

Membrane topology of the *Drosophila* OR83b odorant receptor

Carolina Lundin^a, Lukas Käll^b, Scott A. Kreher^c, Katja Kapp^d, Erik L. Sonnhhammer^b,
John R. Carlson^c, Gunnar von Heijne^{a,b}, IngMarie Nilsson^{a,*}

^a Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, Svante Arrheniusv. 12, SE-106 91 Stockholm, Sweden

^b Stockholm Bioinformatics Center, AlbaNova, SE-106 91 Stockholm, Sweden

^c Department of Molecular, Cellular, and Development Biology, Yale University, New Haven, CT 06520, USA

^d ZMBH (Zentrum für Molekulare Biologie Heidelberg), University Heidelberg, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany

Received 10 October 2007; revised 2 November 2007; accepted 2 November 2007

Available online 20 November 2007

Edited by Maurice Montal

Abstract By analogy to mammals, odorant receptors (ORs) in insects, such as *Drosophila melanogaster*, have long been thought to belong to the G-protein coupled receptor (GPCR) superfamily. However, recent work has cast doubt on this assumption and has tentatively suggested an inverted topology compared to the canonical $N_{\text{out}} - C_{\text{in}}$ 7 transmembrane (TM) GPCR topology, at least for some *Drosophila* ORs. Here, we report a detailed topology mapping of the *Drosophila* OR83b receptor using engineered glycosylation sites as topology markers. Our results are inconsistent with a classical GPCR topology and show that OR83b has an intracellular N-terminus, an extracellular C-terminus, and 7TM helices.

© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Odorant receptor; Membrane protein; Topology; *Drosophila melanogaster*

1. Introduction

In mammals, odorant receptors (ORs) belong to the large superfamily of G-protein coupled receptors (GPCRs) and have a typical 7 transmembrane (TM) topology with an extracellular (EC) N-terminus and an intracellular (IC) C-terminus [1]. Because insects also have an expanded repertoire of GPCRs it has long been assumed that their ORs have the same canonical 7TM topology, yet recent work on the *Drosophila* OR83b OR tentatively placed its N-terminus intracellularly rather than extracellularly [2,3]. This unexpected result is in accordance with theoretical topology predictions [1]. OR83b is a ubiquitously expressed and highly conserved member of the insect OR family and heteromerizes with other ORs, forming active receptor complexes [2].

Here, we report a detailed study of the membrane topology of OR83b inserted into *Drosophila* rough microsomes (DRMs), using both an endogenous and engineered acceptor

sites for N-linked glycosylation as topological markers. Our results support a 7TM $N_{\text{in}} - C_{\text{out}}$ topology for OR83b.

2. Materials and methods

2.1. Enzymes and chemicals

Unless otherwise stated, all enzymes, plasmid pGEM1, and the TNT[®] Quick transcription/translation system were from Promega (Madison, WI). [³⁵S]-Methionine, [¹⁴C]-methylated marker proteins and deoxynucleotides were from GE Healthcare (Uppsala, Sweden). BigDye Terminator v1.1 Cycle Sequencing Kit was from AB Applied Biosystems (Foster City, CA) and oligonucleotides were from CyberGene AB (Stockholm, Sweden).

2.2. Plasmid construction

Fragments from full-length *Or83b* cDNA prepared *Drosophila* heads was modified in two ways during PCR amplification: (i) by the introduction of a 5' XbaI site, and (ii) by changing the context of the region immediately upstream of the initiator ATG codon to a Kozak consensus ribosome binding sequence, GCCACCATGG [4]; both changes were encoded within the 5'PCR primer. The reverse primer encoded the 3'-end of the selected OR gene, two stop codons, and a SmaI site for cloning. The *Or83b* gene was amplified by PCR using the Expand High Fidelity PCR system from Roche Diagnostics GmbH (Mannheim, FRG) and cloned into pGEM1 downstream of the SP6 promoter as an XbaI–SmaI fragment. The amplified DNA products were purified using the QIAquick PCR Purification kit from QIAGEN (Hilden, FRG).

2.3. DNA manipulations

Glycosylation acceptor sites were designed as described previously [5], i.e. by replacing or insertion of one or more appropriately positioned codons for the acceptor tripeptide Asn-Ser-Thr (NST). To destroy the endogenous glycosylation acceptor site (Asn¹⁶⁹-Ser-Ser (N169)), it was mutated to Gln-Ser-Ser (QSS). To create glycosylation acceptor sites the sequence was changed to N²⁰NST¹²¹ (N21), V¹¹⁴NSTH¹¹⁴ (N112), E¹¹⁹NSTD¹²¹ (N120), V¹⁷⁴NSTE¹⁷⁵ (N175), A²⁶⁴NSTK²⁶⁷ (N265), and K⁴³⁹NSTF⁴⁴¹ (N440). To introduce the C-terminal glycosylation acceptor site, the C-terminal end of OR83b was extended with the sequence K⁴⁸⁶PQSIYQKTMSFDK-LIENSTQKT (C-term NST).

