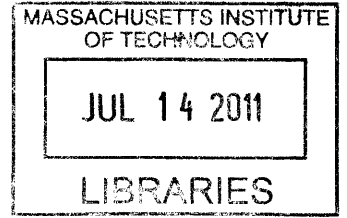


**OLFACTORY-RELATED RECEPTORS: METHODS TOWARDS ENABLING STRUCTURAL AND FUNCTIONAL STUDIES**

**Karolina Corin**  
**B.S. Mechanical Engineering**  
**Massachusetts Institute of Technology, 2003**  
**B.S. Biology**  
**Massachusetts Institute of Technology, 2003**  
**M.S. Mechanical Engineering**  
**Massachusetts Institute of Technology, 2005**



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

Doctor of Philosophy in Biological Engineering

**ARCHIVES**

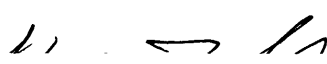
at the

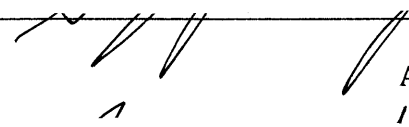
Massachusetts Institute of Technology

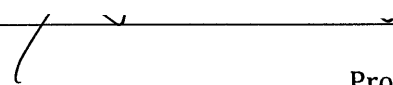
June 2011

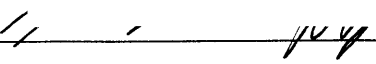
©2011 Karolina Corin. All rights reserved.

The author hereby grants 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: \_\_\_\_\_  
  
Karolina Corin  
Department of Biological Engineering  
May 13<sup>th</sup>, 2011

Certified by: \_\_\_\_\_  
  
Shuguang Zhang  
Associate Director, Center for Biomedical Engineering  
Thesis Supervisor

Certified by: \_\_\_\_\_  
  
Bruce Tidor  
Professor of Biological Engineering and Computer Science  
Co-advisor

Accepted by: \_\_\_\_\_  
  
K. Dane Wittrup  
Professor of Chemical Engineering and Biological Engineering  
Chair, Course XX Graduate Program committee

## Thesis Committee:



**Shuguang Zhang**

Associate Director, Center for Biomedical Engineering  
Massachusetts Institute of Technology



**Bruce Tidor**

Professor of Biological Engineering and Computer Science  
Massachusetts Institute of Technology



**K. Dane Wittrup**

Professor of Chemical and Biological Engineering  
Massachusetts Institute of Technology



**Alexander Rich**

William Thompson Sedgwick Professor of Biophysics  
Massachusetts Institute of Technology

# **OLFACTORY-RELATED RECEPTORS: METHODS TOWARDS ENABLING STRUCTURAL AND FUNCTIONAL STUDIES**

By Karolina A. Corin

Submitted to the Department of Biological Engineering on May 19<sup>th</sup>, 2011  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Biological Engineering

## **ABSTRACT**

Mammalian noses can detect and distinguish an inestimable number of odors at minute concentrations. Four classes of G protein-coupled receptors (GPCRs) are responsible for this remarkable sensitivity: olfactory receptors (ORs), vomeronasal receptors (VNRs), trace amine-associate receptors, and formyl peptide receptors.

Structural knowledge of these receptors is necessary to understand the molecular basis of smell. However, no structure exists for three main reasons. First, milligrams of protein are needed for crystallization screens, but most are expressed at low levels endogenously or in heterologous expression systems. Second, detergents capable of solubilizing and stabilizing these proteins in aqueous solution must be found. Third, the flexible nature of GPCRs can inhibit crystal lattice formation.

Methods for overcoming each obstacle were developed. Milligrams of a VNR were expressed in HEK293 cells, and milligrams of 13 GPCRs were expressed in a cell-free system. All could be purified to >90%. The purified receptors had correct secondary structures, and could bind their ligands. The HEK293 and cell-free receptors had nearly identical structures and binding affinities, demonstrating that cell-free expression can be used for GPCR production and mutational studies. To demonstrate this, six variants of mOR103-15 with single amino acid substitutions were expressed. Ligand-binding measurements indicated which residues were involved in ligand recognition. The choice of detergent used in the cell-free system was critical, and significantly affected expression levels. A class of amphiphilic peptide detergents was designed and tested with the receptors. These detergents could be used to express milligrams of functional receptors. The peptide tail and head group properties did not significantly affect their function, suggesting that they may be a class of surfactants usable with multiple olfactory-related receptors, and even other membrane proteins. Lastly, the protein T4 Lysozyme (T4L) was fused in the 3<sup>rd</sup> intracellular loop of two receptors to increase potential crystal lattice contact points. Purified T4L variants had correct secondary structures, and could bind their ligands and initiate intracellular signaling. The methods described generated sufficient quantities of pure receptors for crystal screens. The large number of functionally expressed GPCRs indicates that these techniques can be applied to other olfactory-related receptors, and even other membrane proteins.

Thesis Supervisor (Advisor): Shuguang Zhang  
Title: Associate Director, MIT Center for Biomedical Engineering  
Thesis Supervisor (Co-Advisor): Bruce Tidor  
Title: Professor of Biological Engineering and Computer Science

## ACKNOWLEDGEMENTS

I would like to wholeheartedly thank my advisor Dr. Shuguang Zhang for his support throughout my entire project. His unwavering enthusiasm was a constant inspiration, and through his mentorship I learned about far more than just doing research. I am truly grateful and indebted to him for his confidence in me. My thesis committee, Professor Bruce Tidor, Professor Alexander Rich, and Professor K. Dane Wittrup, offered valuable suggestions and posed difficult questions, for which I am deeply thankful. I feel honored to have had the chance to work with the members of the Zhang lab. I learned so much from my predecessor Brian Cook. I value the informative discussions with him, as well as with Andreas Mershin, Peter Carr, Dave Kong, Luca Turin, Sotiris Koutsopoulos, Iftach Yacoby, Lotta Tegler, Jennifer Brookes, Sergei Pochekailov, and Oz Wassie. They provided valuable advice throughout my project on any number of disparate subjects. I would also like to recognize the work done by my UROPs Junyao Song, Emily Brown, and Deepali Ravel, as well as their eagerness and willingness to jump into whatever project I may throw their way. I truly appreciate the assistance offered by my collaborators. With Horst Pick's help and expertise, we were able to do calcium influx measurements. I am especially indebted to Philipp Baaske, Stefan Duhr, Christoph Wienken, Moran Jerabek- Willemsen, and Sandra Geissler. Without their help and guidance, as well as gracious permission to use their instrument, it wouldn't have been possible to do any ligand-binding measurements. Finally, I would like to thank my family and friends, and especially my parents. Their encouragement, guidance, and constant support were invaluable throughout my graduate school journey. This work was supported in part by the Defense Advanced Research Projects Agency HR0011-09-0012.

# TABLE OF CONTENTS

List of Figures .....	9
List of Tables .....	10
<b>Chapter 1: Introduction and Background</b> .....	11
1.1 G Protein-Coupled Receptors.....	11
1.1.1 Overview.....	11
1.1.2 Receptors Involved in Olfaction .....	12
1.1.3 Olfactory Receptors.....	13
1.1.4 Trace Amine-Associate Receptors.....	16
1.1.5 Vomeronasal Receptors .....	17
1.1.6 Formyl Peptide Receptors.....	18
1.1.7 Technological Motivations in Studying Olfaction .....	19
1.2 Challenges in Membrane Protein Research .....	19
1.2.1 Overview.....	19
1.2.2 Heterologous GPCR Expression .....	21
1.2.2.1 Escherichia Coli Systems.....	21
1.2.2.2 Yeast Systems .....	22
1.2.2.3 Insect Cell Systems .....	22
1.2.2.4 Mammalian Cell Systems .....	23
1.2.2.5 Cell-Free Systems.....	23
1.2.3 Detergent Selection for GPCR Solubilization.....	25
1.2.4 Strategies for Facilitating GPCR Crystallization.....	25
1.2.4.1 Monoclonal Antibody Method .....	26
1.2.4.2 T4-Lysozyme Insertion Method.....	27
1.2.4.3 Targeted Mutation Methods.....	28
1.2.4.4 Lipidic Cubic Phase Method.....	29
1.3 Detergents for Use with Membrane Proteins.....	31
1.3.1 Overview.....	31
1.3.2 Detergent Properties.....	31
1.3.3 Detergent Classes.....	33
1.3.4 The Protein Detergent Complex (PDC).....	34
1.3.5 Current Problems with Traditional Detergents.....	36
1.3.6 Novel Amphiphile Systems for Membrane Protein Solubilization .....	37
1.3.7 Problems with Novel Surfactant Strategies.....	38
1.4 References .....	39
<b>Chapter 2: Large-Scale Expression of hVN1R1 in a Stable Inducible HEK293 Cell Line</b> ....	49
2.1 Introduction .....	49
2.1.1 HEK293 Cells as a GPCR Expression Platform .....	49
2.1.2 The Human Vomeronasal Organ and Receptors .....	50
2.2 Methods.....	52
2.2.1 hVN1R1 Gene Design and Construction.....	52
2.2.2 Construction of Stable Inducible hVN1R1 HEK293G Cell Lines .....	53

## TABLE OF CONTENTS (CONTINUED)

2.2.3	Immunocytochemistry .....	54
2.2.4	Cell Extract Preparation .....	55
2.2.5	Detergent Screening .....	55
2.2.6	Receptor Purification .....	55
2.2.7	Secondary Structure Analysis Using Circular Dichroism Spectroscopy... 56	
2.2.8	Ligand Binding Measurements Using Microscale Thermophoresis .....	57
2.3	Results .....	58
2.3.1	Induction of hVN1R1 Expression in a Stable HEK293 Cell Line .....	58
2.3.2	Immunohistochemical Staining of Induced Cells .....	59
2.3.3	Systematic Detergent Screening for Receptor Solubilization .....	60
2.3.4	Purification of HEK293-Expressed hVN1R1 .....	63
2.3.5	Structural Characterization of Purified hVN1R1 .....	64
2.3.6	Ligand Binding Characterization of Purified hVN1R1 .....	65
2.4	Discussion .....	67
2.4.1	Existence of Functional Human Vomeronasal Receptors .....	68
2.4.2	Technological Importance of Membrane Proteins .....	69
2.5	References .....	71

### **Chapter 3: Investigation of T4 Lysozyme as a Strategy Toward GPCR Crystallization .....**

3.1	Introduction .....	74
3.2	Methods .....	76
3.2.1	hOR17-4T4L and hVN1R1-T4L Gene Design and Construction .....	76
3.2.2	Expression, Purification, and Structural and Functional Analyses of T4L Mutants .....	77
3.2.3	Calcium Influx Imaging .....	77
3.3	Results .....	78
3.3.1	Induction of hOR17-4T4L and hVN1R1-T4L Expression in Stable HEK293 Cell lines .....	78
3.3.2	Immunohistochemical Staining of Induced Cells .....	79
3.3.3	Functional Characterization of hOR17-4T4L in HEK293 Cells .....	82
3.3.4	Systematic Detergent Screening for Receptor Solubilization .....	84
3.3.5	Purification of hOR17-4T4L and hVN1R1-T4L .....	87
3.3.6	Structural Characterization of Purified hOR17-4T4L and hVN1R1-T4L .....	88
3.3.7	Ligand Binding Analysis of Purified hOR17-4T4L and hVN1R1-T4L .....	90
3.4	Discussion .....	91
3.5	References .....	95

### **Chapter 4: A Robust and Rapid Method of Producing Soluble, Stable, and Functional G Protein-Coupled Receptors .....**

4.1	Introduction .....	97
4.2	Methods .....	99

## TABLE OF CONTENTS (CONTINUED)

4.2.1	GPCR Choice .....	99
4.2.2	GPCR DNA Design and Synthesis.....	101
4.2.3	Cell-Free GPCR Production Using Commercial Kits.....	102
4.2.4	HEK293 GPCR Production.....	103
4.2.5	GPCR Detection and Purity Analysis.....	103
4.2.6	Immunoaffinity Purification.....	104
4.2.7	Secondary Structural Analysis Using Circular Dichroism .....	105
4.2.8	Ligand Binding Measurements Using Microscale Thermophoresis .....	105
4.3	Results .....	106
4.3.1	Cell-Free Detergent Screening.....	106
4.3.2	GPCR Expression and Solubility .....	107
4.3.3	GPCR Purification and Yields.....	108
4.3.4	Secondary Structural Analysis Using Circular Dichroism .....	111
4.3.5	Ligand Binding Analysis Using Microscale Thermophoresis.....	112
4.4	Discussion .....	115
4.5	References.....	118

### **Chapter 5: Designer Lipid-Like Peptides: A Class of Detergents for Producing, Solubilizing, and Stabilizing Functional Olfactory Receptors .....**

5.1	Introduction .....	122
5.2	Methods.....	125
5.2.1	Peptide Synthesis.....	125
5.2.2	Cell-Free Olfactory Receptor Expression and Purification.....	125
5.3	Results .....	126
5.3.1	Detergent Screening .....	126
5.3.2	Determination of Receptor Yields.....	129
5.3.3	Olfactory Receptor Purification and Purity Analysis.....	131
5.3.4	Secondary Structural Analysis Using Circular Dichroism .....	132
5.3.5	Ligand Binding Analysis Using Microscale Thermophoresis.....	134
5.4	Discussion .....	136
5.5	References.....	140

### **Chapter 6: Evaluation of the Effect of Surfactant Peptide Physical and Chemical Properties on Olfactory Receptor Solubility .....**

6.1	Introduction .....	143
6.2	Methods.....	145
6.2.1	Amino Acid Selection.....	145
6.3	Results .....	146
6.3.1	Effects of Residue Hydrophobicity, Flexibility, and Volume .....	146
6.3.2	Effects of Peptide Length and Head Group Properties .....	148
6.4	Discussion .....	152
6.5	References.....	154

## TABLE OF CONTENTS (CONTINUED)

<b>Chapter 7: <i>In Vitro</i> Mutational Analysis of the Olfactory Ligand Binding Pocket .....</b>	<b>155</b>
7.1 Introduction .....	155
7.2 Methods.....	157
7.2.1 Rationale for Receptor Choice.....	157
7.2.2 Rationale for Residue Choice .....	157
7.2.3 Gene Design, and Receptor Expression and Analysis.....	160
7.3 Results .....	160
7.3.1 Protein Expression and Purification.....	160
7.3.2 Secondary Structure Analysis.....	161
7.3.3 Ligand Binding Analysis.....	162
7.4 Discussion .....	166
7.4.1 Molecular Interactions Involved in Olfactory Receptor Ligand Binding.....	166
7.4.2 An Efficient Method for Parallel Screening of Multiple GPCRs .....	171
7.5 References.....	172



## LIST OF FIGURES

Figure 1.1. Schematic Representation of a GPCR.....	12
Figure 1.2. Combinatorial Coding of Olfactory Receptors.....	14
Figure 1.3. Olfactory Receptor Signal Transduction.....	15
Figure 1.4. Vomeronasal Type 1 and Type 2 Receptor Structures.....	17
Figure 1.5. Crystal Structures of the $\beta$ 2-Adrenergic Receptor and Adenosine A <sub>2A</sub> Receptor with T4 Lysozyme Fusions .....	28
Figure 1.6. Lipidic Cubic Phases for Protein Crystallization .....	30
Figure 1.7. Detergent Micelle Models.....	35
Figure 2.1. Immunohistochemical Staining of HEK293 Cells Expressing hVN1R1.....	60
Figure 2.2. Detergent Screen for Solubilizing hVN1R1 from HEK293 Cells.....	61
Figure 2.3. Purification of hVN1R1 .....	64
Figure 2.4. Circular Dichroism Spectra of Purified hVN1R1 .....	65
Figure 2.5. Ligand Binding of Purified hVN1R1 Using Microscale Thermophoresis.....	67
Figure 3.1. Expression of hOR17-4T4L and hVN1R1-T4L in HEK293 Cells.....	79
Figure 3.2. Immunohistochemical Staining of Native and T4L Clones.....	81
Figure 3.3. Ca <sup>+2</sup> Influx Profiles for hOR17-4T4L and hOR17-4 Expressed in Stable Inducible HEK293 Cells .....	83
Figure 3.4 280nm UV Traces Through a Size Exclusion Column and a Silver Stain of the T4L Fractions.....	88
Figure 3.5. Circular Dichroism Spectra of Native and T4L GPCRs.....	89
Figure 3.6. Ligand Binding Measurements of Native and T4L GPCRs.....	91
Figure 4.1. Detergent Screen for Cell-Free hOR17-210 Production .....	107
Figure 4.2. Silver Stains of Purified GPCRs .....	110
Figure 4.3. Circular Dichroism Spectra of Cell-Free Expressed GPCRs .....	112
Figure 4.4. Microscale Thermophoresis Measurements of Purified GPCRs .....	114
Figure 5.1. Molecular Models of Peptide Surfactants at Neutral pH.....	124
Figure 5.2. Olfactory Receptor Solubility in Brij-35 and Peptide Surfactants.....	128
Figure 5.3. Detergent Peptides Can Yield Milligram Quantities of Solubilized Olfactory Receptors.....	130
Figure 5.4. Silver Stain of Four Purified Olfactory Receptors Produced in Either Brij-35 or a Peptide.....	132
Figure 5.5. CD Spectra of Brij-35 and Peptide Surfactant-Produced Olfactory Receptors ..	133

Figure 5.6. Binding of Olfactory Receptor Olfr226 to its Ligand 2,4-DNT .....	135
Figure 6.1. Peptide Volume, Flexibility, and Hydrophobicity Do Not Affect GPCR Solubility.....	148
Figure 6.2. Peptide Tail Length and Head Group Composition Negligibly Affect OR Solubility.....	151
Figure 7.1. Sequence and 2D-Topology of mOR103-15 .....	159
Figure 7.2. Silver Stain of the Native and Mutant mOR103-15 Receptors.....	161
Figure 7.3. Circular Dichroism Spectra of Native and Mutant mOR103-15 .....	162
Figure 7.4. Ligand Binding Measurements of the Native mOR103-15 Receptor and the Mutants .....	164
Figure 7.5. Fluorescence Signal as a Function of Heptanal Concentration.....	165
Figure 7.6. Alignment of mOR103-15 with 9 Olfactory Receptors and 4 GPCRs with Known Structures.....	167

## LIST OF TABLES

Table 1.1. Protein Structure Statistics as of April 2011 .....	20
Table 1.2. Detergent Properties.....	33
Table 3.1. Detergent Structures.....	85
Table 3.2. Detergent Solubilization of hVN1R1, hVN1R1-T4L, hOR17-4, and hOR17-4T4L .....	86
Table 4.1. Experimental GPCRs and Their Properties .....	101
Table 4.2. The Solubility and Maximum Yields of 13 GPCRs.....	111
Table 4.3. Ligand Binding Measurements of Selected GPCRs.....	115
Table 4.4. GPCR Yields in Large-Scale Expression Platforms .....	116
Table 5.1. Surfactant Peptide Properties and Experimental Concentrations.....	126
Table 5.2. Ligand Binding Affinities for Peptide-and Brij-35-Produced Olfactory Receptors.....	136
Table 6.1. Amino Acid Properties .....	146
Table 6.2. Size Ratios Between Peptide Surfactant Heads and Tails .....	150
Table 7.1. Measured EC <sub>50</sub> Values for mOR103-15 Mutants.....	165

# CHAPTER 1

## INTRODUCTION AND BACKGROUND

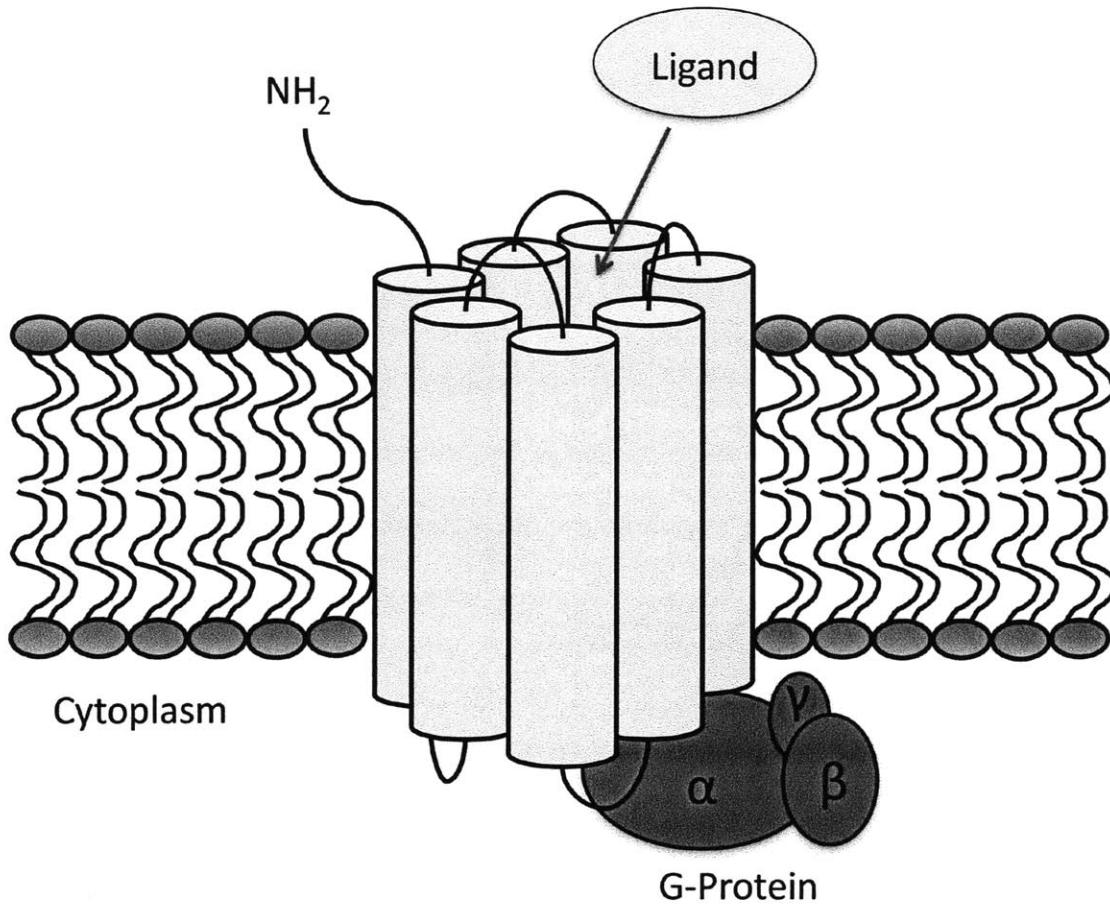
### 1.1 G Protein-Coupled Receptors

#### 1.1.1 Overview

G protein-coupled receptors (GPCRs) are the largest class of membrane proteins. They are involved in many physiological processes, including sight, smell, inflammation, and regulation of the nervous system. Because of the broad range and significance of their functions, GPCRs are involved in many diseases and are the target of over 50% of pharmaceutical drugs [1, 2]. In spite of their importance, relatively little is known about their structure and function.

The GPCR family of proteins is characterized by a common three-dimensional structure and similar signaling pathways. GPCRs have seven transmembrane  $\alpha$ -helical domains joined by flexible loops, an extracellular N-terminus, and an intracellular C-terminus. The transmembrane domains are arranged in a barrel-like configuration (**Figure 1.1**). Ligands bind primarily in the pocket formed from the helix barrel, but they can also bind to the N-terminus or the extracellular loops [3]. Upon ligand binding, the receptor changes conformation and activates a G-protein. The activated G-protein initiates a signaling cascade, where the specific pathway depends on the specific G-protein.  $\beta$ -arrestins may also act as signal transducers instead of a G-protein [4-7]. The structural and

signaling similarities among GPCRs are surprising, considering that sequence identity among individual family members can be less than 20% [8, 9].



**Figure 1.1. Schematic Representation of a GPCR.** GPCRs have seven transmembrane helices arranged in a barrel-like configuration, represented by cylinders. The ligand binds in the pocket formed by the helices on the extracellular side. Upon binding, the GPCR undergoes a conformational change and initiates a signaling cascade through a G-protein on the intracellular surface.

### 1.1.2 Receptors Involved in Olfaction

There are four main classes of GPCRs that are involved in olfaction: olfactory receptors (ORs) [10], trace amine-associate receptors (TAARs) [11], vomeronasal receptors (VNRs) [12, 13], and formyl peptide receptors (FPRs) [14, 15]. All are GPCRs.

Together, they comprise the largest subclass of GPCRs, and account for ~3% of the mammalian genome.

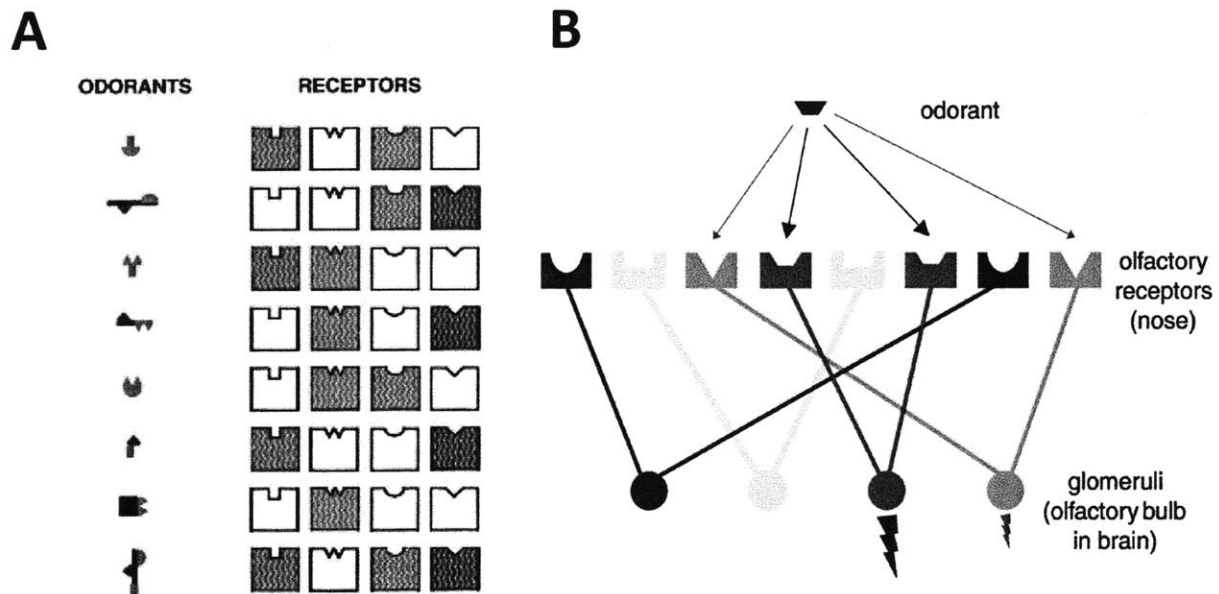
### 1.1.3 Olfactory Receptors

Olfactory receptors are the largest of the four classes of receptors, and are in fact the largest gene family in vertebrates [16]. Humans have up to ~1000 receptors, of which 391 are functional and the remaining are pseudogenes. In contrast, mice and dogs are predicted to have ~1000 functional receptors, and only 200-300 pseudogenes [8, 17-22].

Olfactory receptors are divided into two classes based on sequence homology. Class I receptors are genetically similar to amphibian ORs, and typically bind water-soluble ligands. Class II receptors constitute the majority of mammalian ORs, and bind volatile hydrophobic molecules. All ORs are ~30-35 kDa, while their ligands are all less than 300 Da [21, 23, 24].

Studies indicate that mammalian noses can detect a seemingly infinite number of odorants in minute concentrations as small as parts per trillion [25]. This amazing sensitivity is made possible through the organization of the olfactory system. Olfactory receptors are expressed on the cilia of olfactory neurons in the main olfactory epithelium (MOE) [26]. The MOE is divided into 4 sections. Each neuron expresses only one OR allele, and all neurons expressing the same allele are spread throughout the same MOE section [27-29]. Axons of neurons expressing the same OR meet at foci called glomeruli [29, 30]. Because each receptor recognizes specific chemical groups on odorants, one odorant can activate many receptors, and one receptor can bind several odorants [31, 32] (**Figure 1.2**). Signals from activated neurons converge onto the appropriate glomeruli, creating a spatial

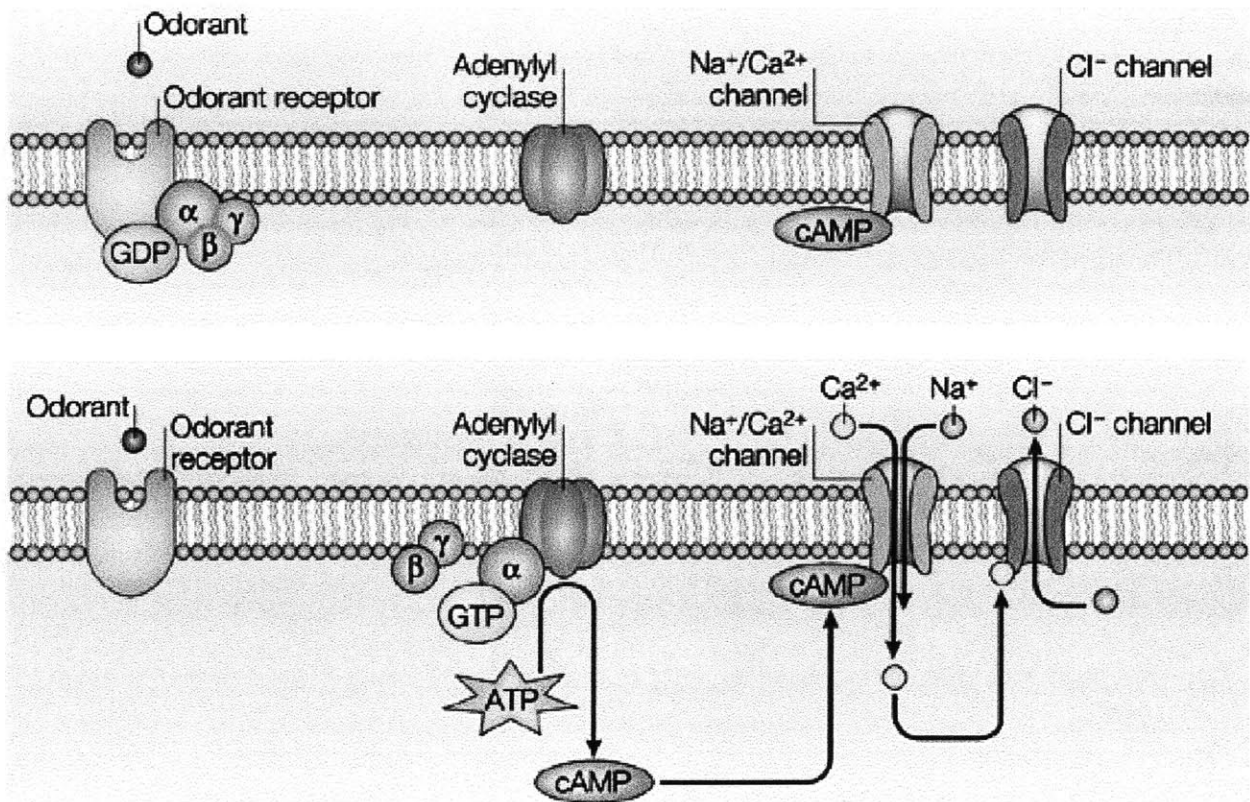
map that is believed to encode a scent's identity [8, 31]. This combinatorial method of signal detection allows the number of detectable odors to greatly exceed the number of ORs in the genome. Because only portions of an odorant are identified by each receptor, the olfactory system can detect new molecules that have not been previously encountered in the course of evolution.



**Figure 1.2. Combinatorial Coding of Olfactory Receptors.** A) Each receptor can recognize specific chemical groups on an odorant. Thus, one odorant can activate multiple receptors, and each receptor can bind to several odorants. Reprinted from [31] with permission from Elsevier. B) The axons of neurons that express the same receptor converge into foci called glomeruli. The spatial map of activated glomeruli encodes a unique odor. From [132], by permission of Oxford University Press.

Olfactory receptors transduce signals through the G-protein  $G_{olf}$ . Upon ligand binding,  $G_{olf}$  is activated and released. The released G-protein subunit stimulates Adenylate Cyclase III to increase cAMP production, which further triggers ion-channel mediated  $Ca^{+2}$  release and firing of an action potential [26, 33-36] (**Figure 1.3**). Olfactory receptors are

capable of interacting with other G-proteins, thus enabling receptor activity to be measured *in vitro* using heterologous cell expression [8, 32, 37].



**Figure 1.3. Olfactory Receptor Signal Transduction.** Upon odorant binding, the G-protein is activated and released. Cyclic-AMP production is increased, which induces an influx of  $\text{Ca}^{2+}$  and firing of an action potential. Reprinted from [36] with permission from Macmillan Publishers Ltd, ©2004.

Although olfactory system morphology and signal transduction have been studied, very little is known about the molecular basis of olfaction. Only a handful of ORs have been deorphaned [8]. Due to the difficulty of working with this class of receptors, only three mutational studies investigating potential ligand-binding sites have been performed [38-40]. All other studies have been computational. Several studies have developed OR

structures based on homology modeling to bacteriorhodopsin or vertebrate rhodopsin [41-47]. Others are based on comparisons between OR orthologs, homologs, and conserved or variable sequence motifs [48-50]. Although the mutational and computation studies yield insights into OR function, it is not possible to predict with certainty which ligands will bind a specific receptor. It is also not possible to predict the “scent” of a molecule. In order to understand this molecular basis of smell, knowledge of the structure of an OR is critical, as well as a deeper understanding of the combinatorial code.

#### **1.1.4 Trace Amine-Associate Receptors**

Trace amine associate receptors (TAARs) are a small class of GPCRs that detects small volatile amines implicated in social signaling [11, 51]. Most species have less than 30 genes, with humans having 7 functional receptors and mice having 15. In mice, TAARs are likely involved in detecting social cues [15].

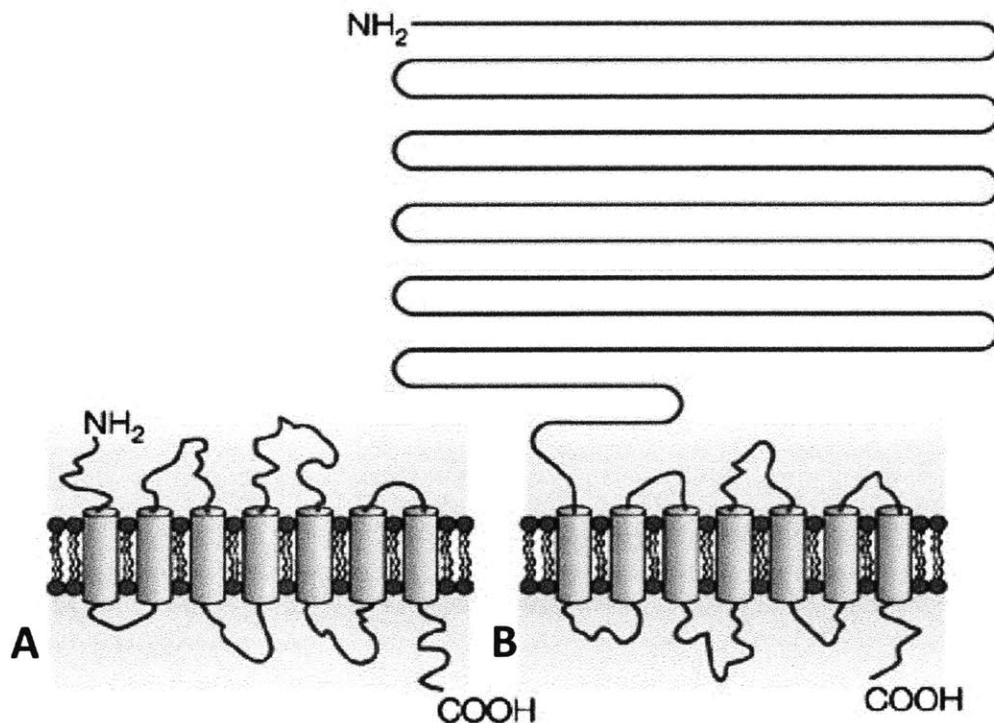
Trace-amine associate receptors are expressed in the main olfactory epithelium alongside ORs. Like ORs, each TAAR gene appears to be expressed within a specific epithelial zone. Evidence also indicates that each neuron expresses only one TAAR gene, and that TAARs are not co-expressed with ORs [11]. Labeling experiments demonstrated that the G-protein  $G_{olf}$  is co-expressed with TAARs. Ligand activation of TAARs increases cAMP levels, suggesting that TAARs utilize the same signaling pathway as ORs [11, 52, 53].

Although both ORs and TAARs detect small volatile compounds, exhibit similar expression patterns, and activate the same signaling messengers, they are unique GPCR families. Sequence homology between the two GPCR classes does not exceed 16% [8, 11]. Also, TAARs do not have the same sequence motifs seen in the OR family.



### 1.1.5 Vomeronasal Receptors

Vomeronasal receptors (VNRs) detect pheromones and molecules involved in social signaling. They are divided into two classes based on their structure and ligands. Type 1 VNRs (VN1Rs) have short N- and C-termini, and bind small volatile ligands. Type 2 VNRs (VN2Rs) have an extended N-terminus, and bind peptides [12, 13, 36, 54, 55] (**Figure 1.4**). Like ORs, only one unique VNR allele is expressed in a neuron, and neurons expressing the same VNR converge to the same glomeruli [36].



**Figure 1.4. Vomeronasal Type 1 and Type 2 Receptor Structures.** A) Type 1 receptors have structures similar to ORs, while B) Type 2 receptors have an extended N-terminus. Reprinted from [36] with permission from Macmillan Publishers Ltd, ©2004.

Human and rodent VNRs likely serve different roles. Mice have 100-200 potentially functional VN1Rs, and about 100 VN2Rs. All are expressed in the vomeronasal organ:

VN1Rs are expressed in the apical zone with the  $G_{\alpha i}$  G-protein subtype, and VN2Rs are expressed in the basal zone with the  $G_{\alpha o}$  subtype. Both families probably signal through phospholipase C (PLC) and inositol 1,4,5-triphosphate (IP3) via TRPC2 [56-60]. In contrast, humans have only 5 potentially functional VN1Rs, and all VN2Rs are pseudogenes [61]. When expressed in HeLa/Olf cells, all 5 human VN1Rs responded to volatile molecules via  $G_{\alpha olf}$  and cAMP [58]. The existence of a functional human vomeronasal organ is questionable [62-70], but mRNA of one human VNR was detected in the main olfactory epithelium, brain, lung, and kidney tissues [71]. This raises the possibility that human VN1Rs may function more like ORs in the MOE, or may have non-olfactory functions.

### **1.1.6 Formyl Peptide Receptors**

Formyl peptide receptors (FPRs) were originally discovered in the immune system. Because these receptors recognized formylated peptides from bacteria or mitochondria, it was widely believed that they functioned in recruiting phagocytes to sites of infection or tissue damage [72].

Recent evidence suggests that this family of receptors may have an olfactory role [15]. Five of seven mouse FPRs were selectively expressed in the mouse VNO [14, 15]. Like VNRs, FPRs either co-expressed solely with  $G_{\alpha i}$  or  $G_{\alpha o}$ , and can signal through PLC and IP3 [14, 15, 72]. As with the ORs, TAARs, and VNRs, only one gene was expressed by a neuron, and no other chemosensory neurons were expressed. Although the location and expression patterns of mouse FPRs indicate that they are chemosensory receptors, more work needs to be done. However, it has been hypothesized that FPRs may aid in the detection of predators or sick conspecifics, or may aid in the discrimination between

members of the same species [14]. An olfactory role for the three human FPRs has not yet been indicated.

### **1.1.7 Technological Motivations for Studying Olfaction**

The mammalian olfactory system can instantly detect odorants at minute concentrations and distinguish between them. Studies also indicate that smell can be used to fingerprint individuals and detect cancer at early stages [73-76]. Although much is known about morphology and signal transduction, the molecular basis of olfaction is not understood. Ligands for the majority of the receptors are not yet known, and the role of each of the four GPCR classes in olfaction has not been elucidated. If the mechanisms of receptor-ligand interactions were understood, as well as the interactions between each of the four systems, it might be possible to engineer highly sensitive and versatile detectors for anything ranging from biological threats to terrorists to early stages of cancer.

## **1.2 Challenges in Membrane Protein Research**

### **1.2.1 Overview**

Despite their large numbers and biological importance, membrane protein research lags far behind that of soluble proteins. Knowledge about GPCRs in particular is sparse. This is due primarily to the difficulty of crystallizing membrane proteins, a precondition to determining their structure. As of April 2011, over 72,000 protein structures have been determined. Only 281 are unique membrane proteins (less than 0.4%!). Of these, only 6 are GPCRs (**Table 1.1**) [77, 78]. The first GPCR structure, that of rhodopsin, was solved in 2000. The second, the  $\beta$ 2-adrenergic receptor, was solved seven years later. It was

followed by two more in 2008: the  $\beta$ 1-adrenergic receptor and the A<sub>2A</sub>-Adenosine receptor. The last two, the CXCR4 chemokine and Dopamine D3 receptors, were crystallized in 2010.

GPCRs are difficult to crystallize for four main reasons. First, abundant quantities of protein are needed to set up crystallization trials, but most are endogenously expressed at low levels. Only rhodopsin, the first crystallized GPCR, is easily obtained in sufficient quantities from native tissues. Second, suitable methods must be found to extract, solubilize, and purify GPCRs. Third, GPCRs must be functionally stabilized for long periods of time, as protein crystals can take weeks or even months to grow. Because GPCRs have a hydrophobic transmembrane region bounded by hydrophilic ends, they aggregate and precipitate out of aqueous solutions when removed from their native membrane environment. Detergents that mimic the lipid bilayer chemistry must therefore be used to maintain GPCRs in a stable, non-aggregated form. Fourth, the flexible nature of GPCRs, and the materials used to stabilize them in aqueous environments, can inhibit crystal lattice formation. Often, the GPCR to be crystallized must be cleverly engineered to increase potential protein-protein contact sites. Each bottleneck must be sequentially overcome. However, no universal method exists: optimal expression, purification, solubilization, and engineering strategies must be empirically determined for each protein of interest.

**Table 1.1: Protein Structure Statistics as of April 2011**

<b>Protein Type</b>	<b>Total Number of Structures Determined</b>	<b>Year of First Structure</b>
All proteins	72,717	1958
Membrane proteins	281	1985
GPCRs	6	2000
Olfactory-related GPCRs	0	n/a

## **1.2.2 Heterologous GPCR Expression**

Systems capable of expressing GPCRs must be developed before any structural or functional studies can be performed. This is a non-trivial task, as most GPCRs are expressed at low levels in their native tissues, and are not easily amenable to expression in heterologous systems. However, overexpression in heterologous systems can be accomplished, and four main platforms have been developed towards this goal. It should be noted that an expression methodology that works well with one protein may not yield sufficient quantities of another.

### **1.2.2.1 Escherichia Coli Systems**

*Escherichia coli* are an attractive receptor-expression platform because of their low cost and fast doubling time (20-40 minutes). Plasmid DNA encoding the desired gene can be synthesized quickly, and *E. coli* clones expressing the receptor of interest can easily be generated. These characteristics allow expression of many GPCR variants to be tested and optimized within a short span of time. However, *E. coli* expression systems suffer a number of disadvantages. Often, the protein of interest must be fused to a bacterial leader sequence to facilitate sufficient expression [79-82]. Because they are prokaryotic cells, their intracellular environment is different from that found in eukaryotic cells. As a result, overexpressed GPCRs can form inclusion bodies. While it is possible to isolate the inclusion bodies, devising procedures to refold the proteins in an active state is a nontrivial task. The reductive environment of the bacterial cytoplasm poses a particular challenge in the functional expression and folding of GPCRs, as disulfide bridges necessary for proper folding may not be able to form [81]. *E. coli* also lack the machinery capable of performing

post translational modifications like glycosylation. While this may yield a homogenous protein sample, the synthesized receptors may be misfolded or non-functional. For example, when the N-terminal glycosylation site of rhodopsin is blocked, the protein loses its function [83, 84]. In ORs, loss of glycosylation may result in improper folding or loss of function [85].

#### **1.2.2.2 Yeast Systems**

Yeast have been used as an expression system because they have short generation times (2 hours), and are also inexpensive to use. Unlike *E. coli*, yeast are eukaryotes that are capable of performing post-translational modifications like glycosylation [81, 82, 86, 87]. However, the composition of the N-glycans in yeast differs from that in mammalian cells. Misfolded or non-functional proteins can result when a specific type and number of carbohydrate chains must be added. Moreover, although yeast have the capability of performing glycosylation, they often do not perform this modification on heterologously expressed GPCRs. Additional difficulties that are often encountered include sequestering of proteins in internal compartments, and proteolytic degradation of the synthesized receptors. Lastly, the strength of the yeast cell wall can make purification of the expressed receptors difficult [81, 82, 86].

