Published in final edited form as:

Protein Expr Purif. 2007 April ; 52(2): 348–355. doi:10.1016/j.pep.2006.10.017.

Expression and Purification of Milligram Levels of Inactive G-Protein Coupled Receptors in *E. coli*

Steven E. Bane, Javier E. Velasquez, and Anne Skaja Robinson

Department of Chemical Engineering, University of Delaware, 150 Academy Street, Newark, Delaware 19716 USA

Abstract

G-protein coupled receptors (GPCRs) are seven transmembrane helical proteins involved in cell signaling and response. They are targets for many existing therapeutic agents, and numerous drug discovery efforts. Production of large quantities of these receptors for drug screening and structural biology remains challenging. To address this difficulty, we sought to express genes for several human GPCRs in *E. coli*. For most of the receptors, expression was poor, and was not markedly improved even in strains designed to compensate for differences in codon bias between human and *E. coli* genes. However, the gene for human NK₁ receptor (hNK₁R) was expressed in large quantities as inclusion bodies in *E. coli*. The inclusion bodies were not soluble in chemical denaturants such as guanidine chloride or urea, but were soluble in ionic detergents such as SDS, and the zwitterionic detergent fos-choline. Using immobilized metal affinity chromatography, we purified milligram amounts of hNK₁R. Although inactive in ligand binding assays, purified hNK₁R in fos-choline micelles appeared to have a high content of alpha-helix, and was well-behaved in solution. Thus this protein is suitable for additional biophysical characterization and refolding studies.

Introduction

Integral membrane proteins are believed to represent more than 50 % of all potential drug targets and to comprise approximately 30 % of all proteins, yet they account for only 0.3 % of available high-resolution protein structures [1, 2]. Among the many obstacles to achieving X-ray or NMR structures of integral membrane proteins, the ability to produce large quantities of active protein is often cited as the most challenging [1]. This difficulty also hampers drug screening, as well as biochemical biophysical analyses of function and conformation. The first roadblock is achieving high levels of expression. To date, studies of membrane protein expression have not yielded any systematic approach. Current methods for membrane protein expression rely mostly on trial and error [1, 3].

© 2006 Elsevier Inc. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

When expressing genes for mammalian membrane proteins in *E. coli*, there are two general strategies. One is to produce the membrane protein in its functional form; this approach usually results in low-level expression [4]. The second is to express the membrane protein as inclusion bodies; these must then be solubilized and then refolded to produce the native, functional form of the protein. This second method has some advantages. Expression of inclusion bodies reduces the toxicity to the cell during expression. Inclusion bodies are easier to purify and are resistant to proteolytic degradation. Most importantly, *E. coli* has the potential to produce as much as hundreds of milligram of inclusion bodies per liter of culture. Some outer membrane proteins (omp) such as the *E. coli* ompF and ompX and some small transmembrane α -helical proteins such as the *E. coli* diacylglycerol kinase, a three transmembrane α -helical protein, have been successfully produced using this approach [5–10]. Using members of the G-protein coupled receptor (GPCR) superfamily, we sought to test whether expression of membrane proteins as inclusion bodies in *E. coli* was routine, and whether sufficient quantities of receptor could be purified to enable refolding studies.

GPCRs are seven transmembrane α -helical proteins that are involved in signaling and response to external stimuli in nearly all human cell types [11]. They are attractive and proven drug targets in a wide range of therapeutic areas. GPCRs have also been recombinantly expressed in *E. coli* using both strategies. Expression of functional GPCRs in *E. coli* has been achieved by fusing the *E. coli* maltose-binding protein (MBP) (including its signal peptide) to the N-terminus of the GPCR; approximately 150 micrograms of functional human adenosine A_{2a} receptor were produced per liter of culture [12]. The same method produced 50 micrograms of functional rat neurotensin receptor per liter of culture [13–17]. The leukotriene B₄ receptor (hBLT₄) was expressed as inclusion bodies with yields of 10 mg of inclusion bodies per liter of culture prior to refolding [18]. In this case, successful expression required optimizing the codon usage in the gene to match *E. coli* codon bias [18].

In order to test whether a general strategy to express GPCRs as inclusion could be developed, we tested expression of nine GPCRs in three *E. coli* strains. For only two receptors, human NK₁ receptor (hNK₁R) and mouse NK₁ receptor (mNK₁R) were we able to expressed large quantities of inclusion bodies. In no case, even for hBLT₄ receptor, was expression of the GPCR genes improved by the use of *E. coli* Rosetta™ strains, which are designed to compensate for differences in codon usage by the expression of rare tRNAs.

