

## Insertion of T4-lysozyme (T4L) can be a useful tool for studying olfactory-related GPCRs†

Karolina Corin,<sup>a</sup> Horst Pick,<sup>b</sup> Philipp Baaske,<sup>c</sup> Brian L. Cook,<sup>a</sup> Stefan Duhr,<sup>c</sup> Christoph J. Wienken,<sup>d</sup> Dieter Braun,<sup>d</sup> Horst Vogel<sup>b</sup> and Shuguang Zhang<sup>\*a</sup>

Received 12th December 2011, Accepted 15th March 2012

DOI: 10.1039/c2mb05495g

The detergents used to solubilize GPCRs can make crystal growth the rate-limiting step in determining their structure. The Kobilka laboratory showed that insertion of T4-lysozyme (T4L) in the 3rd intracellular loop is a promising strategy towards increasing the solvent-exposed receptor area, and hence the number of possible lattice-forming contacts. The potential to use T4L with the olfactory-related receptors hOR17-4 and hVN1R1 was thus tested. The structure and function of native and T4L-variants were compared. Both receptors localized to the cell membrane, and could initiate ligand-activated signaling. Purified receptors not only had the predicted alpha-helical structures, but also bound their ligands canthoxal ( $M_w = 178.23$ ) and myrtenal ( $M_w = 150.22$ ). Interestingly, the T4L variants had higher percentages of soluble monomers compared to protein aggregates, effectively increasing the protein yield that could be used for structural and function studies. They also bound their ligands for longer times, suggesting higher receptor stability. Our results indicate that a T4L insertion may be a general method for obtaining GPCRs suitable for structural studies.

### Introduction

Although membrane proteins comprise 20–30% of cellular proteins and have significant biological importance, membrane protein research lags far behind that of soluble proteins.<sup>1,2</sup> Knowledge about GPCRs in particular is sparse. This is due primarily to the difficulty in determining their structure. As of October 2011, over 76 000 protein structures have been determined. Only 303 are unique membrane proteins. Of these, only 7 are GPCRs (<http://www.pdb.org/pdb/home/home.do>).

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 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. Each bottleneck must be sequentially overcome. However, no universal method exists: optimal protocols for expression, purification, solubilization, and crystal growth must be empirically determined for each protein of interest.<sup>2</sup>

The last bottleneck is usually the rate-limiting step, as the difficulty in predicting crystal growth conditions necessitates screening through thousands of possibilities. Several strategies have been developed to facilitate membrane protein crystal growth.<sup>3–10</sup> To increase the surface area available to form a crystal lattice, T4-lysozyme (T4L) fusions have been synthesized,<sup>3–8</sup> and antibody fragments against specific portions of the membrane protein have been used.<sup>9</sup> 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.<sup>3–5,7–10</sup> To improve protein stability, sequence mutations have been introduced.<sup>4,6,10</sup>

<sup>a</sup> *Laboratory of Molecular Design, Center for Bits and Atoms, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA. E-mail: Shuguang@MIT.edu; Fax: +1-617-258-5239; Tel: +1-617-258-7514*

<sup>b</sup> *Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland*

<sup>c</sup> *NanoTemper Technologies GmbH, Floessergasse 4, 81369 München, Germany*

<sup>d</sup> *Systems Biophysics, Functional Nanosystems, Department of Physics, Ludwig-Maximilians University München, Amalienstrasse 54, 80799 München, Germany*

† This work is supported in part from **Defense Advanced Research Program Agency-HR0011-09-C-0012**. KC is a Yang Trust Fund Fellow.

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 the 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 five of the seven crystallized GPCRs used this approach.<sup>3-9</sup> Unlike the development of antibody fragments, generation of DNA templates is fast using current cloning techniques, and many 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 structure and function of various GPCRs.

This study examines the ability of T4L to be used as a general insert in olfactory-related GPCRs. Two GPCRs were chosen: hOR17-4 and hVN1R1. These receptors belong to two of the four families of GPCRs involved in olfaction. Structural knowledge of these different families can elucidate the different roles they play. Moreover, both receptors are expressed in other tissues, indicating non-olfactory functions as well.<sup>11,12</sup> Native and T4L-variants 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 imaging suggested that they could initiate second messenger signaling upon activation by their specific agonist. Circular dichroism analyses of purified receptors showed that the native and T4L-variants had alpha-helical structures. Microscale thermophoresis showed that the purified receptors bound their small molecule ligands canthoxal and myrtenal. Interestingly, the T4L-variants yielded higher percentages of soluble monomers compared to aggregates, indicating that the T4L insert stabilized the protein structure. These T4L variants also had lower ligand-binding affinities, but yielded more consistent binding results over time with less noise, thus suggesting longer receptor stability. These results suggest that a T4L insertion may be a general method when working with GPCRs and other 7-transmembrane proteins (7TM).

## Materials and methods

### 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, N183Q, N198Q, and N256Q mutations in hVN1R1 to facilitate crystallization. The genes were constructed by Geneart and ligated into the pcDNA4/To vector (Invitrogen, Carlsbad, CA). The plasmids were 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<sup>13</sup> and hVN1R1<sup>14</sup> were synthesized as previously described.

