Small expression tags enhance bacterial expression of the first three transmembrane segments of the apelin receptor

Aditya Pandey^a, Muzaddid Sarker^a, Xiang-Qin Liu^a, and Jan K. Rainey^{a,b,*}

^aDepartment of Biochemistry & Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

^bDepartment of Chemistry, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

Abstract

G-protein coupled receptors (GPCRs) are inherently dynamic membrane protein modulators of various important cellular signaling cascades. The apelin receptor (AR or APJ) is a class A GPCR involved in numerous physiological processes, implicated in angiogenesis during tumour formation and as a CD4 co-receptor for entry of human immunodeficiency virus type 1 (HIV-1) to cells. Due to the lack of efficient methods to produce full-length GPCRs enriched with nuclear magnetic resonance (NMR) active ¹⁵N, ¹³C and/or ²H isotopes, small GPCR fragments typically comprising 1-2 transmembrane segments are frequently studied using NMR spectroscopy. Here, we report successful overexpression of transmembrane segments 1-3 of AR (AR_TM1-3) in the C41(DE3) strain of *Escherichia coli* using an AT-rich gene tag previously reported to enhance cell-free expression yields. The resulting protein, with 6 additional N-terminal residues due to the expression tag, was purified using high performance liquid chromatography (HPLC). Far-ultraviolet circular dichroism spectropolarimetry demonstrates that AR_TM1-3 has the predicted ~40% α -helical character in membrane-mimetic environments. ¹H-¹⁵N HSQC NMR experiments imply amenability to high-resolution NMR structural characterization and stability in solution for weeks. Notably, this small expression tag approach may also be generally applicable to other membrane proteins that are difficult to express in *E. coli*.

Keywords

Apelin receptor; membrane protein expression and purification; GPCR; CD spectroscopy; NMR spectroscopy

Introduction

Membrane proteins are highly under-represented among available protein structures, at <2% of all structures in the protein data bank (PDB) (<http://blanco.biomol.uci.edu/mpstruc/#Latest>). In contrast, almost 26% of the protein-coding genes in the human genome code for membrane proteins (Fagerberg et al. 2010). Approximately 800 of these membrane proteins belong to the G-protein coupled receptor (GPCR) family of proteins. GPCRs have seven α -

*Corresponding author (jan.rainey@dal.ca).

helical membrane spanning segments connected by alternating intracellular and extracellular loops with an extracellular N-terminal tail and intracellular C-terminal tail (e.g. Figure 1). GPCRs serve as gatekeepers, modulating numerous intracellular signaling cascades in response to a wide variety of extracellular stimuli including peptides, proteins, amino acids, lipids, nucleotides, photons and neurotransmitters. Corresponding to their physiological import, GPCRs are the targets of 50–60% of currently marketed drugs (Lundstrom 2009). Elucidating the structure of GPCRs is key both in terms of understanding this important class of proteins and because of the indispensable role this may play in the field of drug discovery (Allen and Roth 2011).

In last 5 years, there has been an exponential growth in the number of solved human GPCR structures (Katritch et al. 2013). To date, 24 X-ray crystal structures have been solved (<http://gpcr.scripps.edu/index.html>) and one solid-state NMR derived backbone structure of CXCR1 has been reported (Park et al. 2012b). NMR spectroscopy is a very well suited technique to study inherently dynamic proteins such as GPCRs (Kim et al. 2013; Nygaard et al. 2013). It is potentially able to address a wide range of questions related to structure, dynamics, and intermediate states involved in receptor activation that are thought to play an important role in receptor function (Bouvier 2013; Kay 2011).

Multi-milligram quantities of a protein are typically required for structural characterization using experimental techniques such as X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. The amount of a GPCR that is produced in native tissues is very limited; therefore, there methodologies to overexpress these proteins are typically required (Zhao and Wu 2012). Furthermore, studies of proteins of ~100 amino acids or larger by NMR spectroscopy typically requires heteronuclear experiments relying upon enrichment in ^{15}N and ^{13}C nuclei at levels well above natural abundance (Cavanagh et al. 2007). Among the five conventional expression systems (*Escherichia coli*, yeast, insect cells, mammalian cells and cell-free expression systems), expression in *E. coli* is the most economically feasible method to obtain an isotope-enriched protein suitable for structural characterization by NMR spectroscopy. However, obtaining sufficient quantities of protein is typically a bottleneck in structural characterization of GPCRs, or any other class of membrane proteins, using NMR spectroscopy (Gautier et al. 2010). As a result of the difficulties in expression of isotope-labeled full-length GPCRs, fragments of receptors ranging in size from 20–120 amino acids have been studied (Langelaan et al. 2013; Tiburu et al. 2011; Tikhonova and Costanzi 2009; Tyukhtenko et al. 2009; Zhao et al. 2006; Zou et al. 2008). Ste2p, a yeast GPCR, is the only example where larger fragments (4 and 5 transmembrane (TM) segments) have been reported for structural characterization by NMR spectroscopy (Cohen et al. 2014; Potetinova et al. 2012).

The rationale behind studying a subset of TM segments in membrane-mimetic environments, rather than the full-length protein, stems from the two-stage model for membrane protein folding (White et al. 2001). This model justifies the study of fragments comprised of single or multiple TM segments, the so-called “divide and conquer” approach, as it states that an isolated TM helix represents a distinct folding subunit that can be separated from its parent protein without losing its inherent structural characteristics (Bordag and Keller 2010). Beyond general applicability to membrane proteins when

judiciously applied (Bordag and Keller 2010), this approach has been quite extensively applied for studies of GPCRs (Cohen et al. 2014).

The apelin receptor (AR or APJ) is a class A, rhodopsin-like GPCR (O'Dowd et al. 1993) activated by the peptidic ligand apelin (Tatemoto et al. 1998). AR and apelin are widely expressed in the human body and have been implicated in modulation of various physiological and homeostatic mechanisms (O'Carroll et al. 2013). The apelin/AR system has been shown to have roles in regulation of fluid homeostasis (De Mota et al. 2004), angiogenesis during tumour formation (Sorli et al. 2007), adipoinular axis function (Boucher et al. 2005), regulation of the cardiovascular (Kagiyama et al. 2005) and central nervous systems (Kralisch et al. 2007), and as a co-receptor in human immunodeficiency virus type-I infection (Cayabyab et al. 2000). To date, the only high-resolution structural information available for AR is for the AR55 fragment (the 55 N-terminal amino acids, including the first TM segment (Langelaan et al. 2013)), shedding light on the positioning of N-terminal amino acids previously reported to be involved in apelin binding and activation.