The hNK₁R inclusion bodies were unusual in that they could not be solubilized with urea or guanidine hydrochloride; typically inclusion bodies are readily dissolved in these chaotropic agents. However, the inclusion bodies were solubilized by SDS, as well as by the zwitterionic surfactant fos-choline 16. The protein was purified to homogeneity by immobilized metal affinity chromatography (IMAC). The purified hNK₁ receptor was inactive in ligand binding studies, but found to be highly helical based on circular dichroism (CD) spectroscopy. Thus, in this work we have developed a method for isolating large quantities of inactive hNK₁R that can be used for refolding trials and further conformational analysis.

Materials and methods

Materials

n-dodecyl β -D-maltoside, octyl glucoside, and Sarkosyl were obtained from Sigma (St. Louis, MO); Empigen (dodecyl-N,N-dimethylglycine, Calbiochem, San Diego, CA) and Fos-choline 16 were obtained from Anatrace (Maumee, OH). Sodium Dodecyl sulfate (SDS) was obtained from Bio-Rad (Hercules, CA). Ni-NTA resin was obtained from Qiagen (Valencia, CA). Urea, 99 % purity, was obtained from Fisher Scientific (Pittsburgh, PA) and Guanidine-HCl, 99.9 % purity, was obtained from MP Biomedical (Auro, OH).

Plasmids and strains

Escherichia coli strains BL21(DE3), BL21(DE3)/pLysS, BL21(DE3)/pLysE, Rosetta(DE3), Rosetta(DE3)/pLysS, Rosetta(DE3)/placI (Novagen) and C43(DE3) [19] were used as expression hosts. *E. coli* strain DH5 α (Novagen) was used for subcloning and plasmid amplification. The pET23b (Novagen) was used as an expression plasmid.

Genes

The human adenosine receptors genes (hA₂aR, hA₃R), and the human follicle stimulating hormone receptor gene (hFSHR) were gifts from Dr. Marlene Jacobson (Merck). The human chemokine receptor genes (hCCR₅ and hCXCR₄) were gifts from Dr. Joseph Sodroski (Harvard). The genes for mouse neurokinin receptor and human neurokinin receptor (mNK₁R and hNK₂R) genes were obtained from the American Type Culture Collection. A form of human neurokinin receptor gene with a 42-amino-acid C-terminal truncation (hNK₁R) and the human leukotriene receptor gene (hLTB₄R) were purchased from the University of Missouri at Rolla cDNA Resource Center.

Subcloning into *Escherichia coli* expression vectors

The hA₂aR, hA₃R, hNK₁R, mNK₁R, hCCR₅R, hCXC₄R and hFSHR genes were subcloned into the pET23b vector using the *Nde*I and *Xho*I sites. The hNK₂R and hLTB₄R genes were subcloned into the pET23b vector using *Nde*I and *Not*I; using the *Nde*I site eliminated the T7 tag. The genes were placed in frame to allow the encoding of a C-terminal hexa-histidine tag to facilitate purification.

Analysis of GPCR Expression

Chemically competent strains of each *E. coli* cell line were transformed with the pET23b vector containing the various GPCR genes. Successful transformants were selected by growth on Luria-Bertani (LB) plates containing 100 μ g/ml ampicillin for BL21 and C43 strains, or 100 μ g/ml Ampicillin and 34 μ g/ml Chloramphenicol for Rosetta pLysS and pLysE strains. Cells were subsequently grown overnight at 37 °C in 5 ml LB media containing 50 μ g/ml ampicillin, and containing 34 μ g/ml chloramphenicol when appropriate for the strain in a 37 °C water shaker. The overnight culture was used to inoculate a fresh 5 ml culture at an OD₆₀₀ of ~ 0.1 and grown at 37 °C. When the optical density reached 0.6, isopropyl- β -D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce expression. Cells were allowed to grow for 4 additional hours and were then

harvested by centrifugation for 5 min at 10,000g. The cell pellet was resuspended in 250 μ l of lysis buffer (50 mM sodium phosphate 10 mM EDTA, 0.1 % (w/v) Triton X-100, 1 mM PMSF and 100 μ g/ml lysozyme, pH 8.0). The cell suspension was treated with three freeze/thaw cycles ($-80^{\circ}\text{C}/37^{\circ}\text{C}$). After the third thaw, MgCl_2 was added to a final concentration of 20 mM, and DNase (Deoxyribonuclease I, Sigma, St. Louis MO) was added to a final concentration of 20 μ g/ml. The samples were incubated at 25°C for 5 min, and then centrifuged at 16,000g for 10 min to separate the inclusion bodies from the soluble protein. The pellet (containing inclusion bodies and cell debris) was treated with 250 μ l solubilization buffer (20 mM Tris-HCl, 1.0 % (w/v) sodium dodecyl sulfate (SDS), pH 8.0). Samples were analyzed by SDS-PAGE and Western Blot.

