An expression and purification system for the biosynthesis of adenosine receptor peptides for biophysical and structural characterization

Zachary T. Britton\textsuperscript{a}, Elizabeth I. Hanle\textsuperscript{a}, and Anne S. Robinson\textsuperscript{a,b,*}
\textsuperscript{a}Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, DE 19716, United States
\textsuperscript{b}Department of Chemical and Biomolecular Engineering, 300 Lindy Boggs Laboratory, Tulane University, New Orleans, LA 70118, United States

Abstract

Biophysical and structural characterization of G protein-coupled receptors (GPCRs) has been limited due to difficulties in expression, purification, and \textit{in vitro} stability of the full-length receptors. “Divide and conquer” approaches aimed at the NMR characterization of peptides corresponding to specific regions of the receptor have yielded insights into the structure and dynamics of GPCR activation and signaling. Though significant progress has been made in the generation of peptides that are composed of GPCR transmembrane domains, current methods utilize fusion protein strategies that require chemical cleavage and peptide separation via chromatographic means. We have developed an expression and purification system based on fusion to ketosteroid isomerase, thrombin cleavage, and tandem affinity chromatography that enables the solubilization, cleavage, and characterization in a single detergent system relevant for biophysical and structural characterization. We have applied this expression and purification system to the production and characterization of peptides of the adenosine receptor family of GPCRs in \textit{Escherichia coli}. Herein, we demonstrate using a model peptide that includes extracellular loop 3, transmembrane domain 7, and a portion of the carboxy-terminus of the adenosine A\textsubscript{2a} receptor that the peptide is sufficiently pure for biophysical characterization, where it adopts \(\alpha\)-helical structure. Furthermore, we demonstrate the utility of this system by optimizing the construct for thrombin processing and apply the system to peptides with more complex structures.

Keywords

Ketosteroid isomerase; Peptide; G-protein coupled receptor; Human A\textsubscript{2a} receptor; Human A\textsubscript{2b} receptor; \textit{Escherichia coli}

Introduction

G protein-coupled receptors (GPCRs)\textsuperscript{1} are \(\alpha\)-helical seven transmembrane domain proteins that play critical roles in cellular signaling in response to diverse extracellular ligands.
Approximately 30% of current small molecule and biological drugs target GPCRs [1], where the effectiveness and specificity of these pharmaceutical therapies could be vastly improved by structure-based drug design. To date, medium to high resolution crystal structures of eight GPCRs have been determined, including rhodopsin [2], the β₁ and β₂ adrenergic receptors [3–7], the adenosine A₂a receptor [8–10], the CXCR4 chemokine receptor [11], the dopamine D₃ receptor [12], the histamine H₁ receptor [13], and the M₂ muscarinic receptor [14]. Opportunities for further development of pharmaceuticals that target GPCRs are considerable, as the sequencing of the human genome has identified upwards of 1000 GPCRs [15]; however, the expression, purification, and in vitro stability of full-length GPCRs have impeded the biophysical and structural characterization of this class of membrane proteins.

Recently, “divide and conquer” approaches, which involve the generation and analysis of peptides corresponding to GPCR transmembrane (TM) domains, have been utilized to overcome obstacles that prohibit characterization of full-length GPCRs. For example, structural characterization by NMR is more easily achieved and less constrained in the membrane mimetic environment than by crystallography. Furthermore, peptides derived from rhodopsin and characterized by NMR were shown to adopt secondary structures similar to those observed by X-ray crystallography [16–19]. In addition, studies of GPCR-derived peptides offer a unique opportunity: specific regions of the receptor, i.e. regions responsible for ligand binding, receptor activation or cellular signaling, can be characterized independently. Since these regions are often quite mobile and poorly resolved, this approach complements high-resolution crystallography. In fact, this piecemeal approach to GPCR characterization has led to insights into the structure of full-length GPCRs, including rhodopsin [16–19], Ste2p, the α-factor receptor of Saccharomyces cerevisiae [20–31], the neurokinin 1 (NK₁) receptor [32], the adenosine A₂a receptor [33–37], the μ-opioid receptor [38,39], and the cannabinoid CB₂ receptor [40].

Only recently have GPCR peptides that are composed of more than a single TM been characterized by this approach, where the cost of chemical synthesis of isotopically labeled peptides (²H, ¹³C, ¹⁵N) and the inherent hydrophobic character of TM peptides have been the major obstacles. However, characterization of multiple transmembrane domains of the μ-opioid receptor, the cannabinoid CB₂ receptor, and yeast Ste2p have been achieved through expression in Escherichia coli of fusion proteins containing the desired peptide [30,38–40]. In fact, large quantities of the isotopically labeled fusion protein/peptide can be generated relatively inexpensively by expression in media containing ²H, ¹³C, and ¹⁵N as sole sources for hydrogen, carbon, and nitrogen. After purification and liberation of the GPCR peptide from the fusion protein, these multi-TM GPCR peptides have shown highly α-helical secondary structure in trifluoro-ethanol/water and a variety of detergents [30,38–40] and provide motivation for further investigations into the “divide and conquer” approach.

In general, expression of GPCR peptides in E. coli is achieved by constructs with multiple peptides in tandem or by fusion to carrier proteins that protect the peptide from intracellular degradation [41]. A number of fusion protein expression systems are commercially available, including those with maltose binding protein, glutathione S-transferase, small ubiquitin-like modifier, and thioredoxin. These carrier proteins not only promote solubility and high expression levels but in some cases also provide convenient means for purification by affinity chromatography. Other carrier proteins, including ketosteroid isomerase [42], the N-terminal domain of S. nuclease (SFC120) [43], fusion to E. coli TrpE [44], and the anti-apoptotic Bcl-2 family protein, Bcl-XL [45], promote insoluble expression and readily form inclusion bodies. Inclusion body expression often results in the highest expression yields [46,47], provides a simple procedure for recovery through lysis and centrifugation, and even
permits the production of toxic peptides [48,49]. Due to their inherent hydrophobic character, GPCR TM peptides are almost exclusively expressed as fusion proteins with carriers that drive expression to inclusion bodies.

