Genetic Code

DOI: 10.1002/anie.201107211

Genetically Encoded Photocrosslinkers as Molecular Probes To Study G-Protein-Coupled Receptors (GPCRs)**

Annette G. Beck-Sickinger and Nediljko Budisa**

amino acids · photoactivation · photochemistry · receptors · receptor–ligand interactions

Crystal structures of various receptors and receptor/ligand complexes provide a major impetus for the work on dynamic features of these systems. However, receptors are usually integral membrane proteins, which are notoriously difficult to crystallize.[1] In addition, in most cases structures usually capture only a snapshot of a ligand-bound ground state and accordingly are limited to a single static image of one particular conformation state. As numerous hydrophilic ligands including pheromones, hormones, neurotransmitters, small molecules, peptides, and even large proteins mediate signals from the environment to the intracellular compartment via receptors, this class of proteins plays a major role in intercellular communication. Therefore, biophysical, biochemical, spectroscopic, and functional studies are urgently required to get more insight into the specificity, affinity of ligand recognition, and binding as well as the ligand/receptor conformational dynamics.

G-protein-coupled receptors (GPCRs) are classical examples of dynamic membrane proteins involved in triggering and regulating cellular signalling pathways.[2] Not surprisingly, the GPCR family comprises about 4% of the entire protein-encoding part of the human genome. GPCRs are grouped into distinct classes named A, B, C, or GRAFS according to major representatives (glutamate, rhodopsin, adhesion, frizzled, secretin families).[2] As currently 30% of all approved drugs target GPCRs, precise chemical information about the molecular mechanism of transmembrane signaling and the understanding the molecular pharmacology of this process is of outstanding interest.[3]

[*] Prof. Dr. A. G. Beck-Sickinger
AK Biochemie und Biooranische Chemie, Institut für Biochemie
Universität Leipzig
Brüderstrasse 34, 04103 Leipzig (Germany)
E-mail: beck-sickinger@uni-leipzig.de
Homepage: http://www.biochemie.uni-leipzig.de/agbs
Prof. Dr. N. Budisa
AK Biokatalyse, Insitut für Chemie
Technische Universität Berlin
Franklinstrasse 29, 10587 Berlin (Germany)
E-mail: budisa@biocat.tu-berlin.de
Homepage: http://www.biocat.tu-berlin.de

[**] We are grateful to Dr. Michael Hoesl for the critical reading of this manuscript.

The incorporation of noncanonical amino acids (ncAAs) into proteins has become an important tool to study the biochemical properties of proteins in a sophisticated manner and can be used to compensate for the lack of structural information. The advantage of the site-directed incorporation of ncAAs is the possibility of introducing specific reactive amino acid side chains. These can be used to study specific structural constraints by derivatization, for example labeling with a fluorescent tag or with biotin, or by site-specific crosslinking to interaction partners.

Different techniques have been developed to introduce ncAAs, including native chemical ligation[4] and expansion of the genetic code by using suppressor-based technologies[5] or auxotrophic bacteria (Figure 1).[6] Photolabile ncAAs such as *p*-benzoyl-L-phenylalanine (Bpa) which covalently bind side