Either 10 or 12 % polyacrylamide gels were cast for SDS-PAGE and Western blots analysis. 20 μ l of sample was added to 10 μ l of 3X sample buffer (163 mM Tris-HCl pH 6.8, 0.25 mg/ml Bromophenol blue, 0.5 % SDS, 50 % glycerol). Samples were not boiled prior to loading as aggregation of GPCRs was observed, as is typical for hydrophobic proteins. The gels were electrophoresed at 25°C at 150 V for 1.5 hours in SDS running buffer (25 mM Tris-HCl, 192 mM glycine, 1 % SDS, pH 8.3) followed by Coomassie staining (10 % acetic acid, 10% 2-propanol, 0.25 mg/ml coomassie blue for 4 hours and then 10 % acetic acid, 0.25 mg/ml coomassie overnight). The gels were destained with 10 % acetic acid. Images detected from the staining were digitized at a resolution of 300 dpi using an Epson flatbed scanner.

For Western blots, after electrophoresis the gel was first incubated in transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% methanol, pH 8.2) for 20 minutes, then proteins were transferred from the gel to nitrocellulose paper at 25 V at 4°C overnight. The proteins were detected using a primary antibody of anti-His₆ IgG antibody (Covance, Princeton, NJ) at a 1:1000 dilution in Tris buffered saline (20 mM Tris-HCl, 150 mM NaCl, pH 7.6) containing 0.1 % Tween. The secondary antibody was anti-mouse IgG horseradish peroxidase (GE Bioscience, Piscataway, NJ), diluted to 1:10,000 in Tris buffered saline containing 0.1% Tween. Immunoreactive bands were detected using the ECL Chemiluminescence Detection Kit (GE Bioscience, Piscataway, NJ) and imaged with a Typhoon 9600TM variable mode imager (GE Bioscience, Piscataway, NJ).

Solubilization of NK₁R Inclusion Bodies

To test solubility of the inclusion bodies under various conditions, 5 ml cultures were used to express the human neurokinin receptor in the RosettaTM (DE3)/pLysS cell line. The inclusion bodies were treated by mixing vigorously 250 μ l solubilization buffer (20 mM Tris-HCl pH 8.0) with 8 M Urea or 1 (w/v) % of different detergents. The samples were incubated for 30 minutes at 25°C before centrifuging at 16,000g for 10 minute to remove insoluble material. The samples were analyzed by SDS-PAGE as described above.

Expression and Purification of hNK₁R

In order to express and purify hNK₁R, a single colony of freshly transformed cells was used to inoculate a 5 ml culture of LB media containing 50 μ g/ml Ampicillin, and 34 μ g/ml Chloramphenicol and grown at 37°C overnight. This overnight culture was used to

inoculate a fresh 50 ml culture of the same media to an OD₆₀₀ of ~ 0.1, and grown at 37 °C. When the OD₆₀₀ reached 0.6, expression was induced by the addition of 1 mM IPTG. Induction was carried out for 4 hrs and cells were harvested by centrifuging for 10 min at 10,000g. The cell pellet was resuspended in 1 ml of lysis buffer, and the cells were lysed by freeze/thaw cycles as described above. The samples were centrifuged for 10 min at 16,000g, and the pellet (containing inclusion bodies and cell debris) was collected. The pellets were treated with 1 ml of binding buffer (50 mM Tris-HCl, 300 mM NaCl, 10 % glycerol, pH 8.0) containing 1 % (w/v) fos-choline 16 for 30 minutes at 25 °C to solubilize the inclusion bodies. After centrifuging for 10 min at 16,000g, the supernatant containing solubilized protein was collected, and the pellet (containing cell debris and any remaining insoluble material) was discarded.

Prior to initiating purification, 250 µl of packed Ni-NTA resin (Qiagen) was preequilibrated for 5 minutes at 25 °C with 1 ml binding buffer containing 0.1 % fos-choline 16. The resin was then centrifuged briefly (4,000g for 5 min) to remove the binding buffer. The resin was incubated in batch mode with the soluble protein fraction in 1 mL of binding buffer containing 1 % fos-choline 16 and tumbled gently at 4 °C overnight. The resin was centrifuged briefly (4,000g for 5 min) to remove any protein that was not bound to the resin. The resin was washed twice in batch with 1 ml of binding buffer containing 0.1% fos-choline 16 and 10 mM imidazole and then twice in batch with 1 ml of binding buffer containing 0.1% fos-choline 16 and 20 mM imidazole to remove non-specifically bound proteins. Washes were removed by centrifuging the resin briefly (4,000g for 5 min). Purified hNK₁R protein was eluted using 1 ml of binding buffer containing 0.1 % fos-choline 16 and 500 mM imidazole. Samples were taken after each step for SDS-PAGE analysis.