Site-specific mutagenesis was performed using the QuickChange[™] Site-Directed Mutagenesis kit from Stratagene (La Jolla, USA). All mutants were confirmed by sequencing of plasmid DNA at BM labbet AB (Furulund, Sweden). All cloning steps were done according to standard procedures using restriction enzymes from Promega (Madison, USA).

2.4. Preparation of *Drosophila* rough microsomes

Drosophila S2 cells (ATCC: CRL-1963) from cultures growing in logarithmic phase were washed twice with PBS and once with buffer H (50 mM HEPES-KOH pH 7.4, 165 mM KOAc, 2 mM Mg(OAc)₂) with centrifugation steps for 3 min at 200 × g, 20 °C. Cells were

*Corresponding author. Fax: +46 8 153679.

E-mail address: ingmarie@dbb.su.se (I. Nilsson).

Abbreviations: OR, odorant receptor; GPCR, G-protein coupled receptor; TM, transmembrane; DRM, *Drosophila* rough microsome; CRM, column-washed dog pancreas rough microsomes; IC, intracellular; EC, extracellular; ER, endoplasmic reticulum

resuspended in 3 volumes buffer H containing 0.01% saponin (Sigma) and incubated 10 min at 20 °C. Buffer H was added to decrease the saponin concentration to 0.002%, and the DRMs were pelleted for 1 min at 2500 g, 20 °C. DRMs were adjusted to 20 A₂₈₀/ml with RM buffer (50 mM HEPES-KOH pH 7.6, 50 mM KOAc, 2 mM Mg(OAc)₂, 250 mM sucrose, 1 mM DTT), CaCl₂ and PMSF were added to a final concentration of 1 mM and 0.2 mg/ml, respectively. DRMs were incubated with 150 U/ml micrococcal nuclease (Nuclease S7 from *Staphylococcus aureus*, Roche) for 10 min at 25 °C, and the reaction was stopped by the addition of EGTA to a final concentration of 2 mM. DRMs were layered on a sucrose cushion (50 mM HEPES-KOH pH 7.6, 50 mM KOAc, 2 mM Mg(OAc)₂, 500 mM sucrose, 1 mM DTT) and separated by centrifugation for 30 min at 35,000 × g, 4 °C. The pellet was resuspended in RM buffer and DRMs adjusted to 100 A₂₈₀/ml (i.e. 2 equiv./μl).

2.5. Expression in vitro

Constructs in pGEM1 were transcribed and translated in the TnT Quick systems from Promega. 1 μg DNA template, 1 μl [³⁵S]-Met (15 μCi) and 1 equiv. of DRM or 2 equiv. of dog pancreas rough microsomes (CRM) [6] were added at the start of the reaction, and samples were incubated for 90 min at 30 °C [7]. Samples were analyzed by SDS-PAGE, and proteins were visualized in a Fuji FLA-3000 phosphorimager using the Image Reader V1.8J/Image Gauge V 3.45 software.

3. Results

3.1. Prediction models and topology assay

The Phobius [8], TMHMM [9], HMMTOP [10], Memsat 2.0 [11] and Toppred [12] algorithms all predict the same 7TM

N_{in} – C_{out} topology for OR83b, Fig. 1A. In an attempt to experimentally map the topology of OR83b, we took advantage of a potential acceptor site for N-linked glycosylation (N¹⁶⁹SS) present in the second predicted EC loop (EC2). N-linked glycosylation is a reliable topology marker, as the endoplasmic reticulum (ER)-resident oligosaccharide transferase enzyme can only transfer glycans to lumenally exposed parts of membrane proteins inserted into the ER [13]. OR83b was cloned behind the SP6 promoter in the pGEM1 vector and transcribed and translated in vitro in the absence or presence of DRMs. In the presence of DRMs, a more slowly migrating form of the protein was observed (Fig. 1B, compare lanes 1 and 2), which disappeared when the glycosylation site was mutated from N¹⁶⁹SS (N169) to Q¹⁶⁹SS (QSS) (lanes 3 and 4). We conclude that the EC2 loop is located in the lumen of the microsomes. This result also shows that the other two potential glycosylation sites (N³³FT (N33) and N¹⁸⁸AS (N188)) in OR83b are not utilized, possibly because they are located in a cytoplasmic part of the protein (N³³FT) or too close to a TM segment (N¹⁸⁸AS) [14].

