LARGE-SCALE PRODUCTION AND CHARACTERIZATION OF AN ENGINEERED HUMAN OLFACTORY RECEPTOR

Brian Lee Cook
B.S. Biological Engineering Sciences
Washington University in Saint Louis, 1999

Submitted to the Department of Biological Engineering in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biological Engineering
at the
Massachusetts Institute of Technology

February 2009

©2008 Brian L. Cook. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: ____________________________

Brian L. Cook
Department of Biological Engineering
October 6th, 2008

Certified by: ____________________________

Shuguang Zhang
Associate Director, Center for Biomedical Engineering
Thesis Supervisor

Certified by: ____________________________

Angela M. Belcher
Professor of Materials Science and Engineering and Biological Engineering
Co-Advisor, Thesis Supervisor

Accepted by: ____________________________

Alan J. Grodzinsky
Professor of Mechanical, Electrical, and Biological Engineering
Graduate Program Chairman
Thesis Committee:

Shuguang Zhang
Associate Director, Center for Biomedical Engineering
Massachusetts Institute of Technology

Angela M. Belcher
Germehausen Professor of Materials Science and Engineering and Biological Engineering
Massachusetts Institute of Technology

Horst Vogel
Professor of Physical Chemistry
Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Harvey F. Lodish
Professor of Biology and Biological Engineering
Member, Whitehead Institute
Massachusetts Institute of Technology
Thesis Chair
LARGE-SCALE PRODUCTION AND CHARACTERIZATION OF A SYNTHETIC HUMAN OLFACTORY RECEPTOR

by Brian L. Cook

Submitted to the Department of Biological Engineering on October 6th, 2008 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biological Engineering

ABSTRACT

Animal noses have evolved the ability to rapidly detect a seemingly infinite array of odors at minute concentrations. The basis of this sensitivity are the olfactory (smell) receptors – a large, highly related class of sensory G-protein coupled receptors that function together combinatorially to allow discrimination between a wide range of volatile and soluble molecules. However, the structural and functional mechanisms of these amazing receptors are not currently known.

In order to begin to investigate the molecular mechanism(s) of olfaction, I have developed a mammalian expression system for the large-scale production and purification of functional olfactory receptor (OR) proteins in milligram quantities. Expressed OR genes were fabricated from scratch using PCR-based gene synthesis, which facilitated codon optimization and attachment of different affinity tags for purification. Established methods for the production and purification of rhodopsin were adapted to olfactory receptors through extensive optimization (including a full-spectrum screening of over 45 detergents). Key to the efficient extraction and solubilization of olfactory receptors tested is the use of novel zwitter-ionic fos-choline detergents.

Following initial experiments on the inducible expression of a human olfactory receptor (hOR17-4) in adherent HEK293S cell cultures, the system was successfully scaled up using a suspension bioreactor. Large-scale culture allowed the purification of >10 milligrams of hOR17-4 monomer at >90%, which was suitable for subsequent X-ray crystallization screening trials. The purified protein was also characterized using several spectroscopic methods and shown to possess the correct secondary structure and several predicted post-translational modifications. To assay the functionality of purified (non-membrane-bound) hOR17-4, we successfully developed an in vitro assay method using surface plasmon resonance (SPR) to demonstrate that the receptor retains functional selectivity in binding specific odorant ligands in a concentration-dependent manner.

The application of these techniques to other olfactory receptors already shows promise and could lead to a generalized method for obtaining large quantities of any olfactory receptor in a rapid and simple manner. Such methods could prove extremely useful in elucidating the structural and functional mechanism(s) of olfactory receptors and in their integration into OR-based biosensor devices.

• Thesis Supervisor (Advisor): Shuguang Zhang
  Title: Associate Director, MIT Center for Biomedical Engineering
• Thesis Supervisor (Co-Advisor): Angela M. Belcher
  Title: Professor of Materials Science and Engineering and Biological Engineering
ACKNOWLEDGEMENTS

First and foremost, I want to thank my advisor Shuguang Zhang for his wonderful mentorship throughout my experience at MIT. His scientific verve and enthusiasm were a constant inspiration, and I owe him so much for his confidence in me and willingness to pursue what was initially a somewhat risky thesis project, which after much hard work ended up succeeding on many levels. My thesis committee, Angela Belcher, Harvey Lodish, and Horst Vogel, have provided vital guidance and suggestions that significantly shaped my work. I am extremely grateful to have had the privilege of working with the excellent scientists in the Zhang lab over the years, especially Karin Ernberg, Dirk Steuerwald, and Melanie Vanberghem. I also wish to recognize my wonderful UROP students Hyeyoun Chung and Allison Berke, who have both moved on to graduate school themselves. My officemate Peter Carr was a most valuable resource for information on gene assembly, DNA techniques, and just a great scientist to talk to. I also wish to acknowledge former lab members Liselotte Kaiser and Joanne Graveland-Bikker, who together with Kara Herlihy at Biacore™ were instrumental in developing the SPR-based assay by which to measure solubilized OR functionality. A special thank you is in order to my parents, my family, and all my friends in Course 20 who gave me support and friendship throughout these five years, especially David Appleyard, Hyung-Do Kim, and, most of all, Andreas Mershin, a lab colleague, comrade and fellow adventurer. I gratefully acknowledge the MIT laboratory of H. Gobind Khorana, especially Philip J. Reeves and Prashen Chelikani for their instruction regarding rho1D4 purification and for providing the parental HEK293S cell lines. I also wish to thank Horst Pick (EPFL) and Ioannis Papayannopoulos (MIT Proteomics Facility) for their assistance with calcium-release assays and mass spectrometry analysis, respectively. I am extremely grateful to the MIT Center for Bits & Atoms for the fellowship that funded me throughout graduate school. This work was also supported in part from a generous grant from ROHM Corporation, Kyoto, Japan. Lastly, I’d like to thank Luca Turin for his valuable collaborations as well as for being the subject of the book “The Emperor of Scent”, which got me interested in smell in the first place.
# TABLE OF CONTENTS

List of Figures ............................................................................................................. 7

List of Tables ............................................................................................................... 8

Table of Abbreviations and Terms ........................................................................... 9

Chapter 1: Introduction and Background ................................................................... 10
  1.1 Olfactory receptors ......................................................................................... 10
  1.2 Detergent solubilization ............................................................................... 16
  1.3 Inducible cell lines ......................................................................................... 16
  1.4 Bioreactor culture .......................................................................................... 19
  1.5 Circular dichroism spectroscopy ..................................................................... 21
  1.6 Surface plasmon resonance .......................................................................... 23
  1.7 References ..................................................................................................... 26

Chapter 2: Fabrication of Custom Olfactory Receptor Genes ................................... 32
  2.1 Overview of PCR-based gene synthesis ....................................................... 32
  2.2 Selection of OR genes for construction ....................................................... 35
  2.3 Fabrication of the hOR17-4-rho synthetic gene ........................................... 39
  2.4 References ..................................................................................................... 42

Chapter 3: Production of hOR17-4 in Adherent Cell Cultures ................................. 45
  3.1 Abstract ......................................................................................................... 46
  3.2 Introduction .................................................................................................... 47
  3.3 Results .......................................................................................................... 50
  3.4 Discussion ..................................................................................................... 59
  3.5 Methods ........................................................................................................ 64
  3.6 References ..................................................................................................... 70

Chapter 4: Production of hOR17-4 in Suspension Cell Cultures ............................... 74
  4.1 Abstract ......................................................................................................... 75
  4.2 Introduction .................................................................................................... 75
  4.3 Results .......................................................................................................... 77
  4.4 Discussion ..................................................................................................... 84
  4.5 Methods ........................................................................................................ 87
  4.6 References ..................................................................................................... 93
TABLE OF CONTENTS (CONTINUED)

Chapter 5: Characterization of purified hOR17-4 ................................................. 96
  5.1 Abstract ........................................................................................................... 97
  5.2 Introduction .................................................................................................. 97
  5.3 Results ......................................................................................................... 98
    5.3.1 Circular dichroism and fluorescence spectroscopy .................................. 98
    5.3.2 Calcium- release assays of cell surface expressed hOR17-4 .................... 100
    5.3.3 SPR-based functional assay .................................................................... 102
    5.3.4 Quantitation of hOR17-4 activity ............................................................ 105
    5.3.5 Post-translational modifications ............................................................... 108
  5.4 Discussion ...................................................................................................... 109
  5.5 Methods ....................................................................................................... 113
  5.6 References .................................................................................................... 117

Appendix A: Additional Studies ............................................................................. 120
  A.1 Crystal Screening .......................................................................................... 120
  A.2 Additional Cell Lines Constructed ................................................................ 122
  A.3 Detection Using Custom Antibodies ............................................................. 127

Appendix B: Experimental Protocols
  B.1 Gene Design Using DNAWorks .................................................................... 129
  B.2 PCR-based Gene Assembly (PCA) ................................................................ 137
  B.3 Stable HEK293S Colony Subcloning and Screening ..................................... 139
  B.4 Immunoblotting Protocol ............................................................................. 141
  B.5 Bioreactor Media Preparation ....................................................................... 144
  B.6 Bioreactor Operation .................................................................................... 145
  B.7 Rho1D4-bead Coupling ............................................................................... 147
  B.8 Immunoaffinity Purification ......................................................................... 149

Appendix C: Sequence info for all OR genes constructed ..................................... 150
LIST OF FIGURES

Figure 1.1 – Cartoon of olfactory receptor showing the seven transmembrane domains and loop regions .............................................................. 11

Figure 1.2 – Diagram of olfactory receptor downstream signaling in olfactory neurons. 12

Figure 1.3 – Neuronal wiring of olfactory system .............................................................. 13

Figure 1.4 – Detergent solubilization ........................................................................... 17

Figure 1.5 – Diagram of the T-REx tetracycline-inducible system ............................... 18

Figure 1.6 – Bioreactor operation ............................................................................... 20

Figure 1.7 – Far-UV circular dichroism spectra characteristic of the major protein secondary structures ............................................................. 22

Figure 1.8 – Overview of surface plasmon resonance (SPR) ........................................ 25

Figure 2.1 – Overview of PCR-based Gene Assembly .................................................. 34

Figure 2.2 – Olfactory receptor phylogenetic tree ...................................................... 35

Figure 2.3 – Partial odorant ligand spectrum for hOR17-4 ........................................... 36

Figure 2.4 – Fabrication of custom hOR17-4-rho gene .............................................. 41

Figure 3.1 – Codon-optimized hOR17-4 sequence .................................................... 51

Figure 3.2 – Induction of hOR17-4 in stable HEK293S cell lines ................................. 53

Figure 3.3 – High-level induction correlates with a band shift of hOR17-4 .............. 55

Figure 3.4 – Immunoaffinity purification of hOR17-4 ................................................ 57

Figure 3.5 – Full purification of hOR17-4 extracted from 50 x 150 mm culture plates. 58

Figure 4.1 – Construction of hOR17-4-inducible HEK293S GnT1/- cell lines for use in liquid bioreactor culture ......................................................... 78

Figure 4.2 – Detergent screen for optimal solubilization of hOR17-4 expressed in HEK293S cells ........................................................................ 80

Figure 4.3 – Full purification of hOR17-4 from 2.65 liters of bioreactor cultured cells. 83
LIST OF FIGURES (CONTINUED)

Figure 5.1 – Circular dichroism analysis of purified hOR17-4........................................... 99
Figure 5.2 – Fluorescence spectroscopy on purified hOR17-4 monomer .................. 100
Figure 5.3 – Calcium- release assays of cell surface-expressed hOR17-4 ................. 101
Figure 5.4 – Preparation of SPR chip surface for odorant binding assay ............... 103
Figure 5.5 – SPR determination of hOR17-4 affinity for bourgeonal...................... 104
Figure 5.6 – Functional assay of purified hOR17-4 functionality using surface plasmon resonance (SPR) ........................................................................ 105
Figure 5.7 – Depiction of surfactant-like peptide solubilization and stabilization of membrane proteins ................................................................. 112
Figure 5.8 – Molecular model of surfactant-like peptides ....................................... 113
Figure A.1 – Hanging drop crystallization of purified monomeric hOR17-4 ............ 121
Figure A.2 – Analysis of His-FLAG-tagged hOR17-4............................................... 124
Figure A.3 – Analysis of untagged hOR17-4 expression ....................................... 126
Figure A.4 – Detection of tagged hOR17-4 variants using custom polyclonal antibodies ........................................................................................................... 128

LIST OF TABLES

Table 1 – Summary of previous OR expression methods.............................................. 15
Table 2 – Primers used for the PCR-based synthesis of the engineered hOR17-4-rho gene.............................................................................................................. 40
Table 3 – Quantitation of purified hOR17-4 receptor functionality............................ 107
Table 4 – Analysis of purified hOR17-4 by mass spectrometry .................................. 109
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CMC</td>
<td>Critical micelle concentration</td>
</tr>
<tr>
<td>DDM</td>
<td>Dodecyl maltoside</td>
</tr>
<tr>
<td>FC14</td>
<td>Fos-choline-14 (a.k.a., n-Tetradecylphosphocholine)</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylglucosamine</td>
</tr>
<tr>
<td>GPCR</td>
<td>G protein-coupled receptor</td>
</tr>
<tr>
<td>HEK</td>
<td>Human embryonic kidney</td>
</tr>
<tr>
<td>HEK293S</td>
<td>Suspension-adapted HEK293S cell line</td>
</tr>
<tr>
<td>HEK293S GnTI</td>
<td>GlcNAc transferase I-deficient HEK293S cell line</td>
</tr>
<tr>
<td>hOR17-4</td>
<td>Human olfactory receptor 17-4 (a.k.a., human OR1D2)</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>NaBu</td>
<td>Sodium butyrate</td>
</tr>
<tr>
<td>OR</td>
<td>Olfactory receptor</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline solution</td>
</tr>
<tr>
<td>PCA</td>
<td>Polymerase construction and amplification</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>rho1D4</td>
<td>Mouse monoclonal antibody against rho-tag</td>
</tr>
<tr>
<td>rho-tag</td>
<td>Rho peptide tag (TETSQVAPA); epitope for rho1D4</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Olfactory Receptors

Introduction

The mammalian nose has evolved the ability to rapidly detect small airborne molecules at minute concentrations. Furthermore, the range of odorants (smellable molecules) detected is chemically diverse and seemingly infinite. This is due to the fact that the olfactory (smell) receptors recognize only parts of the scent molecule, and thus one odorant can in turn activate many different types of receptor [1]. The combinatorial signal generated is then interpreted by the brain and identified. Likewise, since every receptor only recognizes key molecular groups, multiple odorants will activate the same receptor [2]. This gives the mammalian nose the striking ability to detect new chemicals never encountered in evolution.

The detection range of the human nose is on the order of parts per million (ppm) down to parts per billion (ppb). As an analogy, detecting one part per billion is the equivalent of finding one particular penny in 10 million dollars worth of pennies! The sensitivity of the mammalian nose is exemplified in the canine and mouse, where many compounds can be detected down to below the parts per trillion (ppt) threshold [3]. By comparison, most current artificial noses can only detect in the ppm range. Additionally, other mammalian species possess even greater olfactory capability. For example, it is
estimated that the grizzly bear has a sense of smell seven times more sensitive than a bloodhound’s.

**Olfactory Receptor (OR) Background**

![Figure 1.1](image-url)  
*Figure 1.1* — Cartoon representation of an olfactory receptor showing the seven transmembrane domains and loop regions. Reprinted from [5] (figure 3) with kind permission from Springer Science+Business Media, ©2003.

Olfactory receptors (or odorant receptors) are a large, highly related class of G-Protein Coupled Receptors (GPCRs) that function together to allow discrimination between a wide range of volatile molecules [4]. All GPCRs are integral membrane proteins with seven transmembrane domains arranged in a barrel-like conformation (see Figure 1.1). In olfactory receptors, it is thought that this configuration forms a funnel-shaped pocket for odorant binding [5].
The olfactory receptor gene family constitutes the largest single class of genes in the vertebrate genome (2-3% in the human) [6]. Current estimates put the number of human olfactory receptor genes at 636, with only 339 being functional and the rest being non-functional pseudogenes [1, 7-9]. This is considerably less than the mouse OR gene family of nearly 1300 (913 functional) [10, 11] or the canine OR gene family of roughly 1200 (~1000 functional) [12].

The olfactory receptor gene family has been broken into two classes, Class I and Class II. Class II receptors constitute the bulk of mammalian olfactory genes and have affinity for small, volatile hydrophobic compounds. The Class I “fish-like” receptors show genetic similarity to many amphibious ORs and have affinity for small water-soluble compounds. Olfactory receptors are additionally subdivided into a number of

Figure 1.2 – Diagram of olfactory receptor downstream signaling in olfactory neurons. Reprinted from [15] with permission from Elsevier, ©1998.
Figure 1.3 – Neuronal wiring of olfactory system. Signals from neurons that express the same receptor later converge, thus the brain can tell which receptors were activated by analyzing a spatial pattern of activity. Reprinted from [19] with permission of Oxford University Press, ©2002.

subfamilies (172 total for humans) [8, 13].

All ORs (with key exceptions) are expressed solely on the cilia of olfactory neurons within the nasal cavity [14]. Odorant binding and recognition leads to activation and release of the olfactory G-protein G_{olf}, which triggers cyclic-AMP production, ion-channel-mediated Ca^{2+} release, and finally the firing of an action potential into the olfactory bulb to be interpreted by the brain (see Figure 1.2) [15-17]. Through an unknown mechanism of allelic inactivation, every olfactory neuron chooses a single OR gene to express [18]. Signals from neurons that express the same olfactory receptor later converge downstream at neural foci called glomeruli [19, 20]. As the same odorant will stimulate multiple ORs (and to various strengths), the brain receives a spatial map of receptor activity through these glomeruli (see Figure 1.3) [21]. Odorants are thought to be recognized by matching a specific spatial pattern (a combinatorial code) [1, 22].
Despite the genetic and signaling knowledge of the OR family, very little is known regarding how these receptors function and what determines their odorant specificity [23]. Currently, ligand specificities have only been determined in 26 mammalian ORs and a handful of ORs from other non-mammalian species (for a summary, see [8, 11]). The method of measuring odorant specificity currently relies on several techniques for measuring downstream OR activation in cells:

- calcium-influx imaging assays of single olfactory neurons
- calcium-release assays in heterologous expression systems (such as HEK293 cells)
- voltage clamping
- cAMP-responsive element (CRE)-driven luciferase reporters
- detection of cAMP via radiolabeling

There have been a host of previous studies that have expressed and studied olfactory receptors in native and heterologous systems (see Table 1). However, to date there has not been a case where olfactory receptors have been overexpressed and shown to be functionally purified. While purification attempts have been made in bacterial and insect systems, these organisms lack mammalian post-translational machinery and thus purified receptors may be missing critical modifications. It is known that most, if not all, GPCRs are glycosylated, and indeed the olfactory receptors have conserved N-linked glycosylation sites (Asn-X-Ser/Thr) at their N-terminus [24-26]. Several studies have
indicated that loss of glycosylation can lead to improper folding and targeting, resulting in decreased function and compromised structure [25, 27, 28]. Loss of either N-terminal glycosylation site (Asn-2 or Asn-15) of rhodopsin is sufficient to cause loss of signal transduction despite no apparent change in localization or folding [29, 30]. This indicates that purification of olfactory receptors from mammalian expression systems might be crucial for functional expression and purification.

Table 1 – Summary of previous OR expression methods

<table>
<thead>
<tr>
<th>Expression System:</th>
<th>Functionality:</th>
<th>Purification:</th>
<th>Reference(s):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria (E. Coli)</td>
<td>Not measured</td>
<td>6x-His</td>
<td>[31-33]</td>
</tr>
<tr>
<td>Insect (Sf9)</td>
<td>Measured by radiolabeled cAMP</td>
<td>No</td>
<td>[34]</td>
</tr>
<tr>
<td>Insect (Sf9)</td>
<td>Not measured</td>
<td>6x-His / FLAG</td>
<td>[35-37]</td>
</tr>
<tr>
<td>Xenopus Oocyte</td>
<td>Voltage clamp</td>
<td>No</td>
<td>[38-41]</td>
</tr>
<tr>
<td>Olfactory Epithelium</td>
<td>Measured by [Ca2+] influx</td>
<td>No</td>
<td>[1, 42-46]</td>
</tr>
<tr>
<td>Mammalian (COS-1)</td>
<td>Measured by [Ca2+] influx</td>
<td>No</td>
<td>[41, 47]</td>
</tr>
<tr>
<td>Mammalian (HeLa)</td>
<td>Measured by [Ca2+] influx</td>
<td>No</td>
<td>[48]</td>
</tr>
<tr>
<td>Mammalian (HEK293)</td>
<td>Measured by [Ca2+] influx</td>
<td>No</td>
<td>[2, 39, 41, 43, 47, 49-52]</td>
</tr>
<tr>
<td>Mammalian (HEK293)</td>
<td>Measured by CRE-responsive luciferase</td>
<td>No</td>
<td>[53]</td>
</tr>
</tbody>
</table>

There have been numerous reports in the literature citing difficulties in expressing functional olfactory receptors in heterologous systems [54]. This problem stems from improper membrane targeting and resulting cytoplasmic localization of the majority of ORs [27, 55-57]. A recent study by Saito et. al. [53] discovered several genes exclusively expressed in olfactory neurons that dramatically increased OR expression levels and membrane targeting when co-transfected with olfactory receptors into heterologous
mammalian expression systems. These were genes were dubbed RTPs (Receptor Targeting Proteins).

1.2 Detergent Solubilization

When purifying and studying integral membrane proteins, it is essential to select optimal detergents and surfactants that efficiently solubilize receptors from cell membrane and also maintain protein stability and allow high concentrations of receptor without protein aggregation and precipitation [59]. When detergents are used at concentrations above their critical micelle concentration (CMC), they form micelles which can disrupt cellular lipid bilayers and extract membrane proteins (Figure 1.4A). In this thesis we performed a full spectrum detergent screen that included representatives from the non-ionic, zwitter-ionic, polar, and ionic detergent classes (Figure 1.4B).

1.3 Inducible Cell Lines

Initial OR expression trials in HEK293S cells using transient transfections had several disadvantages for large-scale protein purification: i) the transfection alone caused significant cellular toxicity; ii) a significant fraction of the cells remain untransfected and do not produce OR; iii) the cost of transfection reagents required for milligram-scale production would be prohibitive. One solution is to generate cell lines which have a stable copy of the OR gene integrated into the genome, maintained by drug selection. However, this would result in constant production of the OR protein, which we know to have toxic effects. Additionally, harvesting cells with constitutive production would yield OR protein at various stages of degradation by the cell, leading to inhomogeneity and
Figure 1.4 – Detergent Solubilization. (A) Depiction of detergent extraction membrane proteins from native lipid bilayers. Detergent is introduced at a concentration above its CMC so that detergent micelles break up the cell membrane (left). Solubilized membrane proteins will have their hydrophobic portions coated with detergent monomers, such that integral membrane proteins (such as ORs) will have a “belt” of solubilizing detergent surrounding their transmembrane area (right). Reprinted from [59] with permission from Elsevier, ©2003. (B) Molecular structure of several representatives of different detergent classes. FC14 was later found to be the optimal detergent for solubilizing hOR17-4.
Figure 1.5 – Diagram of the T-REx tetracycline-inducible system. (A) The pcDNA6/TR plasmid (blasticidin resistance) expresses the tetR protein, which forms homodimers. (B) The gene of interest is inserted into the pcDNA4/To plasmid (zeocin resistance) and expression is repressed by tetR homodimers binding to TetO2 sites present in the CMV promoter. (C) Upon addition, tetracycline binds to tetR homodimers. (D) The resulting conformational change of tetracycline-bound tetR homodimers causes them to release from the TetO2 sites, which causes derepression and expression of the gene of interest from upstream CMV promoter.
functionally compromised receptor. Therefore we used the tetracycline-inducible T-REx system [60] from Invitrogen (see Figure 1.5) to construct cell lines in which production of OR could be initiated on demand. Thus the cells can be grown to the desired concentration and, when desired, production of fresh OR triggered by the addition of tetracycline (a so called “tet ON” system). Such a system is advantageous as it is cheap (no further transfection required) and efficient (100% of the cells will express OR).

1.4 Bioreactor Culture

Although a confluent 150 mm tissue culture plate can contain 50-70 million HEK293S cells, the low levels of expression seen for most membrane proteins result in a yield of tens of micrograms of GPCR per plate, and even less following subsequent purification. Therefore, in order to obtain milligram levels of a GPCR using mammalian cell culture, hundreds of tissue culture plates would be required, and such experiments would be both extremely time consuming and expensive to perform. One solution is the use of liquid bioreactors to culture mammalian cells in suspension. For instance the suspension-adapted HEK293 cell line (HEK293S) can achieve cell densities in excess of 5 million cells per milliliter in suspension culture [61], and thus one liter of bioreactor medium can contain the cellular equivalent of 100 tissue culture plates.

A bioreactor is essentially a self-regulating sterile fluid chamber (see Figure 1.6). In my research I used the Celligen Plus stirred-tank bioreactor (New Brunswick Scientific). A 4-gas mixture (Air, N₂, O₂, CO₂) is bubbled through the media and the mixture automatically adjusted to regulate the pH and dissolved oxygen content. A
Figure 1.6 – Bioreactor operation. (A) The Celligen Plus liquid bioreactor during a suspension culture run. (B) Close-up of reactor vessel showing sensors, impeller, and gas sparge. (C) Bottle equipped with a sterile siphon system to aseptically add liquid to the reactor vessel. (D) Cell suspension harvested after a successful bioreactor run.
circulating external water jacket maintains culture temperature (37°C). The cells are kept in suspension by an impeller blade that rotates throughout the run and breaks up clumps. The reactor housing has four inlet ports for the sterile addition of solutions or cell suspensions to the reactor, as well as one total drain line in order to evacuate the vessel. For general operation and set-up of the bioreactor, please see [62].

1.5 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy can be used to characterize molecules in solution based on their differential absorption of circularly polarized light [63, 64]. In contrast to linearly polarized light, where the electric field vector component only oscillated in one plane, in circularly polarized light this electric field vector has constant magnitude and rotates in a circle about the direction of propagation. Thus as the light moves through space, the electric field vector forms a helix, which can either be left-handed (left circularly polarized) or right-handed (right circularly polarized). For an excellent depiction of circularly polarized light, see [65]. As this light is “handed” (or chiral), it will preferentially be absorbed by chiral molecules. To obtain a CD spectrum, equal amounts of left- and right-handed circularly polarized light are applied to a solution containing chiral molecules, and the differential absorption is measured with respect to wavelength.

Chiral biomolecules such as nucleic acids, peptides, and proteins can be analyzed in one of several methods [66]. To obtain secondary structure, the far-ultraviolet (190-250 nm) region of the spectrum can be used to estimate the fraction of alpha-helix, beta-sheet, and random coil present in a protein [67], as all these structures have characteristic spectra (Figure 1.7). This technique has been successfully applied not only to soluble
proteins but detergent-solubilized membrane proteins as well [68, 69]. As olfactory receptors possess seven alpha-helical transmembrane domains, their far-UV CD spectra should be dominated by the 208 nm and 222 nm peaks. To obtain information on tertiary structure, the near-UV (250-300 nm) range can be analyzed, although this requires a much higher protein concentration as the absorption signals are quite weak compared to secondary structures. Signals in the near-UV range are due to highly structured aromatic amino acids and disulfide bonds as well as their surrounding environment [70]. One disadvantage to near-UV CD is that, unlike far-UV, spectra cannot be quantified and assigned to any particular structure. Thus, it is only useful if one compares related spectra, such as mutations or denatured versions of the same protein.

Figure 1.7 – Far-UV circular dichroism spectra characteristic of the major protein secondary structures. Secondary structures shown are alpha-helix (1), beta-sheet (2), random coil (3), coiled coil (4), and fully denatured (5). Reprinted from [67] with permission from Macmillan Publishers Ltd, ©2006.
1.6 Surface Plasmon Resonance

Surface plasmon resonance (SPR) is a proven method to analyze molecular interactions in real time. For a review of SPR principles and analysis, please see [71, 72]. A major advantage of the system is that it is label-free, meaning none of the analytes need to be modified with such additions as fluorophores. The assay relies on the phenomena that under total reflection conditions, polarized light bouncing off a glass-liquid interface will produce an electrical field in the liquid (called an evanescent field). If the glass-liquid interface is coated with a thin layer of gold, the evanescent field will interact with gold electrons and produce a surface plasmon (a charge density wave). At a critical angle (called the SPR angle), energy transfer between the plasmon and evanescent field occurs, causing an intensity minimum in the reflected light. The SPR angle varies with the index of refraction of the liquid near the gold surface (within about 200 nm) due to this energy coupling, and thus changes in mass bound or near the surface will cause changes in the SPR angle [73]. Flowing molecules past the detector and constantly measuring the shifts in SPR angle as a function of time gives an SPR data readout, called a sensorgram.

When studying receptor-ligand interactions, the receptor can be immobilized onto the SPR chip surface using any of several methods such as via covalent amine-coupling to a gold-linked dextran matrix. A mobile ligand is then injected and run along the flow cell, whereupon binding to receptor causes a change in the local index of refraction and a resulting shift in the SPR angle (Figure 1.8A) [74]. The resulting sensorgram displays several phases (Figure 1.8B): i) an association phase dependent on kinetics (k_on), ii) a steady-state which depends on the dissociation constant (K_d), and, after the ligand is no
longer applied, iii) a dissociation phase depending on kinetics ($k_{off}$). By analyzing sensorgrams from multiple concentrations of ligand it is possible to estimate these rate constants [75]. The technique has already been applied to several GPCRs including $\alpha_2$-adrenergic receptor [76], neurotensin receptor [77], and the CCR5 and CXCR4 chemokine receptors [78]. While the ligands for these receptors are large polypeptides, other studies have been able to detect the mass increase associated with the binding of molecules as small as 200 Daltons [79]. As most odorants are in this range, the use of the latest SPR machines like the Biacore™ T100 (sensitive to less than 0.1 pg/mm²) has a high potential to measure odorant-OR binding events in such an assay.
Figure 1.8 – Overview of surface plasmon resonance (SPR). (A) Diagram of an SPR biosensor surface. Binding interactions of the soluble ligand with immobilized protein is detected via a change in the SPR angle. (B) Resulting sensorgram readout of the experiment in part A. Sample injection results in association and steady state phases. A dissociation phase results when ligand is no longer injected. Reprinted from [74] with permission from Elsevier, ©2007.
1.7 References


64. [http://www.enzim.hu/~szia/cddemo/edemo1.htm](http://www.enzim.hu/~szia/cddemo/edemo1.htm)


70. [http://www.ap-lab.com/circular_dichroism.htm](http://www.ap-lab.com/circular_dichroism.htm)


CHAPTER 2

FABRICATION OF CUSTOM OLFACTORY RECEPTOR GENES

2.1 Overview of PCR-based gene synthesis

The initial stage of the project will be the construction of selected olfactory receptor gene constructs and insertion into specific expression plasmids. For this I chose the method of polymerase construction and amplification (PCA). This utilized a PCR-based method of gene synthesis [1] that involves parsing the gene sequence into a set of small overlapping oligonucleotides. During an initial assembly PCR, these oligos function as both primer and template, while the DNA polymerase successively builds longer and longer fragments with each round of PCR. A second amplification PCR then enriches for the full-length gene. For a detailed illustration of the mechanism of PCR-based gene synthesis, please see Figure 2.1.

