Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 15 (2005) 3654–3657

Structural biology of G protein-coupled receptors

Kenneth Lundstrom

BioXtal, Chemin des Croisettes 22, CH-1066 Epalinges, Switzerland

Received 21 March 2005; revised 29 April 2005; accepted 11 May 2005
Available online 2 June 2005

Abstract—More than 60% of the current drugs are based on G protein-coupled receptors. Paradoxically, high-resolution structures are not available to facilitate rational drug design. Difficulties in expression, purification, and crystallization of these transmembrane receptors are the reasons for the low success rate. Recent individual and network-based technology development has significantly improved our knowledge of structural biology and might soon bring a major breakthrough in this area.
© 2005 Elsevier Ltd. All rights reserved.

G protein-coupled receptors (GPCRs) mediate many important cellular signal transduction events related to differentiation, proliferation, angiogenesis, cancer, development, and cell survival.[1] The sales of drugs based on GPCRs were $47 billion in 2003.[2] However, structure-based drug design has not been possible, as the three-dimensional structure has only been solved for a single non-therapeutic GPCR, the bovine rhodopsin.[3] The reasons for the low success rate in structure resolution on GPCRs can be found in the complications related to expression of recombinant receptors due to the inefficient transport and insertion of these seven-transmembrane (7TM) receptors in the plasma membrane and the toxic effect on host cells. Additionally, GPCR purification requires detergents, which have a negative effect on receptor yields and stability. The crystallization process is substantially compromised by the presence of detergents, the inherent flexibility of GPCRs, and the relatively small hydrophilic loops providing fewer potential crystal contacts. Intensive technology development has been performed in individual laboratories and more recently in large networks.

Prior to efforts for large-scale purification of GPCRs, a multitude of studies have been carried out to learn more about their structure, particularly by site-directed mutagenesis approaches based on molecular modeling of rhodopsin. Two excellent reviews have been published[4,5] and this topic is described here only briefly. The high-resolution structure of bacteriorhodopsin, a light-driven proton pump from *Halobacterium halobium*,[6] was extensively used to develop several tertiary models based on extensive analysis of multiple GPCRs. More recently, three-dimensional models have been based on the bovine rhodopsin structure as a template.[3] Rhodopsin has been subjected to electron paramagnetic resonance (EPR) spectroscopy and Cys-crosslinking experiments.[7] These studies have provided information on the orientation of helices, protrusions of helices into the cytoplasmic surface, relative flexibility of different loops, and assessment of conformational changes induced by light.[8]

In studies to determine distance constraints between the transmembrane (TM) regions, chimeric muscarinic M_2/M_5 receptors have been engineered.[9] Another approach has been to introduce histidine zinc(II) binding sites in the neurokinin-1[10] (NK1R) and κ opioid[11] receptors, which made it possible to define the proximity and orientation of TM3 in relation to TMs 2 and 5. Distance constraints and flexibility of extracellular loops have been examined by introduction of cysteine mutations resulting in intramolecular disulfide bridges in the M_3 muscarinic receptor.[12] Tertiary structure relationships have been studied by the introduction of a fluorescent unnatural amino acid at known sites in the neurokinin-2 receptor based on the suppression of UAG nonsense codons and application of modified tRNAs.[13] This approach allowed distance determinations of the NK2 receptor in *Xenopus* oocytes between fluorescent peptides by fluorescence resonance energy transfer and defined sites containing the fluorescent amino acid.

Site-directed mutagenesis studies have also contributed to structural understanding of GPCRs to determine the local structure for ligand binding. The presence of a conserved aspartic acid in TM2 and an asparagine in TM7 in nearly all class A rhodopsin-related GPCRs

