Journal of Structural and Functional Genomics **6**: 159–163, 2005.
DOI 10.1007/s10969-005-1917-6
© springer 2005

Large-scale expression and purification of a G-protein-coupled receptor for structure determination – an overview

Reinhard Grisshammer[1,]* Jim F. White[1], Loc B. Trinh[2] & Joseph Shiloach[2]
[1]*Laboratory of Molecular Biology of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, 20892, USA;* [2]*Biotechnology Unit of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, 20892, USA; *Author for correspondence (e-mail: rkgriss@helix.nih.gov; Fax +1-301-480-0597)*

Received 6 August 2004; accepted in revised form 16 January 2005

Key words: automation, *Escherichia coli*, functional expression, G-protein-coupled receptor, large-scale purification

Abstract

Structure determination of G-protein-coupled receptors and other applications, such as nuclear magnetic resonance studies, require milligram quantities of purified, functional receptor protein on a regular basis. We present an overview on expression and purification studies with a receptor for neurotensin. Functional expression in *Escherichia coli* and an automated two-column purification routine allow ongoing crystallization experiments and studies on receptor-bound ligands.

Abbreviations: GPCR – G-protein-coupled receptor; H10 – deca-histidine tag; MBP – *E. coli* maltose-binding protein; NTR – rat neurotensin receptor NTS1; TrxA – *E. coli* thioredoxin

Introduction

G-protein-coupled receptors (GPCRs) are integral membrane proteins involved in many physiological processes such as cell-to-cell communication, mediation of hormonal activity and sensory transduction (for reviews see references [1, 2]). About 1000 GPCRs have been identified, 300–400 of which are found throughout the body binding endogenous ligands (the remainder are chemosensory GPCRs for odors, pheromones or taste) [3]. GPCRs are characterized by having seven hydrophobic transmembrane segments, which are connected by loops of varying lengths. The N-terminus of a GPCR is located at the extracellular side of the membrane. GPCRs are grouped into several classes, with the rhodopsin family being the largest (see GPCRDB: Information system for G-protein-coupled receptors at http://www.gpcr.org/7tm/, and ref. [4]).

GPCRs have been implicated as therapeutic routes to the treatment of human diseases and represent one of the major drug targets. Despite their striking clinical relevance, only one high-resolution structure (rhodopsin) is known [5, 6]. One important factor in the structure determination of rhodopsin was its availability in large quantities from natural sources. In contrast, recombinant overexpression followed by efficient purification methods are needed for GPCRs that occur naturally at low levels. Applications such as nuclear magnetic resonance, to determine for example the conformations of receptor-bound ligands [7–11], likewise require milligram quantities of functional GPCRs.

Structure determination of membrane receptors can be conceptually divided into five steps: (a) availability of the receptor gene or cDNA, (b) system for high level expression, (c) solubilization and purification in the presence of detergents, (d)

characterization and crystallization, (e) structure determination by established crystallographic methods. This summary will focus on aspects of receptor expression and large-scale automated purification, using the rat neurotensin receptor NTS1 (NTR) [12–14] as an example.

Expression

A prediction as to which expression system (bacteria, yeast or *Pichia pastoris*, baculovirus/insect cells, mammalian cells or other expression hosts) is best for the high-level production of a given GPCR is still difficult, despite numerous efforts to improve membrane protein expression systems (see references [15–17]). It is likely that there is no universal expression system suitable for functional production of all GPCRs, and the right combination between the target receptor and expression host must be found empirically. In the following, we discuss expression of functional GPCRs in *Escherichia coli*. For information on other expression systems, see for example references [16, 18, 19].

There are two strategies for the expression of GPCRs in *E. coli*: (a) expression of functional, membrane-inserted receptors (see below) and (b) expression of incorrectly folded, aggregated protein [20–23]. The latter approach requires refolding of the GPCR into its functional form. Although NTR shows little tendency for being deposited as non-functional aggregates [22], other receptors are readily overproduced as 'inclusion bodies'. This may reflect individuality of receptors. However, overexpression in aggregated form of the human leukotriene B4 receptor BLT1 is achieved by utilizing a synthetic BLT1 cDNA with codons preferentially used in *E. coli* [21]; use of the native cDNA did not produce significant amounts of misfolded protein.