To assist in parsing the sequence into oligos, I used the online program DNAWorks (http://helixweb.nih.gov/dnaworks) developed by Hoover and Lubkowski [2]. The amino acid sequence of each olfactory receptor was obtained from GenBank and directly inputted into the software. To adapt the gene for use in specific expression systems and to facilitate purification and detection, the following sequence modifications were made using the DNAWorks software:

a. Codon optimization for expression system of interest
b. Addition of N-terminal and C-terminal tags (along with glycine linkers and signal cleavage sequences) for purification and detection

c. Addition of ribosome binding sites (Kozak / Shine-Delgarno sequences) at 5' end of gene

d. Addition of flanking restriction sites for cloning into expression vector

A number of N-terminal and C-terminal sequence tags were tested, including FLAG and 6x-His (see Appendix A.2). Initially, we focused on the purification tag utilized by the Khorana lab for rhodopsin, namely the nine C-terminal amino acid sequence (TETSQVAPA) for which a specific antibody has been generated (rho1D4). This tag has already facilitated early-stage immunoaffinity purifications of other several other membrane proteins [3, 4].

There have been numerous reports in the literature citing difficulties in expressing functional olfactory receptors in heterologous systems. The main problem seems to be improper membrane targeting and resulting cytosolic localization of the majority of ORs [5-8]. Several studies have been able to alleviate this problem by the addition of putative membrane import signals to the N-terminus of the olfactory receptor [9, 10]. However, more recent studies [11, 12] have managed to express functional olfactory receptors in HEK293 cells without the use of such signal tags. We decided to also investigate whether these putative import sequences could improve expression in the olfactory receptors studied.
Figure 2.1 - Overview of PCR-based Gene Assembly

1st PCR reaction: Assembly

- Parse gene into staggered oligo set (uniform overlap Tm's and minimal hairpin formation), then order oligos and mix

Assembly uses all oligos as both primers and templates

Assembly products increase in length with every PCR cycle

After sufficient cycles of assembly PCR reaction, the full length gene is constructed

However, the final assembly reaction mix contains not only full length product but a mixture of many different product sizes (literally a "smear" when visualized)

2nd PCR reaction: Amplification

Amplifies full length gene

Uses end oligos as primers and small amount of assembly reaction as template
2.2 Selection of OR Genes for Construction

To select genes for study, we took the subset of ORs with known ligand specificity and selected a small group that represented the diversity of olfactory receptors (both genetically and in the class of odorants they respond to) and that also showed strong homology across the well-studied mammalian OR families (canine, mouse, human). The following list of candidate OR genes contains members of both Class I and Class II, and respond to a range of aliphatic and aromatic chemicals (see Figure 2.2).

Figure 2.2 – Olfactory receptor phylogenetic tree: Receptors boxed in green have had synthetic genes constructed using PCA. The receptor boxed in red, hOR17-4, was the candidate receptor chosen for subsequent expression and large-scale purification. The grayed top quadrant of the tree consists of Class I (fish-like) receptors. Reprinted with permission from [13], ©2004, National Academy of Sciences, U.S.A.
**Human hOR17-4 receptor (also known as OR1D2)**

This is the second human olfactory receptor for which ligand specificity has been studied. This receptor is unique in that it is expressed in sperm and appears to govern the sperm’s ability to “home in” on the egg. It has been shown to respond to bourgeonal [14-16], and other lily-like aromatic aldehydes such as lilial and floralzone (Figure 2.3).

![Figure 2.3](image)

**Figure 2.3 – Partial odorant ligand spectrum for hOR17-4.** The hOR17-4 receptor was transiently expressed on HEK293 cells and odorant responses measured by calcium imaging. Active compounds, such as bourgeonal, were aromatic with an aldehyde group (usually two carbons away from the aromatic ring) and a carbohydrate group in the para position. Undecanal was found to be a receptor antagonist. From [16], reprinted with permission from AAAS, ©2003.
**Mouse IG7 receptor (also known as mOR276-1)**

This receptor shows a high response to limonene [9], a well-studied aromatic alkene (hydrocarbon) from citrus fruits that is highly used in industry. The mouse receptor is 88% identical to its human ortholog, hOR2V1.

**Mouse S51 receptor (also known as mOR40-1)**

This receptor is a member of the “fish-like” Class I olfactory receptor family. It has been shown to respond to a range of aliphatic carboxylic acids and alcohols [17]. The mouse receptor is 85% identical to its human ortholog, hOR56A5.

**Mouse I7 receptor (also known as mOR103-15)**

This Class I receptor has been highly studied in both rat and mouse, and shows 90% identity with its human ortholog, hOR6A2. It is highly responsive to aliphatic aldehydes, such as octanal and heptanal [9, 18].

**Mouse M71 receptor (also known as mOR71-2)**

This receptor shows specificity for acetophenone [19], an aromatic ketone. Acetophenone is highly used in the perfuming and flavoring industries for such odors as strawberry and jasmine. The mouse receptor shows 85% identity to its human ortholog hOR8A1.
**Mouse OR73 receptor (also known as mOR-EG and mOR174-9)**

This receptor shows specificity for eugenol [20], an aromatic alcohol / ester. Eugenol has a spiced clove and cinnamon odor and is widely used in industry. The mouse receptor is 82% identical to its human ortholog hOR5D18.

**Mouse IC6 receptor (also known as mOR118-1)**

This receptor shows specificity for the (-) stereoisomer of citronellal [9], a lemon-scented monoterpenoid, which is often used as an insect repellent. The mouse receptor is 81% identical to its human ortholog hOR6E1P.

**Mouse OR23 (also known as mOR267-13)**

This is the mouse homolog to human olfactory receptor 17-4, in that it is expressed in mouse sperm and mediates chemoattraction. It has been shown to respond to the bourgeonal-like odorant lyral [21]. It is 39% identical to hOR17-4 and 87% identical to hOR10J5.

**Nematode ODR-10 receptor**

A well studied olfactory receptor in the worm (C. elegans) which has been shown to have high specificity for diacetyl [22-24], an aliphatic ketone. This receptor is shows little homology to mammalian receptors.
2.3 Fabrication of the hOR17-4-rho synthetic gene

Methods

To adapt the genes for use in mammalian cell expression and purification, the following sequence modifications were made using the DNAWorks software:

i. Human codon optimization (for HEK293S cells)

ii. Addition of C-terminal rhoID4 epitope tag “TETSQVAPA” (preceded by a two glycine linker) to facilitate purification

iii. Addition of NotI restriction site at the 3’ end of gene

iv. Addition of EcoRI restriction site and Kozak sequence (eukaryotic ribosome binding site) at the 5’ end of gene

The designed oligos (Table 2) were purchased from IDT (Coralville, IA) with a maximum length of 45 bp. The assembly PCR was run for a variable number of cycles using a mixture of all oligos at a concentration of 25 nM each. The amplification PCR was run for 30 cycles using 1 µL of the assembly PCR (in a total reaction volume of 50 µL) and each end oligo at a concentration of 300 nM each. PCR reactions were then analyzed by gel electrophoresis and stained with ethidium bromide. Full-length product was excised, extracted, and then digested with the pertinent restriction enzymes. The genes were then ligated into the pcDNA4/To expression plasmid (Invitrogen), sequenced, and a correct clones grown up using a MaxiPrep kit (Qiagen, Valencia, CA).
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>$#$ bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1:</td>
<td>CCTGAATTCGCCGCCACATGGACGGAGGCAA</td>
<td>32</td>
</tr>
<tr>
<td>S2:</td>
<td>CCAACAGCGAGGCCAGGCTTTCTGCTGGTGG</td>
<td>32</td>
</tr>
<tr>
<td>S3:</td>
<td>GCATTTCGGAGCCCGAGCAACACGAGATCCT</td>
<td>34</td>
</tr>
<tr>
<td>S4:</td>
<td>CTTTTGAGTTCTCTGAGCATGTATCTGGTCACCGTGGTCG</td>
<td>41</td>
</tr>
<tr>
<td>S5:</td>
<td>GAAATGTCTTGATATTCTCAGTACGGCAAAGCCAGCACTTCA</td>
<td>44</td>
</tr>
<tr>
<td>S6:</td>
<td>CATACCCCCCTCTCCTTCTTCTGCTAACCTCTCTCTTACAGA</td>
<td>44</td>
</tr>
<tr>
<td>S7:</td>
<td>CCTGTTTTTCTGTCACAAACACACTATCTCAGAAAATGGTCGTCACA</td>
<td>42</td>
</tr>
<tr>
<td>S8:</td>
<td>CCTCCAAACGCAACACAGACTTATAGCTATGCGGGTCG</td>
<td>39</td>
</tr>
<tr>
<td>S9:</td>
<td>CCTACACAAACTCTATTTTCTCTGAGGCTGTCGGC</td>
<td>38</td>
</tr>
<tr>
<td>S10:</td>
<td>TGGAATAATCTGATTTCTGCCCCGCTCATGCTCTTTAGATCAGTGC</td>
<td>41</td>
</tr>
<tr>
<td>S11:</td>
<td>GTGCTATTCTGCTCTTCATACACTAAGCTATAGGCC</td>
<td>43</td>
</tr>
<tr>
<td>S12:</td>
<td>CTAACCGTCTGCTCCTGTGCTCCCTGTGGTGG</td>
<td>36</td>
</tr>
<tr>
<td>S13:</td>
<td>TGCTCTCCGTGTTATGAGATTACACACACTGCTGCA</td>
<td>39</td>
</tr>
<tr>
<td>S14:</td>
<td>TGCAAGATGAGCTTTTTGGTCGTCAGAAAGATCCACTATCTTT</td>
<td>45</td>
</tr>
<tr>
<td>S15:</td>
<td>TCCTCGGAAATCTGCCTCTCCCGATGGCTGTAAG</td>
<td>37</td>
</tr>
<tr>
<td>S16:</td>
<td>CAACATCACATTACATACATCTGCTGATGCTAGGGATCGTTT</td>
<td>45</td>
</tr>
<tr>
<td>S17:</td>
<td>TATTTCTCTCATCCCTTGGATGTCATGTCAGCTACT</td>
<td>43</td>
</tr>
<tr>
<td>S18:</td>
<td>CTCTATACAGACATCCTCTCAAGACACTCGGACTGACTCATA</td>
<td>43</td>
</tr>
<tr>
<td>S19:</td>
<td>ATATATAGATACACCATGCTGCTGAGCATGTCCTCC</td>
<td>39</td>
</tr>
<tr>
<td>S20:</td>
<td>CCGTCAGGCTTTTTAATAGGAAACTGCTAGGTCTATCATCTCAAAC</td>
<td>45</td>
</tr>
<tr>
<td>S21:</td>
<td>CTTCCACACCTACACCGTCGAGACTCCGTCGTA</td>
<td>36</td>
</tr>
<tr>
<td>S22:</td>
<td>CAGTAGTGTAGGGCTGTCGTCACCCCCATGATGAACCC</td>
<td>37</td>
</tr>
<tr>
<td>S23:</td>
<td>CCTTCATCTACTCCCCTCTGaAACAAAGATATGCTAGGGCCTCTC</td>
<td>43</td>
</tr>
<tr>
<td>S24:</td>
<td>GGGAACACTCTCTGGCAAAACACTTTAAAGAAGCTACCGGAGG</td>
<td>41</td>
</tr>
<tr>
<td>S25:</td>
<td>CACAGAGATCTCCAGTCCTGCTCTTGTTAAGC</td>
<td>36</td>
</tr>
<tr>
<td>AS1:</td>
<td>CTCTGCGGCGCCCGCTTTAAGC</td>
<td>25</td>
</tr>
<tr>
<td>AS2:</td>
<td>CGACCTTGGGATTGTCCTGCTGCTGCTCCGGTCAGCTCTTTAAG</td>
<td>41</td>
</tr>
<tr>
<td>AS3:</td>
<td>TGTTTCTTCTATCCGAGGCGATGTCATGTCAGCTACT</td>
<td>38</td>
</tr>
<tr>
<td>AS4:</td>
<td>TTTGTTTCGAGGAGTAAAGGTAGAAGATGGAAGCGCTACTTGAATGGTTT</td>
<td>43</td>
</tr>
<tr>
<td>AS5:</td>
<td>ACAGAACGCGTACAGCTACTGTGCTAGCGAGCTCTTGAC</td>
<td>37</td>
</tr>
<tr>
<td>AS6:</td>
<td>CCTCGTATGGGTGAGAGGTAGGCTAGTACAGACTACAGCAGTG</td>
<td>43</td>
</tr>
<tr>
<td>AS7:</td>
<td>TCCATCAAACAGCTGAGCCGCGCTCCCATGATGGTGC</td>
<td>36</td>
</tr>
<tr>
<td>AS8:</td>
<td>ACAGTGCTCTGAAAGACCTTTATATTTTTTGCTGAGGAAAGGATC</td>
<td>43</td>
</tr>
<tr>
<td>AS9:</td>
<td>CGGAGATGCGCTGTCTATGAAAGGAGCGAGATGAGCTGTGACTGCA</td>
<td>44</td>
</tr>
<tr>
<td>AS10:</td>
<td>TCCGAAAGGGGATAGAAATAATAAGACATCCGGGTAAGGCT</td>
<td>42</td>
</tr>
<tr>
<td>AS11:</td>
<td>CACGTATGTGTTAACTGTGATGTTGCTACAGCCATCCGAGG</td>
<td>42</td>
</tr>
<tr>
<td>AS12:</td>
<td>GAGGACATGATATTTCTGCAAAAATAAGTGTAGGATGACTTCCTTTCCTGAG</td>
<td>44</td>
</tr>
<tr>
<td>AS13:</td>
<td>CACAAAAGGTCACTCTTGTCATGAGCAGTAGTGTGATGATGC</td>
<td>42</td>
</tr>
<tr>
<td>AS14:</td>
<td>CATACACAGCACGGACACCCACAGCAACAGGAC</td>
<td>34</td>
</tr>
<tr>
<td>AS15:</td>
<td>GACGAACTGTAAGGAAGCTTTATAGAGGTGAAGGTGCTAGTT</td>
<td>43</td>
</tr>
<tr>
<td>AS16:</td>
<td>AGAGGGAAACAAGATAGCAGTCAGGCAGTTGTAAGCAGGC</td>
<td>40</td>
</tr>
<tr>
<td>AS17:</td>
<td>CGCAGAAGGTACCATTACAGGAGCCAGCCGACACCGATGC</td>
<td>35</td>
</tr>
<tr>
<td>AS18:</td>
<td>AGCGAAAATAAGAAGTTGAGGAGAGCCCGGGAACATGCTAAATAG</td>
<td>43</td>
</tr>
<tr>
<td>AS19:</td>
<td>CTTCTGTGCTGCTGCTGAGGCTAGGACATTTTGGAGGAT</td>
<td>41</td>
</tr>
<tr>
<td>AS20:</td>
<td>GTCTTGTGAGGAAGGAAAGCAGCTTGCTAAGGAGGATGTCAGCAA</td>
<td>43</td>
</tr>
<tr>
<td>AS21:</td>
<td>GAAAGATGAGGCTAGGAGTGGATGGCTGCTGCTGAGG</td>
<td>39</td>
</tr>
<tr>
<td>AS22:</td>
<td>TAATAGCGAGATATACACACATCTGAGCGCCACCCGACCTGCA</td>
<td>42</td>
</tr>
<tr>
<td>AS23:</td>
<td>GATACATGCTCAAGAAACATCCAAAAGAGGATGCTCTGCTGTCG</td>
<td>44</td>
</tr>
<tr>
<td>AS24:</td>
<td>GGCGTCTCGGACATCGCAGCAGCAAAACACTGCTGCTGCTGCTG</td>
<td>33</td>
</tr>
<tr>
<td>AS25:</td>
<td>CTGCCCCCTGCTTGGTGGTCTCCGTCCTCACTTAGC</td>
<td>31</td>
</tr>
</tbody>
</table>

Table 2 - Primers used for the PCR-based synthesis of the engineered hOR17-4-rho gene. A total of 50 oligonucleotides were used to construct the synthetic gene (25 sense strand oligos, labeled S1-25) and 25 anti-sense strand oligos, labeled AS1-25).
Results

Default protocols for PCR-based gene assembly suggest 20-30 cycles in the assembly reaction to be sufficient for the synthesis of genes up to and beyond 1 kb [2]. However, all of our initial trials using these parameters proved unsuccessful. We hypothesized that, due to possible inefficiencies in oligo gene assembly, more cycles might be necessary to fully elongate the oligo templates. Indeed, we then discovered that the critical parameter governing the success of our PCR-based gene assemblies to be the number of PCR cycles in the initial assembly reaction. As can be seen in Figure 2.4, the

Figure 2.4 – Fabrication of custom hOR17-4-rho gene. Increasing cycles in 1st PCR generate longer assembly products and ultimately enable amplification of full-length OR gene in 2nd PCR. Lane 1 is DNA standards – the bands (from the bottom) are 250, 500, 750, and 1000 bp.
largest species visible in the assembly reaction “smear” increased with number of cycles, and full-length product (~1 kb) was not achieved in the amplification reaction (2nd PCR) until 45 assembly cycles.

Subsequent to construction of our first olfactory receptor, hOR17-4-rho, eight additional olfactory receptors (see Appendix C) were synthesized, most on the first attempt, using our modified procedure. Additionally, we have adapted the system to include generalized tag “modules” which can be swapped into existing oligo gene sets using a custom linker oligo. Three additional tagged versions (R-hOR17-4, HF-hOR17-4, and HF-hOR17-4), incorporating N-terminal “membrane import sequences” or His and FLAG tags, have been generated to date. This demonstrates that our adapted protocol can be generally applied to our project for the rapid fabrication of synthetic olfactory receptor gene constructs.

2.4 References


CHAPTER 3

PRODUCTION OF HOR17-4 IN ADHERENT CELL CULTURES

NOTE: This chapter is a review and expansion of previously published result:

Cook BL, Ernberg KE, Chung H, Zhang S. Study of a synthetic human olfactory receptor 17-4: expression and purification from an inducible mammalian cell line.

3.1 Abstract

In order to begin to study the structural and functional mechanisms of olfactory receptors, methods for milligram-scale purification are required. Here we demonstrate the production and expression of a synthetically engineered human olfactory receptor hOR17-4 gene in a stable tetracycline-inducible mammalian cell line (HEK293S). The OR gene was fabricated from scratch using PCR-based gene-assembly, which facilitated codon optimization and attachment of a 9-residue bovine rhodopsin affinity tag for detection and purification. Induction of adherent cultures with tetracycline together with sodium butyrate led to hOR17-4 expression levels of \(-30\ \mu g\) per 150 mm tissue culture plate. Fos-choline-based detergents proved highly capable of extracting the receptors, and fos-choline-14 (N-tetradecylphosphocholine) was selected for optimal solubilization and subsequent purification. Analysis by SDS-PAGE revealed both monomeric and dimeric receptor forms, as well as higher MW oligomeric species. A two-step purification method of immunoaffinity and size exclusion chromatography was optimized which enabled 0.13 milligrams of hOR17-4 monomer to be obtained at \(>90\%\) purity. This high purity of hOR17-4 is not only suitable for secondary structural and functional analyses but also for subsequent crystallization trials. Thus, this system demonstrates the feasibility of purifying milligram quantities of the GPCR membrane protein hOR17-4.
3.2 Introduction

Membrane proteins are of vital importance to life, as evidenced by the fact that ~30% of the genes in almost all sequenced genomes code for membrane proteins [1-3]. However, our understanding of the structures and functions of membrane proteins has lagged behind the known soluble proteins. As of June 2008, there are only 160 unique membrane protein structures known [http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html], which constitutes less than 1% of all known protein structures. The major bottleneck in obtaining membrane protein structures is the notorious difficulty involved in expressing and purifying the large quantities of membrane protein sample required for X-ray crystallography. In order to accelerate membrane protein structural and function studies, new and simple methods for membrane protein production must be developed.

Olfactory receptors (or odorant receptors) are an extremely large class of G-Protein Coupled Receptors (GPCRs) that function together combinatorially to allow discrimination between a wide range of volatile molecules [4,5]. All GPCRs are integral membrane proteins with seven transmembrane domains arranged in a barrel-like conformation. In olfactory receptors, it is believed that this configuration forms a funnel-shaped pocket for odorant recognition [6]. The olfactory receptor (OR) gene family constitutes the largest single class of genes in the vertebrate genome (2-3% in the human). Current estimates put the number of human olfactory receptor genes at 636, with only 339 being functional and the rest being non-functional pseudogenes [7]. This is considerably less than the mouse OR gene family of 1209 (913 functional) [8] or the canine OR gene family of roughly 1200 (~1000 functional) [9]. Despite the fact that they represent the largest class of known membrane proteins, no detailed structure exists for
any olfactory receptor and the functional mechanisms of these amazing receptors remains unknown. The crucial first step to enable such pivotal studies is to engineer systems with the capacity to generate and purify milligram quantities of an olfactory receptor.

Mammalian olfactory receptors are expressed on the cilia of olfactory neurons within the nasal cavity. Odorant binding and recognition leads to activation and release of the olfactory G-protein Golf, which triggers cyclic-AMP production, ion-channel-mediated Ca2+ influx, and finally the firing of an action potential into the olfactory bulb to be interpreted by the brain [10]. Through an unknown mechanism of allelic inactivation, every olfactory neuron chooses a single OR gene to express. Signals from neurons that express the same olfactory receptor later converge downstream at neural foci called glomeruli [11]. As the same odorant will stimulate multiple ORs (and to various strengths), the brain receives a spatial map of receptor activity through these glomeruli [12]. Odorants are thought to be recognized by matching a specific spatial pattern (a combinatorial code) [5].

The human olfactory receptor 17-4 (hOR17-4, alternately known as OR1D2) is of particular interest since, in addition to olfactory neurons, it is also expressed on the midpiece of human spermatozoa [13]. Sperm expressing hOR17-4 were found to migrate towards known hOR17-4-responsive odorants such as bourgeonal. Thus the receptor serves a dual role in that it recognizes odorants in the nose as well as plays a potential role in sperm chemotaxis and fertilization. As structural studies of hOR17-4 would not only provide information crucial to understanding the molecular mechanism(s) of olfaction but also have application to human reproduction, we selected this receptor as our prototype OR for expression and purification trials.
The GPCR family represents one of the most important known receptor classes as evidenced by the fact that half of all pharmaceutical drugs target GPCRs [14]. Despite their crucial role in mediating such diverse physiological processes as sight, smell, and the response to hormones and neurotransmitters, extremely little is known about these receptors at the structural level. A major breakthrough in 2007 was the determination of only the second GPCR crystal structure – that of a highly engineered human beta2-adrenergic receptor expressed in Sf9 insect cells where intracellular loop 3 was replaced with either antibody or T4 lysozyme to facilitate crystallization [15, 16]. Rhodopsin is perhaps the only GPCR that can be easily extracted from tissue, and this may explain why it was the first GPCR to have a detailed structure determined [17-19]. However, recent advances in the Khorana laboratory have led to the development of specific mammalian HEK293 cell lines for heterologous expression as well as methods for purification that yield milligram quantities of functional rhodopsin which is suitable for functional analysis and structural study [20-22]. We recently used same rho1D4-tag system for a one-step affinity purification of the human tetraspanin membrane protein CD81 [23]. Here we show that this system can be also adapted to facilitate the production and purification of another GPCR, the human olfactory receptor 17-4.

In order to carry out biochemical and structural analyses of olfactory receptors as well as engineer olfactory receptor-based biosensor devices, large quantities of receptors are required. Here we report inducible expression, large-scale production of human olfactory receptor hOR17-4. We bioengineered the synthetic hOR17-4 gene into 50 oligonucleotide fragments, self-assembled them through high cycle PCR and inserted the assembled gene into an inducible human embryonic kidney cell line (HEK293S). We
then induced its production using a combination of tetracycline and sodium butyrate. After systematic detergent screening, the zwitter-ionic detergent fos-choline-14 (FC14) was found to be most effective for solubilization and was subsequently used throughout the entire solubilization and purification. To our knowledge this is the first time an olfactory receptor has been purified from a mammalian cell line.

3.3 Results

Description of synthetic hOR17-4-rho gene

To fabricate synthetic gene constructs we utilized a PCR-based method of gene synthesis [24] that involves parsing the DNA sequence into a set of small overlapping oligonucleotides. During an initial assembly PCR, these oligos function as both primer and template, while the DNA polymerase successively builds longer and longer fragments with each round of PCR. A second amplification PCR then enriches for the full-length gene. This process has more recently become known as PCA, or polymerase construction and amplification. To assist in parsing the sequence into oligos, we used the online program DNAWorks (http://helixweb.nih.gov/dnaworks) developed by Hoover and Lubkowski [25].

The 312 amino acid sequence of wild-type human olfactory receptor hOR17-4 was obtained from GenBank and directly inputted into the software in protein mode. To adapt the resulting gene for use in our specific expression system and to facilitate purification and detection, the reverse-translated DNA sequence was human codon-optimized and appended with a C-terminal rho1D4 epitope tag (TETSQVAPA) [20-22]. The synthetic DNA corresponding to the hOR17-4 gene consisted of 1004 bp, encoding a
receptor protein of 323 amino acids (Figure 3.1). Other than the addition of the rho1D4 tag and linker, the hOR17-4 protein sequence is wild-type and completely full length. Noteworthy in the synthesis procedure was the requirement for a high number of PCR cycles (45) during the assembly PCR, presumably due to inefficiencies resulting from non-productive oligonucleotide mispairings.

Figure 3.1 - Codon-optimized hOR17-4 sequence. The DNA (A) and corresponding amino acid (B) sequence of synthetic hOR17-4 olfactory receptor gene. The DNA sequence was human codon-optimized and a mammalian Kozak ribosome binding site introduced upstream of the translation start site. Translation start and stop sites as well as restriction cloning sites are indicated. The engineered construct also contains a C-terminal tag (underlined) consisting of a glycine linker followed by a nonapeptide epitope for the monoclonal rho1D4 antibody.
Induction of hOR17-4 expression in stable HEK293S cell lines

To minimize the toxic effects of receptor overexpression, stable hOR17-4-inducible HEK293S cell lines were created using the Invitrogen T-REx tetracycline regulation system (Materials and Methods). This allowed large-scale cell culture batches to be grown and then, when desired, concerted production of fresh olfactory receptor to be induced in nearly 100% of the cells. Results for the induction of hOR17-4 expression in two of the subcloned HEK293 lines are shown in Figure 3.2A. Western immunoblotting using a monoclonal antibody against the rho1D4 tag revealed major immunoreactive bands at approximately 32 kD and 60 kD, which correspond in size to monomeric and dimeric forms of the hOR17-4 receptor. This size pattern has been reported previously for several solubilized olfactory receptors expressed in Sf9 and mammalian cells [26-28]. Larger molecular weight complexes were also present, presumably due to aggregation and precipitation of the receptor. As sample boiling only increases the precipitation, it is possible that increased temperatures (above 4°C) caused by the electrophoresis could be causing the receptor to partially aggregate. Thus these high MW species could be a side effect of the SDS-PAGE electrophoresis and not originally present in the solubilized receptor fractions.

The histone deacetylase inhibitor sodium butyrate has been demonstrated to synergistically enhance expression when used with tetracycline-regulated systems [21]. Induction of hOR17-4 in HEK293S cells using tetracycline in conjunction with sodium butyrate increased expression by approximately 4-5 fold over tetracycline alone at all time points tested (Figure 3.2B). There was no detectable expression of hOR17-4 in the
**Figure 3.2 - Induction of hOR17-4 in stable HEK293S cell lines.** (A) Stable inducible cell lines were generated in HEK293S cells using the T-REx system (Invitrogen) expressing hOR17-4 tagged with the rho1D4 tag (TETSQVAPA) at the C-terminus. The pooled cells were then subcloned, expanded, and then tested for induction in media supplemented with (+) or without (-) 1 µg/ml tetracycline for 48 hours. Levels of hOR17-4 were probed via SDS-PAGE western blotting using the rho1D4 antibody. Clones 1 and 5 showed the highest levels of induction while maintaining undetectable background levels in the absence of tetracycline. Clone 5 was selected for all subsequent experiments. (B) Addition of sodium butyrate enhances induced expression of hOR17-4. Inducible HEK293 was subjected to a dosage time course using the indicated concentrations of tetracycline and sodium butyrate. Samples were harvested, subjected to dot blot analysis (western blot against rho1D4), and the results quantified by spot densitometry. Tetracycline in conjunction with sodium butyrate increased expression by approximately 4-5 fold over tetracycline alone at all time points tested. Tetracycline was necessary to cause induction, as no expression was detected if sodium butyrate was used alone. Additionally, increasing tetracycline concentration to 2 µg/ml had no significant effect on induction levels.

absence of tetracycline or with sodium butyrate alone. Significant cell toxicity and death was observed in treatments combining tetracycline and sodium butyrate (5 mM) at the 48 and 72 hours time points. Treatment with sodium butyrate or tetracycline alone did not show this toxicity, indicating it to be a result of the high level expression induced by the drugs in conjunction.
To characterize this effect the samples were subjected SDS-PAGE analysis (Figure 3.3A). The gel showed two monomer band sizes, approximately 30 kD and 32 kD, suggesting distinct monomer forms (and corresponding dimer forms). It is possible these size discrepancies are due to differences in glycosylation of the receptor. Sodium butyrate addition for 24 hours showed a very large increase in expression over tetracycline alone, with the monomer band running at approximately 32 kD. Sodium butyrate addition for longer periods further increased total expression, however this caused the appearance of the additional monomer form running at 30 kD. The 48-hour time point contained roughly equal parts of both forms while the 72 hour time point consisted predominantly of the 30 kD form. To attempt to avoid the aforementioned toxicity we next performed a time course using a range of sodium butyrate concentrations (Figure 3.3B). We discovered that high levels of expression appeared to correlate with the appearance of the 30 kD form (and correspondingly smaller dimer form) and the observed cytotoxicity noted previously.

For subsequent purification experiments, we selected a treatment consisting of sodium butyrate (1 mM) with tetracycline (1 μg/ml) for 48 hours for attempts to purify primarily the 32 kD form. However, to further compare and characterize both forms, increasing the sodium butyrate concentration to 5 mM could be used, as relatively equal amounts of both monomer forms would be present.
Figure 3.3 - High-level induction correlates with a band shift of hOR17-4. (A) Time point samples (from Figure 2) were normalized to total protein content and subjected to SDS-PAGE western blotting against rho1D4. All lanes were treated with 1 μg/ml tetracycline with (+) or without (-) sodium butyrate (5 mM) for the time indicated. (B) hOR17-4-inducible HEK293S cells were subjected to a dosage time course using the indicated concentrations of sodium butyrate. All samples were co-treated with 1 μg/ml tetracycline. Samples were normalized to total protein content and subjected to SDS-PAGE western blotting against rho1D4.

**Purification of heterologously expressed hOR17-4**

We incorporated a C-terminal rho1D4-tag utilized by the Khorana lab for rhodopsin purification and detection [20-22], namely the nine C-terminal amino acid sequence (TETSQVAPA) against which a specific monoclonal antibody has been
generated (rho1D4). This tag has already facilitated early-stage immunoaffinity purifications of several GPCR membrane proteins [29-32]. For the initial immunoaffinity purification, we used CNBr-activated Sepharose 4B beads linked to the mouse monoclonal rho1D4 antibody to capture detergent solubilized receptors [21].