E-mail: klundstrom@mepnet.org

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.05.041

led to a study on TM2–TM7 proximity as the GnRH receptor uniquely has an asparagine in TM2 and an aspartic acid residue in TM7. Replacement of asparagine with aspartic acid in TM2 abolished any detectable binding activity of the GnRH receptor. However, introduction of a second mutation, aspartic acid in TM7, resulted in high-affinity binding, suggesting a close proximity between TM2 and TM7.[14] Ligand-binding sites and their environment have frequently been studied by mutational mapping. In this context, three residues (Asn23, Glu24, and Phe25) at the N-terminus affected substance P binding to NK1R.[15] Moreover, the His108 on the top of TM3 and Tyr287 on the top of TM7 showed interaction with substance P (SP).[16] Interestingly, the binding mode for the nonpeptide antagonist CP 96,345 was significantly different from that of substance P.[17] This could be further confirmed by mutagenesis studies, which demonstrated that the binding pocket for CP 96,345 is not occupied by SP.[18] In another study, molecular modeling and site-directed mutagenesis were applied to analyze antagonist binding of 11 amino acid residues located in the vicinity of the binding pocket. Typically, the NK1R His197 mutant showed a significantly reduced affinity for CP 96,345.[19] Similarly, substantial differences in binding properties of peptide agonists and nonpeptide antagonists have been demonstrated for angiotensin,[20] opioid,[21] cholecystokinin/gastrin,[22] and neurotensin[23] receptors. Recently, homology modeling of the NK1 receptor based on the X-ray structure of bovine rhodopsin allowed virtual screening, which led to selection of seven compounds, one of which showed submicromolar affinity.[24]

Purification of sufficient quantities of GPCRs for structural studies requires high yields of recombinant proteins. In this context, expression has been evaluated in all available systems spanning from bacteria to yeast, insect, and mammalian cells[25] (Table 1). The rat neurotensin receptor expressed as a fusion protein to the maltose-binding protein (MBP) resulted in milligram yields of functional receptor in *Escherichia coli* membranes.[26] In a different approach the human leukotriene BLT1 receptor was produced in large amounts in bacterial inclusion bodies.[27] Additionally, several GPCRs have been expressed at high levels (B_{max} 45 pmol/mg protein) in yeast, and large-scale fermentor production resulted in yields of at least 5 mg/L.[28] Baculovirus-based insect cell expression generated high specific binding of the human NK1R and allowed purification of the functional receptor.[29] Although mammalian expression generally has generated only relatively modest yields, the application of a tetracycline-inducible HEK293 cell line produced 6 mg rhodopsin in a 1.1-L suspension culture.[30] Application of Semliki Forest virus (SFV) vectors has resulted in specific binding of 60–200 pmol/mg protein and yields of 10 mg/L in mammalian suspension cells.[31]

Previously, the rat neurotensin[23] and the human adenosine A2a receptor[32] were purified in a fully functional form for structural studies. Likewise, the leukotriene BLT1 receptor was refolded in an active form from bacterial inclusion bodies.[27] Application of *P. pastoris* expression generated densities of 360 g/L in fermentor cultures, which allowed solubilization of the ETB endothelin receptor with *n*-dodecyl-β-D-maltose-pyranoside followed by successful purification.[28] Several GPCRs have also been produced on a large scale in insect cells using baculovirus vectors. In this context, the human NK1R was purified from Sf9 cells by metal-chelating and gel filtration chromatographies, resulting in pure receptor with properties similar to those of the native NK1R.[29] Similarly, rhodopsin has been expressed from mammalian cells by applying an inducible stable system.[30] Finally, overexpression of several GPCRs in mammalian host cells using SFV vectors has yielded large quantities of purified receptors.[31]

Despite all these efforts on recombinantly expressed GPCRs described above, no high-quality crystals with diffraction properties suitable for high-resolution structures have been obtained, so far. The reasons for this might be the incorrect transport and folding of recombinant GPCRs, the low stability of the solubilized and purified receptors, and the inherent flexibility of GPCRs, which makes crystallization more difficult. To address these issues from a broader perspective, several national and international networks with expertise in expression, biochemistry, purification, and crystallography have been established.[1] The USA and Japan have been at the forefront in structural genomics. Several NIH-sponsored initiatives have been established in the USA. Among them, the Joint Center for Structural Genomics (JCSG) in California works on 180 GPCRs, which are expressed from *E. coli*, baculovirus, and adenovirus vectors. Another large network, the South-East Collaboratory on Structural Genomics (SECSG), runs a program on the *Caenorhabditis elegans* genome, which includes GPCRs to be expressed in insect and mamma-