### Construction of stable inducible HEK293 cell lines

The generation of stable, inducible cell lines capable of expressing the desired proteins is described in ref. 11–13. Briefly, HEK293S *N*-acetylglucosaminyltransferase I-negative cells (HEK293G) containing the pcDNA6/Tr vector 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  $\mu\text{g ml}^{-1}$  of blasticidin and 50  $\mu\text{g ml}^{-1}$  zeocin were added. Cells were re-seeded at low density and grown until individual colonies formed. Colonies were picked and screened for inducible receptor expression. Cells were treated with plain media, media supplemented with 1  $\mu\text{g ml}^{-1}$  tetracycline, and media supplemented with 1  $\mu\text{g ml}^{-1}$  tetracycline and 2.5 mM 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 h at 4 °C. Cell lysates were centrifuged for 30 minutes at 10 000 rpm to remove insoluble debris. Dot blots and Western-blots 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), 15 mM HEPES (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 0.5 mM sodium pyruvate (Invitrogen), 100 units per ml penicillin and 100  $\mu\text{g ml}^{-1}$  streptomycin (Invitrogen). The expanded stable clones were maintained in media that also contained 5  $\mu\text{g ml}^{-1}$  blasticidin and 25  $\mu\text{g ml}^{-1}$  zeocin. All cells were grown at 37 °C, 5% CO<sub>2</sub>, and 95% relative humidity.

### Western blotting and silver staining

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 (Cell Essentials, 1.58 mg ml<sup>-1</sup> stock, 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).

### Immunocytochemistry

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  $1 \mu\text{g ml}^{-1}$  tetracycline and 1 mM 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,  $-20^\circ\text{C}$ ) and non-permeabilized cells were then blocked in PBST (PBS, 0.2% Tween-20, 0.3 M glycine, 4% serum) for 1 hour at room temperature, and incubated with the primary antibody solution ( $1.58 \text{ mg ml}^{-1}$  stock, 1 : 1000, PBS, 0.2% Tween-20, 4% serum) overnight at  $4^\circ\text{C}$ . 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. A Nikon Plan Apo 60x oil immersion lens was used.

### Ca<sup>2+</sup> 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  $\text{ml}^{-1}$ . OR expression was induced with  $1 \mu\text{g ml}^{-1}$  of tetracycline for 48 h. To visualize calcium signaling, cells were loaded with  $10 \mu\text{M}$  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. Ca<sup>2+</sup> signaling in response to  $50 \mu\text{M}$  of the odorant bourgeonal was visualized using confocal fluorescence microscopy (Zeiss LSM 510) with a water immersion objective (Zeiss Achroplan 63x NA 1.2). The cells were excited at 488 nm (Ar + 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  $37^\circ\text{C}$  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.<sup>16</sup>

### Cell extract preparation

Cells were grown on plates as previously described.<sup>13,14</sup> When the appropriate density was reached, cells were induced with tetracycline and sodium butyrate. After two days, cells were scrape harvested. They were pooled, and either used immediately, or snap-frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  until used for future experiments.

### 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  $4^\circ\text{C}$  to solubilize the protein, and were spun at 13000 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.<sup>17</sup>

### Receptor purification

Rho1D4 immunoaffinity purification has been previously described.<sup>13–15</sup> 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 12.5 ml/1 g cells. The protein was solubilized by rotating for 4 hours at  $4^\circ\text{C}$ . The non-solubilized fraction was pelleted by centrifuging for 30 minutes at  $30\,000g$  at  $4^\circ\text{C}$ . 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  $0.7 \text{ mg ml}^{-1}$ ); receptors were captured by rotating the mixture overnight at  $4^\circ\text{C}$ . 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  $4^\circ\text{C}$ , and re-pelleted. Washes were performed until the total protein concentration in the washes was less than  $0.01 \text{ mg ml}^{-1}$  (NanoDrop). One bead volume of elution buffer (PBS + 0.2% w/v FC-14 +  $800 \mu\text{M}$  Ac-TETSQVAPA-NH<sub>2</sub> peptide) was then added to the beads. Elutions were performed until the total protein concentration was less than  $0.01 \text{ mg ml}^{-1}$ .

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 a Äkta 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 1.5–3 ml using a 50000 MWCO filter column (Millipore), loaded on the column, and run with wash buffer at  $0.3 \text{ ml min}^{-1}$ . Fractions exiting the column were automatically collected; the protein content was monitored with UV absorbance at 215 nm, 254 nm, and 280 nm. Peak fractions were pooled, concentrated, and analyzed with Western blotting and silver staining (SilverXpress, Invitrogen).

### Secondary structure analysis using circular dichroism

CD spectra were measured over the wavelengths 200 nm–350 nm with a CD spectrometer (AVIV Biomedical Model 202). Measurements were made at  $15^\circ\text{C}$ , with a step size of 1 nm and an averaging time of 4 seconds. Measurements for each sample were made in triplicate and averaged. The protein spectra were blanked to the spectrum obtained for wash buffer. A QS quartz cuvette (Hellma) with a 1 mm path length was used to perform all experiments.