Here, we report the expression, purification and biophysical characterization of the first three TM segments of AR (AR_TM1-3, amino acids 1-137; Figure 1). Expression was achieved in the C41(DE3) strain of *Escherichia coli* (*E. coli*) using nucleotide tags rich in adenine and thymine base pairs (referred to hereafter as AT-rich gene tags; Table 1) previously reported to enhance expression in a cell-free expression system (Haberstock et al. 2012). Notably, inclusion of these tags extends the length of the protein by only 4–6 amino acids (Table 1), a rather minimal increase in length. Through use of these expression tags, we achieved final purified protein yields of ~8 mg/L following expression in minimal growth medium suitable for stable NMR-active isotope enrichment. Biophysical analysis demonstrates that the resulting protein has the expected α -helical content and is stable in monomeric form in membrane-mimetic environments for at least ~3–4 weeks, making it suitable for future in-depth characterization using multidimensional NMR experiments.

Materials and methods

Materials

The bacterial codon-optimized gene for full-length AR (protein sequence in Figure 1) and DNA primers for cloning were obtained from BioBasic Inc. (Markham, ON). The pET-23b (+) vector was obtained from Novagen (Darmstadt, Germany). The C41(DE3) strain of *E. coli* (Miroux and Walker 1996) was purchased from Lucigen (Middleton, WI). Reagents for polymerase chain reaction (PCR) amplification of DNA were purchased from Invitrogen (Burlington, ON). Acetonitrile (high performance liquid chromatography (HPLC) grade), 2-propanol, trifluoroacetic acid (TFA), isopropyl β -D-1-thiogalactopyranoside (IPTG), dithiothreitol (DTT), ethylenediaminetetraacetic acid disodium salt dehydrate (EDTA- Na_2), ampicillin, sodium dodecyl sulfate (SDS), sodium acetate, sodium azide, 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) and reagents for making lysogeny broth (LB) medium were purchased from Fisher Scientific (Ottawa, ON). 15% Next gel solution for SDS-PAGE was purchased from Amresco (Solon, OH). Uniformly ^{15}N -labeled ammonium chloride, ^{13}C labeled D-glucose, HFIP- d_2 , dimethyl sulfoxide- d_6 (DMSO- d_6), NaOD and SDS- d_{25} were purchased from Cambridge Isotope Laboratories (Tewksbury, MA). DTT- d_6 , DCl, D_2O

(99.8 atom % D) and D₂O containing 1% (w/w) 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) were obtained from C/D/N Isotopes (Pointe-Claire, QC). Dodecylphosphocholine (DPC or Fos-Choline-12) was obtained from Anatrache (Maumee, OH). Columns for reverse phase (RP) HPLC were purchased from Agilent Technologies Canada (Mississauga, ON). Nitrocellulose membrane (0.2 µm) and Clarity Western ECL Substrate for western blotting were purchased from Bio-Rad (Mississauga, ON). HisProbe-HRP was purchased from Thermo Fisher Scientific Inc. (Waltham, MA). Quartz cuvettes for circular dichroism (CD) spectropolarimetry were purchased from Hellma (Vaughn, ON). 5 mm high quality NMR tube were purchased from Bruker Biospin (Fällanden, Switzerland). All other materials were purchased from Sigma-Aldrich (Oakville, ON). Unless otherwise specified, all reagents were used without further purification.

Plasmid construction

The plasmid for expression of AR_TM1-3 was obtained by sub-cloning the appropriate nucleotide into the pET-23b (+) vector using the *NdeI* and *XhoI* restriction enzyme sites. Four open reading frames with different N-terminal tags (Table 1) were constructed and tested for expression of the AR_TM1-3 protein. Forward and reverse primers used to amplify AR_TM1-3 gene from a pUC 57 plasmid containing the sequence for the full-length AR are listed in Table 2. Forward primers for each construct had different overhangs at the 5' end for incorporation of different N-terminal tags. Each plasmid construct thus obtained had an N-terminal expression enhancement tag and a hexahistidine-tag (6xHis-tag) at the C-terminus (Table 2).

Expression

Expression of AR_TM1-3 with and without N-terminal tag was tested in *E. coli* BL21(DE3) and C41(DE3) strains. For test expression of each construct, a single colony was grown in 4 mL LB medium with ampicillin (100 µg/mL) to an optical density at 600 nm (OD₆₀₀) of 0.6. 3 mL of this culture was then induced with IPTG to a final concentration of 0.5 mM and grown at 37°C for 3 hours after induction. Based on OD₆₀₀, equal numbers of cells were pelleted from each sample and re-solubilized in 1× loading dye. Samples were heated at 90°C for 20 min before analysis with SDS-PAGE (15% Next gel; 180 V for 60 min). The gels were stained with Coomassie Brilliant Blue R-250, destained and imaged using a VersaDoc 4000 (Bio-Rad).

Expression was also analyzed using western blot, where the samples were first resolved by SDS-PAGE as described above and then transferred on to a nitrocellulose membrane at 115 V for 60 min. Membranes were incubated in blocking buffer (TBST (Tris-buffered saline with 0.05% w/v tween 20) and 2% w/v BSA) for 60 min at ~23°C. Membranes were then washed twice with TBST for 10 min each before incubation with HisProbe-HRP (1:5000 dilution in blocking buffer) for 60 min at ~23°C. Membranes were washed again four times with TBST for 10 min each and then incubated for 5 min at ~23°C with Clarity Western ECL Substrate before imaging using a VersaDoc 4000.

For optimization prior to large-scale protein expression, the effects of varying IPTG concentration (0.25 mM, 0.5 mM, 0.75 mM and 1 mM) and temperature (at 22°C for 14

hours, at 28°C for 14 hours and at 37°C for 2, 3 and 4 hours) were tested. No significant difference was observed with variation of IPTG concentration, whereas expression was slightly enhanced at 28°C. For large-scale expression, the C41(DE3) strain of *E. coli* was transformed with pAR_TM1-3_AT and was grown for 14 hours at 37°C in a 250 mL baffled flask containing 15 mL of LB medium with ampicillin (100 µg/mL). This was used as a starter culture to inoculate 1L of LB medium with ampicillin (100 µg/mL). Expression was induced by addition of IPTG (to a final concentration of 0.5 mM) to the culture at OD₆₀₀ of 0.6. After induction, the cells were grown for 14 hours at 28°C and then harvested by centrifugation (6,000 × g at 4°C for 20 min).

For expression of uniformly ¹⁵N- and ¹³C-enriched AR_TM1-3, C41(DE3) *E. coli* cells were grown in LB medium with ampicillin (100 µg/mL) to OD₆₀₀ of 0.6. Pelleted cells were re-suspended in a ½-equivalent volume of minimal medium (100 mM NaH₂PO₄, 40 mM K₂HPO₄, 4 mM MgSO₄, 1.8 µM FeSO₄, 2 g/L ¹³C₆ D-glucose, 1 g/L ¹⁵NH₄Cl and 100 mg/L ampicillin, titrated with NaOH to pH 7.3 (Marley et al. 2001)) and grown for 14 hours at 28°C following induction.