Though required for the generation of peptides in *E. coli*, removing the carrier protein is not always a trivial process. The most common methods for liberating peptides from fusion proteins involve chemical cleavage through CNBr [42] or acid hydrolysis [50] and are limited to peptides that do not contain methionine or the acid labile Asp-Pro sequence, respectively, or require their modification by site-directed mutagenesis. Additionally, the harsh conditions under which chemical cleavage is achieved requires purification of the cleavage products by HPLC, where hydrophobic peptides are known to result in low yield [44,51], poor resolution, and low purity [51]. Alternatively, the peptide may be liberated from the carrier protein through incorporation of a protease-specific cleavage site, including those for enterokinase [52], factor Xa 11, tobacco etch virus (TEV) protease [53], and thrombin [54,55], and combined with purification strategies that utilize detergents or chaotropes.

We have developed a robust system for the high-level expression and purification of GPCR peptide fragments in *E. coli* that exploits methods previously developed for solubilizing and recovering inclusion bodies with the zwitterionic detergent fos-choline 16 [56]. The engineered expression system utilizes an N-terminal ketosteroid isomerase (KSI) domain, redundant Strep-Tactin affinity tags, a thrombin cleavage site, and a nickel affinity tag to enable the expression and rapid purification of receptor peptides for biophysical and structural characterization.

**Results**

A peptide representing extracellular loop 3, transmembrane domain 7, and a portion of the C-terminus of hA₂aR was used as a model peptide for developing peptide expression and purification protocols (Fig. 1). Unlike many other G protein-coupled receptors (GPCRs), extensive biophysical and structural characterization of the human adenosine A₂a receptor have been accomplished through biophysical characterization of hA₂aR TM peptides [33,35–37] and the full-length receptor [57,58] as well as through X-ray crystallography [9,10,59]. Completing the biophysical characterization of the hA₂aR TM 7 peptide and comparing the results with the literature would validate the peptide expression and purification system and motivate its application to other GPCR-derived peptides.

**Designing the expression construct**

There have been a number of fusion proteins used for the expression and purification of GPCR-derived peptides, including those that promote solubility or insolubility of the fusion protein [22–31,38–40]. Due to the high hydrophobic character of GPCR TM peptides, carrier proteins designed to promote soluble expression of the fusion protein oftentimes fail in this regard and instead readily form inclusion bodies. For example, initial attempts in our lab at expressing GPCR TM peptides that utilized thioredoxin, a carrier protein that promotes soluble expression, showed that the fusion protein readily formed inclusion bodies in *E. coli* (data not shown). Therefore, in lieu of choosing a carrier protein that promotes soluble expression, we pursued fusion proteins that utilize ketosteroid isomerase (KSI), which promotes insoluble expression in inclusion bodies in *E. coli*.

The commercially available vector, pET-31b (Merck KGaA, Germany), was used as a starting point for developing the expression system; this construct encodes for an N-terminal KSI leader and contains an in-frame His₆ tag for facile immobilized metal affinity chromatography (IMAC) purification. Directing the fusion protein to inclusion bodies
exploits our previous success with the recovery of the human neurokinin NK1 receptor, a GPCR, from inclusion bodies in *E. coli* [56], where protocols for expression, detergent solubilization, and IMAC purification have already been optimized.

Unlike our previous work, however, fusion proteins carrying GPCR-derived peptides require additional preparation and purification steps: the GPCR peptide must be liberated from the fusion protein and then isolated. As previously described, common methods for liberating the peptide include chemical cleavage through CNBr or acid hydrolysis at susceptible sites. This approach, however, does not utilize the detergent-solubilized state of the fusion protein and requires purification by HPLC, which has confounded the recovery of highly hydrophobic peptides [44,51]. Thus, we sought alternatives to chemical cleavage that would permit the cleavage and recovery of fusion proteins in a single detergent system. Our previous work with solubilizing the NK1 receptor [56] and recent evidence from a systematic screen of 96 detergents identified that the fos-choline series of detergents (FC9–16) was the most effective at solubilizing human chemokine receptors from *E. coli* [60]. Therefore, we used fos-choline 16 as the detergent of choice and the previously developed expression and purification protocols [56] as the basis for developing peptide expression and purification methods.

Site-specific proteases, which cleave at specific sequences that are designed into the expression construct through recombinant DNA methods, are an alternative to chemical cleavage. And, importantly, thrombin was determined to be insensitive to a panel of 94 detergents, including the fos-choline series (FC10–16) [61]. Thus, a thrombin cleavage site (LVPRGS) was incorporated between the KSI carrier protein and the peptide of interest in the expression construct. In order to clarify the peptide of interest from undigested fusion protein, we have incorporated two Strep-tag II sequences into the construct design: one on the N-terminus of KSI and one immediately upstream of the thrombin cleavage site. Any undigested fusion protein carrying the Strep-tag II is bound upon exposure to Strep-Tactin resin, whereas the GPCR peptide passes unimpeded through the column and is collected in the flow-through fraction. This clarification step ensures the purity of the GPCR peptide for biophysical and structural characterization.

Initial expression studies using this construct resulted in IMAC purification yields of ~15–20 mg/L fusion protein per liter of culture, which is common for KSI-driven fusion protein expression [42]. However, liberation of the hA2aR TM 7 peptide from the fusion protein limited overall peptide yields and further optimization of this processing step was investigated to improve yields.