## Ligand binding measurements

Microscale thermophoresis was used to measure the binding interactions between purified receptors and their ligands using a setup similar to that previously described.<sup>18–20</sup> 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  $\mu\text{M}$ . Approximately 1.5  $\mu\text{l}$  of each sample was loaded in a fused silica capillary (Polymicro Technologies, Phoenix, USA) with an inner diameter of 300  $\mu\text{m}$ . An infrared laser diode was used to create a 0.12 K  $\mu\text{m}^{-1}$  temperature gradient inside the capillaries (Furukawa FOL1405-RTV-617-1480, wavelength  $\lambda = 1480$  nm, 320 mW maximum power, AMS Technologies AG, München Germany). Tryptophan fluorescence was excited with a UV-LED (285 nm), and was measured with a 40 $\times$  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).

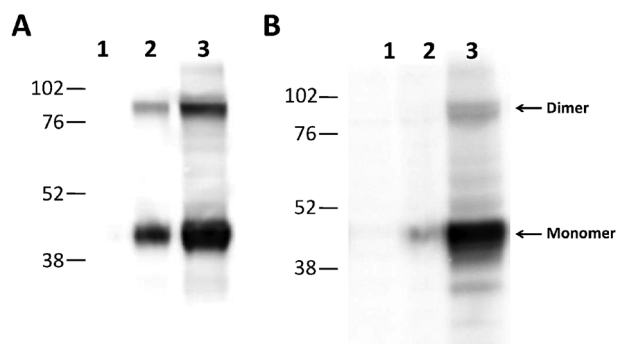
## Results

### Induction of hOR17-4T4L and hVN1R1-T4L expression in stable HEK293 cell lines

The native forms of hOR17-4 and hVN1R1 were previously expressed.<sup>13–15</sup> 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 1  $\mu\text{g ml}^{-1}$  of tetracycline with 5 mM sodium butyrate, and for hVN1R1 were 1  $\mu\text{g ml}^{-1}$  of tetracycline with 1 mM 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 (Fig. 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.

### 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.<sup>21–27</sup> 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 (Fig. 2). Non-induced cells yielded no signal, while



**Fig. 1** Expression of hOR17-4T4L and hVN1R1-T4L in HEK293 cells. (A) Western blot of hOR17-4T4L, and (B) Western blot of hVN1R1-T4L solubilized from HEK293 cells using FC-14. Lane 1 shows receptor solubilized from uninduced cells, lane 2 shows receptor solubilized from cells induced with 1  $\mu\text{g ml}^{-1}$  of tetracycline for 48 hours, and lane 3 shows receptor solubilized from cells induced with 1  $\mu\text{g ml}^{-1}$  of tetracycline and 5 mM 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 5 mM for hOR17-4T4L, and 1 mM for hVN1R1-T4L.

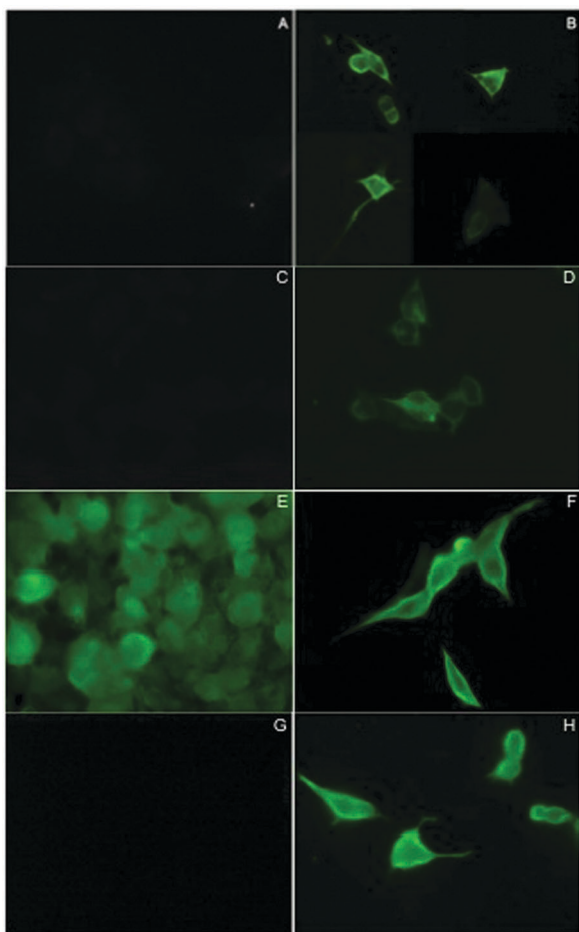
induced cells showed receptors that were localized to the cell membrane. Permeabilized and non-permeabilized cells both demonstrated membrane localization, suggesting that not all receptors are inserted in the correct orientation. However, visualization of both native and T4L-variants in the membrane suggests that the T4L insert did not affect GPCR trafficking. The results also suggest that the glycosylation sites are not necessary for localization of the protein to the membrane.

### Functional characterization of hOR17-4T4L in HEK293 cells

The functional activity of hOR17-4T4L was probed by measuring intracellular  $\text{Ca}^{2+}$  signaling in response to the specific odorant bourgeonal.<sup>12,28</sup> In HEK293 cells, hOR17-4, like other olfactory receptors, can couple to the promiscuous G protein  $G_{\alpha q}$  to initiate a signal through the inositol triphosphate ( $\text{IP}_3$ ) pathway.<sup>13,15,29</sup> Similar results were observed with hOR17-4T4L: application of 50  $\mu\text{M}$  bourgeonal resulted in a transient increase in intracellular  $\text{Ca}^{+2}$  concentration (Fig. 3). However, this  $\text{Ca}^{+2}$  response was 60–70% lower than the native protein. Also, the time needed to return to baseline  $\text{Ca}^{+2}$  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 in a reduced capacity. The receptors hVN1R1 and hVN1R1-T4L were not tested due to the higher levels of cell toxicity after expression.