Inclusion body preparation

Inclusion bodies were purified as described by (Garboczi et al. 1996), with minor changes as detailed by Björkman and co-workers (<http://www.its.caltech.edu/~bjorker/protocols.html>). Briefly, cell pellets were re-suspended in 12 mL of solution buffer (50 mM Tris-Cl, 25% sucrose, 1 mM EDTA-Na₂, 0.1% NaN₃, 10 mM DTT) and then mixed with 11 mL of lysis buffer (50 mM Tris-Cl, 1% Triton X-100, 1% sodium deoxycholate, 100 mM NaCl, 0.1% NaN₃, 10 mM DTT). The cells were lysed using a French pressure cell press (American Instrument Company; Silver Springs, MD). The lysate was centrifuged at 11,000 × g for 20 min at 4°C. The supernatant was discarded and the pellet was re-suspended in 15 mL of wash buffer with Triton (50 mM Tris-Cl, 0.5% Triton X-100, 100 mM NaCl, 0.1% NaN₃, 10 mM DTT) using sonication. Centrifugation was again carried out at 11,000 × g for 20 min at 4°C. The pellet was again re-suspended in 15 mL of wash buffer without Triton (50 mM Tris-Cl, 1 mM EDTA-Na₂, 100 mM NaCl, 0.1% NaN₃, 10 mM DTT) using sonication. The inclusion body pellet containing partially pure AR_TM1-3 was obtained by centrifuging again at 11,000 × g for 20 min at 4°C.

RP-HPLC and mass spectrometry

The inclusion body pellet was re-suspended either in acetonitrile:isopropanol:water (2:1:2, by volume) or in TFA:water (4:1, by volume) at high protein concentration prior to RP-HPLC. Linear binary gradients (solvent A – H₂O and isopropanol; solvent B – acetonitrile and isopropanol; A and B both contain 0.1% TFA and 10% isopropanol for analytical separation, 20% isopropanol for preparative separation) were employed for both analytical and preparative RP-HPLC with the column temperature maintained at 50°C on a Prostar HPLC equipped with a diode array ultraviolet (UV)-visible detector (Varian Canada Inc.; Mississauga, ON). For analytical separation (ZORBAX Rapid resolution 300SB-C3 column, 4.6 mm × 150 mm, 3.5 µm), 100 µL of solubilized inclusion body preparation was subjected to a linear gradient of 30–90% solvent B over 40 min at a flow rate of 1 mL/min (chromatograms at 210 (analytical), 220 (preparative) and 280 nm typically monitored).

Each peak was collected and immediately lyophilized. Preparative RP-HPLC (ZORBAX PrepHT 300SB-C3 column, 21.2mm × 50mm, 5µm) employed 1 mL of solubilized inclusion body preparation with a flow rate of 10 mL/min and a linear gradient of 35–50% solvent B over the first 8 min and 50–75% solvent B over the next 5 min. Due to the highly hydrophobic nature of the protein, the column had to be washed 3–4 times to recover all of the injected protein. Each wash was performed by injecting TFA:water (4:1; volume equal to the volume of the sample injected) in the column and subjecting it to the same gradient as the sample. The purified protein fractions were pooled, lyophilized and stored at –20°C. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry of lyophilized protein samples was performed either at the SPARC Biocentre (Hospital For Sick Children, Toronto, ON) or at the Genomics and Proteomics Facility (GaP, Memorial University of Newfoundland, St. John's, NL).

CD spectropolarimetry

Far-UV CD spectra for all samples of AR_TM1-3 were recorded at 25°C using a J-810 spectropolarimeter (Jasco, Easton, MD) at 20 nm/min with a data pitch of 0.1 nm from 260 nm to 190 nm. Cuvettes of path length (*l*) 0.1 cm and 0.01 cm were used for HFIP and detergent micelle samples, respectively. The HFIP sample contained 8.20±0.05 µM protein in HFIP:H₂O (1:1 by volume); micelle samples contained 82.0±0.5 µM protein, with 20 mM SDS in 20 mM sodium phosphate buffer at pH 6.00±0.05 and pH 7.00±0.05 or 10 mM DPC in 50 mM sodium acetate buffer at pH 4.00±0.05. All samples were prepared from the same protein stock solution, with concentration (*c*) determined using the Beer-Lambert law ($c = A \cdot e^{-1} \cdot l^{-1}$) by measuring absorbance (*A*) at 280 nm using a molar absorptivity (*ε*; calculated following (Pace et al. 1995)) of 42860 M⁻¹cm⁻¹. All data were collected in triplicate, averaged, blank subtracted and converted to mean residue ellipticity ([θ]). Data deconvolution was performed using CDSSTR (Johnson 1999) with the SMP180 dataset (Abdul-Gader et al. 2011) as implemented in the DichroWeb server (Whitmore and Wallace 2008).

Solution-state NMR spectroscopy

1D ¹H (Hwang and Shaka 1995) and 2D ¹H-¹⁵N heteronuclear single quantum correlation (HSQC) (Kay et al. 1992; Palmer Iii et al. 1991; Schleucher et al. 1994) spectra of AR_TM1-3 were acquired in two organic solvent and one detergent micelle membrane-mimetic conditions. Organic solvent samples were prepared in DMSO-d₆ (0.6 mM AR_TM1-3) and HFIP-d₂:H₂O:D₂O (5:4:1 by volume; 0.7 mM AR_TM1-3; 1 mM DSS; 5 mM DTT-d₆). An SDS micelle sample was prepared (0.6 mM AR_TM1-3; 150 mM SDS-d₂₅; 1 mM DSS; 10 mM DTT-d₆; 20 mM sodium phosphate; and, 1 mM NaN₃ in H₂O:D₂O 95:5) and the pH of the sample was adjusted to 6.00±0.05 using DCl and NaOD without accounting for deuterium isotope effects. NMR spectra of the SDS sample were collected at 27, 37 and 47°C using a Bruker Avance III 700 MHz spectrometer (Milton, ON) equipped with a cryogenically cooled 5 mm triple resonance inverse (TCI) probe and a z-axis gradient (Biomolecular Magnetic Resonance Facility (BMRF) housed at the National Research Council, Halifax, NS). NMR spectra of the DMSO sample were collected at 22 and 37°C using a Bruker Avance 500 MHz spectrometer equipped with a room temperature 5 mm triple resonance inverse (TXI) probe and a z-axis gradient (NMR³ Facility at Dalhousie

University). NMR spectra of the 50% HFIP sample were collected at 22, 37 and 47°C using the 700 MHz system and at 37 and 47°C using the 500 MHz system. Both 1D ^1H (16 scans, sweep width 12 or 16 ppm, recycle delay 1 or 2 s) and 2D ^1H - ^{15}N HSQC (16 scans, sweep width 12-16/30-32 ppm in F2/F1, 128–320 increments in F1, recycle delay 1.5 s) spectra were processed using Bruker Topspin 3.1. Spectra were plotted using CcpNmr Analysis 2.2.2 (Vranken et al. 2005) for 2D experiments and Bruker Topspin 3.1 for 1D experiments.