**Thrombin processing efficiency**

Even in cases where the detergent does not influence protease activity, ineffective processing of fusion proteins by the protease can limit high peptide yields. Protease processing yields are commonly in the range of 60–70% [62] but may be improved by incorporation of glycine rich linkers flanking the cleavage site [63]. A fusion protein of hA2aR TM 7 that lacked a glycine-rich linker was produced and purified by IMAC; this peptide was tested for thrombin processing and found to be 65% cleaved after 24 h. at 4 °C (Fig. 2). In order to improve the cleavage yields, we incorporated glycine rich linkers G(GSG)nG upstream and downstream of the thrombin cleavage site. We then determined the effect of different linker lengths on the thrombin processing efficiency of purified fusion proteins that differed only in the downstream glycine rich linker, including G(GSG)nG linkers of n = 8,4,2,1,0 (Fig. 2), as thrombin processing would remove the upstream linker.

Achieving high thrombin processing yields with only very short glycine rich linkers ensured that contributions of the artificial linker to biophysical and structural characterization are
minimized. We, therefore, pursued expression and purification of Strep-KSI-G\(_8\)G-
Strep-Thrombin-GGSGG-hA\(_2\)aR TM 7 that when released from the carrier protein would
result in a peptide corresponding to GSLKGGSGGVD-hA\(_{259-316}\)-LEHHHHHHH.

**Purification of hA\(_2\)aR TM 7**

We purified hA\(_2\)aR TM 7 for biophysical characterization and validated protocols for
expressing and purifying GPCR peptides in *E. coli*. One liter of *E. coli* Rosetta(DE3)/pLysS
culture expressing Strep-KSI-G\(_8\)G-Strep-Thrombin-GGSGG-hA\(_2\)aR TM 7 was grown
and harvested. Inclusion bodies were extracted, solubilized with fos-choline 16, and purified
by immobilized metal affinity chromatography (IMAC) through selective affinity of the
His\(_6\)-tag for Ni–NTA resin. Samples were taken after each step of the IMAC purification
and analyzed by SDS–PAGE and Coomassie staining (Fig. 3). Following elution, Strep-
KSI-G\(_8\)G-Strep-Thrombin-GGSGG-hA\(_2\)aR TM 7 appeared as a single band on
Coomassie-stained SDS–PAGE. Western blot analysis using an anti-His\(_6\) antibody
confirmed the identification of the ~29 kDa band.

4 mg Strep-KSI-G\(_8\)G-Strep-Thrombin-GGSGG-hA\(_2\)aR TM 7 was digested in 10 mL
thrombin cleavage buffer supplemented with 5 mM \(\beta\)-ME and 0.0106% fos-choline 16 at 4
\(^\circ\)C with 25 U/mL thrombin. Thrombin digestion resulted in 43% cleavage of the fusion
protein in this case. Other 1L scale studies showed that fresh thrombin and detergent
concentration were critical for best yields (data not shown). Thus, we believe the
uncommonly low cleavage yield compared to our earlier success was the result of the
thrombin preparation or ineffective reduction of the detergent concentration for this
particular preparation. Despite the low cleavage yield, the uncut fusion protein and hA\(_2\)aR
TM 7 -His\(_6\) peptide that contained the His\(_6\) tag were readily separated from thrombin and
the KSI-containing digestion product by IMAC. His-tagged proteins were eluted in thrombin
cleavage buffer supplemented with 5 mM \(\beta\)-ME and 0.0106% fos-choline 16 and 500 mM
imidazole (Fig. 4A). Fractions enriched in uncut Strep-KSI-G\(_8\)G-Strep-Thrombin-
GGSGG-hA\(_2\)aR TM 7 -His\(_6\) and hA\(_2\)aR TM 7 -His\(_6\) peptide were combined and separated
by differential affinity for Strep-Tactin resin (Fig. 4B). Flow-through fractions were
collected and analyzed by SDS–PAGE and Coomassie staining. Fractions enriched in pure
peptide were combined and concentrated using a centrifugal concentrator. The hA\(_2\)aR TM 7
peptide was >95% pure as judged by SDS–PAGE and permitted biophysical characterization
by circular dichroism (CD) spectroscopy. A summary of the processing and purification
steps leading to the final peptide yield appears in Table 3.

\(\beta\)-mercaptoethanol (\(\beta\)-ME) was included throughout the purification of hA\(_2\)aR TM 7 to
reduce disulfide bonds and promote fusion protein and peptide monomer formation, since
the hA\(_2\)aR TM 7 peptide contains two cysteines known to form a disulfide bond. However,
\(\beta\)-ME has been known to confound secondary structure characterization by CD spectroscopy
due to inherent optical properties, and was therefore removed prior to CD characterization of
the peptide. Eliminating the reducing agent resulted in the formation of reversible peptide
dimers (Fig. 5A). Peptides dimers were reduced by treatment with DTT indicating that
peptide dimerization is the result of intermolecular disulfide bonds.

**Secondary structure**

The secondary structure of hA\(_2\)aR TM 7 was investigated using circular dichroism (CD)
spectroscopy. In order to determine the secondary structure of hA\(_2\)aR TM 7 of both the
monomeric, reduced and dimeric peptide by CD spectroscopy, samples were obtained in
phosphate buffer and fos-choline 16 with or without the reducing agent TCEP (Fig. 5B),
which previously has been shown to fully reduce the disulfide bonds in the full-length
hA\(_2\)aR at the concentration tested [64]. CD spectra show that hA\(_2\)aR TM 7 exhibits a-
helical structure that is signified by two distinct minima at 208 and 222 nm, in either the presence or absence of TCEP (Fig. 5B). Addition of TCEP did not result in a significant change to spectra, indicating that the secondary structure of intermolecular disulfide-bonded hA2aR TM 7 dimers occupy similar secondary structure to their monomeric counterparts. Using this spectrum, the α-helical content of the hA2aR TM 7 peptide was estimated to be ~17% (Eq. (2), Materials and Methods) for both buffer conditions.