Expression of functional, membrane-inserted NTR at levels sufficient for purification and crystallization experiments depends on the optimization of many parameters, rather than on one key factor. We found that transcription and translation efficiency (codon usage) did not seem critical for the overproduction of NTR in *E. coli*. In contrast, receptor insertion into the cytoplasmic membrane seems to be rate-limiting [24]. Correct insertion and folding can be aided by fusing the *E. coli* maltose-binding protein (MBP) with its own signal peptide to the N-terminus of NTR [24]. The efficient translocation of MBP across the *E. coli* cytoplasmic membrane positions the receptor N-terminus in the periplasm (which corresponds to the outer side of a eukaryotic plasma membrane), thus facilitating correct folding of the receptor [25]. By using this approach, correct transmembrane topology of the receptor is favored according to the 'positive-inside' rule [26]. In addition, production of functional receptors necessitates that the *E. coli* translocase (see reference [27]), the probable site of membrane insertion for MBP-NTR fusion proteins, is not overloaded by the nascent receptor chain. This can be achieved by using a low-copy number plasmid with a weak promoter [28] and low temperature (22 °C) during production. Fewer protein molecules are made at any given moment, but accumulation of correctly folded receptors is observed over time (40 h). Surprisingly, the nature of affinity tags at the receptor C-terminus has a substantial effect on the expression levels of MBP-NTR fusion proteins [28]. Fusing the compactly folded *E. coli* thioredoxin (TrxA) moiety to the receptor C-terminus results in the highest expression levels. Likewise, the specific amino acid sequence of the region between MBP and the receptor N-terminus has some influence on receptor expression levels, albeit to a lesser effect (unpublished results). Using optimal conditions, the NTR fusion protein accumulates to ~1000 receptors/cell as determined by [^{3}H]neurotensin binding (0.3–0.5 mg of NTR fusion protein per liter of culture, ~24 pmol/mg of crude membrane protein [29]). Fermentation at the 200-liter scale yields about 1.1 kg of wet bacterial cell paste with 90 mg of functional NTR fusion protein [30].

Is the above-described bacterial expression system applicable to all GPCRs ? As with other host systems, successful functional production depends on the nature of the respective receptor [17]. For example, a cannabinoid CB1 receptor is highly susceptible to *in vivo* proteolysis when expressed in *E. coli* [31]. Removal of defined protease-sensitive sites was part of the strategy for successful expression and large-scale purification of an adenosine A2a receptor [32], a M1 muscarinic acetylcholine receptor [33] and a neurotensin receptor [28]. Functional opioid receptors are

produced in *E. coli* at low levels [34]. Substance K receptors, neuropeptide Y1 receptors and β 2-adrenegric receptors are produced in *E. coli* at moderate levels [35–37], whereas a cannabinoid CB2 receptor is expressed much higher [31]. Extensive optimization of growth conditions and receptor modifications have contributed to enhanced production of a M2 muscarinic acetylcholine receptor [38]. A careful comparative analysis of many GPCRs, studied under standardized conditions, is needed to reveal parameters (such as charge distribution i.e. positive-inside rule, length and hydrophobicity of transmembrane helices etc.) that may allow predicting expression levels from a receptor sequence (see also reference [39] for thermodynamics and folding of membrane protein). In this respect, the Membrane Protein Network (MePNet, http://www.mepnet.org/), which has selected 100 GPCRs for study of expression, purification and structure determination, may be a valuable source for the statistical evaluation of the parameters described above.

Purification

Careful optimization of the detergent and the buffer composition is critical for achieving maximum solubilization efficiency and stability of functional receptors in solution [28, 40]. For NTR, a combination of the detergents 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and *n*-dodecyl-β-D-maltoside (LM), in the presence of cholesteryl hemisuccinate (CHS) and glycerol (30%), is most effective in maintaining receptor functionality. All steps must be performed at 4 °C [28].