3.2. Topology mapping Or83b

Additional acceptor sites for N-linked glycosylation were engineered into the N-terminal tail of OR83b, as well as into loops IC1, IC2, IC3, and EC2. A 22-residue peptide (PQSIYQKTMSFDKLIENSTQKT) containing a glycosylation acceptor site (underlined) was also fused to the C-terminus. The EC1 and EC3 loops were not targeted, as they are

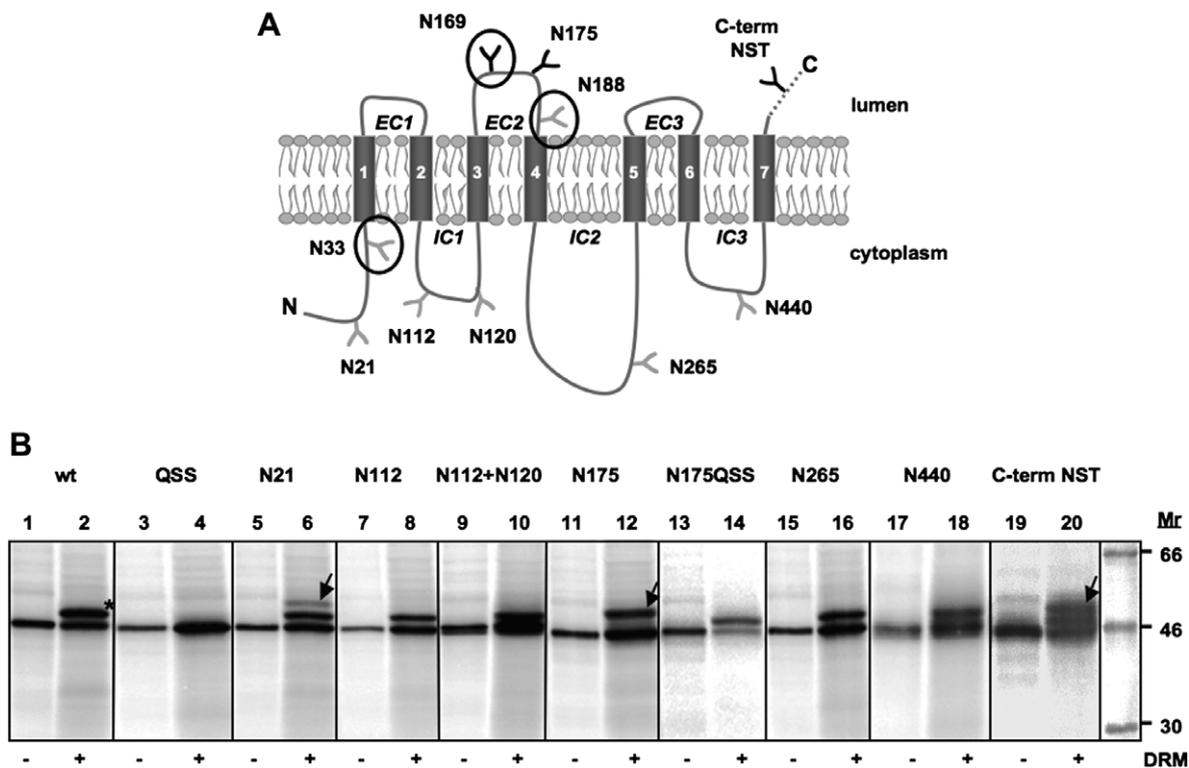


Fig. 1. The topology of the Or83b protein. (A) Predicted topology of OR83b. Phobius [8] predicts an N_{in} – C_{out} topology with the following 7TM helices: 50–68, 80–98, 137–158, 194–213, 350–371, 391–413, 462–483. Engineered and endogenous (N³³FT, N¹⁶⁹SS and N¹⁸⁸AS encircled) putative glycosylation sites are indicated; sites that become modified upon insertion into DRMs are in black, non-modified sites are in gray. (B) Glycosylation mapping of OR83b. In vitro translation of OR83b wild-type and variants with engineered glycosylation sites in the absence (–DRM) and presence (+DRM) of *Drosophila* rough microsomes. Molecules glycosylated only on the endogenous N¹⁶⁹SS site are indicated by * and molecules glycosylated on an additional engineered site are indicated by arrows.