Results

Expression optimization using AT-rich tags

Four small AT-rich 5' gene tags (Table 1) were tested for the ability to enhance protein expression of recombinant AR_TM1-3 containing the first three TM segments of AR in the BL21(DE3) and C41(DE3) strains of *E. coli*. The BL21(DE3) strain was not suitable for expression of AR_TM1-3, whether untagged or with any of the four tags, as a large amount of cell death was observed after induction (Figure S1 and Table S1, Supplementary Material). In C41(DE3), no target protein expression was observed for the AR_TM1-3 construct without AT-rich tags (Figures 2–3); conversely, SDS-PAGE (Figure 2) and western blot analysis (Figure 3) demonstrate that all four AT-rich tagged AR_TM1-3 constructs allowed AR_TM1-3 expression in the C41(DE3) strain. Differences in tagged AR_TM1-3 protein migration are clearly apparent upon western blot visualization, likely arising from the differences in tag net charge and length. Notably, although the G-tag was capable of enhancing protein expression over the untagged protein, the amount of protein produced with this tag was clearly lower than any of the other 3 tags (Figure 3). Western blot quantitation (Figure 3) showed that the construct with the H-tag provided the best enhancement of AR_TM1-3 expression. However, the comparative enrichment of the AR_TM1-3 in comparison to background protein expression levels evident from SDS-PAGE analysis (Figure 2) led us to pursue the AR_TM1-3 construct with the AT-tag for overexpression in both rich and minimal medium (relative levels illustrated in Figure S2A, Supplementary Material) and subsequent purification.

Purification

Overexpressed AR_TM1-3 was found to form inclusion bodies; hence, the first step in purification was isolation and washing of the inclusion bodies. Very little to no loss of AR_TM1-3 was observed while washing the inclusion bodies (Figure S2B, Supplementary Material). To further purify AR_TM1-3, solubilized inclusion bodies were subjected to RP-HPLC (Figure 4). Despite having a 6xHis-tag, RP-HPLC was chosen as the method of purification due to an inability to find conditions providing both protein solubility and sufficient affinity of the 6xHis-tag to the Ni-NTA column (Figure S2B, Supplementary Material). The identity of the AR_TM1-3 RP-HPLC eluent peak and the incorporation of ^{15}N and ^{13}C stable isotopes were confirmed by MALDI-TOF (Figure 5; expected mass of unlabeled AR_TM1-3 17684.69 Da, observed mass 17684.52 Da; expected mass of 100% $^{15}\text{N}/^{13}\text{C}$ -enriched AR_TM1-3 18690.55 Da, observed mass 18627.94 Da indicative of ~92% isotope incorporation efficiency.) It should be noted that the pure protein shows evidence of conformational heterogeneity evidenced by the appearance of multiple eluent peaks in RP-HPLC (Figure S3, Supplementary Material, analogous to anomalous SDS-PAGE migration

behaviour of membrane proteins, e.g. (Rath et al. 2013; Rath et al. 2009)). The final yield of lyophilized AR_TM1-3 powder following purification was ~ 8 mg/L of minimal medium, with a purity of >99% as confirmed by analytical HPLC (Figure S3, Supplementary Material).

CD spectropolarimetry

Far-UV CD spectropolarimetry was used to evaluate the secondary structure of purified AR_TM1-3. This was carried out in 50% HFIP, in SDS micelles at pH 6.00±0.05 and pH 7.00±0.05 and in DPC micelles at pH 4.00±0.05. (The DPC conditions differ from SDS as AR_TM1-3 was not soluble in DPC micelles at pH > 4.) Far-UV CD spectra for all of the samples except DPC were characteristic (Cantor and Schimmel 1980) of an α -helix-rich protein, with a positive band at ~192 nm ($\pi \rightarrow \pi^*$ transition) and negative bands at ~222 nm ($n \rightarrow \pi^*$ transition) and 208 nm ($\pi \rightarrow \pi^*$ transition) (Figure 6A). The signal intensity was higher in 50% HFIP; therefore, the data are also shown in a normalized manner (with the minimum $[\theta]$ value for each spectrum used as the normalization factor; Figure 6B) to allow evaluation of relative secondary structure content between samples. This normalization implies that AR_TM1-3 had slightly higher helical content in 50% HFIP, as is evident from higher signal intensity for the $\pi \rightarrow \pi^*$ transitions at 192 nm and 208 nm (Figure 6B).

Estimation of secondary structure composition through CD spectral deconvolution using the CDSSTR algorithm (Johnson 1999) and employing the SMP180 (Abdul-Gader et al. 2011) mixed soluble/membrane protein reference dataset provided reliable fits based on low normalized root mean square deviation (NRMSD) values (Table 3). The α -helical proportion predicted for AR_TM1-3 by deconvolution in SDS micelles (~42–43%) is highly comparable to the ~40% expected on the basis of the topology predicted based on canonical GPCR structuring (Figure 1). It is also evident that the secondary structure composition of the protein in SDS micelles is essentially the same at both pH 6.00±0.05 and 7.00±0.05 (Figure 6 and Table 3). Despite the suitability of DPC for solubilization of AR55 in our previous structural study (Langelaan et al. 2013), the CD spectrum for the DPC sample of AR_TM1-3 resembles that of a protein rich in β -strand secondary structure, as upheld by deconvolution which estimates a proportion of only ~3% helical structure. This is indicative of improper folding and/or aggregation; hence, DPC micelle conditions were not pursued further.

Solution-state NMR spectroscopy

Using uniformly ^{15}N and ^{13}C -enriched AR_TM1-3, we performed 1D ^1H and 2D ^1H - ^{15}N HSQC NMR experiments in different solvent and micellar conditions and at a variety of temperatures in order to gauge feasibility of a future full-scale high-resolution NMR structural study. Consistent with the far-UV CD results, both 1D ^1H and 2D ^1H - ^{15}N HSQC spectra (Figures 7 and 8) indicate that AR_TM1-3 is appropriately structured in 50% HFIP as well as in SDS, as indicated by the dispersion of signals and the number of peaks in the HSQC spectra. The SDS micelle NMR sample runs as a single oligomeric state upon SDS-PAGE as visualized by western blotting for its 6xHis-tag (Figure 3A). This species migrates consistently with the AR_TM1-3 protein in cell pellet and mass spectrometry has

consistently shown only the monomeric species (Figure 5). Based on previous studies of SDS micelle-solubilized TM domain oligomerization (Tulumello and Deber 2009) and the observation of oligomers of a larger yeast GPCR fragment by SDS-PAGE (Cohen et al. 2014; Potetinova et al. 2012), it is likely that any oligomeric state in SDS micelles would be retained during SDS-PAGE. Finally, the observed line widths and suitability of standard HSQC experiments rather than relaxation optimized experiments such as TROSY (Cavanagh et al. 2007) are consistent with AR_TM1-3 being monomeric in both 50% HFIP and SDS micelles.