We first performed a small-scale purification from six 150 mm culture plates of hOR17-4-inducible HEK293S cells. The plates were treated as to have equal amounts of the 30 kD and 32 kD bands upon harvesting (tetracycline plus 5 mM sodium butyrate, 48 hours). Following a thorough wash procedure to remove non-specific impurities, the bound receptors were eluted by the addition of an excess of epitope peptide (TETSQVAPA). Fractions were subjected to SDS-PAGE followed by either western immunoblotting (Figure 3.4A) or total protein staining using highly sensitive SYPRO-Ruby (Figure 3.4B). The receptor was completely captured by the bead matrix, as no hOR17-4 was detected in the flow through by western blot. The bound OR eluted primarily in the first and second elution fractions. Total yield of hOR17-4 was approximately 30 µg per 150 mm plate. Mass spectrometry analysis on samples isolated from SDS-PAGE gel bands confirmed the identity of putative monomer and dimer protein bands as hOR17-4 receptor (see Materials and Methods).

To expand the system to potential milligram scale, fifty 150 mm culture plates of hOR17-4-inducible HEK293S cells were used. The plates were treated tetracycline plus 5 mM sodium butyrate, 48 hours). The total yield of hOR17-4 following immunoaffinity purification was 1.5 milligrams. To further purify the receptor and to remove the elution peptide, the hOR17-4 was subjected to size exclusion chromatography (SEC) using a gel
Figure 3.4 - Immunoaffinity purification of hOR17-4. Six 150 mm tissue culture plates of were grown to 90% confluence then induced and treated with tetracycline (1 μg/ml) and sodium butyrate (5 mM) for 48 hours. Plates were scrape harvested, solubilized, and the processed lysate subjected to immunoaffinity purification using rho1D4 antibody linked to sepharose beads for capture. Bound proteins were washed and then eluted using the nonapeptide TETSQVAPA. Samples were subjected to SDS-PAGE followed by either western immunoblotting with rho1D4 antibody (A) or total protein staining with Sypro Ruby (B). Black triangles indicate the 30 and 32 kD monomer forms. Tot, total lysate; FT, flow through; W, wash; E, elution.
Figure 3.5 - Full purification of hOR17-4 extracted from 50 x 150 mm culture plates. (A) Size exclusion chromatography (SEC) on immunoaffinity-purified hOR17-4. Absorbance was simultaneously recorded at 280 nm (black line, values on left axis) and 215 nm (grey line, values on right axis). The peaks indicated by numbers were pooled and concentrated. Peak 5 consists of the elution nonapeptide from the immunoaffinity purification. (B) Total protein staining of SEC peak fractions. Load is the original immunoaffinity purified sample applied to the chromatography column. Peak numbers refer to those designated in (A). Peak 3 contains monomeric hOR17-4 at >90% purity. Total monomer yield was 2.6 μg per 150 mm culture plate.
filtration column on an Äkta HPLC system. Column flow through was monitored by UV absorption (280 nm and 215 nm) and separated into fractions by an auto-fraction collector. As seen in Figure 3.5A, five distinct peaks were observed. The peak fractions were then pooled, concentrated and subjected to SDS-PAGE followed by total protein staining. As seen in Figure 3.5B, peak 3 contains monomeric hOR17-4 (>90% purity) while earlier peaks contained largely dimeric (peak 2) and aggregated/oligomerized (peak 1) forms. Peak 4 showed no visible protein and peak 5 corresponds to the residual elution peptide (TETSQVAPA) from the immunoaffinity purification. The final yield of purified hOR17-4 monomer was 0.13 milligrams (2.6 µg per plate). Thus, using a two-step procedure we have successfully obtained significant amounts of hOR17-4 in a highly pure form.

3.4 Discussion

In this study, we successfully developed methods for the construction of inducible mammalian cell lines that generate large quantities of olfactory receptor on demand. To our knowledge this is the first olfactory receptor to be purified from a mammalian cell line. Currently, we have demonstrated the production of the human olfactory receptor hOR17-4 in a stable tetracycline-inducible human embryonic kidney cell line (HEK293S). Expressed OR genes were fabricated from scratch using PCR-based gene synthesis (polymerase construction and amplification, a.k.a. PCA), which facilitated codon optimization for high level expression and attachment of affinity tags for detection and purification. The HEK293S cells can be grown and OR expression induced in adherent cultures (yield of ~30 micrograms/150-mm plate). Using methods originally
adapted from the production and purification of the GPCR rhodopsin [20-22] and further optimized (including the full-spectrum screening of over 70 detergents), the olfactory receptor is solubilized and extracted from the cells using fos-choline-14. The OR protein is then isolated using a two-step purification method (immunoaffinity followed by size-exclusion chromatography) which yields hOR17-4 monomer at greater than 90% purity (Figure 3.5B).

There have been a host of previous studies that have expressed and studied olfactory receptors in native and heterologous systems. However, to date there has not been a case where olfactory receptors have been overexpressed and purified to homogeneity in significant quantities. While purification of ORs has been attempted in bacterial [33] and Sf9 insect systems [26-27], these were unable to produce large quantities of native full-length olfactory receptor. Additionally, these organisms lack mammalian post-translational machinery and thus purified receptors may be improperly folded or missing critical modifications. It is known that most, if not all, GPCRs are glycosylated, and indeed the olfactory receptors have conserved N-linked glycosylation sites (Asn-X-Ser/Thr) at their N-terminus [34-35]. Several studies have indicated that loss of glycosylation can lead to improper folding and targeting, resulting in decreased function and compromised structure [36-37]. Loss of either N-terminal glycosylation site (Asn-2 or Asn-15) of rhodopsin is sufficient to cause loss of signal transduction despite no apparent change in localization or folding [38-39]. This indicates that purification of olfactory receptors from mammalian systems might be crucial for functional expression and purification.
**Characterization of purified hOR17-4**

It is interesting to note that the engineered hOR17-4-rho protein (with theoretical molecular mass of 36.2 kD) migrates slightly faster than expected on SDS-PAGE gels (30 kD and 32 kD). One possible explanation for this discrepancy is incomplete receptor denaturation by SDS since boiling the samples results in aggregation. Indeed, many other membrane proteins have been found to migrate faster than their actual size on SDS-PAGE, and this has been reported for other olfactory receptors as well [27, 33]. We believe both monomer forms of the expressed hOR17-4 receptor to be intact, full-length proteins as evidenced by: 1) detection by the C-terminal rho1D4 monoclonal antibody, 2) detection by custom anti-hOR17-4 polyclonal antibodies raised against the hOR17-4 N-terminus or C-terminus, 3) detection of a 30 kD N-terminally 6xHis-tagged hOR17-4 variant using anti-His antibodies (data not shown). Additionally, we see only the 32 kD form at low levels of induction while the 30 kD form only begins to appear at higher levels of induction. One explanation is that the receptor is glycosylated at low induction but that the toxicity associated with higher expression causes the receptor to begin to accumulate in the ER and not be properly processed.

The appearance of two distinct hOR17-4 monomer bands following purification could pose a problem for structural studies using X-ray crystallography, since typically a high degree of protein homogeneity is required for protein crystallization. We initially believed that it was possible to obtain primarily 32 kD form using lower levels of induction (Figure 3.3B). However, we have observed that the rho1D4 immunoaffinity purification appears to increase the proportion of 30 kD monomer form relative to 32 kD
form given the original treatment. For example, the treatment dosages for the small-scale (Figure 3.4) and large-scale (Figure 3.5) purifications were chosen (based on data in Figure 3.3b) to result in equal amounts of both forms or primarily 32 kD form, respectively. However, following immunoaffinity purification the resulting elution samples were enriched in the 30 kD monomer form. One hypothesis is that the 30 kD (potentially non-glycosylated form) binds more readily to the rho1D4-coupled bead matrix. Therefore, to obtain truly homogeneous hOR17-4 monomer it might be necessary to use even lower levels of induction where no 30 kD form exists. Another option would be to mutate the asparagine in the consensus N-glycosylation site of the receptor, located at amino acid position 5. We predict such a change to result in exclusive production of the 30 kD monomer form. However, the functional effect of abolishing glycosylation on the receptor is unknown.

There have been numerous reports in the literature citing difficulties in expressing functional olfactory receptors in heterologous systems. The main problem seems to be improper membrane targeting and resulting cytosolic localization of the majority of ORs [35-36, 40-41]. Several studies have been able to alleviate this problem by the addition of putative membrane import signals to the N-terminus of the olfactory receptor [42-43]. However, more recent studies [44-45] have expressed functional olfactory receptors in HEK293 cells without the use of such signal tags. We also investigated the use of an N-terminal “membrane import” sequence composed of the first twenty amino acids of bovine rhodopsin but found no significant increase in receptor yield or localization (results not shown).
Our largest adherent culture experiment consisted of fifty 150 mm tissue culture plates. While the total yield of crude hOR17-4 following immunoaffinity purification was 1.5 milligrams, the subsequent size exclusion chromatography and associated concentration steps reduced this yield to 0.13 milligrams of purified hOR17-4 monomer (2.6 μg per plate). The monomer was >90% pure, with the only other major contaminating band consisting of hOR17-4 dimer. While this represents a significant milestone, using adherent culture for milligram-scale purification of the receptor monomer poses a substantial challenge. However, the HEK293S cell line is capable of suspension culture, and we plan to scale-up purification yields by adapting the system to culture in a large volume (5-10 liter) liquid bioreactor. The high cell densities allowed should allow the production and subsequent purification of milligram quantities of olfactory receptors, which will be necessary for future experiments in determining receptor structure and function.

**General application for the purification of other olfactory receptors**

Here we show that a GPCR olfactory receptor gene can be designed, synthesized, placed into an inducible mammalian expression system and the resulting full-length protein purified to near homogeneity in a two-step process. In addition to our experiments on hOR17-4, we have constructed optimized synthetic genes for several mouse olfactory receptors and initial expression trials have proven successful. The small size of the rho1D4 tag and its extremely mild elution conditions provide a minimum of disruption to the purified protein. In contrast to previous attempts at OR purification using bacterial systems which required fusion to GST and truncated or mutated OR protein sequences.
[33], the system described here allows for the production of full-length wild-type OR. The application of this technique to other olfactory receptors could feasibly lead to a generalized method for obtaining large quantities of any olfactory receptor in a rapid and simple manner. Such methods could prove extremely useful in elucidating the structural and functional mechanism(s) of olfactory receptors and in their integration into OR-based biosensor devices.

3.5 Methods

Gene Construction

The protein sequence for hOR17-4 (also known as OR1D2) was obtained from GenBank (NCBI Accession # NP002539). DNAWorks online software (http://helixweb.nih.gov/dnaworks) was used in protein mode to design the synthetic gene and parse it into an oligonucleotide set. To adapt the synthetic hOR17-4 olfactory receptor gene for use in mammalian cell expression and purification, the following sequence modifications were made: i) human codon optimization; ii) addition of a C-terminal rho1D4 epitope tag (TETSQVAPA) preceded by a two glycine linker to facilitate detection and purification; iii) addition of a Kozak consensus sequence (GCCACCACC) immediately 5’ to the ATG start codon; iv) addition of an EcoRI restriction site at the 5’ end and a NotI restriction site at 3’ end of the gene to enable cloning into expression vectors. The synthetic hOR17-4 gene consisted of 1004 bp, of which 969 bp code for the 323 amino acid hOR17-4-GlyGly-rho1D4 protein. The designed oligonucleotide primers were purchased from IDT (Coralville, IA) with a maximum length of 45 bp. PCR-based gene synthesis was performed using a 2-step
assembly/amplification protocol [46] with the exception that the assembly PCR was run for 45 cycles. PCR reactions were then analyzed by gel electrophoresis and stained with ethidium bromide. Full-length product was excised, extracted, and then digested with the pertinent restriction enzymes. The genes were then ligated into the T-REx pcDNA4/To inducible expression plasmid (Invitrogen, Carlsbad, CA), sequenced, and a correct clones grown up using a MaxiPrep kit (Qiagen, Valencia, CA). The plasmid containing the optimized hOR17-4 gene was designated pcDNA4/To-hOR17-4-rho1D4.

**Generation of Stable Inducible Cell Lines**

HEK293S (suspension adapted HEK293 cells) containing the stable expression of pcDNA6/Tr (Invitrogen) that encodes the Tet repressor protein (TetR) had previously been generated and cloned [21]. HEK293S cell monolayers were grown in DMEM/F12 with GlutaMAX (Invitrogen catalog # 10565-042) supplemented with fetal bovine serum (10%), HEPES (15 mM), non-essential amino acids (0.1 mM), sodium pyruvate (0.5 mM), penicillin (100 units/ml), streptomycin (100 μg/ml) and grown at 37°C at 5% CO₂. All tissue culture and media components were purchased from Invitrogen unless otherwise noted. The pcDNA4/To-hOR17-4-rho1D4 plasmid was then transfected into these cells using Lipofectamine 2000 and after 48 hours cells were subjected to drug selection in 5 μg/ml blasticidin and 250 μg/ml zeocin for 2 weeks and then subcloned. 28 colonies were expanded and screened for inducible expression using media supplemented with or without 1 μg/ml tetracycline for 48 hours. Samples were then scrape harvested, solubilized in phosphate buffered saline (PBS) with 2% w/v Fos-Choline-14 (Anatrace, Maumee, OH) and Complete Protease Inhibitor Cocktail (Roche, Basel, CH) for 1 hour at
4°C. Expression was assayed via dot blotting and SDS-PAGE western blotting using the mouse monoclonal antibody rho1D4. Clone 5, the colony showing the best expression of hOR17-4 under induction conditions while maintaining undetectable expression without induction, was selected and expanded into large-scale culture and used for all subsequent experiments. The hOR17-4-inducible HEK293S cell line was maintained using 5 µg/ml blasticidin and 250 µg/ml zeocin.

**Cell Extract Preparation**

Buffers used were as follows: PBS buffer: 137 mM NaCl, 2.7 mM KCl, 1.8 mM KH₂PO₄, 10 mM Na₂HPO₄ (pH 7.4); Solubilization buffer: PBS containing Complete Protease Inhibitor Cocktail (Roche, Basel, Switzerland) and 2% wt/vol FC14; and Wash buffer: PBS containing 0.2% FC14; Elution buffer: Wash buffer containing 100 µM Ac-TETSQVAPA-CONH₂ elution peptide. The detergent FC14 was purchased from Anatrace (Maumee, OH). Sodium butyrate was purchased from Sigma (Saint Louis, MO).

For initial dosage and time course experiments, hOR17-4-inducible HEK293S cells were grown to 80-90% confluency at 37°C in 6-well tissue culture plates, treated as indicated and then scrape harvested into ice-cold PBS containing Complete Protease Inhibitor Cocktail. The hOR17-4 was then solubilized by resuspending the cell pellets in 150 µl solubilization buffer and rotating for 1 hour at 4°C. The non-solubilized fraction was then pelleted using at 13,000g for 30 minutes. The supernatant was then removed and analyzed by SDS-PAGE.
For purification experiments, up to fifty 150 mm tissue culture plates were used per experiment. Briefly, hOR17-4-inducible HEK293S cells were seeded at a density of 5 x 10^6 cells per 150 mm dish and grown for 72 hours at 37°C, at which point they reached 80-90% confluency. The cells were then induced with medium containing tetracycline (1 μg/ml) plus sodium butyrate (as indicated). After 48 hours, the cells were harvested by scraping (at 4°C) each plate into 2 ml PBS containing Complete Protease Inhibitor Cocktail. The cells were then pooled and snap frozen in liquid nitrogen and stored at -80°C until purification was carried out. On the day of purification, cells were thawed on wet ice and spun down by centrifugation at 4000g for 1 minute. All further steps were performed at 4°C unless noted. The hOR17-4 was then solubilized by resuspending the cells in solubilization buffer (1-2 ml per 150 mm plate) and rotating for 4 hours. The non-solubilized fraction was then pelleted using an ultracentrifuge at >100,000g for 30 minutes. The resulting supernatant was removed and put at 4°C. A small amount of supernatant (100 μl) was set aside, labeled “total lysate” and stored at -20°C. The remainder was directly applied to immunoaffinity purification.

**Immunoblotting and Total Protein Staining**

Samples were assayed via polyacrylamide gel electrophoresis (SDS-PAGE) under both reducing and denaturing conditions. Samples were prepared and loaded according to standard Novex gel protocols with the exception that the samples were incubated at room temperature prior to loading, as boiling caused membrane protein aggregation. Full Range Rainbow (GE Healthcare, Waukesha, WI) molecular weight marker was loaded as the protein size standard. Samples were resolved on Novex 10% Bis-Tris SDS-PAGE
gels (Invitrogen) were run using NuPAGE MOPS buffer at 100V and were subsequently transferred to a 0.45 μm nitrocellulose membrane and subjected to western immunoblotting using the rho1D4 as primary antibody, followed by a secondary HRP-linked goat anti-mouse IgG (Pierce, Rockford, IL) and detection using the ECL-Plus Kit (GE Healthcare). For total protein staining, SDS-PAGE gels were run as above, stained using SYPRO-Ruby (a more sensitive alternative to Coomassie; Invitrogen), and visualized by fluorescence using UV transillumination (excitation wavelength 300 nm). All western blot and SYPRO-Ruby images were captured using a Fluor Chem gel documentation system (Alpha Innotech, San Leandro, CA).

**Immunoaffinity Purification**

For immunoaffinity purification we utilized rho1D4 monoclonal antibody (Cell Essentials, Cambridge, MA) chemically linked to CNBr-activated Sepharose 4B beads (GE Healthcare). The rho1D4 elution peptide Ac-TETSQVAPA-CONH₂ was synthesized by CBC Scientific (San Jose, CA). Rho1D4-sepharose immunoaffinity purification has been described previously [21]. Briefly, the cell extract supernatant was mixed with rho1D4-coupled sepharose bead slurry (binding capacity 0.7 mg/ml) and rotated overnight at 4°C to capture the rho1D4-tagged olfactory receptors. The beads were then pelleted by centrifugation at 2000g for one minute and the supernatant collected, labeled as “flow-through” and saved for future analysis. The beads were then resuspended in 100 bead volumes of cold wash buffer (PBS + 0.2% Fos-Choline-14), rotated for 10 minutes at 4°C, then repelleted. A total of five washes were carried out and 100 μl of each sequential wash was saved for subsequent analysis. After the final wash, the beads were
pelleted again and transferred to a new tube for elution. A series of five elutions (each rotated 1 hour at room temperature) was then carried out, each using 1 bead volume of elution buffer (PBS + 0.2% FC14 + 100 µM TETSQVAPA peptide). Total protein concentration was measured using BCA assay (Pierce).

Mass Spectrometry

Immunoaffinity-purified samples of hOR17-4 were separated via SDS-PAGE, stained with SYPRO-Ruby and gel bands at 30 kD, 32 kD, and 60 kD were excised into sterile, methanol-rinsed microcentrifuge tubes. The samples subjected to trypsin digestion and the resulting fragments analyzed by Ion Trap LCMS for protein identification by the MIT Biopolymers Laboratory (Cambridge, MA). All bands were identified as hOR17-4, indicating monomeric and dimeric forms.

Size Exclusion Chromatography

For further purification, hOR17-4 proteins were subjected to gel filtration chromatography using a HiLoad 16/60 Superdex 200 column on an Äkta Purifier HPLC system (GE Healthcare). The column was first equilibrated using wash buffer (PBS + 0.2% w/v Fos-Choline-14). Pooled hOR17-4 elution fractions from the rho1D4 immunoaffinity purification were concentrated to 0.75 mg/ml using a 10kD MWCO filter column (Millipore, Billerica, MA) and then applied to the Äkta system. After loading, the column was run with wash buffer at 1 ml/min and column flow through monitored via UV absorbance at 280 nm and 215 nm. Protein fractions were collected using an automated fraction collector. Peak fractions were then pooled, concentrated and subjected
to SDS-PAGE and analysis via Sypro Ruby staining. Total protein concentration was measured using BCA assay (Pierce).

### 3.6 References


CHAPTER 4

PRODUCTION OF hOR17-4 USING SUSPENSION CELL CULTURES

NOTE: This chapter is an expansion of a manuscript submitted to Proceedings of the National Academy of Sciences (PNAS):

4.1 Abstract

In order to investigate the elusive biochemistry and molecular mechanism(s) of olfaction, we have developed a mammalian expression system for the large-scale production and purification of a functional olfactory receptor (OR) protein in milligram quantities. Here we report the study of human OR17-4 purified from a HEK293S tetracycline-inducible system. Scale-up of production yield was achieved through suspension culture in a bioreactor, which enabled the preparation of >10mg of monomeric hOR17-4 receptor following immunoaffinity and size exclusion chromatography, with expression yields reaching 3 mg per liter of culture medium. Final monomeric hOR17-4 protein was >90% pure, with the only major contaminant being dimeric receptor, and was thus directly applicable for subsequent characterization studies and crystallization screening for X-ray diffraction experiments. Thus this system constitutes a viable method for producing the large quantities of olfactory receptor necessary for structural and functional analyses.

4.2 Introduction

Animal noses have evolved the ability to rapidly detect a seemingly infinite array of odors at minute concentrations. The basis of this sensitivity are the olfactory (smell) receptors – a large, highly related class of sensory G-protein coupled receptors (GPCRs) that function together combinatorially to allow discrimination between a wide range of volatile and soluble molecules [1, 2]. As GPCRs, all olfactory receptors (ORs) are integral membrane proteins with seven transmembrane domains arranged in a barrel-like conformation. Despite the fact that olfactory receptors represent the largest class of known membrane proteins, no detailed structure exists for any OR and the functional
mechanisms of these amazing receptors remains unknown. The crucial first step to enable such pivotal biochemical and structural analyses is to engineer systems with the capacity to generate and purify milligram quantities of an olfactory receptor.

The human olfactory receptor 17-4 (hOR17-4, alternately known as OR1D2) is of particular interest since, in addition to olfactory neurons, it is also expressed on the midpiece of human spermatozoa [3]. Sperm expressing hOR17-4 were found to migrate towards known hOR17-4-responsive odorants such as bourgeonal, lilial, and floralozone [3]. Thus the receptor serves a dual role in that it recognizes odorants in the nose as well as plays a potential role in sperm chemotaxis and fertilization. Structural studies of hOR17-4 would not only provide information crucial to understanding the molecular mechanism(s) of olfaction but also have application to human reproduction.

To date, crystal structures exist for only three GPCR proteins: retina-purified bovine opsin [4-6] and highly engineered versions of human beta2-adrenergic receptor [7, 8] and turkey beta1-adrenergic receptor [9], both expressed in Sf9 insect cells. We have recently developed an OR expression system [10] using stably-inducible mammalian HEK293S cell lines by optimizing methods originally developed in the Khorana lab to generate milligram quantities of functional rhodopsin [11-13]. In adherent culture, this adapted rho-tag system was used express and purify monomeric hOR17-4 to >90% purity [10]. Here we show that this system can be scaled up using bioreactor culture to facilitate the production and purification of milligram amounts of hOR17-4. Key to the efficient extraction of hOR17-4 was a comprehensive screen of diverse detergents and the selection of zwitter-ionic fos-choline detergents as solubilizing agents, as all non-ionic detergents proved ineffective. The purified hOR17-4 protein was also structurally and
functionally characterized using several spectroscopy methods. To our knowledge this is the first time milligram quantities of functional olfactory receptor have been purified from any heterologous system.

4.3 Results

Construction of stable hOR17-4-inducible HEK293S GnTT cell lines

In Chapter 3, I described the fabrication of a synthetic hOR17-4 gene using PCR-based gene synthesis and the subsequent construction of stable HEK293S cell lines with tetracycline-inducible expression of the hOR17-4 receptor protein [10]. When induced with a combination of tetracycline and sodium butyrate, these HEK293S cells could generate over 30 micrograms of hOR17-4 per 15-cm plate. However, when assayed by SDS-PAGE, the receptor monomer migrated as a doublet at approximately 30 kD and 32 kD (the full-length rho-tagged hOR17-4 protein, with theoretical molecular mass of 36.2 kD, migrates slightly faster on SDS-PAGE gels). Our initial hypothesis was that the 32 kD band constituted a glycosylated form of the receptor. As heterogeneity could potentially interfere with future structural analysis and crystallization, we sought to achieve a homogeneous glycosylation pattern by porting the hOR17-4-inducible expression system into a HEK293S N-acetylglucosaminyltransferase I-negative (GnTI) cell line shown to produce homogeneously glycosylated rhodopsin [13]. During colony screening we isolated subclonal strains that exclusively expressed the slower migrating (32 kD) form of the receptor even under high-level expression (Figure 4.1a).
Figure 4.1 - Construction of hOR17-4-inducible HEK293S GnT1-/- cell lines for use in liquid bioreactor culture. Clones were tested for induction after 48 hours in plain media (-) or media supplemented with 1 ug/mL tetracycline (+) or tetracycline plus 5 mM sodium butyrate enhancer (++). Arrows indicate the position of the 32 kD and 30 kD monomer forms. (A) Levels of hOR17-4 were probed via SDS-PAGE western blotting against the rho tag (rho1D4 monoclonal antibody). Clones 3 and 8 show high levels of induction following the addition of sodium butyrate but have low levels of the potentially unglycosylated 30 kD monomer form of hOR17-4, unlike Clone 11 and previous clones in the HEK293S system. Clone 3 was selected for subsequent bioreactor experiments. (B) In order to investigate the potential N-linked glycosylation of hOR17-4, the consensus glycosylation sequence (-Asn-Gln-Ser-) was altered using site directed mutagenesis to change the asparagine at position 5 to glutamine (N5Q mutation). Following the generation of new stable hOR17-4(N5Q) inducible clones, the SDS-PAGE migration pattern of receptor monomer was compared to wild-type following induction. Mutation of the glycosylation site (N5Q) eliminated the upper form (32kD) of hOR17-4 monomer and only lower form (30kD) is present, indicating the size discrepancy is indeed due to glycosylation on Asn5.

Olfactory receptors possess a conserved N-linked glycosylation consensus sequence (Asn-X-Ser/Thr) at their N-termini [14], and the resulting glycosylation may be important for receptor functionality and proper folding [15], as is the case for other GPCRs [16]. To investigate whether the observed size discrepancy was due to N-linked glycosylation, we generated a stable cell line that expressed a mutated form of hOR17-4 (N5Q) where the consensus asparagine was replaced by a glutamine. This hOR17-4 N5Q
mutant ran solely at 30 kD with no 32 kD form present (Figure 4.1b), indicating that the 32 kD form of wild-type hOR17-4 is N-glycosylated on Asn5. As lack of glycosylation could potentially compromise receptor function, all subsequent experiments were performed using the wild-type hOR17-4 inducible cell line (Clone 3, Figure 4.1a).

**Detergent Screen and optimization of hOR17-4 solubilization from HEK293S cells**

Initial transient transfections into HEK293S cells and subsequent solubilization using the detergent dodecyl maltoside (DDM), which has been successfully used to solubilize several other GPCRs [12, 13, 17], revealed low levels of hOR17-4 protein yield in comparison to constructs encoding opsins. To investigate whether DDM was insufficient to solubilize hOR17-4, we performed a large detergent screen that included representatives from the non-ionic, zwitter-ionic, polar, and ionic detergent classes. Immunoblot analysis showed that the majority of commercially available detergents were poor choices for extracting the hOR17-4 GPCR protein from HEK293S cells (Figure 4.2). However, the fos-choline class of detergents proved highly effective and showed a clear relationship between chain length and solubilization yield, with fos-choline-16 (FC16) showing a >10-fold increase over DDM. However, the critical micelle concentration (CMC) of FC16 is so low (0.00053%) as to make any subsequent detergent exchange nearly impossible. Therefore, fos-choline-14 (FC14), with a CMC nearly 10x higher (0.0046%) was selected as the ideal solubilization agent. Importantly, FC14 showed greater hOR17-4 yield than solubilization with harsher ionic detergents such as sarcosine and deoxycholate. The fos-choline detergents are structurally related to
Figure 4.2. Detergent screen for optimal solubilization of hOR17-4 expressed in HEK293S cells. Expression of hOR17-4 was induced with tetracycline (1 μg/ml) and sodium butyrate (5 mM) for 48 hours and receptors solubilized in PBS containing detergents for 4 hours at 4°C. All detergents were used at a concentration of 2% wt/vol unless otherwise indicated. Relative solubilization corresponds to the fold increase over dodecyl maltoside (DDM) in solubilizing hOR17-4 monomer/dimer. Detergent abbreviations used are: DM, decyl maltoside; DDM, dodecyl maltoside; C/C, CHAPS (1%) and Cholesterol hemisuccinate (0.2%); OG, octyl glucoside; NG, nonyl glucoside; NP40, nonidet P40; DAO, n-decyl-N,N-dimethylamine-N-oxide; DDAO, n-dodecyl-N,N-dimethylamine-N-oxide; TDAO, n-tetradecyl-N,N-dimethylamine-N-oxide; DMDPO, dimethyldicyphosphate oxide; DDMG, n-decyl-N,N-dimethylglycine; DDDMG, n-dodecyl-N-N-dimethylglycine; sarcosine, sodium dodecanoyl sarcosine; DOC, deoxycholate.

Phosphatidylcholine (PC), a phospholipid and major constituent of the lipid bilayer of mammalian cells. Additional solubilization yield studies were carried out to determine the optimal FC14 concentration and extraction buffer (Materials and Methods). Addition of glycerol or increasing salt concentration was found to substantially decrease receptor yield (data not shown).
After solubilization of hOR17-4 from native cell membranes, we attempted to exchange the zwitter-ionic fos-choline for the milder non-ionic dodecyl maltoside (DDM) during the immunoaffinity bead immobilization, as DDM has been successfully used to keep many other GPCRs soluble. However, this resulted in a near total loss of hOR17-4 yield due to aggregation (data not shown), indicating that FC14 is crucial not only for OR extraction but to maintain the solubility of OR proteins in solution.

**Milligram-scale bioreactor production of hOR17-4 and subsequent purification**

The use of adherent culture for milligram-scale purification of the receptor monomer poses a substantial challenge, as many hundreds of plates would be required. As the HEK293S GnTI cell line is capable of suspension culture at high cell densities, we chose to scale-up production using a bioreactor and methods previously optimized for the milligram-scale production of bovine rhodopsin [12]. Each bioreactor run consisted of 1.25 – 1.40 liters of culture media (see Methods) inoculated with hOR17-4-inducible HEK293S GnTI cells at an initial density of 6 – 8 x 10^5 cells/mL. The media was supplemented on day 5 and hOR17-4 expression induced on day 6 using tetracycline and sodium butyrate (see Methods) and harvested 40 hours later (day 8). Cell density at time of induction was 6.0 x 10^6 cells/mL and had increased to 9.6 x 10^6 cells/mL at the time of harvest. Thus a single 1.25 liter bioreactor run produced 12 billion cells (a cell pellet of 16.3 grams), the equivalent of approximately 200 x 15-cm tissue culture plates.