#### **1.2.2.3 Insect Cell Systems**

Insect cells are eukaryotes capable of expressing high levels of functional GPCRs [81, 82]. Indeed all of the crystallized GPCRs except rhodopsin were expressed in Sf<sub>9</sub> insect cells [88-93]. As eukaryotes, they are capable of glycosylation. However, GPCRs can be poorly

glycosylated [94], or may not contain the appropriate sugars [95]. Additionally, incomplete glycosylation results in heterogeneous protein samples, which are difficult to crystallize [81, 82, 86]. Because some GPCRs either cannot be expressed in insect cells, or are non-functional, an insect platform must be evaluated for each individual protein of interest. Their long generation time (24 hours) makes this a time consuming process, and evaluation of multiple GPCRs or variants of a single GPCR can be impractical.

#### **1.2.2.4 Mammalian Cell Systems**

Mammalian cells represent the most natural expression environment for most GPCRs. They are closest to the tissues in which GPCRs are natively expressed, and are capable of all necessary post-translational modifications. However, overexpression of GPCRs can lead to incomplete or differential glycosylation, resulting in a heterogeneous protein sample. Mammalian cells have the slowest doubling times (24-48 hours), are prone to infection, and are expensive to handle. Overexpression of GPCRs in mammalian cells is possible, but is usually toxic to the cells. Inducible stable cell lines can overcome this problem [37, 96-99]. However, the process of creating a stable clone lasts several months, and many potential clones must be screened to find a line that maximizes expression, while minimizing toxicity or sources of heterogeneity (e.g. differential glycosylation or degradation).

#### **1.2.2.5 Cell-Free Systems**

Cell-free protein expression offers a potential solution to the problems of low yields, cell-toxicity, misfolding, protein heterogeneity, and receptor aggregation typical in

cell-based platforms. In these systems, plasmid DNA encoding the protein of interest is added to a mixture of amino acids and transcriptional and translational machinery directly in a tube. Because the system is open, accessory reagents like lipids or detergents can be easily added. Although cell-free expression is a mature technology for soluble proteins, it has not been well established for membrane proteins. Several strategies have nevertheless been developed for cell-free membrane protein synthesis.

The easiest method of expressing soluble membrane proteins is to add a detergent directly to the synthesis reaction. However, the detergent must be carefully chosen, as it can interfere with transcription or translation. Although numerous detergents must be screened before an optimal one is found for a particular protein, mild non-ionic detergents are usually preferred [100-103]. Because detergent micelles do not completely mimic the mechanical and chemical properties of the lipid bilayer, solubilized proteins may not be in their native conformation.

Because liposomes and membrane vesicles better mimic the lipid bilayer, several studies have used them instead of detergents during cell-free expression. Two general strategies have been employed. In the first, a membrane protein is expressed without a surfactant. The resulting aggregate is purified and reconstituted in a liposome [104]. The reconstitution step can also be performed with a mild detergent [104-107]. In the second, liposomes or vesicles are added directly to the reaction. Membrane proteins are then incorporated into them as they are expressed [108, 109]. Although liposomes and vesicles can facilitate membrane protein expression, they are not a viable option for structural studies due to the heterogeneity of the resulting samples. Liposomes have been reported to range from 30-200 nm in diameter, and have varying numbers of incorporated proteins or



no protein at all. Additionally, reconstitution conditions must be empirically determined for each protein [104].

Nanolipoprotein particles (NLPs) have been developed to overcome some of the shortcomings of liposomes and vesicles. NLPs are discoidal particles with a lipid bilayer core surrounded by scaffold proteins [110]. Because the scaffold protein limits the size of the disc, NLPs are relatively homogeneous in size (<5% deviation in mass). Also, NLPs have been shown to enhance protein yields, achieving up to 0.9 mg/ml. They can be used in cell-free extracts from multiple organisms, and are commercially available [111]. However, the usefulness of cell-free technology with GPCRs has yet to be reproducibly demonstrated. Indeed, current cell-free methods typically either do not produce sufficient GPCRs, or produce GPCRs that are inactive or misfolded [82, 104].

### **1.2.3 Detergent Selection for GPCR Solubilization**

Because GPCRs have a hydrophobic core bounded by two hydrophilic surfaces, detergents must be used to stabilize them once they are removed from their native membrane environment. Selection of an appropriate detergent is critical, and is a difficult and laborious task. These difficulties and the current state of the field are reviewed in section 1.3.

### **1.2.4 Strategies for Facilitating GPCR Crystallization**

Detergents used to solubilize membrane proteins form a protective belt, shielding the hydrophobic transmembrane regions from the aqueous solution. This detergent belt severely reduces the number of potential protein-protein contact sites, thereby reducing

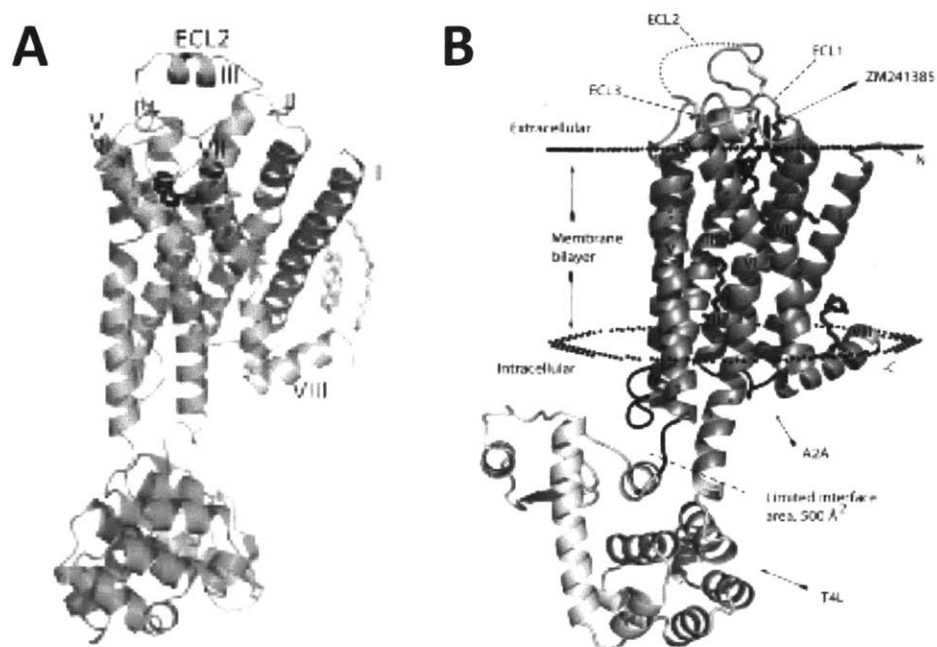
the likelihood of crystal lattice formation. GPCRs present additional difficulties in crystallization. A crystal is a well-ordered, repeating unit structure. The flexible nature of GPCRs can introduce heterogeneity that either reduces the resolution of any crystals or inhibits their formation. In order to facilitate crystal growth, four general strategies for increasing the hydrophilic surface area or increasing GPCR stability have been developed: 1) Monoclonal antibodies, 2) T4-Lysozyme insertions, 3) Stabilizing mutations or deletions, and 4) Lipidic cubic phases.

#### **1.2.4.1 Monoclonal Antibody Method**

Monoclonal antibodies have been used to increase the polar surface area of membrane proteins [112-115]. This strategy has been used successfully with several membrane proteins, including one of the 6 crystallized GPCRs: the  $\beta$ 2-adrenergic receptor [91]. The advantages of using antibody fragments for crystallization are that they are highly soluble, and can be engineered to bind with extremely high affinity to a protein of interest. A relatively homogenous sample can be produced in *E. coli*, and it does not destabilize the resulting antibody-membrane protein complex [116]. Although antibodies can facilitate crystal-lattice formation, they suffer several drawbacks. The creation of a high-affinity antibody requires considerable expertise and lasts at least several months. Moreover, because they only bind to specific epitopes, non-stabilized portions of a protein may be too flexible to be visualized [91].

#### 1.2.4.2 T4-Lysozyme Insertion Method

Insertion of T4-Lysozyme in the 3<sup>rd</sup> intracellular loop of GPCRs is an alternative strategy that has been used to increase the polar surface area and potential number of lattice-forming contacts [88-90, 93]. T4-Lysozyme is a bacteriophage enzyme that is soluble and easy to crystallize. The full protein has 164 amino acids and is about 18.6 kDa, roughly half the size of olfactory-related GPCRs. Because it is highly soluble, it folds into a distinct domain away from the transmembrane region (**Figure 1.5**). Four of the 6 crystallized GPCRs have used this strategy, which suggests that it could be used for additional GPCRs. The ease of generating plasmid DNA templates and predicting transmembrane regions makes this strategy more appealing than the use of antibodies. Crystal structures of ligand-bound GPCRs indicate that the T4-Lysozyme insertion does not significantly affect the transmembrane or binding pocket structure. However, because the intracellular loops interact with G-proteins, the insertion likely disrupts signal transduction [93].



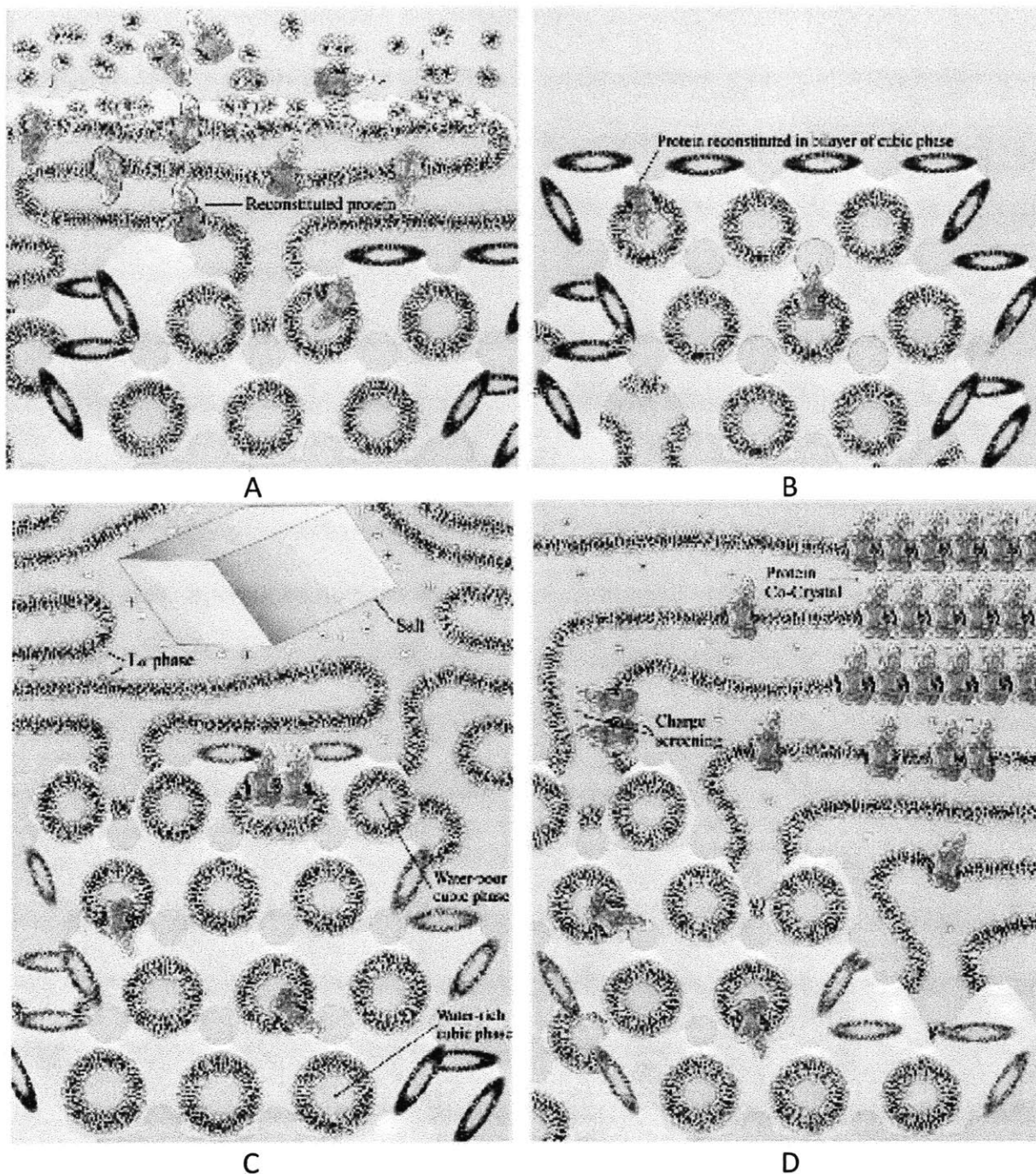
**Figure 1.5. Crystal Structures of the  $\beta$ 2-Adrenergic Receptor and Adenosine  $A_{2A}$  Receptor with T4 Lysozyme Fusions.** A) The  $\beta$ 2-adrenergic receptor is grey, and the T4 lysozyme fusion is green. The transmembrane helices are numbered I-VII, and the second extracellular loop is labeled ECL2. From [88]. Reprinted with permission from AAAS. B) The adenosine  $A_{2A}$  receptor is gold, and the T4 lysozyme fusion is blue. The loops, lipid bilayer boundaries, and ligand are labeled. From [90]. Reprinted with permission from AAAS. The T4 lysozyme domain folds away from the native protein, thereby minimally disturbing the structure of the transmembrane regions.

### 1.2.4.3 Targeted Mutation Methods

Targeted mutations have been introduced in order to sufficiently stabilize GPCRs for crystallization. The intent of such mutations is to stabilize the receptor's structure, or to facilitate formation of a well-ordered crystal. The most common mutation is a deletion of large and flexible loop or terminal regions [88-92]. Point mutations that confer thermal stability [89, 92, 93] or that remove palmitoylation sites [92] have also been performed. Experiments indicate that these mutations do not significantly alter receptor structure or binding.

#### 1.2.4.4 Lipidic Cubic Phase Method

Lipidic cubic phases (LCPs) are hypothesized to facilitate crystallization by mimicking the cell-membrane better than detergent micelles [116]. LCPs are bicontinuous nanoscale structures formed between lipids and an aqueous solution. Because of their immiscible nature, the lipids pack together to form a continuous bilayer that separates 2 distinct but interpenetrating aqueous phases (Figure 1.6). Diffusion coefficients in the lipid bilayer are similar to those in cell membranes, and diffusion in the aqueous compartments is comparable to that in bulk water [116, 117]. Membrane proteins incorporated into the LCP network can freely diffuse in the lipid layer, while soluble additives can diffuse in the aqueous compartments. This structure allows lattice-forming contacts to be made with both hydrophilic and hydrophobic protein regions. Moreover, the planar nature of the lipid bilayer allows it to exert mechanical stresses comparable to those in the cell membrane, thereby maintaining protein structure more efficiently than detergent micelles [88, 116, 118]. Four GPCRs have been crystallized using LCPs [88-90, 93], indicating their efficacy. However, LCPs still suffer similar drawbacks as traditional crystallization methods. The number of parameters that must be tested is large, and include protein concentration, lipid content, detergent type, salts, temperature, and pH. Optimal crystallization conditions must be empirically determined for each tested protein. Also, a protein must be solubilized in detergent before it can be incorporated into an LCP matrix. But, high detergent concentrations can cause the cubic phase to become lamellar [116], which is not conducive to crystal formation.



**Figure 1.6. Lipidic Cubic Phases for Protein Crystallization.** A) Detergent-solubilized protein (top) is added to a lipidic cubic phase (bottom). B) The protein is incorporated into the LCP. C) Salts or precipitants are added to facilitate crystal nucleation and growth. D) As the water content decreases upon the addition of salt, the bilayer curvature increases. The proteins are free to diffuse in the bilayer, and can form lattice-forming contacts. Reprinted from [117], with permission from Elsevier.

## **1.3 Detergents for Use with Membrane Proteins**

### **1.3.1 Overview**

Membrane proteins have evolved to stably exist in biological membranes, which are fluid, amphipathic, and anisotropic structures. To maintain this stable integration, membrane proteins are comprised of a hydrophobic core embedded within the lipid bilayer, and hydrophilic ends that remain in the polar portion of the bilayer or extend into the extracellular or cytosolic space. As a result, they are not soluble in either completely polar or completely hydrophobic solvents. Amphipathic molecules that mimic the properties of biological membranes must be used to handle this class of proteins outside their native environment.

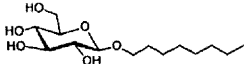
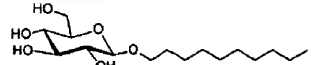

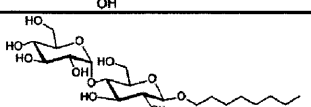
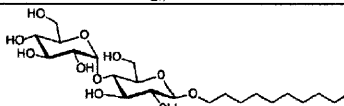
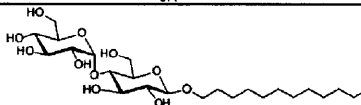
### **1.3.2 Detergent Properties**

In aqueous solution, detergent molecules exist as monomers that can transition to two basic phases [119]. One phase is crystallization, though this does not typically occur with detergents used for membrane proteins. The second phase is micellarization. At a concentration above a critical value, detergent monomers self-aggregate into spherical micelles where the non-polar tails are sequestered in the center and the polar head groups interface with the aqueous solvent. This concentration is denoted the critical micelle concentration (CMC). Detergent solutions at or above their CMC are capable of solubilizing membrane proteins: the detergent micelles are able to encapsulate the hydrophobic protein core and substitute for the plasma membrane. At significantly high concentrations, detergents are able to disrupt cellular membranes and extract membrane proteins.

The size and shape of detergent micelles depends on the properties and chemistry of the monomer head and tail groups, as well as the temperature, pH, ionic strength, and general nature of the solvent. The head group typically affects the CMC and average number of monomers in a micelle. As a general rule, in a series of detergents with the same head group, the CMC decreases by a factor of 10 for every 2 methylenes added to an alkyl tail. Shorter chain detergents tend to have higher CMC values, while longer chain detergents can be less soluble, thereby having lower CMC values. Detergent families with larger or more soluble headgroups are more soluble (**Table 1.2**) [120, 121]. In general, detergents with shorter chain lengths, smaller head groups, and charged head groups are harsh and can lead to protein denaturation or instability. However, they can be better at extracting proteins from the lipid bilayer or crystallizing membrane proteins. Detergents with longer alkyl tails, larger head groups, or head groups with a neutral charge are milder, and are better able to stabilize membrane proteins [120]. Detergent monomers in a micelle are in equilibrium with the free monomers in the solvent. This property allows one detergent to be exchanged for another. Because detergents with longer alkyl chains typically have a slower rate of exchange, only detergents with a lower CMC or smaller micelle size can be replaced.



**Table 1.2: Detergent Properties**

Detergent	Formula	Structure	CMC* (mM)	Aggregation Number*	Solubility*
N-octyl- $\beta$ -D-glucopyranoside	C <sub>14</sub> H <sub>28</sub> O <sub>6</sub>		18-20	27-100	>20% (0-5°C)
N-decyl- $\beta$ -D-glucopyranoside	C <sub>16</sub> H <sub>32</sub> O <sub>6</sub>		2.2	NA	>0.1% (20°C)
N-dodecyl- $\beta$ -D-glucopyranoside	C <sub>18</sub> H <sub>36</sub> O <sub>6</sub>		0.19	NA	>0.008% (20°C)
N-octyl- $\beta$ -D-maltopyranoside	C <sub>20</sub> H <sub>38</sub> O <sub>11</sub>		19.5**	6	>20% (0-5°C)
N-decyl- $\beta$ -D-maltopyranoside	C <sub>22</sub> H <sub>42</sub> O <sub>11</sub>		1.8	45	>20% (0-5°C)
N-dodecyl- $\beta$ -D-maltopyranoside	C <sub>24</sub> H <sub>46</sub> O <sub>11</sub>		0.17	78-149	>20% (0-5°C)

\* All CMC, aggregation number, and solubility values were determined in water unless otherwise specified.

\*\* Water with 100mM NaCl, 20mM HEPES, pH 7.5

Data was obtained from Anatrace's online product information [121]

### 1.3.3 Detergent Classes

Chemical detergents are amphipathic molecules that have traditionally been used to solubilize, stabilize, and crystallize membrane proteins. Three general categories of detergents have been used: ionic, nonionic, and zwitterionic. Ionic detergents are typically harsh, and can often completely denature proteins or dissociate complex proteins into their constitutive components. The ionic strength and nature of the counterion in solution also affect ionic detergent properties; charge-shielding effects can diminish repulsion between the charged head groups and lead to a lower CMC and larger average micelle size. Temperature and pH are also important factors. Particularly, such surfactants can only solubilize proteins at a pH above (for anionic) or below (for cationic) the pKa of the ionizable head group. Because of these variables, it is necessary to demonstrate that native protein function is maintained under experimental conditions when using ionic detergents.

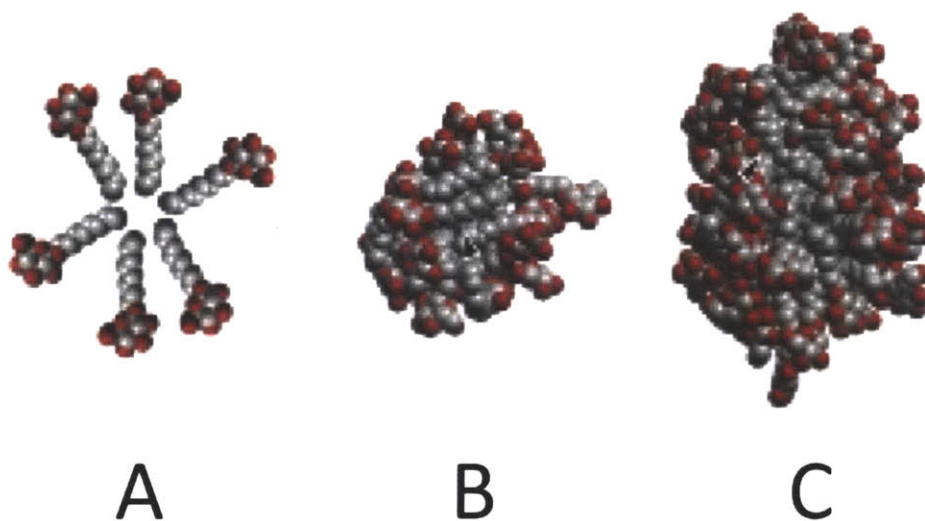
Zwitterionic detergents have an overall electrically neutral charge, and do not suffer from the same drawbacks as ionic detergents. They are better able to protect the native state of the membrane protein, and are used more often than ionic detergents to solubilize and stabilize them. Nonionic detergents, particularly those with polyoxyethylene or sugar head groups, are perhaps the gentlest class of detergents. Because they are strong enough to disrupt lipid-lipid or lipid-protein interactions, but not protein-protein interactions, they are used most often to solubilize and stabilize membrane proteins, and to study their functions. This class of detergents has also been most successfully used to crystallize membrane proteins. Indeed, nearly 40% of membrane protein structures have been determined with the nonionic detergents LDAO (N,N-dimethyldodecylamine-N-oxide), DDM (N-dodecyl- $\beta$ -D-maltopyranoside), and OG (N-octyl- $\beta$ -D-glucopyranoside) [122, 123].

#### **1.3.4 The Protein Detergent Complex (PDC)**

Solubilized membrane proteins must exist in a complex bound with detergent. This complex (PDC) has properties different from either the protein or detergent micelle alone, which must be considered during crystallization experiments.

Detergents bind to membrane proteins in a torus-like shape, shielding the hydrophobic core from the aqueous solution [124, 125]. The detergent, protein, and solvent properties affect the average size and shape of this detergent “belt”. However, micelle packing is not perfectly spherical, as is usually assumed and illustrated. Molecular dynamics simulations indicate that the detergent tails pack in a more compact and disordered configuration (**Figure 1.7**). As a result, the micelle surface is rough and heterogeneous, and portions of the hydrophobic tails can have significant contact with the

surrounding solution [119]. The degree of contact between the tails and solution can influence protein aggregation; if a detergent is not able to sufficiently shield the transmembrane portions of a protein, these regions may self-aggregate and precipitate out of solution. Typically, this aggregation is irreversible [120].



**Figure 1.7. Detergent Micelle Models.** A) The classical spherical micelle model. B) A 20-monomer micelle model after a 40ns molecular dynamics simulation C) A 50-monomer micelle model after a 40ns molecular dynamics simulation. The polar oxygens are colored red; the hydrophobic carbon tails are grey. Contrary to simple classical illustrations, portions of the hydrophobic micelle core are exposed to the solvent. This research was originally published in [119]. ©The American Society for Biochemistry and Molecular Biology.

The molecular weight of the PDC can be twice that of the native protein, or more. The number of detergent molecules bound to the protein depends not only on the detergent aggregation number and protein size, but also on the detergent concentration. For example, at 103mM, about 40 molecules of OG were bound to prostaglandin H synthase-I (PGHS-I). At 23.9mM, the number of bound OG molecules was almost double.

Because the protein-bound detergent will form most of the non-solvent mass in a protein crystal, its properties can often dominate over those of the protein. Thus, the detergent properties need to be carefully considered, in tandem with the protein properties, when setting up crystallization screens [120, 124].

### **1.3.5 Current Problems with Traditional Detergents**

Many problems exist with current detergent-membrane protein protocols. First, a bewilderingly wide range of detergents is available. Yet, no single detergent or class of detergents is universally useful for any given experiment or family of proteins. For example, although LDAO is one of the most successfully used detergents in crystal screens, it is estimated that only 20% of membrane proteins are soluble and stable in it [126]. The optimal detergent for a given protein must therefore be determined empirically. The optimal concentration must also be tested, as low concentrations may not sufficiently solubilize the protein and high concentrations could denature it. Second, the detergent that works best for one application may not work at all in another. For example, detergents with smaller micelles can increase protein solubility and the number of potential protein-protein lattice-forming contacts, but also lead to protein denaturation or instability. Thus, while perhaps aiding in protein solubilization and crystal formation, a detergent could adversely affect the protein's structure or function. An additional confounding factor is that the order of surfactant use can be a critical concern. For example, PGHS crystals that were grown in OG were of better quality if the protein was first solubilized in C10M instead of Tween-20 before exchanging surfactants [124]. Third, surfactants can exhibit complex phase behavior under specific environmental conditions (i.e. temperature or

concentration). This behavior can affect the experimental outcome or state of the solubilized protein.

### **1.3.6 Novel Amphiphile Systems for Membrane Protein Solubilization**

Because detergents play a crucial role in membrane protein experiments, and because it is difficult to find the “perfect detergent”, research efforts have focused on developing new amphiphilic systems. The primary goal is to develop a more universal surfactant that can both maintain protein stability and facilitate crystallization.

Several novel surfactant systems have been developed. The first novel non-detergent surfactant was designed in 1993 [127]. It was comprised of a 24-amino acid peptide that folded into an alpha-helix with a hydrophilic face and a hydrophobic face. The peptide monomers assembled into an anti-parallel four-helix bundle. This bundle was able to maintain 85% of a bacteriorhodopsin sample and 60% of a rhodopsin sample soluble for 2 days. However, it was unable to solubilize PhoE porin. A lipopeptide detergent (LPD) that used the same basic framework was designed to improve protein stability, and address the problem of imperfect micelle packing [126]. This surfactant was comprised of an alpha helix that had fatty acids coupled to each end. LPD monomers formed a wedge-like shape, and assembled into a cylindrical structure with the fatty acid chains sequestered in the core. LPDs with fatty acid chains of various lengths were able to solubilize bacteriorhodopsin, lactose (lac) permease, PagP, and a beta barrel protein. They were also able to maintain the stability of solubilized bacteriorhodopsin and lac permease over one month at room temperature. Although LPDs may offer many advantages over traditional detergents, they are significantly more expensive to synthesize and purify. It is not

practical to use LPDs during any of the protein purification steps, and it may not even be economically feasible to produce enough material to set up the numerous screens necessary for crystallization experiments [126]. Other developed systems include tripod amphiphiles, and surfactant polymers or amphipols [128-131].

### **1.3.7 Problems with Novel Surfactant Strategies**

Although several novel surfactant systems have been developed and tested, none is ideal. The bio-surfactants are expensive to obtain at high purity, and require proper folding before they become useful. The chemical surfactants can require extensive screening or yield heterogeneous populations that may interfere with crystal growth. Furthermore, most have only been tested on rhodopsin or bacteriorhodopsin. These proteins are easy to obtain or express in large quantities, and their handling, working, and crystallization conditions are well known. Of the handful of other proteins tested, *none are GPCRs*, and all are membrane proteins that have known experimental working conditions. Success with these “easy to handle” membrane proteins does not imply success with more difficult proteins, even of the same family. Limited testing of these surfactants also does not demonstrate their general use as a class among a broad range of proteins. Economical *de novo* surfactants capable of solubilizing and stabilizing membrane proteins that are more difficult to express or handle are needed.

## 1.4 References

1. Klabunde T, and Hessler G. Drug design strategies for targeting G-protein-coupled receptors. *ChemBioChem* **3**, 928-944 (2002).
2. Lundstrom K. Structural biology of G protein-coupled receptors. *Bioorg Med Chem Lett* **15**, 3654-3657 (2005).
3. Kristiansen K. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacol Therapeut* **103**, 21-80 (2004).
4. Violin JD, and Lefkowitz RJ.  $\beta$ -Arrestin-biased ligands at seven-transmembrane receptors. *Trends Pharmacol Sci* **28**(8), 416-422 (2007).
5. Rakesh K, *et al.*  $\beta$ -Arrestin-biased agonism of the angiotensin receptor induced by mechanical stress. *Sci Signal* **3**(125), ra46 (2010).
6. Wei H, *et al.* Independent  $\beta$ -arrestin 2 and G protein-mediated pathways for angiotensin II activation of extracellular signal-regulated kinases 1 and 2. *Proc Natl Acad Sci USA* **100**(19), 10782-10787 (2003).
7. Luttrell LM, and Lefkowitz RJ. The role of  $\beta$ -arrestins in the termination and transduction of G-protein-coupled receptor signals. *J Cell Sci* **115**, 455-465 (2002).
8. Krautwurst D. Human olfactory receptor families and their odorants. *Chem Biodivers* **5**, 842-852 (2008).
9. Zhao H, and Firestein S. Vertebrate odorant receptors. *Cell Mol Life Sci* **56**, 647-659 (1999).
10. Buck L, and Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175-187 (1991).
11. Liberles SD, and Buck LB. A second class of chemosensory receptors in the olfactory epithelium. *Nature* **442**, 645-650 (2006).
12. Dulac C, and Axel R. A novel family of genes encoding putative pheromone receptors in mammals. *Cell* **83**, 195-206 (1995).
13. Herrada G, and Dulac C. A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. *Cell* **90**, 763-773 (1997).

14. Liberles SD, *et al.* Formyl peptide receptors are candidate chemosensory receptors in the vomeronasal organ. *Proc Natl Acad Sci USA* **106**(24), 9842-9847 (2009).
15. Riviere S, Challet L, Fluegge D, Spehr M, Rodriguez I. Formyl peptide receptor-like proteins are a novel family of vomeronasal chemosensors. *Nature* **459**, 574-577 (2009).
16. Crasto C, Singer MS, Shepherd GM. The olfactory receptor family album. *Genome Biol* **2**(10), reviews1027.1-1027.4 (2001).
17. Malnic B, Godfrey PA, Buck LB. The human olfactory receptor gene family. *Proc Natl Acad Sci USA* **101**(8), 2584-2589 (2004).
18. Olender T, *et al.* The canine olfactory subgenome. *Genomics* **83**(3), 361-372 (2004).
19. Safran M, *et al.* Human gene-centric databases at the Weizmann Institute of Science: GeneCards, UDB, CroW 21 and HORDE. *Nucleic Acids Res* **31**(1), 142-146 (2003).
20. Godfrey PA, Malnic B, Buck LB. The mouse olfactory receptor gene family. *Proc Natl Acad Sci USA* **101**(7), 2156-2161 (2004).
21. Zhang X, and Firestein S. The olfactory receptor gene superfamily of the mouse. *Nat Neurosci* **5**(2), 124-133 (2002).
22. Zozulya S, Echeverri F, Nguyen T. The human olfactory receptor repertoire. *Genome Biol* **2**(6), research0018.1-0018.12 (2001).
23. Glusman G, *et al.* The complete human olfactory subgenome. *Genome Res* **11**, 685-702 (2001).
24. Turin L, and Yoshii F. Structure-odor relations: A modern perspective. *Handbook of Olfaction and Gustation*, R Doty, Editor New York Marcel Dekker (2001).
25. Laska M, *et al.* Detecting danger – or just another odorant? Olfactory sensitivity for the fox odor component 2, 4, 5-trimethylthiazoline in four species of mammals. *Physiol Behav* **84**(2), 211-215 (2005).
26. Ebrahimi FA, and Chess A. Olfactory G proteins: Simple and complex signal transduction. *Curr Biol* **8**(12), R431-3 (1998).
27. Chess A, *et al.* Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**(5), 823-834 (1994).
28. Ressler KJ, Sullivan SL, Buck LB. A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* **73**, 597-609 (1993).



29. Mombaerts P. Molecular biology of odorant receptors in vertebrates. *Annu Rev Neurosci* **22**, 487-509 (1999).
30. Mombaerts P, et al. Visualizing an olfactory sensory map. *Cell* **87**, 675-686 (1996).
31. Malnic B, et al. Combinatorial receptor codes for odors. *Cell* **96**(5), 713-723 (1999).
32. Gaillard I, et al. A single olfactory receptor specifically binds a set of odorant molecules. *Eur J Neurosci* **15**(3), 409-418 (2002).
33. Schild D, and Restrepo D. Transduction mechanisms in vertebrate olfactory receptor cells. *Physiol Rev* **78**(2), 429-466 (1998).
34. Touhara K. Odor discrimination by G protein-coupled olfactory receptors. *Microsc Res Tech* **58**(3), 135-141 (2002).
35. Mombaerts P. Seven-transmembrane proteins as odorant and chemosensory receptors. *Science* **286**(5440), 707-711 (1999).
36. Mombaerts P. Genes and ligands for odorant, vomeronasal and taste receptors. *Nat Rev Neurosci* **5**, 263-278 (2004).
37. Cook B, et al. Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4. *Proc Natl Acad Sci USA* **106**(29), 11925-11930 (2009).
38. Kuang D, et al. Molecular similarities in the ligand binding pockets of an odorant receptor and the metabotropic glutamate receptors. *J Biol Chem* **278**(43), 42551-42559 (2003).
39. Katada S, Hirokawa T, Oka Y, Suwa M, Touhara K. Structural basis for broad but selective ligand spectrum of a mouse olfactory receptor: Mapping the odorant binding site. *J Neurosci* **25**(7), 1806-1815 (2005).
40. Baud O, et al. The mouse eugenol odorant receptor: structural and functional plasticity of a broadly tuned odorant binding pocket. *Biochemistry* **50**(5), 843-853 (2011).
41. Lai PC, Singer MS, Crasto CJ. Structural activation pathways from dynamic olfactory receptor-odorant interactions. *Chem Sens* **30**, 781-792 (2005).
42. Kurland MD, et al. Discrimination of saturated aldehydes by the rat I7 olfactory receptor. *Biochemistry* **49**(30), 6302-6304 (2010).

43. Hall SE, Floriano WB, Vaidehi N, Goddard WA III. Predicted 3-D structures for mouse I7 and rat I7 olfactory receptors and comparison of predicted odor recognition profiles with experiment. *Chem Sens* **29**(7), 595-616 (2004).
44. Floriano WB, Vaidehi N, Goddard WA III. Making sense of olfaction through predictions of the 3-D structure and function of olfactory receptors. *Chem Sens* **29**(4), 269-290 (2004).
45. Singer MS. Analysis of the molecular basis for octanal interactions in the expressed rat I7 olfactory receptor. *Chem. Sens* **25**, 155-165 (2000).
46. Hummel P, Vaidehi N, Floriano WB, Hall SE, Goddard WA III. Test of the binding threshold hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93. *Prot Sci* **14**, 703-710 (2005).
47. Floriano WB, Vaidehi N, Goddard WA III, Singer MS, Shepherd GM. Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. *Proc Natl Acad Sci USA* **97**(20), 10712-10716 (2000).
48. Krautwurst D, Yau K-W, Reed RR. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* **95**, 917-926 (1998).
49. Liu AH, Zhang X, Stolovitzky GA, Califano A, Firestein SJ. Motif-based construction of a functional map for mammalian olfactory receptors. *Genomics* **81**, 443-456 (2003).
50. Man O, Gilad Y, Lancet D. Prediction of the odorant binding site of olfactory receptor proteins by human-mouse comparisons. *Prot Sci* **13**, 240-254 (2004).
51. Liberles SD. Trace amine-associated receptors are olfactory receptors in vertebrates. *Ann NY Acad Sci* **1170**, 168-172 (2009).
52. Zucchi R, Chiellini G, Scanlan TS, Grandy DK. Trace amine-associated receptors and their ligands. *Brit J Pharmacol* **149**, 967-978 (2006).
53. Hu LA et al. Human and mouse trace amine-associated receptor 1 have distinct pharmacology towards endogenous monoamines and imidazoline receptor ligands. *Biochem J* **424**, 39-45 (2009).
54. Krieger J, et al. Selective activation of G protein subtypes in the vomeronasal organ upon stimulation with urine-derived compounds. *J Biol Chem* **274**(8), 4655-4662 (1999).
55. Pantages E, and Dulac C. A novel family of candidate pheromone receptors in mammals. *Neuron* **28**, 835-845 (2000).

56. Holy TE, Dulac C, Meister M. Responses of vomeronasal neurons to natural stimuli. *Science* **289**, 1569-1572 (2000).
57. Runnenburger K, Breer H, Boekhoff I. Selective G protein  $\beta\gamma$ -subunit compositions mediate phospholipase C activation in the vomeronasal organ. *Eur J Cell Biol* **81**, 539-547 (2002).
58. Shirokova E, Raguse JD, Meyerhof W, Krautwurst D. The human vomeronasal type-1 receptor family – detection of volatiles and cAMP signaling in HeLa/Olf cells. *FASEB J* **22**, 1416-1425 (2008).
59. Wekesa KS, Miller S, Napier A. Involvement of  $G_{q/11}$  in signal transduction in the mammalian vomeronasal organ. *J Exp Biol* **206**, 827-832 (2003).
60. Lucas P, Ukhanov K, Leinders-Zufall T, Zufall F. A diacylglycerol-gated cation channel in vomeronasal neuron dendrites is impaired in TRPC2 mutant mice: Mechanism of pheromone transduction. *Neuron* **40**, 551-561 (2003).
61. Rodriguez I, and Mombaerts P. Novel human vomeronasal receptor-like genes reveal species-specific families. *Curr Biol* **12**, R409-411 (2002).
62. Trotier D, *et al.* The vomeronasal cavity in adult humans. *Chem Senses* **25**, 369-380 (2000).
63. Bossy J. Development of olfactory and related structures in staged human embryos. *Anat Embryol* **161**, 225-236 (1980).
64. Humphrey T. The development of the olfactory and the accessory olfactory formations in human embryos and fetuses. *J Comp Neurol* **73**(3), 431-468 (1940).
65. Meisami E, and Bhatnagar KP. Structure and diversity in mammalian accessory olfactory bulb. *Microsc Res Tech* **43**, 476-499 (1998).
66. Boehm N, and Gasser B. Sensory receptor-like cells in the human foetal vomeronasal organ. *NeuroReport* **4**, 867-870 (1993).
67. Takami S. *et al.* Vomeronasal epithelial cells of the adult human express neuron-specific molecules. *Neuroreport* **4**, 375-378 (1993).
68. Grosser BI, Monti-Bloch L, Jennings-White C, Berliner DL. Behavioral and electrophysiological effects of androstadienone, a human pheromone. *Psychoneuroendocrino* **25**(3), 289-99 (2000).
69. Berliner DL, Monti-Bloch L, Jennings-White C, Diaz-Sanchez V. The functionality of the human vomeronasal organ (VNO): Evidence for steroid receptors. *J Steroid Biochem Molec Biol* **58**(3), 259-265 (1996).

70. Meredith M. Human vomeronasal organ function: A critical review of best and worst cases. *Chem Sens* **26**, 433-445 (2001).
71. Rodriguez I, Greer CA, Mok MY, Mombaerts P. A putative pheromone receptor gene expressed in human olfactory mucosa. *Nat Genet* **26**, 18-19. (2000).
72. Le Y, Murphy PM, Wang JM. Formyl-peptide receptors revisited. *TRENDS Immunol* **23**(11), 541-548 (2002).
73. Horvath G, Jarverud GA, Jarverud S, Horvath I. Human ovarian carcinomas detected by specific odor. *Integr Cancer Ther* **7**(2), 76-80 (2008).
74. Kuhn F, and Natsch A. Body odour of monozygotic human twins: a common pattern of odorant carboxylic acids released by a bacterial aminoacylase from axilla secretions contributing to an inherited body odour type. *J R Soc Interface* **6**, 377-392 (2009).
75. Penn DJ, Oberzaucher E, Grammer K, Fischer G, Soini HA, Wiesler D, Novotny MV, Dixon SJ, Xu Y, and Brereton RG. Individual and gender fingerprints in human body odour. *J R Soc Interface* **4**, 331-340 (2007).
76. McCulloch M, et al. Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers. *Integr Cancer Ther* **5**(1), 30-39 (2006).
77. Membrane Proteins of Known Structure Website: [http://blanco.biomol.uci.edu/Membrane Proteins xtal.html](http://blanco.biomol.uci.edu/Membrane%20Proteins%20xtal.html)
78. Protein Data Bank Website: <http://www.pdb.org/pdb/home/home.do>
79. Grisshammer R, Duckworth R, Henderson R. Expression of a rat neurotensin receptor in *Escherichia coli*. *Biochem J* **295**, 571-576 (1993).
80. Hampe W, et al. Engineering of a proteolytically stable human  $\beta_2$ -adrenergic receptor/maltose-binding protein fusion and production of the chimeric protein in *Escherichia coli* and baculovirus infected insect cells. *J Biotechnol* **77**, 219-234 (2000).
81. Sarramegna V, Talmont F, Demange P, Milon A. Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification. *Cell Mol Life Sci* **60**, 1529-1546 (2003).
82. McCusker EC, Bane SE, O'Malley MA, Robinson AS. Heterologous GPCR expression: A bottleneck to obtaining crystal structures. *Biotechnol Prog* **23**, 540-547 (2007).
83. Kaushal S, Ridge KD, Khorana HG. Structure and function in rhodopsin: The role of asparagine-linked glycosylation. *Proc Natl Acad Sci USA* **91**, 4024-4028 (1994).

84. Zhu L, et al. A naturally occurring mutation of the opsin gene (T4R) in dogs affects glycosylation and stability of the G protein-coupled receptor. *J Biol Chem* **279**(51), 53828-53839 (2004).
85. Katada S, Tanaka M, Touhara K. Structural determinants for membrane trafficking and G protein selectivity of a mouse olfactory receptor. *J Neurochem* **90**, 1453-1463 (2004).
86. Tate CG, and Grisshammer R. Heterologous expression of G-protein-coupled receptors. *Trends Biotechnol* **14**, 426-430 (1996).
87. Reilander H, and Weiss HM. Production of G-protein-coupled receptors in yeast. *Curr Opin Biotechnol* **9**, 510-517 (1998).
88. Cherezov V, et al. High resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science* **318**(5854), 1258-1265 (2007).
89. Chien et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091-1095 (2010).
90. Jaakola V-P, et al. The 2.6 Å crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**(5905), 1211-1217 (2008).
91. Rasmussen SGF, et al. Crystal structure of the human  $\beta_2$  adrenergic G-protein-coupled receptor. *Nature* **450**, 383-388 (2007).
92. Warne T, et al. Structure of a  $\beta_1$ -adrenergic G protein-coupled receptor. *Nature* **454**(7203), 486-491 (2008).
93. Wu B, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071 (2010).
94. Boundy VA, Lu L, Molinoff PB. Differential coupling of rat D2 dopamine receptor isoforms expressed in *Spodoptera Frugiperda* insect cells. *J Pharmacol Exp Ther* **276**(2), 784-794 (1996).
95. Reilander H, et al. Purification and function characterization of the human  $\beta_2$ -adrenergic receptor produced in baculovirus-infected insect cells. *FEBS* **282**(2), 441-444 (1991).
96. Cook B, Ernberg KE, Chung H, Zhang S. Study of a synthetic human olfactory receptor 17-4: Expression and purification from an inducible mammalian cell line. *PLoS One* **3**(8): e2920. (2008).
97. Reeves PJ, Thurmond RL, Khorana HG. Structure and function in rhodopsin: high level expression of a sythetic bovine opsin gene and its mutants in stable mammalian cell lines. *Proc Natl Acad Sci USA* **93**, 11487-11492 (1996).