### Systematic detergent screening for receptor solubilization

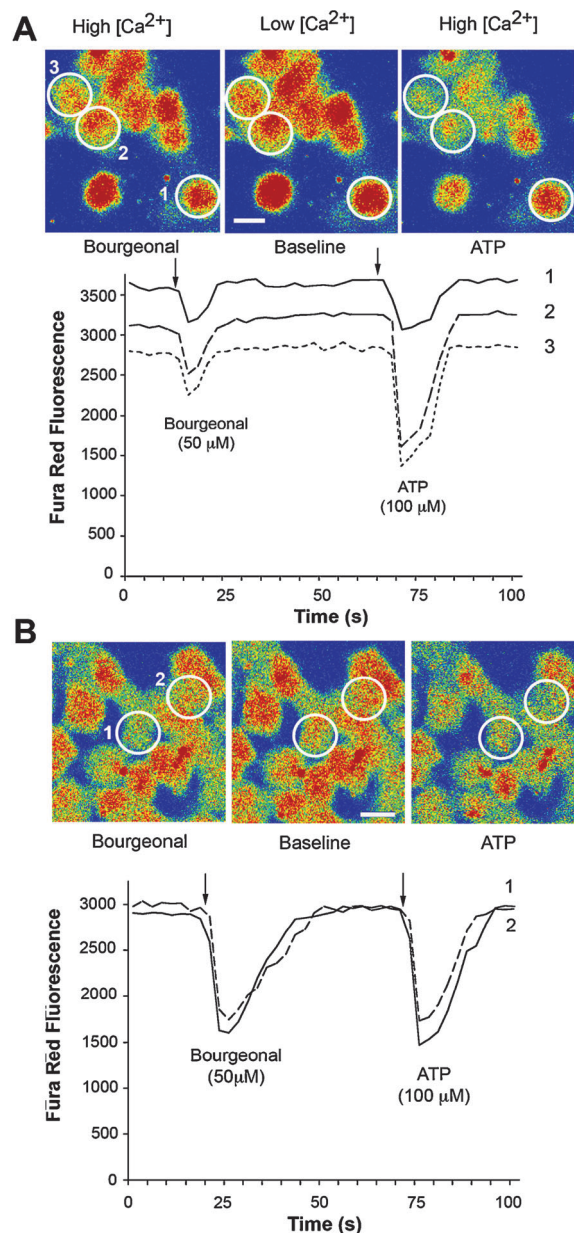
An appropriate detergent must be used to solubilize receptors from cell membranes, and keep them stable and functional in solution. A confounding factor is that different detergents may need to be used even for similar or related proteins, and must therefore be empirically determined. Thus, although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



**Fig. 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. (A) and (B) are from ref. 14, and are shown here for comparison.

determined,<sup>13,14</sup> 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 1). 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.<sup>14,17</sup> 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 cyclo-fos 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



**Fig. 3**  $\text{Ca}^{+2}$  imaging 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  $\text{Ca}^{+2}$  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  $\text{Ca}^{+2}$  concentration is recorded as a function of time (A and B, bottom). The arrows indicate the times at which, initially the odorant bourgeonal, and subsequently ATP were applied. 100  $\mu\text{M}$  ATP was used as a control for cell viability. 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  $\mu\text{m}$ .