Once suitable conditions for solubilization have been identified, a robust and effective purification scheme is required with the number of column steps kept to a minimum to preserve receptor integrity at maximum yield. Automation of sample loading and successive processing of chromatography columns is desirable to make this process most time-efficient [30, 40, 41]. To date, only a few GPCRs have been purified in milligram quantities [30, 32, 33, 40, 42–45] (using bacteria, the baculovirus system and mammalian cells as expression hosts).

The large-scale purification of the NTR fusion protein starts with the solubilization of receptors from whole *E. coli* cells rather than from crude membranes, which are time-consuming to prepare. The use of a deca-histidine tag (H10) (rather than six histidine residues) allows efficient enrichment of receptors from total cell lysate by immobilized metal affinity chromatography [46], because column loading and washing is possible in the presence of a high imidazole concentration (50 mM). In contrast to *E. coli* contaminants, the NTR fusion protein still binds to the affinity matrix under these conditions. Solubilized GPCRs are surrounded by a detergent belt, and it is important to realize that the bound detergent may mask the affinity tag if placed too close to a transmembrane helix, leading to reduced column binding with concomitant low receptor recovery. Batch loading, as opposed to column loading, increases the contact time of the tag with the affinity resin, resulting in better recovery of receptors with potentially not fully exposed affinity tags [32]; however, batch loading is not easily amenable to automation. The NTR fusion protein contains TrxA as a spacer between the receptor C-terminus and the H10 tail, rendering the affinity tag accessible to the column matrix, allowing column loading and hence automation. A subsequent neurotensin column selects for functional receptors from the metal affinity column eluate [28, 47]. The described two-column purification is robust and easy to scale up. Ten milligrams of pure, functional NTR fusion protein can be obtained from 250 grams of *E. coli* cell paste in an automated manner (see Figure 1) [30].

Generation of neurotensin receptor devoid of its fusion partners

The MBP and TrxA-H10 moieties at the receptor N- and C-termini may be flexible in respect to NTR. This flexibility may preclude the formation of high-quality three-dimensional crystals for structure determination. The presence of tobacco etch virus (Tev) recognition sites at either end of NTR allows removal of the fusion partners by Tev protease treatment. However, relatively high concentrations of Tev protease are needed to achieve complete cleavage at both receptor ends.

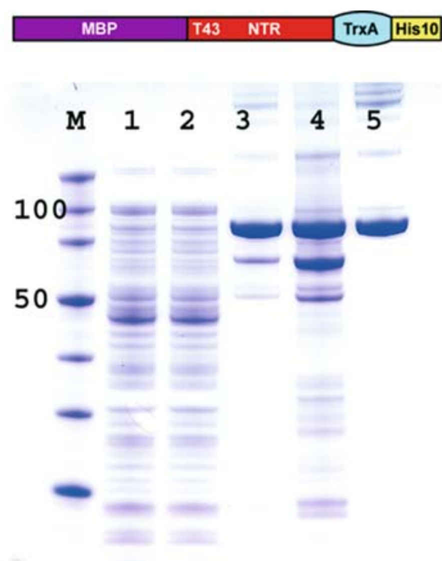

Figure 1. Automated [30] 10-mg preparation of a NTR fusion protein. The fusion protein was purified on a 100-ml Ni-NTA column followed by a 20-ml NT column, starting from 250 gram of *E. coli* cells. The progress of purification was monitored by SDS-PAGE (NuPAGE 4–12% Bis–Tris gel, Invitrogen, 1 × Mes buffer) and Coomassie R-250 staining. Lane M: Novagen Perfect Protein Marker (15–150 kDa); lane 1: 10 μg of supernatant; lane 2: 10 μg of Ni-NTA column flowthrough; lane 3: 5 μg of Ni-NTA column eluate; lane 4: 10 μg of NT column flowthrough; lane 5: 5 μg of NT column eluate. The fusion protein (NTS1-624, see [30]) is schematically shown at the top of the figure. Adapted from reference [30] with permission [50].