Cell pellets from two separate bioreactor runs were combined and subjected to solubilization using FC14 followed by immunoaffinity purification using the rho1D4 monoclonal antibody conjugated to sepharose beads in order capture the rho-tagged
hOR17-4 protein. The eluate (containing 10.3 milligrams of hOR17-4) was then subjected to size exclusion chromatography to isolate the monomeric receptor fraction using gel filtration. Several peaks were observed (Figure 4.3a) and were found to correspond to aggregate, dimeric, and monomeric receptor (Figure 4.3b). The dimeric and monomer peaks eluted at 63.4 mL (0.51 CV) and 72.1 mL (0.58 CV), respectively, which was identical to that observed in our hOR17-4 purification using adherent culture [10]. Using gel filtration standards we estimated the apparent masses of the hOR17-4-detergent complexes at 140 kD (monomer) and 275 kD (dimer), indicating that each hOR17-4 protein is solubilized in a complex with approximately 270 molecules of FC14 (2.8 g of FC14 / g of protein). While this might seem high, the mass of bound detergent is greater than the micellar mass (47 kD for FC14 [18]) for most membrane proteins, including monomeric rhodopsin/DDM complexes (126 kD) [19]. Were our 140 kD complex to contain dimeric hOR17-4, each would be complexed with approximately 90 molecules of FC14, which is well below the aggregation number of 120 [18].

Peak fractions were collected, pooled and concentrated and subjected to SDS-PAGE (Figure 4.3b). The final yield of hOR17-4 monomer was 2.68 milligrams at >90% purity, the only other band visible being dimeric hOR17-4. Additionally, the putative dimer peak contained a total of 2.45 mg of largely dimeric receptor. Importantly, the appearance of monomer form in earlier peaks is likely due to the effects of SDS dissociating the dimeric and oligomeric receptor forms.

Following the initial milligram scale purification, we repeated the experiment using two additional bioreactor runs (2.5 liters of suspension culture). However, a large excess of rho1D4-sepharose beads (60 mL bead slurry, total binding capacity 42 mg) was
added to ensure complete capture of solubilized hOR17-4. Total yield of receptor following immunoaffinity chromatography was 30.5 milligrams, which led to the purification of 7.5 milligrams of hOR17-4 monomer following gel filtration chromatography. This constitutes nearly a three-fold increase over the first run and a yield of 3 mg/L for purified hOR17-4 monomer. As the resulting protein is at sufficiently high purity (>90%), it can immediately be concentrated and subjected to crystallization screening.

Figure 4.3 - Full purification of hOR17-4 from 2.65 liters of bioreactor cultured cells. (A) Size exclusion chromatography (SEC) on immunoaffinity-purified hOR17-4. Absorbance was recorded at 280 nm (black), 254 nm (red), and 215 nm (blue). Peaks 1-5 (indicated by numbers) were pooled and concentrated. The predicted monomer peak was pooled into an early fraction (4) and a late fraction (5). Peak 6 contains the nine amino acid elution peptide TETSQVAPA from the immunoaffinity purification. (B) Total protein staining of SEC peak fractions. Column fractions were collected and subjected to SDS-PAGE followed by staining with Sypro Ruby. Load is the original immunopurified sample applied to the chromatography column. Peak numbers refer to those designated in (A). Peaks 4 and 5 contain monomeric hOR17-4 at >90% purity.
4.4 Discussion

The major obstacle to structural and functional studies on membrane proteins is the notorious difficulty involved in expressing and purifying the large quantities of receptor protein sample required for such techniques as X-ray crystallography. A testament to this is the fact that as of September 2008, there are only 171 unique membrane protein structures known [20], which constitutes less than 1% of all known protein structures. There have been a host of previous studies that have expressed and studied olfactory receptors in native and heterologous systems. While purification of ORs has been attempted in bacterial [21] and Sf9 insect [22, 23] systems, these were unable to produce large quantities of native full-length olfactory receptor. The methods and results presented here constitute the first cell-based platform for the production of milligram quantities of purified olfactory receptor.

The appearance of two distinct hOR17-4 monomer bands following purification could pose a problem for structural studies using X-ray crystallography, since typically a high degree of protein homogeneity is required for protein crystallization. We initially believed that it was possible to obtain primarily 32 kD form using the new HEK293S GnTI cell line clones (Figure 4.1a). However, the rho1D4 immunoaffinity purification appears to significantly increase the proportion of 30 kD monomer form relative to 32 kD form, as seen in Figure 4.3b. One hypothesis is that the 30 kD (potentially non-glycosylated form) binds more readily to the rho1D4-coupled bead matrix. Therefore, to obtain truly homogeneous hOR17-4 monomer it might be necessary to perform purifications using the hOR17-4(N5Q) mutant cell line (Figure 4.1b), where only the 30 kD monomer form is produced. However, the functional effect of abolishing
glycosylation on this receptor is unknown. Several studies have indicated that loss of GPCR glycosylation can lead to improper folding and targeting, resulting in decreased function and compromised structure and stability [16]. Loss of either N-terminal glycosylation site (Asn-2 or Asn-15) of rhodopsin is sufficient to cause loss of signal transduction despite no apparent change in localization or folding [24, 25]. Additionally, mutating out the glycosylation site of a mouse OR (mOR-EG) was found to completely abolished its ability to localize to the membrane [15], indicating that this modification may be necessary for OR function. As other heterologous expression systems (bacterial, insect, etc) lack mammalian post-translational machinery, purification of olfactory receptors from mammalian systems might be crucial for functional production. It is worthwhile to note that retina-derived bovine opsin, which has been crystallized, is not heterogeneously glycosylated but contains at least three N-glycan variants (70% Man₃GlcNAc₃, 10% Man₄GlcNAc₃, and 20% Man₅GlcNAc₃) [26].

In addition to the N-glycosylation site at Asn5, hOR17-4 has a potential site at Asn195. However, this residue does not appear to be heavily glycosylated in this system as evidenced by: i) the N5Q mutation causes a complete shift in mobility from 32 kD (glycosylated) to the 30 kD form, which corresponds to the mass of the deglycosylated hOR17-4 protein [23]; ii) mass spectrometry analysis did not detect glycosylation at this site but did detect unglycosylated peptides containing Asn195; iii) the consensus sequence is at the hypothetical EC2/TM5 border and thus is not likely to have the flexibility required for N-linked glycosylation. We did see a minor band running at ~33 kD for both the wild-type and N5Q mutant versions of hOR17-4 (Figure 4.3b, Figure 4.1b), which we have not yet ruled out as due to potential Asn195 glycosylation. Since
this band does not appear to shift with N5Q mutation, were this band glycosylated it would be on Asn195 alone (and not both sites). Should this be the case, a double mutant (N5Q, N195Q) might be advantageous, as it would eliminate both forms.

Here we demonstrated the utility of the fos-choline-based detergents (most notably FC14) in extracting and solubilizing olfactory receptors. Fos-choline-12 was recently found to refold the integral membrane protein diacylglycerol kinase and maintain its functional state [27]. Additionally, Fos-choline-14 was used in the crystallization of the E. Coli mechanosensitive ion channel MscS for a successful structure [28]. The promise of this detergent class in future membrane protein research is underscored by our recent findings which identified the fos-choline series as the best detergent class for extracting and solubilizing the human chemokine receptors CCR3, CCR5, CXCR4 and CX3CR1 [29].

In summary, we have developed methods for the construction of inducible mammalian cell lines that generate milligram quantities of olfactory receptor on demand. Currently, we have demonstrated the production of >10 milligrams of full-length human olfactory receptor hOR17-4 in a stable tetracycline-inducible human embryonic kidney cell line (HEK293S GnTI). Using methods originally adapted from the production and purification of the GPCR rhodopsin [11-13] and further optimized (including the full-spectrum screening of over 45 detergents), the olfactory receptor is solubilized and extracted from the cells using fos-choline-14. The OR protein is then isolated using a two-step purification method which yields hOR17-4 monomer at greater than 90% purity (Figure 4.3b). The system described is capable of generating up to 3 mg / liter of monomeric hOR17-4, the yield of which approaches that of rhodopsin and rhodopsin
mutants when similarly expressed [12]. The application of this technique to other olfactory receptors could feasibly lead to a generalized method for obtaining large quantities of any olfactory receptor in a rapid and simple manner. Such methods could prove extremely useful in elucidating the structural and functional mechanism(s) of olfactory receptors and in their integration into OR-based biosensor devices.

4.5 Methods

Buffers and solutions

Buffers used were as follows: PBS buffer: 137 mM NaCl, 2.7 mM KCl, 1.8 mM KH₂PO₄, 10 mM Na₂HPO₄ (pH 7.4); Solubilization buffer: PBS containing Complete Protease Inhibitor Cocktail (Roche, Basel, Switzerland) and 2% wt/vol FC14; Wash buffer: PBS containing 0.2% FC14; Elution buffer: Wash buffer containing 100 µM Ac-TETSQVAPA-CONH₂ elution peptide. The detergent FC14 was purchased from Anatrace (Maumee, OH). All tissue culture and media components were purchased from Invitrogen unless otherwise noted. Sodium butyrate was purchased from Sigma (Saint Louis, MO).

Site Directed Mutagenesis

The pcDNA4/To-hOR17-4-rho plasmid, containing a synthetic, codon optimized hOR17-4-rho gene construct [10], was mutated to eliminate the hOR17-4 N-linked glycosylation consensus sequence. The codon corresponding to hOR17-4 residue Asn5 was converted from AAC (Asn) to CAG (Gln) by site directed mutagenesis using the Quikchange kit (Stratagene) according to the manufacturer’s instructions. The oligonucleotides used were: 5’-CCATGGACGGAGGCAGCAAAGCGAGGGCAG-3’
and 5’-CTGCCCTCGCTTTGCTGGCCTCCGTCCATGG-3’ (bottom). The successfully mutated plasmid was designated pcDNA4/To-hOR17-4(N5Q)-rho.

Generation of Stable hOR17-4-Inducible Cell Lines

HEK293S GnTI (suspension adapted, N-acetylglucosaminyltransferase I-negative HEK293) cells containing the stable expression of pcDNA6/Tr (Invitrogen) that encodes the Tet repressor protein (TetR) had previously been generated and cloned [13]. Adherent HEK293S GnTI cell monolayers were grown in DMEM/F12 with GlutaMAX (Invitrogen catalog # 10565-042) supplemented with fetal bovine serum (10%), HEPES (15 mM), non-essential amino acids (0.1 mM), sodium pyruvate (0.5 mM), penicillin (100 units/ml), streptomycin (100 μg/ml) and grown at 37°C at 5% CO₂. Either the pcDNA4/To-hOR17-4-rho plasmid, containing a synthetic, codon optimized, rho-tagged hOR17-4 gene construct (hOR17-4-GG-TETSQVAPA) [10] or the pcDNA4/To-hOR17-4(N5Q)-rho plasmid were then transfected into these cells using Lipofectamine 2000 and after 48 hours cells were subjected to drug selection in 5 μg/ml blasticidin and 50 μg/ml zeocin for 2 weeks and then subcloned. For each variant, at least 25 colonies were expanded and screened for inducible expression in 6-well plates using plain media (control) or media supplemented with 1 μg/mL tetracycline or tetracycline plus 5 mM sodium butyrate. 48 hours post-treatment, samples were scrape harvested into ice-cold PBS containing Complete Protease Inhibitor Cocktail (Roche, Basel, CH). The hOR17-4 was then solubilized by resuspending the cell pellets in 150 μl solubilization buffer and rotating for 1 hour at 4°C. The non-solubilized fraction was then pelleted by centrifugation at 13,000g for 30 minutes. The supernatant was then removed and OR
expression assayed via dot blotting and SDS-PAGE western blotting using the mouse monoclonal antibody rho1D4. For wild-type hOR17-4, Clone 3 was selected and expanded into large-scale culture and used for all subsequent experiments. The hOR17-4-inducible HEK293S GnTI cell lines were maintained using 5 µg/ml blasticidin and 25 µg/ml zeocin.

**Detergent Screening**

For initial solubilization trials, the wild-type pcDNA4/To-hOR17-4-rho plasmid was transiently transfected into 150mm tissue culture plates of HEK293S cells using Lipofectamine 2000. After 48 hours, cells were scrape harvested and pooled. Cells were spun down and resuspended in ice cold PBS containing Complete Protease Inhibitor Cocktail at a density of $2 \times 10^7$ cells/ml and then aliquotted into microcentrifuge tubes (120µl each). 30µl of detergent was then added from stock solutions (10% wt/vol) such that the final concentration was 2%, except where noted. Care was taken not to vortex or pipette-mix the samples after detergent was added in order to avoid breaking cell nuclei. Samples were then rotated at 4°C for 4 hours before being centrifuged at 13,000g for 30 minutes to pellet insoluble material. Supernatants were then removed and subjected to dot blot and SDS-PAGE analysis using the rho1D4 mAb. As dot blotting also detects aggregated/oligomerized receptor, the solubilization was quantified via SDS-PAGE western blotting as the total amount of monomeric and dimeric hOR17-4 present, as determined by spot densitometry. Relative solubilization corresponds to the fold increase over dodecyl maltoside (DDM). All detergents were purchased from Anatrace (Maumee, OH) except digitonin, which was purchased from Sigma.
Growth of HEK293S GnTr suspension cultures in a bioreactor

Suspension culture was carried out in a Celligen Plus bioreactor (New Brunswick Scientific, Edison, NJ) according to the protocol developed by Reeves, et al [12]. Media formulation was identical except that Primatone RL-UF used at 3.0 g/L. On day 0, the bioreactor media was inoculated with cells trypsinized from 6-9 confluent 15-cm tissue culture plates such that the inoculation density was $6 \times 10^5$ cells/mL. The four-gas mixture (air, O$_2$, N$_2$, and CO$_2$) was supplied by direct sparge only. Gas flow rate was initially 21 ml/min but was increased as needed throughout the run to regulate the pH and dissolved oxygen. If required, 20 mL of 8% sodium bicarbonate was added to increase the buffering capacity of the media if the gas mixture was unable to adequately regulate the pH. On day 5, the reactor was supplemented with 30 mL of 10% Primatone RL-UF and 10 mL of 20% glucose. On day 6, the expression of hOR17-4 was induced by the addition of tetracycline (2 μg/mL) and sodium butyrate (2 mM) and the cells harvested 40 hours post-induction. Harvested cells were pelleted and washed with cold PBS containing complete protease inhibitor cocktail. The cell pellets were then weighed and snap-frozen in liquid nitrogen and stored at -80°C until purification was carried out. On the day of purification, cells were thawed on wet ice and spun down by centrifugation at 4000g for 1 minute. All further steps were performed at 4°C unless noted. The hOR17-4 was then solubilized by resuspending the cells in solubilization buffer (12.5 ml per gram of cell pellet) and rotating for 4 hours. The non-solubilized fraction was then pelleted using an ultracentrifuge at >100,000 g for 45 minutes. The resulting supernatant was removed and put at 4°C. A small amount of supernatant (100 μl) was set aside, labeled
“total lysate” and stored at -20°C. The remainder was directly applied to immunoaffinity purification.

**Immunoblotting and Total Protein Staining**

Samples were assayed via polyacrylamide gel electrophoresis (SDS-PAGE) under both reducing and denaturing conditions. Samples were prepared and loaded according to standard Novex gel protocols with the exception that the samples were incubated at room temperature prior to loading, as boiling caused membrane protein aggregation. Full Range Rainbow (GE Healthcare, Waukesha, WI) molecular weight marker was loaded as the protein size standard. Samples were resolved on Novex 10% Bis-Tris SDS-PAGE gels (Invitrogen) using NuPAGE MOPS buffer, run at 100V and were subsequently transferred to a 0.45 μm nitrocellulose membrane and subjected to western immunoblotting using the rho1D4 as primary antibody, followed by a secondary HRP-linked goat anti-mouse IgG (Pierce, Rockford, IL) and detection using the ECL-Plus Kit (GE Healthcare). For total protein staining, SDS-PAGE gels were run as above, stained using SYPRO-Ruby (a more sensitive alternative to Coomassie; Invitrogen), and visualized by fluorescence using UV transillumination (excitation wavelength 300 nm). All western blot and SYPRO-Ruby images were captured using a Fluor Chem gel documentation system (Alpha Innotech, San Leandro, CA).

**Immunoaffinity Purification**

For immunoaffinity purification we utilized rho1D4 monoclonal antibody (Cell Essentials, Cambridge, MA) chemically linked to CNBr-activated Sepharose 4B beads
(GE Healthcare). The rho1D4 elution peptide Ac-TETSQVAPA-CONH₂ was synthesized by CBC Scientific (San Jose, CA). Rho1D4-sepharose immunoaffinity purification has been described previously [12]. Briefly, the cell extract supernatant was mixed with rho1D4-coupled sepharose bead slurry (binding capacity 0.7 mg/ml) and rotated overnight at 4°C to capture the rho-tagged olfactory receptors. The beads were then pelleted by centrifugation at 2000g for one minute and the supernatant collected, labeled as "flow-through" and saved for future analysis. The beads were then resuspended in 100 bead volumes of cold wash buffer (PBS + 0.2% fos-choline-14), rotated for 10 minutes at 4°C, then repelleted. The procedure was repeated until the supernatant had a UV absorbance at 280nm of less than 0.01. After the final wash, the beads were pelleted again and transferred to a new tube for elution. A series of elutions (each rotated 1 hour at room temperature) was then carried out, each using 1 bead volume of elution buffer (PBS + 0.2% FC14 + 100 μM TETSQVAPA peptide). Elutions were repeated until the UV absorbance at 280nm was less than 0.1. Total protein concentration was measured by BCA assay (Pierce).

**Size Exclusion Chromatography**

For further purification, hOR17-4 proteins were subjected to gel filtration chromatography using a HiLoad 16/60 Superdex 200 column on an Äkta Purifier HPLC system (GE Healthcare). The column was first equilibrated using wash buffer (PBS + 0.2% w/v fos-choline-14). Pooled hOR17-4 elution fractions from the rho1D4 immunoaffinity purification were concentrated to 3 mg/ml using a 10kD MWCO filter column (Millipore, Billerica, MA) and then applied to the Äkta system. After loading, the
column was run with wash buffer at 0.3 ml/min and column flow through monitored via UV absorbance at 280 nm, 254 nm and 215 nm. Protein fractions were collected using an automated fraction collector. Peak fractions were then pooled, concentrated and subjected to SDS-PAGE and analysis via Sypro Ruby staining. Total protein concentration was measured using BCA assay (Pierce). The molecular mass of hOR17-4-detergent complexes was estimated by calibrating the column with gel-filtration standard mixture (Bio-Rad). Molecular mass was correlated to retention volume using a power law curve-fit.

4.6 References


CHAPTER 5

CHARACTERIZATION OF PURIFIED hOR17-4

NOTE: This chapter is an expansion of a manuscript submitted to Proceedings of the National Academy of Sciences (PNAS):

5.1 Abstract

To characterize the previously purified hOR17-4 receptor protein, we subjected it to several spectroscopic modes of analysis in order to determine if the solubilized receptor was correctly folded and retained functional activity. Analysis by far-UV circular dichroism showed the receptor to have correct secondary structure (~50% alpha-helical), while near-UV spectra were indicative of a defined tertiary structure. Using a surface plasmon resonance (SPR) assay, extracted hOR17-4 specifically bound its known odorants lilial and floralozone in vitro, while denatured receptor had no binding activity. At least 50% of the receptors bound odorants as estimated from bound SPR mass. The synthetic hOR17-4 also recognized specific odorants on the surface of HEK293S cells as determined by calcium-release assay. Additionally, analysis using mass spectrometry confirmed the presence of several post-translational modifications. Thus, the engineered mammalian expression system and optimized purification protocols developed during this thesis are capable of producing milligram quantities of correctly folded and functionally active olfactory receptor.

5.2 Introduction

No method currently exists with which to ascertain the functionality of detergent solubilized (non-membrane bound) olfactory receptor proteins. We subjected our purified hOR17-4 receptor to a range of spectroscopic and analytical techniques in order to confirm correct folding, function, and post-translational modifications. Additionally, we tested the functionality of the protein in the native cell membrane in order to confirm that
the rho-tagged, synthetic gene truly encoded a functional protein capable of normal G protein-mediated signal transduction. The techniques applied were:

i. Circular dichroism (CD) spectroscopy — quantitates secondary structural components and provides information on tertiary structure through differential absorption of circularly-polarized light (see section 1.5). Changes in CD spectra can indicate changes in secondary structure and protein folding.

ii. Fluorescence spectroscopy — light reemitted from excited states of tryptophan and other aromatic amino acid residues at characteristic wavelengths. Alterations in emission spectra can be used to measure changes in protein conformation.

iii. Surface plasmon resonance (SPR) — detects molecular interactions in real-time due to changes in refractive index near the surface (see section 1.6).

iv. Liquid Chromatography-Mass Spectrometry — protein is digested using a protease and resulting fragments are size separated and the masses precisely determined so that the peptides can be identified. Post-translational modifications can also be determined.

5.3 Results

5.3.1 Circular dichroism and fluorescence spectroscopy

We asked if the purified, FC14-solubilized hOR17-4 retained proper structure and functionality. Our prediction for hOR17-4 secondary structure, based on structural modeling [1] and transmembrane domain calculations [2], was 47% alpha-helix. When
Figure 5.1 – Circular dichroism analysis of purified hOR17-4. Purified hOR17-4 monomer was analyzed by both circular dichroism spectroscopy and fluorescence spectroscopy. (A) Far-UV CD spectrum of hOR17-4 displaying correct secondary structure (49% alpha helix). Spectrum shown is the average of 5 replicate scans. Mean residue ellipticity $[\theta]$ has units of degree$\cdot$cm$^2$*dmol$^{-1}$. (B) Near-UV CD spectrum of hOR17-4 showing distinct tertiary structure peaks. Spectrum shown is the average of 3 replicate scans. Functional bovine rhodopsin has a similar peak in this region, while non-functional opsin mutants show flat spectra characteristic of a misfolded globular state.

subjected to far-UV circular dichroism spectroscopy, the monomeric hOR17-4 displayed a spectrum characteristic of alpha helix with minima at 208 nm and 222 nm (Figure 5.1a). Analysis of the spectrum using the K2D algorithm [3] returned values of 49% alpha helix, 18% beta sheet, and 33% random coil content, confirming a correct hOR17-4 secondary structure.

To probe tertiary structure, we subjected hOR17-4 to near-UV circular dichroism spectroscopy. Since this technique requires milligram quantities of pure protein, this is to our knowledge the first instance of a near-UV spectrum for an olfactory receptor (Figure 5.1b). Several significant peaks were observed which suggest a defined tertiary structure. Wild-type opsin has similar near-UV peaks in this region, whereas functionally inactive opsin mutants showed flat spectra, indicating misfolded protein [4]. Characterization by tryptophan fluorescence spectroscopy using excitation at 280 nm showed an emission
maximum at 335 nm (Figure 5.2), which is similar to the value experimentally
determined for rat OR5 of 328 nm [5].

![Fluorescence spectroscopy on purified hOR17-4 monomer](image)

**Figure 5.2** - **Fluorescence spectroscopy on purified hOR17-4 monomer.** Excitation at 280 nm resulted in a clear emission peak at 335 nm. The hOR17-4 amino acid sequence has 2 tryptophans, 17 tyrosines, and 19 phenylalanines. The spectrum shown is the average of 3 replicate scans. This agrees with previous characterization of rat OR5 receptor showing a 328 nm emission peak from excitation at 280 nm.

### 5.3.2 Calcium- release assays of cell surface expressed hOR17-4

We also confirmed that our synthetic hOR17-4 displayed wild-type function in the HEK293S cell membranes. The functional activity and specificity of induced hOR17-4 was measured on the surface of HEK293S cells by calcium- release assay and visualized using time-lapse confocal microscopy. In our heterologous HEK293S system, ORs can signal through the inositol triphosphate (IP₃) pathway to release intracellular Ca²⁺ from the ER by signaling through the “promiscuous” G-protein Gₒ₅. Induced cells responded to the specific odorant bourgeonal at concentrations as low as 1 µM (Figure 5.3). Odorant response could be blocked by co-application of the hOR17-4 antagonist undecanal. No response seen for the non-specific odorants octanal and anithole (data not shown).
Figure 5.3 - Calcium-release assays of cell surface-expressed hOR17-4. hOR17-4 expressed in a stable inducible HEK293S cell line exhibits specific activation by its cognate ligand bourgeonal. (A) Transient changes of the cytosolic Ca\(^{2+}\) concentration were recorded with confocal microscopy using Fura-Red (Ex 488 nm/Em 560 nm) as a fluorescent Ca\(^{2+}\) indicator. The decrease of the fluorescence signal induced by receptor activation in response to bourgeonal (100 μM) corresponds to an increase of the cytosolic Ca\(^{2+}\) concentration. The application of 200 μM adenosine triphosphate (ATP) served as a control of HEK293S cell excitability. (B) In a randomly selected field of view, Fura Red fluorescence intensities of odorant-induced Ca\(^{2+}\) responses were recorded on four individual cells (1, 2, 3, 4) as a function of time. hOR17-4 induces transient Ca\(^{2+}\) signaling to consecutive stimulations by bourgeonal (1 μM; 100 μM). Arrows indicate the time point of odorant application. The preincubation (black bar) with the hOR17-4 antagonist undecanal (100 μM) inhibited hOR17-4 activation by bourgeonal (100 μM) during co-application (arrow) with undecanal (100 μM). After subsequent odorant washout, cells were again excitable with bourgeonal (100 μM).
5.3.3 Analysis of OR17-4 odorant binding using surface plasmon resonance

To assay the functionality of purified (non-membrane-bound) hOR17-4, we developed an assay method using surface plasmon resonance (SPR) to demonstrate that the solubilized receptor retains selectivity in binding odorant ligands in a concentration-dependent manner. The rho1D4 monoclonal antibody was covalently attached to the dextran surface of a Biacore CM4 chip using standard amine-coupling chemistry. The hOR17-4 receptor protein was then non-covalently bound to the antibody via its C-terminal rho tag (TETSQVAPA). Figure 5.4 displays the SPR readout during the surface assembly on the CM4 chip. Odorant ligands were then applied and odorant binding detected in real time via the associated mass increase. Solubilized hOR17-4 receptors bound the odorant bourgeonal in a concentration-dependent manner (Figure 5.5), and a dissociation constant was fitted to the data using the steady-state equation [6]:

\[
y = \frac{\frac{[L]}{max}}{\frac{[L]}{max} + \frac{[L]}{0}}
\]

Where \( y \) is the fractional saturation and \([L]\) denotes the odorant concentration. This equation also assumes no ligand depletion ( \([L] \approx [L]_0\) ), which is valid as the concentration of odorant far exceeds that of the immobilized olfactory receptor. The fitted \( K_d \) was 9.5 \( \mu \text{M} \), which is a rather weak interaction, as many other receptors have ligand affinities in the nanomolar range. Solubilized hOR17-4 receptors also bound the specific odorants.
Figure 5.4 – Preparation of SPR chip surface for odorant binding assay. (A) Chip surface is activated, followed by covalent attachment of rho1D4 mAb by amine chemistry, followed by a blocking step to react all remaining ester groups. The OR protein is then applied and immobilized on the surface via its C-terminal rho-tag epitope (B). The asterisk (*) denotes the zero RU point for the subsequent odorant binding assays. The inset shows a depiction of the assembled chip surface.
lilial and floralozone in a dose-dependent manner (Figure 5.6). Due to low odorant solubility, the dissociation constant for these odorants could not be rigorously determined, but was approximately in the low micromolar range. No binding was detected for the non-specific odorant sulfurol acetate and no odorants were bound when the receptor was denatured using guanidium HCl (data not shown). Thus, these results indicate that hOR17-4 receptor retains its functionality in the solubilized state.

Figure 5.5 – SPR determination of hOR17-4 affinity for bourgeonal. (A) Immobilized hOR17-4 binds increasing concentrations of bourgeonal in a dose-specific manner. These measurements were used to fit a dissociation constant of $K_d = 9.5 \, \mu M$ using the steady-state relation (B).
Figure 5.6. Functional assay of purified hOR17-4 functionality using surface plasmon resonance (SPR). Detection of functional OR-odorant interactions was monitored in real time using a Biacore A100. The hOR17-4 was adsorbed onto the SPR chip surface using the rho1D4 monoclonal antibody, and spectra were recorded following odorant application. Immobilized receptor bound increasing concentrations of the specific odorant ligands lilial (A) and floralozone (B). Odorant binding curves shown are: blank control (black), 5 μM (red), 10 μM (light blue), 20 μM (dark blue), and 40 μM (green). No response was seen for the control odorant sulferol acetate (result not shown). Guanidium HCl-denatured hOR17-4 did not bind any odorants (result not shown). All results are simultaneously subtracted from a reference channel containing mAb and blank buffer.

5.3.4 Quantitation of hOR17-4 activity

Although we confirmed the purified hOR17-4 to be functional via SPR, it is possible that only a small fraction of the receptor is truly functional and binding odorants while the remainder is inert. Such concerns are warranted given the fact that the receptor is solubilized by a zwitter-ionic detergent, which, as a class, are generally harsher than non-ionic detergents, and it is possible that the bulk of the receptor is not functionally active. Thus we performed quantitative analysis on the SPR data in an effort to determine what percentage of the purified hOR17-4 receptor is binding specific odorants, based on
bound mass detected by SPR. Such techniques have been previously used to quantitate ligand binding for the GPCR chemokine receptors CXCR4 and CCR5 [7].

In SPR, the response units (RUs) are a measurement of the dielectric constant in the region near the surface, which is directly related to the index of refraction (n). Thus, SPR signals are not truly proportional to mass bound but instead are governed by the equation [8]:

\[ RU = X \left( \frac{\partial n}{\partial C} \right) \cdot C \]  

Where X is a proportionality constant, C is concentration, and the quantity (dn/dc) represents the refractive index increment (also known as differential refractive index). In practical terms, (dn/dc) represents how much the refractive index changes for a given change in concentration. In SPR, it is known that for proteins, 1 RU ≈ 1 pg/mm² and that the (dn/dc) is 0.18 mL/g [8]. This relation hold for a large range of protein types and sizes. However, it does not necessarily hold for small molecules such as odorants. This protein relation can, however, be used to rearrange the above RU equation to a more useful form:

\[ C = \left( \frac{\partial n}{\partial C} \right)_{\text{odorant}} \cdot \frac{RU_{\text{odorant}}}{RU} \]  

Using this equation, we only need the (dn/dc) for the odorant ligand in order to calculate its concentration (and thus the number of molecules bound by the immobilized hOR17-4). Since this data is unavailable for the odorants in question, we can obtain a rough estimate by using the Dale-Gladstone relation [9], which is typically accurate to within a few percent for a variety of small molecules, including odorants such as benzene [10] and n-alkanes [9]:

\[ ]
Where \( n \) is the index of refraction of the odorant, \( n_{\text{solvent}} \) is the index of refraction of the solvent (1.33 for the HBS saline buffer [11]), and \( \rho \) is the odorant density. These odorant properties are readily available on web-based chemical databases, and values were obtained from [12] for lilial and [13] for floralozone. The values obtained were used to calculate the odorant (dn/dc) values using equation <4> and these were in turn used in equation <3> to calculate the concentration of odorant bound. The concentration of immobilized hOR17-4 was also calculated using equation <2> using the standard protein relation. Results for lilial and floralozone are shown in Table 3.