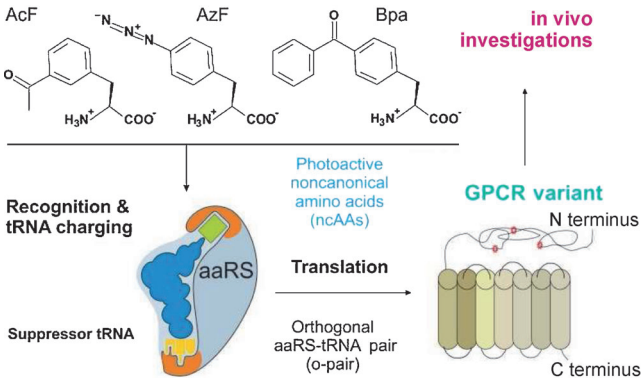

Figure 1. In vivo incorporation of ncAAs in mammalian cells for near-native studies of GPCRs. GPCRs are dynamic membrane proteins that trigger and regulate cellular signaling pathways. Most commonly used ncAAs in GPCR structure–function studies are *p*-acetyl-L-phenylalanine (AcF), which can react with a hydrazide or hydroxylamine derivative, and *p*-benzoyl-L-phenylalanine (Bpa), a useful photoactivatable crosslinker upon exposure to UV light. *p*-Azido-L-phenylalanine (azF) is useful as an IR probe and as a tag for bioorthogonal conjugations such as click chemistry and the Staudinger–Bertozzi ligation. Orthogonal pairs are derived from evolutionary distant cells and organisms by exploiting species specificities in the aminoacylation reaction (i.e. amino acid activation and tRNA charging). These ncAAs introduced into GPCRs facilitate studies of ligand–receptor interactions as well as intra- and intermolecular conformational dynamics and signal transduction.

chains of interacting partners at a distance of 3–4 Å upon irradiation are especially useful. This in situ noninvasive chemistry, however, is only possible if the reprogrammed translational machinery is available in the studied cells. Furthermore, incorporated keto amino acids such as *p*-acetyl-L-phenylalanine (AcF) allow the specific labeling with hydrazine or hydroxylamine derivatives.[7]

Owing to the low abundance of GPCR in cells, and the difficulties in recombinant bacterial expression and solubilization, few examples of the efficient expression of GPCR containing ncAAs have been reported up to now. In 2008, Sakmar and co-workers incorporated two distinct ncAAs (AcF and Bpa) into two different GPCRs: CCR5 (a chemokine receptor and major co-receptor for the human immunodeficiency virus) and rhodopsin.[8] Both AcF and Bpa were introduced into GPCRs in *HEK293T* mammalian cells by using two orthogonal pairs (o-pairs) consisting of the *Bacillus stearothermophillus* suppressor tRNA (*Bst-Yam*) and evolved *Escherichia coli* tyroysl-tRNA synthetase (TyrRS) variants AcFRS and BpaRS. These o-pairs were previously developed for ncAA translation by read-through of an Amber termination codon (UAG) in yeast mRNA. Both GPCRs, which harbor reactive keto groups at specific positions, were generated in sufficient amounts such that a variety of spectroscopic probes could be introduced (e.g. rhodopsin could be noninvasively labeled by fluorescein hydrazide).[9] Similarly, Becker and co-workers used Bpa to probe the potential binding site of the pheromone receptor *Ste2* (a yeast GPCR). They were able to identify positive crosslinks as potential binding sites for its natural peptide α-factor ligand.[10] In the meantime, AzF was also successfully incorporated into rhodopsin to measure its dynamics by Fourier transform infrared (FTIR) spectroscopy.[11]

These first experiments fully confirmed that site-specific ncAA insertion within the receptor is an invaluable noninvasive tool to study the mechanisms and dynamics of GPCR ligand binding and signal transmission. It is of particular importance that the repertoire of ncAAs makes it possible to gain information about GPCR structure and function by, for example, trapping a receptor–ligand complex using photoactivatable crosslinkers or by monitoring its mobility with fluorescent probes. In addition, the labeling of a receptor using o-pairs might be combined with the use of various synthetic ligands. This year, Sakmar and co-workers employed an in situ photocrosslinking approach by using o-pairs for AzF and AcF to study the binding interactions of inhibitor T140 specific for CXC chemokine receptor 4 (CXCR4). This receptor is important in directed cell migration, cancer metastasis, and HIV entry. They tested eight amino acid positions at the receptor interface and mapped a unique UV-light-dependent crosslink in vivo—the site of specific interaction between CXCR4 residue 189 and T140 (HIV-1 co-receptor blocker). Importantly, their findings were consistent with crystal-structure-based molecular modeling with peptide inhibitors.[12]