structures. 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

**Table 1** Detergent solubilization of hVN1R1, hVN1R1-T4L, hOR17-4, and hOR17-4T4L\*

#	Detergent	Type*	CMC(%)**	Spot Intensity Average			
				hVN1R1	hVN1R1-T4L	hOR17-4	hOR17-4T4L
1	ANAMEG®-7	N	0.65				
2	ANAPOE®-20	N	0.0072				
3	ANAPOE®-35	N	0.001				
4	ANAPOE®-58	N	0.00045				
5	ANAPOE®-80	N	0.0016				
6	ANAPOE®-C10E6	N	0.025				
7	ANAPOE®-C10E9	N	0.053				
8	ANAPOE®-C12E8	N	0.0048				
9	ANAPOE®-C12E9	N	0.003				
10	ANAPOE®-C12E10	N	0.2				
11	ANAPOE®-C13E8	N	0.0055				
12	ANAPOE®-X-100	N	0.015				
13	ANAPOE®-X-114	N	0.011				
14	ANAPOE®-X-305	N	-				
15	ANAPOE®-X-405	N	0.16				
16	ANZERGENT® 3-10	Z	1.2				
17	ANZERGENT® 3-12	Z	0.094				
18	ANZERGENT® 3-14	Z	0.007				
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-DODECAFOS™	Z	0.77				
26	CYCLOFOS™-4	Z	0.45				
27	CYCLOFOS™-5	Z	0.15				
28	CYCLOFOS™-6	Z	0.094				
29	CYCLOFOS™-7	Z	0.022				
30	CYLU®-3	N	0.86				
31	CYMAL®-4	N	0.37				
32	CYMAL®-5	N	0.12				
33	CYMAL®-6	N	0.028				
34	CYMAL®-7	N	0.0099				
35	2,6-Dimethyl-4-heptyl-β-D-maltopyranoside	N	1.2				
36	2-Propyl-1-pentyl maltopyranoside	N	1.9				
37	FOS-CHOLINE®-9	Z	1.2				
38	FOS-CHOLINE®-10	Z	0.35				
39	FOS-CHOLINE®-11	Z	0.062				
40	FOS-CHOLINE®-12	Z	0.047				
41	FOS-CHOLINE®-13	Z	0.027				
42	FOS-CHOLINE®-14	Z	0.0046				
43	FOS-CHOLINE®-15	Z	0.0027				
44	FOS-CHOLINE®-16	Z	0.00053				
45	FOS-CHOLINE®-ISO-9	Z	0.99				
46	FOS-CHOLINE®-ISO-11	Z	0.9				
47	FOS-CHOLINE®-ISO-11-6U	Z	0.87				
48	FOS-CHOLINE®-UNSAT-11-10	Z	0.21				
49	FOS-MEA®-6	A	0.59				
50	FOS-MEA®-10	A	0.15				
51	FOSFEN™-9	Z	0.014				
52	MEGA-8	N	2.5				
53	NOPOL-FOS™	Z	1.4				
54	PMAL™-CS	Z	-				
55	PMAL™-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-α-D-maltopyranoside	N	-				
65	n-Decyl-β-D-maltopyranoside	N	0.087				
66	n-Undecyl-α-D-maltopyranoside	N	0.029				
67	n-Undecyl-β-D-maltopyranoside	N	0.029				
68	n-Dodecyl-α-D-maltopyranoside	N	0.0076				
69	n-Dodecyl-β-D-maltopyranoside	N	0.0087				
70	n-Tridecyl-β-D-maltopyranoside	N	0.0017				
71	n-Heptyl-β-D-thioglucopyranoside	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-Undecyl-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 dodecanoyl sarcosine	A	0.42				
81	Hexaethylene glycol monoethyl ether (C8E6)	N	0.39				
82	Octaethylene glycol monododecyl ether (C12E8)	N	0.0048				
83	Pentaethylene glycol monododecyl ether (C10E5)	N	0.031				
84	Tetraethylene glycol monoethyl ether (C8E4)	N	0.25				
85	Sucrose monododecanate	N	0.016				
86	Dimethyldecylphosphine 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	Digitonin	N	-				
92	Digitonin (8%)+Cholate (2%)	N	-				
93	CHAPS (6%)+CHS (1.2%)+DDM (10%)	N	-				
94	CHAPS (10%)+CHS (2%)+DDM (2%)	N	-				
95	CHAPS (10%)+CHS (2%)	N	-				
96	CHAPS (10%)+OG (1%)	N	-				

\*Detergent type, A (anionic detergent), C (cationic detergent), N (non-ionic detergent) and Z (zwitter-ionic detergent), \*\*CMC (critical micelle concentration) in H<sub>2</sub>O from the Detergent Properties List (<http://www.aquatec.com>), \*\*\*concentration of detergent used in the screening.

the choline nitrogen have been replaced with hydrogens. The anzergerents 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.

#### 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.<sup>13–15</sup> While UV absorption at 280 nm typically showed three distinct peaks for the native proteins, it usually showed only two for the T4L variants. Western blot analysis demonstrated that the three native peaks corresponded to aggregated, dimerized, and monomeric forms of the protein.<sup>13,14</sup> Similar analyses of the T4L-variants showed that the peaks corresponded to dimerized and monomeric protein forms (Fig. 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 a 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-variant batch will be usable for crystallization screens. Both native and T4L-variant proteins were able to yield about 1 mg 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

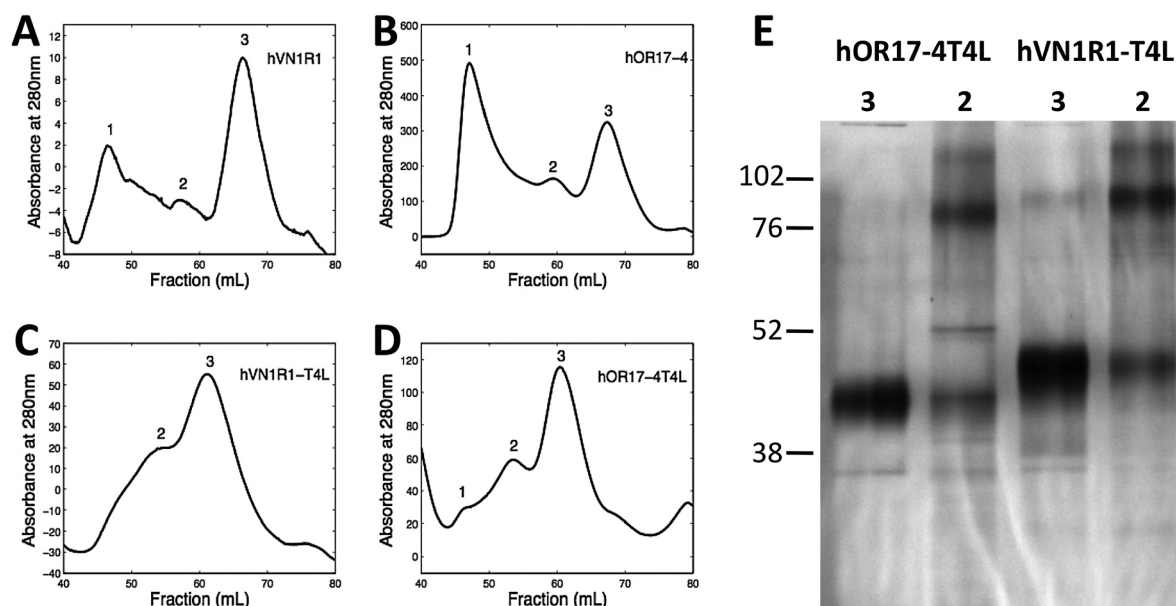
the soluble monomer for the T4L variants. In contrast, typically ~10–40% of the protein recovered after SEC was the soluble monomer for the native ones.