98. Corin K, et al. The human vomeronasal type 1 receptor 1: large-scale purification, and secondary structure and ligand binding analyses. *Scientific Reports* (submitted).
99. Wang X, Corin K, Rich C, Zhang S. Study of two G-protein coupled receptor variants of human trace amine-associated receptor 5. *Scientific Reports* (submitted).
100. Ishihara G, et al. Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Express Purif* **41**, 27-37 (2005).
101. Gourdan P, et al. Optimized *in vitro* and *in vivo* expression of proeiorhodopsin: A seven-transmembrane proton pump. *Protein Express Purif* **58**, 103-113 (2008).
102. Kaiser L, et al. Large-scale production and study of a synthetic G protein-coupled receptor: human olfactory receptor 17-4. *Proc Natl Acad Sci USA* **105**: 15726-15731 (2008).
103. Klammt C, et al. Cell-free production of G protein-coupled receptors for functional and structural studies. *J Struct Biol* **158**, 482-493 (2007).
104. Katzen F, Peterson TC, Kudlicki W. Membrane protein expression : no cells required. *Trends Biotechnol* **27**(8), 455-460 (2009).
105. Klammt C, et al. High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568-580 (2004).
106. Klammt C, et al. Evaluation of detergents for the soluble expression of  $\alpha$ -helical and  $\beta$ -barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* **272**, 6024-6038 (2005).
107. Keller T, et al. Cell free expression and functional reconstitution of eukaryotic drug transporters. *Biochemistry* **47**, 4552-4564 (2008).
108. Nomura SM, et al. Direct preparation of giant proteo-liposomes by *in vitro* membrane protein synthesis. *J Biotechnol* **133**, 190-195 (2008).
109. Kalmbach R, et al. Functional cell-free synthesis of a seven helix membrane protein: *In situ* insertion of bacteriorhodopsin into liposomes. *J Mol Biol* **371**, 639-648 (2007).
110. Nath A, Atkins WM, Sligar SG. Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **46**(8), 2059-2069 (2007).

111. Katzen F, et al. Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J Proteome Res* **7**, 3535-3542 (2008).
112. Ostermeier C, Iwata S, Ludwig B, Michel H. Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nat Struct Biol* **2**, 842-846 (1995).
113. Ostermeier C, Harrenga A, Ermler U, Michel H. Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody Fv fragment. *Proc Natl Acad Sci USA* **94**, 10547-10553 (1997).
114. Hunte C, Koepke J, Lange C, Rossmann T, Michel H. Structure at 2.3 Å resolution of the cytochrome bc<sub>1</sub> complex from the yeast *Saccharomyces cerevisiae* co-crystallized with an antibody Fv fragment. *Structure* **8**, 669-684 (2000).
115. Uysal S, et al. Crystal structure of full-length KcsA in its closed conformation. *Proc Natl Acad Sci USA* **106**(16), 6644-6649 (2009).
116. Iwata S (editor). *Methods and Results in Crystallization of Membrane Proteins*. International University Line (2003).
117. Caffrey M. Membrane protein crystallization. *J Struct Biol* **142**(1), 108-132 (2003).
118. Landau EM, and Rosenbusch JP. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci USA* **93**, 14532-14535 (1996).
119. Garavito RM, and Ferguson-Miller S. Detergents as tools in membrane biochemistry. *J Biol Chem* **276**(35), 32403-32406 (2001).
120. Prive GG. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
121. Anatrace website:  
<http://www.affymetrix.com/browse/brand/anatrace/anatrace-overview.jsp?category=35858&categoryIdClicked=35858&rootCategoryId=35677&navMode=35858&parent=35858&ald=anatraceNav>
122. Arnold T, and Linke D. The use of detergents for purify membrane proteins. *Curr Protoc Protein Sci* **53**, 4.8.1-4.8.30 (2008).
123. Helenius A, McCaslin DR, Fries E, Tanford C. Properties of detergents. *Methods Enzymol* **56**, 734-749 (1979).
124. Garavito RM, Picot D, Loll PJ. Strategies for crystallizing membrane proteins. *J Bioenerg Biomembr* **28**(1), 13-27 (1996).

125. Pebay-Peyroula E, Garavito RM, Rosenbusch JP, Zulauf M, Timmins PA. Detergent structure in tetragonal crystals of OmpF porin. *Structure* **3**(10), 1051-1059 (1995).
126. McGregor CL, et al. Lipopeptide detergents designed for the structural study of membrane proteins. *Nat Biotechnol* **21**, 171-176 (2003).
127. Schafmeister CE, Miercke LJW, Stroud RM. Structure at 2.5Å of a designed peptide that maintains solubility of membrane proteins. *Science* **262**, 734-738 (1993).
128. Popot JL, et al. Amphipols: polymeric surfactants for membrane biology research. *Cell Mol Life Sci* **60**, 1559-1574 (2003).
129. Theisen MJ, Potocky TB, McQuade DT, Gellman SH, Chiu ML. Crystallization of bacteriorhodopsin solubilized by a tripod amphiphile. *Biochim Biophys Acta* **1751**, 213-216 (2005).
130. Tribet C, Audebert R, Popot JL. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci USA* **93**(26), 15047-15050 (1996).
131. Yu SM, et al. An improved tripod amphiphile for membrane protein solubilization. *Protein Sci* **9**, 2518-2527 (2000).
132. Young JM, and Trask BJ. The sense of smell: genomics of vertebrate odorant receptors. *Hum Mol Genet* **11**(10), 1153-1160 (2002).



# CHAPTER 2

## LARGE-SCALE EXPRESSION OF hVN1R1 IN A STABLE HEK293 CELL LINE

*This chapter is an expansion of a manuscript submitted to Scientific Reports:*

*Corin, K., Baaske, P., Geissler S., Wienken C.J., M., Duhr, S., Braun, D. & Zhang, S. (2011) Biochemical study of a bioengineered functional G-protein coupled receptor: the human vomeronasal type 1 receptor 1. Scientific Reports (submitted).*

### 2.1 Introduction

#### 2.1.1 HEK293 Cells as a GPCR Expression Platform

Large-scale expression of GPCRs is necessary before structural or functional studies can be performed. However, this presents a daunting task. GPCRs are endogenously expressed at low levels, and high levels of expression in heterologous systems typically result in cell toxicity, protein misfolding, or aggregation.

Mammalian cells are arguably the most ideal cells in which to express GPCRs, as the environment they provide is closest to the native environment of most GPCRs. They also have the necessary machinery to perform glycosylation and other post-translational modifications. However, yields are typically low, and decrease with each cell passage after transient transfections. Moreover, the expressed proteins can be in various stages of

expression or degradation. Such an inhomogeneous sample can inhibit crystallization, and may also interfere with other assays.

Large-scale expression of rhodopsin and an olfactory receptor in HEK293 cells has previously been reported [1-3]. The authors were able to generate stable inducible cell lines that yielded milligram quantities of homogenous receptors. They were also able to purify the proteins using size exclusion chromatography alone, and show that the purified protein was properly folded and functional. However, the expressed proteins retained their glycosylation sites, and were glycosylated [2]. In order for purified protein to be useful for crystallization screens, the carbohydrate chains must be removed without altering the receptors' structure or function.

Success with two receptors suggests that HEK293 cells may be able to produce additional GPCRs. Little information is available on human vomeronasal receptors. This is partly due to the ambiguity of whether humans have a functional vomeronasal organ, and partly due to the difficulty of expressing these receptors. A human vomeronasal receptor is thus an attractive candidate for expression in HEK293 cells.

### **2.1.2 The Human Vomeronasal Organ and Receptors**

Although the existence and function of a human vomeronasal organ (VNO) is controversial, almost all studies agree that a vomeronasal organ is present in the developing fetus. However, many studies report that this organ is vestigial, and function is lost in adults, while others report development of a mature organ [4-9]. Interestingly, one study found that the ability to detect the adult VNO can vary; the investigators could not

find VNOs in subjects that had clearly identifiable organs on prior inspections, and vice versa [4].

In most mammals, signals from the VNO are directed to the accessory olfactory bulb (AOB). However, studies in humans suggest that degenerative changes occur in the fetal AOB, and that it isn't present in adults [7, 8]. Most studies that used neuron-specific stains failed to find evidence of nerve bundles [4, 9]. Others suggest the presence of neurons in the adult vomeronasal organ, albeit at a low density [10]. In spite of this, some studies indicate pheromonal responses in humans directly through the VNO. Androstadienone applied directly to the female VNO decreased nervousness, tension, and negative feeling states. Decreases in cardiac frequency and respiratory frequency, and increases in body temperature and alpha-cortical activity, were also measured [11]. When a steroidal vomeropherin was applied to the male VNO, a decrease in follicle-stimulating hormone and luteinizing hormone pulsatility was observed. Changes in cardiac frequency, respiratory frequency, and other physiological responses were also observed [12].

Taken together, current evidence suggests that humans may have a functional vomeronasal organ that functions differently than in other mammals. However, until recently, it was believed that human vomeronasal receptors (VNRs) were all pseudogenes. In 2002, Rodriguez and Mombaerts found 5 VNR type 1 (VN1R) genes with open reading frames. They were successfully expressed in HeLa/Olf cells [14]. Interestingly, all five were found to respond to volatile molecules and signal through the canonical olfactory signaling pathway. Even more curiously, mRNA of one receptor – hVN1R1 – was detected in several other tissues, including the main olfactory epithelium (MOE), brain, lung, and

kidney [15]. This raises the possibility that human VN1Rs may function more like olfactory receptors in the MOE, or may have non-olfactory functions.

Studying human vomeronasal receptors at the molecular level may shed light on both their function and that of the human vomeronasal organ. However, the difficulty of functionally expressing and purifying these G protein-coupled receptors in sufficient quantities is a major obstacle towards research. Only one report has produced functional receptors in cells, and identified potential ligands [14]. However, in order to study their structure and function, it is necessary to produce and purify milligram-scale quantities. This study shows that it is possible to generate a stable, inducible HEK293 cell line capable of expressing milligram quantities of the human vomeronasal receptor hVN1R1. A novel finding is that removal of glycosylation sites does not impede membrane localization of the expressed receptors. Circular dichroism indicated that the purified receptor was properly folded with stable  $\alpha$ -helical domains. Microscale thermophoresis showed that the purified receptor was able to bind its reported ligand. These results provide the basis for carrying out further structural analyses of human vomeronasal receptors, and enable screens for crystallization trials. They could also aid in future studies aimed at identifying the biological role of human VNRs.

## **2.2 Methods**

### **2.2.1 hVN1R1 Gene Design and Construction**

The protein sequence for hVN1R1 was obtained from GenBank (AAG10698). To enable expression and purification from mammalian cells, the following modifications were made: 1) addition of a C-terminal rho tag (TETSQVAPA) preceded by a two glycine linker;

2) human codon optimization; 3) addition of a Kozak sequence 5' to the start codon; 4) addition of a 5' EcoRI site and a 3' XhoI site to facilitate subcloning into expression vectors; 5) addition of an N-terminal strep tag (ASWSHPQFEK) followed by a GSSG linker for further purification; and 6) N117Q, N151Q, N183Q, N198Q, and N256Q mutations to facilitate crystallization. The genes were constructed by Geneart and ligated into the pcDNA 4/To vector (Invitrogen, Carlsbad, CA). The plasmid was amplified in subcloning efficiency DH5 $\alpha$  E. coli (Invitrogen) and purified using MiniPrep or MaxiPrep kits (Quiagen, Valencia, CA). The transmembrane and loop domains were predicted using the TMHMM Server v 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). N-linked glycosylation sites were predicted using the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

### **2.2.2 Construction of Stable Inducible hVN1R1 HEK293G Cell Lines**

HEK293S N-acetylglucosaminyltransferase I-negative cells (HEK293G) containing the pcDNA6/Tr vector [2] were transfected with the pcDNA4/To hVN1R1 vector using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Forty-eight hours after transfection, selective media containing 5 $\mu$ g/ml of blasticidin and 50 $\mu$ g/ml zeocin was added. Cells were re-seeded at low density and grown until individual colonies formed. Forty-eight colonies were picked and screened for inducible hVN1R1 expression. Cells were treated with plain media, media supplemented with 1 $\mu$ g/ml tetracycline, and media supplemented with 1 $\mu$ g/ml tetracycline and 2.5mM sodium butyrate. Two days after induction, cells were scrape-harvested and solubilized in PBS with 2% w/v Fos-Choline 14 (FC-14) (Anatrace) and protease inhibitors (Roche #04693132001) for 1 hr at 4°C. Cell lysates were centrifuged for 30 minutes at 10,000 rpm to remove insoluble debris.

Dot blots and Western-blotting were used to compare protein expression among clones. The clone with the highest expression when induced, the least detectable expression when not induced, and least toxicity upon induction, was expanded for future experiments. All cultures were grown in DMEM F12 with GlutaMAX (Invitrogen #10565-042) supplemented with 10% fetal bovine serum (Invitrogen #16000-044), 15mM HEPES (Invitrogen), 0.1mM non-essential amino acids (Invitrogen), 0.5mM sodium pyruvate (Invitrogen), 100Units/ml penicillin and 100µg/ml streptomycin (Invitrogen). The expanded stable hVN1R1 clone was maintained in media that also contained 5µg/ml blasticidin and 25µg/ml zeocin. All cells were grown at 37°C, 5% CO<sub>2</sub>, and 95% relative humidity.

### **2.2.3 Immunocytochemistry**

Vomer nasal receptors were visualized using a rho1D4 primary antibody. Cells were seeded at low density on poly-L-lysine (Sigma-Aldrich) coated glass coverslips. After one day, cells were induced with 1µg/ml tetracycline and 1mM sodium butyrate. One day after induction, the media was removed. Cells were gently washed with PBS and fixed for 20 minutes in 10% neutral buffered formalin (Sigma-Aldrich) at room temperature. Permeabilized (1:1 acetone:methanol, 3 minutes, -20°C) and non-permeabilized cells were then blocked in PBST (PBS, 0.2% tween-20, 0.3M glycine, 4% serum) for 1 hour at room temperature, and incubated with the primary antibody solution (1:1000, PBS, 0.2% tween-20, 4% serum) overnight at 4°C. The labeled protein was visualized with Alexa-fluor-488 goat-anti-mouse secondary antibody conjugate (1:3000, PBS, 1 hour, room temperature). Slides were mounted using ProLong Gold Antifade with DAPI.

#### **2.2.4 Cell Extract Preparation**

Cells were grown on plates as previously described [2]. When the appropriate density was reached, cells were induced with 1 $\mu$ g/ml tetracycline and 2.5mM sodium butyrate. Two days after induction, cells were scraped harvested, pooled, snap frozen in liquid nitrogen, and stored at -80°C until used for future experiments.

#### **2.2.5 Detergent Screening**

Frozen cell pellets were thawed on ice and resuspended in PBS containing protease inhibitors (Roche). Detergents were added to a final concentration of 2% w/v. The suspensions were rotated for 1 hour at 4°C to solubilize the protein, and were spun at 13,000 rpm for 30 minutes to remove insoluble fractions. Relative protein solubilization in each detergent was assayed with a dot blot. Ninety-six detergents were selected for screening as previously described [16].

#### **2.2.6 Receptor Purification**

Rho1D4 immunoaffinity purification has been previously described [2, 3]. Briefly, frozen cell pellets were thawed on ice. Cells were resuspended in PBS containing protease inhibitors. PBS containing FC-14 was added to a final concentration of 2% w/v FC-14. The final liquid:cell ratio was 12.5ml/1g cells. The protein was solubilized by rotating for 4 hours at 4°C. The non-solubilized fraction was pelleted by centrifuging for 30 minutes at 30,000g at 4°C. The solubilized fraction was incubated with DNase (1:2000) and RNase (1:1000) for 15 minutes on ice. Rho1D4-coupled CNBr-activated Sepharose 4B beads (GE Healthcare) were added to the cell extract supernatant (binding capacity 0.7mg/ml);

receptors were captured by rotating the mixture overnight at 4°C. The beads were collected by centrifuging at 1400 rpm for 1 minute, or filtering the supernatant through a filter column (Biorad). The supernatant was saved for future analysis and labeled as “flow-through”. The beads were resuspended in 1 bead volume of wash buffer (PBS + 0.2% w/v FC-14), rotated for 10 minutes at 4°C, and re-pelleted. Washes were performed until the total protein concentration in the washes was less than 0.01mg/ml (NanoDrop). One bead volume of elution buffer (PBS+0.2% w/v FC-14 + 800uM Ac-TETSQVAPA-NH<sub>2</sub> peptide) was then added to the beads. Elutions were performed until the total protein concentration was less than 0.01mg/ml.

Size exclusion chromatography was used to separate the monomeric and higher molecular-weight forms of the receptor. A Hi-Load 16/60 Supradex 200 column with an Akta Purifier HPLC system (GE Healthcare) was used. The column was first equilibrated with at least 1 column volume of wash buffer. Protein samples were concentrated to 1.5-3ml using a 50,000 MWCO filter column (Millipore), loaded on the column, and run with wash buffer at 0.3ml/min. Fractions exiting the column were automatically collected; protein content was monitored with UV absorbance at 215nm, 254nm, and 280nm. Peak fractions were pooled, concentrated, and analyzed with Western blotting and silver staining (SilverXpress, Invitrogen).

### **2.2.7 Secondary Structure Analysis Using Circular Dichroism Spectroscopy**

CD spectra were measured over the wavelengths 200nm-350nm with a CD spectrometer (AVIV Biomedical Model 202). Measurements were made at 15°C, with a step size of 1nm and an averaging time of 4 seconds. Measurements for each sample were made



in triplicate and averaged. Protein samples were concentrated to at least 0.1mg/ml for far UV readings, and 1.7mg/ml for near UV readings; concentrations were determined by averaging ten spectroscopic measurements (Nanodrop). The protein spectra were blanked to the spectrum obtained for wash buffer. A QS quartz cuvette (Hellma) with a 1mm path length was used to perform all experiments.

### **2.2.8 Ligand Binding Measurements Using Microscale Thermophoresis**

Thermophoresis was used to measure the binding interactions between purified receptors and their ligands using a setup similar to that previously described [17, 18]. To eliminate artifacts caused by labeling or modifying proteins, the fluorescence of native GPCR tryptophans was used to monitor the local receptor concentration. For each sample, a titration series with constant receptor concentration and varying ligand concentrations was prepared in a final solution of 10% DMSO and 0.2% FC-14 in PBS. Potential autofluorescence of each ligand was checked: no fluorescence signal was detected from the ligands in the tryptophan fluorescence channel. The final receptor concentration was 1-2 $\mu$ M. Approximately 1.5 $\mu$ l of each sample was loaded in a fused silica capillary (Polymicro Technologies, Phoenix, USA) with an inner diameter of 300 $\mu$ m. An infrared laser diode was used to create a 0.12K/ $\mu$ m temperature gradient inside the capillaries (Furukawa FOL1405-RTV-617-1480, wavelength  $\lambda$ =1480nm, 320mW maximum power, AMS Technologies AG, M $\ddot{u}$ nich Germany). Tryptophan fluorescence was excited with a UV-LED (285nm), and was measured with a 40x SUPRASIL synthetic quartz substrate microscope objective, numerical aperture 0.8 (Partec, Goerlitz, Germany). The local receptor concentration in response to the temperature gradient was detected with a photon counter

PMT P10PC (Electron Tubes Inc, Rockaway, NJ, USA). All measurements were performed at room temperature. Fluorescence filters for tryptophan (F36-300) were purchased from AHF-Analysentechnik (Tübingen, Germany).

## **2.3 Results**

### **2.3.1 Induction of hVN1R1 Expression in a Stable HEK293 Cell Line**

Producing milligram quantities of homogenous membrane protein samples is notoriously difficult. Over-expression of these proteins in mammalian cells is often reported to have toxic effects, which reduce the total protein yield. Additionally, constitutive expression of such proteins yields non-homogeneous samples, as proteins can be in various stages of synthesis, degradation, or post-translational modification. To minimize toxic effects, and to generate a homogenous protein sample, the Invitrogen T-REx system was used to create stable hVN1R1- inducible cell lines [2]. This system allows large batches of cells to be grown and simultaneously induced to express protein.

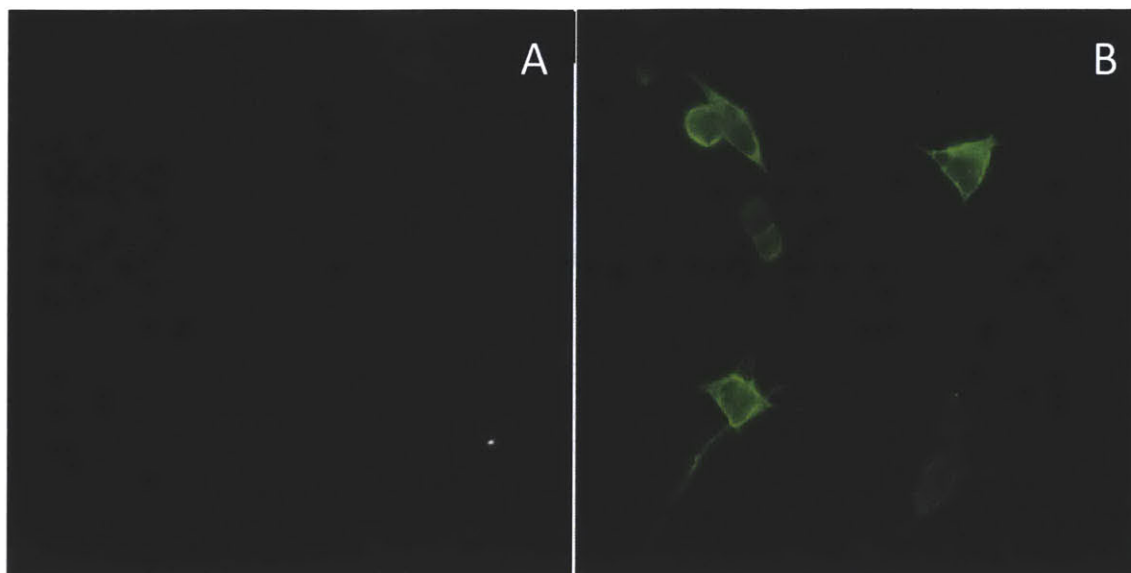
Sodium butyrate has previously been used to enhance tetracycline-regulated protein expression [1, 2]. Induction of VN1R1 expression was thus tested with tetracycline alone, and tetracycline combined with sodium butyrate at various concentrations. No detectable protein expression was seen in the absence of induction agents. Sodium butyrate combined with tetracycline enhanced protein expression 4-5 fold over induction with tetracycline alone. Significant cell death was observed in cultures treated with both tetracycline and sodium butyrate, indicating the toxic effects of protein expression. The clone with the highest expression, the least toxicity, and least degradation products was

selected and expanded. For all subsequent experiments, the following induction condition was used: 1 µg/ml of tetracycline and 2.5 mM of sodium butyrate for 48 hours.

SDS-PAGE and Western blotting analysis were used to characterize the protein samples. Immunoblotting against the rho1D4 tag revealed two bands, which correspond in size to monomeric and dimeric forms of the receptors. This size pattern has been reported for several olfactory receptors [2, 19, 20], and is the first such reporting for a human vomeronasal receptor. The hVN1R1 sample also showed evidence of a degradation product.

### **2.3.2 Immunohistochemical Staining of Induced Cells**

Numerous studies report the difficulty of expressing GPCRs in heterologous systems, and particularly of targeting proteins to the membrane [21-24]. Glycosylation, particularly of specific conserved N-terminal sites, may be necessary for proper protein folding and localization [25]. However, glycosylation can lead to inhomogenous samples and make crystal packing difficult. Here, the native protein was used with potential glycosylation sites removed. To determine whether removal of predicted glycosylation sites affected protein localization, induced cells were stained with antibodies against the rho1D4 tag. **(Figure 2.1)** Non-induced cells showed negligible fluorescence. Permeabilized cells stained with the rho1D4 tag demonstrated membrane-localization of both proteins. These results indicate that the protein is trafficked to the cell membrane. They further demonstrate that the hVN1R1 glycosylation sites are not necessary for membrane localization in HEK293 cells, and indicate that they are not necessary for appropriate protein folding.



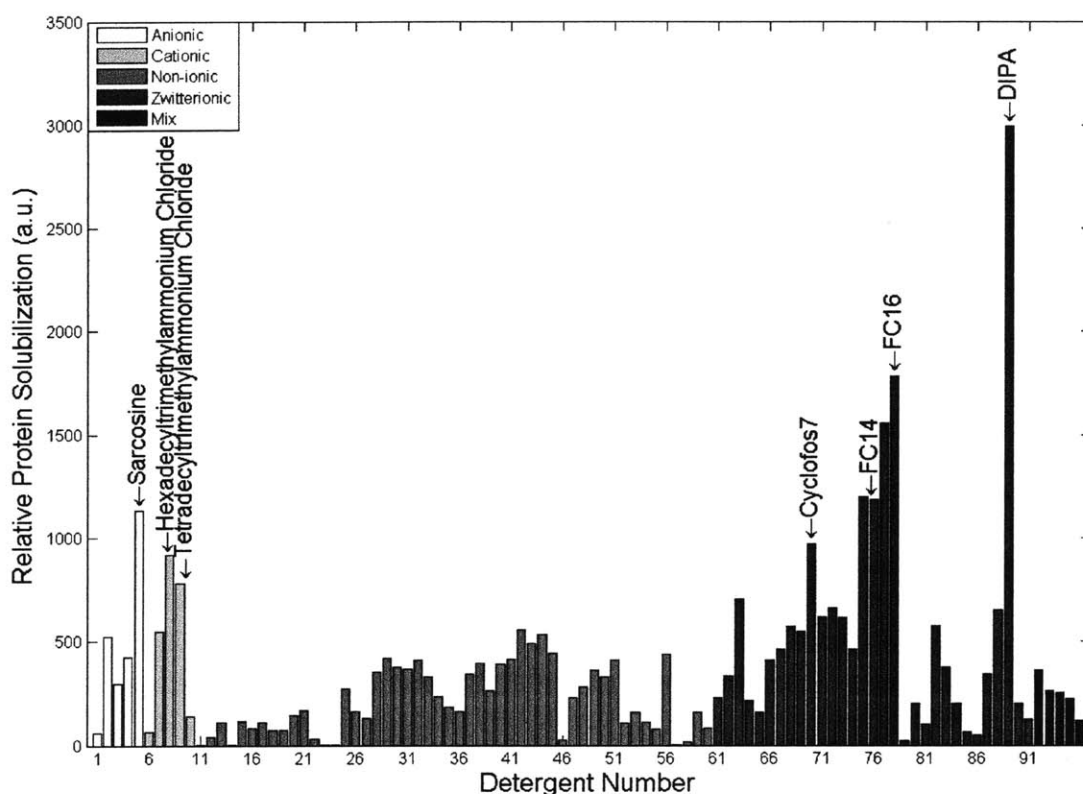
**Figure 2.1. Immunohistochemical Staining of HEK293 Cells Expressing hVN1R1.** A) Non-induced HEK293 cells show no staining. B) Induced HEK293 cells stained with the rho1D4 antibody show that the expressed hVN1R1 is localized to the cell membrane.

### 2.3.3 Systematic Detergent Screening for Receptor Solubilization

Selection of an appropriate detergent is critical for the successful solubilization of a GPCR from a cell membrane, and for the subsequent purification and stabilization of the receptor. Since the optimal detergent can vary even between proteins in the same family, it must be empirically determined. Thus, a systematic detergent screen was performed (**Figure 2.2**).

Ninety-six detergents were screened. These detergents belonged to one of four general classes: anionic, cationic, non-ionic, and zwitterionic. The detergents were chosen from a commercial Solution Master Detergent Kit [16], and included detergents that have been successfully used to purify, solubilize, or crystallize GPCRs. Additionally, several

detergent mixtures were tested as previously described [16]. All detergents were used at a concentration of 2%, which was above the critical micelle concentration, with the exception of MEGA-8.



**Figure 2.2. Detergent Screens for Solubilizing hVN1R1 from HEK293 Cells.** HEK293 cells were induced with tetracycline and sodium butyrate for 48 hours. The cells were then scrape-harvested and incubated with the shown detergents in order to disrupt the membrane and extract the expressed protein. A dot blot was used to compare the relative amounts of solubilized receptor. Zwitterionic detergents typically solubilized a greater portion of the expressed protein, though several cationic and anionic detergents were able to solubilize significant amounts. The fos-choline detergents were the sub-class of detergents that solubilized the greatest amount of protein. The names corresponding to the detergent numbers are listed in Chapter 3.

The most effective detergents for hVN1R1 solubilization were the zwitterionic fos-choline series (#71-78) and n-Dodecyl  $\beta$ -iminodpropionic acid (DIPA, #89). The zwitterionic cyclofos-7, cationic hexadecyl- and tetradecyltrimethylammonium chloride

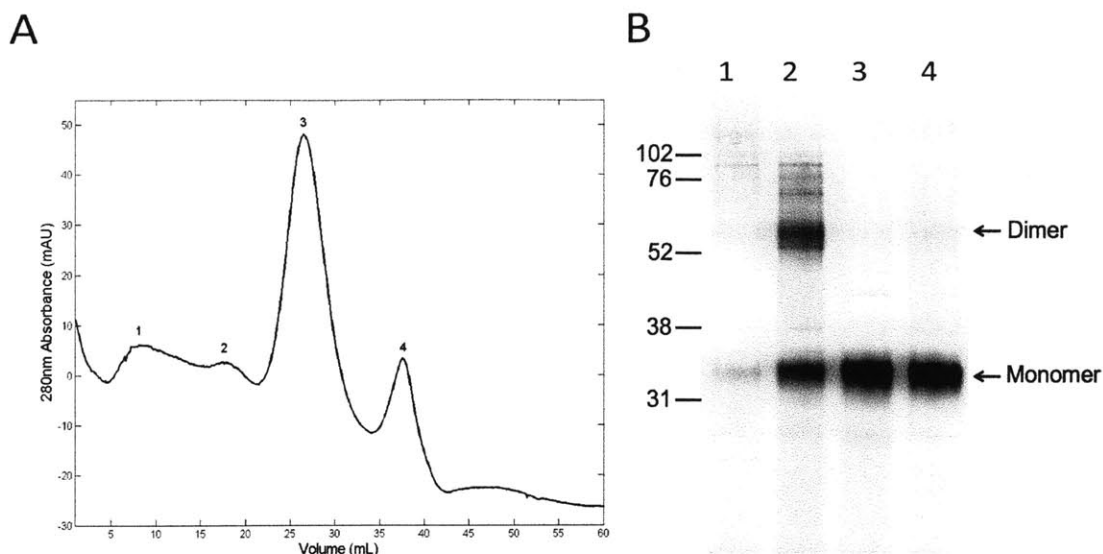
detergents (#8, 9), and anionic sodium dodecanoyl sarcosine (#5) were also reasonably effective, while most other detergents failed to solubilize a significant amount of receptor. Similar results with other GPCRs have previously been reported [3, 16]. The effectiveness of the fos-choline (FC) series to solubilize membrane proteins is not surprising: they are structurally related to phosphatidylcholine, a constituent of the phospholipid bilayer. Although FC15, FC16, and DIPA were able to solubilize more protein, FC14 was chosen for all subsequent experiments: The CMC for DIPA has not been established, and FC15 and FC16 have much lower CMCs, which could inhibit detergent substitution in downstream experiments. Furthermore, previous experiments indicate that the fos-choline family may promote proper folding as well as stabilize the protein structure [26]. Additionally, FC14 has been used for membrane protein solubilization and purification [2, 3, 16, 27], and has been used to crystallize and solve two high-resolution protein structures [28, 29].

Because the optimal detergent for a given membrane protein must currently be determined empirically, finding a detergent that can stabilize a class of membrane proteins would be an important contribution. This would save time and resources, and also shorten the path towards studying more membrane proteins. These results contribute to a growing body of evidence that FC14 is a promising detergent for solubilizing diverse GPCRs from cells, and stabilizing the receptors. In addition to hVN1R1, it has solubilized and stabilized various other GPCRs, including several chemokine receptors (CCR3, CCR5, CXCR4, CX3CR1), several olfactory receptors, and the trace-amine associated receptor 5 [2, 16, 27, 30]. FC14 thus seems like a promising choice in membrane protein studies, and particularly in crystallization trials.

#### 2.3.4 Purification of HEK293-Expressed hVN1R1

A two-step process using immunoaffinity chromatography and size-exclusion chromatography (SEC) was used to purify the expressed hVN1R1 receptors (**Figure 2.3**) [2, 3].

A small-scale purification was performed to determine the potential of using heterologous expression for large-scale protein production. Solubilized protein was first bound to rho1D4-tagged Sepharose 4B beads. After thoroughly washing the column to remove impurities, the protein was eluted with an excess of the rho1D4 epitope peptide Ac-TETSQVAPA-NH<sub>2</sub>. To further purify the protein, remove the elution peptide, and separate the monomeric and dimeric forms, the receptor sample was subjected to SEC. The column flow through was monitored with UV absorption at 280 nm and 215 nm, and was automatically collected in separate fractions. The UV spectrum showed 4 distinct peaks (**Figure 2.3A**). The fractions corresponding to these peaks were pooled, concentrated, and analyzed with western blotting and silver staining. As seen in **Figure 2.3B**, peaks 3 and 4 correspond to the monomeric form, which could be obtained at >90% purity. Earlier peaks primarily corresponded to dimerized and aggregated protein, though the monomeric form was present throughout. Western blotting and total protein stains demonstrated that all of the expressed receptor had been captured by the beads. One gram of cells yielded approximately 1 mg of total protein. Of the protein recovered after SEC, up to 66% was the monomer. These yields are sufficient to obtain large quantities of highly pure receptor for biochemical and analyses and crystal screens. Bioreactor suspension cultures could also potentially be used to increase the protein yield [3].



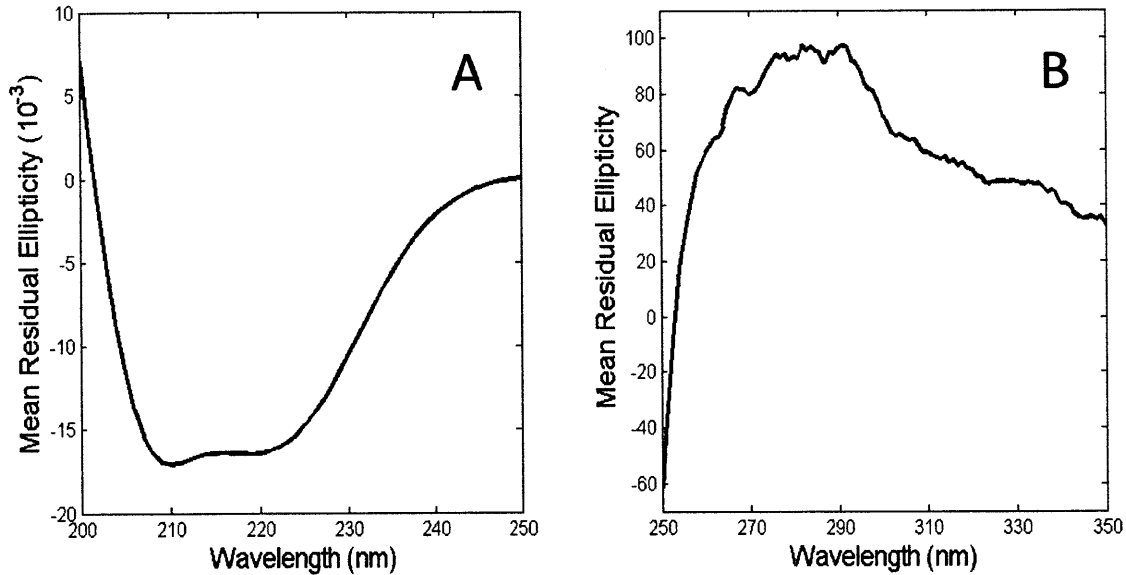
**Figure 2.3. Purification of hVN1R1.** A) 280nm UV trace of hVN1R1 through a size exclusion column (SEC). Four distinct peaks were observed. The fractions corresponding to the peaks were pooled and run on a gel. B) Silver stain of the SEC fractions shown in A). The first peak in A) primarily corresponded to higher molecular weight aggregates. The second peak contained dimerized receptor. The third and fourth peaks contained a monomeric version of hVN1R1. This form of the receptor constituted about 70% of the protein recovered after SEC, and was present in all of the samples.

### 2.3.5 Structural Characterization of Purified hVN1R1

Circular dichroism (CD) was used to determine whether the FC14-solubilized and purified hVN1R1 was properly folded (**Figure 2.4**). Far and near UV were used to probe the secondary and tertiary structure of the purified receptor. **Figure 2.4A** shows the far UV spectrum. The CD spectrum showed a characteristic  $\alpha$ -helical shape with valleys at about 208 and 220 nm. This is expected, since hVN1R1 is a GPCR with 7 predicted transmembrane helices. **Figure 2.4B** shows the near UV spectrum. The near UV spectrum indicates that the purified VN1R1 has a well-defined tertiary structure, as misfolded receptors are not expected to yield any peaks or show a CD trace [3, 31]. Together, the CD spectra suggest that HEK293-expressed hVN1R1 is properly folded, and that FC14 is able to



maintain this structure once the receptor has been extracted from the cell membrane and purified.



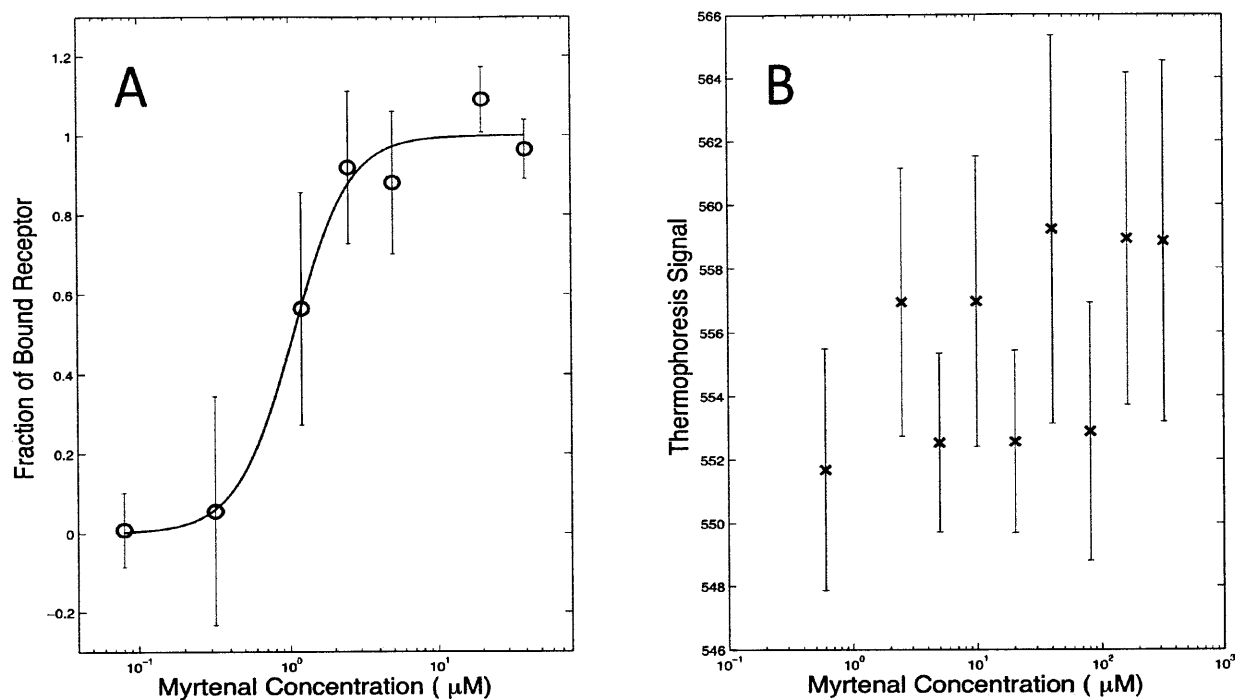
**Figure 2.4. Circular Dichroism Spectra of Purified hVN1R1.** A) Far UV trace showing a curve characteristic of an alpha-helix. This shape is expected, as hVN1R1 is a GPCR with seven predicted transmembrane helices. B) Near UV trace. The peaks indicate that the protein has a well-defined tertiary structure.

### 2.3.6 Ligand Binding Characterization of Purified hVN1R1

Microscale thermophoresis (MST) was used to determine whether the FC14 solubilized receptor could bind its ligand myrtenal (MW 152.23). MST is the directed movement of molecules along a spatial temperature gradient. This movement is sensitive to changes in the hydration shell surrounding the molecule. Ligand binding alters this shell in a way that measurably changes the molecules' thermophoretic movement [17, 18, 32]. Unlike SPR or other surface-based techniques, MST does not require immobilization. The

molecules are monitored in free solution. Additionally, proteins can be tracked by detecting the fluorescence of native tryptophans. Coupling chemistries or other modifications that could potentially alter a receptor's function are thus not necessary.

The purified monomeric form of hVN1R1 was used for ligand-binding measurements. Samples were titrated with varying concentrations of myrtenal. Myrtenal has been reported as a ligand with an expected  $EC_{50}$  of  $\sim 22\mu\text{M}$  in HeLa/Olf cells [14]. **Figure 2.5** shows the MST titration curves normalized to the fraction of bound receptor. The monomer exhibits a plateau at low concentrations and at high concentrations. Fitting the Hill equation to the data yields a  $K_D$  of  $\sim 1\mu\text{M}$ . This result is similar to the result reported by Shirokova et al, and is also in the same range as  $K_D$  values measured for olfactory receptors [3, 33]. Boiled controls under otherwise identical conditions did not exhibit any plateaus, demonstrating that the measured changes in thermophoretic amplitude resulted from myrtenal binding to hVN1R1. Future studies must be performed to determine the effect dimerization has on hVN1R1 function: because the monomeric form is present in all receptor samples even after SEC, it was not possible to obtain sufficient quantities of pure dimerized hVN1R1 for ligand binding analysis.



**Figure 2.5. Ligand Binding of Purified hVN1R1 Using Microscale Thermophoresis.** A) A sample of hVN1R1 was titrated against varying concentrations of myrtenal, and its thermophoretic mobility was measured. The plateaus at low and high myrtenal concentrations indicate that the receptor is binding its ligand with an  $\text{EC}_{50} \sim 1 \mu\text{M}$ . B) A boiled control did not exhibit the characteristic sigmoidal binding curve, indicating that the assay was indeed measuring binding between hVN1R1 and myrtenal. Since no binding was observed, the boiled control was not normalized to a fraction of bound receptor.

## 2.4 Discussion

This study shows that HEK293 cells can be used to generate milligram quantities of a human vomeronasal receptor. A pure and homogeneous protein sample was obtained through controllable, inducible expression. The purified receptors had alpha helical secondary structures, indicating that they were properly folded. They were also able to bind their reported ligand, and the measured  $\text{EC}_{50}$  value matched what has been previously reported.

A growing body of evidence suggests that HEK293 cells are a suitable platform for expressing GPCRs, and that FC14 is a universally effective detergent for solubilizing and stabilizing them. HEK293 cells have successfully been used in previously reported studies to produce rhodopsin and the olfactory receptor hOR17-4 [1-3]. All of the tested cultures yielded large quantities of receptors that could be easily purified using the rho1D4 epitope and antibody. This study further showed that removal of predicted glycosylation sites does not inhibit membrane localization or ligand binding. Although this study only used cells grown on tissue culture plates, previous reports have shown that HEK293 cells can be grown in suspension to yield even larger quantities of receptors [3]. Numerous studies have shown that the fos-choline series of detergents best solubilize GPCRs. In large screens, the fos-choline series consistently was equal to or outperformed other detergents for the GPCRs OR17-4, CCR3, CCR5, CXCR4, CX3CR1, hVN1R1, and hTAAR5 [3, 16, 30]. Although FC15 and FC16 were able to solubilize a greater quantity of expressed receptor than FC14, all three detergents were able to solubilize comparable amounts. Since FC14 has a CMC that is high enough to allow for detergent exchange, has been used to obtain high resolution protein structures [28, 29], and can be used for both the solubilization and purification processes, it appears to be an excellent detergent that may be broadly useful. However, both the HEK293 platform and the detergent must still be evaluated for every GPCR to verify their effectiveness.

#### **2.4.1 Existence of Functional Human Vomeronasal Receptors**

The results reported here suggest that humans have functional VNR genes, while previous reports suggest that human VN1Rs may function more like olfactory receptors.

This is the first report showing direct binding between a vomeronasal receptor and a volatile ligand. This study and previous reports show that the receptors are likely to have binding affinities in the  $\mu\text{M}$  range [14], that they can couple to the canonical OR signaling pathway via  $G_{\text{olf}}$  and  $G_{\text{cs}}$  and cAMP [14], and that hVN1R1 mRNA is expressed in the human MOE [15]. Indeed, mouse VN1Rs signal through TRP, which is a pseudogene in humans. If human VN1Rs are biologically functional, they must have a different signaling mechanism that still needs to be elucidated. The HEK293 cells used in this study provide an excellent mechanism to perform such studies. These cells have been successfully used to express and purify other GPCRs including the olfactory receptor OR17-4 [1-3]. They can thus be used as a platform to compare OR and VNR function and signaling. Additionally, although the human VNO may not be a functional organ, expression of hVN1R1 mRNA in other tissues in addition to the MOE raises the possibility that human vomeronasal receptors may have non-olfactory functions. Indeed, OR17-4 functions in both olfaction and sperm migration [34].

#### **2.4.2 Technological Importance of Membrane Proteins**

Membrane proteins are critical for communication between the cellular and external environments, and have potential uses in medicine and the design of biologically-based devices. They comprise nearly 25-30% of genes in sequenced genomes, and are involved in various processes ranging from chemical transport, signal transduction, and photosynthesis. GPCRs, the largest subclass of membrane proteins, are the target of nearly 50% of pharmaceutical drugs [35-38]. Olfactory-related proteins (ORs, VNRs, TAARs, and

FPRs) are the largest subclass of GPCRs, and could potentially be used to develop highly sensitive detectors.