Analysis of expression for TM peptides derived from the adenosine family of GPCRs

In order to determine whether the expression system could be easily applied to more complex GPCR peptides, GPCR peptides corresponding to hA2aR TM 7, hA2aR TM 6–7, hA2aR TM 5–7, and hA2aR TM 6 were subcloned into the pET-31b-Strep-KSI-G(GSG)8G-Strep-Thrombin-G(GSG)8G vector. Each was transformed into the Rosetta(DE3)/pLysS cell line and grown in 250 mL cultures. For each culture, samples of cellular lysate, including soluble and insoluble proteins, were analyzed by SDS–PAGE (Fig. 6A). Of the five fusion proteins containing hA2aR transmembrane peptides, only fusion proteins containing peptides of hA2aR TM 7 and hA2aR TM 6 were expressed in Rosetta(DE3)/pLysS at levels sufficient for detection by SDS–PAGE and Coomassie staining (Fig. 6A). Western blot analysis using an anti-His6 antibody confirmed that bands identified by Coomassie staining are indeed the fusion proteins, and also shows that fusion proteins corresponding to hA2aR TM 6–7 and hA2aR TM 5–7 peptides are expressed, albeit at much lower levels (Fig. 6B). In addition to lower per cell expression levels, expression of fusion proteins corresponding to hA2aR TM 6–7, hA2aR TM 5–7 and to a moderate extent hA2aR TM 6 peptides exhibited severe growth inhibition and in some cases modest cell death, as determined by evaluating doubling times based on changes to optical density post-induction.

Expression of particularly flexible or disordered peptides in E. coli may have ill effects despite KSI-driven expression to inclusion bodies. For example, structural flexibility of the peptide may promote soluble expression and inhibit KSI-driven inclusion body formation or alternatively, elicit cellular responses that ultimately lead to degradation of the fusion protein. Mutations to hA2aR created by alanine-scanning mutagenesis and selected for improvements in thermal stability and either agonist or antagonist conformational homogeneity have previously been identified [65]. Rant 5 (an antagonist-stabilized mutant of hA2aR) that contains a single mutation within TM 6 (V239A) was used to generate a TM 6–7 peptide, with the hypothesis that we might observe improved fusion protein expression yields by reducing TM 6 structural flexibility. However, the fusion protein for Rant 5 TM 6–7 was not observed by SDS–PAGE and only modestly improved expression yields over wild-type hA2aR TM 6–7 as determined via Western blot analysis (Fig. 6). Therefore, fusion protein yields of hA2aR TM 6–7 and TM 5–7 may not lead to the yields required for structural characterization.

The limited expression yields observed here for fusion proteins of hA2aR TM 6–7 may be GPCR-specific and not a limitation of the expression system as multi-TM [30,38–40,66] and full-length GPCRs [56,62,66] have been expressed in E. coli with high yields. To determine whether limitations to expression yields of fusion proteins of hA2aR TM 6–7 were specific to hA2aR, fusion protein constructs were generated for TM 7 and TM 6–7 peptides of hA1R and hA3bR. Analysis of expression via SDS–PAGE and Coomassie staining (Fig. 7A) as well as Western blot assay (Fig. 7B) showed that all fusion proteins for TM 7 peptides tested exhibited high expression yields; however, only the fusion protein corresponding to hA3bR TM 6–7 maintained high expression yields. The fusion protein corresponding to hA1R TM 6–7 exhibited severely diminished expression yields and cells expressing this fusion protein grew poorly post-induction, where the OD600 at harvest was ~0.5. Small-scale studies suggested thrombin cleavage of these peptides was efficient in the designed module, so long
as expression was sufficient, indicating expression was the limiting step in the final peptide yields.

**Discussion**

We have developed an expression and purification system for the generation of GPCR-derived peptides of the adenosine family of receptors based on fusion to ketosteroid isomerase. The developed approach utilizes the solubilization, thrombin cleavage, and tandem affinity purification of GPCR peptides in a single detergent system and offers an alternative to current methods for purifying GPCR-derived peptides that are based on chemical cleavage and HPLC.

In all cases, thrombin processing efficiency of fusion proteins with glycine rich linkers was improved over fusion proteins without the linker and resulted in efficiencies up to 95% in 24 h. or less on small-scale preparations (Fig. 2). Furthermore, this thrombin-based strategy for liberating peptides from the KSI carrier appeared to be general, as all adenosine TM peptides tested were efficiently cleaved by this method (data not shown). Because the purification strategy avoids the dependence on chemical cleavage for liberating the GPCR peptide from the fusion protein, GPCR peptides in this system are ensured to represent domains within the wild-type receptor without the need for modifying the primary sequence, which adds an additional advantage over currently used peptide production systems.

We have expressed and purified a model peptide of the human adenosine A<sub>2a</sub> receptor comprised of extracellular loop 3, transmembrane domain 7, and a portion of the C-terminus that contains helix 8 at larger scale to validate the expression and purification system. IMAC purification resulted in a yield of ~17 mg pure fusion protein per liter of culture and enabled further processing and purification steps prior to characterization of the hA<sub>2a</sub>R TM 7 peptide. In this case, yields were sufficient for further processing, but may be improved by sonication of recovered inclusion bodies in the presence of detergent, as batch-to-batch purification yields varied significantly, from 10 to 50 mg per liter of culture.