This may reflect steric hindrance of the protease to reach the engineered recognition sites, or suboptimal enzyme performance in buffer containing detergents (see reference [48]). The receptor can be separated from MBP, Tev protease and TrxA-H10 by gel filtration [30].

Conclusions

Functional NTR is produced in *E. coli* as a MBP fusion protein. Large-scale fermentation and an automated two-column purification scheme allow ten milligrams of purified, functional receptor fusion protein to be obtained on a weekly basis. Each step is kept simple and robust. Fusion partners can be removed by controlled protease treatment. Purified NTR devoid of its fusion partners has been used to generate receptor-specific aptamers [49], and to raise monoclonal antibodies (unpublished work). Three-dimensional crystallization experiments are ongoing. The backbone conformation of a neurotensin analog, NT(8–13), has been determined by solid-state NMR methods using lipid-reconstituted receptor fusion protein [8]. This study provided insight into the structure of a high-affinity agonist bound to its GPCR.

Acknowledgements

Our work on large-scale expression, automated purification and receptor structure determination is supported by the intramural program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH, DHHS.

References

1. Ji, T.H., Grossmann, M. and Ji, I. (1998) *J. Biol. Chem.* **273**, 17299–17302.
2. Gether, U. and Kobilka, B.K. (1998) *J. Biol. Chem.* **273**, 17979–17982.
3. Vassilatis, D.K. et al. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 4903–4908.
4. Baldwin, J.M., Schertler, G.F.X. and Unger, V.M. (1997) *J. Mol. Biol.* **272**, 144–164.
5. Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E.M. and Shichida, Y. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 5982–5987.
6. Palczewski, K. et al. (2000) *Science* **289**, 739–745.
7. Inooka, H. et al. (2001) *Nat. Struct. Biol.* **8**, 161–165.
8. Luca, S., White, J.F., Sohal, A.K., Filippov, D.V., van Boom, J.H., Grisshammer, R. and Baldus, M. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 10706–10711.
9. Williamson, P.T.F., Bains, S., Chung, C., Cooke, R. and Watts, A. (2002) *Febs Lett.* **518**, 111–115.
10. Reeves, P.J., Klein-Seetharaman, J., Getmanova, E.V., Eilers, M., Loewen, M.C., Smith, S.O. and Khorana, H.G. (1999) *Biochem. Soc. Trans.* **27**, 950–955.
11. Furukawa, H. et al. (2002) *Mol. Pharmacol.* **62**, 778–787.
12. Vincent, J.-P., Mazella, J. and Kitabgi, P. (1999) *Trends Pharmacol. Sci.* **20**, 302–309.
13. Carraway, R. and Leeman, S.E. (1973) *J. Biol. Chem.* **248**, 6854–6861.
14. Tanaka, K., Masu, M. and Nakanishi, S. (1990) *Neuron* **4**, 847–854.
15. Issue on Expression and Purification of Membrane Proteins (1999) *Biochem. Soc. Trans.* **27**, 883–960.
16. Special Issue on Overexpression of Integral Membrane Proteins (2003) *Biochim. Biophys. Acta* **1610**, 1–153.
17. Grisshammer, R. and Tate, C.G. (1995) *Quart. Rev. Biophys.* **28**, 315–422.
18. Reinhart, C., Weiss, H.M. and Reiländer, H. (2002) In A Practical Guide to Membrane Protein Purification (Eds., Hunte, C., von Jagow, G. and Schägger, H.), Academic Press, pp. 167–178.