The process was readily scaled up to a 1-L of culture. The expression procedure was similar; a single colony of freshly transformed cells was used to inoculate a 25 ml culture in a 250 ml flask of LB media containing 50 µg/ml Ampicillin, and 34 µg/ml Chloramphenicol and grown at 37 °C overnight. This overnight culture was used to inoculate a fresh 1 L culture in a 3 L flask and grown at 37 °C. When the OD₆₀₀ reached 0.6, expression was induced by the addition of 1 mM IPTG. Induction was carried out for 4 hrs and cells were harvested by centrifuging for 10 min at 10,000g. The cell pellet (~20 g wet weight) was resuspended in 20 ml of lysis buffer, and the cells were lysed by freeze/thaw cycles as described above. The samples were centrifuged for 10 min at 16,000g, and the pellet (containing inclusion bodies and cell debris) was collected. The pellets were treated with 20 ml of binding buffer (50mM Tris-HCl, 300 mM NaCl, 10 % glycerol, pH 8.0) containing 1 % (w/v) fos-choline 16 for 30 minutes at 25 °C to solubilize the inclusion bodies. After centrifuging for 10 min at 16,000g, the supernatant containing solubilized protein was collected, and the pellet (containing cell debris and any remaining insoluble material) was discarded. The ratio of 1 ml of binding buffer to 50 ml of culture was empirically determined to obtain the maximal solubilized hNK₁R. The purification was carried out in similar fashion to that described above with an optimized ratio of 250 µl of resin to 50 ml of culture. Prior to spectroscopic measurements, the protein was dialyzed in 20 mM phosphate buffer pH 7.4 containing 0.01 % fos-choline 16. Protein concentration was measured by UV spectroscopy. The extinction coefficient at 280 nm (ϵ_{280}) was determined to be 83130 cm⁻¹ M⁻¹ based on the Gill and von Hippel method [20].

Radioligand Binding Activity Assays

The ligand binding procedure was adapted from a method described previously for adenosine receptors [21]. All experiments were conducted in Millipore (Bedford, MA) 96 well glass fiber filter B plates, using the agonist [³H] Substance P (Perkin Elmer Biosciences). For determination of ligand binding, ~ 3.0 µg of total protein that was immobilized on Ni-NTA resin and was added to 200 µl ligand binding buffer (50 mM Tris-Cl, pH 7.4, 120 mM NaCl, 0.1% BSA) containing 4.4 nM of [³H] Substance P and substance P concentrations ranging from 6.8 pM to 700 nM. Samples were incubated at room temperature for 1.5 hrs. Unbound ligands were removed by washing 3 times at 4 °C in ligand binding buffer. Radioactivity was measured in counts per minute using a Perkin-Elmer 1450 Microbeta liquid scintillation and luminescence counter.

Circular Dichroism

Far-UV CD spectra were collected using a Jasco J-810 spectropolarimeter with Peltier control. The concentration of hNK₁R was 2–10 µM in 20 mM phosphate buffer, pH 7.4 containing 0.1% (w/v) fos-choline 16. Since the critical micelle concentration is ~ 0.013 mM (0.00053%; Anatrace), this value is ~ 185 times the CMC. A 0.1 cm path-length quartz cuvette was used for all measurements were taken at 25 °C. Spectra were collected from 190 nm to 260 nm using 1 nm step resolution with 3 s integrations. Background spectra containing buffer at the same composition were subtracted from all sample spectra to obtain corrected spectra. CD spectra are expressed in mean molar ellipticities [θ] (deg cm² dmol⁻¹):

$$[\theta] = \frac{\theta}{10lcn} \quad (1)$$

Where θ is the observed ellipticity (mdeg), l is the optical path length (cm), c is the protein concentration (M) and n is the number of residues in the protein. The percentage of α -helical content was estimated from the observed molar ellipticity at 222 nm ($[\theta]_{222}$) as follows:

$$\% \alpha - helix = \frac{[\theta]_{222}}{[\theta]_{222}^{\max}} \left(1 - \frac{k}{n} \right) \quad (2)$$

$[\theta]_{222}^{\max}$ is the theoretical mean molar ellipticity at 222 nm for a helix of infinite length (−39,500 deg cm² dmol⁻¹), n is the number of residues in the protein and k is a wavelength-dependent constant (2.57 for 222 nm).

Results and Discussion

Analysis of GPCR Expression

In order to determine whether GPCRs could be easily expressed as inclusion bodies, nine GPCRs were subcloned into the pET23b vector. Each was transformed into BL21(DE3), Rosetta (DE3), Rosetta(DE3)/pLysS and BL21(DE3)/pLysS cell lines and grown in 5 mL cultures. For each culture, samples of the insoluble fraction that were obtained from low speed centrifugation (< 30,000g) (which would contain cell debris and any putative inclusion bodies) were analyzed by SDS-PAGE (Figure 1a). Of the nine receptors, only the

human and mouse NK₁ receptors were expressed in Rosetta(DE3)/plysS cells at high enough levels to produce a prominent band on Coomassie-stained SDS-PAGE (Figure 1a). Western blot analysis using an anti-His₆ antibody supported the conclusion that the bands are indeed NK₁ receptors (Figure 1b). Western blot analysis did not detect expression of any of the other receptors (Figure 1b), indicating that they are present at very low levels, if at all. Similar results were observed using BL21(DE3)/pLysS cell lines. None of the receptors were able to express to any extent in the BL21 (DE3) and Rosetta (DE3) cell lines alone (data not shown).