Significant peak overlap is apparent in all conditions tested, as is typically observed in α -helical membrane proteins (Kim et al. 2009). The global conformation and/or local dielectric constant experienced by a given region of the protein appears to vary as a function of the solvent or micelle system employed, as indicated by some peak motion in the HSQC spectra (Figure 7D). In particular, AR_TM1-3 in DMSO showed much poorer peak dispersion than in either 50% HFIP or SDS micelles. On the basis of these results, DMSO was not pursued further. The overall structuring of AR_TM1-3 under a particular solvent or micelle condition does not appear to have substantial temperature-dependent change, on the basis of monitoring HSQC peak patterns as the temperature changes (Figure 8). Furthermore, unchanged spectral appearance even after weeks also demonstrates sample stability over time, which is essential for pursuing triple-resonance NMR structural studies for full chemical shift assignment, structure determination and dynamics characterization.

Full resonance assignment has not yet been performed. However, distinct cross-peaks arising from the side chain H ϵ -N ϵ spin-pairs of each of the four tryptophan residues in AR_TM1-3 are clearly visible in all of the tested conditions (zooms in Figure 7). In the 50% HFIP and SDS micelle samples, four major Trp H ϵ -N ϵ cross-peaks of ~equal intensity are clearly observable. Some additional peaks of much lower intensity are observed in the HFIP (1 peak) and SDS (3 peaks) samples, indicating the likely presence of a minor conformation in a subset of tryptophan residues. As a whole, the total number of cross-peaks observed in the HSQC spectra in both the HFIP and SDS samples at 47°C (Figure 8) is appropriate on the basis of the AR_TM1-3 sequence (Figure 1) and these sample conditions appear quite amenable for future high-resolution structural studies of AR_TM1-3.

Discussion

Herein, we have shown that the use of AT-rich gene tags makes possible the expression of a three TM segment fragment of a human GPCR. In *E. coli*, the nucleotide sequence around the ribosome binding site has been shown to play a critical role in the regulation of the amount of mRNA and protein (Kudla et al. 2009). This may be related to the fact that an increase in translation efficiency has been shown to be directly related to better initiation of translation conferred by AT-rich codons (Qing et al. 2003). Haberstock *et al.* (2012) showed that having AT-rich codons next to the translation initiation site had a significant positive effect on translation efficiency in an *E. coli* based cell-free expression system. Our findings demonstrate that the AT-rich tags work not only in an *E. coli* based cell-free expression system but also in the case of protein overexpression in *E. coli* cells in culture media. Without the use of AT-rich tags, the level of AR_TM1-3 expressed was undetectable even by

western blotting for the C-terminal His-tag (Figure 3). It should also be noted that use of the C41(DE3) *E. coli* strain, which has increased tolerance to toxic protein expression (Miroux and Walker 1996), was key for enhanced expression – this was far less effective in BL21(DE3) *E. coli*.

When compared to other expression tags, the small AT-rich tags also have unique advantages, especially for structural studies of proteins. Fusion protein tags have proved to be instrumental in heterologous expression and purification of recombinant proteins (Sorensen and Mortensen 2005; Walls and Loughran 2011). In many instances, tags such as Glutathione S-transferase (Park et al. 2012a) and maltose-binding protein (Buck et al. 2003) have been used to enhance the expression and facilitate purification. These, however, are sizable proteins that must be removed before subsequent study. This requires post-translational cleavage, often achieved through trial-and-error with a wide variety of proteolytic and/or chemical strategies. The highly hydrophobic nature of membrane proteins means that this approach is not always appropriate. Even if expression is enhanced, the removal of the tag protein from the membrane protein can be extremely problematic because the fusion protein is usually insoluble under the conditions needed for protease activity. The need for post-translational tag removal can be avoided by the use of very small tags that can still enhance expression, such as the AT-rich tags used here (Haberstock et al. 2012). In particular, the very small size of these tags makes them highly appealing for structural studies using solution-state NMR spectroscopy (see, e.g., (Zhou and Wagner 2010)).

In summary, we have successfully expressed and purified multi-milligram quantities of a protein comprising the first three TM segments of AR. This fragment is larger (by at least one TM segment) than the vast majority of other GPCR fragments that have been produced and studied to date (Cohen et al. 2014). Biophysical studies in an organic solvent/water mixture (50% HFIP in water) and in SDS micelles indicate that this GPCR fragment has appropriate secondary structuring and is suitable for long-term characterization by solution-state NMR spectroscopy. There is compelling evidence in the literature to suggest that the N-terminal tail and extracellular loops of class A GPCRs play a crucial role in ligand binding and receptor function (Wheatley et al. 2012). AR_TM1-3 will therefore also provide a useful probe for the structural features and dynamics of the N-terminal tail and first extracellular loop of AR upon ligand binding. NMR studies will also be carried out on AR_TM1-3 to evaluate the effect of other membrane-mimetic environments upon the structure of this protein.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a Canadian Institutes of Health Research (CIHR) Operating Grant (MOP-111138 to J.K.R.); a Nova Scotia Health Research Foundation (NSHRF) Scotia Support Grant (to J.K.R.); and, a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant (to X.Q.L.). Key equipment were provided by NSERC Research Tools & Instruments and Canadian Foundation for Innovation Leaders Opportunity Fund grants (to J.K.R.) and by a Dalhousie Medical Research Foundation equipment grant (to J.K.R. and X.Q.L.). J.K.R. is supported by a CIHR New Investigator Award. The He-cooled probe for the 700 MHz spectrometer at the NRC

BMRF was provided by Dalhousie University through an Atlantic Canada Opportunities Agency Grant. We thank Bruce Stewart for technical support; Dr. David Waisman for access to his CD spectropolarimeter; Dr. Mike Lumsden for 500 MHz spectrometer support and maintenance at the NMR³ Facility; Ian Burton and Drs. Nadine Merkley and Ray Syvitski for 700 MHz spectrometer support at the NRC BMRF; Hiren Parmar and Dr. David Langelaan for their contributions in designing the preliminary construct of full-length AR; Ellie Scott, Sam Kerzner and Danielle LeBlanc for assistance during various stages of this work; and, Nigel Chapman and Kyungsoo Shin for their constructive comments on the manuscript.