Characterization of the model peptide by CD spectroscopy showed that the peptide adopted α-helical structure in fos-choline 16 micelles consistent with the full-length receptor. Based on the spectra obtained, the α-helical content of the hA<sub>2a</sub>R TM 7 peptide was estimated to be ~17% (Eq. (2), Materials and Methods) for both buffer conditions investigated. This differs from the expected content of the hA<sub>2a</sub>R TM 7 peptide based on the hA<sub>2a</sub>R crystal structure (49%) [9] and studies of hA<sub>2a</sub>R TM 7 peptides in SDS micelles (>70%) and TFE (~70%) [37]. However, SDS and TFE both promote helix formation and TM 7 in the context of other helices may retain more helical content. Alternatively, these differences may result from an inability of the hydrophobic potential of the fos-choline detergent alone to stabilize helical structure. This observation motivates production and biophysical characterization of peptides with greater complexity, i.e. peptides that contain multiple α-helical TM domains, where helix-helix interactions have been shown to promote helix stability [67].

The expression system was extended to peptides of A<sub>2a</sub>R containing multiple transmembrane domains. Of the five fusion proteins containing hA<sub>2a</sub>R transmembrane peptides, expression of multidomain peptides containing either helix 5 or 6 resulted in much lower expression yield and growth inhibition and sometimes cell death.

It remains unclear as to why the addition of hA<sub>2a</sub>R TM 6 to the well-expressing peptide hA<sub>2a</sub>R TM 7 caused severely decreased expression levels and cell viability, where the final OD<sub>600</sub> was ~0.5; however, the recent crystal structures of hA<sub>2a</sub>R [9,10,59] and biophysical characterization of synthesized peptides corresponding to TM peptides of hA<sub>2a</sub>R [33,35–37] yield some clues. These studies suggest that hA<sub>2a</sub>R TM 6, especially the intracellular
portion of this TM, is inherently flexible. This structural flexibility contributes to GPCR activation by accommodating structural rearrangements within the GPCR helical core, but may contribute to instability or affect folding into inclusion bodies in vivo.

In contrast to the observations for A2aR, the fusion protein corresponding to hA2bR TM 6–7 maintained high expression yields. However, hA1R TM 6–7 also exhibited severely diminished expression yields and cells expressing this fusion protein grew poorly post-induction. Peptides corresponding to hA1R TM 6–7 and hA2aR TM 6–7 share the conserved CPXC motif [68] shown to form a disulfide bond in hA2aR [9,10,59], whereas hA2bR TM 6–7 does not. Mutations that eliminate this conserved motif (APXA) did not improve fusion protein yields for the pET-31b construct of these peptides (data not shown). It remains unclear why expression yields of KSI fusion proteins corresponding to hA1R and hA2aR TM 6–7 are severely limited whereas other peptides and full-length GPCRs expressed in this system are unaffected (data not shown).

These results highlight the importance of exploring multiple fusion protein constructs for optimal peptide expression yields; yet, the general approach developed here for liberating and purifying the peptide of interest are applicable in any system.

**Materials and methods**

**Generating the expression system**

DNA encoding a Strep-tag II affinity tag (WSHPQFEK), which enables purification by affinity for Strep-Tactin affinity resin, was introduced on the N-terminus of ketosteroid isomerase (KSI) by polymerase chain reaction (PCR) of KSI from pET-31b and cloned into pET-23b after NdeI and BamHI restriction digest to form pET-23b-Strep-KSI-His6 (primers listed in Table 1). A second Strep-tag II affinity tag and a thrombin cleavage site (LVPRGS) were incorporated on the C-terminus of Strep-KSI by ligating annealed oligos into KpnI and AflIII digested pET-23b-Strep-KSI-His6 to form pET-23b-Strep-KSI-Strep-Thrombin-His6. DNA encoding the glycine-rich linker G(GSG)8G was then incorporated between the Strep-KSI and Strep-Thrombin cleavage site by cloning annealed oligos into BamHI and KpnI digested pET-23b-Strep-KSI-Strep-Thrombin-His6 to form pET-23b-Strep-KSI-G(GSG)8G-Strep-Thrombin-His6. DNA encoding Strep-KSI-G(GSG)8G-Strep-Thrombin was shuttled into pET-31b by NdeI and XhoI digest. A second G(GSG)8G protein linker was incorporated after the thrombin cleavage site in pET-31b-Strep-KSI-G(GSG)8G-Strep-Thrombin-His6 by ligating annealed oligos into AflIII and SalI restriction sites to form pET-31b-Strep-KSI-G(GSG)8G-Strep-Thrombin-G(GSG)8G-His6. A summary of the cloning strategy appears in Table 2. DNA encoding for peptides of the adenosine A2a receptor, adenosine A1 receptor, and adenosine A2b receptor were then introduced into pET-31b-Strep-KSI-G(GSG)8G-Strep-Thrombin-G(GSG)8G-His6 by SalI and XhoI restriction digest. Tables 4–6 contain details of the cloning strategy and the peptides developed.