### Table 3 – Quantitation of purified hOR17-4 receptor functionality.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>RUs bound</th>
<th>(dn/dc)</th>
<th>fmol/mm² bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>hOR17-4</td>
<td>2800</td>
<td>0.18</td>
<td>73</td>
</tr>
<tr>
<td>Lilial</td>
<td>8.1</td>
<td>0.172</td>
<td>41</td>
</tr>
<tr>
<td>Floralozone</td>
<td>7.6</td>
<td>0.187</td>
<td>38</td>
</tr>
</tbody>
</table>

We can see that the calculated odorant (dn/dc) values are nearly equal to that of protein. The majority of immobilized hOR17-4 receptors bound specific odorants (56% bound lilial and 52% bound floralozone). Additionally, since odorant solubility limits our these ligand concentrations, we are not capable of completely saturating the receptors (see equation <1>) and our estimate is that only ~80% saturation is possible if the receptors were 100% active, given a \( K_d \approx 10 \) μM. Thus we feel confident in stating that
the majority (>50%) of the purified, detergent-solubilized hOR17-4 receptors are functionally active.

5.3.5 Characterization of hOR17-4 post-translational modifications

Olfactory receptors are believed to possess two disulfide bonds between four conserved cysteine residues located in extracellular loops 1 and 2 (EC1 and EC2) [14]. For hOR17-4, these two bonds are Cys97-Cys179 (EC1-EC2) and Cys169-Cys189 (EC2-EC2). Analysis by mass spectrometry (chymotrypsin digest followed by LC-MS) confirmed the presence of one of the two disulfide bonds (Cys169-Cys189) predicted for olfactory receptors (Table 3). No corresponding unlinked peptides were detected, indicating homogeneity. Presence of the remaining disulfide bond (Cys97-Cys179) was unconfirmed, as no corresponding peptides (disulfide-linked or unlinked) were detected, presumably due to the resistance of the detergent-solubilized OR to complete protease digestion.

Additionally, the N-linked glycosylation present on Asn-5 was determined to be Man₅GlcNac₃. To our knowledge this is the first characterization of olfactory receptor glycosylation. This hexasaccharide is also the predominant glycosylation seen in retina-derived bovine opsin on Asn-2 and Asn-15 [15]. However, it is worthwhile to note that rhodopsin heterologously prepared using the HEK293S GnTI- cell line was found to have the N-glycan Man₅GlcNac₂ [16], indicating hOR17-4 follows a different glycosylation pathway in this system.
Table 3 - Analysis of purified hOR17-4 by mass spectrometry. List of chymotryptic hOR17-4 peptides identified by LC-MS. Amino acids in parentheses are those before and after the assigned peptide in the protein sequence. "Ox" indicates oxidation (e.g. of Met).

5.4 Discussion

This work has confirmed that the engineered synthetic hOR17-4-rho gene encodes a receptor protein with wild-type function in the cell membrane and the protein retains this functionality, as ascertained by an SPR-based odorant binding assay, in the FC14-solubilized state. The purified protein also had correct secondary structure and possessed a near-UV CD spectrum indicative of a defined tertiary structure. Predicted post-
translational modifications were confirmed, including a homogeneous disulfide bond in the second extracellular loop and an N-linked hexasaccharide glycosylation on Asn5.

Despite the success of the SPR assay, a simpler, faster alternative would be desirable as the assay is time consuming and requires significant optimization and is thus not easily applicable to screening many odorant-OR combinations. However, now that one method to measure solubilized-OR functionality is in place, we can, by correlating with the SPR assay, begin to investigate more rapid methods of ascertaining olfactory receptor functionality. Other membrane proteins have been shown to display changes in spectroscopic characteristics as they degrade, unfold and lose function. Two such examples are photosystem-I [17] and rhodopsin [18]. Despite the fact that this might be expected as these are both light-sensitive proteins, spectroscopic shifts could correlating with odorant binding or loss of receptor function.

To explore this, a number of techniques can be investigated, such as shifts in tryptophan fluorescence upon odorant binding. Another technique that holds promise would be to perform far-UV CD measurements on OR secondary structure with the addition of odorants. Such assays have already been shown to detect ligand binding in the GPCR 5-HT4a, a result of disulfide bond shifting [19].

**OR stabilization using surfactant-like peptides**

An addition avenue to explore is the functional stabilization of ORs in solution. While the detergent-solubilized hOR17-4-rho protein was functionally active in our SPR assay, the lifespan of this activity has not been tested yet. However, the SPR and CD assays will allow us to investigate means of stabilizing the olfactory receptors in solution.
using other detergents and surfactant-like peptides. Previous work of this lab has successfully kept membrane proteins functionally stable in both wet (aqueous) and dry (powdered) states for periods as long as several weeks with the use of the peptide detergents such as A₆K and V₆D [17, 20]. Stabilized proteins were additionally found to be resistant to high temperature thermal denaturation.

Self-assembling surfactant-like peptides are key tools in the study of membrane proteins. These short peptides consist of a charged or polar head group followed by a string of hydrophobic amino acids, thus mimicking an amphipathic lipid. These molecules show the remarkable ability to spontaneously self-assembly into macromolecular structures such as micelles, nanotubes, and nanovesicles [21]. The designed peptide detergents also have several advantages for study of diverse membrane proteins: 1) Their biochemical properties resemble common detergents with similar critical aggregation concentrations (CACs) and appear not to denature membrane proteins and membrane protein complexes. 2) They are chemically and structurally simple and can be custom tailored for a variety of uses. 3) They are obtainable at high purity, readily soluble in water, and stable at ambient temperature. 4) They can be used together in a combinatorial manner and in conjunction with commercial detergents.

Due to their detergent properties, the small peptides can extract integral proteins from membranes and enhance their solubility in aqueous solutions (see Figure 5.7). This peptide coating can effectively mimic the natural environment of the lipid membrane, and work by the Zhang lab has demonstrated that membrane proteins can be stabilized and remain functional both in liquid as well as dry (powdered) environments for many weeks. Additionally, self-assembling surfactant peptides have been shown to facilitate
crystallization of membrane proteins, potentially leading to X-ray diffraction data and ultimately to valuable structural information. All of these aspects are directly applicable to our study on olfactory receptors.

Figure 5.7 – Depiction of surfactant-like peptide solubilization and stabilization of membrane proteins.

Surfactant-like peptide can be varied in several parameters such as the type (charged / polar) and number of hydrophilic residue(s) used, hydrophobic tail length, and type of hydrophobic residues used. See Figure 5.8 for a diagram of surfactant-like peptide structure. Additionally, multiple types of surfactant-like peptides can be used in conjunction to synergistically stabilize the receptors. We hypothesize that such a combinatorial screening method would further assist in the solubilization and purification of olfactory receptors.
Figure 5.8 – Molecular model of surfactant-like peptides. A6K is cationic while V6D is anionic. Ac refers to acetylation of the N-terminus, which aids in peptide stability.

5.5 Methods

Circular Dichroism Spectroscopy

Spectra were recorded at 15°C using a CD spectrometer (Aviv Associates, Model 202). Far-UV CD spectra were measured over the wavelength range of 195-260 nm with a step size of 1 nm and an averaging time of 5 s. Near-UV CD spectra were measured over the wavelength range of 250-350 nm with a step size of 1 nm and an averaging time of 10 s. All spectra were the average of 5 replicate scans. Spectra shown for purified hOR17-4 were blanked to wash buffer (concentrated to same extent as hOR17-4 sample) to remove effects of the detergent FC14. All measurements are reported in mean residue ellipticity (degrees * centimeter squared per decimole). Protein concentration was
determined from the aromatic absorption in 6 M Guanidinium HCl, pH 6.5 [22]. The secondary structural content was estimated using the online resource K2D (http://www.embl-heidelberg.de/%7Eandrade/k2d.html) [3].

**Fluorescent Spectroscopy**

Spectra were recorded at 25°C in a 1 cm path length semi-micro quartz cuvette (Hellma) using a Fluoromax-2 spectrofluorometer (Instruments SA, Inc., Jobin Yvon-Spex). Samples were excited at 280 nm and emission spectra collected from 300 – 500 nm with a step size of 1 nm and an integration time of 0.1 s. All spectra were the average of 3 replicate scans. Purified hOR17-4 was diluted in wash buffer to a concentration of 25 μg/ml to avoid signal saturation, and spectra were blanked a control wash buffer to remove effects of the detergent FC14.

**Calcium Signaling**

hOR17-4 expression in the stable HEK293S cell line was induced by 48 hour incubation in DMEM/F12 medium with GlutaMAX (Invitrogen), supplemented with 1 μg/ml tetracycline. Then cells were loaded at 37 °C for 30 min with 10 mM Fura-Red-AM (Molecular Probes) in serum-free DMEM/F12 medium and were subsequently washed with PBS and incubated in DMEM/F12 medium containing 10% FCS for 30 min to allow complete hydrolysis of intracellular Fura-Red-AM. Cytosolic Ca²⁺ responses were recorded by confocal fluorescence microscopy (Zeiss LSM 510) using a water immersion objective (Zeiss Achromplan 63 NA 1.2). Excitation was at 488 nm (Ar⁺ laser); the 650 nm long pass emission filter was used to image Fura-Red at a rate of 1 image per
second. Stock solutions of the tested odorants were prepared freshly in dimethylsulphoxide and diluted 1000-fold into phosphate-buffered saline (PBS) to give the desired concentrations.

**Surface plasma resonance (Biacore A100) odorant binding assay**

All odorant binding experiments were performed on a Biacore™ A100 (GE Healthcare, Uppsala, Sweden) at 25°C. A Biacore™ A100 was used in the present studies since it has a parallel flow configuration thus allowing assay development (e.g., solubilization conditions) to be tested and optimized in parallel in a multiplexed format. The sensor chip CM4, amine coupling kit, HBS (10 mM Hepes, 0.15 M NaCl, pH 7.4) and PBS were from (GE Healthcare, Uppsala, Sweden).

The rho1D4 monoclonal antibody (40 µg/ml in 10 mM sodium acetate, pH 5.5) was immobilized onto a series S sensor chip using standard amine-coupling chemistry in HBS running buffer at 10 µl/min [23]. The amount of coupled rho1D4 was about 5500 RU. Control surfaces were prepared similarly without protein derivatization and utilized as a reference surfaces for odorant binding experiments.

HEK293S hOR17-4-rho cells were induced with 1 µg/mL tetracycline plus 5 mM sodium butyrate for 48 hours and then scrape harvested 48 hours post-induction. The cells (5 x 10^6 cells/ml) were lysed with 1.5% FC-14 for 90 min at 4°C. The lysed cell suspension was centrifuged for 10 minutes at 14000 x g at 4 °C to remove cell debris. The supernatant, containing the solubilized olfactory receptor, was immediately captured on the surface plasmon resonance (SPR) chip using PBS, 1 mM Anzergent 3-14
(5xCMC), 5 mM TCEP, and 1.5 % (v/v) ethanol as running buffer at 10 μl/min. A four-minute injection resulted in a surface density of about 2800RU.

Fresh odorant solutions were made as follows. Pure odorant was diluted in ethanol to 0.5 M. This solution was diluted 67 times in PBS, 1 mM Anzergent 3-14, 5 mM TCEP, to obtain a 7.5 mM odorant solution in running buffer with 1.5 % (v/v) ethanol. Further dilutions were made in running buffer containing 1.5% (v/v) ethanol to obtain a concentration series of 5μM, 10μM, 20μM and 40μM of the odorants.

For the actual odorant binding measurements, the odorant concentration series were injected from low to high concentration over control and derivatized surfaces for 30 seconds with a flow rate of 60 μl/min. Zero concentration blank buffer cycles were included as negative control samples. Solvent correction procedures were included to compensate for any ethanol related bulk refractive index variations and performed as described previously [24]. Non-specific binding effects to sensor surface CM4 were absent for all analyses reported. Data analysis was carried out using Biacore A100 evaluation software. Data were prepared by subtraction of reference surface data and blank buffer sample data, a procedure commonly referred to as double referencing [25]. Solvent correction was then applied as described previously [24].

**Mass Spectrometry**

Liquid samples of purified monomeric hOR17-4 were subjected to chymotrypsin digestion and the resulting fragments analyzed using LC-MS by the MIT Koch Institute Proteomics Core Facility (Cambridge, MA). Proteolytic peptides were separated using 0.075 mm x 15 cm C18 column on a nano-HPLC system (TEMPO from Applied
Biosystems). Gradient elution with a water-acetonitrile-formic acid solvent system of peptides was carried out at a flow rate of 250 nL/min over 87 min. Electrospray mass spectra were acquired with a quadrupole time-of-flight mass spectrometer (QSTAR Elite from Applied Biosystems).

5.6 References


2. http://www.uniprot.org/uniprot/P34982


APPENDIX A
ADDITIONAL EXPERIMENTS

A.1 Crystallization Screening

No detailed structure for any olfactory receptor has ever been resolved by X-ray diffraction. For an excellent review of the current state of membrane protein purification and crystallization, see [26] on the previous page. Current structures are rough estimations based on computational modeling and comparisons to rhodopsin. In order to find suitable crystallization conditions, many thousands of solutions must be screened. Following the initial bioreactor purification, we sent 1.5 mg (300 µL of 5 mg/mL hOR17-4 monomer) to Dr. Duncan McRee at Active Sight in San Diego. The protein was then subjected to automated crystal screening using the Phoenix system, which is capable of sub-microliter condition volumes. Over 1200 conditions were tested, using Qiagen kits (MB I & II, MEMFAC, JCSG, JCSG CORE 1 - 4, PEG I & II, PEG ION, INDEX). Crystals were grown at 4°C and 20°C. Additionally, I set up 148 conditions here in conventional hanging drop plates at room temperature for potential diffraction-grade crystals. Several promising conditions were discovered, as seen in Figure A.1. Notably, all used the precipitant polyethylene-glycol-400 (PEG-400) and alkaline pH.

Following the second bioreactor purification, we sent 7.5 mg (750 µL of 10 mg/mL hOR17-4 monomer) to Dr. Eva Pebay-Peyroula at the Institute for Structural Biology in Grenoble, France for high-throughput crystal screening (started 7/16/08). They set up screening trials using both classical detergent-solubilized methods as well as cubic lipid phase crystallization. Additionally, the PEG-400 and high pH conditions are
Figure A.1 – Hanging drop crystallization of purified monomeric hOR17-4. Control consists of the same solubilization buffer used during OR purification (solution was concentrated to same extent as OR sample to ensure equivalent detergent concentration).

being more thoroughly explored due to the results from my hanging drop crystallization experiments. As crystallization can take many months to generate results, we are awaiting the latest screening data and hope that a suitable condition can be found so that we can proceed to X-ray diffraction in the hopes of obtaining the first high-resolution structure for an olfactory receptor.

Early studies in collaborating laboratories have shown that surfactant-like peptides can facilitate and enhance membrane protein crystallization (Joanne Yeh, Brown University, unpublished data). These surfactant-like peptides could prove vitally important in future attempts at the crystallization of hOR17-4 and other ORs. Additionally, the use of ORs known to bind to certain odorants will yield the possibility of getting structures for receptor-odorant complexes.
A.2 Additional Cell Lines Constructed

In addition to the previously discussed HEK293 stable cell lines constructed to express either rho-tagged hOR17-4 or hOR17-4(N5Q), we constructed and subcloned HEK293S stable cell lines which had tetracycline-inducible expression of several differently tagged versions of hOR17-4 to explore alternate purification. In each case, the corresponding gene was designed, synthesized and inserted into the pcDNA4/To T-REx vector, and stable cell lines were grown, subcloned, and screened as described earlier in Chapter 3 and in Appendix B.3. All subcloned stable-inducible lines have been cryopreserved for future use.

His-tagged hOR17-4

The His-tag (comprising a stretch of six histidine residues) has been ubiquitous in general protein purification as it is efficient, simple, and low cost. As the rho1D4-based purification method relies on a rather expensive antibody ($6000 per gram), we also investigated whether His-tag purification could be used as an alternative. A custom hOR17-4 gene was synthesized by PCR-based gene assembly to contain an N-terminal fusion to both a His-tag (for purification) and a FLAG-tag (for immunodetection) epitope. The FLAG tag (DYKDDDDK) can also be used for purification using the anti-FLAG M2 antibody affinity resin available from Sigma-Aldrich. The peptide sequence was MRGSHHHHHHGGDYKDDDDK followed by a two-glycine linker sequence. No C-terminal rho-tag was used in this construct. The construct was designated HF-hOR17-4 (HF standing for His-FLAG). Stable cell lines in HEK293 cells were constructed as described in Chapter 3.
When expression of HF-hOR17-4 was induced using either tetracycline alone or tetracycline plus sodium butyrate enhancer and analyzed via SDS-PAGE western blotting, we discovered that only the 30 kD monomer form of the receptor was present. This could be due to the N-terminal fusion tag interfering with the N-linked glycosylation known to produce the 32 kD monomer form. We also noticed that, when compared to hOR17-4-rho, the same induction treatment led to significantly greater toxicity in the HF-hOR17-4 construct, even when comparing multiple subclones. Whereas significant toxicity was only observed for hOR17-4-rho under very high induction conditions (1 \( \mu g/mL \) tetracycline plus 5 mM sodium butyrate), a similar amount of toxicity was observed for HF-hOR17-4 under tetracycline alone. Thus it is possible that the His-FLAG-tagged variant is inherently more toxic to the cells, and this effect precludes its further processing by glycosylation.

Nevertheless, we attempted to purify the HF-hOR17-4 using immobilized metal ion affinity chromatography (IMAC) followed by gel filtration chromatography. As can be seen in Figure A.2, the IMAC eluate contained a significant number of contaminating bands, due to non-specific interactions of other proteins with the nickel column. Even after a two-step purification (IMAC followed by size exclusion chromatography), the monomeric HF-hOR17-4 could not be separated from the contaminants and only 70% purity could be obtained (whereas >90% was possible using the rho-tagged construct). The yield of HF-hOR17-4 was approximately 40 micrograms per 150 mm tissue culture plate and we were able to purify 0.4 milligrams of the monomeric fraction, though still impure. The HF-hOR17-4 protein also readily precipitated when concentrated, unlike the
rho-tagged version. We therefore chose to abandon this method in favor of the much more successful rho1D4 immunoaffinity method described in the body of this thesis.

Figure A.2 – Analysis of His-FLAG-tagged hOR17-4. (A) SDS-PAGE anti-FLAG immunoblot of HF-hOR17-4 induction in stable HEK293S cell line (Clone 23). Only the 30 kD monomer form is seen, even under low levels of induction (tet only). (B) Comparison of affinity purified hOR17-4 variants, either rho-tagged (rho1D4 immunoaffinity chromatography) or his-flag-tagged (IMAC). (C) Comparison of variants after two-step purification (affinity followed by size exclusion chromatography). Gel filtration fractions loaded were early aggregate (A), dimeric (D), and monomeric (M). Even after two stages of purification, the his-tagged “monomeric” fraction still contains a large amount of contaminants.
Untagged hOR17-4

We have also generated stable inducible HEK293S cells that express an untagged version of hOR17-4, should future research show that the C-terminal rho-tag interferes with structure or function. This required the use of custom anti-hOR17-4 polyclonal antibodies in order to detect the wild-type receptor. Both anti-N-terminal and anti-C-terminal antibodies were raised by rabbit immunization using synthesized peptides corresponding to hOR17-4 amino acid sequences. These epitope sequences needed to be easily accessible and so we selected sequences at the extreme termini and made sure they ended at least five amino acids away from the predicted transmembrane regions. The peptides used were DGGNQSEGSEFC for the N-terminus and CRLLDKHFKRTL for the C-terminus. The rabbit inoculation and antibody purification (by affinity) were performed by Genscript (www.genscript.com).

Following construction of the untagged hOR17-4 gene and subsequent establishment of a subcloned stable-inducible HEK293S cell line, the expressed hOR17-4 was characterized by SDS-PAGE immunoblotting using the custom antibodies against both termini of hOR17-4. As was the case for the His-FLAG-tagged variant, the untagged version showed significantly higher cellular toxicity than the hOR17-4-rho construct, and cell death was seen even when inducing with tetracycline alone. The immunoblots were considerably dirtier than the monoclonal rho1D4 western blots, and numerous non-specific bands were seen. Importantly, only the 30 kD monomer form was seen (Figure A.3), as was the case for HF-hOR17-4-rho. Another observation is that both antibodies reacted to a 15 kD band, which could be a cleaved form of the receptor (essentially cut in half). This band is also seen in His-FLAG-tagged hOR17-4 (see Figure A.2b and A.2c),
but not in rho-tagged hOR17-4 (see Figure A.3). Thus it is reasonable to conclude that is
the increased toxicity (and likely ER accumulation and stalling of post-translational
processing) of both the untagged and His-FLAG-tagged versions of hOR17-4 that
prevents their glycosylation on Asn5 and potentially causes degradation. Exactly why the
rho-tagged version is less toxic remains to be seen, as the protein still retains normal
signal transduction and functional activity when assayed by cell surface calcium imaging.

Figure A.3 – Analysis of untagged hOR17-4 expression. Three separate hOR17-4
expressing clones are shown (w3, w2, and w6), either untreated (-) or treated for 48 hours
with 1 μg/mL tetracycline (+) or tetracycline plus 1 mM sodium butyrate (++). SDS-
PAGE was followed by immunoblotting with both anti-N-terminal and anti-C-terminal
custom polyclonal antibodies.

Additional hOR17-4 mutant versions

In addition to the hOR17-4(N5Q) mutant cell line, which was deficient for N-
linked glycosylation on Asn5, we constructed three additional inducible HEK293S cell
lines carrying mutated versions of hOR17-4. As hOR17-4 contains another potential
glycosylation site at Asn195, we thought to eliminate this site as well, in order to ensure
completely homogeneous receptor monomer should future research require it. Thus we constructed genes and subsequently produced cell lines that inducible express either hOR17-4(N195Q) or hOR17-4(N5Q,N195Q). Expression in these cell lines have not been fully characterized yet, but we would expect the N195Q variant to have little effect on the 32 kD monomer form, but could Asn195 glycosylation could perhaps be responsible for the faint band we frequently see at ~33 kD. The double mutant (N5Q,N195Q) should therefore eliminate both of these forms, resulting in only the 30 kD form being expressed.

After constructing the original hOR17-4-rho gene, additional human genome sequence data showed a polymorphism to exist at amino acid residue 25, with the two major variants in the population being glutamine (Q) or arginine (R). Our original construct had a glutamine at this position, so for future research we also created the alternate variant hOR17-4(Q25R) gene and stably-inducible cell line.

**A.3 Detection Using Custom Antibodies**

In section A.2, I described the generation and use of custom anti-hOR17-4 polyclonal antibodies raised against the N-terminus or C-terminus of the hOR17-4 protein. In addition to using these to construct and characterize untagged hOR17-4-expressing cell lines, these antibodies were also used to probe both the rho-tagged and His-FLAG-tagged variants of the receptor via SDS-PAGE immunoblotting. Prior to the experiment, I hypothesized that the anti-C-terminal antibody might have trouble detecting hOR17-4-rho, as the C-terminal rho-tag could presumably interfere with antibody binding. Similarly, the anti-N-terminal antibody might have difficulty detecting the HF-hOR17-4 protein due to its N-terminal His-FLAG tag. Additionally, the glycosylation
seen for the hOR17-4-rho protein should prevent the anti-N-terminal antibody from recognizing the 32 kD monomer band since Asn5 comprises part of its epitope.

When the immunoblots were performed on induced cell samples (Figure A.4), all these hypotheses were confirmed, with the exception that the anti-N-terminal antibody had no trouble recognizing the HF-hOR17-4 protein. The 15 kD band, which could potentially be a cleaved form of hOR17-4, was seen for the His-FLAG tagged protein but was absent for the rho-tagged version. This is further evidence that the hOR17-4-rho protein is more stable in this cell line and can be correctly processed via glycosylation and targeted to the cell membrane.

Figure A.4 – Detection of tagged hOR17-4 variants using custom polyclonal antibodies. Either HF-hOR17-4 (HisFlag) or hOR17-4-rho (Rho) expressing clones were either untreated (-) or treated for 48 hours with 1 µg/mL tetracycline (+) or tetracycline plus 5 mM sodium butyrate (++). SDS-PAGE was followed by immunoblotting with the indicated antibody. The anti-C-terminal antibody had difficulty recognizing the rho-tagged variant, while the anti-N-terminal antibody did not effectively bind the 32 kD (glycosylated) monomer form of hOR17-4-rho, despite equal amounts of the two forms (see anti-rho blot).
APPENDIX B: EXPERIMENTAL PROTOCOLS

B.1 Gene Design Using DNAWorks

Go to:
http://helixweb.nih.gov/dnaworks/

1. Enter job name (such as hOR174)
2. Enter your email address
3. Do not check mutant run
4. Codon Frequency – check either “H. Sapiens” (for mammalian cell expression) or “E. Coli Class II” (for bacterial expression)
5. Parameters:
   a. Annealing Temp: 60°C
   b. Oligo Length: 40-50 nt
      i. Try 40 first (shorter usually gives lower error rate in synthesis)
      ii. You can also enter a range (from 40-50) and just put 60°C in both temperature fields
      iii. Or, you can try length of 50 and selecting the “random” option (this will allow the lengths to be non-uniform)
   c. Codon Frequency Threshold: 19% (human) / 15% STRICT (e. coli)
   d. Oligo concentration: 25E-9 M (25 nM)
   e. Na+/K+ concentration: 0.05 M (50 mM) – default value
   f. Mg2+ concentration: 0.002 M (2 mM) – default value
   g. Number of solutions: 10
   h. TBIO mode: do not check
6. Restriction site screen:
   a. Non-degenerate, select EcoRI and NotI (or whatever sites you are using to clone in your gene)
   b. Be sure to hold down command to select BOTH sites!
7. Sequences:
   a. Click “Add sequence field” twice (for the 5’ and 3’ headers)
   b. Sequence 1: 5’ DNA sequence leader (restriction sites, ribosome binding site)
      (example: EcoRI/Kozak header CCTGAATTCGCCGCCACC)
   c. Sequence 2: your protein sequence with any tags included in single letter amino acid format
      i. Be sure to end protein coding sequence with an X (for stop codon)
      ii. for rho1D4 tag, add GGTETSQVAPA to C-terminus (BEFORE the stop codon, of course)
      iii. also add other tags to N- or C- terminus (be sure to add glycine linkers!)
   d. Sequence 3: 3’ DNA sequence flanker
8. Click “Design Oligos”
9. Results will be emailed to you
   a. Lower scores are better (and lower Tm range). An score of less than 1.000 is desirable.
   b. Examine and compare the lowest scoring oligo sets and choose the best one to order.
   c. Note that the NotI site will always give you some repeat, GC-rich, and pattern score

(example: NotI flanker GCGGCCGCGAGAAG)

Sample of a DNAWorks3.1 output trial

Trial shown was used in the design of the olfactory receptor M71-rho gene.