References

- Abdul-Gader A, Miles AJ, Wallace BA. A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics*. 2011; 27(12):1630–1636. DOI: 10.1093/bioinformatics/btr234 [PubMed: 21505036]
- Allen JA, Roth BL. Strategies to discover unexpected targets for drugs active at G protein-coupled receptors. *Annu Rev Pharmacol Toxicol*. 2011; 51:117–144. DOI: 10.1146/annurev-pharmtox-010510-100553 [PubMed: 20868273]
- Bordag N, Keller S. Alpha-helical transmembrane peptides: a “divide and conquer” approach to membrane proteins. *Chem Phys Lipids*. 2010; 163(1):1–26. DOI: 10.1016/j.chemphyslip.2009.07.009 [PubMed: 19682979]
- Boucher J, Masri B, Daviaud D, Gesta S, Guigne C, Mazzucotelli A, Castan-Laurell I, Tack I, Knibiehler B, Carpenne C, Audigier Y, Saulnier-Blache JS, Valet P. Apelin, a newly identified adipokine up-regulated by insulin and obesity. *Endocrinology*. 2005; 146(4):1764–1771. DOI: 10.1210/en.2004-1427 [PubMed: 15677759]
- Bouvier M. Unraveling the structural basis of GPCR activation and inactivation. *Nat Struct Mol Biol*. 2013; 20(5):539–541. DOI: 10.1038/nsmb.2584 [PubMed: 23649361]
- Buck B, Zamoan J, Kirby TL, DeSilva TM, Karim C, Thomas D, Veglia G. Overexpression, purification, and characterization of recombinant Ca-ATPase regulators for high-resolution solution and solid-state NMR studies. *Protein Expr Purif*. 2003; 30(2):253–261. [PubMed: 12880775]
- Cantor, CR., Schimmel, PR. *Biophysical Chemistry, Part 2: Techniques for the Study of Biological Structure and Function*. W.H. Freeman and Company; San Francisco: 1980.
- Cavanagh, J., Fairbrother, WJ., Palmer, AG., Skelton, NJ. *Protein NMR spectroscopy principles and practice*. 2. Academic Press; Boston: 2007.
- Cayabyab M, Hinuma S, Farzan M, Choe H, Fukusumi S, Kitada C, Nishizawa N, Hosoya M, Nishimura O, Messele T, Pollakis G, Goudsmit J, Fujino M, Sodroski J. Apelin, the natural ligand of the orphan seven-transmembrane receptor APJ, inhibits human immunodeficiency virus type 1 entry. *J Virol*. 2000; 74(24):11972–11976. [PubMed: 11090199]
- Cohen LS, Fracchiolla KE, Becker J, Naider F. GPCR Structural Characterization: Using Fragments as Building Blocks to Determine a Complete Structure. *Biopolymers*. 2014; In press. doi: 10.1002/bip.22490
- De Mota N, Reaux-Le Goazigo A, El Messari S, Chartrel N, Roesch D, Dujardin C, Kordon C, Vaudry H, Moos F, Llorens-Cortes C. Apelin, a potent diuretic neuropeptide counteracting vasopressin actions through inhibition of vasopressin neuron activity and vasopressin release. *Proc Natl Acad Sci U S A*. 2004; 101(28):10464–10469. DOI: 10.1073/pnas.0403518101 [PubMed: 15231996]
- Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L. Prediction of the human membrane proteome. *Proteomics*. 2010; 10(6):1141–1149. DOI: 10.1002/pmic.200900258 [PubMed: 20175080]
- Garboczi DN, Utz U, Ghosh P, Seth A, Kim J, VanTienhoven EA, Biddison WE, Wiley DC. Assembly, specific binding, and crystallization of a human TCR-alpha-beta with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2. *J Immunol*. 1996; 157(12):5403–5410. [PubMed: 8955188]
- Gautier A, Mott HR, Bostock MJ, Kirkpatrick JP, Nietlispach D. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat Struct Mol Biol*. 2010; 17(6):768–774. DOI: 10.1038/nsmb.1807 [PubMed: 20512150]

- Haberstock S, Roos C, Hoevels Y, Dotsch V, Schnapp G, Pautsch A, Bernhard F. A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems. *Protein Expr Purif.* 2012; 82(2):308–316. DOI: 10.1016/j.pep.2012.01.018 [PubMed: 22342679]
- Hwang TL, Shaka AJ. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J Magn Reson A.* 1995; 112(2):275–279. <http://dx.doi.org/10.1006/jmra.1995.1047>.
- Johnson WC. Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins.* 1999; 35(3):307–312. [PubMed: 10328265]
- Kagiyama S, Fukuhara M, Matsumura K, Lin Y, Fujii K, Iida M. Central and peripheral cardiovascular actions of apelin in conscious rats. *Regul Pept.* 2005; 125(1–3):55–59. DOI: 10.1016/j.regpep.2004.07.033 [PubMed: 15582714]
- Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. *Annu Rev Pharmacol Toxicol.* 2013; 53:531–556. DOI: 10.1146/annurev-pharmtox-032112-135923 [PubMed: 23140243]
- Kay L, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc.* 1992; 114(26):10663–10665. DOI: 10.1021/ja00052a088
- Kay LE. NMR studies of protein structure and dynamics - a look backwards and forwards. *J Magn Reson.* 2011; 213(2):492–494. DOI: 10.1016/j.jmr.2011.08.010 [PubMed: 21885309]
- Kim HJ, Howell SC, Van Horn WD, Jeon YH, Sanders CR. Recent Advances in the Application of Solution NMR Spectroscopy to Multi-Span Integral Membrane Proteins. *Prog Nucl Magn Reson Spectrosc.* 2009; 55(4):335–360. DOI: 10.1016/j.pnmrs.2009.07.002 [PubMed: 20161395]
- Kim TH, Chung KY, Manglik A, Hansen AL, Dror RO, Mildorf TJ, Shaw DE, Kobilka BK, Prosser RS. The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J Am Chem Soc.* 2013; 135(25):9465–9474. DOI: 10.1021/ja404305k [PubMed: 23721409]
- Kralisch S, Lossner U, Bluher M, Paschke R, Stumvoll M, Fasshauer M. Growth hormone induces apelin mRNA expression and secretion in mouse 3T3-L1 adipocytes. *Regul Pept.* 2007; 139(1–3): 84–89. DOI: 10.1016/j.regpep.2006.10.009 [PubMed: 17126924]
- Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science.* 2009; 324(5924):255–258. DOI: 10.1126/science.1170160 [PubMed: 19359587]
- Langelaan DN, Reddy T, Banks AW, Dellaire G, Dupre DJ, Rainey JK. Structural features of the apelin receptor N-terminal tail and first transmembrane segment implicated in ligand binding and receptor trafficking. *Biochim Biophys Acta.* 2013; 1828(6):1471–1483. DOI: 10.1016/j.bbamem.2013.02.005 [PubMed: 23438363]
- Lundstrom K. An overview on GPCRs and drug discovery: structure-based drug design and structural biology on GPCRs. *Methods Mol Biol.* 2009; 552:51–66. DOI: 10.1007/978-1-60327-317-6_4 [PubMed: 19513641]
- Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR.* 2001; 20(1):71–75. [PubMed: 11430757]
- 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. *J Mol Biol.* 1996; 260(3):289–298. DOI: 10.1006/jmbi.1996.0399 [PubMed: 8757792]
- Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK. The dynamic process of beta(2)-adrenergic receptor activation. *Cell.* 2013; 152(3):532–542. DOI: 10.1016/j.cell.2013.01.008 [PubMed: 23374348]
- O'Carroll AM, Lolait SJ, Harris LE, Pope GR. The apelin receptor APJ: journey from an orphan to a multifaceted regulator of homeostasis. *J Endocrinol.* 2013; 219(1):R13–35. DOI: 10.1530/JOE-13-0227 [PubMed: 23943882]
- O'Dowd BF, Heiber M, Chan A, Heng HH, Tsui LC, Kennedy JL, Shi X, Petronis A, George SR, Nguyen T. A human gene that shows identity with the gene encoding the angiotensin receptor is located on chromosome 11. *Gene.* 1993; 136(1–2):355–360. [PubMed: 8294032]

- Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 1995; 4(11):2411–2423. DOI: 10.1002/pro.5560041120 [PubMed: 8563639]
- Palmer AG III, Cavanagh J, Wright PE, Rance M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J Magn Reson* (1969). 1991; 93(1): 151–170. [http://dx.doi.org/10.1016/0022-2364\(91\)90036-S](http://dx.doi.org/10.1016/0022-2364(91)90036-S).
- Park SH, Casagrande F, Chu M, Maier K, Kiefer H, Opella SJ. Optimization of purification and refolding of the human chemokine receptor CXCR1 improves the stability of proteoliposomes for structure determination. *Biochim Biophys Acta.* 2012a; 1818(3):584–591. DOI: 10.1016/j.bbame.2011.10.008 [PubMed: 22024025]
- Park SH, Das BB, Casagrande F, Tian Y, Nothnagel HJ, Chu M, Kiefer H, Maier K, De Angelis AA, Marassi FM, Opella SJ. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature.* 2012b; 491(7426):779–783. DOI: 10.1038/nature11580 [PubMed: 23086146]
- Potetinova Z, Tantry S, Cohen LS, Caroccia KE, Arshava B, Becker JM, Naider F. Large multiple transmembrane domain fragments of a G protein-coupled receptor: biosynthesis, purification, and biophysical studies. *Biopolymers.* 2012; 98(5):485–500. DOI: 10.1002/bip.22122 [PubMed: 23203693]
- Qing G, Xia B, Inouye M. Enhancement of translation initiation by A/T-rich sequences downstream of the initiation codon in *Escherichia coli*. *J Mol Microbiol Biotechnol.* 2003; 6(3–4):133–144. doi: 77244. [PubMed: 15153766]
- Rath A, Cunningham F, Deber CM. Acrylamide concentration determines the direction and magnitude of helical membrane protein gel shifts. *Proc Natl Acad Sci U S A.* 2013; 110(39):15668–15673. DOI: 10.1073/pnas.1311305110 [PubMed: 24019476]
- Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A.* 2009; 106(6):1760–1765. DOI: 10.1073/pnas.0813167106 [PubMed: 19181854]
- Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletzky O, Glaser SJ, Sorensen OW, Griesinger C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J Biomol NMR.* 1994; 4(2):301–306. [PubMed: 8019138]
- Sorensen HP, Mortensen KK. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol.* 2005; 115(2):113–128. DOI: 10.1016/j.jbiotec.2004.08.004 [PubMed: 15607230]
- Sorli SC, Le Gonidec S, Knibiehler B, Audigier Y. Apelin is a potent activator of tumour neoangiogenesis. *Oncogene.* 2007; 26(55):7692–7699. DOI: 10.1038/sj.onc.1210573 [PubMed: 17563744]
- Tatemoto K, Hosoya M, Habata Y, Fujii R, Kakegawa T, Zou MX, Kawamata Y, Fukusumi S, Hinuma S, Kitada C, Kurokawa T, Onda H, Fujino M. Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor. *Biochem Biophys Res Commun.* 1998; 251(2):471–476. DOI: 10.1006/bbrc.1998.9489 [PubMed: 9792798]
- Tiburu EK, Tyukhtenko S, Zhou H, Janero DR, Struppe J, Makriyannis A. Human cannabinoid 1 GPCR C-terminal domain interacts with bilayer phospholipids to modulate the structure of its membrane environment. *AAPS J.* 2011; 13(1):92–98. DOI: 10.1208/s12248-010-9244-7 [PubMed: 21234731]
- Tikhonova IG, Costanzi S. Unraveling the structure and function of G protein-coupled receptors through NMR spectroscopy. *Curr Pharm Des.* 2009; 15(35):4003–4016. [PubMed: 20028318]
- Tulumello DV, Deber CM. SDS micelles as a membrane-mimetic environment for transmembrane segments. *Biochemistry.* 2009; 48(51):12096–12103. DOI: 10.1021/bi9013819 [PubMed: 19921933]
- Tyukhtenko S, Tiburu EK, Deshmukh L, Vinogradova O, Janero DR, Makriyannis A. NMR solution structure of human cannabinoid receptor-1 helix 7/8 peptide: candidate electrostatic interactions and microdomain formation. *Biochem Biophys Res Commun.* 2009; 390(3):441–446. DOI: 10.1016/j.bbrc.2009.09.053 [PubMed: 19766594]

- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59(4):687–696. DOI: 10.1002/prot.20449 [PubMed: 15815974]
- Walls D, Loughran ST. Tagging recombinant proteins to enhance solubility and aid purification. *Methods Mol Biol*. 2011; 681:151–175. DOI: 10.1007/978-1-60761-913-0_9 [PubMed: 20978965]
- Wheatley M, Wootten D, Conner MT, Simms J, Kendrick R, Logan RT, Poyner DR, Barwell J. Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol*. 2012; 165(6):1688–1703. DOI: 10.1111/j.1476-5381.2011.01629.x [PubMed: 21864311]
- White SH, Ladokhin AS, Jayasinghe S, Hristova K. How membranes shape protein structure. *J Biol Chem*. 2001; 276(35):32395–32398. DOI: 10.1074/jbc.R100008200 [PubMed: 11432876]
- Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*. 2008; 89(5):392–400. DOI: 10.1002/bip.20853 [PubMed: 17896349]
- Zhao J, Zheng H, Xie XQ. NMR characterization of recombinant transmembrane protein CB2 fragment CB2(180–233). *Protein Pept Lett*. 2006; 13(4):335–342. [PubMed: 16712507]
- Zhao Q, Wu BL. Ice breaking in GPCR structural biology. *Acta Pharmacol Sin*. 2012; 33(3):324–334. DOI: 10.1038/aps.2011.187 [PubMed: 22286917]
- Zhou P, Wagner G. Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies. *J Biomol NMR*. 2010; 46(1):23–31. DOI: 10.1007/s10858-009-9371-6 [PubMed: 19731047]
- Zou C, Naider F, Zerbe O. Biosynthesis and NMR-studies of a double transmembrane domain from the Y4 receptor, a human GPCR. *J Biomol NMR*. 2008; 42(4):257–269. DOI: 10.1007/s10858-008-9281-z [PubMed: 18937032]

Figure 2.

Coomassie blue stained gel showing test expression in C41(DE3) strains of *E. coli* for the AR_TM1-3 constructs containing specified tags (sequences in Table 1) and a construct without any N-terminal tag. After expression, equal number of cells (based on OD₆₀₀; detailed in Table S1, Supplementary Material) were pelleted from induced (I) and un-induced (UI) cultures of cells grown for 3 hours after induction with IPTG. Samples were made by mixing the cell pellet from each culture with 100 μ L 1x loading dye and resolved using 15% SDS-PAGE. The arrow indicates the position of the AR_TM1-3 band as confirmed by western blotting (Figure 3).

Figure 3.

(A) Western blot probed with HisProbe-HRP comparing expression of each construct of AR_TM1-3 (SDS-PAGE conditions as in Figure 2). The AR_TM1-3 NMR sample in SDS micelles (pH 6.00) was loaded in the lane labeled “sample”. (B) Quantitation (using ImageJ) of difference in expression observed by western blotting (n=3). Mean (bars) and standard error (error bars) for expression level are plotted, with the intensity of the AT-tag on each blot used as the normalization factor. Untagged AR_TM1-3 is not plotted in (B) since no protein was observable by western blotting for this construct.

Figure 4. (Solid blue line) Analytical RP-HPLC chromatogram (milliabsorbance units (mAU) at 210 nm) for the solubilized inclusion body containing AR_TM1-3. (Dashed red line) Binary gradient of solvent A-B (A: 90% H₂O; B: 90% acetonitrile; both containing 10% isopropanol and 0.1% TFA). The * denotes the peak corresponding to AR_TM1-3. (*Colour online.*)

Figure 5. MALDI-TOF mass spectra of isolated AR_TM1-3 expressed in (A) unlabeled rich medium (expected mass 17684.69 Da) and (B) $^{15}\text{N}/^{13}\text{C}$ isotope-enriched M9 minimal medium (expected mass with 100% isotope-enrichment 18690.55 Da).

Figure 6.

(A) Far-UV CD spectra of AR_TM1-3 in indicated solvent or micellar environment, allowing evaluation of secondary structure composition. (B) Normalized representation (normalized to absolute minimum $[\theta]$), allowing direct comparison of relative helical content. HFIP refers to 50% HFIP in H_2O . (*Colour online.*)

Figure 7.

^1H - ^{15}N HSQC spectra for AR_TM1-3 in: (A) 50% HFIP at 47°C (blue); (B) SDS micelles at 47°C (red); and, (C) DMSO at 22°C (green). (D) Overlay of the three conditions in (A–C) to allow for direct comparison. Zoomed regions illustrating peaks arising from tryptophan side chain H ϵ -Ne are shown in A, B and C. In (C), the zoomed region covers 10.69–10.93 ppm and 131.00–132.75 ppm in ^1H and ^{15}N dimensions, respectively. (*Colour online.*)

Figure 8.

Effect of temperature on 1D ^1H and 2D ^1H - ^{15}N HSQC spectra for AR_TM1-3 in 50% HFIP (A,C; red: 47°C, green 37°C, blue 22°C) and SDS micelles (B,D; red: 47°C, green 37°C, blue 27°C). (Colour online.)

Table 1

Nucleotide and amino acid sequences of the tested expression enhancement tags.

Tag name	Nucleotide sequence	Amino acid sequence	% of A and T
AT	AAATATTATAAATATTAT		100
G	AAAAGTAAAGGAGAAGAA		72
H	AAACCATACGATGGTCCA		55
AT4	AAATATTATAAA		100

Table 2

Primers and constructs.

Plasmid name	Construct	Forward primer				Reverse primer			
pAR_TMI-3_AT	AT-tag	ATA	CTA	AAA	AGC	TTA			
		TAT	TAT	AAA	TAT	ATG	ATG	ATG	ATG
pAR_TMI-3_G	G-tag	ATA	CTA	AAA	AGC	TTA			
		AGT	AAA	GGA	GAA	ATG	ATG	ATG	ATG
pAR_TMI-3_H	H-tag	ATA	CTA	AAA	AGC	TTA			
		CCA	TAC	GAT	GGT	CCA	ATG	ATG	ATG
pAR_TMI-3_AT4	AT4-tag	ATA	CTA	AAA	AGC	TTA			
		TAT	TAT	AAA	TAT	ATG	ATG	ATG	ATG

Table 3

Deconvolution of AR_TM1-3 far-UV CD data using the CDSSTR algorithm and SMP180 reference dataset.

Conditions	Secondary structure ^a
HFIP:H ₂ O (1:1 by volume) <i>Deconvolution NRMSE: 0.010</i>	Helix 1- 0.55, Helix 2- 0.18 Strand 1- 0.01, Strand 2- 0.01 Turn - 0.09, Unordered- 0.14
SDS micelles (pH 7.00±0.05) <i>Deconvolution NRMSE: 0.022</i>	Helix 1- 0.23, Helix 2- 0.20 Strand 1- 0.06, Strand 2- 0.05 Turn - 0.17, Unordered- 0.30
SDS micelles (pH 6.00±0.05) <i>Deconvolution NRMSE: 0.019</i>	Helix 1- 0.24, Helix 2- 0.18 Strand 1- 0.08, Strand 2- 0.06 Turn - 0.14, Unordered- 0.29
DPC micelles (pH 4.00±0.05) <i>Deconvolution NRMSE: 0.012</i>	Helix 1- 0.02, Helix 2- 0.01 Strand 1- 0.43, Strand 2- 0.20 Turn - 0.03, Unordered- 0.29

^a Helix 1 - regular α -helix; Helix 2 - distorted α -helix; Strand 1- regular β -strand; Strand 2- distorted β -strand

^b NRMSE - normalized root mean square deviation