Table 1. Expression of GPCRs in heterologous expression systems

System	Host cells	GPCR (example)	Expression level (mg/L)
Bacterial	*E. coli*	Neurotensin	1
		Leukotriene BLT1	10
Yeast	*S. cerevisiae*	Dopamine D1A	1
	S. pombe	Dopamine D2	1
	P. pastoris	β_2 Adrenergic	5
Baculovirus	Sf9 insect cells	Neurokinin-1	1
Mammalian (T)	CHO, HEK cells	Serotonin 5-HT1E	0.1
Mammalian (S)	HEK cells	Rhodopsin	6
Viral (SFV)	BHK, CHO cells	α_2 Adrenergic	10

T, transient; S, stable; SFV, Semliki Forest virus.

lian cells from baculovirus and lentivirus vectors, respectively. In Japan, the RIKEN Structural Genomics Initiative studies the whole mouse genome with the intention of selecting specific GPCRs for structural characterization. The Biological Information Research Center (AIST) and the Japan Biological Informatics Consortium (JBIC) have established a joint project using bacteriorhodopsin and some mammalian GPCRs as targets. For instance, wild-type and mutant turkey β_2-adrenergic receptors have been expressed in large quantities from baculovirus vectors in insect cells for structural studies (Table 2).

In Europe, several national networks studying GPCRs have been established. The Swiss NCCR (National Center of Competence in Research) has member research teams from the universities of Basel and Zurich and the Paul Sherrer Institute (PSI) in Villigen. Several GPCRs are expressed in *E. coli* and insect cells in their program. The Postgenome Research and Technology Program SweGene in South West Sweden deals also with GPCRs. The hosts for overexpression are mainly *E. coli* and yeast cells. A newly established EU-funded network called E-MeP studies 100 prokaryotic and 200 eukaryotic membrane proteins. Among the eukaryotic targets 100 are GPCRs, of which the expression of selected targets will initially be evaluated in *E. coli*, *Lactococcus lactis*, *Saccharomyces cerevisiae*, *P. pastoris*, baculovirus-infected insect cells, and SFV-infected mammalian cells. Additionally, cell-free translation is verified for GPCR expression. E-MeP consists of 18 laboratories in six European countries.

The largest dedication to structural biology of GPCRs has so far been through the privately funded consortium called MePNet (Membrane Protein Network), which has evaluated expression levels of 101 GPCRs in bacteria, yeast, and mammalian cells in parallel. Immunodetection indicated that approximately 50% of the GPCRs were expressed in *E. coli* applying either pET15 or Gateway vectors. The success rate was significantly higher for Gateway vector-based expression. Expression in yeast cells from *P. pastoris* vectors generated positive signals in 95% of the targets. Similarly, high success rates (96%) were obtained in mammalian cells infected with recombinant SFV particles. Although the expression levels varied, overall 60 GPCRs were considered to be structural biology compatible, generating 1 mg/L or higher quantities of GPCRs. As the GPCRs were uniquely expressed in inclusion bodies in *E. coli*, further refolding was necessary to restore the functionality of receptors. Selected GPCRs were produced on a large scale in fermentor culture resulting in yields up to 375 mg/L. Currently several GPCRs have been purified and subjected to refolding exercises.