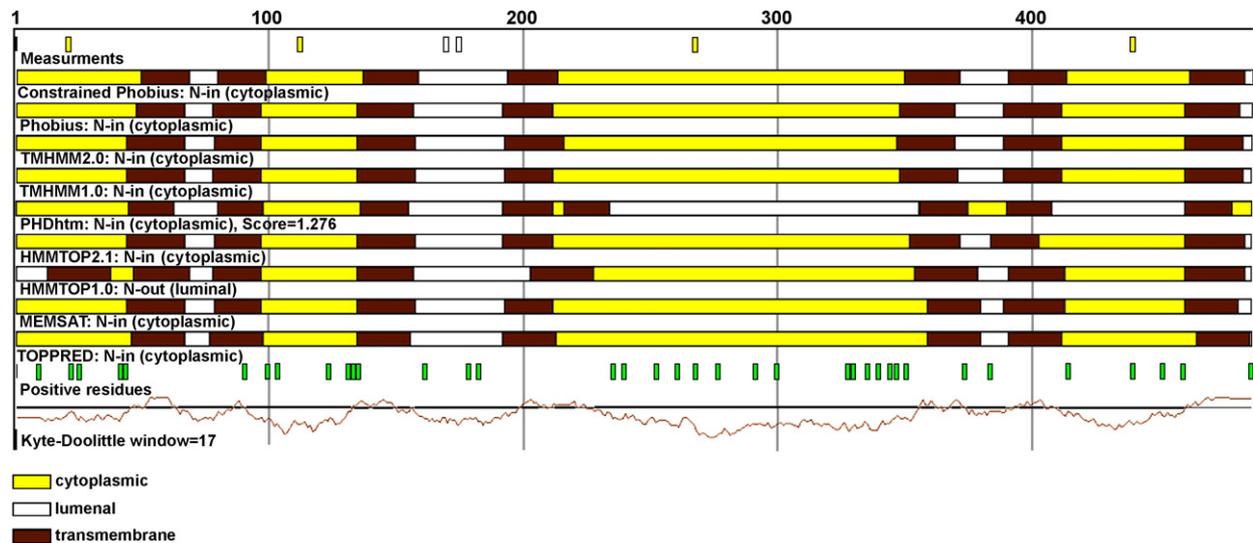


Fig. 2. Predictions of OR83b transmembrane topology, adapted from the Sfnx metaserver (<http://sfnx.cgb.ki.se>; [19]). A Phobius prediction obtained by using the full set of glycosylation data as constraints is shown on top, below are the default predictions of the Sfnx server: unconstrained predictions from Phobius [8], TMHMM (1.0 and 2.0) [9], HMMTOP (2.1 and 1.0) [10], PHDhtm [20], Memsat 2.0 [11], and Toppred [12]. OR83b is predicted to have 7TM helices with the N-terminus in the cytoplasm by eight of the nine methods. Predicted transmembrane segments (brown segments), cytoplasmic loops (yellow segments), and luminal loops (white segments) are indicated above the respective prediction method.

too short to be modified by the oligosaccharyl transferase [14,15]. The engineered glycosylation site (N175) in loop EC2 was efficiently modified in the presence of DRMs, resulting in a protein carrying two N-linked glycans (Fig. 1B, lanes 11 and 12). When the natural glycosylation site N¹⁶⁹SS was mutated to N¹⁶⁹QSS in this construct (N175QSS), only the N175 site was modified (lanes 13 and 14). The extended C-terminal tail (C-term NST) was also modified in the presence of DRMs (lanes 19 and 20); the somewhat lower modification efficiency of this site is probably caused by its location very close to the C-terminus of the protein [16]. In contrast, none of the acceptor sites in loops IC1 (N112), IC2 (N265), or IC3 (N440) were modified. The engineered site (N21) in the N-terminal tail was mostly non-glycosylated, although a faint doubly glycosylated product (modified on both N²¹ and N¹⁶⁹) was also seen for this construct (lanes 5 and 6). Since no faint doubly glycosylated bands were seen for the N112 and N112+N120 constructs (lanes 7–10), loop IC1 faces the cytosol in all molecules. Molecules glycosylated on N²¹ thus represent a minor, probably misfolded, fraction of the protein in which the first or second TM segment does not span the membrane. Earlier work using GFP and YFP fusions and epitope staining is consistent with a cytosolic location of the N-terminus and a luminal location of the EC2 loop [2].

As a control, the constructs were also translated in the presence of mammalian (dog pancreas) RMs (Supplementary Fig. S1). The results were the same as those obtained with the *Drosophila* RMs, except that the glycosylated bands were hard to resolve for the construct with an extended C-terminal tail.