Detailed structural knowledge of membrane proteins is therefore critical for developing new technologies and medicines. Yet, in spite of their enormous potential, very little is known about membrane protein structures. Of the >72,000 protein structures known, only 281 are unique membrane proteins, and only 6 are unique GPCRs as of April 2011. To facilitate protein structure determination, fast and reliable methods of protein expression and purification must be developed. In addition to hVN1R1, the methods used in this study have been successfully used with bovine rhodopsin and OR17-4 [1-3]. This study further showed that the expressed proteins are functional and targeted to the membrane even when potential glycosylation sites have been removed. This further underscores the potential for this system to be used with a wide variety of membrane proteins in preparation for crystallization trials. These efforts to purify hVN1R1 may not only help understand the role and function of the human vomeronasal system, but could also contribute toward other membrane protein studies and the development of biologically-inspired technologies like artificial noses.

## 2.5 References

1. Reeves PJ, Thurmond RL, Khorana HG. Structure and function in rhodopsin: high level expression of a synthetic bovine opsin gene and its mutants in stable mammalian cell lines. *Proc. Natl. Acad. Sci. USA* **93**, 11487-11492 (1996).
2. Cook B, Ernberg KE, Chung H, Zhang S. Study of a synthetic human olfactory receptor 17-4: Expression and purification from an inducible mammalian cell line. *PLoS One* **3**(8): e2920. (2008).
3. Cook B. *et al.* Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4. *Proc. Natl. Acad. Sci. USA* **106**(29), 11925-11930 (2009).
4. Trotier D. *et al.* The vomeronasal cavity in adult humans. *Chem. Senses* **25**, 369-380 (2000).
5. Bossy J. Development of olfactory and related structures in staged human embryos. *Anat. Embryol.* **161**, 225-236 (1980).
6. Boehm N, Roos J, Gasser B. Luteinizing hormone-releasing hormone (LHRH)-expressing cells in the nasal septum of human fetuses. *Dev. Brain Res.* **82**, 175-180 (1994).
7. Humphrey T. The development of the olfactory and the accessory olfactory formations in human embryos and fetuses. *J. Comp. Neurol.* **73**(3), 431-468 (1940).
8. Meisami E, and Bhatnagar KP. Structure and diversity in mammalian accessory olfactory bulb. *Microscopy Research and Technique* **43**, 476-499 (1998).
9. Boehm N, and Gasser B. Sensory receptor-like cells in the human foetal vomeronasal organ. *NeuroReport* **4**, 867-870 (1993).
10. Takami S. *et al.* Vomeronasal epithelial cells of the adult human express neuron-specific molecules. *Neuroreport* **4**, 375-378 (1993).
11. Grosser BI, Monti-Bloch L, Jennings-White C, Berliner DL. Behavioral and electrophysiological effects of androstadienone, a human pheromone. *Psychoneuroendocrino* **25**(3), 289-99 (2000).
12. Berliner DL, Monti-Bloch L, Jennings-White C, Diaz-Sanchez V. The functionality of the human vomeronasal organ (VNO): Evidence for steroid receptors. *J. Steroid Biochem. Molec. Biol.* **58**(3), 259-265 (1996).
13. Rodriguez I, and Mombaerts P. Novel human vomeronasal receptor-like genes reveal species-specific families. *Curr. Biol.* **12**, R409-411 (2002).

14. Shirokova E, Raguse JD, Meyerhof W, Krautwurst D. The human vomeronasal type-1 receptor family – detection of volatiles and cAMP signaling in HeLa/Olf cells. *FASEB J.* **22**, 1416-1425 (2008).
15. Rodriguez I, Greer CA, Mok MY, Mombaerts P. A putative pheromone receptor gene expressed in human olfactory mucosa. *Nat. Genet.* **26**, 18-19. (2000).
16. Ren H, Yu D, Ge B, Cook B, Xu Z, Zhang S. High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS One* **4**(2), e4509. doi: 10.1371/journal.pone.0004509 (2009).
17. Baaske P, Wienken CJ, Reineck P, Duhr S, & Braun D. Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed* **49**, 2238-2241 (2010).
18. Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Comm.* **1**, 100 (2010).
19. Gat U, Nekrasova E, Lancet D, Natochin M. Olfactory receptor proteins. Expression, characterization and partial purification. *Eur. J. Biochem.* **225**, 1157-1168 (1994).
20. Nekrasova E, Sosinskaya A, Natochim M, Lancet D, Gat U. Overexpression, solubilization and purification of rat and human olfactory receptors. *Eur. J. Biochem.* **238**, 28-37 (1996).
21. Katada S, Tanaka M, Touhara K. Structural determinants for membrane trafficking and G protein selectivity of a mouse olfactory receptor. *J. Neurochem.* **90**(6), 1453-1463 (2004).
22. Gimelbrant AA, Haley SL, McClintock TS. Olfactory receptor trafficking involves conserved regulatory steps. *J. Biol Chem.* **276**(10), 7285-7290 (2001).
23. Lu M, Staszewski L, Escheverri F, Xu H, Moyer, B. D. Endoplasmic reticulum degradation impedes olfactory G-protein coupled receptor functional expression. *BMC Cell Biol.* **5**:34 (2004).
24. Lu M, Echeverri F, Moyer BD. Endoplasmic reticulum retention, degradation, and aggregation of olfactory G-protein coupled receptors. *Traffic* **4**, 416-433 (2003).
25. Jayadev S, Smith RD, Jagadeesh G, Baukal AJ, Hunyady L. N-linked glycosylation is required for optimal AT1a angiotensin receptor expression in COS-7 cells. *Endocrinology* **140**, 2010-2017 (1999).
26. Gorzelle BM, et al. Reconstitutive refolding of diacylglycerol kinase, an integral membrane protein. *Biochemistry* **38**, 16373-16382 (2002).



27. Leck K-J, Zhang S, Hauser CAE. Study of bioengineered zebra fish olfactory receptor 131-2: receptor purification and secondary structure analysis. *PLoS ONE* **5**, e15027 (2010).
28. Bass RB, Strop P, Barclay M, Rees DC. Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science* **298**, 1582-1587 (2002).
29. Wang W, *et al.* The structure of an open form of an E. coli mechanosensitive channel at 3.45 Å resolution. *Science* **321**, 1179-1183 (2008).
30. Wang X, Corin K, Rich C, Zhang S. Study of two G-protein coupled receptor variants of human trace amine-associated receptor 5. *Scientific Reports* (submitted).
31. Liu X, Garriga P, Khorana HG. Structure and function in rhodopsin: Correct folding and misfolding in two point mutants in the intradiscal domain of rhodopsin identified in retinitis pigmentosa. *Proc. Natl. Acad. Sci. USA* **93**, 4554-4559 (1996).
32. Duhr S, & Braun D. Why molecules move along a temperature gradient. *Proc Natl Acad Sci USA* **103**, 19678-19682 (2006).
33. Kaiser L, *et al.* Large-scale production and study of a synthetic G protein-coupled receptor: human olfactory receptor 17-4. *Proc Natl Acad Sci USA* **105**, 15726-15731 (2008).
34. Spehr M, *et al.* Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* **299**, 2054-2058 (2003).
35. Stevens TJ and Arkin IT. Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins* **39**, 417-420 (2000).
36. Kumari T, Pant B, and Pardasani KR. A model for the evaluation of domain based classification of GPCR. *Bioinformation* **4**(4), 138-142 (2009).
37. Klabunde T, and Hessler G. Drug design strategies for targeting G-protein-coupled receptors. *ChemBioChem* **3**, 928-944 (2002).
38. Lundstrom K. Structural biology of G protein-coupled receptors. *Bioorg Med Chem Lett* **15**, 3654-3657 (2005).

# CHAPTER 3

## INVESTIGATION OF T4 LYSOZYME AS A STRATEGY TOWARD GPCR CRYSTALLIZATION

### 3.1 Introduction

Once sufficient quantities of membrane proteins have been produced and purified to allow for crystallization screens, and a suitable detergent has been found, other obstacles to crystallization must be overcome. Three such difficulties are primarily encountered. First, the detergent used to solubilize the protein forms a protective belt, severely decreasing the number of potential protein-protein contacts available to form a crystal. Second, large regions of the protein may be disordered as a result of its secondary structure or post-translational modifications. This heterogeneity of structure prevents a regular lattice from forming. Third, the protein may not remain stable for the amount of time necessary for crystal formation to occur.

Several novel strategies have been developed to facilitate membrane protein crystal growth. To increase the surface area available to form a crystal lattice, T4-lysozyme (T4L) fusions have been synthesized [1-4], and antibody fragments against specific portions of the membrane protein have been used [5]. These antibody or T4L fragments are soluble proteins that effectively increase the solvent-exposed receptor area, thereby facilitating

protein-protein contacts needed for crystal formation. To increase the structural homogeneity of a protein sample, loops and other large protein segments without a defined and stable secondary structure have been deleted, and post-translational modifications like glycosylation have been removed [1-3, 5, 6]. To improve protein stability, sequence mutations have been introduced [2, 4, 6].

Current strategies are rationally designed, but must be empirically tested with each membrane protein. The most beneficial strategy in membrane protein structural research would be one that could be used with many proteins. Insertion of T4L in the third intracellular loop seems to be a promising technique for GPCRs, as four of the six crystallized GPCRs used this approach [1-4]. Generation of DNA templates is fast and easy using current cloning techniques. Unlike the development of antibody fragments, it isn't necessary to spend months to find a specific variant for the protein of interest. Instead, many DNA templates can easily be generated in parallel. This method also allows the entire protein or the majority of it to be crystallized. However, insertion of T4L near the putative G-protein binding domain can disrupt receptor function. It can also potentially interfere with proper folding, limiting the useful information that could be obtained from the crystal structure. It is thus necessary to evaluate how a T4L insert can affect the function and structure of various GPCRs.

This study examines the ability of T4L to be used as a general insert in olfactory-related GPCRs. Native and T4L-mutants of hOR17-4 and hVN1R1 were expressed in HEK293 cells. The structure and function of the T4L variants were compared to the native forms. Immunocytochemistry showed that both protein forms were localized to the cell surface, and calcium influx assays indicated that they could initiate signaling. Circular

dichroism and microscale thermophoresis analyses of purified receptors showed that the native and T4L-mutants had alpha-helical structures and bound their ligands.

Interestingly, the T4L-mutants yielded higher percentages of soluble monomers compared to aggregates, indicating that the T4L insert stabilized the protein structure. The mutants also had lower ligand-binding affinities, but yielded more consistent results over time with less noise, further suggesting higher receptor stability. These results indicate that a T4L insertion may be a viable option when working with GPCRs with the goal of obtaining a crystal.

## **3.2 Methods**

### **3.2.1 hOR17-4T4L and hVN1R1-T4L Gene Design and Construction**

The protein sequences for hOR17-4 (NCBI Accession #NP002539) and hVN1R1 (AAG10698) were obtained from GenBank. To enable expression and purification from mammalian cells, the following modifications were made to both genes: 1) addition of a C-terminal rho tag (TETSQVAPA) preceded by a two glycine linker; 2) human codon optimization; 3) addition of a Kozak sequence 5' to the start codon; 4) addition of a 5' EcoRI site and a 3' XhoI site to facilitate subcloning into expression vectors. 5) addition of an N-terminal strep tag (ASWSHPQFEK) followed by a GSSG linker for further purification; 6) insertion of T4 Lysozyme residues 2-161 in the predicted third intracellular loop after residue S232 in hOR17-4 and after residue L261 in hVN1R1; and 7) N5Q and N195Q mutations in hOR17-4, and N117Q, N151Q, and N183Q, N198Q, and N256Q mutations in hVN1R1 to facilitate crystallization. The genes were constructed by Geneart and ligated into the pcDNA 4/To vector (Invitrogen, Carlsbad, CA). The plasmid was amplified in

subcloning efficiency DH5 $\alpha$  *E. coli* (Invitrogen) and purified using MiniPrep or MaxiPrep kits (Quiagen, Valencia, CA). The transmembrane and loop domains were predicted using the TMHMM Server v 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The DNA sequences for hOR17-4 [8] and hVN1R1 (Chapter 2) were synthesized as previously described.

### **3.2.2 Expression, Purification, and Structural and Functional Analyses of T4L**

#### **Mutants**

The generation of stable, inducible cell lines capable of expressing the desired proteins is described in Chapter 2. Methods for purifying the receptors and analyzing their structure and function are also described.

### **3.2.3 Calcium Influx Imaging**

HEK293 cells expressing hOR17-4 or hOR17-4T4L were seeded on 0.18 mm thick cover-glass slides at a density of  $10^5$  cells/ml. OR expression was induced with 1  $\mu$ g/ml of tetracycline for 48 h. To visualize calcium signaling, cells were loaded with 10  $\mu$ M Fura-Red-AM (Invitrogen) for 30 min in serum-free DMEM/F12 medium and washed in PBS. The cells were then incubated an additional 30 min in DMEM/F12 supplemented with 10% FCS to allow intracellular Fura-Red-AM to completely hydrolyze. Calcium influx in response to 50  $\mu$ M of the odorant bourgeonal was visualized using confocal fluorescence microscopy (Zeiss LSM 510) with a water immersion objective (Zeiss Achroplan 63 x NA 1.2). The cells were excited at 488 nm (Ar<sup>+</sup> laser), and fluorescence emission of Fura-Red-AM was monitored at 650 nm using a long pass emission filter. Images were collected

every 2 seconds for a total of 100 seconds. ATP was used as a control. Unless noted otherwise, cells were incubated at 37°C and 5% CO<sub>2</sub> in DMEM (Dulbecco's modified Eagle's medium, Invitrogen, San Diego, CA) supplemented with 10% FCS (Invitrogen). Calcium imaging of hVN1R1 and hVN1R1-T4L cells was not performed due to the higher toxicity of expressing the proteins, and because it is not known whether the receptors are capable of coupling to the endogenous HEK293 G-proteins [7]. Any results obtained with these cell lines would thus be ambiguous. These experiments were performed with the help of Dr. Horst Pick from the Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL).

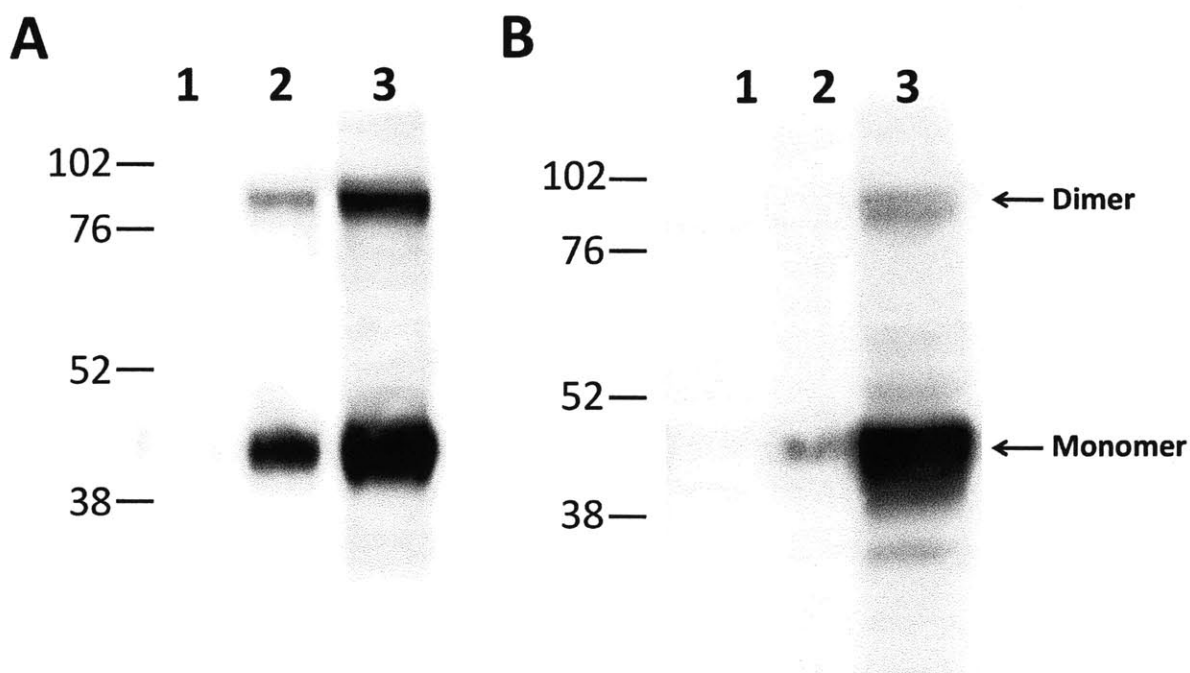
## **3.3 Results**

### **3.3.1 Induction of hOR17-4T4L and hVN1R1-T4L Expression in Stable HEK293 Cell Lines**

The native forms of hOR17-4 and hVN1R1 were previously expressed [8, 9, Chapter 2]. The T-Rex system was used to make the T4L variants in the same manner as the native receptors. Final induction conditions for hOR17-4T4L were 1µg/ml of tetracycline with 5mM sodium butyrate, and for hVN1R1 were 1µg/ml of tetracycline with 1mM sodium butyrate. All inductions were performed for 48 hours prior to harvesting the cells for receptor purification.

The expressed protein was analyzed using SDS-PAGE and Western blotting. Both T4L clones had two bands, corresponding to the monomeric and dimeric forms of the receptor (**Figure 3.1**). The presence of this characteristic size pattern indicates that the

T4L insert does not significantly alter receptor folding or function, as both GPCRs are capable of forming dimers.



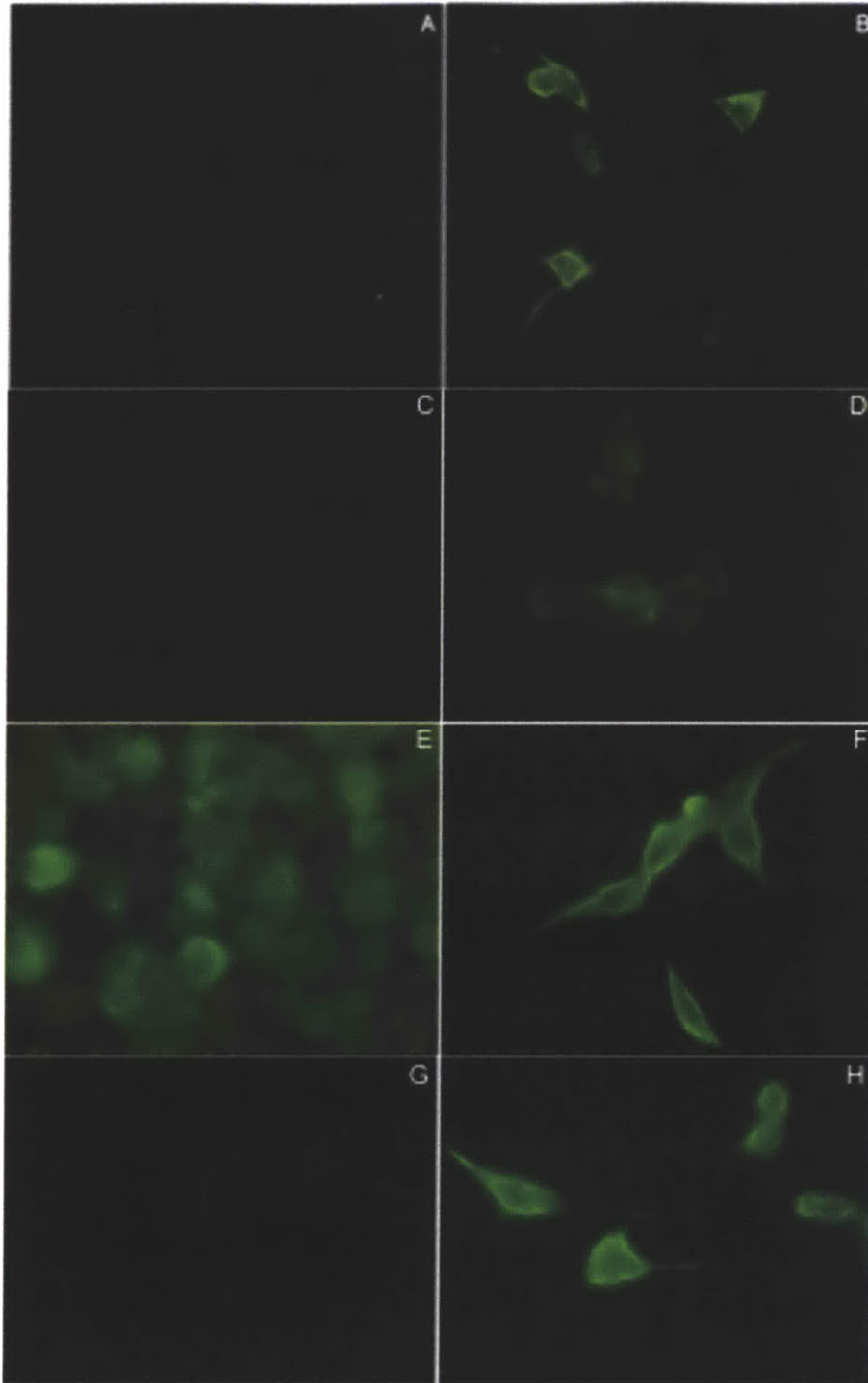
**Figure 3.1. Expression of hOR17-4T4L and hVN1R1-T4L in HEK293 Cells.** A) hOR17-4T4L, and B) hVN1R1-T4L. Lane 1 shows receptor solubilized from uninduced cells, lane 2 shows receptor solubilized from cells induced with 1µg/ml of tetracycline for 48 hours, and lane 3 shows receptor solubilized from cells induced with 1µg/ml of tetracycline and 5mM sodium butyrate for 48 hours. No protein was detected in uninduced cells. Induction with tetracycline and sodium butyrate resulted in maximum expression. Both receptors had monomeric and dimeric forms. Subsequent experiments determined that the optimal sodium butyrate concentration was 5mM for hOR17-4T4L, and 1mM for hVN1R1-T4L.

### 3.3.2 Immunohistochemical Staining of Induced Cells

GPCRs often become trapped in the cell when expressed in heterologous systems. This is particularly true when post-translational modifications like glycosylation are removed, suggesting that the conserved N-terminal glycosylation site of ORs may be necessary for appropriate localization [10-16]. To determine whether the stably expressed hOR17-4T4L or hVN1R1-T4L receptors were trafficked to the cell membrane, induced and

non-induced cells were stained with the rho1D4 antibody (**Figure 3.2**). Non-induced cells yielded no signal, while induced cells showed receptors that were localized to the cell membrane. Permeabilized and non-permeabilized cells both demonstrated membrane localization, suggesting that not all receptors are inserted in the correct orientation. However, visualization of both native and T4L-mutant proteins in the membrane indicates that the T4L insert does not affect GPCR trafficking. The results also indicate that the glycosylation sites are not necessary for localization of the protein to the membrane.

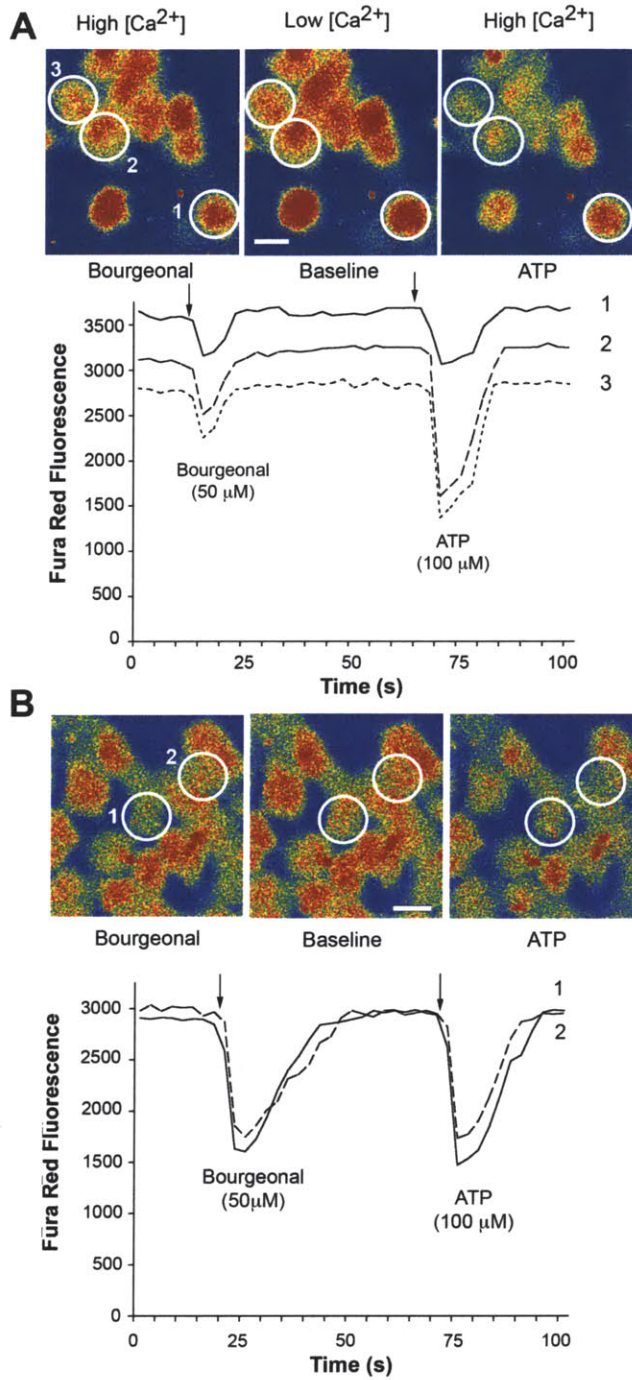




**Figure 3.2. Immunohistochemical Staining of Native and T4L Clones.** A) Uninduced hVN1R1; B) Induced hVN1R1; C) Uninduced hVN1R1-T4L; D) Induced hVN1R1-T4L; E) Uninduced hOR17-4; F) Induced hOR17-4; G) Uninduced hOR17-4T4L; H) Induced hOR17-4-T4L. The noninduced cells express undetectable amounts of receptor. The expressed receptors are localized to the cell membranes. Panel E was overexposed to visualize the cells; some leaky expression can be seen in a few cell membranes, but otherwise no surface-localization can be seen.

### 3.3.3 Functional Characterization of hOR17-4T4L in HEK293 Cells

The functional activity of hOR17-4T4L was probed by measuring the influx of calcium ions in response to the specific odorant bourgeonal [17, 18]. In HEK293 cells, hOR17-4 can couple to the promiscuous G protein  $G_{\alpha q}$  to initiate a signal through the inositol triphosphate (IP<sub>3</sub>) pathway [8, 9]. Similar results were observed with hOR17-4T4L: application of 50 $\mu$ M bourgeonal resulted in a transient increase in intracellular Ca<sup>+2</sup> concentration (**Figure 3.3**). However, this Ca<sup>+2</sup> response was 60-70% lower than the native protein. Also, the time needed to return to baseline Ca<sup>+2</sup> levels after induction was shorter (5-10 seconds for hOR17-4T4L, and 25 seconds for hOR17-4). These results show that hOR17-4T4L is capable of signal transduction, although perhaps in a reduced capacity. The receptors hVN1R1 and hVN1R1-T4L were not tested due to the higher levels of cell toxicity.



**Figure 3.3.  $Ca^{2+}$  Influx Profiles for A) hOR17-4T4L and B) hOR17-4 Expressed in Stable, Inducible HEK293 Cells.** Responses from three individual cells (nos. 1, 2, 3) are shown for hOR17-4T4L, and two individual cells (nos. 1, 2) for hOR17-4. The intracellular  $Ca^{2+}$  concentration was monitored with confocal microscopy using Fura-Red as an indicator. The intensities are displayed in pseudo color (A and B, top). The cytosolic  $Ca^{2+}$  concentration is recorded as a function of time (A and B, bottom). The arrows indicate the time at which the odorant bourgeonal was applied. 100 $\mu M$  ATP was used as a control. The profiles show that both proteins initiate signal transduction, though the response from hOR17-4T4L is lower in amplitude and shorter in duration. The size bar is 10 $\mu m$ .

### 3.3.4 Systematic Detergent Screening for Receptor Solubilization

An appropriate detergent must be used to solubilize receptors from cell membranes, and keep them stable and functional in solution. A confounding factor is that different detergents may need to be used even for similar or related proteins, and must therefore be empirically determined. Thus, although optimal detergents for hOR17-4 and hVN1R1 have been experimentally determined [9, Chapter 2], new detergent screens were performed for hOR17-4T4L and hVN1R1-T4L.

The optimal detergents for hVN1R1-T4L and hOR17-4T4L were again the fos-choline series (**Table 3.2**). The receptors hVN1R1, hVN1R1-T4L, and hOR17-4 had nearly identical detergent profiles, which were also similar to the profiles reported for several other GPCRs [9, 19]. In contrast, hOR17-4T4L exhibited a slightly different profile. For hVN1R1, hVN1R1-T4L, and hOR17-4, receptor solubility was vastly improved in the fos-choline detergents compared to most of the other detergents. Only a handful of other detergents were able to solubilize comparable amounts of protein. While receptor solubility for hOR17-4T4L was higher in the fos-cholines than in any other detergent class, many other detergents were able to solubilize sufficient quantities of receptor. The cyclofos and anzerger families yielded only slightly less soluble protein than the fos-cholines, as well as Fos-MEA-10 and FOSFEN-9. This similarity is not surprising, as these detergents have similar structures (**Table 3.1**). The only difference between the fos-cholines and cyclofos detergents is a cyclohexane ring at the end of the cyclofos carbon chain. Similarly, FOSFEN detergents have a phenyl ring attached to the carbon chain. Fos-MEA detergents are similar to the fos-cholines, except that 2 methyl groups on the choline nitrogen have been replaced with hydrogens. The anzerger exhibit slightly greater differences: they

have a sulfate group instead of a phosphate group, and the order of the negatively and positively charged groups is reversed. All of these detergents are structurally or chemically similar to phosphatidylcholines, phospholipids that are major constituents of cell membranes. Thus, although hOR17-4T4L has a different detergent profile than the other tested GPCRs, these results reinforce increasing evidence that detergents that are structurally related to biological phospholipids are optimal detergents for GPCRs. It should be noted that several other detergents, including the maltosides, were able to solubilize noticeable amounts of protein. However, they were not as effective as the fos-cholines or related detergents.

**Table 3.1: Detergent Structures**

Detergent Family	General Structure [20]
Fos-Choline	$\text{CH}_3-(\text{CH}_2)_n-\text{O}-\text{P}(=\text{O})(\text{O}^-)-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}^+(\text{CH}_3)_3$
Cyclofos	$\text{Cyclohexane ring}-(\text{CH}_2)_n-\text{O}-\text{P}(=\text{O})(\text{O}^-)-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}^+(\text{CH}_3)_3$
FOSFEN	$\text{CH}_3(\text{CH}_2)_n-\text{C}_6\text{H}_4-\text{O}-\text{P}(=\text{O})(\text{O}^-)-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}^+(\text{CH}_3)_3$
Fos-MEA	$\text{CH}_3-(\text{CH}_2)_n-\text{O}-\text{P}(=\text{O})(\text{O}^-)-\text{O}-\text{CH}_2-\text{CH}_2-\text{N}(\text{H})\text{CH}_3$
Anzergent	$\text{CH}_3(\text{CH}_2)_n-\text{N}^+(\text{CH}_3)_2(\text{CH}_2)_3-\text{S}(=\text{O})_2\text{O}^-$

**Table 3.2: Detergent Solubilization of hVN1R1, hVN1R1-T4L, hOR17-4, and hOR17-4T4L\***

#	Detergent	Type <sup>a</sup> CMC(%) <sup>b,c</sup>	Spot Intensity Average			
			hVN1R1	hVN1R1-T4L	hOR17-4	hOR17-4T4L
1	ANAGE <sup>d</sup> -7	N 0.65				
2	ANAFOS <sup>d</sup> -20	N 0.0072				
3	ANAFOS <sup>d</sup> -35	N 0.001				
4	ANAFOS <sup>d</sup> -38	N 0.00045				
5	ANAFOS <sup>d</sup> -50	N 0.0016				
6	ANAFOS <sup>d</sup> -CL059	N 0.025				
7	ANAFOS <sup>d</sup> -CL059	N 0.053				
8	ANAFOS <sup>d</sup> -CL258	N 0.0048				
9	ANAFOS <sup>d</sup> -CL259	N 0.003				
10	ANAFOS <sup>d</sup> -CL2510	N 0.2				
11	ANAFOS <sup>d</sup> -CL358	N 0.0055				
12	ANAFOS <sup>d</sup> -X-100	N 0.015				
13	ANAFOS <sup>d</sup> -X-114	N 0.011				
14	ANAFOS <sup>d</sup> -X-305	N -				
15	ANAFOS <sup>d</sup> -X-405	N 0.16				
16	AKZERGBIT <sup>d</sup> 3-10	Z 1.2				
17	AKZERGBIT <sup>d</sup> 3-12	Z 0.094				
18	AKZERGBIT <sup>d</sup> 3-14	Z 0.097				
19	Big CHAP	N 0.25				
20	Big CHAP, deoxy	N 0.12				
21	CHAPS	Z 0.49				
22	CHAPSO	Z 0.5				
23	Deoxycholic acid, sodium salt	A 0.24				
24	Sodium cholate	A 0.41				
25	C-DODECAPOS <sup>g</sup>	Z 0.77				
26	CYCLOPOS <sup>g</sup> -4	Z 0.45				
27	CYCLOPOS <sup>g</sup> -5	Z 0.15				
28	CYCLOPOS <sup>g</sup> -6	Z 0.094				
29	CYCLOPOS <sup>g</sup> -7	Z 0.022				
30	CYGLI <sup>d</sup> -3	N 0.86				
31	CYHAL <sup>d</sup> -4	N 0.37				
32	CYHAL <sup>d</sup> -5	N 0.12				
33	CYHAL <sup>d</sup> -6	N 0.028				
34	CYHAL <sup>d</sup> -7	N 0.0059				
35	2,6-Dimethyl-4-heptyl-β-D-maltopyranoside	N 1.2				
36	2-Propyl-1-pentyl maltopyranoside	N 1.9				
37	POS-CHOLDS <sup>d</sup> -4	Z 1.2				
38	POS-CHOLDS <sup>d</sup> -4.0	Z 0.35				
39	POS-CHOLDS <sup>d</sup> -4.1	Z 0.052				
40	POS-CHOLDS <sup>d</sup> -4.2	Z 0.047				
41	POS-CHOLDS <sup>d</sup> -4.3	Z 0.027				
42	POS-CHOLDS <sup>d</sup> -4.4	Z 0.0046				
43	POS-CHOLDS <sup>d</sup> -4.5	Z 0.0027				
44	POS-CHOLDS <sup>d</sup> -4.6	Z 0.00053				
45	POS-CHOLDS <sup>d</sup> -CS0-9	Z 0.99				
46	POS-CHOLDS <sup>d</sup> -CS0-11	Z 0.9				
47	POS-CHOLDS <sup>d</sup> -CS0-11-RU	Z 0.07				
48	POS-CHOLDS <sup>d</sup> -UNSAT-11-10	Z 0.21				
49	POS-HSA <sup>d</sup> -8	A 0.39				
50	POS-HSA <sup>d</sup> -10	A 0.15				
51	POS-FEN <sup>d</sup> -9	Z 0.044				
52	HEGA-8	N 2.5				
53	NOFOL-FOB <sup>g</sup>	Z 1.4				
54	PHAL <sup>g</sup> -CS	Z -				
55	PHAL <sup>g</sup> -C10	Z -				
56	Decyltrimethylammonium chloride	C 0.07				
57	Dodecyltrimethylammonium chloride	C 0.0012				
58	Hexadecyltrimethylammonium chloride	C 0.000102				
59	Tetradecyltrimethylammonium chloride	C 0.0009				
60	n-Octyl-β-D-glucopyranoside	N 0.53				
61	n-Nonyl-β-D-glucopyranoside	N 0.2				
62	n-Octyl-β-D-maltopyranoside	N 0.89				
63	n-Nonyl-β-D-maltopyranoside	N 0.28				
64	n-Decyl-β-D-maltopyranoside	N -				
65	n-Dodecyl-β-D-maltopyranoside	N 0.087				
66	n-Uldecyl-β-D-maltopyranoside	N 0.029				
67	n-Uldecyl-β-D-maltopyranoside	N 0.029				
68	n-Dodecyl-β-D-maltopyranoside	N 0.0076				
69	n-Dodecyl-β-D-maltopyranoside	N 0.0087				
70	n-Tridecyl-β-D-maltopyranoside	N 0.0817				
71	n-Heptyl-β-D-thiogalactopyranoside	N 0.85				
72	n-Octyl-β-D-thiomaltopyranoside	N 0.4				
73	n-Nonyl-β-D-thiomaltopyranoside	N 0.15				
74	n-Decyl-β-D-thiomaltopyranoside	N 0.045				
75	n-Uldecyl-β-D-thiomaltopyranoside	N 0.011				
76	n-Dodecyl-β-D-thiomaltopyranoside	N 0.0026				
77	n-Decyl-N,N-dimethylglycine	Z 0.46				
78	n-Dodecyl-N,N-dimethylglycine	Z 0.041				
79	n-Dodecyl-β-iminodipropionic acid, monosodium salt	Z -				
80	Sodium dodecanyl sarcosine	A 0.42				
81	Hexaethylene glycol monooleyl ether (C8E6)	N 0.39				
82	Octaethylene glycol monododecyl ether (C12E8)	N 0.0048				
83	Pentaethylene glycol monododecyl ether (C12E5)	N 0.031				
84	Tetraethylene glycol monooleyl ether (C8E4)	N 0.25				
85	Sucrose monododecanosate	N 0.016				
86	Dimethyldiethylphosphine oxide	N 0.1				
87	n-Tetradecyl-N,N-dimethylamine-N-oxide (TDAO)	Z 0.0075				
88	n-Dodecyl-N,N-dimethylamine-N-oxide (DDAO)	Z 0.023				
89	HEGA-10	N 0.26				
90	NP-40	N 0.05-0.3				
91	Diglycine	N -				
92	Diglycine (8%)+Cholate (2%)	N -				
93	CHAPS (8%)+CHS (1.2%)+DDM (10%)	N -				
94	CHAPS (10%)+CHS (2%)+DDM (2%)	N -				
95	CHAPS (10%)+CHS (2%)	N -				
96	CHAPS (10%)+CG (1%)	N -				

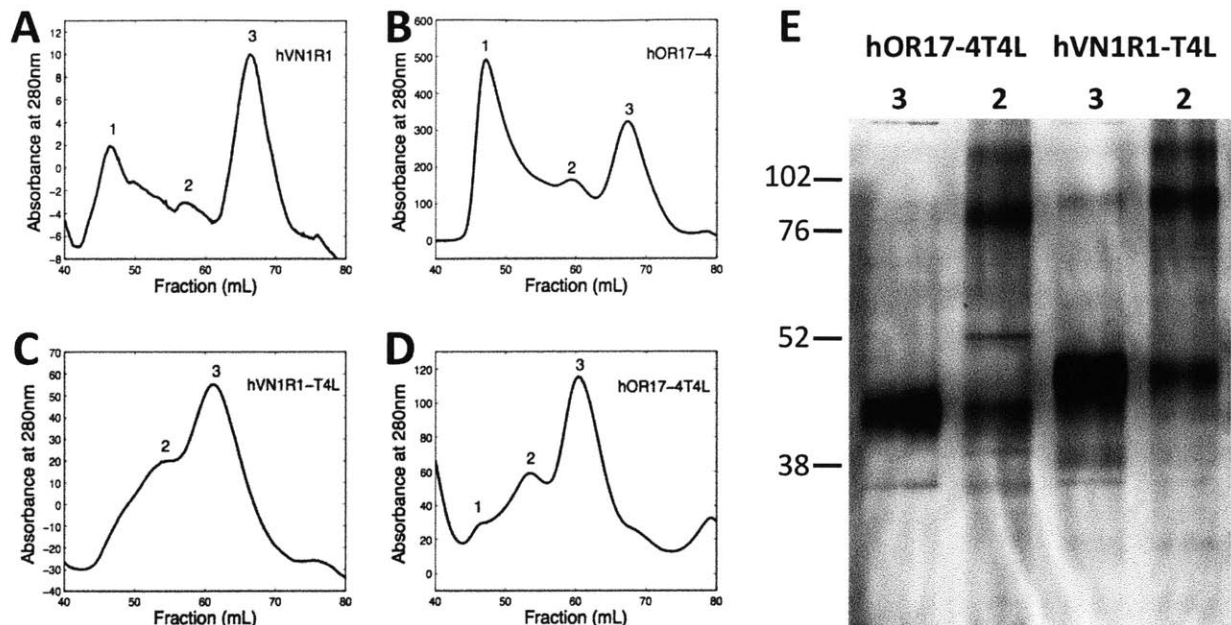
<sup>a</sup>Detergent type, A (anionic detergent), C (cationic detergent), N (non-ionic detergent) and Z (zwitterionic detergent). <sup>b</sup>CMC (critical micelle concentration) in H<sub>2</sub>O from the Detergent Properties List (<http://www.amerchem.com>). <sup>c</sup>Non-ionic detergents used in the experiment. <sup>d</sup>CMC (critical micelle concentration) in H<sub>2</sub>O from the Detergent Properties List (<http://www.amerchem.com>).

\*Detergent Screening of hOR17-4 has been reported in [9].

### 3.3.5 Purification of hOR17-4T4L and hVN1R1-T4L

A two-step purification procedure using immunoaffinity chromatography and size-exclusion chromatography (SEC) was used to purify the expressed receptors [8, 9, Chapter 2].

While UV absorption at 280 nm typically showed three distinct peaks for the native proteins, it usually showed only two for the T4L mutants. Western blot analysis demonstrated that the three native peaks corresponded to aggregated, dimerized, and monomeric forms of the protein [8, 9, Chapter 2]. Similar analyses of the T4L-mutants showed that the peaks corresponded to dimerized and monomeric protein forms (**Figure 3.4**). There was little or no evidence of aggregated receptor. The UV readings also indicated that a higher percentage of the total protein obtained was soluble monomer. These results suggest that the T4L insert may help stabilize the receptor in a soluble state, and that a higher proportion of protein purified from a T4L-mutant batch will be usable for crystallization screens. Both native and T4L-mutant proteins were able to yield about 1mg of total protein per gram of cells. However, depending on the purification batch, at least 50% and up to 60-70% of the protein recovered after SEC was soluble monomer for the T4L clones. In contrast, ~10-40% of the protein recovered after SEC was soluble monomer for the native clones.



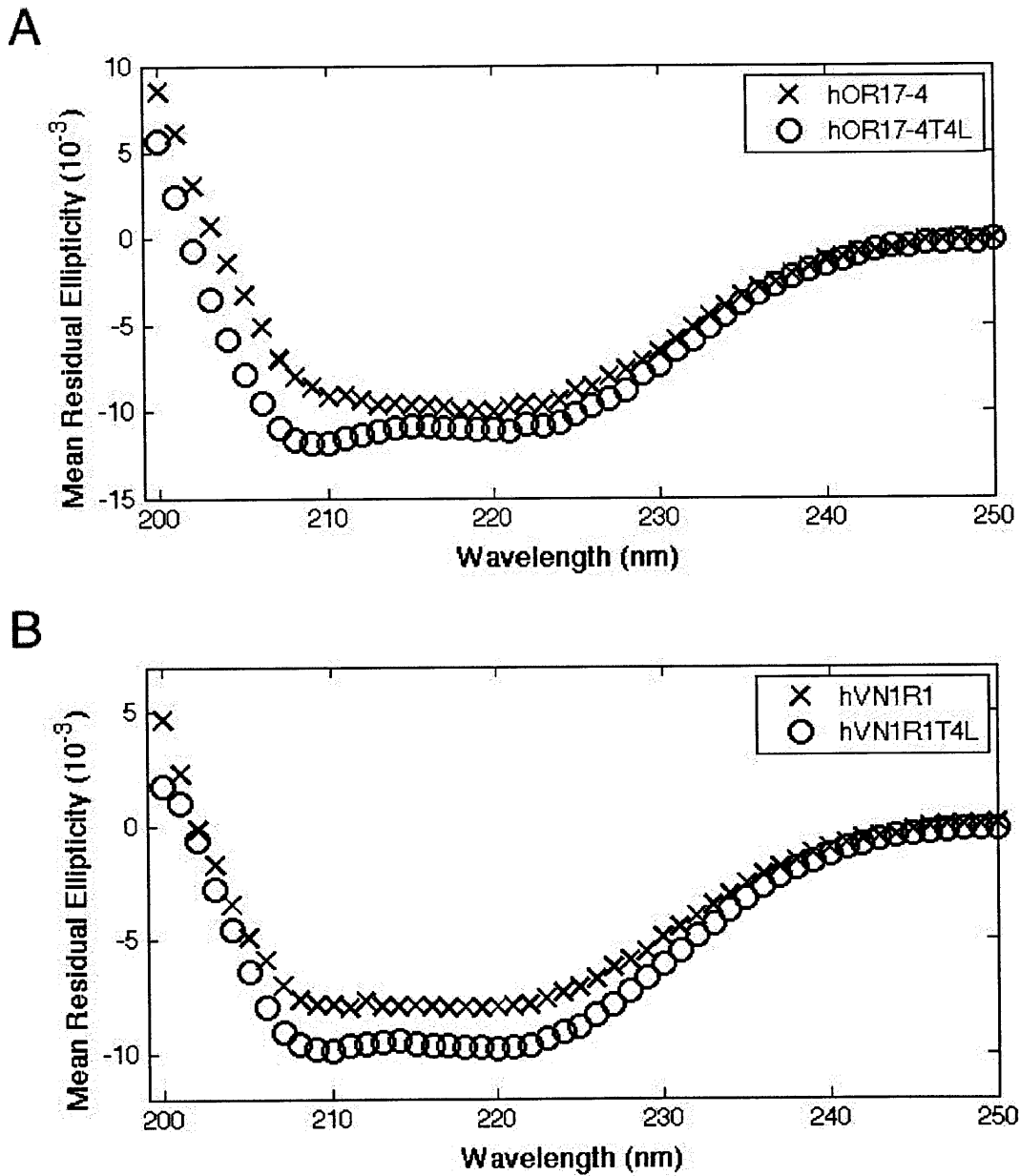
**Figure 3.4. 280 nm UV Traces Through a Size Exclusion Column and a Silver Stain of the T4L Fractions.** A) Typical 280 nm UV trace of hVN1R1, B) Typical 280 nm UV trace of hOR17-4, C) Typical 280 nm UV trace of hVN1R1-T4L, D) Typical 280 nm UV trace of hOR17-4T4L, and E) Silver stain of the purified samples in C) and D). Peak 1 from D) is not shown, but contained primarily aggregated protein. The native proteins yield more aggregated receptors. The T4L mutants are primarily a monomeric population. This suggests that the T4L insert prevents aggregation by keeping the receptors soluble in the monomeric or, secondarily, dimerized state. Number code: 1) Aggregate; 2) Dimer; 3) Monomer. Blots of peak fractions in A) and B) have been reported in [9] and Chapter 2.