PARAMETERS FOR TRIAL 67

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size Of Gene</td>
<td>995 nt</td>
</tr>
<tr>
<td>Protein Residues</td>
<td>321</td>
</tr>
<tr>
<td>Mutatable Residues</td>
<td>307</td>
</tr>
<tr>
<td>Fixed Nucleotides</td>
<td>74 nt</td>
</tr>
<tr>
<td>Oligo Size</td>
<td>46 nt</td>
</tr>
<tr>
<td>Annealing Temp</td>
<td>60 +/- 1*C</td>
</tr>
<tr>
<td>Oligo Concentration</td>
<td>2.50E-8 M</td>
</tr>
<tr>
<td>Sodium Concentration</td>
<td>5.00E-2 M</td>
</tr>
<tr>
<td>Mg2+ Concentration</td>
<td>2.00E-3 M</td>
</tr>
<tr>
<td>Codon Frequency Threshold</td>
<td>19%</td>
</tr>
<tr>
<td>Repeat Threshold</td>
<td>8 nt</td>
</tr>
<tr>
<td>Mispriming Threshold</td>
<td>8/18 (6 exact) nt</td>
</tr>
</tbody>
</table>

The DNA sequence # 67 is:

```
1 CCTGAATTCGCCGCCACCATGACCGCTGAAAACCAAAGCACCGTGACCGAGTTTATCCTC
61 GGAGGGCTCACAATAATTGGCTGAACTCGACCTCCCCCTCTCTCTCTGTGTTTCTGGGGATC
121 TACGCTGGTGAATGGTGGGCAATCTGGGCATGATCACTCTGATTGGCCTCAACAGCCAA
181 CTGCACACCCCCATGTACTTTTTTCGAGCAACCTGAGCTGGTGGACCTGTGCTATTCC
241 AGGCTCATCACACAAAAATGCTGAACCTTGCTGAGCCAAAGGAATCTCATACGCTATT
301 GTGGGATGATAGCCAGCCGAGCTGTATTTCTTCTGTGTTTCTGTGATCGCGAGTGCTACATG
361 CTGACAGTGATGGCTATAGCAAGTGATCTGCGCCATCTGCAAGCCCTGCTCTACAACATT
421 ATCTACGTCGCTCGGGAACATCTGCCACTCAATGAAACTCTGCTGCTCCAGACATACAGCT
481 GGCAGAGCAGCAAGCGAGAGCTGATGCTGAAGCTCAACTACAGGACCTACATCTCC
541 CACTACCTTCTGCGGAACATCTGCCACTCAATGAAACTCTGCTGCTCCAGACATACAGCT
601 GAGATGCCTGCCTTTCTTCTGCGCTGCGCTGAGCTACATGCTCACCGCCCTACGGCTCTGTG
661 ATCTCCTACGCTCTTCTATCTTCGTCGTCACTCATCGTACGGTTACCTTGCAACCCGCGAG
721 AAAAGCTTTTGGACACCTCGGCACTGGCTGCTGCTGCCCGCTGCGCTGCGCTGCTGCTGC
781 GCATTTGATGATGACGAGCCACCCCGCCTCCTGCTGCTGCCCGCGAGAGCGAGCAGCAGCG
841 GTCTTCTATACGCCGGTCCTGGCTATGCTCAGCCTATGCTATGCTATGCTATGCTATGCTA
901 GAGCTGGAAGGAGCCGCTGTGACGACGACCTCTGCGCGCGGAGAGGTCCTCTGGCGGCCCG
961 AGCCAGGGCGCTCTCGCTGAGCCGCGCGCGAGAAG
```

The oligonucleotide assembly is:

```
1 10 20 30 40 50 60
| | | | | | |
1 CCTGAATTCGCCGCCACCATGACCGCTGAAAACCAAAGCACCGTGACCGAGTTTATCCTC 3 ttagcttct
GACCTTTTGGTTTCGCTGCCACTGGCTCAAAATGGAG
```

M T A E N Q S T V T E F I L

```
| | | | | | |
5 --->
```
961 agccaggtcgctcc

** repeat
** GC rich

tcgctcagcgagggactcgccggcgctcttc

S Q V A P A X

--- 36

The total codon usage score ............ 0.000
The total length score .................. 0.000
The total melting temperature score ... 0.000
The total repeat score .................. 0.322
The total pattern score .................. 0.322
The total mispriming score .............. 0.000
The total AT content score ............. 0.000
The total GC content score ............. 0.322
The OVERALL score ............ 0.965

DETAILED CODON FREQUENCY REPORT
[ Codon, AA, Frequency, # of times used in coding sequence ]

<table>
<thead>
<tr>
<th>Codon, AA, Frequency, # of times used in coding sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTT F 0.46 7 TCT S 0.19 0 TAT Y 0.44 5 TGT C 0.45 0</td>
</tr>
<tr>
<td>TTC F 0.54 15 TCC S 0.22 13 TAC Y 0.56 13 TGC C 0.55 9</td>
</tr>
<tr>
<td>TTA L 0.08 0 TCA S 0.15 0 TAA X 0.30 0 TGA X 0.47 1</td>
</tr>
<tr>
<td>TTG L 0.13 0 TCG S 0.06 0 TAG X 0.23 0 TGG W 1.00 0</td>
</tr>
<tr>
<td>CTT L 0.13 0 CCT P 0.28 2 CAT H 0.42 0 CGT R 0.08 0</td>
</tr>
<tr>
<td>CCT L 0.20 14 CCC P 0.33 8 CAC H 0.58 3 CGC R 0.19 0</td>
</tr>
<tr>
<td>CTA L 0.07 0 CCA P 0.28 1 CAA Q 0.26 3 CGA R 0.11 0</td>
</tr>
<tr>
<td>CTG L 0.40 32 CCG P 0.11 0 CAG Q 0.74 5 CGG R 0.20 3</td>
</tr>
<tr>
<td>ATT I 0.36 6 ACT T 0.25 2 AAT N 0.47 3 AGT S 0.15 0</td>
</tr>
<tr>
<td>ATC I 0.47 17 ACC T 0.36 14 AAC N 0.53 10 AGC S 0.24 20</td>
</tr>
<tr>
<td>ATA I 0.17 0 ACA T 0.28 6 AAA K 0.43 3 AGA R 0.21 0</td>
</tr>
<tr>
<td>ATG M 1.00 14 ACG T 0.11 0 AGG K 0.57 6 AGG R 0.21 5</td>
</tr>
<tr>
<td>GTT V 0.18 0 GCT A 0.26 5 GAT D 0.46 0 GGT G 0.16 0</td>
</tr>
<tr>
<td>GTC V 0.24 10 GCC A 0.40 13 GAC D 0.54 7 GGC G 0.34 10</td>
</tr>
<tr>
<td>GTA V 0.12 0 GCA A 0.23 3 GAA E 0.42 3 GGA G 0.25 3</td>
</tr>
<tr>
<td>GTG V 0.47 16 GCG A 0.11 0 GAG E 0.58 8 GGG G 0.25 3</td>
</tr>
</tbody>
</table>

Frequency Range Number of Codons
-----------------------------------
<p>| 0% - 4% | 0 |
| 5% - 9% | 0 |
| 10% - 14% | 0 |
| 15% - 19% | 17 |
| 20% - 24% | 51 |
| 25% - 29% | 25 |
| 30% - 34% | 18 |
| 35% - 39% | 65 |</p>
<table>
<thead>
<tr>
<th>Tm Range</th>
<th># of Overlaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;57</td>
<td>0</td>
</tr>
<tr>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>59</td>
<td>18</td>
</tr>
<tr>
<td>60</td>
<td>17</td>
</tr>
<tr>
<td>61</td>
<td>0</td>
</tr>
<tr>
<td>62</td>
<td>0</td>
</tr>
<tr>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>&gt;=65</td>
<td>0</td>
</tr>
</tbody>
</table>

Tm Range = 1.7

<table>
<thead>
<tr>
<th>Ovrlap Len Range</th>
<th># of Oligos</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;15</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>&gt;=25</td>
<td>0</td>
</tr>
</tbody>
</table>

Lowest Overlap = 16

<table>
<thead>
<tr>
<th>Length Range</th>
<th># of Oligos</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;37</td>
<td>0</td>
</tr>
<tr>
<td>37-38</td>
<td>1</td>
</tr>
<tr>
<td>39-40</td>
<td>0</td>
</tr>
<tr>
<td>41-42</td>
<td>0</td>
</tr>
<tr>
<td>43-44</td>
<td>0</td>
</tr>
<tr>
<td>45-46</td>
<td>35</td>
</tr>
<tr>
<td>47-48</td>
<td>0</td>
</tr>
<tr>
<td>49-50</td>
<td>0</td>
</tr>
<tr>
<td>51-52</td>
<td>0</td>
</tr>
<tr>
<td>53-54</td>
<td>0</td>
</tr>
<tr>
<td>55-56</td>
<td>0</td>
</tr>
<tr>
<td>&gt;=57</td>
<td>0</td>
</tr>
</tbody>
</table>

Longest = 46
There are 1 repeats greater than 8 nt:

<table>
<thead>
<tr>
<th>PR Pos1 = 982</th>
<th>Pos2 = 982</th>
<th>Size = 8</th>
<th>Seq1 = GCGGCCGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq2 = GCGGCCGC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequence Patterns Screened (As Supplied By User)

<table>
<thead>
<tr>
<th>Name</th>
<th>Seq</th>
<th>Pos</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>GAATTC</td>
<td>4</td>
<td>forward</td>
</tr>
<tr>
<td>NotI</td>
<td>GCGGCCGC</td>
<td>982</td>
<td>forward</td>
</tr>
</tbody>
</table>

36 oligonucleotides need to be synthesized
B.2 PCR-based Gene Assembly (PCA)

We highly recommend oligos to be ordered from IDT (idtdna.com) as they consistently have the lowest error rates. Oligos from IDT come in plate format at 50uM concentration. For the PCR reactions we recommend Pfu Turbo Hotstart polymerase (Stratagene).

Combine oligos for assembly in equal proportions (for example, ~20 ul each). Vortex well to mix. Final concentration will be 50uM total, but 50/(# oligos) uM for each oligo. Thus, having a pool of 34 oligos means each is at 1.47 uM.

Also combine end oligos (10 ul each) and add 8 parts milliQ H20 (80 ul), so that final concentration of each is 5 uM. This is the primer end mix used in PCR 2.

**PCR 1: Assembly, 50 ul reactions**

<table>
<thead>
<tr>
<th>volume</th>
<th>reagent</th>
<th>initial conc</th>
<th>final conc</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 uL</td>
<td>10x cloned buffer (for Pfu Turbo Hotstart)</td>
<td>10x</td>
<td>1x</td>
</tr>
<tr>
<td>1.25 uL</td>
<td>dNTP mix</td>
<td>10 mM each</td>
<td>250 uM each</td>
</tr>
<tr>
<td>0.5 uL</td>
<td>pooled oligos</td>
<td>50/# uM each (50 uM total)</td>
<td>500/# nM each (500 nM total)</td>
</tr>
<tr>
<td>1 uL</td>
<td>polymerase (Pfu Turbo Hotstart)</td>
<td></td>
<td>2.5 Units</td>
</tr>
<tr>
<td>42.25 uL</td>
<td>distilled water</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 uL</td>
<td>total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Note: some reactions appear to work better if individual oligo conc is constant (25 nM each oligo) and not total oligo conc. Thus, for 34 oligos you would add 0.85 uL pooled oligos (instead of 0.5 uL) – in this case, total oligo conc would then be 850 nM. Note that in this method, there is a risk of dNTP depletion.

Average oligo extension is ~1/2 gene length
In 50uL rxn, have 50 nmol total dNTPs and 25 pmol total oligos (if 500 nM total)
Thus about 2000 fold excess of dNTPs
If gene is 1kB, then FULL extension will only use about 25% of dNTPs
If keep each oligo constant, then will use up more dNTPs (43% if have 34 oligos)

PCR1 program:
94 C hot-start – 2 minutes
30-50 cycles (we use 40 by default; occasionally, more are needed)
   94 C denaturation 30 seconds
   55 C annealing 30 seconds
   72 C extension 1.5 minutes
72 C final extension 2 minutes
4 C hold
If the products of this reaction are run on a gel, the normal result is a smear, with the lower edge running about the same as the lengths of the oligos, and the upper edge slightly higher than the expected product length.

**PCR 2: Amplification, 50 ul reactions**

The 2\textsuperscript{nd} PCR is more like normal PCR, using two primers to selectively amplify the correct length fragment from the mixture created by assembly PCR above. These primers can simply be the oligos used at the beginning (5' end) of each strand for gene assembly in PCR 1. **CAUTION:** the small amount of PCR product from the previous reaction will contain some Pfu polymerase which has already been activated by the hot-start procedure. Thus, it is a good idea to add the 1 ul of PCR product last, just before starting the program.

<table>
<thead>
<tr>
<th>amount</th>
<th>reagent</th>
<th>initial conc</th>
<th>final conc</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 uL</td>
<td>10x cloned buffer (for Pfu Turbo Hotstart)</td>
<td>10x</td>
<td>1x</td>
</tr>
<tr>
<td>1.25 uL</td>
<td>dNTP mix</td>
<td>10 mM each</td>
<td>250 uM each</td>
</tr>
<tr>
<td>3 uL</td>
<td>Primer ends mix</td>
<td>5 uM</td>
<td>300 nM</td>
</tr>
<tr>
<td>1 uL</td>
<td>product from assembly PCR</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>1 uL</td>
<td>polymerase (Pfu Turbo Hotstart)</td>
<td></td>
<td>2.5 Units</td>
</tr>
<tr>
<td>38.75 uL</td>
<td>distilled water</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 uL</td>
<td>total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PCR2 program:
94 Celsius hot-start – 2 minutes
30 cycles:
94 C denaturation 30 seconds
58 C annealing 30 seconds (varies depending on the particular primers, we’ll typically use 50 C)
72 C extension 1 minute (for sequences > 1 kb, use 1 minute per kb)
72 C incubation 2 minutes
4 C hold
B.3 Stable HEK293S Colony Subcloning and Screening

Cells were transfected using Lipofectamine 2000 (Invitrogen) for 16 hours, according to the manufacturer’s protocol for HEK293 cells. Two days later the media was changed to selective plate media containing blasticidin and zeocin:

<table>
<thead>
<tr>
<th></th>
<th>Blasticidin</th>
<th>Zeocin (selection)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEK293S cells:</td>
<td>5 μg/mL</td>
<td>250 μg/mL</td>
</tr>
<tr>
<td>HEK293 GnTI- cells:</td>
<td>5 μg/mL</td>
<td>50 μg/mL*</td>
</tr>
</tbody>
</table>

*: Once the cell line has been established, use 25 μg/mL zeocin (half as much) in HEK293 GnTI- cells to maintain the stable cells.

After three weeks, a stable cell line had been established. In order to establish syngeneic colonies, the cell culture was split to yield a very low density of one single cell per low-power microscope field (approximately 50-100 cells per 150mm culture plate). Colonies were grown in selective media until they reached a size of 1-3 millimeters, whereupon the media was changed to PBS. Single colonies were aspirated using the a sterile pipette, and transferred to a new well of a 24-well plate containing 100µL trypsin and broken up and resuspended by pipetting up and down. After 1 minute, 1 mL of selective plate media was added to inactivate the trypsin, and cells were incubated overnight to allow attachment. Media was then changed again and cells were grown until confluence, whereupon they were split for inducible-expression screening.

To split grown-up colonies for treatment:

(note: it’s generally easier to do the following procedure 3 wells at a time (thus filling up a complete 12-well plate with each set))

1. Aspirate media and add 1mL PBS to well
2. Aspirate and add 100uL trypsin to well
3. Place in incubator for several minutes
4. Resuspend with 1mL of media (will be 1.1mL with the trypsin) and distribute to 4 wells of a 12-well plate as follows:
   a. Top 3 wells get 340uL
   b. Last well gets remainder (will use later to grow up the cell line)
   c. Then add 1mL media to each of those 4 new wells
   d. Rock plate (or bonk against table) to distribute cells evenly
5. Repeat for rest of colonies (can thus put 3 colony trials on one 12-well plate)

6. Once cells are near confluence, treat with drugs:
   a. [-] Untreated, [+] tet only (1 ug/mL), [++] tet + NaBu 2.5-5mM, untreated (grow up to establish the cell line)
   b. Be sure to change the media on the untreated wells too!
   c. Harvest the 3 treated wells (-, +, ++) after 48 hours (see above)
   d. Keep the final well growing. After analyzing the samples, pick the best ones and split them into 100mm plates to establish the cell line. Maintain
the presence of the T-REx system by growing in media containing blasticidin and zeocin. Keep growing up for purification experiments and be sure to cryopreserve some aliquots for future use.

**To harvest treated samples:**
(note: if many samples, easiest to do in “shifts” of 18 since cold room centrifuge can only hold 18 tubes)

1. Aspirate media and add 1mL PBS to well
2. Pipette up and down over entire well surface to resuspend cells, then transfer to a labeled microcentrifuge tube
3. Spin down cells at 10,000 rpm for 5 minutes at RT
4. Remove PBS
5. Resuspend cell pellet in 120uL of cold PBS + protease inhibitors
6. Add 30uL of cold 10% FC14 detergent (in PBS) – final concentration will be 2%
   - Ensure cells are suspended when detergent is added (vortex if necessary)
   - Do not pipette up and down after detergent addition!
7. Invert tube to mix and solubilize receptors by rotating for 60 minutes at 4°C
   - Use the rotator in the cold room
   - If many tubes, can pack 6 per 50mL conical and then rubber band those in place on the rotator
8. Spin in cold-room centrifuge at 13,000rpm for 30 minutes to pellet insoluble material
9. Transfer 100uL of supernatant to a new tube (PCR tube strips work well)
   - Take 100uL from the top and avoid the pellet!
10. Freeze solubilized supernatant samples at -20°C (freezer) and discard pellets
B.4 Immunoblotting Protocol

Gel electrophoresis (NuPAGE)

- Thaw samples to run on ice
- Prepare 1L of NuPAGE MOPS buffer (50mL of the stock + 950mL diH20)
  - MOPS buffer stock is under the electrophoresis bench area in a white box
  - After making, put in cold room so it gets to 4°C
- Prepare samples according to Novex NuPAGE procedure card
  - Usually run the following:
    - 1-4 uL of sample
    - 2.5 uL of 4x LDS
    - 1 uL of 10x reducing agent
    - milliQ H20 to 10 uL total volume
  - For ladder there is no need to add anything else
- Let samples sit at RT for >15 minutes (don’t heat samples as it causes aggregation)
- Prepare Novex NuPAGE 10% Bis-Tris gel and place into an assembled Novex gel box. Pour about 100mL of MOPS buffer into bottom of gel box
- Put 200mL of MOPS buffer in new bottle and add 500uL antioxidant, then fill inner chamber with this solution
- Add regular MOPS buffer to outer chamber until half full
- With a 200uL pipettor, briefly rinse out each well
- Apply samples to wells using gel loading tips (20uL pipettor)
  - Be careful of bubbles! (watch the tip and stop when sample is at the end)
  - Best method is to keep tip close to the well bottom as you apply sample.
    - This results in less mixing of sample with the MOPS buffer.
  - Load 10uL of Rainbow marker for all ladder lanes
- Move gel box to cold room and run at 100V (constant voltage) until the blue dye runs off the bottom. This will take several hours. For best resolution of hOR17-4, you can run until the green marker (25kD) is about 2cm from the bottom.
- After the SDS-PAGE is complete:
  - For immunoblotting, perform transfer protocol (below) followed by western protocol (below)
  - For total protein staining, stain gel with Sypro Ruby (Molecular Probes) using the manufacturer protocol.

Transfer (Bio-Rad MiniTrans Blot)

1) Get out a Nitrocellulose Membrane sandwich (from Bio-Rad). They are in a small box under the gel electrophoresis bench. They are precut squares about the same size as the gels we use. (Filter paper | Membrane | Filter paper)
2) Get out transfer case (green) with the red/black/white stuff inside.
3) Fill ice tray (pack down) and place in case
4) Fill up transfer case ~2/3 full with 1x transfer buffer (newer)
5) Ensure small magnetic stir bar is in bottom of case
6) Get out plastic tub (above my bench) as well as glass roller
7) In tub, place one transfer cassette open, black on right side, put 1 pad (white sponge) on each side, and put membrane sandwich on left (clear side)
8) Pour a layer of 1x transfer buffer (older) into tub and wet pads and membrane sandwich
9) Take gels out of NuPAGE box and crack open perimeter using tool above the sink.
10) Gel will stick to one of the plates – just throw out the other one
  a. Note – if gel sticks to the taller plate it will effectively be flipped over (left to right) from how you loaded it. Be sure to flip it over again with your fingers during step 12
11) Use a razor blade to trim off bottom “lip” of gel as well as the well combs.
12) Use water and fingers to loosen gel from plate (flip it over with fingers, if necessary)
13) Use tweezers to move top piece of filter paper (on membrane sandwich) over to right (black) side, exposing the membrane. Ensure membrane is wetted with transfer buffer.
14) Place trimmed gel (right side up!) on top of membrane & gently roll out bubbles with glass roller.
15) Place filter paper (from black side) back on top and gently roll out bubbles
16) Place right hand side pad on top, fold over black side and secure the clamp
  a. Note: make sure gel is between membrane and black side, or else proteins will transfer OUT INTO THE BUFFER!
17) Place transfer sandwich in transfer case with black side to black side
18) Repeat for 2nd gel, or else just place empty sandwich in 2nd slot
19) Fill case with 1x transfer buffer until ~1cm from lip and put on lid
20) Run in cold room (4°C) while stirring at 0.25 amps = 250 milliamps (constant amps) for 1 hour
  a. Note: can also run overnight at 30V
  b. Transfer buffer can be reused several times)
21) Disassemble transfer cassettes and ensure that protein has transferred to gel
  a. can tell by looking at membrane for reflection of protein bands AND by transfer of colored protein ladder
  b. note that smaller proteins transfer faster than larger ones (so should increase transfer time for very large (>100kD) proteins)
22) Mark (dot with lab marker) protein ladder bands (since they can fade in subsequent steps)

**Dot Blotting:**

For dot blots (useful for comparing the relative expression levels of many samples), simply pipette 1 μL of each sample onto a nitrocellulose membrane in a grid pattern. Allow to air dry for 20 minutes, then follow western blotting procedure below.
Western Blotting:

- Block membrane in 50mL blocking buffer (TBST + 5% milk)
  - Shake 1hr at RT or overnight at 4°C
- Dilute primary antibody in new blocking buffer
  - For my square membrane containers, 30mL total is fine
  - For rho1D4 (anti-rho tag), dilute antibody 1:2000 (so 15uL into 30mL)
  - For anti-N-term, anti-C-term, and anti-FLAG, dilute antibody 1:3000 (so 10uL into 30mL)
- Incubate membranes in primary antibody (shaking) for 2 hours at RT or overnight at 4°C
  - Can cut up membranes if necessary (e.g., one side rho, other side N-term)
- Pour off antibody solution and save in fridge
- Wash membrane briefly in about 20mL TBST (no milk)
- Wash 5x with 50mL TBST, 5 minutes each (shake)
- Dilute secondary antibody in new blocking buffer
  - For my square membrane containers, 50mL total is fine
  - Dilute antibody 1:5000 (so 10uL into 50mL)
  - Goat anti-**mouse** antibody HRP for mouse primaries (like the anti-rho and anti-FLAG)
  - Goat anti-**rabbit** antibody HRP for rabbit primaries (like the anti-N-term and anti-C-term)
- Incubate membranes in secondary antibody (shaking) for 1 hour at RT
- Discard secondary and wash membrane as before (5 x 50mL, 5 mins each, with TBST)
- Detect using ECL Plus Kit (GE Healthcare)
B.5 Bioreactor Media Preparation

For 1 Liter:

Dissolve the following in 1 liter of milliQ-filtered water:

- 13.7g DMEM powder (calcium free, from Atlanta Biologicals)
- 3.7g Sodium Bicarbonate (NaHCO3)

pH to 7.0 and add:

- 0.3g Dextran Sulfate (Mr~5000, from Sigma)
- 3g Primatone RL-UF (from Quest International)

Sterilize by filtering through 0.2um filter into a sterile bottle.

Now add the following sterile components:

- 100mL FBS (to 10% vol/vol)
- 10mL Pen-Strep (to 1% vol/vol)
- 10mL L-glutamine (to 2mM)
- 10mL 10% Pluronic F-68 (to 0.1% wt/vol)

To make sterile 10% Pluronic F-68 solution (wt/vol):

- Dissolve 10g Pluronic F-68 in 100mL milliQ H20
- Sterilize by filtering through 0.2um filter into a sterile bottle
- Store at RT

Also make sterile supplements same way:

- 2mM L-glutamine (2.92g in 100mL, store at 4°C)
- 8% Sodium Bicarbonate (8g in 100mL, store at RT) add 20mL if necessary...
- 20% glucose (20g in 100mL, store at RT) add 10mL for d5 supplement
- 10% Primatone RL-UF (3g in 30mL, make fresh) add 30mL for d5 supplement

To induce cells:

- We use tetracycline and sodium butyrate at slightly higher levels [Prashen uses tetracycline at 2 ug/mL (2x) and sodium butyrate at 7.5mM (1.5x)]

**For ORs, use 2 ug/mL tetracycline and 2.5mM NaBu**

To induce 1L of cells:

- 20mL of 50x induction = 0.1 mg/mL tetracycline and 125mM NaBu (= 13.76 mg/mL)

  - Dissolve the following in 18mL of milliQ H20:
  - 2 mL of 1 mg/mL tetracycline stock (total of 2 mg)
  - 0.275g of sodium butyrate
  - Gives 20mL total
B.6 Bioreactor Operation

(optimized for 1.25L final culture volume):

- **Day 0:** Make up 1.1L of media (recipe). With FBS and other ingredients the final volume will come out to about 1.25L total volume. Put 100mL in a separate sterile bottle to use cell resuspension, and add rest to reactor using sterile siphon into **INPUT PORT 1** (after draining PBS).

  Split 6-8 confluent 150mm plates of cells, resuspend in remaining media (100mL) and inoculate reactor with cells using siphon into **INPUT PORT 2** (final cell density should be $3-5 \times 10^5$ cells/mL).

- **Day 1-4:** Monitor reactor conditions and adjust as necessary. If pH begins dropping and increased gas flow (up to flow rate of 25) cannot compensate, add 20mL of the sterile 8% NaHCO3 solution through **INPUT PORT 4** by siphon.

- **Day 5:** Growth supplement

  Add 3.4g of Primatone RL-UF to 34mL of PBS and dissolve by heating in water bath to ~50°C. Use a stir bar to get it completely dissolved so that solution is totally clear. In the sterile hood, add 11mL of the sterile 20% glucose solution and mix again – we want the Primatone to be completely dissolved or else it will clog the filter later! Once it’s dissolved, sterile filter through a 0.2µA zap cap into a sterile 250mL bottle. Then siphon transfer into reactor using **INPUT PORT 3**.

- **Day 6:** OR induction.

  Induce expression using 2 µg/mL tetracycline and 2.5mM sodium butyrate final concentration.

  To 23.4mL of PBS add 0.286g of sodium butyrate powder. Stir to dissolve. Then inside sterile hood, add 2.6mL of tetracycline stock (1 mg/mL). Stir to mix, then sterile filter. It might be more practical to use a syringe filter (0.2µm) than a Zap cap – just be careful to keep it sterile and filter into a sterile container like a 50mL conical. Then transfer to sterile 250mL bottle and siphon to reactor using **INPUT PORT 4**.

- **Day 8:** Harvesting cells. → Harvest @ 40hr post-induction

  Make up 100mL of PBS plus protease inhibitors. Thus drop 2 of the Complete EDTA-free tablets (NOT mini) into 100mL total. Once made, put on ice to cool to 4°C.

  At reactor, use total drain line to siphon all cells into a clean/sterile 2L bottle. (note – we don’t really need to be absolutely sterile from this point on, but still try to be clean!)
***Take 1mL of the reactor harvest and put in a microcentrifuge tube for later cell-counting analysis. Spin down cells in centrifuge for 2 minutes at 10000 rpm and keep on ice until freezing in liquid nitrogen later.***

Transfer cells to 6 big centrifuge bottles (the kind we use for bacteria – be sure to thoroughly clean them out first!!!). Balance the bottles, then spin in the centrifuge (Sorvall, GSA rotor) for 20 minutes at 8000 rpm at 4°C. Afterwards, all cells should be pelleted and media supernatant totally clear – if not, centrifuge again. Once OK, pour off supernatant. Resuspend cells in cold PBS + inhibs solution by vortexing thoroughly (with lid on) and then pipetting to break up clumps. Pool together into two 50mL conicals and centrifuge for 20 minutes at 5000 g at 4°C.

You can then snap freeze the pellets in liquid nitrogen and store at -80°C until the day of purification.
B.7 Rho1D4-bead Coupling

The antibody is rho1D4 monoclonal antibody (from Hybridoma 1B4-1) from Cell Essentials. They will produce and purify it for you (protein A purification) at a cost of roughly $6000 per gram.

**NOTE:** to avoid the hassle of dialyzing the antibody, be sure to have Cell Essentials ship the antibody in the coupling buffer (0.25M NaHCO3, 0.5M NaCl, pH 8.3) and NOT in phosphate buffer.

If antibody is not in coupling buffer, dialyze as follows:

1) Before coupling to beads, antibody must be dialyzed into **coupling buffer** (0.25M NaHCO3, 0.5M NaCl, pH 8.3)!
   a. pH must be 8-10 for efficient coupling
   b. Bicarbonate or borate buffers must be used, as Tris and other buffer salts will couple to the gel!
   c. High salt content (0.5M NaCl) is crucial as it prevents aggregates and minimizes protein-protein adsorption

2) Wash dialysis tube (“Spectra/Por 2” RC membrane w/ 12-14kD MWCO [could be higher]) thoroughly with distilled water, then fill with antibody solution

3) Dialyze against 5-10 L **coupling buffer** (at least 100 volumes, if possible) overnight at 4°C. Change dialysis buffer at least 3 times.

4) Determine protein concentration using A280 (1.383 AU == 1 mg/mL)
   a. (GE Healthcare recommends 5-10 mg/mL (2-8 is optimal, NOT HIGHER!)

**Beads:** CNBr-activated Sepharose 4B (GE Healthcare)

1g of dried bead powder will equate to 3.5 mL of final gel volume (bead slurry)

1) Suspend beads in 1mM HCl (beads swell → gel)
2) Wash for 15 minutes with 1mM HCl on a sintered glass filter
   a. Wash procedure: Use sintered glass funnel and vacuum flask. Resuspend without suction. Wait ~1 minute. Apply vacuum until a dry cake is obtained. Do not over dry.
   b. Use about 200mL wash per gram of powder, added in several aliquots
3) After washing, add bead slurry to antibody (now in coupling buffer)
   a. Add 20 mL bead slurry to 130-200 mg antibody equivalent
   b. Thus, 50 mL slurry is plenty for our 326.70 mg antibody
   c. (Rule from GE Healthcare is 5 mL coupling solution per gram of dried bead powder) – 5 mL solution / 3.5 mL bead slurry
4) Mix and rotate at 4°C (4 hours to overnight)
   a. Monitor antibody concentration in supernatant using A280 (spin down beads for 5 mins first, 2000 rpm)
   b. Can stop when antibody concentration is below 5% of starting value
5) Discard supernatant (spin down beads first)
6) (optional) Wash away excess antibody with 5 gel volumes of coupling buffer
7) Block any remaining active groups on the beads
   a. Use 1M ethanolamine, pH 8.0
   b. Volume should be equal to original supernatant volume (antibody solution)
   c. Rotate overnight at 4°C (16 hours) or 2h at RT
8) Remove excess of uncoupled antibody
   a. Wash beads on sintered glass filter 4 times with alternating solutions of coupling buffer and acetate buffer (0.1M NaOAc, 0.5M NaCl, pH 4.0)
   b. Wash volume should be at least 5 times original gel volume (recall 1 g dry beads = 3.5 mL gel slurry)
9) Suspend beads in 1 gel volume of PBS pH 7.2 + 0.05% NaN3
   a. Will be 50% slurry
   b. Aliquot and store at 4°C

SOLUTIONS:

**HCl buffer:** (1mM HCl) – for 1 L
For HCl, 36, 37, 38% concentration stocks are 11.64M, 12.02M, and 12.39M

85.9 uL 36% HCl stock into 1 L milliQ H20
83.3 uL 37% HCl stock
80.7 uL 38% HCl stock

**Coupling buffer:** (0.25M NaHCO3, 0.5M NaCl, pH 8.3) – for 1 L
29.221 g NaCl
21 g Sodium Bicarbonate (NaHCO3)
900 mL milliQ H20
pH to 8.3 (w/ NaOH), then bring volume up to 1 L

**Blocking buffer:** (1M ethanolamine, pH 8.0) – for 100 mL
6 mL ethanolamine (99.5% liquid stock is 16.5 M)
93 mL milliQ H20
pH to 8.0 w/ HCl

**Acetate buffer:** (0.1M NaOAc, 0.5M NaCl, pH 4.0) – for 1 L
29.221 g NaCl
8.206 g Anhydrous Sodium Acetate (CH3COONa)
(13.608 g if trihydrous CH3COONa*3(H2O))
900 mL milliQ H20
pH to 4.0, then bring up to 1 L
B.6 Immunoaffinity Purification

1) Solubilize cells in PBS + 2% FC-14 + protease inhibitors (Roche Complete tabs) for 4 hr (rotate at 4°C)
   a. Use 1-2mL per 150mm plate, or 12.5mL per gram of bioreactor pellet
2) Spin down at 100,000g for 30 mins at 4°C
3) Transfer supernatant to new tube (discard nuclear pellets)
4) SAVE a small amount (~50uL) for western blot later – label as “total lysate”
5) Mix bead-buffer solution (either by tapping or pipetting) to form a slurry
6) Add r1D4-coupled bead slurry (Pharmacia Sepharose-4B, CNBr activated) to remainder of supernatant [Binding capacity of bead slurry is ~0.7 mg/mL (so 100uL of slurry will bind ~70ug protein)]
7) Rotate at 4°C (>4 hrs, **overnight is better**)
8) Spin down beads 2K rpm for 1-2 min at 4°C
9) Remove supernatant, SAVE a small amount (~100uL) for western blot (label as “flow thru” (probably better to save all FT in case doesn’t work)
10) Wash beads with >100 bead volumes (100mL for every 1mL bead slurry) of PBS + 0.2% FC-14 (without protease inhibitors, since proteases will wash away). Make sure beads are suspended and mixing during the rotation.
    a. Do 5 washes for 10 mins each (rotate at 4°C) – wash until absorbance 280nm is less than 0.01.
    b. (SAVE ~100uL of each supernatant (label as “wash 1”, “wash 2”, etc) and discard remainder of supernatant)
    c. For large scale purifications, use disposable filter columns for washing. After rotating, simply drain the column instead of centrifuging beads.
    d. For small batches, spin down beads (2K rpm, 1-2 min) between washes
11) Make elution buffer (PBS + 0.2% FC-14 + 100uM peptide (TETSQVAPA), without protease inhibitors) – keep at ROOM TEMPERATURE
    a. To elute, add 1 bead volume elution buffer to beads and incubate 1 hr (rotate at ROOM TEMPERATURE). Make sure beads are suspended and mixing during the rotation.
    b. Repeat elutions until absorbance of eluate at 280nm is less than 0.1 (usually 6-7 elutions are required).
    c. For large batches, continue to use disposable filter column. Drain and collect elution as above.
    d. For small batches, a microcentrifuge or PCR tube works best. Spin down beads (2K rpm, 1-2 min) between elutions fractions.
12) Used beads can be regenerated for re-use by washing the medium with 2–3 column volumes of alternating high pH (0.1 M Tris-HCl, 0.5 M NaCl, pH 8.5) and low pH (0.1 M sodium acetate, 0.5 M NaCl, pH 4.5) buffers. This cycle should be repeated 3 times followed by re-equilibration in binding buffer.
APPENDIX C:
OR GENE SEQUENCES

The following is a list of all OR genes (and corresponding protein sequences) constructed during this thesis.