#### Structural characterization of purified hOR17-4T4L and hVN1R1-T4L

The secondary structure of the T4L variants was assayed using circular dichroism (CD). Far UV spectra of both proteins have characteristic alpha-helical spectra, similar to the native proteins (Fig. 5). However, the peaks at 208 nm and 220 nm are more defined in the T4L variants. 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 nm and 220 nm, while a protein that has a higher percentage of loops and random coils will have peaks that are less pronounced.<sup>30</sup> These results thus suggest that the T4L insert may help to structurally stabilize the proteins.

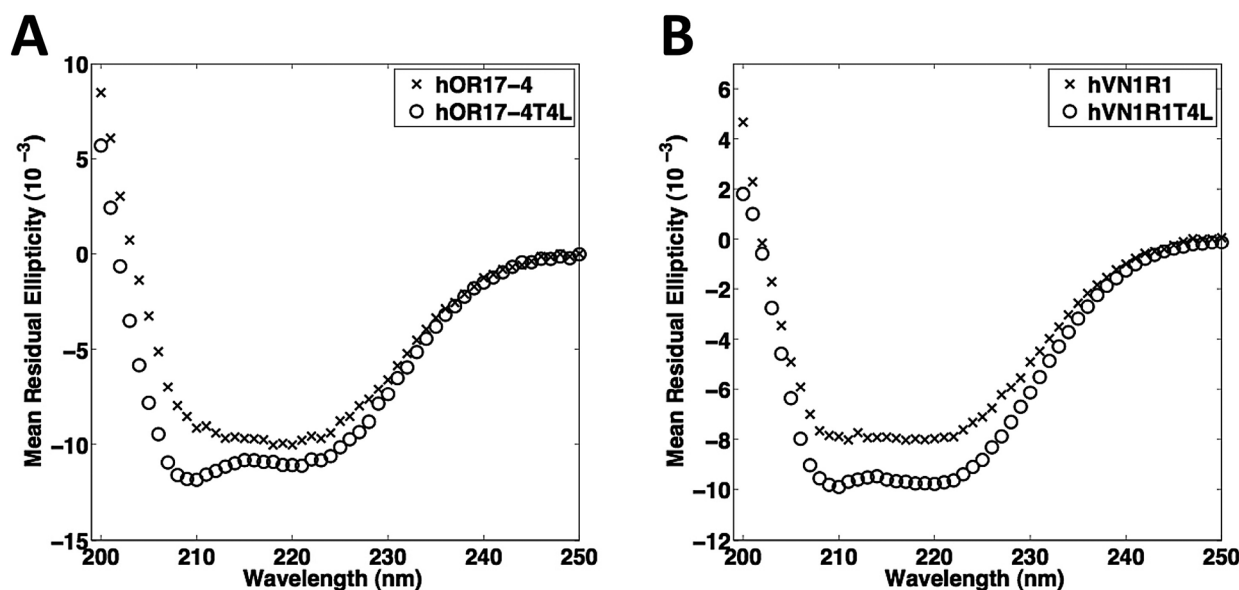
#### Ligand binding analysis of purified hOR17-4T4L and hVN1R1-T4L

Microscale thermophoresis (MST) was used to determine whether the native proteins and T4L variants could bind their ligands. Microscale thermophoresis is a technique that induces a local spatial temperature gradient. Proteins in the gradient migrate along it. Free receptors have a different migration speed compared to receptors that are bound to their ligands, likely due to changes in their hydration shell.<sup>18,31</sup> MST can measure binding of ligands as small as individual ions,<sup>19</sup> which is critical because odorants are less than 300 Da while their receptors are larger than 30 kDa. MST can also be used in receptors in solution, and can detect fluorescence from native



**Fig. 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 variants 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 ref. 14 and 15.