Cells were induced at an OD₆₀₀ $\approx$ 0.6. For each receptor, and in every cell line, there was little or no growth following induction, regardless of whether expression of the receptor was observed by SDS-PAGE. This observation suggests that some cell death occurred in the cells. In all cases, we used plasmid rescue and sequencing to verify that the plasmids were retained and the GPCR genes were still present after induction and growth.

For human and mouse NK₁R, the only receptors for which expression was observed, the final OD₆₀₀ was $\sim$ 1.5. This finding indicates that even expressing NK₁R as inclusion bodies severely stressed the cells.

Expression of hNK₁ and hLTB₄ receptors in various cell lines

To help elucidate what factors govern whether a receptor is expressed well in *E. coli*, we characterized the expression behavior of two receptors, hNK₁R and hLTB₄R, in *E. coli* in more detail. We evaluated expression in C43(DE3) cells, in BL21 and Rosetta cells without pLysS.

BL21(DE3)/pLysS, Rosetta(DE3)/pLysS and C43(DE3) cell lines all overexpressed hNK₁R as inclusion bodies, and at similar levels (Figure 2). However, expression of the hLTB₄ receptor could not be detected in any of the cell strains, either by Coomassie-stained SDS-PAGE or Western blot (data not shown).

Differences in codon bias between humans and *E. coli* are thought to make expression of human proteins difficult when the human gene contains many rare *E. coli* codons, especially when rare codons are repeated, or when there are many rare arginine codons [22]. The Rosetta™ (EMD Biosciences) strains are designed to address this problem by expressing additional tRNAs for rare *E. coli* codons. Since the hNK₁R and hLTB₄R genes contain many such rare *E. coli* codons (19 in hNK₁R, 21 in hLTB₄R) we expected that expression of the two proteins might be improved in the Rosetta strains. However, there was no noticeable difference in expression between the BL21(DE3)/pLysS and Rosetta(DE3)/pLysS strains; both strains overexpressed hNK₁R to high levels (Figure 2). Thus, it appears that the presence of rare codons, or lack of sufficient tRNAs for these rare codons, does not limit hNK₁R expression in *E. coli* BL21.

Interestingly, expression of hLTB₄R was also not affected by the use of the Rosetta strain; expression of hLTB₄R was undetectable in both the Rosetta and BL21 strains, as well as in C43(DE3) cell lines (not shown). This result was surprising, since previous studies showed expression of this receptor was greatly enhanced when the rare codons in the hLTB₄R gene

were replaced with the codons preferred in *E. coli* [18]. Among the 21 rare *E. coli* codons in the hLTB₄R gene are two (non-repeating) rare arginine codons and one repeating set of rare proline codons; it was expected that the tRNAs provided by the Rosetta™ strains would compensate for the effects of these and other rare codons.

In BL21 and Rosetta strains lacking the pLysS plasmid, expression of hNK₁R decreased to undetectable levels (Figure 2); both strains appear to require the pLysS plasmid to overexpress hNK₁R. pLysS or pLysE plasmids have often been shown to enable expression of otherwise toxic proteins. The gene for T7 lysozyme is encoded by these plasmids; when expressed it acts as a natural inhibitor of the T7 RNA polymerase. This inhibition reduces any leaky expression of the target protein in the uninduced state [23]. The C43 cell strain is thought to aid cells in production of toxic proteins by decreasing the rate of mRNA production, although the specific mutations that cause this effect are not known [19].

Optimizing solubilizing agents for hNK₁R inclusion bodies

Chaotropic agents and surfactants are often used to solubilize inclusion bodies of membrane proteins. For example, both 8 M urea and SDS were used to solubilize the inclusion bodies of the photosynthetic light harvesting protein LHCII [24]. Empigen, a zwitterionic surfactant was the most effective in solubilizing the inclusion bodies of diacylglycerol kinase from *E. coli* [8]. As a first step in purification of the hNK₁R protein, we isolated the pellet (cell debris and any inclusion bodies) following expression and tested whether they could be solubilized in various chaotropic agents and surfactants.