All genes were optimized with human codon preferences and with 5’ EcoRI and 3’ NotI restriction sites for cloning into mammalian expression vectors such as pcDNA4/To. A Kozak ribosome binding site (RBS) consisting of the DNA sequence GCCGCCACC was placed immediately before the ATG start codon.

To enable cloning into the wheat germ cell-free expression vector pIVEX1.3, we also incorporated a 5’ NcoI site (CCATGG) at the start ATG. As NcoI requires a G following the start ATG, certain genes had to be modified by inserting either a glycine (GGC) or alanine (GCC) codon after the start ATG. This is noted in the sequences.

The majority of genes constructed contained a C-terminal rho tag sequence (GCTETSQVAPA), designated “-rho” in the gene title.

In the course of the thesis, genes using several other tag sequences were constructed and tested. The abbreviations are:

R-: N-terminal “membrane import sequence” consisting of the first 20 amino acids of bovine rhodopsin (MNGTEGPNFYVPFSNKTGV-)

HF-: N-terminal His/FLAG tag sequence (MRGSHHHHHGDDYKDDDKG-)

HFR-: HF plus an N-terminal rhodopsin “membrane import sequence” MRGSHHHHHGDDYKDDDKG-MNGTEGPNFYVPFSNKTGV-
hOR17-4-rho (OR1D2)

*already contains an NcoI site (no modification necessary)*

**DNA**

```
 1  CCTGAATTCCG   CGCCACCACAG   GGACGGAGGC   AACAAAGCGG   AGGGCAGCCGA
 51  GTTTCGCTGTG  CTTGCAATCTG  CGGAGGAGCT  CGCAAGACAG  CAGATCTCCT
 101  TTTGAGACGTG  TTCTGCTGCTA  CCGTGCTCAG  AAATGTCCTGT
 151  ATTTACCTCG  CTATTGCTCTG  CGACAGCAGA  CTCCTAGGGT  CGGCTACTTT
 201  CTTTCCTGGCT  AACCTCTCTGT  TTACAGACCTGT  GTTPTTTCGTCT  ACAAACACAA
 251  TCCCAAAATAT  GTCGGTCACAC  CTCGAAAGGCC  ACAACAAAGCT  TATTAGCTAT
 301  GCAGCCGCTGC  TACAGACCTGT  GTGAGGCTGTG  TGTGACCTG  TGCCCTGCTG
 351  TAATCTCGATG  CTGGCCCTGTA  TGGCTTACAG  TGGATTGTG  GCTATTGTG
 401  GCCCTCTCCA  CTATACACACA  GCTAGAGGGCT  GCTAACTGTT  CATCTGTCCT
 451  CTGCTCCGTG  GTGGGAGGCT GCAGAGGCGCC  TATCTGGTCA  CGACAGCAGA
 501  GCCGACCATCGCTGGGCATTGT  CTCTGAGCATG  CTATTAGCTC  AACCTCTCCT
 551  GCTCGTCAACCTACACAACTG  CTCTGGGAGCC  AGAGTGACCTGT  CGTACGTCCTG
 601  GCTGATTGCATCAGCAAAAAAT  TCCGTGCTGCTG  GCTATCAGCT  CAGCAAAAAAT
 651  TCAGCCTGTTACCTACAGCG  CCCCCATGATG  GCGCTCTCGG  ACAGAGACAT
 701  EFLLLGMSESFFLANLTSFTDNLILAVMAYLLMTRVTFCGFGFVIISYVLKPLHTYSVKDTPMMNPFIYS
 751  PEQQQILFWMEFLLGMSESFFLANLTSFTDNLILAVMAYLLMTRVTFCGFGFVIISYVLKPLHTYSVKDTP
 801  NVLIILAISAINYAGCLTVICILLYSLCWVNIIQINHTVLIASHLGAVSLLRNKDMHGALENVLIILAIS
 851  NVLIILAISAINYAGCLTVICILLYSLCWVNIIQINHTVLIASHLGAVSLLRNKDMHGALENVLIILAIS
 901  NVLIILAISAINYAGCLTVICILLYSLCWVNIIQINHTVLIASHLGAVSLLRNKDMHGALENVLIILAIS
 951  NVLIILAISAINYAGCLTVICILLYSLCWVNIIQINHTVLIASHLGAVSLLRNKDMHGALENVLIILAIS
1001  NVLIILAISAINYAGCLTVICILLYSLCWVNIIQINHTVLIASHLGAVSLLRNKDMHGALENVLIILAIS

**Protein**

```
 1  MDGGNQSEGS  EFLLLGMSES  PEQQQILFWM  FLSMYLVTVV  GNVLILILAS
 51  SDSLRLHTPVS  FFLANLTSFTD  LFFVTNTIPK  MLVNLQSHNK  AISYAGCLTQ
101  LYFLSSVLVAL  DLNIALVAMA  DRYVAICCP  HYTTAMSPKL  CIIILSLCWV
151  LSVLVLYLHTT  LLMTFRVPCGC  SRKHYIFCEF  MVVLRRMACS  NQQINHTVLI
201  ATGCPPFLIP  FGFPVISYVL  IIRAILRIPS  VSKKYKAFST  CASHLGAUVSL
251  FYGTGRELVMVL  KPLHTYSVKD  SVATVMYAVV  TPMNPFIYS  LRNKDMHGALENVLIILAIS
301  GRLLDKHFKR  LT-GG-TETSQVAPA
```

Mutations: additional mutant versions of this gene were made via site-directed mutagenesis (SDM) to either alter N-linked glycosylation sequences or to generate a SNP variation.

**N5Q:** The asparagine at position 5 (codon AAC) was mutated to a glutamine (CAG)

**N195Q:** The asparagine at position 195 (codon AAC) was mutated to a glutamine (CAG)

**N5Q,N195Q:** Both asparagines positions were mutated

**Q25R:** The glutamine at position 24 (codon CAG) was mutated to an arginine (codon AGG). A single-nucleotide polymorphism (SNP) causes population variation at this amino acid (Q/R).
hOR17-4 (untagged)
*already contains an NcoI site (no modification necessary)

DNA

1  CCTGAATTCCG  CCACCACCCTG  GAGCGCCTG  AACCAGTCTAG  AGGGCAGCGA
51  GTTCTCGCTGG  CTCGCCATGG  CGGAAGGCC  AGAGCAACAG  CAAATGTTCTG
101  TCTGGATGT  TCTGACATGG  TACCCTGCTA  CGGTGATTGG  CAATGCTCTG
151  CATATTCTCG  CCAATTCTCG  CAGATACAGG  CTGACACACC  AGGTGATTCT
201  TTTCTCTGCCT  AACCTGAGCT  TTACCGACAC  CTTCTTTTGCT  ACCAAACACCA
251  TCCCAAGAT  GCTGGTCAAC  CTGCAAGAAC  AACCAAGGCC  AATCAGCTAC
301  GCCGCAATCC  TCACCCACGT  GTACTTTGCTG  GCTCTTCTGTT  TGGCCCTCGA
351  CAACCTGAC  CTGCGACTCA  TGGCTTATGA  CAGGCTACGT  GCCATCTCTT
401  GCCGCTGCA  CTACACACCC  GCCATGTGAC  CTAAGCTCTG  CATCTCTCTG
451  CTCTCTCTGT  GCTGGTCTCC  CAGGCTGCTG  TACCACCTCC  TCAACACCTC
501  GTGAGATGac  AGGGTACCT  TCTGAGCTAC  TAGGAAGATC  CACTACATTT
551  TTTGCCGAAAT  GTATGCTCTC  CTGGCCTAGG  CCGCCCTCAA  TATCCAATTA
601  AACCACACC  TGCTGATGGC  CACCGCTGCT  TTGAGGCCTG  ACTACCTCTT
651  TGGCTCTGCTG  ATGATCTCTT  AGTGGCTCAT  TATCAGGACC  ATCTGGAGGA
701  TCCCTCTGCT  CAGCAAGAAG  TATAAAGCCT  TTAGCTCTTG  GCCGAGCCAC
751  CTGGGGTCAC  TGAGCTGCTG  TACATTTTCC  CAGGAAGATC  CACTACCTCT
801  GCCCCCTCAT  ACCTACAGCC  TGAAGGATAG  AATCCATTTA  CAGGCTCCTG
851  CTGCAGTCGAC  CCCATGTGCT  ATACATATTG  TCTCTCTCTT  CAGGAGGAA
901  GATATGACAGC  GCGCATTGGG  CAGGCTTCTC  GACAAGCATG  TCAAGAGGCT
951  GACATGACCC  GCCGCAAGAA  G

Protein

1  MDGGNQSEGSG  EFLLLGMSES  PEQQQILFWM  FLSMYLVTVV  GNVLIIALIS
51  SDSRLHPTVY  FFLANLSFPTD  LFFVTNTIPK  MLVNLQSHNK  AISYAGCTQ
101  LYFLSLSVLY  DLNLILVAMAY  DRYVAICCPHL  HYTTAMSPKL  CILLLSCWV
151  LSVYGLILHT  LLMTVRVTCG  SKRHYIFCE  MYVLLRMACS  NIQINHTVLI
201  ATGCIFIFIL  FGFPVISYVL  IIIRAILRIPS  VSKZYAFST  CASHLGAVAL
251  FYGTLCMVLY  KPLHTYSVKD  SVATVMYAVV  TPMMPPIYS  LRNKDGHGAL
301  GRLLDKHFKR  LT
**DNA**

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCTGAATTGC CCGCCACCAT GAACGGCACA GAGGGGCCCA ACTTCTACGT</td>
</tr>
<tr>
<td>51</td>
<td>GCCCTTCAGC AACAGAGCCG GCGTGGTGGA CGCGATGGAC GAGGGGACCC</td>
</tr>
<tr>
<td>101</td>
<td>AAAGCGGAGG CAGGAGTTT CTGCTGCTGG GCTATGCGCA GAGCCCGCAG</td>
</tr>
<tr>
<td>151</td>
<td>CAACACGAGA TCTCCTTTTG GATGGTCTGT AGCATGTAAC TGTGACCCGT</td>
</tr>
<tr>
<td>201</td>
<td>GCTGGGAAT GTTCTGTATG TCCTGCTGAT AGCTTCACGC AGCAGACTCC</td>
</tr>
<tr>
<td>251</td>
<td>ATACCCTCCT GTACTTCTTT CTGCTTACAC ACAACCTCTT TTTCTCCTGA</td>
</tr>
<tr>
<td>301</td>
<td>TTCTGCAACA ACACATCTCC AAAATGCTCT GCACACCTCC AAAAGCCACA</td>
</tr>
<tr>
<td>351</td>
<td>CAAAGCCTAT AGCTGCTCCC GCTGCTACAC ACAACCTCTT TTTCTCCTGA</td>
</tr>
<tr>
<td>401</td>
<td>GCTGGCTTGGC CCTGGAATAAT GTGATCTTCC GGTGCTTCTT TGATGACCCG</td>
</tr>
<tr>
<td>451</td>
<td>TACGTGCTTA TTTGGTGCCC TCTCCACATT ACAACAGCTT TGAACCTTAA</td>
</tr>
<tr>
<td>501</td>
<td>ACTGTCATTC CTCGCTCCGT CCGGTCTTCC GTGCTTCTAT TTCTGCTTCA</td>
</tr>
<tr>
<td>551</td>
<td>GACTCATTCA CACACTGCTC ATGACAAGAG TGAACCTTTG TGGCTTCAGA</td>
</tr>
<tr>
<td>601</td>
<td>AAGATCCACT CATTTTCTGT CGAAATGTAC GTCTTCTCCC GGAAGGCGCT</td>
</tr>
<tr>
<td>651</td>
<td>TAGCACATTT CAGATTAACC ATACCGTCTG GATGGCTTACC GGAATCTTTA</td>
</tr>
<tr>
<td>701</td>
<td>TTTCTTCAT CCCCCCGCG TTTGCTGATCA TCGTCCAGTC CTCTATTATC</td>
</tr>
<tr>
<td>751</td>
<td>AGGCCATCC TCCGGATCCC TTCCGTCAGC AAAATAATA AGCTCTTCCA</td>
</tr>
<tr>
<td>801</td>
<td>CACCTGTCG CAGCATTCTC AGGCGTCTG CCGTCTCTTT GTGACTCTTA</td>
</tr>
<tr>
<td>851</td>
<td>GTATGGTCTA TTCTCACACCT CTCCACACCT ACACGTCAAA GGAACCTCTG</td>
</tr>
<tr>
<td>901</td>
<td>GCTACAGTGA TGATGGCCTG CTGACCCACC ATGATGAACC CCTTCTATTA</td>
</tr>
<tr>
<td>951</td>
<td>CTCCCTCAGA AACACAGATA TCGATGGCCC TCTCCGAGAG CTCTGGACCA</td>
</tr>
<tr>
<td>1001</td>
<td>AACACTTATA AAGACTGACC GAGGCACAG AGACATCCCA AGTGGCTTCTT</td>
</tr>
<tr>
<td>1051</td>
<td>GCTTAAGCGG CGCGAGAGG</td>
</tr>
</tbody>
</table>

**Protein**

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MGTEGPNFY VPFSSLGTV GG-MDGGNQGS EGSEPLLLGM SESPEQQQIL</td>
</tr>
<tr>
<td>51</td>
<td>FWMFLSMYLV TVVGNVLLIL AISSDSRLHT PVYFLANLS FTDLFFVTNT</td>
</tr>
<tr>
<td>101</td>
<td>IPKMLVLQG HNKAISYAGC LTQLYFLVSL VALDNLLILAV MAXDRYVAIC</td>
</tr>
<tr>
<td>151</td>
<td>CPLHYTHTMS PKLCLLLLS CWLSSVLGLY IHTLLMTRTV FCGSRIKHYI</td>
</tr>
<tr>
<td>201</td>
<td>FCEMYVLLRM ACSVQINHNT VLIATGCFFIF LIPFGVIISS YVLITRAILR</td>
</tr>
<tr>
<td>251</td>
<td>IPSVSKKYYK FSTCASHLGA VSLFYGLCM YVLKPLHTYS VKDSVATVMY</td>
</tr>
<tr>
<td>301</td>
<td>AVVTPMMNPF IYSLRNDKMH GALGRLLDKH FKRLT-GGTESTQVAPA</td>
</tr>
</tbody>
</table>
HF-hOR17-4 (OR1D2)

DNA

1  CCTGAATTCC  CCGCCACCAC  GAGAGGCTCT  CACCACCAC  AATACCAGTG
51  AGGGCAGTAC  AAGGGCAGATG  AGGCAAGGG  CGGCAAGCA  CGGGAAGACC
101  AGTCGTGGAGG  CAGGCAATTC  CTGTGCTCTG  GCTATGGCAGA  AAGCCCAGAG
151  CAAGCAGCAA  GCCGGTCTGCT  GTGGAACAG  GTGGGCAATT  ACACCCCCGT
201  GGGTGGGCCAAT  ACACCCCCGT  AGGCAACGCA  CTCTGCCACT  AGGACGCGGA
251  TTTTGAACTC  ACAAGGACGAA  CAAGGATGGGT  GTCCATGCTG  AGGACGCGGA
301  CAAATTTACC  CACCACCATC  GCATACCCTG  GACCATCTGG  CTGACGCCAG
351  ATTGTTGAGG  GTGGAATGGG  CTGGCTGCTTGG  ATGACCAGGG  GG-MDGGNQS
401  ATAAGGTTTG  AGGTATTTGGG  ATGGAGGCAG  ACTGGGCAATC  ACGG-DYKDDDG
451  CTCCCTGCTC  GCTGGCTTGGC  AGATGATGACT  ACTGGGCAATC  ACGG-DYKDDDG
501  TTTCTGATCC  GCTGGCTTGGC  AGATGATGACT  ACTGGGCAATC  ACGG-DYKDDDG
551  AAGATCCAGG  CTCCCTGCTC  GCTGGCTTGGC  AGATGATGACT  ACTGGGCAATC

Protein

1  MRGSHHHHHH  GDGYKDDDDK  GG-MDGGNQS  EGSEFLLGLM  SESPEQQQIL
51  FWMDLALNNNL  TVVGNVNLIL  AISSDSRLHT  PYYFLANLS  FDLFFVFVTNT
101  IPMLVNLQG  HNKAIYSAYG  LTQLYFLVSL  VALDNILAVR  MAYDRYVAIC
151  CPLHYTTAMS  PKLICLLLRL  CWLVSVLGLY  IHTLLMTRT  FCGSKRIHYI
201  FECDYVLLRM  ACSSNQINHT  VLCATGCFFF  LIPPGFIIS  YVLIIRAILR
251  IPWSVKKKYYA  FSTCSASHLGA  VSLFYGLTLC  VYLKPLHTYS  VKDSVATVMY
301  AVTPPMNPF  IYSLRNKDMH  GALGRLLDKH  FKRLT
**HFR-hOR17-4 (OR1D2)**

**DNA**

1. CCTGAATTGC  CCGCCACCAC  GAGAGGCTCT  CACCACCAC  ATCACCATGG
2. AGGCCGACGAC  AAGGACGAGG  AGGACAAGGG  GGGCATGAA  GGGACCGAGG
3. GGGCAAAATTC  TCTGGAGATG  TGATGATTCT  GCGCAATCTCC
4. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  TCCGATAGCA
5. AGGCGACTAC  GCCCCAACTT  ATGGACGGCG  201
6. TGTACCTCGT  GCCCCAACTT  ATGGACGGCG  251
7. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  301
8. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  351
9. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  401
10. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  451
11. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  501
12. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  551
13. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  601
14. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  651
15. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  701
16. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  751
17. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  801
18. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  851
19. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT  901
20. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC  951
21. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1001
22. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1051
23. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1101
24. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1151
25. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1201
26. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1251
27. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1301
28. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1351
29. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1401
30. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1451
31. ATGGACGGCG  GTCCGAAAGC  TGTACCTCGT 1501
32. GCCCCAACTT  ATGGACGGCG  GTCCGAAAGC 1551

**Protein**

1. MRGSHHHHHH  GGDYKDDDDK  GGMNGTEGPY  FYVFPFSKTA  VVGG-MDGGN
2. QSEGSEQLLL  GMSESEPEQQQ  ILFWMPFSMY  LVTVGNNLI  ILAISDSRSL
3. HTVYFPLAN  LSFTDFLFFVT  NTIPKMLVNL  QSHNKAISYA  GCLTQLFYLV
4. SLVALDNLIL  AVMNADYRVY  ICCPLHYTTA  MSPKLICILL  SLCWLSLVL
5. GLLHTLAAT  VFTCGSRIKH  YIFCEMVVLL  RMACSNIQIN  HVTLLATGCFC
6. IFLIPFQFVI  ISYVIILRAI  LRPVSVSKKY  KAFSTCASHL  GAVSLFYGTL
7. CMVYKLPLHT  YSVKDSVATV  MYAVVTPMMN  PFIYSLRKNK  MHGALGRRLLD
8. KHPKRLT
S51-rho (mOR40-1)
modified to have an NcoI site — inserted alanine GCC after ATG

DNA
1  CCTGAATTCG CGGCCACCAG CCCATGCGTC GACGCCACCA TGGAGGCCCA
51  GAGCAACACTGCTGCGTC GCCATCCGAG GATGGTGCCA CTGTGAGGTC
101  CCACTACCA GACTGCGGAG CATCGTGGCA GTGCCCGCTC GTCCTCTTCC
151  TTTCTGTGGG CTATGGGCGA CAAAGGCCCT CCTCGTAGCA CCACTCCGGAT
201  GGAAGGCAAGC TGCGACGAGC CCAATGTACG CTTGCTGAGCC GCTCTGAGCC
251  TGCTGGCTGC CGTGGCTGTC GGGACCTGAG CTTGCTGAGCC GCTCTGAGCC
301  TTCTGGCTGC ATACACAAGC TACGGGCTTC AGCAGCTGCT TCTGAGGAT
351  GTGCTGTGATG AACTCTTTC TGACTATGGA GAGCTGACCC TTCACTGAGT
401  TGCCCTAGCG CGGTAGCGTA GGCTCTGCAG GCCATCTTCG TGCCCGCTCG
451  ATCATCCAGG ATCAGGCTGT GGTGCGGAGC GCCATCTTCG TGCCCGCTCG
501  GAACGGCATT CTGACATGC CCACTCCTAT CCTGAGACGC CAGCTGCCTT
551  ACTGGGCCCG GATGACCGGG AACTGCTACT GCAAACACAT GAGCTGAGCC
601  AAGCTGAGCT CGGCCAGCAT CACCTCTGAG AAGCTGCTACT GCAAACACAT GAGCTGAGCC
651  CGGCCCTGAGCT CTCCTGGCGA GCGACCTGAT CCTGATCGTC CTGCTCTACA
701  GCCTCATGCTG GAAAGGAGCT GTCGAGATCT GCCAAGACAT GACGCTGAGCC
751  AAGGCCCTTG AGCCACTGGC AAGGGTCCTC ATCTCTCACG TGTTCCTTCA
801  CACCCTCTTG TCTGCTGTGG TAGACCAAGA CTTGGCCGAG GCAGGATCC
851  CCCCCGACCT GCCATCTCTG TGAACATGC TGACACCCTC GACCTCTCACG
901  GCCCTGACCC CATCGTGTAA GCGGTCGCGG ACCCGGGAGG TTAACCGAGG
951  CATCAGAACC CTGGCTGGGG GGAAGCTGGG GCCACCGGAG ACCACCGGAG
1001  TGGCCCCCGG CGGCAGGAG

Protein
1  MAIRRQHMEA QSNTSSILAP DFLLICFPNY QTWQHWSLTP LSSLFLLLAMG
51  ANATLLITIR MEASLHEPMY YLLSSLSSLIL IVLCLTVIPK VLAIFWPFDNK
101  SIGFSSCFLQ MFMVMFSFIM ESCTFVMVMA DRYVAICKPL QYPSIITDFQ
151  VVRAAFIVAA RNGILTMPIP ILSSLQRYCA RIIPNCTN IMSKSLSCDD
201  ITFNLLEQPF IGWWLLGSDL ILIVLSYISP ILKAVLRKAE GAVAKALTC
251  GSHFLLLIFH STYLLVLVIR NLARERIPPD VPIILNLLH LIPPALNPIV
301  YGVTRREIKQ GINLRLRRRL-GGTETSQVAPA
**IG7-rho (MOR276-1)**

*already contains an NcoI site (no modification necessary)*

**DNA**

```plaintext
1  CCTGAATTCC  CGGCCACCAGT  GGGAATCTGG  CTGAACGAGA  GCAGCGTGGA
51  CGGCTTCATC  CTGCTGGGCA  TCTTCAGCCA  GTATCATCAG  GCCAGTGGCC
101  TCTTCTCCAC  CAGGACGGTC  AAGCTTCGGT  CTGTGGCGCA  GTGCGCCGCA
151  CTGCTGATCCA  TCCGGCCGAA  CTGCTGACCG  CTCCCCGCCG  GCCAGCGCTT
201  TGCTGATTTA  AGCCGACCGA  CAGCGTGGCC  GCTGCTGGGC  GCTGCTGCTG
251  ATCGGCCGCC  TTGCTCTGCC  CTGCTGATCC  ATCGCTGGCA  CGCGGCGCCT
301  CGTGATGCC  CTGCAGAGGT  GCTGCTGCC  GCAGCGTGGA  GCCAGCGCTT
351  AAGCTGGGGC  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
401  GTGCTGGGCC  CAGCGACCGA  CTGCCGCGAC  GCCAGCGCTT
451  TCCTTCAGCCA  TGATCATCAG  GCCAGTGGCC  GTGCGCCGCA  GCCAGCGCTT
501  CTGTGGCGCA  GTGCGCCGCA  GCCAGCGCTT
551  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
601  CGCGGCGCCT  GCCAGCGCTT
651  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
701  CGCGGCGCCT  GCCAGCGCTT
751  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
801  CGCGGCGCCT  GCCAGCGCTT
851  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
901  CGCGGCGCCT  GCCAGCGCTT
951  ATCGCTGGCA  CGCGGCGCCT  GCCAGCGCTT
1001  CGCGGCGCCT  GCCAGCGCTT
```

**Protein**

```plaintext
1  MGIWLNESSV  DGFILLGIFS  QSQTDDLFS  TVNLFVTVAL  CGNVLLILLI
51  YTDPRLHTPM  YFFLSQSLSM  DLMVCNIVP  KMAVNFLSGR  KSISFACGGI
101  QIGFFVSLVG  SEGLLLGLMA  YDRYVAISHP  LHYPIILMSQK  VCLQIAGSSW
151  AFGILGDIIQ  MAVAMSPLYC  GSRYIDHFFC  EVPALLKLAC  ADTSLFDTLL
201  FACCVFMLLL  PFSIIVTSAI  RILGAVLRLHB  SAQSRKKAL  TCSHSLTAVS
251  LFYGAMFIY  LRPRRYRAPS  HDKVVSIFYT  VLTPMLNPLI  YSLRNREVMG
301  ALRKLGLDRCR  VGSQH-GTGETSQVAPA
```

157
M71-rho (mOR171-2)
*modified to have an NcoI site – inserted glycine (GGC) after start ATG

DNA

<table>
<thead>
<tr>
<th></th>
<th>CCTGAATTTCG</th>
<th>CGGCCACCAT</th>
<th>GGCGACCGCT</th>
<th>GAAAAACAAA</th>
<th>GCACCGTGAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>GCAGTTTATGC</td>
<td>CTCGGAGGCA</td>
<td>TCACAAATAG</td>
<td>GCTGAAACTG</td>
<td>CACCTCCCCC</td>
</tr>
<tr>
<td>101</td>
<td>TCTTCTCTCTG</td>
<td>GACGGCGGCA</td>
<td>ATCTACGTGG</td>
<td>TGCAATTGG</td>
<td>GGCAATCTG</td>
</tr>
<tr>
<td>151</td>
<td>GGCATGATCT</td>
<td>CTCGGAGGCA</td>
<td>TCCTCAAGCA</td>
<td>CAACTACAGC</td>
<td>CCCCATATTG</td>
</tr>
<tr>
<td>201</td>
<td>CTCTTCTTCTG</td>
<td>AGCAACCTGA</td>
<td>GCTGACGTGA</td>
<td>CTTCTGCTAT</td>
<td>TCCACGCTCA</td>
</tr>
<tr>
<td>251</td>
<td>TCACACCCAAA</td>
<td>ATATGCTGAT</td>
<td>AACTTCTGGA</td>
<td>GCGAAAGGAA</td>
<td>GCCATAGTAC</td>
</tr>
<tr>
<td>301</td>
<td>TATCTGGGATG</td>
<td>GCTAGAGCGA</td>
<td>CTCGTATTTC</td>
<td>TTCCCTGCTG</td>
<td>TCCTGATGC</td>
</tr>
<tr>
<td>351</td>
<td>CGAGTGTGCTAC</td>
<td>AGTCTGCAAG</td>
<td>TGAATGGCCA</td>
<td>TGCAGGTGAC</td>
<td>GTCGCAATCT</td>
</tr>
<tr>
<td>401</td>
<td>GCCAGCCCTT</td>
<td>GCTCTACAAC</td>
<td>ATATCTATGT</td>
<td>CCAAGATAGC</td>
<td>GTGACAGCC</td>
</tr>
<tr>
<td>451</td>
<td>CGGTGGGCTTT</td>
<td>TCGTCTACGC</td>
<td>TGTGGGCTTA</td>
<td>ATGCGAGCCG</td>
<td>CATGAGAC</td>
</tr>
<tr>
<td>501</td>
<td>CGGACCTGATG</td>
<td>CTGAAGCTCA</td>
<td>ACTACTGCA</td>
<td>GACCTGATC</td>
<td>TCCCACTACT</td>
</tr>
<tr>
<td>551</td>
<td>TCTGGGCACT</td>
<td>CCTGCCACTC</td>
<td>ATGAAATGTT</td>
<td>CCTGCTCCAG</td>
<td>CAGATAGC</td>
</tr>
<tr>
<td>601</td>
<td>GTCAGATGGA</td>
<td>CGGTGGGCTT</td>
<td>TCTGGCAGCC</td>
<td>TCGACATCA</td>
<td>TCGCTACCAG</td>
</tr>
<tr>
<td>651</td>
<td>CCTACGCCTC</td>
<td>CTGTACCTCT</td>
<td>ACGGCTCTAC</td>
<td>TCGTCCAGC</td>
<td>ATCTCAGGA</td>
</tr>
<tr>
<td>701</td>
<td>TCCCCCTTCA</td>
<td>CGAGACGGCGA</td>
<td>AGCAAAGGCT</td>
<td>TTGACACCTG</td>
<td>TGGCCAGGG</td>
</tr>
<tr>
<td>751</td>
<td>TTCGCTGCTG</td>
<td>TCGGACCTTT</td>
<td>TCTGACAGCA</td>
<td>ACTGAAATGC</td>
<td>TCGTACCTAA</td>
</tr>
<tr>
<td>801</td>
<td>GCCGACCCCA</td>
<td>GCTCTACCTC</td>
<td>TGCGACAGGA</td>
<td>GACAGTGCC</td>
<td>TCGACCTTCT</td>
</tr>
<tr>
<td>851</td>
<td>ATACCCCGAT</td>
<td>CATCCTTCTG</td>
<td>TATCACTCCG</td>
<td>TGACCTACG</td>
<td>GCGCTAGGG</td>
</tr>
<tr>
<td>901</td>
<td>AAGGAAGTGGA</td>
<td>AGACCGCCCT</td>
<td>GGACACAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>951</td>
<td>CGGACCCCGA</td>
<td>ACAAACAGGAAA</td>
<td>TGCTCCCGGA</td>
<td>ATGAGCAGGA</td>
<td>GCGAGAAG</td>
</tr>
</tbody>
</table>

Protein

<table>
<thead>
<tr>
<th></th>
<th>MGTAENNQSTV</th>
<th>TEFILGGLTN</th>
<th>RPELQLPFLFL</th>
<th>LFLGIYVVTM</th>
<th>VGNLGMITLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>GLNSQHHTPM</td>
<td>YFFLSNLSTL</td>
<td>DCLYSVITTP</td>
<td>KMLINFVSQ</td>
<td>NLISYVGCMS</td>
</tr>
<tr>
<td>101</td>
<td>QLYFLFLFVI</td>
<td>AEYMLTVMMA</td>
<td>YDRYVAICQP</td>
<td>LLYNIMSPA</td>
<td>LCSLVAFYV</td>
</tr>
<tr>
<td>151</td>
<td>AVGILSAIE</td>
<td>TGLMLKLYNC</td>
<td>EDLISHYJCD</td>
<td>ILPLMLKSCS</td>
<td>STYDVEMAVF</td>
</tr>
<tr>
<td>201</td>
<td>FLAGFDIIVT</td>
<td>SLTVLISYAF</td>
<td>ILSSILRISS</td>
<td>NEGRSFKAST</td>
<td>CSSHFAVGL</td>
</tr>
<tr>
<td>251</td>
<td>FYGSTAFMYL</td>
<td>KPSTASSLFL</td>
<td>ENVASVFYTT</td>
<td>VIMPFNPLY</td>
<td>SLRNKEVKT</td>
</tr>
<tr>
<td>301</td>
<td>LDKTLRRKVF</td>
<td>GGTETSQVAPA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
OR23-rho (mOR267-13)
*modified to have an NcoI site – inserted glycine GGC after ATG

**DNA**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCTGAATTCG</td>
<td>CGGCCACCAT</td>
<td>GGCGCGAGCG</td>
<td>AACAATTTCG</td>
</tr>
<tr>
<td>51</td>
<td>GAGTGTGCGT</td>
<td>TTCCCTCGGCT</td>
<td>TTCCATGGCT</td>
<td>CGGGAAACAC</td>
</tr>
<tr>
<td>101</td>
<td>TGTTGCCTAG</td>
<td>GTTCTGAGCA</td>
<td>ATCTACATCC</td>
<td>TCAACCGTCG</td>
</tr>
<tr>
<td>151</td>
<td>ATTATGTCAG</td>
<td>CAATCCACCA</td>
<td>CATTGCACG</td>
<td>CACCTGCACA</td>
</tr>
<tr>
<td>201</td>
<td>CTTTTTCTTG</td>
<td>TCCATGCTGG</td>
<td>CTAGCTCAGG</td>
<td>GACGCCTTAC</td>
</tr>
<tr>
<td>251</td>
<td>TTGTGCCTAG</td>
<td>AATGCCTGTC</td>
<td>AGCCTGATCT</td>
<td>TTTACACCTC</td>
</tr>
<tr>
<td>301</td>
<td>CTCCGGCGGT</td>
<td>CGCCACACAG</td>
<td>GATGTTTCTC</td>
<td>TTCGAGAGCT</td>
</tr>
<tr>
<td>351</td>
<td>CAACCTGCTT</td>
<td>CTCCTGACCC</td>
<td>CATATGCTGG</td>
<td>TGCCCATGAA</td>
</tr>
<tr>
<td>401</td>
<td>GCCAACCACG</td>
<td>TACGGTACCT</td>
<td>ATCATCATGA</td>
<td>GAAAGCGCAT</td>
</tr>
<tr>
<td>451</td>
<td>CTCGTCTGGC</td>
<td>GAGCGCCGCT</td>
<td>CACCGGCTGA</td>
<td>GTGATGCGCT</td>
</tr>
<tr>
<td>501</td>
<td>CCGCGCCATG</td>
<td>TTCCACCTCC</td>
<td>CTTCTGCTG</td>
<td>TCCGACCTGA</td>
</tr>
<tr>
<td>551</td>
<td>TCCTGCGGGC</td>
<td>GGAGCGGCCG</td>
<td>AGGACGCACT</td>
<td>AACAGGCAAG</td>
</tr>
<tr>
<td>601</td>
<td>AAGCAGGATTA</td>
<td>TCAATTACGG</td>
<td>GTGAGCTCC</td>
<td>TTTGCTACCT</td>
</tr>
<tr>
<td>651</td>
<td>CGGCCCTACG</td>
<td>TTCCATGCGT</td>
<td>ATGTGCTGAT</td>
<td>TCGAGCTCTC</td>
</tr>
<tr>
<td>701</td>
<td>TGGTGCACAC</td>
<td>CGAGGGCGAA</td>
<td>AAGAGGCGTT</td>
<td>TCGACAGCAG</td>
</tr>
<tr>
<td>751</td>
<td>CTCAAGGCTT</td>
<td>TCATCGTCCCA</td>
<td>CTACGGCTGC</td>
<td>GAAACGATTG</td>
</tr>
<tr>
<td>801</td>
<td>GCCCAAGAGC</td>
<td>GAATCCCTGG</td>
<td>TCAGAGAGGA</td>
<td>TCGAGCTTCT</td>
</tr>
<tr>
<td>851</td>
<td>AATACGATCT</td>
<td>TACCCCTCTG</td>
<td>AGGACGCCCT</td>
<td>ACCAGCCAAG</td>
</tr>
<tr>
<td>901</td>
<td>AAGGAGGTGA</td>
<td>AGGACGCCGG</td>
<td>GTGGAGAAGA</td>
<td>ATACCTGCGG</td>
</tr>
<tr>
<td>951</td>
<td>CGGACGCTGA</td>
<td>ACCAGCCAAG</td>
<td>TGGCACCTGC</td>
<td>CTAAGCGGC</td>
</tr>
</tbody>
</table>

**Protein**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MGQRNFTEVE</td>
<td>IEVPLGFSS</td>
<td>FGQHQITLFV</td>
<td>VFLTIYIIIT</td>
</tr>
<tr>
<td>51</td>
<td>HIDHLHTLTPM</td>
<td>YFFLSMLASS</td>
<td>ETQYTVIYV</td>
<td>RPMSSLIFYN</td>
</tr>
<tr>
<td>101</td>
<td>QMFFVVLAT</td>
<td>NNCFLATMAG</td>
<td>YDRYVAICNP</td>
<td>LRYTIIMSKG</td>
</tr>
<tr>
<td>151</td>
<td>GTLVMVAVLH</td>
<td>VPAMPHLPFC</td>
<td>GTVEHFFCDD</td>
<td>IYPVMKLCV</td>
</tr>
<tr>
<td>201</td>
<td>GVSSPVILVP</td>
<td>GLIPISYVL</td>
<td>IVSSILKIVS</td>
<td>TEGQKAFAT</td>
</tr>
<tr>
<td>251</td>
<td>HYGASIAVL</td>
<td>KPKSSSSEVEK</td>
<td>DLLLSVTYTI</td>
<td>ITPLLNPVYY</td>
</tr>
<tr>
<td>301</td>
<td>LCRAVGRNTS-GGTETSQVAPA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**17-rho (mOR103-15)**
*already contains an NcoI site (no modification necessary)*

**DNA**