### 3.3.6 Structural Characterization of Purified hOR17-4T4L and hVN1R1-T4L

The secondary structure of the T4L mutants was assayed using circular dichroism (CD). Far UV spectra of both proteins have characteristic alpha-helical shapes, similar to the native proteins (**Figure 3.5**). However, the peaks at 208 and 220 nm are more defined in the T4L mutants. A CD spectrum is assumed to be a linear composition of the individual spectra of each secondary structure present in the tested protein. A protein that is purely alpha-helical in shape will have sharply defined peaks at 208 and 220 nm, while a protein that has a higher percentage of loops and random coils will have peaks that are less



pronounced [21]. These results thus suggest that the T4L insert may help to structurally stabilize the proteins.

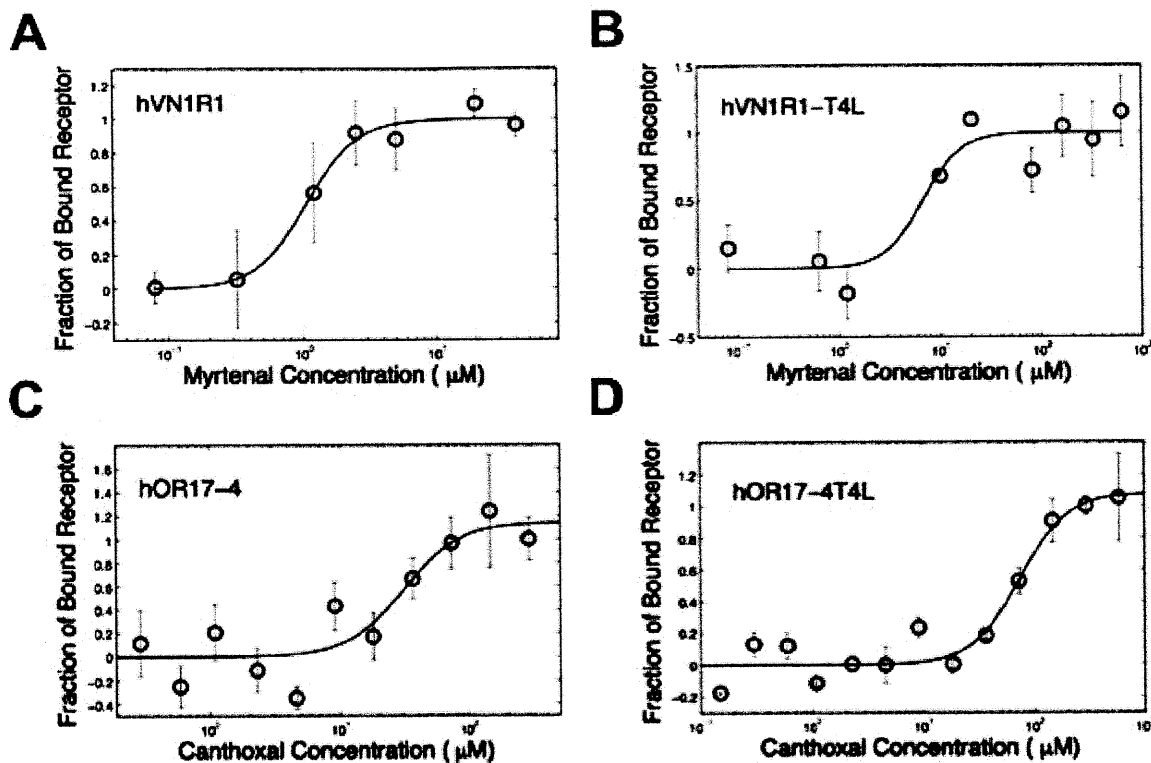


**Figure 3.5. Circular Dichroism Spectra of Native and T4L GPCRs.** A) hOR17-4 and hOR17-4T4L. B) hVN1R1 and hVN1R1T4L. The native and T4L variants have alpha-helical secondary structures. However, the peaks are more pronounced in the T4L mutants, suggesting that the structure is more stabilized.

### 3.3.7 Ligand Binding Analysis of Purified hOR17-4T4L and hVN1R1-T4L

Microscale thermophoresis (MST) was used to determine whether the native proteins and T4L mutants could bind their ligands. Microscale thermophoresis is a technique that induces a local spatial temperature gradient. Proteins in the gradient diffuse along it. Free receptors have a different diffusion pattern compared to receptors that are bound to their ligands, likely due to changes in their hydration shell [22, 23]. MST can measure binding of ligands as small as individual ions [24], which is critical because odorants are less than 300Da while their receptors are larger than 30kDa. MST can also be used on receptors in solution, and can detect fluorescence from native tryptophans, thereby eliminating the need to use surface or coupling chemistries, which can alter the structure or function of the protein being studied.

The native and mutant proteins were titrated against their known ligands. For both proteins, the T4L mutant had a lower ligand-binding affinity than the native protein, but was still able to bind the ligand (**Figure 3.6**). Interestingly, the binding analyses indicated that the T4L insert stabilized the receptor, even though it also probably disrupted the ability of the receptor to bind its ligand. Measurements taken on the same sample of hOR17-4T4L over two months apart yielded the same results, while similar measurements of hOR17-4 were noisier and less consistent.



**Figure 3.6. Ligand Binding Measurements of Native and T4L GPCRs.** A) hVN1R1 binding to myrtenal with an  $EC_{50} \sim 1\mu\text{M}$ . B) hVN1R1-T4L binding to myrtenal with an  $EC_{50} \sim 9\mu\text{M}$ . C) hOR17-4 binding to canthoxal with an  $EC_{50} \sim 30\mu\text{M}$ . D) hOR17-4T4L binding to canthoxal with an  $EC_{50} \sim 70\mu\text{M}$ . The native and T4L mutants were able to bind their ligands, but the T4L variants had lower affinities. Boiled controls did not bind their ligands, and had behavior similar to that observed in Chapters 2, 4, and 5. For clarity, they are not shown. All curves are normalized to the fraction of bound receptor, and are fit to the Hill equation with a coefficient of 2.

### 3.4 Discussion

In spite of their importance, knowledge about the structure and function of membrane proteins in general, and specifically GPCRs, lags far behind that of soluble proteins. This is mainly due to the difficulty of expressing and solubilizing functional receptors, as well as the difficulty of crystallizing any proteins that are produced. Structural knowledge of a protein is necessary in order to fully understand its function. It

is especially important with respect to GPCRs, as a determination of their structure could enable the design of specific drugs or biologically-inspired technologies.

A significant challenge in the field of membrane protein crystallography lies in finding strategies that can be applied to multiple proteins. Indeed, crystallization conditions for even similar or related proteins are unique and must be empirically determined [25]. Insertion of T4 Lysozyme in the third intracellular loop of GPCRs seems to be a promising approach, as 4 of the 6 structures to date were obtained using this strategy. However, to be truly useful, this approach should not only facilitate crystallization, it should also not interfere with the normal structure and function of the protein. This study has examined how addition of the T4L sequence in the third intracellular loop affects expression, solubilization, purification, folding, and function of two olfactory-related GPCRs.

The detergent screens indicated that insertion of T4L can alter the optimal detergent for the protein. The receptors hVN1R1 and hVN1R1-T4L had similar patterns of solubility in the tested detergents, while the solubility pattern of hOR17-4 and hOR17-4T4L was more variable. However, in both cases, the optimal detergents for both native and mutant proteins belonged to the same general class, and had similar structures or chemical properties.

Results obtained during the purification process indicate that insertion of T4L may help increase the yield of soluble protein for crystallization screens. Size-exclusion chromatography showed that ~10-40% of the recovered native protein was soluble monomer. In contrast, at least 50%, and up to 70%, of the recovered mutant protein was soluble monomer. This is an important finding, because milligram quantities of

homogeneous protein are needed for crystallization trials and other structural studies. Aggregation or impurities are common, and severely limit the amount of usable protein typically obtained after a batch purification. Indeed, up to 90% of the receptor obtained from the native proteins was aggregated or impure, and could not be used for subsequent experiments. Addition of T4L increased the yield of usable receptor, likely due to its solubility, which increased the solubility of the fusion protein. In addition, the CD and MST measurements indicated that the T4L insert may help stabilize the protein in which it is inserted; both T4L mutants had more defined CD spectra, and hOR17-4T4L was able to bind its ligand several months after it was purified. This is particularly important, as several months may be needed for crystal growth.

Immunohistochemical data, calcium influx assays, circular dichroism, and microscale thermophoresis indicate that insertion of T4L in the third intracellular loop does not completely disrupt protein structure or function. Both T4L mutants trafficked to the cell membrane. Because membrane localization can be impaired for improperly folded or glycosylated GPCRs, this suggests that the T4L insertion does not adversely affect receptor structure. Circular dichroism showed that the purified proteins had alpha-helical conformations, indicating that they were properly folded. Indeed, the T4L mutants had more defined peaks, suggesting that they might be more stable. Calcium influx assays in HEK293 cells demonstrated that signaling still occurred with the hOR17-4T4L mutant, although it was more limited. Microscale thermophoresis measurements of purified receptors showed that the T4L mutants had higher  $EC_{50}$  values, but were still able to bind their ligands. The latter two results indicate that the insert may interfere with G-protein interactions, as well as with ligand binding. Since GPCRs are known to have many flexible

conformations, it's possible that the T4L insertion may stabilize a particular conformation, making binding of certain ligands more difficult. Thus, although lower binding affinities were measured with canthoxal (hOR17-4) and myrtenal (hVN1R1), it is possible that interactions with other ligands would be less affected, or might even have higher affinities. Future experiments will be carried out to test this hypothesis.

Structural knowledge of GPCRs and other membrane proteins is a prerequisite for the design of specific therapies or biologically-inspired technologies. Insertion of T4 Lysozyme in the third intracellular loop seems to be a promising strategy for GPCR crystallization, as four of the six crystallized GPCRs have a T4L insertion. The results presented here further support this. Comparison of native and T4L mutants of hOR17-4 and hVN1R1 indicates that T4L helps to solubilize and stabilize the proteins, as well as increase the yield of protein that is usable for crystallization screens. Membrane trafficking, calcium influx assays, and ligand binding analysis show that the T4L mutants maintain some biological activity. These results indicate that this may be a promising strategy for a wider range of GPCRs, and potentially other membrane proteins. This strategy may thus help facilitate crystallization of a broader range of GPCRs.

### 3.5 References

1. Cherezov V, et al. High resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science* **318**(5854), 1258-1265 (2007).
2. Chien et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091-1095 (2010).
3. Jaakola V-P, et al. The 2.6 Å crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**(5905), 1211-1217 (2008).
4. Wu B, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071 (2010).
5. Rasmussen SGF, et al. Crystal structure of the human  $\beta_2$  adrenergic G-protein-coupled receptor. *Nature* **450**, 383-388 (2007).
6. Warne T, et al. Structure of a  $\beta_1$ -adrenergic G protein-coupled receptor. *Nature* **454**(7203), 486-491 (2008).
7. Shirokova E, Raguse JD, Meyerhof W, Krautwurst D. The human vomeronasal type-1 receptor family – detection of volatiles and cAMP signaling in HeLa/Olf cells. *FASEB J.* **22**, 1416-1425 (2008).
8. Cook B, Ernberg KE, Chung H, Zhang S. Study of a synthetic human olfactory receptor 17-4: Expression and purification from an inducible mammalian cell line. *PLoS One* **3**(8): e2920. (2008).
9. Cook B. *et al.* Large-scale production and study of a synthetic G protein-coupled receptor: Human olfactory receptor 17-4. *Proc. Natl. Acad. Sci. USA* **106**(29), 11925-11930 (2009).
10. Katada S, Tanaka M, Touhara K. Structural determinants for membrane trafficking and G protein selectivity of a mouse olfactory receptor. *J. Neurochem.* **90**(6), 1453-1463 (2004).
11. Gimelbrant AA, Haley SL, McClintock TS. Olfactory receptor trafficking involves conserved regulatory steps. *J. Biol Chem.* **276**(10), 7285-7290 (2001).
12. Lu M, Staszewski L, Escheverri F, Xu H, Moyer, B. D. Endoplasmic reticulum degradation impedes olfactory G-protein coupled receptor functional expression. *BMC Cell Biol.* **5**:34 (2004).
13. Lu M, Echeverri F, Moyer BD. Endoplasmic reticulum retention, degradation, and aggregation of olfactory G-protein coupled receptors. *Traffic* **4**, 416-433 (2003).

14. Jayadev S, Smith RD, Jagadeesh G, Baukal AJ, Hunyady L. N-linked glycosylation is required for optimal AT1a angiotensin receptor expression in COS-7 cells. *Endocrinology* **140**, 2010-2017 (1999).
15. Gimelbrant AA, Stoss TD, Landers TM, McClintock TS. Truncation releases olfactory receptors from the endoplasmic reticulum of heterologous cells. *J Neurochem* **72**, 2301-2311 (1999).
16. McClintock TS, et al. Functional expression of olfactory-adrenergic receptor chimeras and intracellular retention of heterologously expressed olfactory receptors. *Mol Brain Res* **48**, 270-278 (1997).
17. Spehr M, et al. Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* **299**, 2054-2058 (2003).
18. Spehr M, et al. Particulate adenylate cyclase plays a key role in human sperm olfactory receptor-mediated chemotaxis. *J Biol Chem* **279**(38), 40194-40203 (2004).
19. Ren H, Yu D, Ge B, Cook B, Xu Z, Zhang S. High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS One* **4**(2), e4509(2009).
20. Anatrace Technical Data, obtained from [http://wolfson.huji.ac.il/purification/PDF/detergents/Anatrace\\_Structures.pdf](http://wolfson.huji.ac.il/purification/PDF/detergents/Anatrace_Structures.pdf)
21. Greenfield N, and Fasman GD. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**(10), 4108-4116 (1969).
22. Baaske P, Wienken CJ, Reineck P, Duhr S, & Braun D. Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed* **49**, 2238-2241 (2010).
23. Duhr S, & Braun D. Why molecules move along a temperature gradient. *Proc Natl Acad Sci USA* **103**, 19678-19682 (2006).
24. Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Comm.* **1**, 100 (2010).
25. Garavito RM, Picot D, Loll PJ. Strategies for crystallizing membrane proteins. *J Bioenerg Biomembr* **28**(1), 13-27 (1996).



# CHAPTER 4

## A ROBUST AND RAPID METHOD OF PRODUCING SOLUBLE, STABLE, AND FUNCTIONAL G PROTEIN-COUPLED RECEPTORS

*This chapter is an expansion of a manuscript accepted by PLoS ONE:*

*Corin, K., Baaske, P., Ravel, D.B., Song, J., Brown, E., Wang, X., Wienken, C.J., Geissler, S., Jerabek-Willemsen, M., Duhr, S., Braun, D. & Zhang, S. (2011) A robust and rapid system for studying functional G-protein coupled receptors. PLoS ONE, (In press).*

### 4.1 Introduction

G-Protein Coupled Receptors (GPCRs) are the focus of intense research, as they are the largest class of integral membrane proteins and are the targets of ~50% of pharmaceutical drugs. A critical bottleneck in GPCR studies has been the difficulty of expressing soluble and stable receptors in sufficient quantities to allow for rapid progress in structure and function studies. In order for this to occur, a rapid, simple, cost-effective and high-yield method of producing GPCRs must be found.

Cell-free *in vitro* translation is an alluring technology for the expression of GPCRs because experiments with soluble proteins show that it has the potential to fulfill the above-stated requirements. Unlike expression in eukaryotic or bacterial cells, protein can be produced within several hours directly from plasmid DNA. The typical issues of cell

toxicity, protein degradation, protein inhomogeneity, and aggregation in internal compartments or inclusion bodies do not arise.

Although cell-free *in vitro* translation is a mature technology for expression of soluble proteins, it has only recently been used to express membrane proteins [1-16]. This is due primarily to the necessity of including a detergent capable of solubilizing and stabilizing the newly synthesized proteins without interfering with transcription or translation. Finding an optimal detergent is expensive and laborious. Several studies indicate that mild detergents like polyoxyethylene derivatives may be effective [3, 14], but only a limited number of proteins have been successfully tested [1-15], and most published reports have limited their studies to one or few proteins.

Although these reports are promising, four areas must still be addressed for cell-free expression to become a practical and widely useful technology for producing GPCRs. First, milligram quantities of receptor must be able to be consistently produced and purified. Second, cell-free expression methods must be sufficiently robust to work with a wide variety of proteins. Often, specific expression platforms work well only with a limited number of membrane proteins, and even then they must often be optimized for a given protein. It is thus necessary to demonstrate that cell-free GPCR expression is feasible for numerous GPCRs. Third, cell-free expressed proteins must be directly compared to those produced in mammalian cells or purified from native tissues. Many GPCRs require post-translational modifications for structural stability or biological function. *E. coli* and wheat germ extracts do not contain the necessary machinery. Thus, it must be demonstrated that the structure and ligand-binding affinities of cell-free expressed receptors are similar to those of native receptors. No published study has directly compared a cell-free GPCR to a

counterpart expressed in mammalian cells to verify that the cell-free constructs are indeed viable. Fourth, many studies have only used cell-free systems developed in individual laboratories [1-13], which are thus not available for widespread use. In order for cell-free technology to truly benefit the entire GPCR and membrane protein research community, commercial reagents are optimal to minimize variations in sample preparations from each laboratory. It would therefore be advantageous if commercial cell-free systems currently used for soluble proteins could be optimized for large-scale GPCR expression.

In this study, commercial cell-free translation systems were used with Brij-35 for the rapid, high-yield production of 13 GPCRs (**Table 4.1**), including 9 olfactory receptors (ORs), one human trace-amine receptor (hTAAR5), one human formyl peptide receptor (hFPR3), and 2 human vomeronasal receptors (hVN1R1 and hVN1R5). The expressed GPCRs could be purified to >90%, were properly folded, and were able to bind their ligands. The GPCR hVN1R1 was stably cloned into HEK293 cells. The HEK293 and cell-free expressed hVN1R1 had comparable structures and ligand-binding properties.

## **4.2 Methods**

### **4.2.1 GPCR Choice**

Nine diverse olfactory receptors were chosen for this study (**Table 4.1**). The receptors were of human, mouse or rat origin, and came from both Class I and Class II type receptors. Because of the necessity of measuring OR-ligand interactions to show functionality, only ORs with known ligands were chosen. The chosen ORs have been reported to respond to chemically diverse odorants.

Only human receptors were chosen from the remaining olfactory-related GPCR families. Most studies examine mouse receptors due to the relative ease of studying how these receptors function in their native environment. Cell-free expression can make the human receptors more amenable to study, which may help elucidate differences between the human and mouse sense of smell. Two human vomeronasal receptors with known ligands were chosen (Table 4.1). Although the existence of a functional human vomeronasal organ is debated, the chosen receptors have been expressed in HeLa Olf cells, and have been shown to initiate the canonical G-protein signalling pathway upon exposure to their ligands [17]. They also have intact reading frames in the human genome. The receptor hVN1R1 was chosen because studies have reported mRNA expression in several tissues, including the olfactory epithelium [18]. It also has a high sequence homology to mouse and rat VNRs, and contains the same conserved residues [18]. The receptor hVN1R5 was chosen because it is least homologous to the other 4 human VNRs, as well as VNRs in other species [19], and therefore represents an intriguing point of comparison. The human hTAAR5 gene was chosen because it is phylogenetically unique [20]. Also, it does not bind to the same ligands as its mouse orthologue, suggesting different functions in mice and humans [21]. Although an olfactory role for the human FPRs has not been demonstrated, the receptor hFPR3 was chosen. Because these receptors function in the immune system, expression of an active receptor would indicate that cell-free technology can be expanded to broader classes of GPCRs, and is not limited solely to those involved in olfaction.

**Table 4.1: Experimental GPCRs and Their Properties**

Receptor Name (Accession Number)	Amino Acids	Molecular Weight (Da)	Ligands	Species
mOR33-1 (GenBank AAL60676.1)	331	36686.00	Aliphatic alcohols and carboxylic acids [22]	Mouse, Class I
mOR103-15 (NP_035113.1)	340	37324.02	Heptanal [23-25]	Mouse, Class II
mOR106-13 (NP_001011738.1)	337	37263.62	Aliphatic alcohols [22]	Mouse, Class II
mOR175-1 (SwissProt Q9QY00.1)	332	37097.23	2-Heptanone, isoamyl acetate [26-28]	Mouse, Class II
mOR174-4 (GenBank BAB59038.1)	332	36875.88	Ethyl vanillin [29,30]	Mouse, Class II
mOR174-9 (NP_473431.1)	327	36245.05	Ethyl Vanillin, Eugenol [29-31]	Mouse, Class II
hOR17-209 (NP_003546.1)	326	36097.60	Isoamyl acetate [32]	Human, Class II
hOR17-210 (SwissProt Q8WZA6.2)	357	39748.54	Acetophenone [32]	Human, Class II
Olf226 (SwissProt P23270.2)	339	37398.36	2,4-DNT [33]	Rat, Class II
VN1R1 (AF255342)	40021.02	353	Myrtenal, Carveol, $\beta$ -Ionone [17]	Human
VN1R5 (AY114735)	40778.58	357	Myrtenal, Carveol, $\beta$ -Ionone [17]	Human
hTAAR5 (NP_003958.2)	337	38242.45	None identified	Human
hFPR3 (NP_002021.3)	353	39965.40	F2L, WKYMVM [34, 35]	Human

#### 4.2.2 GPCR DNA Design and Synthesis

The protein sequences of the chosen receptors were obtained from the NCBI online database (Table 4.1). To facilitate purification and detection of synthesized receptors, the

rho1D4 epitope tag (TETSQVAPA) preceded by a GSSG linker, or a GG linker for the VNRs, was added to the C-terminus of each protein. EcoRI and NcoI restriction sites were added to the 3' end, and an XhoI restriction site was added to the 5' end. A second stop codon was inserted after the original codon to ensure completion of transcription. A Kozak sequence was added upstream of the start codon to facilitate expression in mammalian cells, if necessary. The codons for each receptor were optimized for *E. coli* expression. The genes were commercially synthesized by GeneArt and subcloned into the pIVex2.3d vector (Roche Diagnostics Corp.) using the NcoI and XhoI restriction sites. The hVN1R1 gene was also subcloned into the pcDNA4/To vector (Invitrogen) using the EcoRI and XhoI restriction sites. This vector was used for mammalian expression, and is described in Chapter 2. The plasmids were transformed into DH5 $\alpha$  cells, amplified, and then purified using the Qiagen miniprep or maxiprep kits. The final plasmid constructs were verified by DNA sequencing (MIT Biopolymers Labs, Cambridge, MA).

#### **4.2.3 Cell-Free GPCR Production Using Commercial Kits**

*E. coli*-based cell-free expression kits were used to synthesize the GPCRs according to the manufacturer's instructions (Invitrogen K9900-97, Qiagen 32506), with the exception that reactions were performed at 30-33°C. To compensate for the lack of a natural membrane, detergents were added directly to the reactions. A preliminary screen determined that the optimal concentration was 0.2% w/v. A final reaction volume of 50 $\mu$ l was used for all screens. After the reactions were complete, the samples were centrifuged at 10,000 rpm for 5 minutes. The supernatant containing the solubilized protein was removed, and the pellet was resuspended in an equivalent volume of PBS. The relative

quantities of solubilized and precipitated protein were determined with a western or dot blot. ImageJ (<http://rsb.info.nih.gov/ij/>) was used to perform all densitometry analyses. Final reaction volumes of 0.5-1.0ml were used to produce protein that was purified for secondary structure and binding analyses.

#### **4.2.4 HEK293 GPCR Production**

A stable, inducible HEK293 cell line expressing hVN1R1 was generated as previously described [36, 37]. Briefly, plasmid DNA was transfected into HEK293G cells already expressing the pcDNA6/TR vector using Lipofectamine 2000 (Invitrogen). The transfected cells were grown in selective media containing Zeocin and Blasticidin until individual resistant colonies were visible. Resistant colonies were picked and screened for optimal protein expression. The clone with the highest expression level and least toxicity was selected and amplified for subsequent experiments. Detailed procedures are described in Chapter 2.

#### **4.2.5 GPCR Detection and Purity Analysis**

Western blots and silver stains were used to detect proteins and analyze their purity. Samples were prepared and loaded in Novex 10% Bis-Tris SDS-PAGE gels (Invitrogen) according to the manufacturer's protocol, with the exception that the samples were incubated at room temperature prior to loading as boiling causes membrane protein aggregation. For blotting, the gel-resolved samples were transferred to a nitrocellulose membrane, blocked in milk (5% w/v non-fat dried milk in TBST) for 1 hour, and incubated with a rho1D4 primary antibody (1:3000 in TBST, 1 hour at room temperature, or

overnight at 4°C). The GPCRs were then detected with a goat anti-mouse HRP-conjugated secondary antibody (Pierce, Rockford, IL) (1:5000 in TBST, 1 hour, room temperature) and visualized using the ECL-Plus Kit (GE Healthcare). The SilverXpress kit (Invitrogen, LC6100) was used according to the manufacturer's instructions to perform total protein stains of the samples. All images were captured using a Fluor Chem gel documentation system (Alpha Innotech, San Leandro, CA).

#### **4.2.6 Immunoaffinity Purification**

CNBr-activated Sepharose 4B beads (GE Healthcare) chemically linked to the rho1D4 monoclonal antibody (Cell Essentials, Boston, MA) were used for immunoaffinity purification. Solubilized protein from the cell-free reactions was mixed with the bead slurry (binding capacity 0.7mg/ml) and rotated overnight at 4°C to capture the synthesized protein. The beads were then washed with wash buffer (PBS + 0.2% FC-14 w/v) until spectrophotometer readings indicated that all excess protein had been removed (<0.01 mg/ml). The captured GPCRs were eluted with elution buffer (PBS + 0.2% FC-14 + 800µM elution peptide). The elution peptide Ac-TETSQVAPA-CONH<sub>2</sub> was synthesized by CPC Scientific Inc., CA. Elutions were performed until spectrophotometer readings indicated that no more protein was present (<0.01mg/ml). The protein was concentrated using 30kDa or 50kDa MWCO filter columns (Millipore, Billerica MA). All concentrations were measured using the NanoDrop 1000 spectrophotometer (Thermo Scientific). For some samples, the concentration was also measured with a total protein stain by comparing the intensity of the GPCR band to the intensity of a BSA band of known concentration. The



beads were pelleted by centrifugation at 1,400xg for one minute between each wash and elution.

#### **4.2.7 Secondary Structural Analysis Using Circular Dichroism**

Spectra were recorded on a CD spectrometer (Aviv Biomedical, model 410) at 15°C over the wavelength range of 195-250 nm with a step size of 1nm and an averaging time of 4 seconds. Spectra for purified GPCRs were blanked to wash buffer. A 111-QS quartz sample cell with a path length of 1mm (Hellma, USA) was used. 300µl of protein sample was used for each experiment. The spectra were smoothed using an averaging filter with a span of 5.

#### **4.2.8 Ligand Binding Measurements Using Microscale Thermophoresis**

Thermophoresis was used to measure the binding interactions between purified receptors and their ligands using a setup similar to that previously described [38]. To eliminate artifacts caused by labeling or modifying proteins, the fluorescence of native GPCR tryptophans was used to monitor the local receptor concentration. For each tested GPCR, a titration series with constant receptor concentration and varying ligand concentrations was prepared in a final solution of 10% DMSO and 0.2% FC-14 in PBS. Potential autofluorescence of each ligand was checked: no fluorescence signal was detected from the ligands in the tryptophan fluorescence channel. The final receptor concentration was 1-2µM. Approximately 1.5µl of each sample was loaded in a fused silica capillary (Polymicro Technologies, Phoenix, USA) with an inner diameter of 300µm. An infrared laser diode was used to create a 0.12K/µm temperature gradient inside the

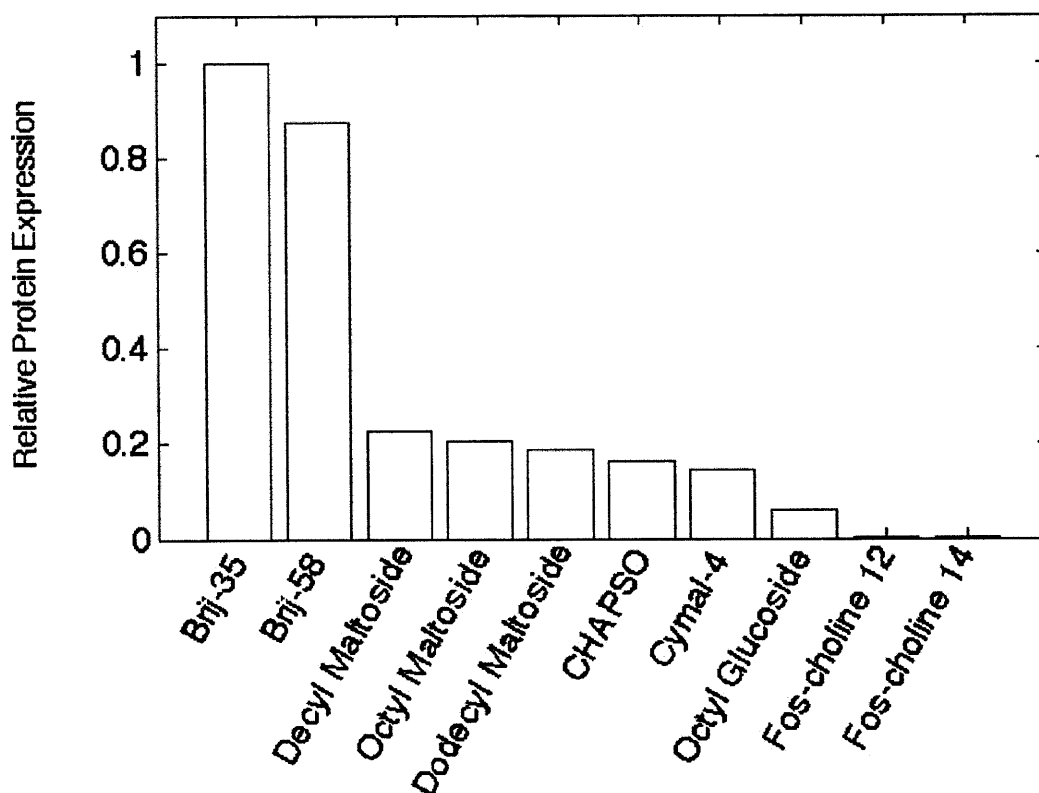
capillaries (Furukawa FOL1405-RTV-617-1480, wavelength  $\lambda=1480\text{nm}$ , 320mW maximum power, AMS Technologies AG, München Germany). Tryptophan fluorescence was excited with a UV-LED (285nm), and was measured with a 40x SUPRASIL synthetic quartz substrate microscope objective, numerical aperture 0.8 (Partec, Goerlitz, Germany). The local receptor concentration in response to the temperature gradient was detected with a photon counter PMT P10PC (Electron Tubes Inc, Rockaway, NJ, USA). All measurements were performed at room temperature. Fluorescence filters for tryptophan (F36-300) were purchased from AHF-Analysentechnik (Tübingen, Germany).

## 4.3 Results

### 4.3.1 Cell-Free Detergent Screening

Numerous detergents were systematically screened to assess their ability to produce and solubilize GPCRs expressed in cell-free systems. The chosen detergents included, among others, OG, DM, DDM, CHAPSO, Brij-35, Brij-58, and fos-choline14 (FC14). These latter detergents have effectively solubilized and stabilized GPCRs produced in mammalian cells [36, 37], and have been used to obtain high-resolution protein structures. Few suitable detergents were found. The gentle non-ionic detergents tended to work best, with polyoxyethylene derivatives proving to be the most efficient. Indeed, Brij-35 and Brij-58 yielded  $\sim 4\text{-}5$  times more protein than the next best detergent, with Brij-35 consistently achieving slightly higher yields. **Figure 4.1** shows the results from selected detergents tested with hOR17-210. These results are representative of results obtained with other receptors, and demonstrate that the choice of detergent is critical. Detergents commonly used in cell-based production may not be optimal for use in cell-free systems. Indeed, while FC14 achieved high receptor yields from HEK293 cells, negligible

protein was detected after cell-free synthesis. This suggests that some detergents may inhibit transcription or translation by interfering with ribosomes or other synthesis machinery, or by interacting with nucleic acids.



**Figure 4.1. Detergent Screen for Cell-Free hOR17-210 Production.** Brij-35 and Brij-58 yielded ~4-5 times more receptor than the next best detergent. Although comparable, Brij-35 consistently had slightly higher yields than Brij-58. Each bar represents the average of 2-3 experiments. The data was normalized to Brij-35, and is representative of results observed with other receptors.

#### 4.3.2 GPCR Expression and Solubility

Because Brij-35 resulted in the highest cell-free expression levels for a subset of the chosen GPCRs, its ability to solubilize all of the GPCRs was assessed. Western and dot blot analyses were used to compare soluble and insoluble protein fractions. Without Brij-35, only ~10% of the produced protein is soluble. With Brij-35, up to ~93% is soluble (Table

4.2). This data demonstrates the necessity of using a detergent to prevent aggregation and potential misfolding of the synthesized proteins. The high solubility level is sufficient to collect the majority of the expressed protein for downstream applications. Reconstitution strategies are thus not necessary.

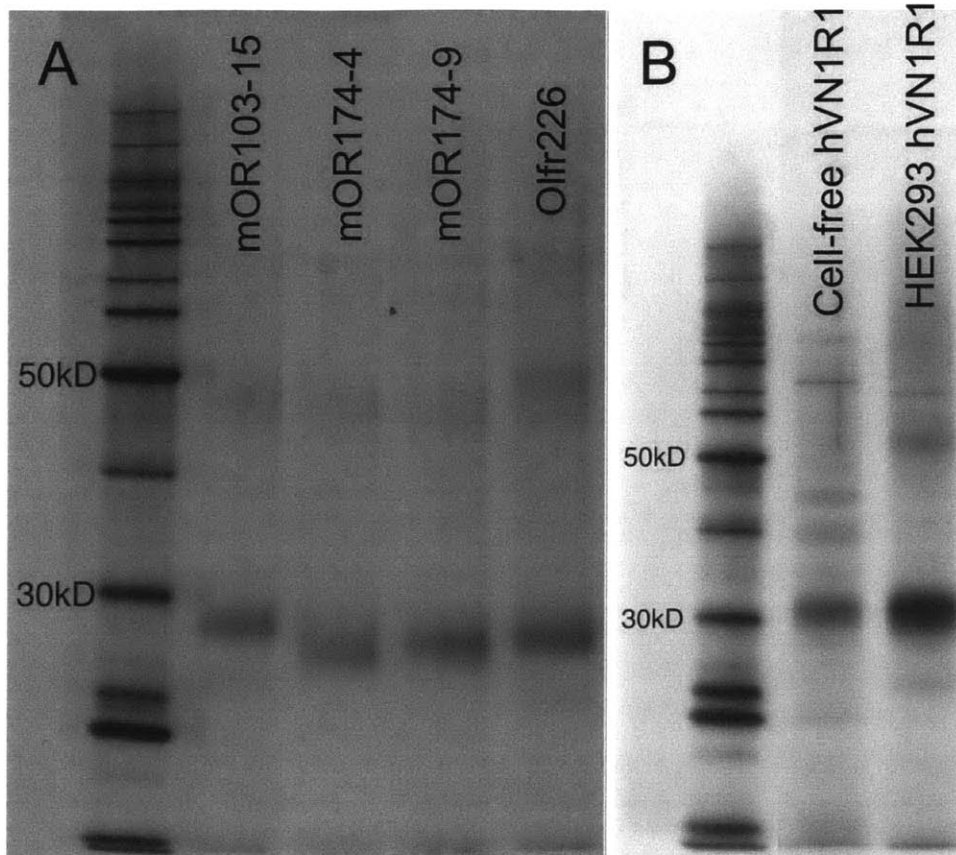
#### 4.3.3 GPCR Purification and Yields

All of the GPCRs were produced in a cell-free reaction, and purified using rho-1D4 antibody tagged beads. The HEK293-expressed hVN1R1 was also purified in a similar manner. Each receptor was purified in the presence of fos-choline 14 (FC14) because it has been shown to be the optimal detergent for GPCR purification [36, 37, 39, 40]. Moreover, this demonstrates the ability to perform a detergent exchange. Since the optimal detergent depends on both the specific protein and the application, the ability to exchange detergents is critical.

**Figure 4.2** shows a silver stain of 5 purified GPCRs expressed in a cell-free environment, as well as the HEK293 counterpart. Two bands are visible for each receptor, corresponding to monomeric and dimeric forms. Dimerization of olfactory receptors has been reported previously [36, 37, 41, 42], and thus indicates that the cell-free reactions are producing viable proteins. Dimerization of hVN1R1 has not previously been reported due to the difficulty of expressing and purifying the protein. However, the occurrence of dimerization in both the cell-free and HEK293 samples suggests that this expression pattern is typical for this receptor. Although not shown, all of the ORs, hTAAR5, and hFPR3 displayed a monomeric and dimeric band when run on a gel.

The purified receptors could be produced at sufficient purity for crystal screens.

**Figure 4.2** shows that the receptors were up to 90% pure using immunoaffinity chromatography alone. Due to the large number of tested GPCRs, 100-1000 $\mu$ L cell-free reactions were performed. The resulting yields were too small to run on a size exclusion column due to typical protein losses. However, previous studies have demonstrated that immunoaffinity chromatography followed by size exclusion chromatography (SEC) can yields receptors over 95% pure. Moreover, SEC can separate the monomer from the dimer, yielding a pure and homogeneous sample [36, 37]. Indeed, this level of homogeneity and purity was obtained for the HEK293 VNR1 sample. This level of purity and homogeneity that can be achieved from cell-free reactions is sufficient for crystallization.



**Figure 4.2. Silver Stains of Purified GPCRs.** A) Four cell-free expressed GPCRs. B) Comparison between cell-free and HEK293 expressed hVN1R1. Most GPCRs could be purified to >90% purity, and all showed two bands characteristic of a monomer and a dimer [36, 37]. The cell-free and HEK293 expressed receptors run at the same size, and have similar purities.

The maximum yield of each GPCR after purification was estimated using spectroscopic measurements at 280nm. **Table 4.2** shows the maximum yield of each GPCR produced in Brij-35. These yields are comparable to those obtained from protein expressed in mammalian cells, and are sufficient for biochemical and structural studies. Unlike cell-based protein production, which requires several months, cell-free systems can produce milligrams of protein within hours directly from plasmid DNA. Cell-free production of GPCRs is thus a promising and attractive technique for membrane protein studies.

**Table 4.2: The Solubility and Maximum Yields of 13 GPCRs**

GPCR	% Solubility	Yield (mg/10ml)*	GPCR	% Solubility	Yield (mg/10ml)*
Olf226	86 ± 8	3.7	hOR17-209	88 ± 4	2.5
mOR33-1	85 ± 2	5.9	hOR17-210	91 ± 2	4.5
mOR103-15	90 ± 4	4.5	hFPR3	83 ± 5	5.5
mOR106-13	86 ± 13	2.4‡	hTAAR5	90 ± 1	4.5
mOR174-4	89 ± 2	2	hVN1R1	88 ± 0.1	0.4
mOR174-9	86 ± 3	6	hVN1R5	85 ± 2	1‡
mOR175-1	81 ± 8	2.5			

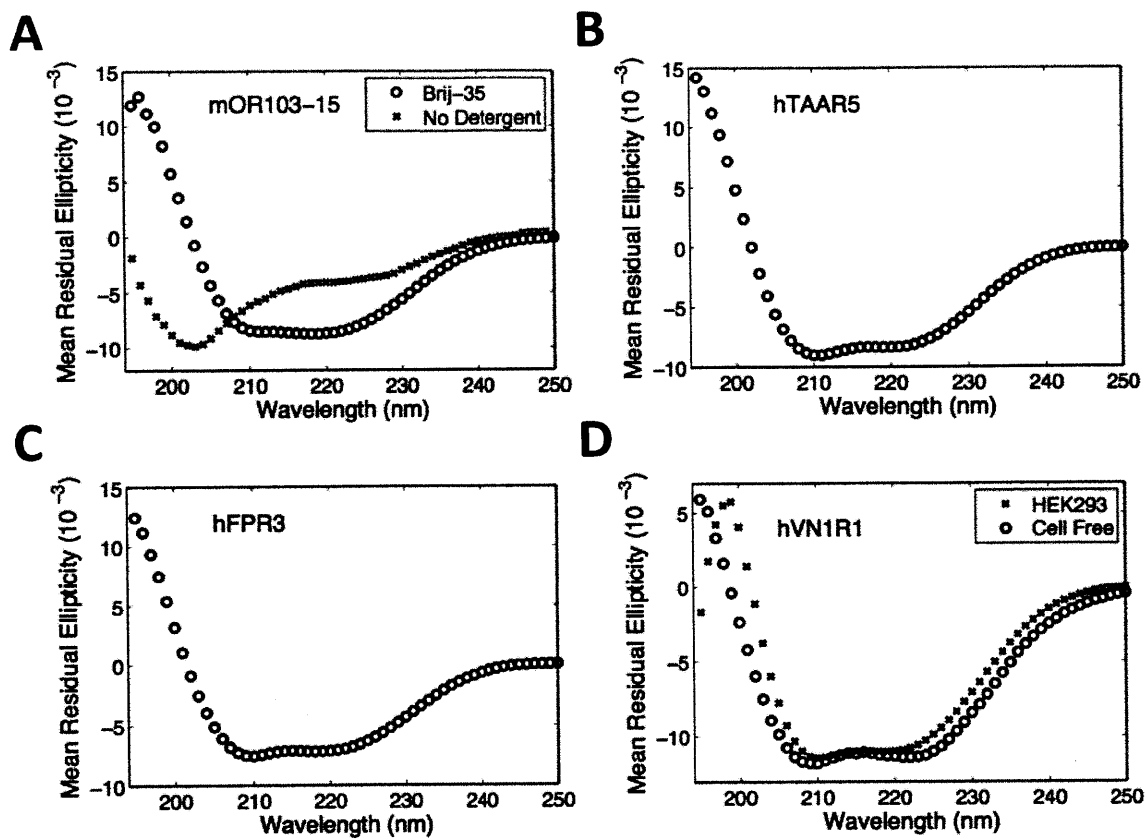
\*Milligrams of receptor that could be produced in a 10ml cell-free reaction. These yields were calculated from smaller batches of protein purified using immunoaffinity chromatography. Experiments showed that up to 1mg/ml of protein could be produced, but that up to half could be lost during the purification process. The yields were determined by spectrophotometer readings.

‡ These proteins were not purified. Yields were calculated by comparing the intensities of the receptor samples against a sample with a known concentration.

#### 4.3.4 Secondary Structural Analysis Using Circular Dichroism

Circular dichroism (CD) was used to assess the secondary structure of the produced GPCRs. **Figure 4.3** shows the CD spectra of 4 GPCRs. All of the receptors have characteristic  $\alpha$ -helical spectra, with signature valleys at 208nm and 220nm. Since GPCRs have 7-transmembrane  $\alpha$ -helical segments, the spectra suggest that the receptors are properly folded. The near overlap of the spectra for the cell-free and HEK293 hVN1R1 samples indicates that the cell-free reactions produce properly folded receptors.