Following harvesting of the cell culture, insoluble fractions containing the inclusion bodies were incubated with chaotropes or surfactants, the samples were centrifuged, and the supernatants analyzed on SDS-PAGE to detect the presence of hNK₁R in the soluble fraction (Figure 3). We found that 8 M urea and 6 M guanidine HCl could not extract hNK₁R from the inclusion bodies. Although non-ionic detergents such as maltosides and glucosides have been effective in solubilizing and stabilizing membrane proteins [25], we found them ineffective in solubilizing hNK₁R inclusion bodies (Figure 3). Ionic detergents such as SDS and Sarkosyl were able to solubilize the inclusion bodies (Figure 3). However, both compounds are ionic detergents and are incompatible with immobilized metal affinity chromatography (IMAC). Treating the inclusion bodies with SDS, then diluting with 8 M urea caused the receptor to precipitate. Therefore, we investigated other surfactants to find an agent compatible with IMAC that would effectively extract hNK₁R from the inclusion bodies.

Since ionic detergents have often been used to solubilize membrane proteins and inclusion bodies, we reasoned that a detergent series with some charge and some lipid-like character might work well in our case. The fos-choline series are lipid-like zwitterionic detergents and are compatible with IMAC. The detergents all have a phosphocholine headgroup, but possess a simple hydrophobic tail:

where n can vary between 7 and 15.

We found that fos-choline 13–16 detergents were all effective at solubilizing hNK₁R inclusion bodies, with fos-choline 16 (where n=15) being the best (Figure 3). Smaller chains in the series (fos-choline 9–12) failed to solubilize the inclusion bodies (data not shown). Similar trends were seen for mNK₁R (data not shown). Our current investigations are aimed at understanding what interactions are present in the inclusion bodies of membrane proteins, and identifying features of surfactants most well suited to disrupting these improper interaction and stabilizing the proteins in solution.

IMAC purification of hNK₁R

We then purified hNK₁R, to enable activity assays, conformational analysis, and ultimately refolding trials. 50 ml of *E. coli* Rosetta(DE3)/pLysS culture expressing hNK₁R was grown and harvested. The inclusion bodies were isolated, solubilized with fos-choline 16, and purified via Ni-NTA resin binding to the His6 tail on the recombinantly expressed receptor. Samples taken after each step were analyzed by SDS-PAGE (Figure 4). Following elution, hNK₁R appeared as a single band on Coomassiestained SDS-PAGE. N-terminal sequencing of the purified protein yielded a sequence of MDNVLPVDS, an exact match for the first nine amino acids of hNK₁R. Western blot analysis using an anti-His antibody confirmed that the 30 kDa band contained the expected C-terminal residues, and supports the conclusion that the protein is full length hNK₁R.

IMAC purification was efficient and resulted in a yield of ~ 5 mg pure protein per liter of culture. Protein concentration as measured UV spectroscopy (the extinction coefficient of 83130 cm⁻¹ M⁻¹) was 20 μM.

The purified hNK₁R receptor migrated at an apparent MW of ~30 kDa MW, smaller than its true molecular weight of 43 kDa. This altered mobility for membrane proteins has been described previously, particular for GPCRs [12, 17]. It probably results from the fact the membrane proteins do not fully denature in SDS-PAGE experiments – thus their mobility is not proportional to their molecular weight; protein conformation, shape, and charge all effect the proteins' mobility. This effect also causes the proteins' mobility to depend upon solution conditions and other aspects of sample composition. Indeed, hNK₁R migrates differently with respect to MW standards when run on SDS-PAGE in the insoluble fraction of cell extracts (Figures 1, 2 and 3) than when purified and solubilized in micelles of fos-choline 16 (Figure 4). Mouse NK₁R also migrated slightly faster than expected based on its molecular weight (not shown).

Analysis of Activity and Secondary Structure

Activity of the human NK₁R was determined by measuring its ability to bind the agonist [³H]-Substance P, as described in Materials and Methods. No specific binding to the ligand was observed (data not shown), indicating that the purified receptor is not functional, and suggesting that it is not correctly folded. Control samples of hNK₁R expressed in mammalian cells were active in our assay (not shown; samples provided by J. Winget).

We have initiated preliminary biophysical characterization of this inactive form of human NK₁R. Knowledge of the conformation of this protein may help us understand the differences between the structure of the active and inactive states, and guide refolding strategies.

The secondary structure of the human NK₁ receptor was investigated using circular dichroism (CD) spectroscopy. The CD spectrum displays a significant amount of α -helical structure indicated by the two distinct minima at 208 nm and 222 nm (Figure 5). Using this spectrum, the percentage of α -helical content of the inactive hNK₁R was estimated to be 38 % (Equation 2, Materials and Methods), close to the value of 41 % derived from secondary structure prediction algorithms (TM assignments from GPCR.org). This agreement might indicate that most of all of the helices have formed. However, the protein is inactive, suggesting that the helices are not in the proper conformation, perhaps due to improper (or missing) long-range interactions, or helix-helix packing.