The first study of a class B GPCR was recently reported by Wang and co-workers,[13] who studied the corticotropin-releasing factor receptor type 1 (CRF-R1), which upon endogenous peptide–ligand binding participates in stress response. Like other members of class B GPCRs (or secretin family), CRF-R1 has a large N-terminal domain (roughly 120 residues) which serves as the major binding site of peptide-based ligands. Since no full-length receptor has been structurally characterized by NMR spectroscopy or X-ray crystallography, the structural information about these receptors is mainly from structure–activity relationship studies and classical crosslinking studies (e.g. derivatization of cysteines with thiol-specific maleimide-based or haloacetamide-based reagents). In this study, CRF-R1 receptors were investigated in 293T cells transfected with the plasmid encoding the o-pair ($tRNA_{CUA}$(Tyr)-AzFRS derived from *E. coli*) for site-specific AzF incorporation in response to the UAG codon. In the experimental setup, the endogenous ligand was radioactively labeled. This made it possible to identify crosslinking points usually not found in traditional crosslinking investigations. Fine and distinct mapping of these interactions was further enabled by studying the crosslinking behavior of different ligands at the same AzF-containing variant receptor set. Four different AzF-containing CRF-R1 variants were tested for crosslinking in parallel with four radiolabeled ligands. Curiously, the number of crosslinking sites was surprisingly small, indicating that the ligands are not tightly bound to the receptors in the native cellular context of the CRF-R1 complexes.

With these recent experiments and breakthroughs, the technology using the expanded genetic code proved to be indispensable for noninvasive in vivo studies in mammalian cells and even whole organisms, since it provides chemical handles as probes for monitoring receptor–ligand interactions, receptor activation, and the real-time interrogation of active signaling complexes. However, routine application of this methodology still faces a number of challenges. Examples are the correct transcription of the tRNAs from o-pairs, and subsequent processing, modification, and export to the cytoplasm of the cell, and the requirement that an o-aaRS should be highly specific for both the cognate orthogonal tRNA and thencAA. Finally, the mRNA encoding a gene of interest and bearing the *in frame* stop codon might become a substrate of nonsense-mediated decay (NMD)[14] which might destroy transcripts with premature termination signals. Results obtained by the use of *p*-azido-L-phenylalanine (AzF) in the biological context should be critically evaluated as it is well known that upon its incorporation into the target protein the reduced forms are observed as well. This could contribute to the chemical reactivity and photoinstability of the aryl-azido group during different analytical procedures such as mass analyses. However, it has been described that, for example, *Saccharomyces cerevisiae* can be used as a biocatalyst to reduce arylazides to arylamines.[15]

Beside the problem of chemical and metabolic stability as well as intracellular uptake of particular ncAAs, there are at least four issues that should be taken into consideration before planning and designing these experiments. First, while the incorporation efficiency of ncAAs is relatively high in *E. coli*, it is generally much lower in yeast and especially in mammalian cells. Namely, low read-through of Amber stop codons might cause an increase in truncated protein fractions, which can be potentially toxic or interfere with the function of

the full-length target protein. Therefore, the application of ncAA mutagenesis in mammalian cells requires more effort to improve the efficiency of specific o-pairs. Second, efficient translation of various ncAAs into membrane proteins in mammalian cells is still a challenging task, not only because of the inefficiency of the o-pairs but also because of low protein yields in systems where these are not highly expressed. Third, the type of ncAA used is always dictated by the biological problem to be studied. For example, photocaged amino acids in living cells allow the photocontrol of protein interactions, protein localization, enzymatic activity, and cellular signaling. In contrast, translation of photocrosslinking ncAAs is useful for the mapping of weak or transient, noncovalent protein interactions. Conformationally restricted ncAAs or amino acid analogues might be incorporated to study the role of local constraints and energy barriers.[16] Finally, for biological studies requiring protein imaging and spectroscopy, the incorporation of ncAAs with bio-orthogonal chemical handles and biophysical probes is the first choice. However, in many cases it may be necessary to combine or replace a suppressor-based methodology with other approaches[17] including SNAP-tag technology, FlaSH-tag approaches, GFP fusions, classical Cys modifications,[18] and even semi-synthetic methods such as expressed enzymatic or native protein ligations.[19] For example, ncAA approaches are not suitable when backbone modifications are examined, pulse-chase experiments are required, large nonnatural segments need to be introduced, when protein immobilization is required or when subsequent modification may not work owing to low concentrations. Therefore, despite the current success, this is just the beginning of an emerging field that requires chemistry compatible with biological systems. There is still much room for improvement, especially when membrane proteins such as receptors have to be examined or intracellular proteins specifically modified.