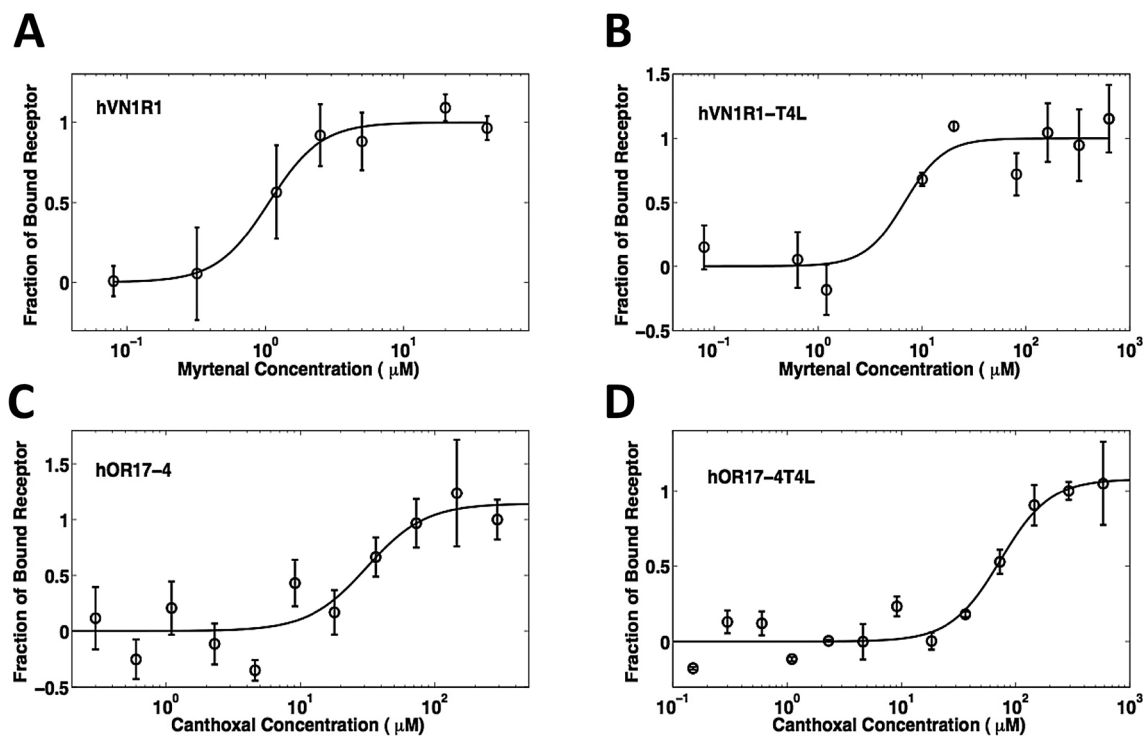


**Fig. 5** Circular dichroism spectra of native and T4L GPCRs. (A) hOR17-4 and hOR17-4T4L. (B) hVN1R1 and hVN1R1-T4L. The native and T4L variants have alpha-helical secondary structures. However, the peaks are more pronounced in the T4L variants, suggesting that the structure is more stabilized.

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 variant proteins were titrated against their known ligands. For both proteins, the T4L variant had a lower ligand-binding affinity than the native receptor, but was still able

to bind the ligand (Fig. 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.



**Fig. 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 variants 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 ref. 14 and 20. 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. (A) is from ref. 14 and is shown here for comparison.



## Discussion

A significant challenge in the field of GPCR structural biology 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.<sup>32</sup> Insertion of T4 lysozyme in the third intracellular loop of GPCRs seems to be a promising approach, as 5 of the 7 structures to date were obtained using this strategy.<sup>3–10</sup> 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. Our 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 suggest that insertion of T4L could 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 fusion 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 receptor was the soluble monomer. In contrast, at least 50%, and up to 70%, of the recovered fusion protein was the 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 variants had more defined CD spectra, and hOR17-4T4L was able to bind its ligand 2 months after it was purified. This is particularly important, as several months may be needed for crystal growth.

Immunohistochemical data, calcium imaging assays, circular dichroism, and microscale thermophoresis suggest that insertion of T4L in the third intracellular loop does not completely disrupt protein structure and function. Both T4L variants 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, suggesting that they were properly folded. Indeed, the T4L variants had more defined peaks, suggesting that they might be more stable. Ca<sup>2+</sup> imaging assays in HEK293 cells demonstrated that signaling still occurred with the hOR17-4T4L variant although it was more limited. This is in stark contrast to the T4L variants of non-olfactory receptors. Of the currently determined structures, signaling assays in cells have only been performed on the

A<sub>2A</sub> adenosine and CXCR4 T4L fusion proteins.<sup>5,6</sup> In both constructs, no downstream signaling was observed. This difference is likely due to deletion of a greater portion of the third intracellular loop in the non-olfactory GPCRs. It may also be caused by the specific orientation of the T4L segment, and the resulting steric hindrance. Microscale thermophoresis measurements of purified receptors showed that the T4L variants had higher EC<sub>50</sub> values, but were still able to bind their small molecular ligands. Together, the Ca<sup>2+</sup> imaging and MST results suggest 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 is 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. Indeed, non-olfactory related GPCRs sometimes exhibited higher affinities in the T4L constructs.<sup>5</sup> Future experiments will be carried out to probe potentially altered receptor coupling to G proteins, as well as changes in second messenger signaling.<sup>33–35</sup>

Structural knowledge of GPCRs and other membrane proteins is a prerequisite for the design of specific therapies or biologically inspired sensing technologies. Insertion of T4 lysozyme in the third intracellular loop seems to be a promising strategy for GPCR studies, as five of the seven crystallized GPCRs have a T4L insertion. The results presented here further support this. Furthermore, they open the possibility that T4L insertion may facilitate structural studies of a wider range of 7TM proteins, and potentially other membrane proteins.