Summary

Although intuitively, one might expect that any GPCR of interest could be readily expressed in inactive forms as inclusion bodies in *E. coli*, from our results it is clear that even inclusion body production of inactive receptors may be feasible only for a limited number of GPCRs. At present, we have been able to produce high levels only of NK₁ receptors. At the liter scale, the level of expression yields ~ 5 mg of purified inactive receptor per liter of cell culture. The other eight receptors we tested could not be expressed at detectable levels. It has been shown that expression of membrane proteins with strong promoters may cause cell death [19]. It will be interesting to identify what features of the genes or amino acid sequences correlate to the receptors' ability to be expressed in *E. coli*.

In all cases, the expression level was not altered in hosts that express tRNA for rare codons, regardless of whether the receptor was highly expressed (NK₁R) or poorly expressed (e.g. hLTB₄R). This result was particularly surprising, since optimizing the codon usage in the hLTB₄R gene has been reported to dramatically increase expression levels in *E. coli* [18, 26]. This result suggests that codon optimization may be more effective than supplementing rare tRNAs when expressing human genes in *E. coli*.

Although there have been several reports of integral membrane proteins expressed as inclusion bodies in *E. coli*, the seven transmembrane helices in GPCRs may contribute to the greater difficulties encountered in their expression; attempts to express *Mycobacterium tuberculosis* membrane proteins in *E. coli* showed that expression was suppressed for proteins containing large number of helices [18].

Although expression of GPCRs as *E. coli* inclusion bodies may be more challenging than originally anticipated, for some receptors, such as hNK₁R, it can generate very high levels of protein that can be solubilized in detergent micelles, making it suitable for further study. Thus, our work provides a platform for production of receptor than can be used in systematic refolding studies and biophysical characterization, which have been limited in the past by poor solubility and expression levels. It is hoped that such analysis of the inactive protein (and comparison to active forms of other GPCRs) will reveal conformational and structural features that will guide refolding efforts.

Acknowledgments

The authors are grateful to Dr. Clifford Robinson for careful reading of the manuscript. We thank Dr. David Raden and Dr. Sara Powers for interesting and helpful discussions, as well as Dr. Bryan Berger and Dr. Clifford Robinson for useful discussions of surfactants. The authors also thank Michelle O' Malley for her assistance with hLBT₄ experiments, and Jason Winget for assistance with ligand binding assays. These studies were supported by NIH RR15588.

References

1. Grishammer R, Tate CG. Preface: Overexpression of integral membrane proteins. *Biochem Biophys Acta*. 2003; 1610:1.
2. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov*. 2002; 1:727–730. [PubMed: 12209152]
3. Loll PJ. Membrane protein structural biology: The high throughput challenge. *J Struct Biol*. 2003; 142:144–153. [PubMed: 12718926]
4. Sarraemagna 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*. 2003; 60:1529–1546. [PubMed: 14513829]
5. Tian CL, Tobler K, Lamb RA, Pinto LH, Cross TA. Expression and initial structural insights from solid-state NMR of the M2 proton channel from influenza A virus. *Biochemistry*. 2002; 41:11294–11300. [PubMed: 12220196]
6. Ma C, Marassi FM, Jones DH, Straus SK, Bour S, Strebel K, Schubert U, Oblatt-Montal M, Montal M, Opella SJ. Expression, purification, and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Science*. 2002; 11:546–557. [PubMed: 11847278]
7. Rogl H, Kosemund K, Kuhlbrandt W, Collinson I. Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *Febs Letters*. 1998; 432:21–26. [PubMed: 9710243]
8. Gorzelle BM, Nagy JK, Oxenoid K, Lonzer WL, Cafiso DS, Sanders CR. Reconstitutive refolding of diacylglycerol kinase, an integral membrane protein. *Biochemistry*. 1999; 38:16373–16382. [PubMed: 10587463]
9. Bannwarth M, Schulz GE. The expression of outer membrane proteins for crystallization. *Biochimica Et Biophysica Acta-Biomembranes*. 2003; 1610:37–45.
10. Buchanan SK. Beta-barrel proteins from bacterial outer membranes: Structure, function and refolding. *Current Opinion in Structural Biology*. 1999; 9:455–461. [PubMed: 10449368]
11. Pierce KL, Premont RT, Lefkowitz RJ. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol*. 2002; 3:639–650. [PubMed: 12209124]
12. Weiss HM, Grishammer R. Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur J Biochem*. 2002; 269:82–92. [PubMed: 11784301]
13. Grishammer R, Averbeck P, Sohal AK. Improved purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem Soc Trans*. 1999; 27:899–903. [PubMed: 10830124]