Received: October 12, 2011
Published online: November 30, 2011

[1] C. Ostermeier, H. Michel, *Curr. Opin. Struct. Biol.* **1997**, *7*, 697.
[2] a) A. G. Gilman, *Annu. Rev. Biochem.* **1987**, *56*, 615; b) H. B. Schiöth, R. Fredriksson, *Gen. Comp. Endocrinol.* **2005**, *142*, 94.
[3] J. P. Overington, B. Al-Lazikani, A. L. Hopkins, *Nat. Rev. Drug Discovery* **2006**, *5*, 993.
[4] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776; b) T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705; c) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.
[5] a) H. H. Chung, D. R. Benson, P. G. Schultz, *Science* **1993**, *259*, 806; b) R. Furter, *Protein Sci.* **1998**, *7*, 419; c) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498.
[6] a) N. Budisa, *Angew. Chem.* **2004**, *116*, 6586; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426; b) A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603; c) M. Hoesl, N. Budisa, *Angew. Chem.* **2011**, *123*, 2948; *Angew. Chem. Int. Ed.* **2011**, *50*, 2896.
[7] K. A. Daggett, T. P. Sakmar, *Curr. Opin. Chem. Biol.* **2011**, *15*, 392.
[8] T. Huber, S. Menon, T. P. Sakmar, *Biochemistry* **2008**, *47*, 11013.
[9] S. Ye, C. Köhrer, T. Huber, M. Kazmi, P. Sachdev, E. C. Yan, A. Bhagat, U. L. RajBhandary, T. P. Sakmar, *J. Biol. Chem.* **2008**, *283*, 1525.
[10] L. Y. Huang, G. Umanah, M. Hauser, C. Son, B. Arshava, F. Naider, J. M. Becker, *Biochemistry* **2008**, *47*, 5638.
[11] S. Ye, E. Zaitseva, G. Caltabiano, G. F. Schertler, T. P. Sakmar, X. Deupi, R. Vogel, *Nature* **2010**, *464*, 1386.
[12] A. Grunbeck, T. Huber, P. Sachdev, T. P. Sakmar, *Biochemistry* **2011**, *50*, 3411.
[13] I. Coin, M. H. Perrin, W. W. Vale, L. Wang, *Angew. Chem.* **2011**, *123*, 8227; *Angew. Chem. Int. Ed.* **2011**, *50*, 8077.
[14] J. Hodgkin, A. Papp, R. Pulak, V. Ambros, P. Anderson, *Genetics* **1989**, *123*, 301.
[15] a) A. Kamal, Y. Damayanthi, B. L. Narayan Reddy, B. S. Praveen Reddy, *Chem. Commun.* **1997**, 1015; b) J. W. Back, O. David, G. Kramer, G. Masson, P. T. Kasper, L. J. de Koning, L. de Jong, J. H. van Maarseveen, C. G. de Koster, *Angew. Chem.* **2005**, *117*, 8160; *Angew. Chem. Int. Ed.* **2005**, *44*, 7946.
[16] a) E. B. van Arnam, H. A. Lester, D. A. Dougherty, *ACS Chem. Biol.* **2011**, PMID: 21776983; b) L. Moroder, N. Budisa, *ChemPhysChem* **2010**, *11*, 1181.
[17] a) M. J. Hinner, K. Johnsson, *Curr. Opin. Biotechnol.* **2010**, *21*, 766; b) I. Böhme, A. G. Beck-Sickinger, *Cell Commun. Signaling* **2009**, *7*, 16; c) C. Jing, V. W. Cornish, *Acc. Chem. Res.* **2011**, *44*, 784.
[18] L. Merkel, L. Moroder, N. Budisa in *Protein Eng* (Eds.: U. RajBhandary, C. Koehrer), Springer, Heidelberg, **2008**, pp. 29–64.
[19] a) Z. Machova, R. von Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, *Angew. Chem.* **2003**, *115*, 5065; *Angew. Chem. Int. Ed.* **2003**, *42*, 4916; b) R. David, R. Günther, L. Baumann, T. Lühmann, D. Seebach, H. J. Hofmann, A. G. Beck-Sickinger, *J. Am. Chem. Soc.* **2008**, *130*, 15311; c) K. Bellmann-Sickert, A. G. Beck-Sickinger, *ChemMedChem* **2011**, *6*, 193.