## Acknowledgements

We thank members of Zhang Laboratory for stimulating discussions.

## References

- 1 E. Wallin and G. von Heijne, *Protein Sci.*, 1998, **7**(4), 1029–1038.
- 2 P. J. Loll, *J. Struct. Biol.*, 2003, **142**(1), 144–153.
- 3 V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens, *Science*, 2007, **318**, 1258–1265.
- 4 E. Y. Chien, W. Liu, Q. Zhao, V. Katritch, G. W. Han, M. A. Hanson, L. Shi, A. H. Newman, J. A. Javitch, V. Cherezov and R. C. Stevens, *Science*, 2010, **330**, 1091–1095.
- 5 V. P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman and R. C. Stevens, *Science*, 2008, **322**, 1211–1217.
- 6 B. Wu, E. Y. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov and R. C. Stevens, *Science*, 2010, **330**, 1066–1071.
- 7 T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G. W. Han, T. Kobayashi, R. C. Stevens and S. Iwata, *Nature*, 2011, **475**, 65–70.
- 8 M. A. Hanson, *et al.*, *Science*, 2012, **335**, 851–855.
- 9 S. G. Rasmussen, H. J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. Schertler, W. I. Weis and B. K. Kobilka, *Nature*, 2007, **450**, 383–388.
- 10 T. Warne, M. J. Serrano-Vega, J. G. Baker, R. Moukhametzianov, P. C. Edwards, R. Henderson, A. G. Leslie, C. G. Tate and G. F. Schertler, *Nature*, 2008, **454**, 486–491.

- 
- 11 I. Rodriguez, C. A. Greer, M. Y. Mok and P. A. Mombaerts, *Nat. Genet.*, 2000, **26**, 18–19.
  - 12 M. Spehr, *et al.*, *Science*, 2003, **299**, 2054–2058.
  - 13 B. Cook, K. E. Ernberg, H. Chung and S. Zhang, *PLoS One*, 2008, **3**, e2920.
  - 14 K. Corin, P. Baaske, S. Geissler, C. Wienken, S. Duhr, D. Bruan and S. Zhang, *Sci. Rep.*, 2011, **1**, 172.
  - 15 B. Cook, D. Steuerwald, L. Kaiser, J. Graveland-Bikker, M. Vanberghem, K. Herlihy, H. Pick, H. Vogel and S. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11925–11930.
  - 16 E. Shirokova, J. D. Raguse, W. Meyerhof and D. Krautwurst, *FASEB J.*, 2008, **22**, 1416–1425.
  - 17 H. Ren, D. Yu, B. Ge, B. Cook, Z. Xu and S. Zhang, *PLoS One*, 2009, **4**, e4509.
  - 18 P. Baaske, C. J. Wienken, P. Reineck, S. Duhr and D. Braun, *Angew. Chem., Int. Ed.*, 2010, **49**, 2238–2241.
  - 19 C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun and S. Duhr, *Nat. Commun.*, 2010, **1**, 100.
  - 20 K. Corin, P. Baaske, D. B. Ravel, J. Song, E. Brown, X. Wang, C. J. Wienken, M. Jerabek-Willemsen, S. Duhr, D. Braun and S. Zhang, *PLoS One*, 2011, **6**(11), e25067.
  - 21 S. Katada, M. Tanaka and K. Touhara, *J. Neurochem.*, 2004, **90**, 1453–1463.
  - 22 A. A. Gimelbrant, S. L. Haley and T. S. McClintock, *J. Biol. Chem.*, 2001, **276**, 7285–7290.
  - 23 M. Lu, L. Staszewski, F. Escheverri, H. Xu and B. D. Moyer, *BMC Cell Biol.*, 2004, **5**, 34.
  - 24 M. Lu, F. Echeverri and B. D. Moyer, *Traffic*, 2003, **4**, 416–433.
  - 25 S. Jayadev, R. D. Smith, G. Jagadeesh, A. J. Baukal and L. Hunyady, *Endocrinology*, 1999, **140**, 2010–2017.
  - 26 A. A. Gimelbrant, T. D. Stoss, T. M. Landers and T. S. McClintock, *J. Neurochem.*, 1999, **72**, 2301–2311.
  - 27 T. S. McClintock, *et al.*, *Mol. Brain Res.*, 1997, **48**, 270–278.
  - 28 M. Spehr, *et al.*, *J. Biol. Chem.*, 2004, **279**, 40194–40203.
  - 29 V. Jacquier, H. Pick and H. Vogel, *J. Neurochem.*, 2006, **97**, 537–544.
  - 30 N. Greenfield and G. D. Fasman, *Biochemistry*, 1969, **8**, 4108–4116.
  - 31 S. Duhr and D. Braun, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 19678–19682.
  - 32 R. M. Garavito, D. Picot and P. J. Loll, *J. Bioenerg. Biomembr.*, 1996, **28**, 13–27.
  - 33 H. Pick, S. Etter, O. Baud, R. Schmauder, L. Bordoli, T. Schwede and H. Vogel, *J. Biol. Chem.*, 2009, **284**, 30547–30555.
  - 34 O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel and H. Pick, *Biochemistry*, 2011, **50**, 843–853.
  - 35 T. Dahoun, L. Grasso, H. Vogel and H. Pick, *Biochemistry*, 2011, **50**, 7228–7235.