4. Discussion

The *Drosophila* OR83b protein is an ubiquitously expressed member of the insect OR family, and it forms functional heteromers with other OR proteins [2]. Mammalian ORs are 7TM GPCRs with an EC N terminus, but there is no detect-

able sequence similarity between mammalian and insect ORs [2]. Recent work has tentatively located the N terminus of OR83b to the cytosol [2], arguing that this protein is not a classical GPCR. Given this rather surprising conclusion, we decided to perform a detailed study on the topology of OR83b using glycosylation mapping, an approach that has been widely applied to eukaryotic membrane proteins [17].

In short, our results confirm the suggested IC location of the N-terminal tail of OR83b and in addition show that the EC2 loop and the C-terminal tail are EC. We also find that glycosylation acceptor sites engineered into loops IC1, IC2, and IC3 are not modified by the lumenally disposed oligosaccharyl transferase, in accordance with the proposed 7TM $N_{in} - C_{out}$ topology [2].

When the available experimental data is used to constrain [18] the Phobius predictor, the predicted topology for OR83b is as shown in Fig. 2, i.e., essentially the same as predicted by the unconstrained Phobius, TMHMM, HMMTOP, Memsat 2.0, and Toppred algorithms (cf., Fig. 1). We conclude that OR83b has 7TM helices, an IC N-terminus and EC C-terminus, and is thus inverted compared to the canonical $N_{in} - C_{out}$ 7TM topology of the GPCR family of mammalian ORs.

Acknowledgements: We thank Prof. Bernard Dobberstein for advice concerning *Drosophila* microsomes and Prof. Arthur E. Johnson for providing dog pancreas microsomes. This work was supported by grants from the Swedish Cancer Foundation to IN and GvH, from the Swedish Research Council to GvH, from the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) to IN, from Magnus Bergvalls Stiftelse, Henrik Granholms Stiftelse, and Carl Tryggers Stiftelse to IN, and from the NIH to JC.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.febslet.2007.11.007](https://doi.org/10.1016/j.febslet.2007.11.007).

References

- [1] Wistrand, M., Käll, L. and Sonnhammer, E.L. (2006) A general model of G-protein coupled receptor sequences and its application to detect remote homologs. *Protein Sci.* 15, 509–521.
- [2] Benton, R., Sachse, S., Michnick, S.W. and Vosshall, L.B. (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol.* 4, e20.
- [3] Vosshall, L.B. and Stocker, R.F. (2007) Molecular architecture of smell and taste in *Drosophila*. *Annu. Rev. Neurosci.* 30, 505–533.
- [4] Kozak, M. (1989) Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell Biol.* 9, 5073–5080.
- [5] Nilsson, I., Whitley, P. and von Heijne, G. (1994) The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J. Cell Biol.* 126, 1127–1132.
- [6] Walter, P. and Blobel, G. (1983) Preparation of microsomal membranes for cotranslational protein translocation. *Method Enzymol.* 96, 84–93.
- [7] Hessa, T. et al. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433, 377–381.
- [8] Käll, L., Krogh, A. and Sonnhammer, E.L. (2004) A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* 338, 1027–1036.
- [9] Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.
- [10] Tusnady, G.E. and Simon, I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17, 849–850.
- [11] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* 33, 3038–3049.
- [12] von Heijne, G. (1992) Membrane protein structure prediction – hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* 225, 487–494.
- [13] Kaplan, H.A., Welply, J.K. and Lennarz, W.J. (1987) Oligosaccharyltransferase: the central enzyme in the pathway of glycoprotein assembly. *Biochim. Biophys. Acta* 906, 161–173.
- [14] Popov, M., Tam, L.Y., Li, J. and Reithmeier, R.A. (1997) Mapping the ends of transmembrane segments in a polytopic membrane protein. Scanning N-glycosylation mutagenesis of extracytosolic loops in the anion exchanger, band 3. *J. Biol. Chem.* 272, 18325–18332.
- [15] Nilsson, I. and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* 268, 5798–5801.
- [16] Nilsson, I. and von Heijne, G. (2000) Glycosylation efficiency of Asn-Xaa-Thr sequons depends both on the distance from the C-terminus and on the presence of a downstream transmembrane segment. *J. Biol. Chem.* 275, 17338–17343.
- [17] von Heijne, G. (2006) Membrane-protein topology. *Nat. Rev. Mol. Cell Biol.* 7, 909–918.
- [18] Melen, K., Krogh, A. and von Heijne, G. (2003) Reliability measures for membrane protein topology prediction algorithms. *J. Mol. Biol.* 327, 735–744.
- [19] Sonnhammer, E.L. and Wootton, J.C. (2001) Integrated graphical analysis of protein sequence features predicted from sequence composition. *Proteins* 45, 262–273.
- [20] Rost, B., Fariselli, P. and Casadio, R. (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci.* 5, 1704–1718.