19. Reiländer, H., Reinhart, C. and Szmolenszky, A. (1999) In *G-Protein-Coupled Receptors* (Eds., Haga, T., Berstein, G.), CRC Press, Boca Raton, pp. 281–322.
20. Lala, A., Sojar, H.T. and DeNardin, E. (1997) *Biochem. Pharmacol.* **54**, 381–390.
21. Banères, J.-L., Martin, A., Hullot, P., Girard, J.-P., Rossi, J.-C. and Parello, J. (2003) *J. Mol. Biol.* **329**, 801–814.
22. Kiefer, H., Vogel, R. and Maier, K. (2000) Receptors & Channels **7**, 109–119.
23. Kiefer, H., Maier, K. and Vogel, R. (1999) *Biochem. Soc. Trans.* **27**, 908–912.
24. Grisshammer, R., Duckworth, R. and Henderson, R. (1993) *Biochem. J.* **295**, 571–576.
25. Bertin, B., Freissmuth, M., Breyer, R.M., Schütz, W., Strosberg, A.D. and Marullo, S. (1992) *J. Biol. Chem.* **267**, 8200–8206.
26. von Heijne, G. (1989) *Nature* **341**, 456–458.
27. Drew, D., Fröderberg, L., Baars, L. and de Gier, J.-W.L. (2003) *Biochim. Biophys. Acta* **1610**, 3–10.
28. Tucker, J. and Grisshammer, R. (1996) *Biochem. J.* **317**, 891–899.
29. Grisshammer, R. and Hermans, E. (2001) *Febs Lett.* **493**, 101–105.
30. White, J.F., Trinh, L.B., Shiloach, J. and Grisshammer, R. (2004) *Febs Lett.* **564**, 289–293.
31. Calandra, B., Tucker, J., Shire, D. and Grisshammer, R. (1997) *Biotechnol. Lett.* **19**, 425–428.
32. Weiß, H.M. and Grisshammer, R. (2002) *Eur. J. Biochem.* **269**, 82–92.
33. Hulme, E.C. and Curtis, C.A.M. (1998) *Biochem. Soc. Trans.* **26**, S361.
34. Stanasila, L., Massotte, D., Kieffer, B.L. and Pattus, F. (1999) *Eur. J. Biochem.* **260**, 430–438.
35. Münch, G., Walker, P., Shine, J. and Herzog, H. (1995) *Receptors Channels* **3**, 291–297.
36. Hampe, W., Voss, R.-H., Haase, W., Boege, F., Michel, H. and Reiländer, H. (2000) *J. Biotechnol.* **77**, 219–234.
37. Grisshammer, R., Little, J. and Aharony, D. (1994) *Receptors Channels* **2**, 295–302.
38. Furukawa, H. and Haga, T. (2000) *J. Biochem. (Tokyo)* **127**, 151–161.
39. White, S.H. (2003) *Febs Lett.* **555**, 116–121.
40. Warne, T., Chirnside, J. and Schertler, G.F.X. (2003) *Biochim. Biophys. Acta* **1610**, 133–140.
41. Laible, P.D., Scott, H.N., Henry, L. and Hanson, D.K. (2004) *J. Struct. Funct. Genomics* **5**, 167–172.
42. Reeves, P.J., Kim, J.-M. and Khorana, H.G. (2002) *Proc. Natl. Acad. Sci.* **99**, 13413–13418.
43. Ohtaki, T. et al. (1998) *J. Biol. Chem.* **273**, 15464–15473.
44. Klaassen, C.H.W., Bovee-Geurts, P.H.M., DeCaluwé, G.L.J. and DeGrip, W.J. (1999) *Biochem. J.* **342**, 293–300.
45. Ratnala, V.R.P., Swarts, H.G.P., VanOostrum, J., Leurs, R., DeGroot, H.J.M., Bakker, R.A. and DeGrip, W.J. (2004) *Eur. J. Biochem.* **271**, 2636–2646.
46. Grisshammer, R. and Tucker, J. (1997) *Protein Expres. Purif.* **11**, 53–60.
47. Grisshammer, R., Averbeck, P. and Sohal, A.K. (1999) *Biochem. Soc. Trans.* **27**, 899–903.
48. Mohanty, A.K., Simmons, C.R. and Wiener, M.C. (2003) *Protein Expres. Purif.* **27**, 109–114.
49. Daniels, D.A., Sohal, A.K., Rees, S. and Grisshammer, R. (2002) *Anal. Biochem.* **305**, 214–226.
50. Reprinted by permission of Federation of the European Biochemical Societies from ‘Automated large-scale purification of a G-protein-coupled receptor for neurotensin’, by White *et al.*, *Febs Lett.* **564**, 289–293, 2004.