A model peptide of A2aR (AA 259–316) composed of extracellular loop 3, transmembrane domain 7, and a portion of the C-terminus that includes the membrane-associated helix 8 (Fig. 1), was devised based on the design of a 73-residue peptide of the S. cerevisiae alpha-factor pheromone receptor (Ste2p) that previously has been characterized by NMR [22]. In previous work on synthetic peptides of hA2aR, TM 7 was shown to have high α-helical content and biophysical characterization of this peptide could serve to validate the expression and purification protocols developed.
Optimizing the protein linker for A<sub>2a</sub>R TM 7

In order to optimize thrombin digestion, constructs of hA<sub>2a</sub>R TM 7 that differ in protein linker length were generated by digesting pET-31b-Strep-KSI-G(GSG)<sub>8</sub>G-Strep-Thrombin-G(GSG)<sub>8</sub>G-A<sub>2a</sub>R TM 7 with AflII and SalI and ligating annealed oligos corresponding to G(GSG)<sub>4</sub>G and G(GSG)<sub>2</sub>G linkers (Fig. 1B). For linkers of GGSGG and GG, annealed oligos of GGSGG and GG were ligated into pET-31b-Strep-KSI-G(GSG)<sub>8</sub>G-Strep-Thrombin-His<sub>6</sub> previously digested with AflII and XhoI restriction sites. DNA encoding for A<sub>2a</sub>R TM 7 was then introduced into pET-31b-Strep-KSI-G(GSG)<sub>8</sub>G-Strep-Thrombin-GGSGG-His<sub>6</sub> and pET-31b-Strep-KSI-G(GSG)<sub>8</sub>G-Strep-Thrombin-GG-His<sub>6</sub> at SalI and XhoI restriction sites.

Expression and immobilized metal affinity chromatography (IMAC) purification

Chemically competent Rosetta(DE3)/pLysS (Novagen) were transformed with the pET-31b vector containing the GPCR transmembrane peptide of interest. Transformants were selected by growth on Luria–Bertani (LB) plates supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Cells were grown overnight at 37 °C in 50 mL LB liquid media supplemented with 50 μg/mL ampicillin and 34 μg/mL chloramphenicol. The overnight culture was used to inoculate 1 L fresh media at an OD<sub>600</sub> of ~0.1 and grown at 37 °C. When the optical density reached ~0.6, isopropyl-β-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce expression. Cells were grown for four hours post-induction and were then harvested by centrifugation for 5 min at 10,000g. The cell pellet was resuspended in 20 mL lysis buffer (50 mM sodium phosphate, pH 8.0; 10 mM EDTA, 0.1% Triton X-100; 100 μg/mL lysozyme; complete EDTA-free protease inhibitors (Roche)). The suspension was treated with three freeze/thaw cycles (~80/37 °C). After the third thaw, the suspension was supplemented with MgCl<sub>2</sub> and DNAsE (Deoxyribonuclease I, Sigma, St. Louis MO) to final concentrations of 20 mM and 20 μg/mL, respectively. Samples were incubated at 25 °C for 20 min and then were centrifuged at 16,000g for 10 min to separate inclusion bodies from soluble proteins. The pellet was treated with 20 mL binding buffer (50 mM Tris–HCl, pH 8.0; 300 mM NaCl; 10% glycerol; 5 mM β-mercaptoethanol (β-ME); 20 mM imidazole) supplemented with 1% (w/v) fos-choline 16 for 1 h. at 25 °C to solubilize inclusion bodies. After centrifuging for 10 min at 16,000g, the supernatant containing solubilized inclusion bodies was collected and the pellet containing insoluble material was discarded.

Prior to purification, 5 mL of packed Ni–NTA resin (Qiagen) was pre-equilibrated for 1 h. at 25 °C with 20 mL binding buffer containing 0.1% fos-choline 16 and 20 mM imidazole. The resin was centrifuged 3,220g for 3 min and the equilibration buffer was removed. The resin was incubated in batch mode with solubilized inclusion bodies at 4 °C and tumbled gently on an end-over-end mixer overnight. The resin was centrifuged at 3,220g for 3 min and the supernatant containing unbound protein was removed. The resin was washed with 20 mL binding buffer containing 0.1% fos-choline 16 and 20 mM imidazole, then with 20 mL binding buffer containing 0.1% fos-choline 16 and 30 mM imidazole. The resin was then washed with 10 mL binding buffer containing 0.1% fos-choline 16 and 50 mM imidazole. Purified fusion proteins were eluted with 5 mL binding buffer containing 0.1% fos-choline 16 and 500 mM imidazole. Samples taken throughout the purification were analyzed by SDS–PAGE.

Thrombin cleavage

As determined by SDS–PAGE analysis and Coomassie staining, elutions enriched in purified fusion protein were pooled and dialyzed against thrombin cleavage buffer (20 mM Tris–HCl, pH 8.0; 150 mM NaCl; 2.5 mM CaCl<sub>2</sub>) supplemented with 5 mM β-ME and 0.0106% fos-choline 16. The concentration of dialyzed fusion proteins were determined by
UV spectroscopy, where the extinction coefficient for the fusion protein at 280 nm ($\varepsilon_{280}$) was calculated based on the Gill and von Hippel method [69]. Fusion proteins were diluted to 0.4 mg/mL in thrombin cleavage buffer supplemented with 5 mM β-ME and 0.0106% fos-choline and digested at 4 °C with 25 U/mL thrombin (Merck KGaA, Germany) on an end-over-end mixer. Samples corresponding to time = 0, 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, 16 and 24 h after addition of thrombin were removed, quenched with 5x non-reducing sample buffer (163 mM Tris–HCl, pH 6.8, 0.25 mg/mL Bromophenol blue, 0.5% SDS, 50% glycerol), and stored at 4 °C until analysis by SDS–PAGE.

Twenty microliters of each sample were loaded on 16% poly-acrylamide gels and electrophoresed at 25 °C at 150 V for 1.5 h in SDS running buffer (25 mM Tris–HCl, 192 mM Glycine, 1% SDS, pH 8.3). Following electrophoresis, gels were subjected to Coomassie staining (10% acetic acid, 10% 2-propanol, 0.25 mg/mL Coomassie blue for 4 h followed by 10% acetic acid, 0.25 mg/mL Coomassie blue for 4 h). Gels were destained with 10% acetic acid. Gel images were recorded at a resolution of 650 dpi using an Epson flatbed scanner. In order to determine thrombin processing efficiency, images were analyzed using ImageJ software, where the percentage of fusion protein remaining at a given time was determined by the ratio of optical densities of the fusion protein at $t = x$, where $x = 0, 0.25, 0.5, 1, 2, 4, 8, 16, 24$ h and $t = 0$ h.