14. Grisshammer R, Tate CG. Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys.* 1995; 28:315–422. [PubMed: 7480624]
15. Grisshammer R, White JF, Trinh LB, Shiloach J. Large-scale expression and purification of a G-protein-coupled receptor for structure determination – an overview. *J Struct Funct Genomics.* 2005; 6:159–163. [PubMed: 16211513]
16. Tucker J, Grisshammer R. Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem J.* 1996; 317(Pt 3):891–899. [PubMed: 8760379]
17. White JF, Trinh LB, Shiloach J, Grisshammer R. Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett.* 2004; 564:289–293. [PubMed: 15111111]
18. Baneres JL, Martin A, Hullot P, Girard JP, Rossi JC, Parello J. Structure-based analysis of GPCR function: Conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant Blt1. *J Mol Biol.* 2003; 329:801–814. [PubMed: 12787679]
19. Miroux B, Walker JE. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology.* 1996; 260:289–298. [PubMed: 8757792]
20. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem.* 1989; 182:319–326. [PubMed: 2610349]
21. Niebauer RT, Robinson AS. Exceptional total and functional yields of the human adenosine (A2a) receptor expressed in the yeast *Saccharomyces cerevisiae*. *Protein Expr Purif.* 2006; 46:204–211. [PubMed: 16289981]
22. Schenk PM, Baumann S, Mattes R, Steinbiss HH. Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. *Biotechniques.* 1995; 19:196–198. 200. [PubMed: 8527135]
23. Studier FW. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol.* 1991; 219:37–44. [PubMed: 2023259]
24. Plumley FG, Schmidt GW. Reconstitution of chlorophyll A/B light-harvesting complexes: Xanthophyll-dependent assembly and energy transfer. *Proc Natl Acad Sci U S A.* 1987; 84:146–150. [PubMed: 16593794]
25. Berger BW, Gendron CM, Robinson CR, Kaler EW, Lenhoff AM. The role of protein and surfactant interactions in membrane protein crystallization. *Acta D Crystallography.* 2005 in press.
26. Baneres JL, Parello J. Structure-based analysis of GPCR function: Evidence for a novel pentameric assembly between the dimeric leukotriene B4 receptor Blt1 and the G-protein. *J Mol Biol.* 2003; 329:815–829. [PubMed: 12787680]

Figure 1. Analysis of the expression of various GPCRs in the Rosetta (DE3) pLysS *E. coli* strain
 A) SDS-PAGE (10% acrylamide gel) shows that the only hNK₁ receptors express at high levels as inclusion bodies. Arrows indicate the positions of the bands corresponding to hNK₁R and mNK₁R. Expression of the empty pET23b vector in the same strain was used as a negative control. The target GPCR is indicated at the top of each lane. The left lane contains MW standards; sizes are shown to the left of the gel. B) Western blot analysis of GPCR expression. The target GPCR is indicated at the top of each lane. The 1^o antibody was anti-His₆ (Covance) used at 1:1000 dilution, and the 2^o antibody was anti-mouse horseradish peroxidase (Amersham Biosciences), used at 1:5000 dilution. The image was detected with a Typhoon 9600 (Amersham Biosciences).

Figure 2. Analysis of the expression of the human NK₁R in various *E. coli* strains
 SDS-PAGE (12% acrylamide) shows expression in the BL21(DE3), Rosetta(DE3), and C43(DE3) parental cell lines or those with plasmids to suppress high-level protein expression, such as pLysS, pLysE, or placI, as indicated. The arrow indicates the band corresponding to hNK₁R. The empty pET 23b vector that lacks the NK₁R gene was used as a negative control (Control). The left lane contains MW standards; sizes are shown to the left of the gel.

Figure 3. Analysis of the ability of various reagents to extract human NK₁R from inclusion bodies

SDS-PAGE (10 % acrylamide) shows proteins recovered in the soluble fraction following incubation of the insoluble fraction (containing hNK₁R inclusion bodies) with various surfactants or chaotropes. The arrow indicates the band corresponding to hNK₁R. Sarkosyl and Fos-choline surfactants can effectively extract hNK₁R from the inclusion bodies present in the insoluble fraction, with Fos-choline 16 being the most effective. SDS, Sodium dodecyl sulfate; β-DM, n-dodecyl β-D-maltoside; OG, Octyl Glucoside; EMP, Empigen; Trix, Triton X-100.

Figure 4. Analysis of human NK₁R purification by IMAC

SDS-PAGE (12 % acrylamide) shows samples from each step of the purification. (IB), solubilized proteins following extraction of insoluble fraction and inclusion bodies with Fos-choline 16; (FT), flow through; W1 and W2, 10 mM imidazole washes; W3 and W4, 20 mM imidazole washes; E1 and E2, Eluted fractions.

Figure 5. Circular dichroism spectroscopy of hNK₁R indicates presence of substantial amount of α -helix

Purified hNK₁R was diluted to 2.4 μ M in 10 mM phosphate buffer pH 7.4, 0.1% Fos-choline 16 (critical micelle concentration $\sim$ 0.00053%; Anatrace). The circular dichroism spectrum was measured as described in Materials and Methods.