```
1  CCTGAATTCCG CGGCCACCAT GGAACGGAGA AATCACACCG GACGGGTGAG
51  CGAGTTTGTC CTGCTCGGAT TCCCCGCACC CGCTCCTTCG CGGGACTGC
101  TCTTTTTCTC GCCAACATGT CTTTTGGA GATCTGTTGAC GTGACAGTCA
151  CACATTATTAGAGAATCCG TTCCCCACCTTCCGCT GCAGTTTGTC
201  CAGCTCATTAGCCGAACTGGGAC CGAGTTTGTC CGAGTTTGTC
251  CCGATTCCCAA CCAGCTCATTA CCGCGGATGC CCGCATCTCATGC
301  CCTCGGATGTC ACGCGGTTGTC CGAGTTTGTC CGAGTTTGTC
351  AGCTTGGCCCAT CTGCTCAGGTC ACCTTTTCTC CCGATTCCCAA
401  CCCTCCTGCTG AGGTGTCAAG CCATCAACCA ACTGATATGA TCTGCTGGGG
451  GCCAACATGT CAGGCTGCTA CTTCATTTAT TGGTCAGCGT TACTGTCTGA
501  GCCAACATGT CACGGGTGAGT GCCGATCCCA GCCTCTATGCT GAAACCAGGA
551  GCCAACATGT CGGGAACCGA PAPLRALLFF EIWYVTVTIP AVMAYDRYVA
601  GCCAACATGT CGGGAACCGA LSLLAYVLVL TENILITIAI KMLAGFIGSE ENHQLISFE
651  GCCAACATGT CGGGAACCGA AATCACACCG GACGGGTGAG TCTCTCTGCC
701  GCCAACATGT CGGGAACCGA TCGCATCTGC AGGTGTCAAG CCATCAACCA
751  GCCAACATGT CGGGAACCGA AATCACACCG GACGGGTGAG TCTCTCTGCC
801  GCCAACATGT CGGGAACCGA LSLLAYVLVL TENILITIAI KMLAGFIGSE ENHQLISFE
851  GCCAACATGT CGGGAACCGA AATCACACCG GACGGGTGAG TCTCTCTGCC
901  GCCAACATGT CGGGAACCGA LSLLAYVLVL TENILITIAI KMLAGFIGSE ENHQLISFE
951  GCCAACATGT CGGGAACCGA AATCACACCG GACGGGTGAG TCTCTCTGCC
1001 GCCAACATGT CGGGAACCGA LSLLAYVLVL TENILITIAI KMLAGFIGSE ENHQLISFE
```

**Protein**

```
1  MERRNHHTGRV SEFVLLGFPA PAPLRALLFF LSLLAYVLVL TENILITIAI
51  RNHPNLHKPM YFPLANMSFL EIYVVTITIP KMLAGFIGSE ENHQLISFE
101  ACMQMDFLPL GLGCCTCVCW AVMAYDRYVA ICHPLYPVI VSSRLCVQMA
151  AGSWMAGFFG SIYKVLISLR LSYGEPNTIN HFFCDVSPLL NLNCTDMSTA
201  ELTDPIFALIF ILLGPLTSVTG ASYMAITGAV MRRSAGRRH KAPSTCSHL
251  TVVFIPFAS IFIYARPKAL SAFDTNKLVS VLYAVIVPPLL NPIIIYLRNQ
301  EVKKALRRTL HLAQGQDANT KKSRRDGGTGTSQAPA
```

160
ODR-10-rho (C. elegans)
*modified to have an NcoI site – inserted glycine GGC after ATG

DNA

1  CCTGAATTCG CGGCCACCAG GGGCAGCGGC GAGCTGTTGA TTACCCCTCG
51  CGACACCGCC GAGTTCGCTG GGTCAACCTC CACATTTCGT GTGGAATATT
101  TCTCCTCTGG CCTGCTGAGA ACTGCGGGGA AGAACCTCGG GACCCTACAG
151  TACTTACTGG CCTGTTCAAC ATCTTTTACG CAATCATCGA
201  GTTCATCTCT AGGCAATTA TCGACATCGA GAACACCCAC TTCTTTCTCA
251  TCAGCGAAA AGGTCGCTAC TACGACACC ACGTCCGCAA GATCAACAC
301  GCCCTTTACT GGGCCTGTTG CGCCACCTCC TTTTCGCTGA GCACCGTGC
351  CCTGCTCTAT AGGTTACCTC CAACCTGCAA GCTAACTCCT CGTGCGTTG
401  TAACCTCCAC CACTGCTGT GCTCAGGCCC TGCGTGCTAG CGTCCGACTG
451  ACCATGTGCG CAGCGCTCTC TACCTTCCGT TAACCAGATA CGGAGTACAC
501  CGAGCGCGCT GTGACCAACAC TGGTAACTAA CCACTACAC TGGAATCAAG
551  AGGAGAACGT GAGCTACATC GCCTACGTGT ATACCAAATA CGAGAAGGCC
601  GTGAGGCGATA TCTATCTGAA AAACCTGCTG GGCTGCTTCC TCCACTACTT
651  CGTGATGAGG ATGACCTTGG TGTTGATGTT CTACTGCGGG TACCCGACTT
701  GGAAGACCAT GAACGACGAC AAGACGCTGT GGACAGCAC CCAGCACCTG
751  CAAAGACGCG TGTTTAAGGC CCTGCTGCTC CAAACTCTCA TCCCACCAC
801  TTTCTGATAC GCCTCCACGC GGGTCGAAGA CATGCGCCCA TTCCTCGACG
851  TGACACTGAA CGTAAATGCT AACCTTATGG TGTTTGTCTC TCTCTCTCAC
901  CTTGCCCTGG ACCAATCTAC TGATACGAGG ATCATACGGG ATCTACGGG
951  GACCATCTTC ATATTCCCTGC GTGCAAGAAG AACAGCGCGT GATGAAACCC
1001  GTGCCACACCG TCGGGCTAAT CTGACAGCAGG TGGACATCGG GCGGACTGAA
1051  ACCAGCCCGCT TGGCCCAAGG CTAGCCGCCC GGAGAAGG

Protein

1  MGSGELWITL VDTADIVGVT LTFCVNVILL GLLKTRGKNL GTYKYLMAFF
51  SVFISFYAII EFILRPMHJ ENTTFPLISR KRFXNSTKLG KINSFYACAF
101  FATSFVVSQG HPVRRYFAPC KLKYNLTHR PNLLRLPNLP TTLWPLGCS
151  VPVTNOWASYS YFLYPDTEYT EAAVTNLNN HYNWIKKENV SYIAVYYQY
201  ENGRHYIYLK NLGCFVHYF VMSMTFVVMF YGYATWKTMD NEHSDKVDRT
251  RALQQLFKA LVQLTIPTF MYAPTGVMP IAPFPDNLNN ANANFIVFCS
301  FLYPLDPLI LILIIRDFRR TIFNFGCGKK NSVDESRTTT RANLSQVPTG
351  GGTETSQVAPA
**OR73-rho (mOR-EG, mOR174-9)**

*modified to have an NcoI site — inserted glycine GGC after ATG*

### DNA

<table>
<thead>
<tr>
<th>1</th>
<th>CCTGAAATTCG</th>
<th>CCAGCCACCAT</th>
<th>GGGCACCCCTG</th>
<th>AGCGACGGGA</th>
<th>ACCACTCTGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>GAGATGTTTC</td>
<td>ACCCTGCTCG</td>
<td>GGTACGCGCA</td>
<td>CTGACCCGAA</td>
<td>CTGACCATCC</td>
</tr>
<tr>
<td>101</td>
<td>GATCTTTTCT</td>
<td>GATCTTTTCT</td>
<td>ACTATCTACA</td>
<td>GATCACCGGT</td>
<td>GATCGGCAC</td>
</tr>
<tr>
<td>151</td>
<td>ATCCGAAATG</td>
<td>TCTGATCATC</td>
<td>AGGATTAAC</td>
<td>CCCAGCTCC</td>
<td>AATCCCCCAT</td>
</tr>
<tr>
<td>201</td>
<td>GTCTTTTCT</td>
<td>CTCAAGGACC</td>
<td>TGTAGCCTGT</td>
<td>GAGCTTCTGC</td>
<td>TACTCTCTCT</td>
</tr>
<tr>
<td>251</td>
<td>CACTCTCTCT</td>
<td>CGAATTTTCC</td>
<td>AAGCTGCTGC</td>
<td>ATTATCCTGC</td>
<td>GTCCGCCAGC</td>
</tr>
<tr>
<td>301</td>
<td>AGCTTTTTTG</td>
<td>GGGCCCTGCT</td>
<td>GAGTTCTTTT</td>
<td>TTTCTTTGCA</td>
<td>CTTGCTGCTC</td>
</tr>
<tr>
<td>351</td>
<td>ACGGAGGCC</td>
<td>TTCTCTCTGT</td>
<td>GAGTCATGGC</td>
<td>ATACAAGAG</td>
<td>TTTGGGCGCA</td>
</tr>
<tr>
<td>401</td>
<td>TCGGGAATCC</td>
<td>CTTCTTCTAC</td>
<td>ACCTGCCCCT</td>
<td>TCAAGGAAAG</td>
<td>ACCTGACCTC</td>
</tr>
<tr>
<td>451</td>
<td>ATCTCTGCTC</td>
<td>TCGAGACTGA</td>
<td>GCTGCTGGGC</td>
<td>GTGCTGCTCA</td>
<td>GCTGAGACTC</td>
</tr>
<tr>
<td>501</td>
<td>GACCTGAGCC</td>
<td>GCCCTCCCAC</td>
<td>TCTCCTCTTC</td>
<td>CGCTCTCAAC</td>
<td>ATGATCAATC</td>
</tr>
<tr>
<td>551</td>
<td>AGCTCTTTCT</td>
<td>CGCAATTTTT</td>
<td>AGCTGCTCTC</td>
<td>CCTGCTCCAG</td>
<td>GTCCGGATAC</td>
</tr>
<tr>
<td>601</td>
<td>AGCGTGTGCC</td>
<td>AACTGCTGCT</td>
<td>CTCTCTTCTT</td>
<td>GCCACTTCTC</td>
<td>AGGAGATACG</td>
</tr>
<tr>
<td>651</td>
<td>CACTCTCTCT</td>
<td>TTATATCTGC</td>
<td>TGCTCTATTG</td>
<td>GCTCGTACAT</td>
<td>GTGACAATCC</td>
</tr>
<tr>
<td>701</td>
<td>TCAAGATGGA</td>
<td>GTTGCCGAGC</td>
<td>GGAGGCCAGA</td>
<td>AAGCTTACAG</td>
<td>CACCTGGCC</td>
</tr>
<tr>
<td>751</td>
<td>CCTCATCCTG</td>
<td>CGCTATACCA</td>
<td>CATCTCCAC</td>
<td>GCCACCATCC</td>
<td>TCTCTCTGTA</td>
</tr>
<tr>
<td>801</td>
<td>CTCGCTGCC</td>
<td>AACAGCAAGA</td>
<td>ATAGCAGGCA</td>
<td>TACTGTAAGG</td>
<td>GTGCGCAGCG</td>
</tr>
<tr>
<td>851</td>
<td>TGCTCTATAC</td>
<td>TGTCGCTCATC</td>
<td>CCTATGTCA</td>
<td>ACCCCCTGAT</td>
<td>TTACTCTCTG</td>
</tr>
<tr>
<td>901</td>
<td>CGCAATAGAG</td>
<td>ATTCATGAGA</td>
<td>CACCTGCAAG</td>
<td>AAGATTATCG</td>
<td>GGAATACAGT</td>
</tr>
<tr>
<td>951</td>
<td>GTCAAGCTCT</td>
<td>GGAGGCCACC</td>
<td>AGACCAGCCA</td>
<td>GTGCGCCACC</td>
<td>CCTGAGGCGG</td>
</tr>
<tr>
<td>1001</td>
<td>GGCGCAGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Protein

<table>
<thead>
<tr>
<th>1</th>
<th>MGTLSDGNHS</th>
<th>AVFTLLGFS</th>
<th>YPELTIPFPL</th>
<th>IFLTLYSITV</th>
<th>VGNIMIVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>RINPKLHMP</td>
<td>YFFLSHLSFV</td>
<td>DFCYSSIVAP</td>
<td>KMLVNLVTMN</td>
<td>RGISFVGCLV</td>
</tr>
<tr>
<td>101</td>
<td>QFFPCTFV</td>
<td>TESFLGVM</td>
<td>YDFVAIRNP</td>
<td>LLYTVAMSQR</td>
<td>LCAMLVLGSY</td>
</tr>
<tr>
<td>151</td>
<td>AWGVCSLL</td>
<td>TCSALNVGM</td>
<td>YFNMNHFFC</td>
<td>EFSLLSLSR</td>
<td>SDTSVSQLLL</td>
</tr>
<tr>
<td>201</td>
<td>FVFATFNEIS</td>
<td>TLLILLLSYV</td>
<td>LVVVTLKMK</td>
<td>SASGRRKAFS</td>
<td>TCASHLTAIT</td>
</tr>
<tr>
<td>251</td>
<td>IFHGTIFLY</td>
<td>CVPSNSRHR</td>
<td>TKVYAVFYT</td>
<td>VVPPMLNPLI</td>
<td>YSLRNKDVK</td>
</tr>
<tr>
<td>301</td>
<td>TVKKIIGTKV</td>
<td>YSS-GTETSQVAPA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HFR-OR73 (mOR-EG, mOR174-9)

DNA

1   CCTGAACTCG  CCACCATG  GCAGGCGAGC  CATCACCACC  ATCATCACGG
51  AGGCGATGAC  AAGGACGAGC  ATGACAAGGG  AGGATGAATG  GGCACCGAAG
101 GCCCACAATT  CTACGTCCCC  TTTAGCAACA  AGAAGCGGCT  CGTGAGGGGC
151 ATAGCCCTGA  GCAGCGGGAA  CCACTCTTGG  GCAGGTTCA  CACTGCTCGG
201 GTTCCAGCGAC  TACCCCAAGG  TGGACATCCC  CCTGTTCTTG  ATCTTTCTGA
251 CTATCTACAG  CATACCCGGT  GTCGCAACA  TGGAATGATG  CGTATCATC
301 AGGATTAACC  CCAAGCTCCA  CATCCCCAGT  TACTTTTTCC  TCAGCCACCT
351 GAGCTTCCGTG  GACTTTCTCT  ACTCCCTAT  CGTGCCCCCC  AAGATGCTCG
401 TCAACCTCGT  GACCATGACC  GCGGCCAGCA  GCCCTCTGGG  GTCCCTGGTG
451 CATTTCTTTT  TCTTTTCGAC  CCTGCTCTGT  AGCGATCTCT  TCTCTGGTG
501 AGTCATGGCA  TAGCAGCGGT  TTGGCGCATC  GCGGATCCCG  CTCTCTCTCA
551 CCGTGGCTAT  GAGCCAGAGA  CTGTCGCGCA  TGCTGTGGCT  CGGAGCTAT
601 GCATGGGCGC  TGCTCTCGAG  CCTGATCCTG  ACTGCAAGCG  CACTGAACCT
651 CTCTTTTTAC  GCCTCAACA  TAGCACTAAT  CTTCTCTGGC  GAATTTTCCA
701 GCCGCTCTCT  CCTGTCAGGG  TCCGATACAA  GCTGTCCCA  ACTGCTGCTC
751 TTGGCTCTCG  CCCATCTCAA  CGAGACGAGC  ACTGCTCTCA  TTATCTGCT
801 GTCTAGTGGT  CTCAGTGGGG  TGAACACCTC  CAAGATGAGT  TCCGCCAGCG
851 GCAGGAGGAA  AGCCCTCAAG  ACCGTGGCCT  CTGATCTAGC  CGCTATCACC
901 ATCTTTCAAG  GCACCATTTA  CTTCTCTGAC  TGCTGTCCCA  AAGCAAGAAA
951 TAGAGGCAC  ACTGTAAGGG  TGCCGCAGGT  GTCTCTACT  GTGCTCATCC
1001 ACTAGCTCAA  CCCCTGTGCA  TACTCTCTCC  GCAGTGGCAA  GTGCAAGAC
1051 ACGCTCAAGA  AGATTTACGG  GACTAAGGGT  TACAGCTTCT  GAGCCGCGGC
1101 GAGAAG

Protein

1   MRGSHHHHHH  GDYKDDDDK  GGMNTEGPN  FYVPFSNKTI  VVGG-MTLSD
51  LHIPMYFFLS  LGSFDPYELT  IPLFIPLFTI  YSITTVVGNIG  MIVIIRINPK
101 CTFVVTEFL  LGMVAYDRFV  AIRNPFLYTV  AMSQRLCAML  VLGSYAWGVV
151 CSILILCSAL  NLSFYGFMNI  NHFFCEFSSL  LLSRSRTSV  SQLLRFVAT
201 FNEISTLLII  LLSYVLIVYT  ILMKMSASGR  RKAFTSCASH  LTAAITFHTG
251 IFLYCVPNS  KNSRHTVKVA  SVFYTVVIPM  LNPLIYSLRN  KDVKDTVKKI
301 IGTKVYSS
351
**DNA**

1. CCTGAATTCG  CGGCCACCAGT  GGCAAACTCTC  ACCAACGTGA  CCGAGTTTAT
2. CCTCTCGGG  CTGAGCGACG  GACCCATGT  ACCAACGTGA  CCGAGTTTAT
3. GATCTCTCTC  GACCTAGGT  TCGAGAACTC  ACCAACGTGA  CCGAGTTTAT
4. TTCTACACAC  TCUGCAGCTG  CAGGCGTAC  ACCAACGTGA  CCGAGTTTAT
5. GAGGACTTCG  GCCATGCCTG  AGATCTGGG  CACCAGGTAG  ATCTTCCCTG
6. AGAGCCCTAC  CAACATCATG  ACTCGCCACA  AAACCAATTAG  CCGGCTCGGC
7. TCTCTCTTCC  AGGGCTTCTC  GTACTTCTCT  CTTGGGACA  CAGAGAAAAC
8. GGAGGTGCGG  TGAGGCGACG  GACCTACTTC  TGGCTGACCG  GCCATGCTCG
9. GTACGAGTTT  GCATGCCATG  AGGCTTCTAC  TCCTGGGTTA  CCCTGGGTAC
10. GTTCCTCGCA  GTGCGGTACGC  TGCAGCTGGA  GTGCACCAAG  GTCCAGCGGA
11. GAGGAACTTC  CCTCCTCGGG  GATTCCTCCT  TTCATCACAC  GAGGAACTTC
12. AGATGCTCAC  CAACATCATG  ACTCGCCACA  AAACCAATTAG  CCGGCTCGGC
13. TGGGTCTCCG  AGGCCTTCTC  GTACTTCTCT  CTTGGGACA  CAGAGAAAAC
14. TGCCGTACGC  CACAATCATG  TCGAGAACTC  ACCAACGTGA  CCGAGTTTAT
15. CGAGAAGAT  CTGGGTTA  CCCTGGGTAC  GTGCACCAAG  GTCCAGCGGA
16. MANSTTVTEF  ILLGLSDACE  LQVFLIFGFL  LTYFLILLGN  PLIIFITLAD
17. RRLYTPMYEF  LRFAMLEIW  FTSVIFPFKL  TNIITGHTIT  SLLGCFLQAF
18. LYFFLGTTEF  FLAVMSFDR  VYAICNPLRY  ATIMSKRCV  QLVFGSWSMG
19. LLIIVPSII  VFQOPFCGPN  IHNFCDNF  PLMELICADT  SLVFLPGFVI
20. ANFSLLGTTLA  VTAQCYHIL  YTLHIPSASK  ERKKAFTSCS  SHIIVVSLFY
21. GSIFMYVRS  GKNPOGEDDN  KVVALLNTVV  TPTLPFIYT  LRNQKVKurations
22. REHVSKFQKF  SQT-GGTETSQVAPA
HFR-IC6 (mOR118-1)

DNA

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>CCTGAATTTCG</th>
<th>CGGCCACCAT</th>
<th>GAGGGGCTCC</th>
<th>CACCATCATC</th>
<th>ACCACCACGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td></td>
<td>CGGAGACTAC</td>
<td>AAGGATGATG</td>
<td>ACGACAAAGG</td>
<td>CGGAATGAA</td>
<td>GGCACCTGAGG</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>GGGCAATTTT</td>
<td>CTACGTGCC</td>
<td>TTTCAACAAC</td>
<td>AGACCGGCT</td>
<td>GGTGCGGAGG</td>
</tr>
<tr>
<td>151</td>
<td></td>
<td>ATGGCAACCT</td>
<td>CTACCACTGG</td>
<td>GACCGAGTTT</td>
<td>ATCTTCTCG</td>
<td>GCTGTAGGCA</td>
</tr>
<tr>
<td>201</td>
<td></td>
<td>CGCCTGTGAG</td>
<td>TCCTCATCCT</td>
<td>CGCAGGCTGT</td>
<td>CGAGATCTGG</td>
<td>TTTACTGGCCA</td>
</tr>
<tr>
<td>251</td>
<td></td>
<td>CTGTACTTCT</td>
<td>CTTCGACAGG</td>
<td>TGTCCAAGAG</td>
<td>CTCCTGCTCA</td>
<td>CGGCCCAAAC</td>
</tr>
<tr>
<td>301</td>
<td></td>
<td>AGCTGATCTG</td>
<td>GCCAACTTTT</td>
<td>ACACATCCTG</td>
<td>AAGCATTTAG</td>
<td>GGCTCCTGCA</td>
</tr>
<tr>
<td>351</td>
<td></td>
<td>GCCAACTTTT</td>
<td>AGGGAGCACG</td>
<td>GAGAAG</td>
<td>MRGSHHHHHH</td>
<td>SDACELQVLI</td>
</tr>
<tr>
<td>401</td>
<td></td>
<td>TTVTEFILLGL</td>
<td>TVTEFILLGL</td>
<td>TVTEFILLGL</td>
<td>TVTEFILLGL</td>
<td>TVTEFILLGL</td>
</tr>
<tr>
<td>451</td>
<td></td>
<td>PMYYFLRNFA</td>
<td>PMYYFLRNFA</td>
<td>PMYYFLRNFA</td>
<td>PMYYFLRNFA</td>
<td>PMYYFLRNFA</td>
</tr>
<tr>
<td>501</td>
<td></td>
<td>GTTEFFLLAV</td>
<td>GTTEFFLLAV</td>
<td>GTTEFFLLAV</td>
<td>GTTEFFLLAV</td>
<td>GTTEFFLLAV</td>
</tr>
<tr>
<td>551</td>
<td></td>
<td>TQGCPPNNFA</td>
<td>TQGCPPNNFA</td>
<td>TQGCPPNNFA</td>
<td>TQGCPPNNFA</td>
<td>TQGCPPNNFA</td>
</tr>
<tr>
<td>601</td>
<td></td>
<td>TNCVYFRKLS</td>
<td>TNCVYFRKLS</td>
<td>TNCVYFRKLS</td>
<td>TNCVYFRKLS</td>
<td>TNCVYFRKLS</td>
</tr>
<tr>
<td>651</td>
<td></td>
<td>GAGGGGCTCC</td>
<td>GAGGGGCTCC</td>
<td>GAGGGGCTCC</td>
<td>GAGGGGCTCC</td>
<td>GAGGGGCTCC</td>
</tr>
</tbody>
</table>

Protein

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>MRGSHHHHHH</th>
<th>GGDYKDDDDD</th>
<th>GGMNGTEGPN</th>
<th>FYVPFSNKTG</th>
<th>VVGG-MANST</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td></td>
<td>TVTEFILLGL</td>
<td>SACEELQLVI</td>
<td>FLGFLLLYFL</td>
<td>ILLGFLIIIF</td>
<td>ITLADRLRTY</td>
</tr>
<tr>
<td>101</td>
<td></td>
<td>PMYYFLRNFA</td>
<td>MLEIFWFTSVI</td>
<td>FPMLTNIIIT</td>
<td>GHTISLLGC</td>
<td>FLQAFLYFL</td>
</tr>
<tr>
<td>151</td>
<td></td>
<td>GTTEFFLLAV</td>
<td>MSFDRYAIC</td>
<td>NPLRYATIMS</td>
<td>KRVQVFLVPC</td>
<td>SWMSGLLLII</td>
</tr>
<tr>
<td>201</td>
<td></td>
<td>VPSSIVFQPQ</td>
<td>FCGPNIINHF</td>
<td>PCDNPLMEL</td>
<td>ICADTSLVME</td>
<td>LGFVIANFLS</td>
</tr>
<tr>
<td>251</td>
<td></td>
<td>LGLTALATAC</td>
<td>YGILYTHLIY</td>
<td>IPSAKERKKA</td>
<td>FSTCSSIIIV</td>
<td>VSLFYGSCIF</td>
</tr>
<tr>
<td>301</td>
<td></td>
<td>MyVRSGKNNQ</td>
<td>GEDRHKVVAL</td>
<td>LNTVVTPTLN</td>
<td>PFIYTLRNKQ</td>
<td>VKQVFREHVS</td>
</tr>
<tr>
<td>351</td>
<td></td>
<td>KFPKFSQQT</td>
<td>KFPKFSQQT</td>
<td>KFPKFSQQT</td>
<td>KFPKFSQQT</td>
<td>KFPKFSQQT</td>
</tr>
</tbody>
</table>