As a control, the receptor mOR103-15 was produced without detergent, and purified using PBS alone. The CD spectrum of this receptor is characteristic of a random coil [43] (**Figure 4.3**), suggesting that it is misfolded. The contrast between this control receptor and those produced and purified with detergent further suggests that the Brij-produced GPCRs are properly folded.



**Figure 4.3. Circular Dichroism Spectra of Cell-Free Expressed GPCRs.** A) mOR103-15 made with no detergent (x) or in the presence of Brij-35 (o). B) hTAAR5 made in Brij-35. C) hFPR3 made in Brij-35. D) Cell-free expressed hVN1R1 made in Brij-35 (o), or hVN1R1 expressed in HEK293 cells (x). The GPCRs expressed in HEK293 cells or in cell-free reactions in the presence of Brij-35 have characteristic alpha helical spectra. Since GPCRs have seven transmembrane helical domains, this suggests that the receptors are properly folded. The near overlap of the HEK293 and cell-free expressed receptors indicates that protein produced in cell-free reactions is comparable to that produced in cells. mOR103-15 made without detergents has a curve characteristic of a random coil. Thus further suggests that the receptors expressed in Brij-35 are properly folded, and demonstrates the necessity of including a detergent.

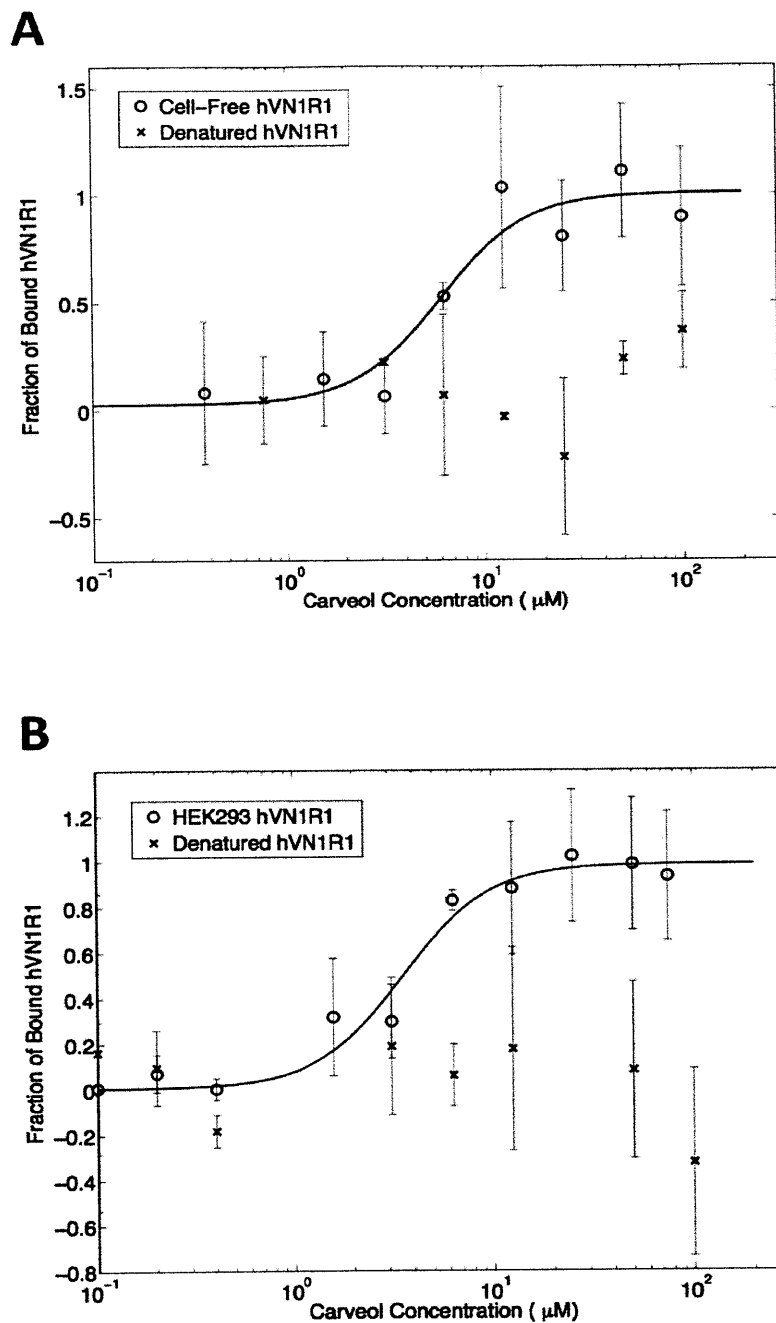
#### 4.3.5 Ligand Binding Analysis Using Microscale Thermophoresis

Microscale thermophoresis was used to detect binding between the purified GPCRs and their ligands. Microscale thermophoresis is based on the ligand-binding-induced change in movement of molecules along a temperature gradient [38, 44]. Unlike surface plasmon resonance (SPR) or other surface-based techniques, thermophoresis is a label-free and surface-



free technology that can be used with sample volumes smaller than 5 $\mu$ L per measurement. Thermophoretic molecular gradients are measured in free solution using the fluorescence of a protein's native tryptophans. Immobilization and other coupling chemistries that could alter protein function are thus avoided. Moreover, thermophoresis can detect ligands as small as 40Da [45]. Most volatile odorants are less than 300Da, whereas their receptors are over 30,000Da (**Table 4.1**). Because of the large mass ratio, these binding interactions are extremely difficult to measure using mass-based technologies, but are possible using microcale thermophoresis [38, 44, 45].

The purified receptors shown in **Figure 4.2** were analyzed for ligand binding. Heat-denatured receptors were used as negative controls, as well as mOR103-15 that was not expressed in the presence of a detergent. Each native receptor exhibited a typical sigmoidal binding curve, while the heat-denatured and detergent-free controls had random amplitudes throughout the ligand titration range (**Figure 4.4**). These results suggest that the cell-free produced GPCRs bound their ligands. All of the measured binding affinities were in the micromolar range, (**Table 4.3**) which is consistent with previous reports [14, 17, 36]. The cell-free and HEK293 expressed proteins had similar binding affinities, further demonstrating that cell-free produced proteins are functional.



**Figure 4.4. Microscale Thermophoresis Measurements of Purified GPCRs.** A) Cell-free expressed hVN1R1 (o) and the heat-denatured control (x). B) HEK293 expressed hVN1R1 (o) and the heat-denatured control (x). The non-denatured receptors show typical sigmoidal binding curves, with plateaus at low and high concentrations. The heat-denatured controls have flat responses or random amplitudes throughout the ligand titration range. Together, these results show hVN1R1 is binding carveol. The similar  $EC_{50}$  values demonstrate that cell-free produced receptors function as well as HEK293 expressed receptors. The curves were normalized to the fraction of bound receptor. Each data point represents the mean of 3 independent experiments; error bars show the standard deviation. The binding curves were fit to the Hill equation. The binding results shown are representative of the data from other receptors.

**Table 4.3: Ligand Binding Measurements of Selected GPCRs**

Receptor	Expression Method Used	Ligand	Measured EC <sub>50</sub> (μM)
mOR103-15	Cell-free	Heptanal	2 ± 0.7
mOR174-4	Cell-free	Ethyl Vanillin	7 ± 2
mOR174-9	Cell-free	Ethyl Vanillin	4.9 ± 3.5
Olf226	Cell-free	2,4-DNT	86 ± 36
hVN1R1	Cell-free	Carveol	6 ± 2
hVN1R1	HEK293	Carveol	3.5 ± 0.7

## 4.4 Discussion

This study shows that cell-free membrane protein production is a useful technology for expressing milligrams of GPCRs. The receptors could be purified to ~90% purity using immunoaffinity chromatography alone. CD measurements on a subset of purified GPCRs showed that they had the predicted secondary structures, which suggests that they were properly folded. Microscale thermophoresis indicated that the cell-free-produced GPCRs were functional by showing that the purified receptors could bind their reported ligands. Comparison of HEK293 and cell-free expressed proteins suggests that cell-free systems are a practical alternative to cell-based platforms for producing GPCRs.

Although cell-free production is a mature technology for soluble proteins, very few membrane proteins have been produced [1-16], largely due to the lack of suitable detergents and laborious detergent screens. In this study, Brij-35 performed consistently as the optimal detergent for olfactory-related GPCRs. Previous reports suggest that, while Brij-35 may not be

optimal for all membrane proteins or GPCRs, the Brij family of detergents may function best with cell-free membrane protein expression [3, 14]. While the best detergent for protein production may not be the best detergent for downstream applications, a single detergent exchange with FC14 is possible without compromising receptor structure and function. Since FC14 has been used to obtain protein structures [46, 47], it should be possible to couple cell-free expression with crystal screens or NMR structural studies.

**Table 4.4: GPCR Yields in Large-Scale Expression Platforms\***

Expression Platform	Receptor Yield
<i>E. coli</i>	~0.15-0.20 mg/L
Yeast	~4 mg/L
Insect Cells	~2mg/L
Mammalian Cells	~1.7-2 mg/L or <1mg/g of cells
Cell-Free	~0.5-1 mg/mL

\*Adapted from [48].

In order to accelerate membrane protein structure and function studies, it is vital to develop simple, straightforward methods of producing sufficient quantities of membrane proteins. Commercial cell free kits offer an attractive alternative to cell-based systems. In this study, milligrams of protein could be produced within hours directly from plasmid DNA. These yields are comparable to those previously reported, and are significantly greater than yields in other systems (Table 4.4) [48]. Because the materials used here are commercially available, the necessary reagents are easily and widely obtainable, and results are reproducible. Furthermore, the produced proteins can be purified quickly using conventional methods, and are amenable to detergent exchange for downstream applications. Although the 13 GPCRs reported here represent a small fraction of all receptors, it is the largest number presented in a single study with a single method. The ability, demonstrated by this study, to produce significant quantities of

these GPCRs using commercial cell-free systems demonstrates in turn the usefulness of this technology in the field. Indeed, the critical production bottleneck in membrane protein studies may potentially be overcome. Structure and function studies of additional GPCRs may be stimulated and accelerated in the coming years.

## 4.5 References

1. Klammt C, Löhr F, Schäfer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C, Bernhard F. (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem.* 271: 568-580.
2. Klammt C, Schwarz D, Fendler K, Haase W, Dötsch V, Bernhard F. (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* 272: 6024-6038.
3. Klammt C, Schwarz D, Eifler N, Engel A, Piehler J, Haase W, Hahn S, Dötsch V, Bernhard F. (2007) Cell-free production of G protein-coupled receptors for functional and structural studies. *J Struct Biol.* 158: 482-493.
4. Klammt C, Srivastava A, Eifler N, Junge F, Beyermann M, Schwarz D, Michel H, Dötsch V, Bernhard F. (2007) Functional analysis of cell-free-produced human endothelin B receptor reveals transmembrane segment 1 as an essential area for ET-1 binding and homodimer formation. *FEBS J*, 274:3257-3269.
5. Schwarz D, Junge F, Durst F, Frölich N, Schneider B, Reckel S, Söbhanifar S, Dötsch V, Bernhard F. (2007) Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nat Protoc.* 2: 2945-2957.
6. Schwarz D, Dötsch V, Bernhard F. (2008) Production of membrane proteins using cell-free expression systems. *Proteomics.* 8: 3933-3946.
7. Junge F, Luh LM, Proverbio D, Schäfer B, Abele R, Beyermann M, Dötsch V, Bernhard F. (2010) Modulation of G-protein coupled receptor sample quality by modified cell-free expression protocols: a case study of the human endothelin A receptor. *J. Struct. Biol.* 172: 94-106.
8. Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M, Yokohama S. (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* 41: 27-37.
9. Liguori L, Marques B, Lenormand JL. (2008) A bacterial cell-free expression system to produce membrane proteins and proteoliposomes: from cDNA to functional assay. *Curr Protoc Protein Sci.* Chapter 5:Unit 5.22.
10. Kamonchanok S, Balog CI, van der Does AM, Booth R, de Grip WJ, Deelder AM, Bakker RA, Leurs R, Hensbergen PJ. (2008) GPCR proteomics: mass spectrometric and functional analysis of histamine H1 receptor after baculovirus-driven and in vitro cell free expression. *J Proteome Res.* 7: 621-629.
11. Maslennikov I, Klammt C, Hwang E, Kefala G, Okamura M, Esquivies L, Mörs K, Glaubitz

- C, Kwiatkowski W, Jeon YH, Choe S. (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. *Proc Natl Acad Sci U S A.* 107: 10902-10907.
12. Kai L, Kaldenhoff R, Lian J, Zhu X, Dötsch V, Bernhard F, Cen P, Xu Z. (2010) Preparative scale production of functional mouse aquaporin 4 using different cell-free expression modes. *PLoS ONE* 5:e12972.
  13. Savage DF, Anderson CL, Robles-Colmenares Y, Newby ZE, Stroud RM. (2007) Cell-free complements in vivo expression of the E. coli membrane proteome. *Protein Sci* 16: 966-976.
  14. Kaiser L, Graveland-Bikker J, Steuerwald D, Vanberghem M, Herlihy K, Zhang S. (2008) Large-scale production and study of a synthetic G protein-coupled receptor: human olfactory receptor 17-4. *Proc Natl Acad Sci USA* 105: 15726-15731.
  15. Deniaud A, Liguori L, Blesneac I, Lenormand JL, Pebay-Peyroula E. Crystallization of the membrane protein hVDAC1 produced in cell-free system. (2008) *Biochim Biophys Acta.* 1798: 1540-1546.
  16. Wang X, Corin K, Wienken CJ, Jerabek-Willemsen M, Duhr S, Braun D, Zhang S. (2011) Designer lipid-like peptide surfactants for cell-free production of diverse functional G-protein coupled receptors. *Proc. Natl. Acad. Sci. USA* 118: (accepted).
  17. Shirokova E, Raguse JD, Meyerhof W, Krautwurst D. The human vomeronasal type-1 receptor family – detection of volatiles and cAMP signaling in HeLa/Olf cells. *FASEB J.* 22, 1416-1425 (2008).
  18. Rodriguez I, Greer CA, Mok MY, Mombaerts P. A putative pheromone receptor gene expressed in human olfactory mucosa. *Nat. Genet.* 26, 18-19. (2000).
  19. Rodriguez I, and Mombaerts P. Novel human vomeronasal receptor-like genes reveal species-specific families. *Curr. Biol.* 12, R409-411 (2002).
  20. Lindemann L, et al. Trace amine-associated receptors form structurally and functionally distinct subfamilies of novel G protein-coupled receptors. *Genomics* 85, 372-385 (2005).
  21. Staubert C, et al. Structural and functional evolution of the trace amine-associated receptors TAAR3, TAAR4, and TAAR5 in primates. *PLoS ONE* 5(6), e11133 (2010).
  22. Malnic B, Hirono J, Sato T, Buck LB. Combinatorial receptor codes for odors. *Cell* 96, 713-23 (1999).
  23. Krautwurst D, Yau KW, and Reed RR. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* 95(7),917-26 (1998).

24. Minic *et al.* Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening. *FEBS J* **272**(2), 524-37 (2005).
25. Araneda RC, Kini AD, Firestein S. The molecular receptive range of an odorant receptor. *Nat Neurosci.* **3**, 1248-55 (2000).
26. Gaillard *et al.* A single olfactory receptor specifically binds a set of odorant molecules. *Eur J Neurosci.* **15**, 409-18 (2002).
27. Gaillard *et al.* Amino-acid changes acquired during evolution by olfactory receptor 912-93 modify the specificity of odorant recognition. *Hum Mol Genet* **13**(7), 771-80 (2004).
28. Rouquier *et al.* A gene recently inactivated in human defines a new olfactory receptor family in mammals. *Hum Mol Genet* **7**(9), 1337-45 (1998).
29. Kajiyama *et al.* Molecular bases of odor discrimination: Reconstitution of olfactory receptors that recognize overlapping sets of odorants. *J Neurosci* **21**(16), 6018-25 (2001).
30. Katada *et al.* Odorant response assays for a heterologously expressed olfactory receptor. *Biochem Biophys Res Commun* **305**(4), 964-9 (2003).
31. Oka *et al.* An odorant derivative as an antagonist for an olfactory receptor. *Chem Senses* **29**(9), 815-22 (2004).
32. Matarazzo *et al.* Functional characterization of two human olfactory receptors expressed in the baculovirus Sf9 insect cell system. *Chem Senses* **30**, 195-207 (2005).
33. Radhika V, et al. Chemical sensing of DNT by engineered olfactory yeast strain. *Nat Chem Biol* **3**, 325-330 (2007).
34. Christophe T, *et al.* The synthetic peptide Trp-Lys-Tyr-Met-Val-Met-NH<sub>2</sub> specifically activates neutrophils through FPRL1/Lipoxin A<sub>4</sub> receptors and is an agonist for the orphan monocyte-expressed chemoattractant receptor FPRL2. *J Biol Chem* **276**(24), 21585-21593 (2001).
35. Migeotte I, et al. Identification and characterization of an endogenous chemotactic ligand specific for FPRL2. *J Exp Med* **201**(1), 83-93 (2001).
36. Cook B, Steuerwald D, Kaiser L, Graveland-Bikker J, Vanberghem M, Herlihy K, Pick H, Vogel H, Zhang S. Large scale production and study of a synthetic G-protein coupled receptor: Human olfactory receptor 17-4. *Proc Natl Acad Sci USA* **106**, 11925-11930 (2009).
37. 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. *PLoS ONE* **3**, e2920 (2008).



38. Baaske P, Wienken CJ, Reineck P, Duhr S, Braun D. Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed Engl* **49**, 2238-2241 (2010).
39. Ren H, Yu D, Ge B, Cook BL, Xu Z, Zhang S. High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS ONE* **4**, e4509 (2009).
40. Leck K-J, Zhang S, Hauser CAE. Study of bioengineered zebra fish olfactory receptor 131-2: receptor purification and secondary structure analysis *PLoS ONE* **5**, e15027 (2010).
41. Gat U, Nekrasova E, Lancet D, Natochin M. Olfactory receptor proteins. Expression, characterization and partial purification. *Eur. J. Biochem.* **225**, 1157-1168 (1994).
42. Nekrasova E, Sosinskaya A, Natochin M, Lancet D, Gat U. Overexpression, solubilization and purification of rat and human olfactory receptors. *Eur. J. Biochem.* **238**, 28-37 (1996).
43. Greenfield N, and Fasman GD. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**(10), 4108-4116 (1969).
44. Duhr S, Braun D. Why molecules move along a temperature gradient. *Proc Natl Acad Sci USA* **103**, 19678-19682 (2009).
45. Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Communications* **1**, 100 (2010).
46. Bass RB, Strop P, Barclay M, Rees DC. Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel. *Science* **298**, 1582-1587 (2002).
47. Wang W, Black SS, Edwards MD, Miller S, Morrison EL, Bartlett W, Dong C, Naismith JH, Booth IR. The structure of an open form of an E. coli mechanosensitive channel at 3.45 Å resolution. *Science* **321**, 1179-1183 (2008).
48. McCusker EC, Bane SE, O'Malley MA, Robinson AS. Heterologous GPCR expression: A bottleneck to obtaining crystal structures. *Biotechnol Prog* **23**, 540-547 (2007).

# CHAPTER 5

## DESIGNER LIPID-LIKE PEPTIDES: A CLASS OF DETERGENTS FOR PRODUCING, SOLUBILIZING, AND STABILIZING FUNCTIONAL OLFACTORY RECEPTORS

*This chapter is an expansion of a manuscript accepted by PLoS ONE, and a manuscript published by PNAS:*

*Corin, K., Baaske, P., Ravel, D.B., Song, J., Brown, E., Wang, X., Wienken, C.J., Jerabek-Willemsen, M., Duhr, S., Braun, D., Luo, Y., & Zhang, S. (2011) Designer lipid-like peptides: A class of detergents for studying functional olfactory receptors using commercial cell-free systems. PLoS ONE, (In press).*

*Wang, X.\*, Corin, K.\*, Baaske, P., Wienken, C.J., Jerabek-Willemsen, M., Duhr, S., Braun, D., & Zhang, S. (2011) Peptide surfactants for cell-free production of functional G protein-coupled receptors. PNAS, doi: 10.1073/pnas.1018185108.*

*\*Equal contributors*

### 5.1 Introduction

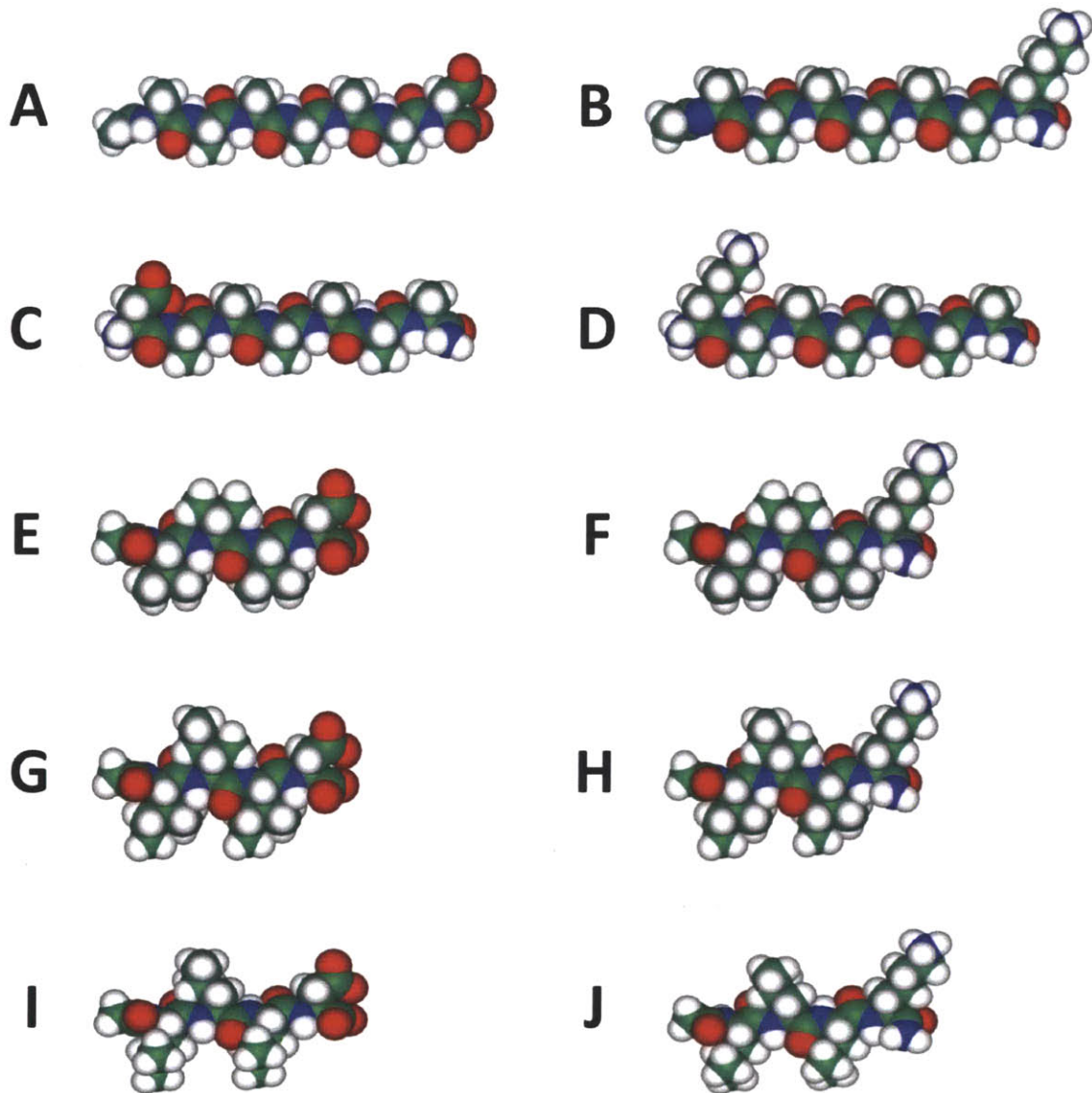
Selecting an appropriate detergent is crucial for membrane protein studies, yet is a daunting task. A bewilderingly large selection of surfactants is available, and the optimal detergent for a protein must be empirically determined [1]. To further complicate matters, detergents that are optimal for one application may not be appropriate for others. For example,

detergents that best solubilize proteins from cell membranes often cause destabilization or denaturation in the long run. Additionally, detergents appropriate for biochemical assays may inhibit protein crystallization [1, 2]. Careful screening is necessary, but is a time-consuming and expensive process. Finding an appropriate detergent has thus become a critical bottleneck not only for GPCR and other membrane protein studies, but also for designing and producing membrane proteins for biotechnological devices.

The limitations and problems encountered with traditional detergents highlight the need for a general class of detergents that can work with diverse membrane proteins. Several attempts have been made, including the design of amphipathic helical peptides, lipopeptides, amphipols, and tripod amphiphiles [3-8]. However, these surfactants are expensive, difficult to manufacture, or heterogeneous. Some, moreover, are ineffective with many proteins, or cannot maintain proteins soluble and functional for sufficient periods of time.

The Zhang Laboratory at MIT has designed a class of simple peptide surfactants as an alternative both to traditional detergents and the innovative strategies mentioned above (**Figure 5.1**). These peptides have a charged hydrophilic head made from one lysine or aspartic acid, and a 3-6 residue hydrophobic tail. All tail residues are composed of one of the following amino acids: glycine, alanine, valine, leucine, or isoleucine. Acetylation at the N-terminus, amidation at the C-terminus, or both, can be used to control the ionic nature of each peptide. This capping strategy allows the construction of cationic, anionic, and zwitterionic peptides. These peptide surfactants have defined critical aggregation concentrations (CAC), and form nanostructures including micelles, nanovesicles and nanotubes [9-13]. They also interact well with lipids to form monoolein bilayers [14]. More importantly, they have been shown to solubilize and

stabilize several diverse multi-transmembrane proteins, including Glycerol-3-phosphate dehydrogenase [15], photosystem I [16, 17], and a handful of GPCRs [18, 19].



**Figure 5.1. Molecular models of peptide surfactants at neutral pH.** A) Ac-AAAAAAD-COOH. B) Ac-AAAAAAK-CONH<sub>2</sub>. C) DAAAAAA-CONH<sub>2</sub>. D) KAAAAAA-CONH<sub>2</sub>. E) Ac-VVVD-COOH. F) Ac-VVVK-CONH<sub>2</sub>. G) Ac-IIID-COOH. H) Ac-IIIK-CONH<sub>2</sub>. I) Ac-LLLD-COOH. J) Ac-LLLK-CONH<sub>2</sub>. Aspartic acid (D) is negatively charged and lysine (K) is positively charged. The hydrophobic tails of the peptide surfactants consist of alanine (A), valine (V), isoleucine (I) and leucine (L). Each peptide is ~1.5--2.5 nm long, similar in size to biological phospholipids. Color code: green, carbon; red, oxygen; blue, nitrogen and white, hydrogen.

The success of using peptide surfactants to stabilize a small number of membrane proteins suggests that they may be capable of functionally stabilizing a larger number of receptors. To test this hypothesis, 12 olfactory receptors were expressed in the presence of 10 different peptide surfactants (**Figure 5.1**) using a cell-free expression system. The peptides' performance was compared to Brij-35, a common detergent that works best with cell-free GPCR expression (Chapter 4). All 10 peptides were able to solubilize all 12 ORs, demonstrating their potential as a class of surfactants for OR and other membrane protein studies.

## **5.2 Methods**

All methods are the same as those reported in Chapter 4, with the following exceptions.

### **5.2.1 Peptide Synthesis**

The peptides shown in **Figure 5.1** were synthesized and purified by CPC Scientific Inc., CA. The peptides, received in powder form, were dissolved in milli-Q water, sonicated, and adjusted to a pH value above 7.0 with NaOH or HCl to increase peptide solubility. The suspension was then filtered through a 0.22 $\mu$ m filter to remove insoluble particles and stored at room temperature.

### **5.2.2 Cell-Free Olfactory Receptor Expression and Purification**

The ORs were expressed in the manner described in Chapter 4, with the exception that peptide surfactants were added directly to the reaction mixture prior to incubation. The peptides were sonicated for 10 minutes prior to adding them to the reactions. The

final concentrations of each peptide used are shown in **Table 5.1**. Because the peptides elicit a circular dichroism (CD) signal that interferes with and overwhelms the protein signal, all purifications were performed using FC14. Because protein can be reused after a CD scan, the same receptor samples were used for CD and thermophoresis measurements.

**Table 5.1. Surfactant Peptide Properties and Experimental Concentrations**

Peptide (1-letter code)	Molecular weight (kDa)	Net charge at pH 7.0	Experimental Concentration	CAC in water (mM) <sup>1</sup>
Ac-VVVD	472.5	-2	0.625mM <sup>2</sup>	2.3
Ac-VVVK-CONH <sub>2</sub>	484.6	+1	2.5mM	1.6
Ac-IIID	514.6	-2	0.625mM <sup>2</sup>	1.0-1.2
Ac-IIIK- CONH <sub>2</sub>	526.7	+1	0.625mM	0.4-0.5
Ac-LLLD	514.6	-2	2.5mM	1.2
Ac-LLLK- CONH <sub>2</sub>	526.7	+1	2.5mM	1.2
Ac-AAAAAAD	601.6	-2	1.7 mM	0.3
Ac-AAAAAAK- CONH <sub>2</sub>	613.7	+1	1.6mM	0.2
DAAAAAA- CONH <sub>2</sub>	558.6	0	2.5mM	0.2
KAAAAAA- CONH <sub>2</sub>	571.7	+2	0.5mM	0.3

1. CAC—critical aggregation concentration. Two conventional methods, electrical conductivity and surface tension measurements, were used to determine the CACs of ten peptide surfactants in water. The peptides typically have 6-fold lower CACs in PBS.

2. Due to peptide solubility, higher concentrations were not possible. However, because the peptides typically have 6-fold lower CACs in PBS, the concentrations used for all experiments were estimated to be above their CAC.

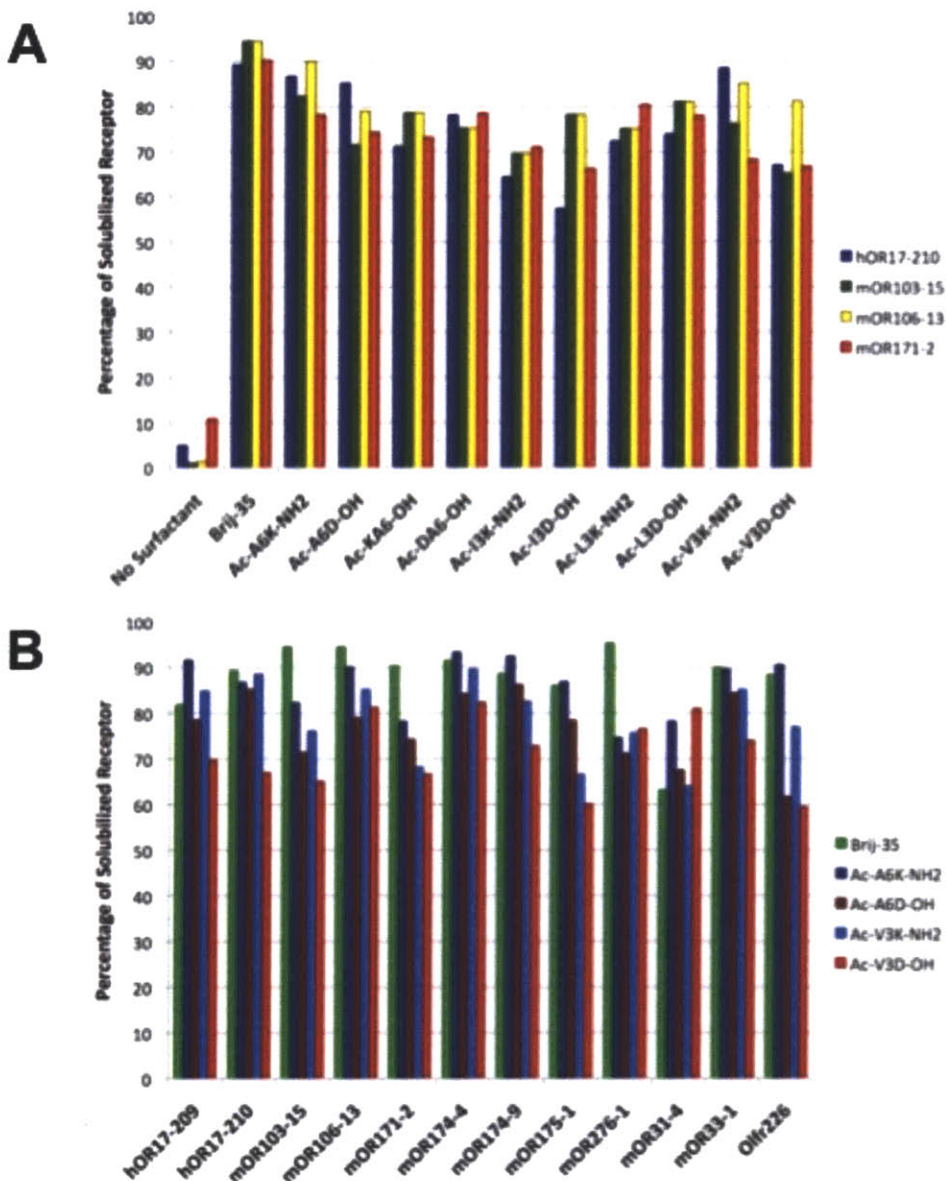
## 5.3 Results

### 5.3.1 Detergent Screening

Systematic screens were used to assess the ability of peptide surfactants to produce and solubilize 12 ORs in a commercial *E.coli*-based cell-free expression system. First, the

ability of diverse peptides to function as surfactants was tested. Four ORs were selected and produced in the cell-free system in the presence of all 10 peptides. The soluble and insoluble protein fractions were compared (**Figure 5.2A**). The surfactant Brij-35 was used as a control because Brij-35 is the optimal surfactant for producing ORs in the cell-free system (Chapter 4). Reactions with no surfactant served as additional controls. Second, the ability of peptides to solubilize a wide variety of ORs was tested by comparing the solubility of all 12 ORs in 4 peptide surfactants and Brij-35 (**Figure 5.2B**).

Western and dot blot analysis was used to compare the soluble and insoluble protein fractions. **Figure 5.2A** shows that a surfactant is necessary to solubilize the ORs. Without surfactant, only ~10% of the produced protein is soluble. With a proper surfactant, up to 95% of the expressed protein becomes soluble. The peptides and Brij-35 maintained similar fractions of ORs soluble. Brij-35 solubilized 63-95% of the expressed ORs, with most receptors having soluble fractions between 80-90%. The peptides solubilized 57-93% of the expressed ORs, with most ORs having fractions between 75-90%. **Figure 5.2B** shows that the 4 peptide surfactants solubilized ~60-90% of all 12 ORs, suggesting their potential use as a general class of surfactants for additional ORs and perhaps a wider range of GPCRs and other membrane proteins. **Figure 5.2B** also suggests that the peptide properties affect their performance as surfactants. The cationic peptides (Ac-A<sub>6</sub>K-CONH<sub>2</sub> and Ac-V<sub>3</sub>K-CONH<sub>2</sub>) solubilized a greater percentage of protein than their anionic counterparts (Ac-A<sub>6</sub>D-COOH and Ac-V<sub>3</sub>D-COOH). Also, peptides with longer hydrophobic tails yielded higher soluble fractions than peptides with shorter tails (Ac-A<sub>6</sub>K-CONH<sub>2</sub> v. Ac-V<sub>3</sub>K-CONH<sub>2</sub> and Ac-A<sub>6</sub>D-COOH v. Ac-V<sub>3</sub>D-COOH).

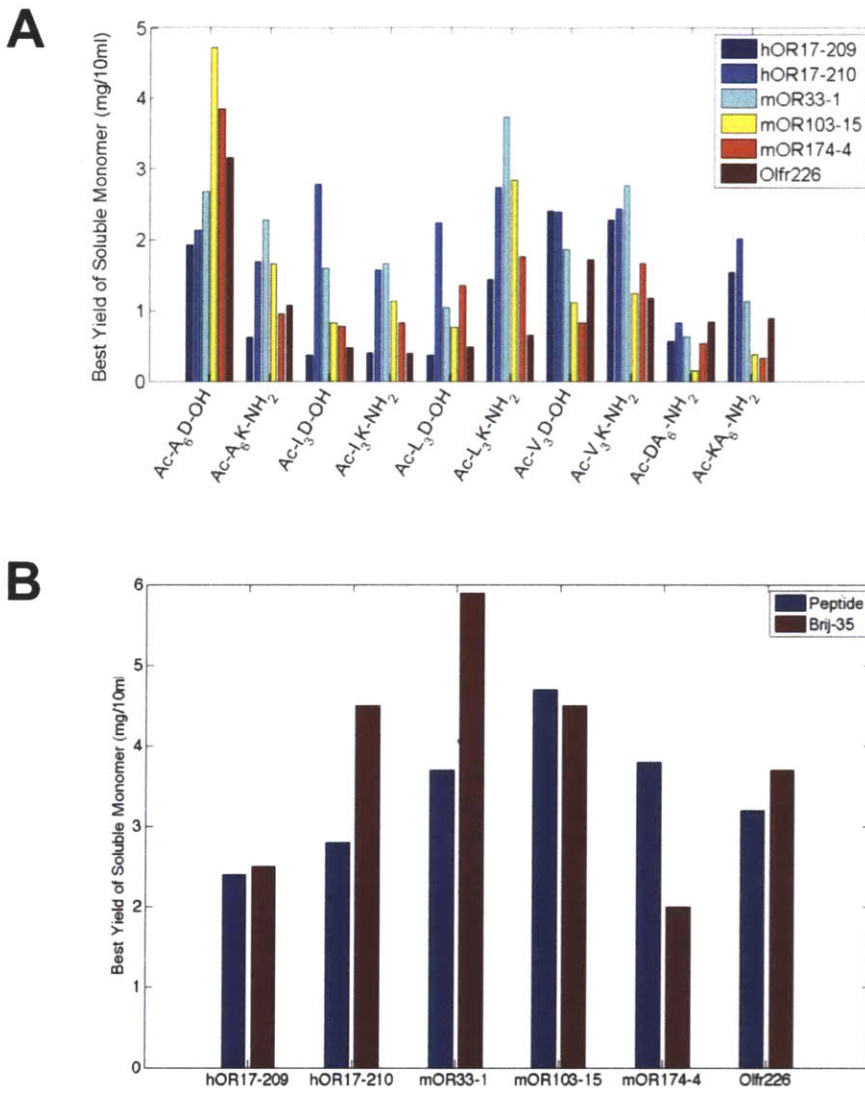


**Figure 5.2. Olfactory Receptor Solubility in Brij-35 and Peptide Surfactants.** Each OR was expressed in the presence of Brij-35 or a peptide surfactant using a commercial *E.coli* cell-free expression system. Upon completion of the reactions, the samples were centrifuged to separate solubilized receptor from insoluble aggregates. The soluble fraction was removed, and the pellet was resuspended in an equal volume of buffer. Soluble and precipitated protein samples were analyzed with Western and dot blots using an anti-rho-tag monoclonal antibody; relative sample intensities were used to calculate the percentage of solubilized receptor. As controls, reactions with no peptide or other detergent were assayed. All bars represent the average of two or three independent experiments. A) The presence of a detergent was necessary to solubilize the GPCRs, and all of the peptide surfactants were able to solubilize four unique ORs. B) The surfactant peptides and Brij-35 were able to solubilize similar fractions of protein. Peptides that were positively charged or had longer tails tended to solubilize higher fractions of receptors.



### 5.3.2 Determination of Receptor Yields

The maximum yields of olfactory receptors produced in the presence of the peptide detergents were estimated by comparing their band intensity to that of a receptor with a known concentration. Milligram quantities of receptor could be produced, demonstrating that cell-free synthesis in the presence of peptide detergents is a good alternative for large-scale olfactory receptor production, and perhaps for others GPCRs as well. However, as with the solubilization results, the protein yield depended on both the specific receptor and the specific peptide detergent used (**Figure 5.3A**). At least 2mg of hOR17-210 and mOR33-1 could be produced in a 10ml *E.coli* cell-free reaction with at least 5 of the peptide detergents. The peptide Ac-A<sub>6</sub>D-OH yielded the largest amount: ~4.8mg of mOR103-15 per 10ml reaction. Six other peptides were able to produce at least ~2.5mg for at least one OR: Ac-A<sub>6</sub>K-NH<sub>2</sub>, Ac-I<sub>3</sub>D-OH, Ac-L<sub>3</sub>D-OH, Ac-L<sub>3</sub>K-NH<sub>2</sub>, Ac-V<sub>3</sub>D-OH, and Ac-V<sub>3</sub>K-NH<sub>2</sub>. Overall, at least 2mg of most receptors could be synthesized in a 10ml reaction in the presence of at least one of the peptide detergents. These amounts were comparable to those that could be obtained using Brij-35 (**Figure 5.3B**).



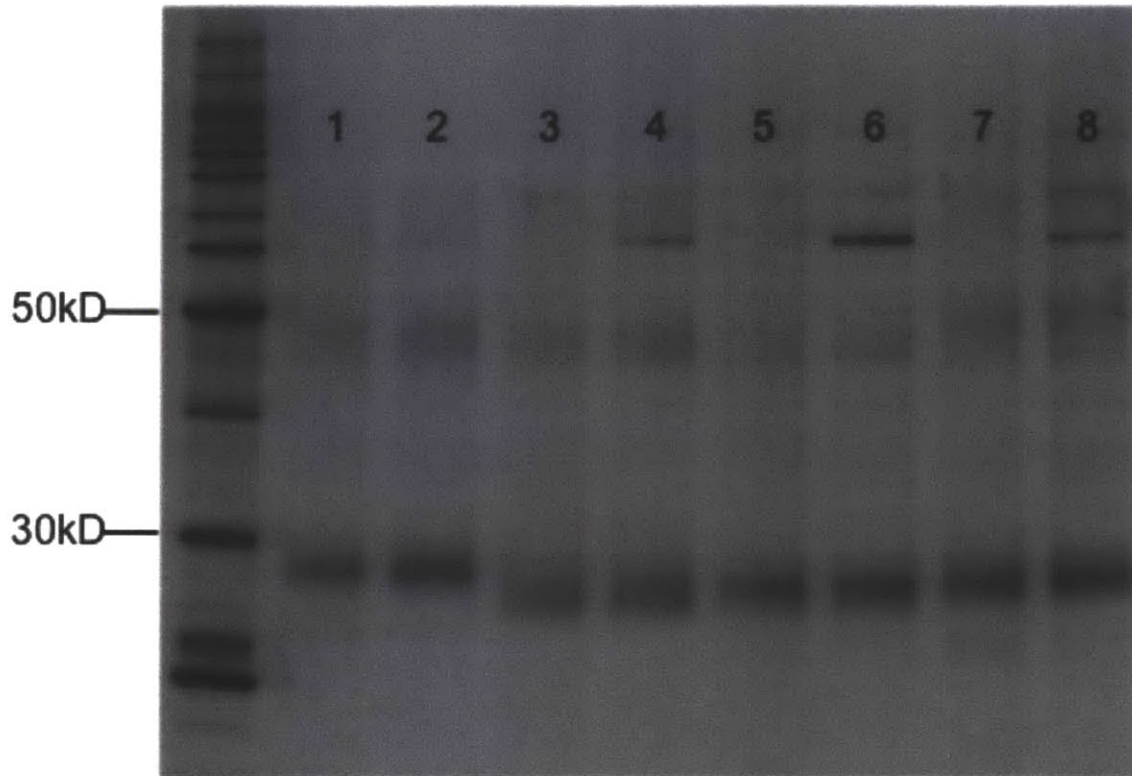
**Figure 5.3. Detergent Peptides Can Yield Milligram Quantities of Solubilized Olfactory Receptors.** The total receptor yield is dependent on the peptide used. A) The maximum expected yields of solubilized monomer for 6 of the receptors in the presence of each peptide. To determine the expected yields, solubilized receptor and protein with a known concentration were compared on a Western blot. The relative intensities of the known protein sample and the test samples were used to calculate the maximum receptor yields. B) The maximum yield of the monomeric form of some tested olfactory receptors expected in a 10ml reaction. Only results from the most effective detergent peptide are shown. The total protein yield is dependent on the peptide detergent used. These results are compared to the yields of receptors made in Brij-35. In most cases, the yields are comparable. For two receptors, Brij-35 resulted in more expressed protein, and for one receptor a peptide resulted in more.

### 5.3.3 Olfactory Receptor Purification and Purity Analysis

Four olfactory receptors were selected for larger scale expression and purification for structural and functional analysis. The olfactory receptors mOR103-15, mOR174-4, mOR174-9, and Olfr226 were expressed using Brij-35 or a peptide detergent, and purified using the rho1D4 monoclonal antibody. The purifications were performed in the presence of fos-choline 14 (FC14) because it has been shown to be the preferred detergent for olfactory receptor purification [20-22], and because a common buffer for the Brij-35- and peptide-produced receptors allows for direct comparison between the two samples. Moreover, a detergent exchange was necessary in order to carry out subsequent analyses, as the peptides elicit signals that are difficult to distinguish from the receptor signals. Because a simple detergent exchange is unlikely to re-fold misfolded receptors, the structures of the olfactory receptors initially produced in Brij-35 or a peptide should not be positively affected by the change to FC14.