**Purification of hA$_{2a}$R TM 7**

Expression of pET31b-Strep-KSI-G(GSG)$_8$G-Strep-Thrombin-GGSGG-hA$_{2a}$R TM 7-His$_6$ in Rosetta(DE3)/pLysS and IMAC purification of Strep-KSI-G(GSG)$_8$G-Strep-Thrombin-GGSGG-hA$_{2a}$R TM 7 -His$_6$ was performed as described previously. IMAC-purified fusion protein was extensively dialyzed against thrombin cleavage buffer supplemented with 5 mM β-ME and 0.0106% fos-choline 16 and the concentration of the dialyzed fusion protein was determined by UV spectroscopy. The fusion protein was then diluted to a concentration of 0.4 mg/mL in 10 mLs of thrombin cleavage buffer supplemented with 5 mM β-ME and 0.0106% fos-choline 16 and digested at 4 °C for 24 h with 25 U/mL thrombin on an end-over-end mixer. Thrombin cleavage products that contained the His$_6$-tag (unchelved fusion protein and hA$_{2a}$R TM 7) were purified by IMAC. The thrombin cleavage products were permitted to bind 1 mL Ni–NTA that had previously been equilibrated with thrombin cleavage buffer supplemented with 5 mM β-ME and 0.0106% fos-choline 16 for 2 h. The resin was centrifuged 3,220g for 3 min and the supernatant containing unbound protein was removed. The resin was washed with 10 mL thrombin cleavage buffer containing 5 mM β-ME, 0.0106% fos-choline 16, and 20 mM imidazole, then with 10 mL thrombin cleavage buffer containing 5 mM β-ME, 0.0106% fos-choline 16, and 30 mM imidazole. The resin was then washed with 10 mL thrombin cleavage buffer containing 5 mM β-ME, 0.0106% fos-choline 16, and 50 mM imidazole. IMAC-purified cleavage products were eluted with 1 mL thrombin cleavage buffer containing 5 mM β-ME, 0.0106% fos-choline 16, and 250 mM imidazole. Samples taken throughout the purification were analyzed by SDS–PAGE and proteins were visualized with Coomassie staining.

Elutions enriched in hA$_{2a}$R TM 7 peptide were combined and further purified from uncleaved fusion protein by differential affinity for Strept-Tactin resin. A 6 mL Kimble Chase Kontes FlexColumn was packed with 3 mL high-capacity Strept-Tactin resin (IBA-GmbH, Germany), washed extensively with water, and equilibrated by gravity flow with 40 mL thrombin cleavage buffer containing 5 mM β-ME and 0.0106% fos-choline 16. Three column volumes from the second IMAC purification that were enriched in hA$_{2a}$R TM 7 peptide were applied directly to the Strept-Tactin column. Fractions representing the column flow-through and additional washes were collected in 3 mL fractions. Samples taken throughout the purification were analyzed by SDS–PAGE and Coomassie staining.
Those fractions enriched in purified hA₂aR TM 7 peptide were pooled and concentrated using an Amicon centrifugal concentrator with a molecular weight cutoff of 3.5 kDa resulting in a final yield of ~ 0.55 mg/mL in 600 μL.

**Circular dichroism**

The secondary structure of GGSGG-hA₂aR TM 7 -His₆ in fos-choline 16 was characterized by circular dichroism spectroscopy. The purified peptide was dialyzed against 20 mM phosphate buffer, pH 7.4 containing 0.1% fos-choline 16 at 4 °C. This sample was then incubated with either buffer alone or buffer containing 5 mM Tris[2-carboxyethyl]phosphine (TCEP) for 1 h. at room temperature. The concentration of hA₂aR TM 7 analyzed by circular dichroism was 41.4 μM.

Far-UV CD spectra were obtained using a Jasco-J-810 spectropolarimeter with Peltier thermal control. All measurements were taken at 25 °C using a 0.1 cm path-length quartz cuvette. Spectra were collected from 190 to 250 nm using one nm step resolution and four s integration time. The appropriate reference CD spectra have been subtracted from the reported CD spectra for GGSGG-hA₂aR TM 7. CD spectra are expressed in mean molar ellipticities [Θ] (deg cm²dmol⁻¹):

\[ [\Theta] = \frac{MW \cdot \Theta_{abs}}{10 \times l \times c \times n} \] (1)

where \( \Theta_{abs} \) 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 ([Θ]₂₂₂):

\[ \%\alpha-\text{helix} = \left( \frac{[\Theta]_{222}}{[\Theta]_{max}} \right)^{1-(k/n)}, \text{where } k=4.0 \] (2)

**Applying the expression and purification system to other GPCR peptides**

In order to confirm that the expression system is a general method for production of GPCR transmembrane peptides, the fusion protein system was extended to peptides consisting of single and multiple transmembrane domains of hA₂aR and further validated on peptides of hA₁R and A₂bR.