GPCR expression in yeast and mammalian membranes allowed determination of functional receptor activity by binding studies with specific radioligands. Specific binding activity was determined for some 60 and 40 GPCRs expressed in yeast and mammalian cells, respectively. A large number of GPCRs showed specific binding higher than 10 pmol/mg protein. The binding activity in *P. pastoris* could be significantly improved by lowering the culture temperature and by supplementation of additives and ligands to the culture medium. Typically, 20-fold increase resulting in more than 100 pmol receptor per milligram protein was observed. SFV-based expression showed large variations in different cell lines (BHK, CHO, and HEK293 cells). The time (post-infection) of harvest also affected the expression levels significantly. Certain GPCRs showed maximum expression levels at 24 h, others at 48 or 72 h post-infection. Initial expression optimization on MePNet GPCRs suggested that addition of ligands to the cell culture medium can enhance the binding activity, as has been previously demonstrated for the rat histamine H2 receptor expressed in SFV-infected COS7 cells.[33] The highest binding activity was observed in SFV-infected BHK and CHO cells resulting in 287 pmol/mg protein. Several GPCRs have been successfully solubilized and purified from both yeast and mammalian cells. At present, the first GPCRs have been subjected to crystallization attempts.

In conclusion, various methods including molecular modeling and site-directed mutagenesis have allowed some preliminary characterization of the structure of GPCRs. For instance, it has been possible to determine the counter-clockwise orientation of the 7TMs in GPCRs. Furthermore, the proximity of individual TMs has been defined. Recent developments in technol-

Table 2. Structural genomics networks on membrane proteins

Network[a]	Targets	Website
USA		
JCSG	Mammalian GPCRs (180)	www.jcsg.org
SECSG	*C. elegans* GPCRs (20)	www.sgce.cbse.uab.edu
Japan		
RIKEN GSC	Murine GPCRs (n)	www.riken.go.jp
Japanese MP	Rhodopsin, β_2-AR	www.rcsb.org
Europe		
E-MeP	Mammalian MPs (200)	www.e-mep.org
MePNet	Mammalian GPCRs (100)	www.mepnet.org
NCCR	Mammalian GPCRs (10)	www.structuralbiology.ethz.ch
SweGene	Mammalian GPCRs (5)	www.swegene.org

[a] JCSG, Joint Center for Structural Genomics; SECSG, South-East Collaboratory on Structural Genomics.

ogy have also made it possible to express a large number of GPCRs in several expression systems (bacterial, yeast, insect, and mammalian cells) at levels that provide sufficient amounts for serious purification efforts. Refolding and solubilization still pose bottle-necks in structural biology of membrane proteins. Excess quantities of GPCRs, however, now allow screening of a large number of conditions for improved refolding efficacy. A variety of detergents under a multitude of conditions can be tested for solubilization. Additionally, novel detergents are evaluated. Furthermore, significant improvements in the area of crystallization, particularly application of methods on a nanoliter scale, have substantially improved the screening abilities of various crystallization conditions without the need of large quantities of precious material. Today, in several cases, purification to high purity of recombinant membrane proteins has been achieved, followed by successful crystallization. However, the obtained crystals have not diffracted, which has raised serious questions pertaining to the quality of not only the purified protein but also the composition of lipids and other materials. Further technology development is therefore essential in most areas, including expression, purification, and crystallography. The establishment of large networks with a broad expertise in areas of protein expression, purification, and crystallography will further enhance the possibilities of success to conquer this last frontier in structural biology. In the meantime, molecular modeling and virtual screening approaches have to rely on the rhodopsin structure, which has proven to be successful in deducing pharmacophore models to obtain high-affinity compounds as demonstrated for the NK1 receptor. These findings should underline the importance of structure-based drug design.