The purified receptors were analyzed on a gel (**Figure 5.4**). The olfactory receptors produced in both the Brij-35 and peptide detergents were the same size, and exhibited the expected monomeric and dimeric bands. These results suggest that the peptide detergents do not interfere with full-length cell-free protein expression. They further suggest that the peptide detergents do not interfere with proper receptor folding or structure, as the expressed protein shows the tendency to dimerize. Silver staining showed that receptors expressed in Brij-35 could be purified up to ~90% purity, while receptors expressed in a peptide could be purified up to >80% purity, using immunoaffinity chromatography alone. The lower purity of the peptide-produced samples is due primarily to the presence of 2 bands around 60 and 80kDa. It is possible that these bands are impurities in the sample, or result from ORs aggregating during the gel electrophoresis. However, since the peptides can form various mesoscale

structures, it is possible that the extra bands represent different states of the peptide-OR complexes.

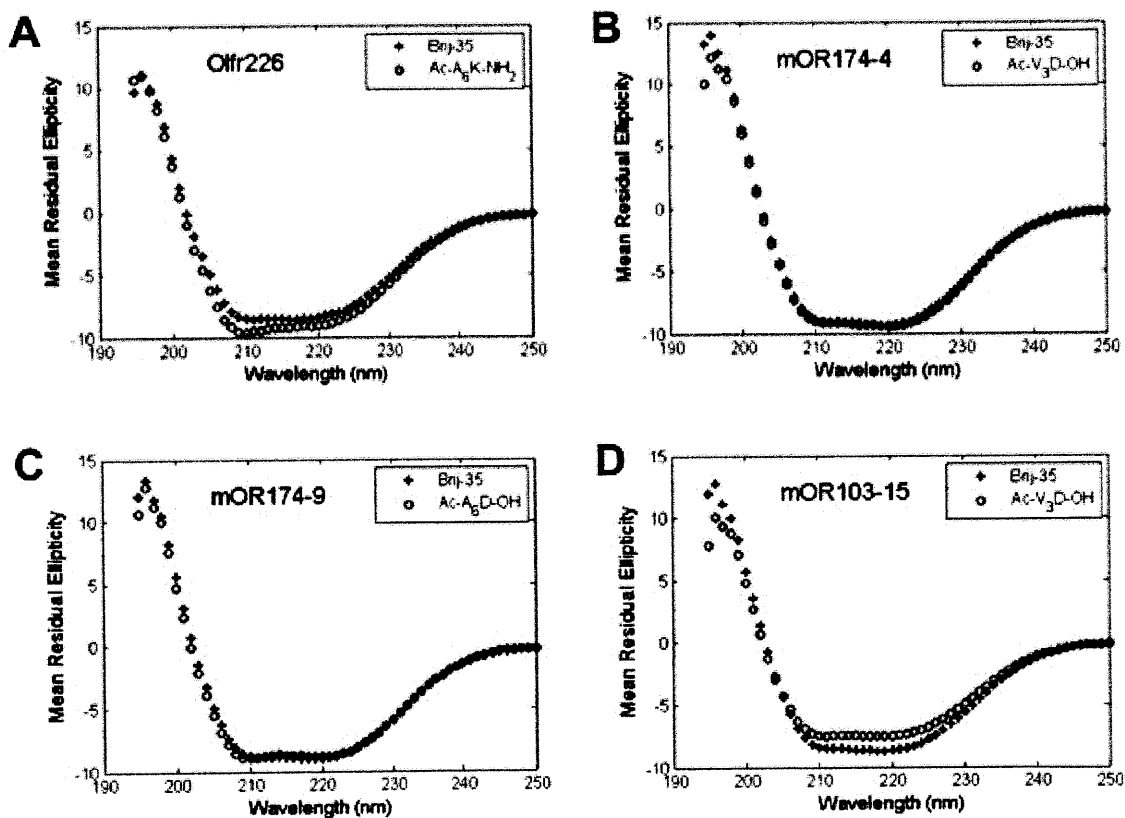


**Figure 5.4. Silver Stain of Four Purified Olfactory Receptors Produced in Either Brij-35 or a Peptide.** Lane 1: mOR103-15 produced in Brij-35. Lane 2: mOR103-15 produced in Ac-V<sub>3</sub>K-CONH<sub>2</sub>. Lane 3: mOR174-4 produced in Brij-35. Lane 4: mOR174-4 produced in Ac-V<sub>3</sub>K-CONH<sub>2</sub>. Lane 5: mOR174-9 produced in Brij-35. Lane 6: mOR174-9 produced in Ac-A<sub>6</sub>D-OH. Lane 7: Olfr226 produced in Brij-35. Lane 8: Olfr226 produced in Ac-A<sub>6</sub>K-CONH<sub>2</sub>. The Brij-35 and peptide-produced samples have similar purities. Also, all of the ORs are the same size show the same tendency to dimerize. This indicates that both types of surfactants function comparably, and that the peptides do not interfere with full-length protein translation, or proper receptor folding or function.

#### 5.3.4 Secondary Structural Analysis Using Circular Dichroism

Circular dichroism was used to assess the secondary structure of the 4 purified olfactory receptors. The purification and CD analyses of olfactory receptors were performed in FC-14 because the peptides themselves have strong CD signals that interfere with and overwhelm

the receptor signals. **Figure 5.5** directly compares the CD spectra of ORs produced in either Brij-35 or a peptide detergent. All purified receptors have characteristic  $\alpha$ -helical spectra, with signature valleys at 220 nm and 208 nm. Because GPCRs have seven transmembrane  $\alpha$ -helical domains, these spectra indicate that the receptors are properly folded. Moreover, the nearly superimposed spectra for ORs produced in Brij-35 or a peptide surfactant indicate that the peptide surfactants are able to structurally stabilize these olfactory receptors equally as well as traditional surfactants.

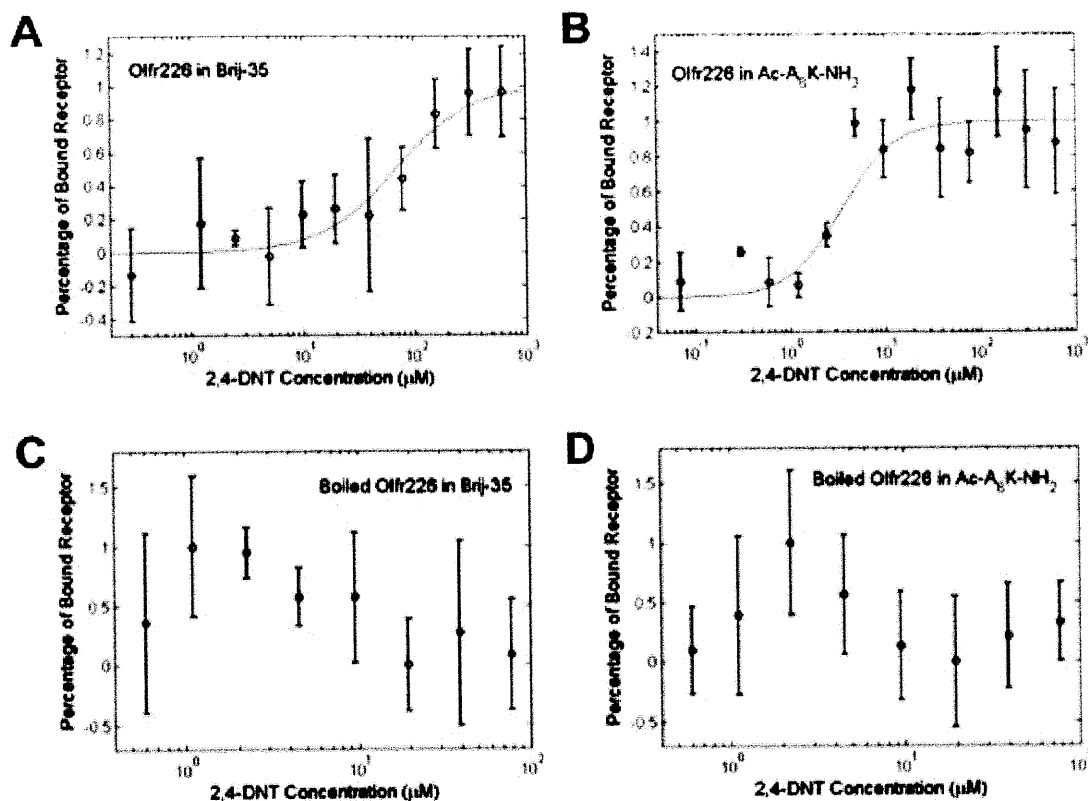


**Figure 5.5 CD spectra of Brij-35 and Peptide Surfactant-Produced Olfactory Receptors.** A) Olfr226, B) mOR174-4, C) mOR174-9, and D) mOR103-15. All eight samples have characteristic secondary  $\alpha$ -helical spectra, suggesting that the receptors are properly folded. The near overlap of the peptide and detergent CD curves suggests that, for olfactory receptors, the peptides function as detergents as effectively as traditional detergents.

### 5.3.5 Ligand-Binding Analysis Using Microscale Thermophoresis

Microscale thermophoresis was used to determine whether the expressed and purified proteins were functional. This approach is based on the ligand binding-induced change in movement of molecules along a temperature gradient [23, 24], and is capable of detecting interactions with ligands as small as calcium ions [25]. Because olfactory receptors are larger than 35kDa and their ligands are smaller than 300Da, it was necessary to use thermophoresis to detect ligand binding instead of less sensitive methods like SPR or quartz crystal microbalance. All 8 receptors used with CD were analyzed with their known odorants (**Table 4.1**). Boiled receptors were used as controls.

**Figure 5.6** and **Table 5.2** show the results of the Brij-35- and peptide surfactant-solubilized ORs, and their boiled controls. All of the protein samples exhibited a typical sigmoidal binding curve. In contrast, the boiled controls had random amplitudes throughout the ligand titration range. These results indicate that all of the ORs bound their respective odorants. The large noise in the boiled controls is probably due to the presence of protein aggregates of different size, and hence various diffusive and thermophoretic properties. Olfr226 produced in peptide exhibited a significantly higher affinity for its ligand than receptor produced in Brij-35. The other ORs had similar binding affinities in both types of surfactants. These results show that both classes of detergents are able to aid in the production and solubilization of functional olfactory receptors, and that peptides may confer more functional stability to some solubilized receptors. The measured  $EC_{50}$  value for each tested olfactory receptor is in the micromolar range, which is consistent with previous reports [21, 22].



**Figure 5.6. Binding of Olfactory Receptor Olfr226 to its Ligand 2,4-DNT:** A) Olfr226 produced in Brij-35, B) Olfr226 produced in Ac-A<sub>6</sub>K-NH<sub>2</sub>, C) Olfr226 produced in Brij-35 and boiled, and D) Olfr226 produced in Ac-A<sub>6</sub>K-NH<sub>2</sub> and boiled. Olfactory receptors produced in Brij-35 or a peptide surfactant exhibit typical sigmoidal binding curves. There is one plateau at low concentrations and another at high concentrations, while the boiled controls have no plateaus. This suggests that the thermophoresis signals are measuring ligand binding. All curves were normalized to the fraction of bound receptor. Open circles show the mean measurements from 3 experiments; the lines through the points are the best-fit curves using the Hill equation to obtain the half maximal effective concentration EC<sub>50</sub>. The binding results shown here are representative of the data from all four OR samples.

**Table 5.2. Ligand Binding Affinities For Peptide- and Brij-35- Produced Olfactory Receptors**

Receptor	Surfactant Used	Ligand	Measured EC <sub>50</sub> (μM)
mOR103-15	Brij-35	Heptanal	2 ± 0.7
mOR103-15	Ac-V <sub>3</sub> D-OH	Heptanal	0.9 ± 0.2*
mOR174-4	Brij-35	Ethyl Vanillin	7 ± 2
mOR174-4	Ac-V <sub>3</sub> D-OH	Ethyl Vanillin	4 ± 2
mOR174-9	Brij-35	Ethyl Vanillin	4.9 ± 3.5
mOR174-9	Ac-A <sub>6</sub> D-OH	Ethyl Vanillin	5.1 ± 2
Olfr226	Brij-35	2,4-DNT	86 ± 36
Olfr226	Ac-A <sub>6</sub> K-NH <sub>2</sub>	2,4-DNT	3 ± 1.6

\*This measurement was obtained at a longer time (25s vs. 15s) and lower IR-laser power (1.5V v. 2.5V) than the other measurements.

## 5.4 Discussion

This study showed that short peptide detergents solubilized and functionally stabilized cell-free produced olfactory receptors equally as well as the detergent Brij-35. All of the tested peptide detergents were able to solubilize all of the tested olfactory receptors. Soluble olfactory receptor fractions were as high as 93%. We previously reported lower solubilities for hOR17-210 and mOR103-15 in the presence of some peptides [19]. This difference in solubility is probably due to differences in the peptide batches, or the dynamic nature of the peptides. Previous reports have noted the peptides' ability to form various mesoscale structures, and to change between structures over time. It is probable that specific structures are better able to solubilize the expressed receptors,



and more research needs to be done to elucidate these effects. However, the comparable results between the peptides and Brij-35 for 12 olfactory receptors reported here demonstrates their potential to be used as detergents for membrane protein studies.

The results in this study show that the efficacy of solubilization primarily depends on the peptide detergent properties. **Figure 5.2B** shows that the cationic peptides usually solubilized a greater fraction of expressed protein than their anionic counterparts. This effect was more pronounced for the longer peptides (12 out of 12 olfactory receptors with Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-A<sub>6</sub>D-OH) than for the shorter peptides (10 out of 12 olfactory receptors with Ac-V<sub>3</sub>K-NH<sub>2</sub> and Ac-V<sub>3</sub>D-OH). **Figure 5.2A** indicates that this effect may depend on the specific receptor. The receptors hOR17-210 and mOR171-2 were consistently more soluble in cationic peptides, while the receptors mOR103-15 and mOR106-13 were often more soluble in anionic peptides. Because mOR103-15 and mOR106-13 had greater solubility in the more hydrophobic anionic peptides (leucine and isoleucine tails), it is probable that the tail composition can alter tendencies caused by the head group properties. Additionally, peptides with longer hydrophobic tails typically had higher soluble fractions than those with the same charge but shorter tails (10 receptors for the cationic peptides, 9 for the anionic) (**Figure 5.2B**). **Figure 5.3A** further suggests that subtle changes in the residue order may affect how the peptides interact with the olfactory receptors expression. Although composed of the same amino acids, Ac-A<sub>6</sub>D-OH yielded significantly more protein than Ac-DA<sub>6</sub>-NH<sub>2</sub>. A similar result was observed with most tested olfactory receptors with Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-KA<sub>6</sub>-NH<sub>2</sub>. These results further indicate the peptide charge or ionic character can greatly affect olfactory receptor solubility or

expression. Additional experiments will be needed to fully characterize and understand the effects of peptide properties on olfactory receptor solubilization and production.

CD and microscale thermophoresis demonstrated that peptide- and Brij-35-produced ORs had similar structures and binding affinities. All of the purified ORs had characteristic  $\alpha$ -helical spectra, which is expected for the 7-transmembrane helix members of the GPCR family. Moreover, the spectra of the Brij-35 and peptide-purified ORs nearly overlapped. Because even small changes in secondary structure are detectable using CD, these spectra indicate that the peptides were able to aid in the proper folding of expressed ORs equally as well as Brij-35. Slight differences in the spectra are probably the result of differences in sample purity, or the presence of residual peptides after the purification process. The peptide-produced proteins had similar or higher-affinity binding constants than the Brij-35-produced proteins. The higher binding constant indicates tighter binding, which could be a result of a more stable receptor conformation. This suggests that the peptide surfactants may be a useful tool in crystallizing GPCRs, and may also be able to aid in the development of GPCR-based technological devices. Further studies are necessary to determine the lifetime of a functional receptor in a surfactant peptide. However, the results reported here, as well as previous studies [15-19], indicate that the peptides may be able to fulfill just such a function, perhaps better than many traditional detergents. Furthermore, these findings are particularly important because they suggest that the peptides are a general class of detergents that can be used with cell-free expression methods. Although cell-free production is a mature technology for soluble proteins [26-28], very few membrane proteins have been produced through this technique, and even then only through laborious surfactant screens [22, 29-32]. The methods reported here

may help accelerate the production of many more membrane proteins for structural studies and biotechnological advancements.

The peptide surfactants described here may be useful for diverse membrane protein studies. They function comparably to traditional surfactants, and offer several advantages over other novel surfactants. Their properties are similar to commonly used detergents, they can be systematically designed and economically produced at high purity, and they remain stable for long periods of time. The ability of every tested peptide to solubilize 12 ORs from three different species, and of 3 peptides to functionally stabilize 4 ORs further suggests that they may be a general class of surfactants capable of functionally solubilizing a wider range membrane proteins. Further studies are needed to characterize surfactant peptide property effects on membrane proteins; it may be possible to rationally design a surfactant optimal for a given protein. Future studies are also needed to analyze the long-term stability of membrane proteins solubilized in the peptide surfactants. However, this study suggests that peptide surfactants are promising for membrane protein studies, that they could lead to a better understanding of olfactory receptors and other GPCRs, and also could be used in the design and fabrication of diverse OR-based biotechnological devices.

## 5.5 References

1. Garavito RM, Picot D, Loll PJ. Strategies for crystallizing membrane proteins. *J Bioenerg Biomembr* **28**, 13-27 (1996).
2. Prive CG. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
3. Schafmeister CE, Miercke LJ, Stroud RM. Structure at 2.5Å of a designed peptide that maintains solubility of membrane proteins. *Science* **262**, 734-738 (1993).
4. McGregor CL, *et al.* Lipopeptide detergents designed for the structural study of membrane proteins. *Nat Biotechnol* **21**, 171-176 (2003).
5. Popot JL, *et al.* Amphipols: Polymeric surfactants for membrane biology research. *Cell Mol Life Sci* **60**, 1559-1574 (2003).
6. Thiesen MJ, Potocky TB, McQuade DT, Gellman SH, Chiu ML. Crystallization of bacteriorhodopsin solubilized by a tripod amphiphile. *Biochim Biophys Acta* **1751**, 213-216 (2005).
7. Tribet C, Audebert R, Popot JL. Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci USA* **93**, 15047-15050 (1996).
8. Yu SM, *et al.* An improved tripod amphiphile for membrane protein solubilization. *Protein Science* **9**, 2518-2527 (2000).
9. Vauthey S, Santoso S, Gong H, Watson N, Zhang S. Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles *Proc Natl Acad Sci USA* **99**, 5355-5360 (2002).
10. Santoso S, Hwang W, Hartman H, Zhang S. Self-assembly of surfactant-like peptides with variable glycine tails to form nanotubes and nanovesicles. *Nano Letters* **2**, 687-691 (2002).
11. von Maltzahn G, Vauthey S, Santoso S, Zhang S. Positively charged surfactant-like peptides self-assemble into nanostructures. *Langmuir* **19**, 4332-4337 (2003).
12. Nagai A, Nagai Y, Qu H, Zhang S. Self-assembling behaviors of lipid-like peptides A<sub>6</sub>D and A<sub>6</sub>K. *J Nanoscience & Nanotechnology* **7**, 2246-2252 (2007).
13. Khoe U, Yanlian Y, Zhang S. Self-assembly of nano-donut structure from cone-shaped designer lipid-like peptide surfactant. *Langmuir* **25**, 4111-4114 (2009).
14. Yaghmur A, Laggner P, Zhang S, Rappolt M. Tuning curvature and stability of monoolein bilayers by designer lipid-like peptide surfactants. *PLoS ONE* **2**, e479 (2007).

15. Yeh JI, Du S, Tordajada A, Paulo J, Zhang S. Peptergent: peptide detergents that improve stability and functionality of a membrane protein glycerol-3-phosphate dehydrogenase. *Biochemistry* **44**, 16912-16919 (2005).
16. Kiley P, Zhao X, Bruce BD, Baldo M, Zhang S. Self-assembling peptide detergents stabilize isolated photosystem I on a dry surface for an extended time. *PloS Biology* **3**: 1181-1186 (2005).
17. Matsumoto K, Koutsopoulos S, Vaughn M, Bruce BD, Zhang S. Designer lipid-like peptide surfactants stabilize functional Photosystem I membrane complex in solution. *J Phys Chem* **115**, 75-83 (2009).
18. Zhao X, et al. Designer lipid-like peptides significantly stabilize G-protein coupled receptor bovine rhodopsin. *Proc Natl Acad Sci USA* **103**, 17707-17712 (2006).
19. Wang X, et al. Peptide Surfactants for Cell-Free Production of Functional G Protein-Coupled Receptors. *Proc Natl Acad Sci USA*: in press (2011).
20. 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. *PLoS One* **3**, e2920 (2008).
21. Cook BL, et al. Large scale production and study of a synthetic G-protein coupled receptor: Human olfactory receptor 17-4. *Proc Natl Acad Sci USA* **106**, 11925-11930 (2009).
22. Kaiser L, et al. Efficient cell-free production of olfactory receptors: Detergent optimization, structure, and ligand binding analyses. *Proc Natl Acad Sci USA* **105**(41), 15726-15731 (2008).
23. Baaske P, Wienken CJ, Reineck P, Duhr S, Braun D. Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed* **49**, 2238-2241(2010).
24. Duhr S, Braun D. Why molecules move along a temperature gradient. *Proc Natl Acad Sci USA* **103**, 19678-19682 (2006).
25. Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein Binding Assays in Biological Liquids using Microscale Thermophoresis. *Nat Comm* **1**,100 (2010).
26. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YBA. Continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162-1164 (1988).
27. Yokoyama S. Protein expression systems for structural genomics and proteomics. *Curr Opin Chem Biol* **7**, 39-43 (2003).

28. Endo Y, Sawasaki T. Cell-free expression systems for eukaryotic protein production. *Curr Opin Biotechnol* **17**, 373-380 (2006).
29. Ishihara G, *et al.* Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* **41**, 27-37 (2005).
30. Klammt C, *et al.* Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* **272**, 6024-6038 (2005).
31. Savage DF, Anderson CL, Robles-Colmenares Y, Newby ZE, Stroud RM. Cell-free complements in vivo expression of the E. coli membrane proteome. *Protein Sci* **16**, 966-976 (2007).
32. Klammt C, *et al.* Cell-free Production of G Protein-Coupled Receptors for Functional and Structural Studies. *J Struct Biol* **158**, 482-493 (2007).

# CHAPTER 6

## EVALUATION OF THE EFFECT OF SURFACTANT PEPTIDE PHYSICAL AND CHEMICAL PROPERTIES ON OLFACTORY RECEPTOR SOLUBILITY

### 6.1 Introduction

Chapter 5 demonstrates the usefulness of peptides as surfactants for GPCR studies. The data in **Figure 5.2** further suggests that the efficacy of solubilization depends in part on the peptide surfactant properties. **Figure 5.2B** shows that the cationic peptides usually solubilized a greater fraction of expressed protein than their anionic counterparts. This effect was more pronounced for the longer peptides (12 out of 12 ORs with Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-A<sub>6</sub>D-OH) than for the shorter peptides (10 out of 12 ORs with Ac-V<sub>3</sub>K-NH<sub>2</sub> and Ac-V<sub>3</sub>D-OH). **Figure 5.2A** suggests that this effect may depend on the specific receptor. The receptors hOR17-210 and mOR171-2 were consistently more soluble in cationic peptides, while the receptors mOR103-15 and mOR106-13 were often more soluble in anionic peptides. Because mOR103-15 and mOR106-13 had greater solubility in the more hydrophobic anionic peptides (leucine and isoleucine tails), it is reasonable to hypothesize that the tail composition can alter tendencies caused by the head group properties. Additionally, peptides with longer hydrophobic tails typically had higher soluble fractions

than those with the same charge but shorter tails (10 receptors for the cationic peptides, 9 for the anionic) (**Figure 5.2B**). **Figure 5.3A** further suggests that subtle changes in the residue order may affect how the peptides interact with OR expression. Although composed of the same amino acids, Ac-A<sub>6</sub>D-OH yielded significantly more protein than Ac-DA<sub>6</sub>-NH<sub>2</sub>. A similar result was observed with most ORs tested with Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-KA<sub>6</sub>-NH<sub>2</sub>. These results further suggest that the peptide charge or ionic character can greatly affect OR solubility or expression.

This study examined the hypothesis that peptide surfactant properties affect GPCR solubility. A set of peptides was designed to test the following potential variables: tail length, tail hydrophobicity, peptide flexibility, residue volume, and head group charge. Six different GPCRs were expressed in the set of peptides, and the insoluble and soluble protein fractions were compared. Surprisingly, the hydrophobicity, flexibility, and volume of the tail residues had no effect on GPCR expression or solubility. Moreover, despite the suggestive results reported above, the tail length and head group charge had negligible effects on GPCR solubility. This data suggests that although some physical properties may affect the peptide surfactants ability to function as detergents, these effects are not significant enough to warrant the careful design of peptides for a specific receptor. This further suggests that these simple peptides could be broadly useful as a general class of mild detergents for membrane protein studies.



## 6.2 Methods

### 6.2.1 Amino Acid Selection

Two sets of peptides were constructed. The first was designed to test the variables of tail hydrophobicity, flexibility, and residue volume. The second tested the variables of tail length and head group charge.

The first set of peptides consisted of Ac-G<sub>3</sub>K-NH<sub>2</sub>, Ac-A<sub>3</sub>K-NH<sub>2</sub>, Ac-V<sub>3</sub>K-NH<sub>2</sub>, Ac-I<sub>3</sub>K-NH<sub>2</sub>, and Ac-L<sub>3</sub>K-NH<sub>2</sub> (**Table 6.1**). The head group amino acid and tail length were held constant, while the residues comprising the tail were varied. All 5 peptides were needed in order to simultaneously test the effects of hydrophobicity, flexibility, and volume because some of the variables are difficult to decouple. In general, as the volume of the side chain increases, the hydrophobicity and rigidity also increase (**Table 6.1**). However, because of the beta-branching of its residue side chain, isoleucine is significantly more flexible than leucine, even though both residues have the same composition and volume (**Figure 5.1, Table 6.1**) [1]. Moreover, although it has a larger volume, isoleucine is also significantly more flexible than the beta-branched valine. Thus, if residue volume is a critical variable, GPCR solubility should increase or decrease as the tail residue is changed from glycine to isoleucine, but should be the same for leucine and isoleucine. However, if tail flexibility is a critical parameter, solubility in isoleucine should be comparable to alanine (**Table 6.1**). Although hydrophobicity generally increases as the residue volume increases, valine has a higher hydrophobicity than leucine, and a similar hydrophobicity to isoleucine [2], possibly due to the bulk of the side chain residing closer to the backbone, thus shielding it from aqueous solution (**Table 6.1**). Thus, if hydrophobicity is a critical parameter, GPCR

solubility should increase or decrease as the residue volume increases, with the exception that solubility in valine should most similar to that in isoleucine.

The second set of peptides consisted of two subsets. The first subset included Ac-A<sub>3</sub>K-NH<sub>2</sub>, Ac-A<sub>4</sub>K-NH<sub>2</sub>, Ac-A<sub>5</sub>K-NH<sub>2</sub>, Ac-A<sub>6</sub>K-NH<sub>2</sub>, while the second subset included Ac-A<sub>4</sub>D-OH, Ac-A<sub>5</sub>D-OH, and Ac-A<sub>6</sub>D-OH. Within each subset the tail length varied, while all other parameters remained constant. Data from the two subsets were compared in order to examine the importance of head group charge. The peptide Ac-A<sub>3</sub>K-NH<sub>2</sub> was included in the first subset as a control, so that the data from all of the peptide sets could be compared. Data from all of the peptides should elucidate not only which parameters affect solubility, but also their relative importance.

All other methods are as described in Chapters 4 and 5.

**Table 6.1. Amino Acid Properties**

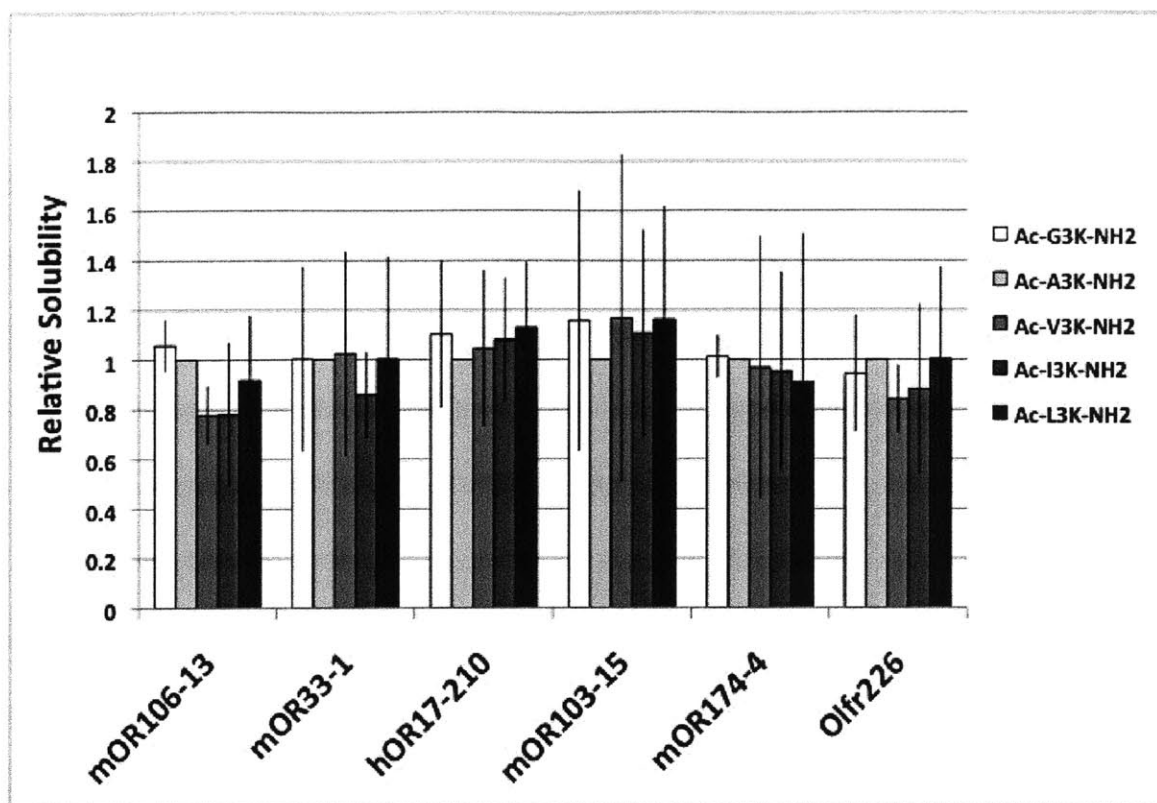
Amino Acid	Volume (A <sup>3</sup> ) [3]	Hydrophobicity Index [2]	Flexibility (10 <sup>6</sup> s <sup>-1</sup> ) [1]
Glycine	60.1	-0.4	39
Alanine	88.6	1.8	18
Valine	140.0	4.2	3
Leucine	166.7	3.8	2.3
Isoleucine	166.7	4.5	10

## 6.3 Results

### 6.3.1 Effects of Residue Hydrophobicity, Flexibility, and Volume

The receptors mOR33-1, hOR17-210, mOR103-15, mOR174-4, mOR106-13, and Olfr226 were tested with the first set of peptides. **Figure 6.1** shows the solubility of each receptor in each peptide. Surprisingly, no difference in solubility was observed for any given receptor in any of the peptides. Because hydrophobicity of the tail influences its

tendency to interact with the GPCR, we would expect an optimal hydrophobicity to exist for a given head group. If the surfactant isn't hydrophobic enough, it shouldn't be able to sufficiently shield the transmembrane segments of the expressed protein from aqueous solution. If the surfactant tail is too hydrophobic, it might not be able to form micelles capable of solubilizing the protein, and might instead form peptide aggregates that precipitate out of solution. However, it is possible that hydrophobicity is an important parameter in general, but that differences in the peptides are not sufficient to yield observable results within the range of parameters tested. Similarly, one could expect flexibility to be an important parameter, as a flexible surfactant may be better able to mold itself to the contours of the expressed GPCR. However, it is possible that flexibility is an important parameter affecting GPCR solubility. Since the peptide tails were only 3 amino acids long, differences in their flexibility may have been negligible. It is possible that differences in solubility may be observed with longer tails. Because no differences were observed in these six receptors, additional receptors were not tested.



**Figure 6.1: Peptide Volume, Flexibility, and Hydrophobicity Do Not Affect GPCR Solubility.** The solubility of 6 different ORs was examined in peptides with different tail compositions. These tails had different volumes, flexibilities and hydrophobicities. The tail length and head group remained constant. All solubilities were normalized to the peptide Ac-A<sub>3</sub>K-NH<sub>2</sub>, which was defined to have a relative solubility of 1. No significant differences in OR solubility were observed across the different peptides.

### 6.3.2 Effects of Peptide Length and Head Group Properties

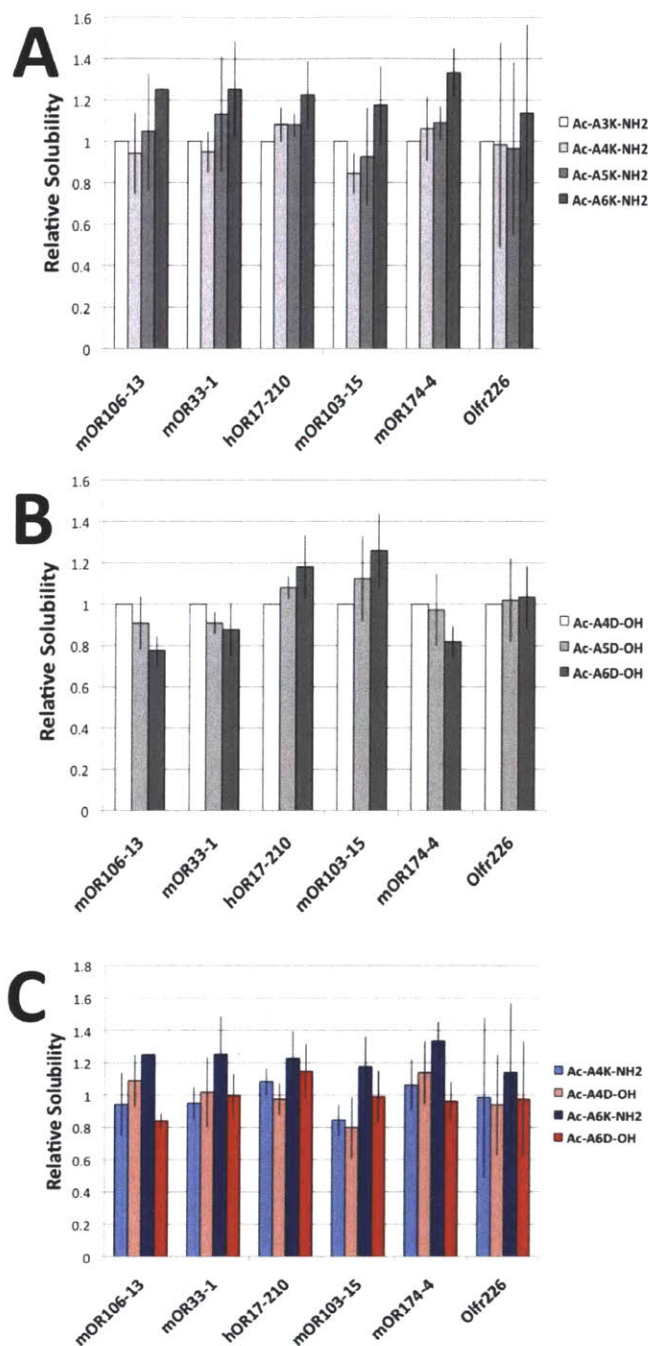
The same 6 GPCRs were tested with the second set of peptides. **Figure 6.2A** shows that solubility tends to increase as the tail length increases for positively charged peptides. In contrast, **Figure 6.2B** shows that solubility for half of the GPCRs tends to decrease with increasing tail length for negatively charged peptides. However, the reverse trend was observed for hOR17-210 and mOR103-15, while no trend was observed for Olfr226. **Figure 6.2C** shows that the cationic peptides tended to solubilize a greater amount of receptor than anion peptides, and that this effect was more pronounced as the tail length increased. This behavior is similar to that of traditional detergents, in that tail length and

head group charge affect protein stability. In general, detergents with longer tails are better at stabilizing proteins, while detergents with larger head groups are more soluble [4, 5, **Table 1.2**]. For the detergents with a lysine head, the increase in solubility with increasing tail length likely reflects the ability of the longer tail to better stabilize and hence solubilize the proteins. The different trend observed with the peptides with an aspartic acid head may reflect the difference in the size of the head group. Lysine is larger, and perhaps is better at keeping the peptide micelles soluble. It may also promote the formation of larger micelles: steric hindrance of the head group may make micelle formation of shorter peptides less favorable. Because the tighter radius of curvature can provide a harsh environment for membrane proteins, their solubility may increase as the micelle size increases. Indeed, it's likely that the ratio between the head group size and tail size may determine whether a particular peptide forms micelles optimal for GPCR solubilization. **Table 6.2** shows that these ratios are similar for Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-A<sub>4</sub>D-OH, and that they either increase or decrease away from these two points. Because these two peptides usually produced the most soluble receptor, these ratios strongly suggest that an optimal mechanical configuration exists. Indeed, the ratio between the head and tail sizes may indicate the tail length likely to result in optimal micelle packing for GPCR studies for a given head group, and any given amino acid used for the tail. However, **Figure 6.2C** further suggests that a peptide surfactants' head group properties will affect its efficacy once its geometry has been optimized.

**Table 6.2. Size Ratios Between Peptide Surfactant Heads and Tails**

<b>Peptide Surfactant</b>	<b>Ratio of Head Group Volume: Tail Length</b>	<b>Ratio of Head Group Volume: Tail Volume</b>
Ac-A <sub>3</sub> K-NH <sub>2</sub>	9.5	0.64
Ac-A <sub>4</sub> K-NH <sub>2</sub>	7.6	0.48
Ac-A <sub>5</sub> K-NH <sub>2</sub>	6.3	0.38
Ac-A <sub>6</sub> K-NH <sub>2</sub>	5.4	0.32
Ac-A <sub>4</sub> D-OH	5.0	0.32
Ac-A <sub>5</sub> D-OH	4.1	0.25
Ac-A <sub>6</sub> D-OH	3.6	0.21

All together, the data shows that solubility increases with tail length with the positively charged peptides, and tends to decrease with the negatively charged peptides. The peptide charge is unlikely to be an important factor independently of tail length. If this were the case, the same trends would be expected in both subsets of peptides, with one set consistently showing greater solubility at each tail length. However, it should be noted that the differences in solubility among all of the tested peptides are slight, and most show no statistically significant differences. This contrasts greatly with the results observed with Brij-35, which was 4-5 times better than the next best detergent. Thus, although tail length and head group size do affect peptide solubility, they are not critical factors of consideration.



**Figure 6.2. Peptide Tail Length and Head Group Composition Negligibly Affect OR Solubility.** A) Peptides with a cationic head group and tails varying in length between 3 to 6 amino acids were tested with six ORs. B) Peptides with an anionic head group and tails varying in length between 4 to 6 amino acids were tested with the same set of ORs. C) Comparison of long and short cationic and anionic peptides. When the head group is positively charged, solubility increases with increasing tail length. When the head group is negatively charged, the solubility either increases or decreases with increasing tail length, and the specific trend depends on the receptor. Cationic peptides tend to solubilize a greater percentage of receptor than their anionic counterparts, and this effect is more pronounced as the tail length increases. However, differences in solubility among all of the samples are small. In A) and B), the data was normalized to the shortest peptide (Ac-A<sub>3</sub>K-NH<sub>2</sub> and Ac-A<sub>4</sub>D-OH, respectively). In C), the data was normalized to Ac-A<sub>3</sub>K-NH<sub>2</sub>.

## 6.4 Discussion

This study indicates that the length of the peptide surfactant tail and the size of the head group can affect GPCR solubility. When the head group was composed of lysine, a residue with a large side chain, GPCR solubility tended to increase with tail length. In contrast, when the head group was composed of aspartic acid, a residue with a smaller side chain, no common trend was observed. Two of the GPCRs were more soluble as the tail length increased, while three were less soluble and one showed no difference. The different trends observed among the peptides with lysine or aspartic acid heads are unlikely to solely result from the head group charge. If the charge was a critical independent parameter, both sets should exhibit similar trends as the tail length is varied, with one subset consistently showing higher solubility. Instead, different solubilities were only consistently observed when the tail length was 6 residues long.

Surprisingly, the hydrophobicity, volume, or flexibility of the tail did not affect peptide solubility. Although the tested peptides had three-amino acid tails, the results in **Figure 6.2** suggest that this short length is not masking any effects. Longer peptides are more flexible due to the greater number of potential conformations, and are more hydrophobic. It is thus possible that the effect of these parameters on GPCR solubility will not become apparent until the peptides have reached a critical length. However, both Ac-A<sub>3</sub>K-NH<sub>2</sub> and Ac-A<sub>6</sub>K-NH<sub>2</sub> have comparable solubilities, suggesting this is not the case. All of the tested peptides are 2-3nm in length, which is similar to traditional detergents. It is not practical to greatly increase the length of the peptides, as this would likely result in a larger “belt” around the solubilized GPCR, which would inhibit protein-protein contacts in a crystal lattice. An increase in peptide length would also decrease their solubility, making



them more difficult to handle. Thus, the parameters of hydrophobicity, volume, and flexibility are not important for GPCR solubility at the length scales that are biologically or experimentally relevant.

The peptide batch or tendency to spontaneously form nanostructures may affect the peptides' ability to solubilize GPCRs. In Chapter 5, solubilities as high as 93% were observed, while solubilities in this Chapter were, on average, lower for each tested GPCR (50-85%; the data is not shown). Different batches of peptides were used for the two studies. This indicates that the manufacturing process, minute impurities, or small variations in the solubilization of the lyophilized peptides may affect their ability to perform as surfactants. Also, several studies have reported that the peptides form dynamic structures, switching from micelles to rods or other structures [6-10]. It is probable that GPCRs are more soluble in some structures than in others. It was not possible within the scope of the present study to simultaneously evaluate the peptide microstructure and GPCR solubility.

As the data in Chapter 5 suggested, the peptide length and head group charge do affect GPCR solubility. However, **Figures 6.1** and **6.2** demonstrate that the effect of these variables is not significant. The ratio of soluble GPCRs in Ac-A<sub>6</sub>K-NH<sub>2</sub> and Ac-A<sub>3</sub>K-NH<sub>2</sub> was between ~1.2-1.3, and these ratios were the largest observed among all of the peptides. This indicates that all of the peptides function comparably as detergents. In contrast, Brij-35 and Brij-58 were 4-5 times better than the next best detergents. This stark difference suggests that the researcher may decide to use the peptide that is easiest to produce, or that is most compatible with the desired experiments.

## 6.5 References

1. Huang F, Nau WM. A conformational flexibility scale for amino acids in peptides. *Angew Chem Int Ed* **42**, 2269-2272 (2003).
2. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132 (1982).
3. The Amino Acid Repository: [http://www.imb-jena.de/IMAGE\\_AA.html](http://www.imb-jena.de/IMAGE_AA.html)
4. Prive CG. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
5. Anatrace Online Technical Documents:  
<http://www.affymetrix.com/browse/brand/anatrace/anatrace-overview.jsp?category=35858&categoryIdClicked=35858&rootCategoryId=35677&navMode=35858&parent=35858&aid=anatraceNav>
6. Vauthey S, Santoso S, Gong H, Watson N, Zhang S. Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles *Proc Natl Acad Sci USA* **99**, 5355-5360 (2002).
7. Santoso S, Hwang W, Hartman H, Zhang S. Self-assembly of surfactant-like peptides with variable glycine tails to form nanotubes and nanovesicles. *Nano Letters* **2**, 687-691 (2002).
8. von Maltzahn G, Vauthey S, Santoso S, Zhang S. Positively charged surfactant-like peptides self-assemble into nanostructures. *Langmuir* **19**, 4332-4337 (2003).
9. Nagai A, Nagai Y, Qu H, Zhang S. Self-assembling behaviors of lipid-like peptides A<sub>6</sub>D and A<sub>6</sub>K. *J Nanoscience & Nanotechnology* **7**, 2246-2252 (2007).
10. Yang S, Zhang S. Self-assembling behavior of designer lipid-like peptides. *Supramolecular Chemistry* **18**, 389-396 (2006).

