730.

[4] a) M. C. Loewen, J. Klein-Seetharaman, E. V. Getmanova, P. J. Reeves, H. Schwalbe, H. G. Khorana, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4888–4892; b) K. Y. Chung, T. H. Kim, A. Manglik, R. Alvares, B. K. Kobilka, R. S. Prosser, *J. Biol. Chem.* **2012**, *287*, 36305–36311; c) J. F. Fay, D. L. Farrens, *J. Biol. Chem.* **2012**, *287*, 33873–33882; d) R. Rahmeh, M. Damian, M. Cottet, H. Orcel, C. Mendre, T. Durroux, K. S. Sharma, G. Durand, B. Pucci, E. Trinquet, J. M. Zwier, X. Deupi, P. Bron, J. L. Baneres, B. Mouillac, S. Granier, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6733–6738; e) T. H. Kim, K. Y. Chung, A. Manglik, A. L. Hansen, R. O. Dror, T. J. Mildorf, D. E. Shaw, B. K. Kobilka, R. S. Prosser, *J. Am. Chem. Soc.* **2013**, *135*, 9465–9474; f) R. Horst, J. J. Liu, R. C. Stevens, K. Wüthrich, *Angew. Chem. Int. Ed.* **2013**, *52*, 10762–10765; *Angew. Chem.* **2013**, *125*, 10962–10965; g) S. Mathiasen, S. M. Christensen, J. J. Fung, S. G. Rasmussen, J. F. Fay, S. K. Jorgensen, S. Veshaguri, D. L. Farrens, M. Kiskowski, B. Kobilka, D. Stamou, *Nat. Methods* **2014**, *11*, 931–934; h) A. Manglik, T. H. Kim, M. Masureel, C. Altenbach, Z. Yang, D. Hilger, M. T. Lerch, T. S. Kobilka, F. S. Thian, W. L. Hubbell, R. S. Prosser, B. K. Kobilka, *Cell* **2015**, *161*, 1101–1111.

[5] T. Didenko, J. J. Liu, R. Horst, R. C. Stevens, K. Wüthrich, *Curr. Opin. Struct. Biol.* **2013**, *23*, 740–747.

[6] D. H. Jones, S. E. Cellitti, X. Hao, Q. Zhang, M. Jahnz, D. Summerer, P. G. Schultz, T. Uno, B. H. Geierstanger, *J. Biomol. NMR* **2010**, *46*, 89–100.

[7] H. Wu, D. Wacker, M. Mileni, V. Katritch, G. W. Han, E. Vardy, W. Liu, A. Thompson, X.-P. Huang, I. Carroll, W. Mascarella, R. Westkaemper, P. Mosier, B. Roth, V. Cherezov, R. Stevens, *Nature* **2012**, *485*, 327–332.

[8] a) V. P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. IJzerman, R. C. Stevens, *Science* **2008**, *322*, 1211–1217; b) G. Lebon, T. Warne, P. Edwards, K. Bennett, C. Langmead, A. Leslie, C. Tate, *Nature* **2011**, *474*, 521–525; c) F. Xu, H. Wu, V. Katritch, G. W. Han, K. A. Jacobson, Z. G. Gao, V. Cherezov, R. C. Stevens, *Science* **2011**, *332*, 322–327; d) W. Liu, E. Chun, A. A. Thompson, P. Chubukov, F. Xu, V. Katritch, G. W. Han, C. B. Roth, L. H. Heitman, A. P. IJzerman, V. Cherezov, R. C. Stevens, *Science* **2012**, *337*, 232–236.

[9] M. Brinkley, *Bioconjugate Chem.* **1992**, *3*, 2–13.

[10] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.

Received: September 10, 2015
Published online: November 6, 2015

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 