References and notes

1. Lundstrom, K. *Trends Biotechnol.* **2005**, *23*, 103.
2. Med Ad News Staff. *Med Ad News* **2004**, *23*, 60
3. Palczewski, K.; Kumasaki, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739.
4. Gether, U. *Endocr. Rev.* **2000**, *21*, 90.
5. Kristiansen, K. *Pharmacol. Ther.* **2004**, *103*, 21.
6. Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. H. *J. Mol. Biol.* **1990**, *213*, 899.
7. Columbus, L.; Hubbell, W. L. *Trends Biochem. Sci.* **2002**, *27*, 288.
8. Altenbach, C.; Cai, K.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1999**, *38*, 7931.
9. Pittel, Z.; Wess, J. *Mol. Pharmacol.* **1994**, *45*, 61.
10. Elling, C. E.; Schwartz, T. W. *EMBO J.* **1996**, *15*, 6213.
11. Thirstrup, K.; Elling, C. E.; Hjorth, S. A.; Schwartz, T. W. *J. Biol. Chem.* **1996**, *271*, 7875.
12. Zeng, F. Y.; Hopp, A.; Soldner, A.; Wess, J. *J. Biol. Chem.* **1999**, *274*, 16629.
13. Turcatti, G.; Nemeth, K.; Edgerton, M. D.; Meseth, U.; Talabot, F.; Peitsch, M.; Knowles, J.; Vogel, H.; Chollet, A. *J. Biol. Chem.* **1996**, *271*, 19991.
14. Zhou, W.; Flanagan, C.; Ballesteros, J. A.; Konvicka, K.; Davidson, J. S.; Weinstein, H.; Millar, R. P.; Sealfon, S. C. *Mol. Pharmacol.* **1994**, *45*, 165.
15. Fong, T. M.; Yu, H.; Huang, R. R. C.; Strader, C. D. *Biochemistry* **1992**, *31*, 11806.
16. Huang, R. R. C.; Yu, H.; Strader, C. D.; Fong, T. M. *Biochemistry* **1994**, *33*, 3007.
17. Gether, U.; Johansen, T. E.; Snider, R. M.; Lowe, J. A.; Nakanishi, S.; Schwartz, T. W. *Nature* **1993**, *362*, 345.
18. Sachais, B. S.; Snider, R. M.; Lowe, J. A.; Krause, J. E. *J. Biol. Chem.* **1993**, *268*, 2319.
19. Lundstrom, K.; Hawcock, A. B.; Vargas, A.; Ward, P.; Thomas, P.; Naylor, A. *Eur. J. Pharmacol.* **1997**, *337*, 73.
20. Schambye, H. T.; Hjorth, S. A.; Bergsma, D. J.; Sathe, G.; Schwartz, T. W. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7046.
21. Kong, H.; Raynor, K.; Yano, H.; Takeda, J.; Bell, G. I.; Reisine, T. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8042.
22. Labbe-Jullie, C.; Barroso, S.; Nicolas-Eteve, D.; Reversat, J. L.; Botto, J. M.; Mazella, J.; Berenassau, J. M.; Kitabgi, P. *J. Biol. Chem.* **1998**, *273*, 16351.
23. Blaker, M.; Ren, Y.; Gordon, M. C.; Hsu, J. E.; Beinborn, M.; Kopin, A. S. *Mol. Pharmacol.* **1998**, *54*, 857.
24. Evers, A.; Klebe, G. *J. Med. Chem.* **2004**, *47*, 5381.
25. Sarramegna, V.; Talmont, F.; Demange, P.; Milon, A. *Cell Mol. Life Sci.* **2003**, *60*, 1529.
26. Luca, S.; White, J. F.; Sohal, A. K.; Filippov, D. V.; van Boom, J. H.; Grisshammer, R.; Baldus, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10706.
27. Baneres, J. L.; Martin, A.; Hullot, P.; Girard, J. P.; Rossi, J. C.; Parello, J. *J. Mol. Biol.* **2003**, *329*, 801.
28. Schiller, H.; Molsberger, E.; Janssen, P.; Michel, H.; Reilander, H. *Receptors Channels* **2001**, *7*, 453.
29. Mazina, K. E.; Strader, C. D.; Fong, T. M. *J. Recept. Signal Transduct. Res.* **2001**, *16*, 191.
30. Reeves, P. J.; Kim, J. M.; Khorana, H. G. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13419.
31. Lundstrom, K. *Biochem. Biophys. Acta* **2003**, *1610*, 90.
32. Weiss, H. M.; Grisshammer, R. *Eur. J. Biochem.* **2002**, *269*, 82.
33. Hoffmann, M.; Verzijl, D.; Lundstrom, K.; Simmen, U.; Alewijnse, A. E.; Timmermann, H.; Leurs, R. *Eur. J. Pharmacol.* **2001**, *427*, 105.