# CHAPTER 7

## ***IN VITRO* MUTATIONAL ANALYSIS OF THE OLFACTORY LIGAND BINDING POCKET**

### **7.1 Introduction**

In order to understand olfaction at the molecular level, as well as rationally design artificial noses, the odorant binding pocket in olfactory-related receptors must be precisely determined. Interactions between specific residues, or even atoms, on receptors and their ligands uniquely determine each receptor's function. For example, the rat I7 olfactory receptor preferentially binds octanal. A valine to isoleucine switch in the mouse I7 ortholog results in preferential binding to heptanal [1]. The human receptor hOR912-13 is inactive even when a nonsense mutation resulting in premature truncation is corrected, while the mouse ortholog strongly binds ketones [2]. Computational analyses indicate that replacing a glycine with a serine in the predicted hOR912-13 binding site should allow ligands to form a critical hydrogen bond, and thereby restore the receptor's function [3]. Similarly, 17 out of 18 residues interacting with the ligand eticlopride are identical in the dopamine D2 and D3 receptors. Modeling studies suggest that subtle differences in the

backbone conformation or residue packing of these receptors can alter their ligand selectivity, even when their binding pockets are nearly identical [4].

Although the detailed structure of the binding pocket in olfactory-related receptors has not been precisely elucidated, its general location is known. Hydrophathy plots, sequence alignments, and residue helix-forming tendencies can be used to predict the transmembrane regions. Crystal structures of 6 GPCRs and mutational studies confirm that the ligand-binding domain is in the pocket of the barrel formed by the transmembrane helices [4-12]. Moreover, alignments of olfactory receptor protein sequences show that specific regions in the transmembrane segments are highly variable, while other sections of the receptors are highly conserved. Several groups of researchers have hypothesized that these variable regions include the binding pocket residues, and it is this variability that allows olfactory receptors to detect an extraordinary range of odors [13-16]. Additional studies indicate that the second extracellular loop may also be involved in ligand binding [14, 17].

There have been extensive modeling and simulation analyses of the binding of olfactory receptors with their odorants [18-25], but only 3 experimental mutagenesis assays [10, 11, 26]. These mutagenesis experiments were performed in cells with calcium influx assays. No biochemical, biophysical or structural studies have been performed on the actual receptors due to the difficulty of producing and purifying them in sufficient quantities. The methods described in Chapters 2 and 4 have overcome this crucial limitation. The ability to work in a cell-free system allows many mutants to be quickly and simultaneously screened. Promising mutants can then be selected, and their activity can be evaluated in HEK293 cells. It is now possible to experimentally verify or discount previous simulations, as well

as explore critical binding sites in multiple ORs. Indeed, it should now be possible to determine whether certain residues or alignment positions are conserved in the majority of OR binding sites.

## **7.2 Methods**

### **7.2.1 Rationale for Receptor Choice**

The receptor mOR103-15 was chosen for mutational experiments because it was a receptor with one of the highest cell-free expression yields, and because several groups of researchers have performed simulations to determine key residues involved in binding. However, no experimental data has yet been generated. Since no OR crystal structure exists, the actual binding site residues remain unknown. The simulations thus give a starting point, and the results described here should also indicate how effective various models are at predicting critical residues in OR binding sites. This also forms a basis that can be expanded to other ORs.

### **7.2.2 Rationale for Residue Choice**

Two approaches were used to predict residues in the ligand-binding pocket (**Figure 7.1**). 1) Residues were chosen if they were implicated by multiple computational analyses specific for mOR103-15. 2) A sequence alignment of mOR103-15 with the ORs analyzed in Chapter 4, the ORs published by [10, 15, 18-24], and the GPCRs with known structures was performed. Residue positions that mutational or computational studies suggested may form the pocket were examined. Residues were chosen for mutational analysis if their position was involved in binding in multiple studies, or if a position reported in a single

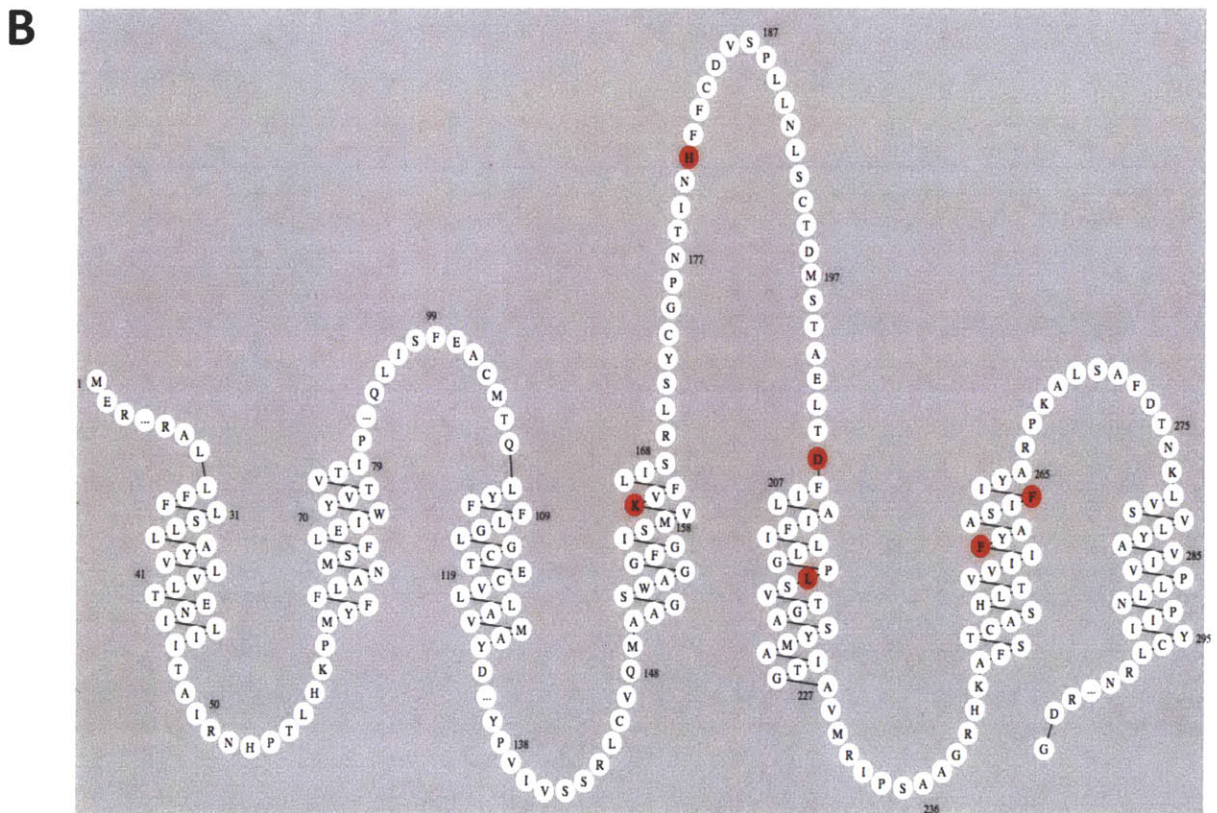
study belonged to a residue with potentially critical or highly conserved properties. To test whether the chosen residues were involved in ligand binding, they were mutated to residues with opposite physical or chemical properties.

Several independent simulations indicate that K164 is a critical residue involved in hydrogen bonding and ligand recognition, and that D204 is a residue that may participate in this interaction [18, 21, 19]. Because any ligand binding interactions would be due to their charged, hydrophilic nature, these residues were mutated to the hydrophobic residue leucine. Some reports indicate that the ligand binding pocket is lined with hydrophobic residues like F262 whose volume limits the size of the entering odorants [25]. To preserve the relatively hydrophobic nature of this residue while changing its shape and flexibility, this residue was also changed to leucine. The residues L216 and F256 have been implicated in binding in other ORs. The residue F256 is relatively conserved throughout the OR family, while L216 is in a hypervariable region. Mutational analysis has shown that that both residues are important for ligand binding in mOR174-9 [10]. Hydrophobic binding pocket residues probably interact with odorants through hydrophobic or van der Waals interactions. Because the latter in particular is sensitive to distance, the residue L216 was changed to the smaller residue valine. Because a phenylalanine-to-leucine substitution severely diminished binding activity in mOR174-9, this strategy was also employed with F256: this residue was mutated to leucine.

It has been proposed that metal binding is important for odorant receptor activation. The conserved histidine H181 has been proposed as a metal-binding site [17]. This histidine is located in a conserved motif common to the OR family, suggesting that this sequence may impart a specific olfactory function. When expressed as an independent

peptide, this 5-residue motif can bind copper and zinc ions, and this binding induces an  $\alpha$ -helical conformation [17]. Metal ions are present in the buffers used to synthesize and purify the olfactory receptors. To test whether this residue might be involved in ligand recognition, and to abolish potential metal-coordinating abilities, it was mutated to a leucine.

**A** MERRNHTGRV SEFVLLGFPA PAPLRALLEF LSELLAYVLVL TENILIITAI RNHPTLHKPM  
YFFLANMSFL EIWYVTVTIP KMLAGFIGSE ENHGQLISFE ACMTQLYFFL GLGCTECVLL  
AVMAYDRYVA ICHPLHYPVI VSSRLCVQMA AGSWAGGFGI SMVKVFLISR LSYCGPNTIN  
HFFCDVSPLL NLSCTDMSTA ELTDFILAIF ILLGPLSVTG ASYMAITGAV MRIPSAAGRH  
KAFSTCASHL TVVIIFYAAS IFIYARPKAL SAFDTNKLVS VLYAVIVPLL NPIIYCLRNQ  
EVKKALRRTL HLAQQQDANT KKSSRDGGSS GTETSQVAPA (340 amino acids)



**Figure 7.1. Sequence and 2D-Topology of mOR103-15.** A) The sequence of the native mouse olfactory receptor mOR103-15. The proposed mutation sites are highlighted in red, the hypothesized metal-binding region is highlighted in blue, the 7 predicted transmembrane domains are underlined, and the rho1D4 epitope used for purification and GSSG linker are highlighted in green. B) The two-dimensional topology of mOR103-15 showing the predicted transmembrane regions and ligand-binding sites (red). The snake plot was generated using [33].

### **7.2.3 Gene Design, and Receptor Expression and Analysis**

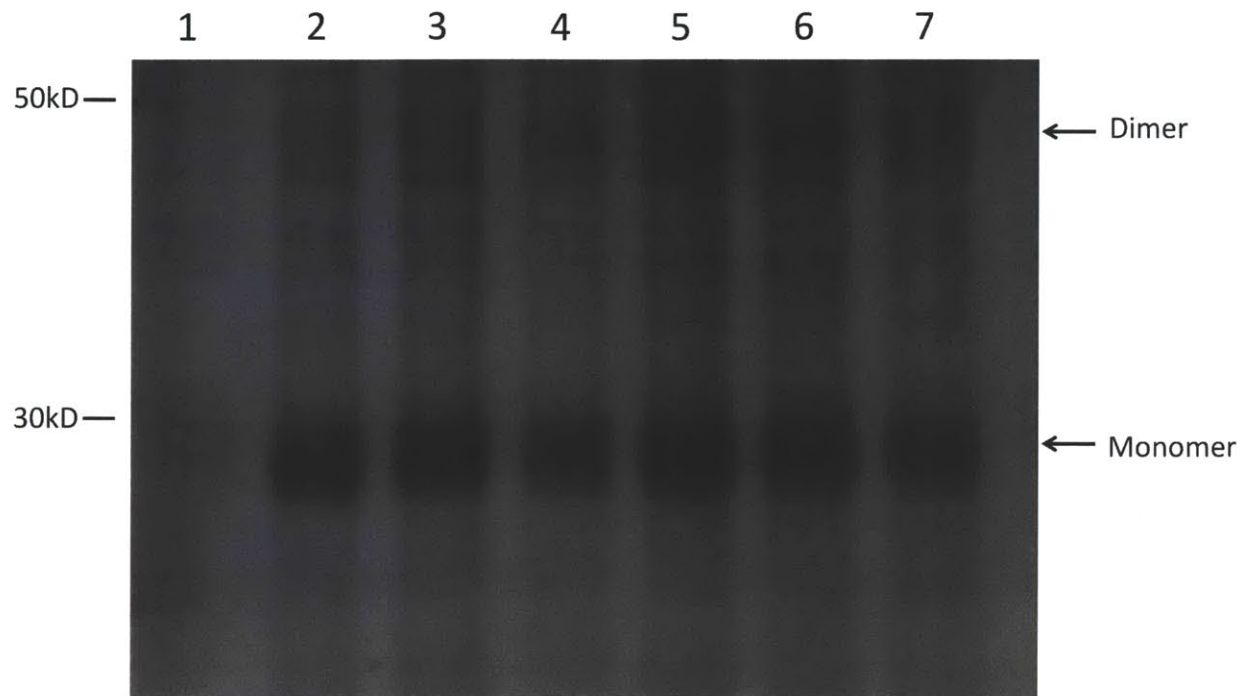
Six single-amino acid variants of mOR103-15 were designed with the mutations described above. The genes were synthesized by GeneArt, and subcloned into the pIVex2.3 vector using the NcoI and XhoI restriction sites. The protocols for expressing and purifying receptors are described in Chapter 4, as well as the methodologies used to assay their secondary structure and ligand-binding ability.

## **7.3 Results**

### **7.3.1 Protein Expression and Purification**

The six mOR103-15 mutants and native protein were expressed in cell-free reactions, and purified using immunoaffinity chromatography. The purified samples were analyzed on a western blot and silver stain (**Figure 7.2**). Both gels showed bands that ran at the expected monomer and dimer sizes. The western blot confirmed the identity of the proteins, while the silver stain showed that they could be purified to the same level. Because all of the mutants ran at the same size as the native protein, and they all showed the same tendency to dimerize, their global structure and backbone conformation were probably not compromised as a result of the mutations. Instead, any conformational changes that may affect receptor function are probably more localized.



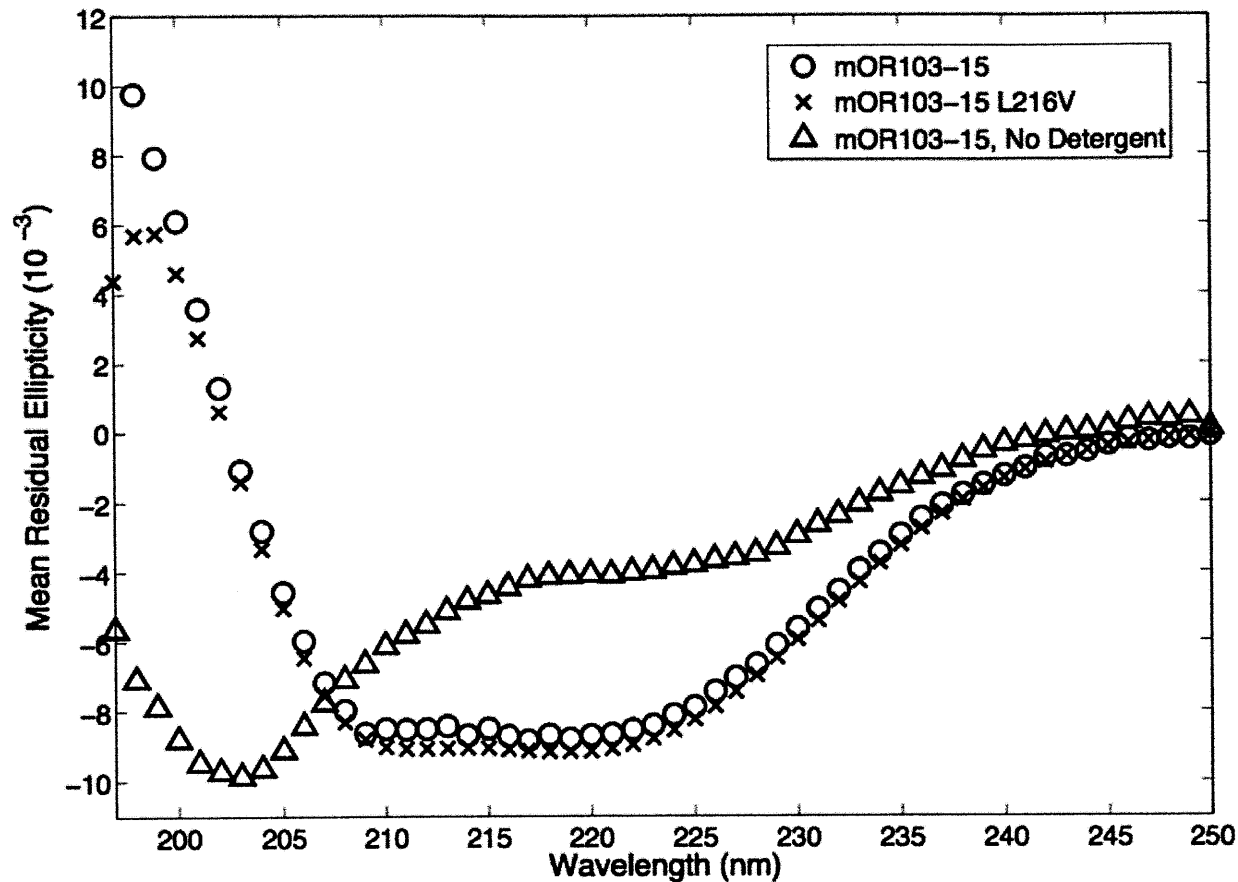


**Figure 7.2: Silver Stain of the Native and Mutant mOR103-15 Receptors.** Lane 1: mOR103-15, Lane 2: mOR103-15 K164L, Lane 3: mOR103-15 H181L, Lane 4: mOR103-15 D204L, Lane 5: mOR103-15 L216V, Lane 6: mOR103-15 F256L, Lane 7: mOR103-15 F262L. All of the receptors are the same size and show the same tendency to dimerize. They can also be purified to the same level of purity.

### 7.3.2 Secondary Structure Analysis

Circular dichroism was used to analyze the secondary structure of the purified mutants and native protein. All of the samples had characteristic  $\alpha$ -helical spectra, indicating that they were properly folded (**Figure 7.3**). Moreover, the spectra of the native and mutant proteins nearly overlapped, indicating that they were folded in the same manner. This data and the spectrum of mOR103-15 expressed without detergent further suggest that the mutations do not result in misfolded receptors. Lack of receptor function could be due to either improper protein folding, or a physical or chemical change in key binding-pocket residues. The mutant CD spectra and gel images suggest that the receptors

are folded into 7 transmembrane  $\alpha$ -helices, and that any changes observed in ligand binding result from local disturbances introduced through the mutations.



**Figure 7.3: Circular Dichroism Spectra of Native and Mutant mOR103-15.** The native receptor and the L216V mutant have overlapping spectra that are characteristic of an  $\alpha$ -helical secondary structure, while the native receptor made without detergent has a spectrum reminiscent of a random coil. This suggests that the native and mutant receptors made with detergent are properly folded, and that the mutation does not affect protein folding. The spectrum of mOR103-15 L216V is characteristic for all of the mutants.

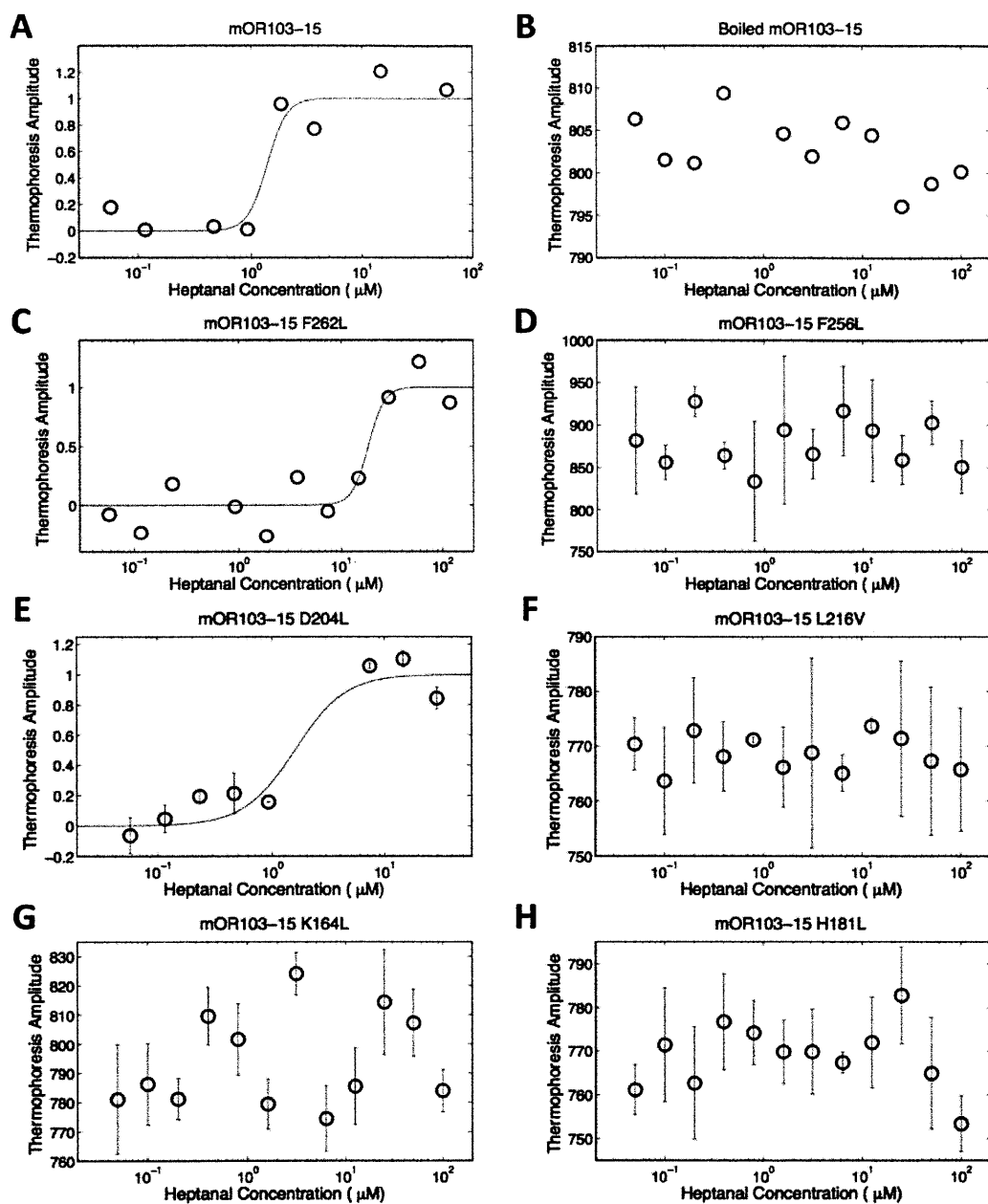
### 7.3.3 Ligand Binding Analysis

The same receptor samples that were analyzed using circular dichroism were also analyzed using microscale thermophoresis. As a control, a boiled sample of mOR103-15 was also analyzed.

The native mOR103-15 and boiled control exhibited the same binding behavior observed in Chapter 4. The native receptor bound its ligand heptanal with an  $EC_{50}$  of  $1.4\mu\text{M}$ , while the denatured sample did not exhibit any binding (**Figure 7.4**). The similar binding behavior in two different mOR103-15 samples reinforces the conclusion that the measurements are detecting binding between the receptor and its ligand.

Two of the mutants were able to bind heptanal, while no binding was seen in others (**Figure 7.4, Table 7.1**). The mutant D204L bound heptanal with an  $EC_{50}$  similar to that of the native protein, indicating that this residue is not critical for ligand binding. The mutant F262L was also able to bind heptanal, albeit with an affinity that was an order of magnitude lower. This significant difference suggests that this residue is involved in binding the ligand, though perhaps indirectly. All ligand-binding was abolished in the mutants K164L, F256L, H181L, and L216V, suggesting that these residues directly interact with the ligand.

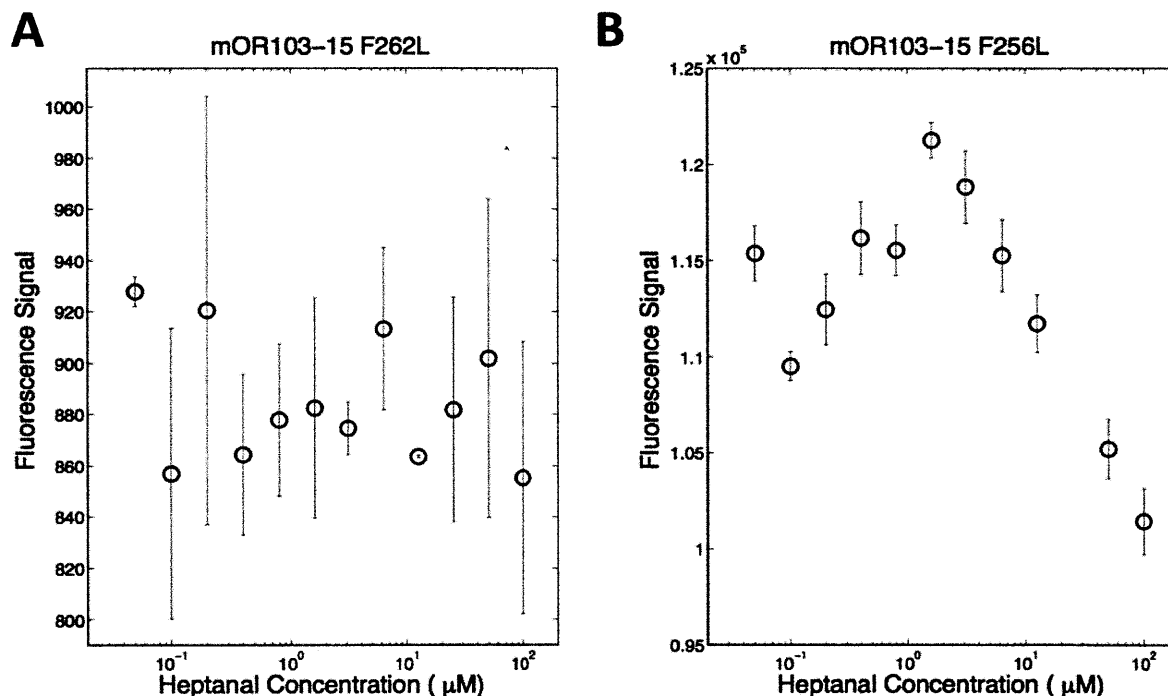
Total fluorescent measurements in each capillary further suggested how the ligand may be interacting with the receptor in the binding pocket. For most of the samples, the fluorescence in each capillary remained constant as the ligand concentration was increased. However, for the F256L variant, the fluorescence decreased linearly as the heptanal concentration increased (**Figure 7.5**). Because all measured fluorescent signals came from the protein tryptophans and are thus proportional to the receptor concentration, these results demonstrate that the ligand is quenching fluorescence by interacting with a tryptophan.



**Figure 7.4: Ligand-Binding Measurements of the Native mOR103-15 Receptor and the Mutants.** A) Native mOR103-15. Only one measurement was made to verify that this sample had the same properties as the sample of mOR103-15 that was tested in Chapter 4. B) Heat-denatured mOR103-15. C) mOR103-15 F262L. Only one measurement was obtained. D) mOR103-15 F256L. E) mOR103-15 D204L. F) mOR103-15 L216V. G) mOR103-15 K164L. H) mOR103-15 H181L. Most of the mutated receptors did not demonstrate binding to heptanal, indicating that the substituted residues interact with the ligand. Two mutants did show binding: D204L, F262L. Because the observed half-maximal concentration was an order of magnitude higher in the F262L mutant compared to the native protein, the residue F262 is also likely to be involved in receptor-ligand interactions.

**Table 7.1: Measured EC<sub>50</sub> Values for mOR103-15 Mutants**

Mutation in mOR103-15	EC <sub>50</sub> (μM)
Native Protein	1.4 ± 0.7
Denatured Native Protein	No binding
K164L	No binding
H181L	No binding
D204L	1.6 ± 0.7
L216V	No binding
F256L	No binding
F262L	18.5 ± 3.8



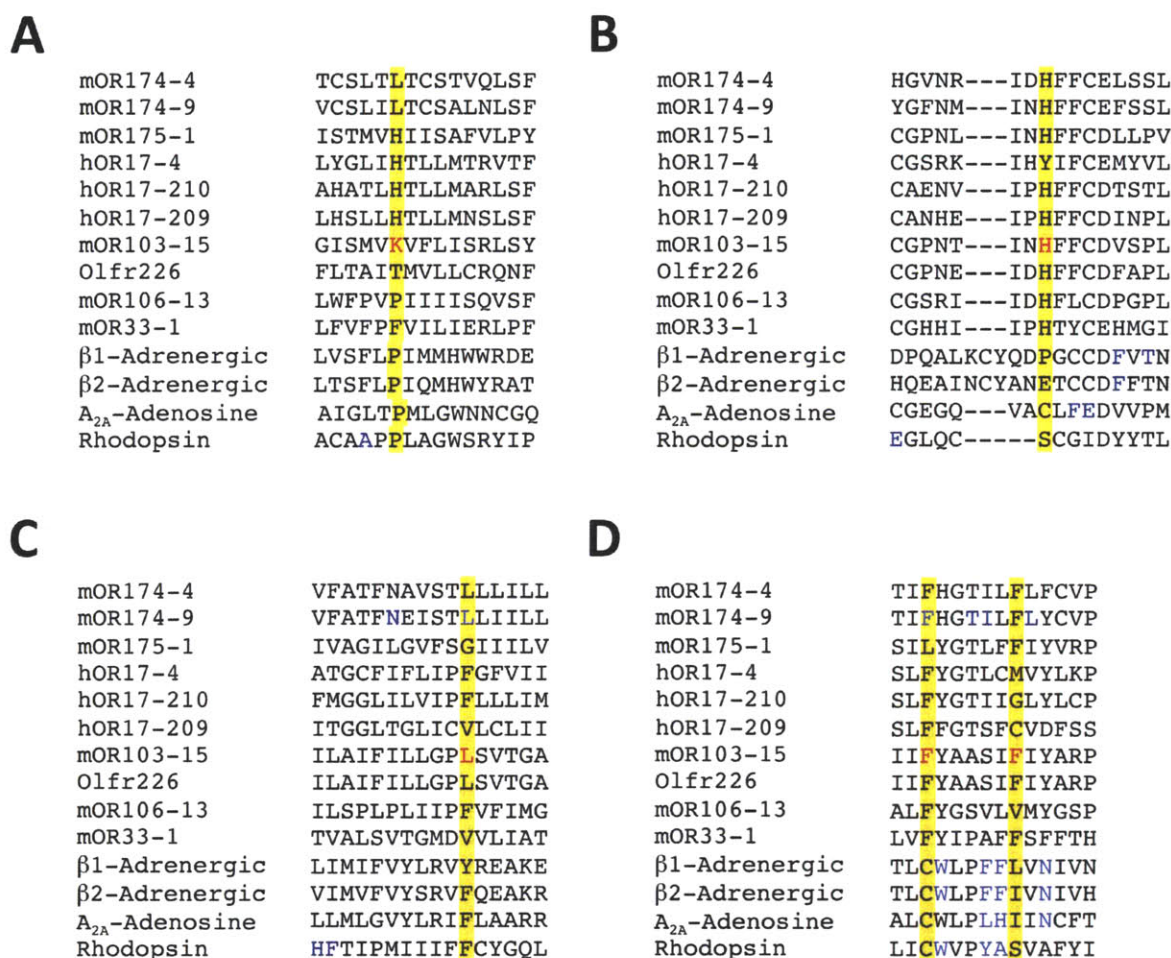
**Figure 7.5: Fluorescence Signal as a Function of Heptanal Concentration.** A) The signal for the F262L Mutant. B) The signal for the F256L Mutant. In the F256L mutant samples, the fluorescence signal is quenched as the heptanal concentration increases. This indicates that the ligand is directly interacting with a tryptophan in the receptor. The fluorescence signal in the mutant F262L does not vary with heptanal concentration, and this behavior is representative of the remaining mutants as well as the native protein. This further demonstrates that the quenching observed with the F256L mutant results from a heptanal-tryptophan interaction that depends on the residue position 256, and not the amino acid.

## 7.4 Discussion

### 7.4.1 Molecular Interactions Involved in Olfactory Receptor Ligand Binding

The results presented in this study suggest which residues may be involved in ligand binding in mOR103-15, as well as other olfactory receptors.

Based on computational analyses, several groups of researchers have reported that residue K164 recognizes odorants through hydrogen bonds, or potentially by forming a Schiff base [18, 19, 21]. One study also suggested that the residue D204 participates in this interaction, either through repulsive forces that prevent a strong interaction between the ligand and K164, or by stabilizing the putative Schiff base [21]. Modeling in this latter study further predicted that ligand affinity would decrease if K164 were mutated to an alanine, while the affinity would increase if D204 were mutated to an alanine. The data in **Figure 7.4** suggests that K164 is indeed involved in ligand recognition, as mutation to a leucine abolished a response to heptanal. However, mutation of D204 to a leucine did not affect the ligand affinity (**Table 7.1**), indicating that this residue is not involved in binding heptanal. Interestingly, the alignment position of K164 in other olfactory receptors often has either a histidine or glutamine, or a straight-chain hydrophobic residue. It is likely that this represents a key position in many olfactory receptors that is involved in ligand recognition. Indeed, several computational studies predict that the histidine at this alignment position is involved in ligand binding [16, 21]. Lysine, histidine, and glutamine can form hydrogen bonds, and are probably involved in recognizing specific functional groups on odorants. In contrast, receptors with a hydrophobic residue at this alignment position have probably evolved to recognize different odorant functional groups.



**Figure 7.6: Alignment of mOR103-15 with 9 Olfactory Receptors and 4 GPCRs with Known Structures.** A) Region Surrounding K164, B) Region Surrounding H181, C) Region Surrounding L216, and D) Region Surrounding F256 and F262. The position to the right of F256 contains a tryptophan that is conserved in most GPCRs, and a tyrosine that is conserved in ORs. The residues mutated in this study are in red, the alignment position is highlighted in yellow and residues experimentally shown to bind ligands in other receptors are in blue. Alignments were performed by ClustalW [34].

The results in **Figures 7.4** and **7.5** indicates that the residue F256 interacts with the ligand heptanal, and may even be a conserved ligand-recognition site in the OR family. Mutation of this phenylalanine to a leucine abolished a binding response that could be detected by thermophoresis. Moreover, heptanal was able to interact with a tryptophan in the mutated protein and quench its fluorescence. The receptor mOR103-15 has two

tryptophans. One is located in the second transmembrane helix (W73), and the other is in the fourth transmembrane helix (W154). Several groups predict that W73 faces the cell membrane, while W154 faces the binding pocket [14, 21]. Also, W154 is conserved not only in olfactory receptors, but in the GPCR family as well, suggesting that it may play a critical structural or functional role. Taken together, this suggests that heptanal is interacting with W154, and that F256 and W154 are vertically aligned or in close proximity. Perhaps F256 acts as a gate; upon ligand binding, it is somehow able to change conformation and interact with W154, enabling cell-signaling. When this molecular “gate” is removed by substituting in a leucine, the ligand can freely and directly access underlying tryptophan. Interestingly, a tryptophan that is generally conserved among GPCRs in the sixth transmembrane region is believed to form a “toggle” switch necessary for receptor activation [8, 27-29]. In olfactory receptors, a tyrosine typically occupies this conserved position. It is possible that this tyrosine functions as a toggle switch in olfactory receptors, but it is also possible that a different molecular mechanism exists. It should be noted that the proposed mechanism does imply ligand-receptor binding, and that none was detected. However, it’s possible that the fluorescence quenching is masking any thermophoresis signal. Also, because no olfactory receptor structure exists, the possibility of an interaction with W73 cannot be discounted. The transmembrane regions are based on modeling predictions, and a difference of even a few residues in the transmembrane helix boundaries can change the relative orientations of these residues. It is thus possible that W73 does face the binding pocket. Interestingly, the residue F256 in mOR103-15 corresponds to the residue F252 in mOR174-9, and is conserved throughout the OR family (**Figure 7.6**). A mutational study demonstrated that substituting F252 in mOR174-9 with a leucine



decreased its responsiveness to its ligand two-fold [10]. Because this residue is involved in ligand recognition in at least one olfactory receptor, and because the residue is conserved in the OR family, it is likely to be involved in ligand-binding in other olfactory receptors.

The residue F262 is also conserved in ORs, and is aligned near two phenylalanines in the  $\beta$ 1- and  $\beta$ 2-adrenergic receptors that are involved in ligand-binding (**Figure 7.6**). Although the adenosine  $A_{2A}$  receptor and rhodopsin do not have phenylalanines at this location, the residues that do align with adrenergic receptor phenylalanines also interact with the ligands of these receptors [5, 7, 8, 12]. Because ligand-binding residues are perfectly aligned in transmembrane 6 in GPCRs whose crystal structures are known, it seems probable that other GPCRs may have ligand-binding residues in the same region. Thermophoresis data suggests that F262 is one such residue: the measured dissociation constant increased by an order of magnitude when this residue was substituted with a leucine (**Table 7.1**), suggesting weaker binding. Because binding was not completely abolished, it is likely that this residue interacts with heptanal through van der Waals interactions, or by physical interference. It is also possible that an F262L mutation alters the local conformation of the receptor backbone, or the orientation of neighboring residues, such that binding of heptanal is more constricted or less favorable. Indeed, several reports predict that this residue faces the lipid bilayer [14, 21]. However, since no OR structure exists, it is possible that F262 does actually face the binding pocket. Its close alignment with ligand-binding residues in other GPCRs suggests that this may be the case.

Although it is located in the second extracellular loop (ECL2) instead of the ligand-binding pocket (**Figure 7.1**), the residue H181 seems to be involved in ligand binding. Indeed, ECL2 residues in other GPCRs are involved in ligand-receptor interactions, and

H181 is closely aligned with these residues. It is well known that strong odors are good ligands for metal ions [17, 30], and that histidines can coordinate metals and typically comprise metal-binding domains in proteins [31]. These observations suggest that olfactory receptors are metalloproteins, and that a conserved histidine like H181 may coordinate metal ion interactions that in turn affect ligand binding or sensitivity. Indeed, one study has shown that the conserved motif in which H181 is located (HFFCD/E) is capable of binding metal ions [17]. Taken together, this suggests that H181 may affect OR-odorant sensitivity by coordinating with metal ions. Replacing H181 with a leucine prevents this interaction. It is also possible that H181 directly interacts with the ligand, and future studies must be performed to elucidate its role. Since the ECL2 ligand-binding residues in other GPCRs are primarily phenylalanines or glutamic acid, and these residues are also located near H181, it is possible that H181 only neighbors the actual ligand-binding residue and that an H181L substitution alters the loop conformation in a way that inhibits binding.

The residue L216 also appears to be involved in ligand recognition, as mutation to valine prevents binding to heptanal (**Figure 7.4**). This residue is in a hypervariable region [14], and is aligned to a leucine in mOR174-9 that has been shown to be involved in ligand binding [10]. This suggests that this particular residue position may be critical for ligand recognition in other ORs as well, and that it likely either interacts through hydrophobic or van der Waals interactions instead of recognizing functional groups on odorants.

Taken together, the binding results presented here, as well as sequence alignments and published mutational studies, suggest that some residues or key alignment positions may be involved in ligand binding in all or most ORs. Most of the residues implicated in

binding in this study are either conserved throughout the entire OR family, or are aligned with residues with similar properties. The specific physical or chemical properties of these amino acids may thus determine which functional groups are recognized on odorants, and how the odorants are situated within the binding pocket. This, in turn, would determine the extent of odorant-activated signal transduction.

#### **7.4.2 An Efficient Method for Parallel Screening of Multiple GPCRs**

The results reported in this chapter demonstrate that the methods developed in Chapter 4 constitute a relatively high-throughput method for producing ORs and studying their function. The large number of receptors that could be produced further suggests that the technique can be extended to other GPCRs, and potentially other membrane proteins. However, in order to verify binding measurements made with purified receptors, and to completely understand their biological function, calcium influx assays must also be conducted. It is possible for a ligand to bind to a receptor, yet not initiate a signal. This can happen if the ligand is an antagonist [32], or if a minimal energy barrier has not been reached [22]. Conversely, it may be possible for a ligand to initiate a signal, yet not bind in a detectable manner due to a low affinity, or high noise-to-signal ratio. Also, because the lipid bilayer exerts mechanical forces on proteins found within it, it is possible that the conformation of a receptor may change slightly yet significantly when it is transferred to a detergent environment, thus changing its binding affinity. Nevertheless, it is now possible to simultaneously express and purify multiple mutants using cell-free technology, and screen these mutants for potential ligand binding activity. Promising mutants can then be selected, and further tested in cells.

## 7.5 References

1. Krautwurst D, Yau K-W, Reed RR. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* **95**, 917-926 (1998).
2. Gaillard I, *et al.* A single olfactory receptor specifically binds a set of odorant molecules. *Eur J Neurosci* **15**(3), 409-418 (2002).
3. Hummel P, Vaidehi N, Floriano WB, Hall SE, Goddard WA III. Test of the binding threshold hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93. *Prot Sci* **14**, 703-710 (2005).
4. Chien *et al.* Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091-1095 (2010).
5. Cherezov V, *et al.* High resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science* **318**(5854), 1258-1265 (2007).
6. Rasmussen SGF, *et al.* Crystal structure of the human  $\beta_2$  adrenergic G-protein-coupled receptor. *Nature* **450**, 383-388 (2007).
7. Warne T, *et al.* Structure of a  $\beta_1$ -adrenergic G protein-coupled receptor. *Nature* **454**(7203), 486-491 (2008).
8. Jaakola V-P, *et al.* The 2.6 Å crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**(5905), 1211-1217 (2008).
9. Wu B, *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071 (2010).
10. Katada S, Hirokawa T, Oka Y, Suwa M, Touhara K. Structural basis for broad but selective ligand spectrum of a mouse olfactory receptor: Mapping the odorant binding site. *J Neurosci* **25**(7), 1806-1815 (2005).
11. Baud O, *et al.* The mouse eugenol odorant receptor: structural and functional plasticity of a broadly tuned odorant binding pocket. *Biochemistry* **50**(5), 843-853 (2011).
12. Palczewski K, *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745 (2000).
13. Buck L, and Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175-187 (1991).

14. Liu AH, Zhang X, Stolovitzky GA, Califano A, Firestein SJ. Motif-based construction of a functional map for mammalian olfactory receptors. *Genomics* **81**, 443-456 (2003).
15. Zozulya S, Echeverri F, Nguyen T. The human olfactory receptor repertoire. *Genome Biol* **2**(6), research0018.1-0018.12 (2001).
16. Pilpel Y, Lancet D. The variable and conserved interfaces of modeled olfactory receptor proteins. *Protein Sci* **8**, 969-977 (1999).
17. Wang J, Luthey-Schulten ZA, Suslick KS. Is the olfactory receptor a metalloprotein? *Proc Natl Acad Sci* **100**(6), 3035-3039 (2003).
18. Kurland MD, et al. Discrimination of saturated aldehydes by the rat I7 olfactory receptor. *Biochemistry* **49**(30), 6302-6304 (2010).
19. Hall SE, Floriano WB, Vaidehi N, Goddard WA III. Predicted 3-D structures for mouse I7 and rat I7 olfactory receptors and comparison of predicted odor recognition profiles with experiment. *Chem Sens* **29**(7), 595-616 (2004).
20. Floriano WB, Vaidehi N, Goddard WA III. Making sense of olfaction through predictions of the 3-D structure and function of olfactory receptors. *Chem Sens* **29**(4), 269-290 (2004).
21. Singer MS. Analysis of the molecular basis for octanal interactions in the expressed rat I7 olfactory receptor. *Chem. Sens* **25**, 155-165 (2000).
22. Hummel P, Vaidehi N, Floriano WB, Hall SE, Goddard WA III. Test of the binding threshold hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor 912-93. *Prot Sci* **14**, 703-710 (2005).
23. Floriano WB, Vaidehi N, Goddard WA III, Singer MS, Shepherd GM. Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. *Proc Natl Acad Sci USA* **97**(20), 10712-10716 (2000).
24. Man O, Gilad Y, Lancet D. Prediction of the odorant binding site of olfactory receptor proteins by human-mouse comparisons. *Prot Sci* **13**, 240-254 (2004).
25. Lai PC, Singer MS, Crasto CJ. Structural activation pathways from dynamic olfactory receptor-odorant interactions. *Chem Sens* **30**, 781-792 (2005).
26. Kuang D, et al. Molecular similarities in the ligand binding pockets of an odorant receptor and the metabotropic glutamate receptors. *J Biol Chem* **278**(43), 42551-42559 (2003).
27. Park JH, Scheerer P, Hofmann KP, Choe H, Ernst OP. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183-188.

28. Shi L, *et al.* B2 adrenergic receptor activation: modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J Biol Chem* **277**(43), 40989-40996 (2002).
29. Crocker E, *et al.* Location of Trp265 in metarhodopsin II: Implications for the activation mechanism of the visual receptor rhodopsin. *J Mol Biol* **357**, 163-172 (2006).
30. Rakow NA, Suslick KS. A colorimetric sensor array for odour visualization. *Nature* **406**, 710-713 (2000).
31. Branden C, Tooze J. Introduction to Protein Structure, Second Edition. Garland Publishing, Inc, NY, USA (1999).
32. Oka Y, Omura M, Kataoka H, Touhara K. Olfactory receptor antagonism between odorants. *EMBO J* **23**(1), 120-126 (2004).
33. Konvicka K, Campagne F, Weinstein H. Iterative construction of residue-based diagrams of proteins: the RbDe web service. *Prot Eng* **13**(6), 395-396 (2000).
34. Chenna R, *et al.* Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**(13), 3497-3500 (2003).