Constructs that differed in the GPCR peptide were generated by digesting pET-31b-Strep-KSI-G(GSG)₈-Strep-Thrombin-G(GSG)₈ with Sall and Xhol and ligating similarly digested PCR products representing transmembrane peptides of hA₂aR, hA₁R, and A₂bR. Constructs confirmed by DNA sequencing were transformed into Rosetta(DE3)/pLysS and selected for growth on LB plates supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Control Rosetta(DE3)/pLysS and Rosetta(DE3)/pLysS containing a plasmid for GPCR peptide expression were grown overnight in 10 mL LB media containing either 34 μg/mL chloramphenicol or 50 μg/mL ampicillin and 34 μg/mL chloramphenicol, respectively, in a 37 °C water shaker. The overnight culture was used to inoculate 250 mL fresh LB media supplemented with the appropriate antibiotics at an OD₆₀₀ of ~0.1 and grown at 37 °C. When the optical density reached 0.6, expression was induced by the addition of 1 mM IPTG and permitted to grow for 4 h post-induction. Aliquots corresponding to 100 mL of the culture were harvested by centrifugation for 5 min at 10,000g. Cell pellets were then resuspended in 1 mL solubilization buffer (20 mM Tris–HCl, pH 8.0 and 1% SDS), sonicated to homogeneity, and centrifuged at 16,000g for 10 min to remove any remaining insoluble material. Samples were analyzed by SDS–PAGE and subsequent Western blot analysis.
Western blot analysis

Following electrophoresis, proteins were transferred from the gel to a 0.2 μm nitrocellulose membrane at 50 V and 4 °C for 2 h in transfer buffer (25 mM Tris–HCl, 192 mM glycine, 20% methanol, pH 8.2). Proteins present on the membrane were detected using a primary antibody of anti-His\textsubscript{6} IgG antibody (Covance, Princeton, NJ) at a dilution of 1:1000 in Tris-buffered saline (20 mM Tris–HCl, 150 mM NaCl, pH 7.6) containing 0.1% Tween 20 and 5% non-fat milk incubated overnight at 4 °C. The secondary antibody used for detection was Alexa Fluor\textsuperscript{®} 633 goat anti-mouse IgG diluted to 1:5000. Resulting data were imaged with a Typhoon 9600\textsuperscript{TM} variable mode imager (GE Bioscience, Piscataway, NJ).

Acknowledgments

The authors would like to thank Dr. Asokhan Anbanandam, Professor Tatyana Polenova, and Dr. David Raden for useful discussions during the initial development of this project. The authors would also like to thank Drs. Chris Tate (University of Cambridge, UK) and Paula Booth (University of Bristol, UK) for the Rant 5 plasmid. These studies were supported by NIH P20 RR15588 (ASR) and NIH T32 GH08550 (ZTB).

References


Fig. 1.
(A) Topology diagram of hA2aR denoting the model peptide hA2aR TM 7. The hA2aR TM 7 peptide is highlighted in gold, and includes residues 259–316. (B) Schematic of the expression construct highlighting design elements, including redundant Strep-tag II affinity tags, the ketosteroid isomerase carrier protein, the thrombin cleavage site, the GPCR peptide, and a His6 affinity tag. The glycine-rich linker downstream of the thrombin cleavage site has been varied by the number of GSG repeats, where $n$ is 0, 1, 2, 4, or 8.
Thrombin processing efficiency differs significantly for fusion proteins with different linkers. SDS–PAGE (16% acrylamide) shows thrombin processing of Strep-KSI-Strep-Thrombin-hA$_{2a}$R TM 7 -His$_{6}$ (A), Strep-KSI-G(GSG)$_{8}$G-Strep-Thrombin-GG-hA$_{2a}$R TM 7 -His$_{6}$ (B), and Strep-KSI-G(GSG)$_{8}$G-Strep-Thrombin-GGSGG-hA$_{2a}$R TM 7 -His$_{6}$ (C). Quantification of fusion protein cleavage for Strep-KSI-Strep-Thrombin-hA$_{2a}$R TM 7 -His$_{6}$ (yellow), Strep-KSI-G(GSG)$_{8}$G-Strep-Thrombin-GG-hA$_{2a}$R TM 7 -His$_{6}$ (grey), and Strep-KSI-G(GSG)$_{8}$G-Strep-Thrombin-GGSGG-hA$_{2a}$R TM 7 -His$_{6}$ (green) (D) shows that thrombin processing efficiency is improved with a glycine-rich linker flanking the cleavage site. Thrombin (T); fusion protein (FP); KSI-containing cleavage product (K); and hA$_{2a}$R TM 7 peptide (P) are indicated with arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 3.
Analysis of Strep-ksi-G(GSG)₈-G-Strep-Thrombin-GGSGG-hA₂aR TM 7-His₆ purification by IMAC. SDS–PAGE (12% acrylamide) shows samples from each step of the purification corresponding to 1 OD cells. Solubilized protein following extraction of insoluble fraction and inclusion bodies with Fos-choline 16 (IB); flow through (FT), 20 mM imidazole wash (W1); 30 mM imidazole wash (W2); 50 mM imidazole wash (W3); 500 mM imidazole elutions (E1, E2, E3). Full-length fusion protein (FP) is indicated with an arrow.
Fig. 4.
SDS–PAGE (16% acrylamide) shows the purification of GGSGG-hA$_2$aR TM 7 after thrombin cleavage by a second IMAC step (A) and Strep-Tactin affinity (B). Dialyzed fusion protein was digested with thrombin and the cleavage products were separated by IMAC. Samples of each purification step were analyzed including undigested fusion protein (primary IMAC), digestion products, flow through, 20 mM imidazole wash (W1); 30 mM imidazole wash (W2); 50 mM imidazole wash (W3); 250 mM imidazole elutions (E1, E2, E3). IMAC elutions consisting of uncleaved fusion protein and hA$_2$aR TM 7 peptide from the pooled eluate (E1–E3) were separated by affinity to high-capacity Strep-Tactin resin. Fractions collected from the flow-through and column washes W1 and W2 representing 3 mL thrombin cleavage buffer supplemented with FC16 are enriched in peptide and not uncleaved fusion protein. Thrombin (T); fusion protein (FP); KSI-containing cleavage product (K); and hA$_2$aR TM 7 peptide (P) are indicated with arrows.
Fig. 5.
Circular dichroism spectroscopy of A2aR TM 7 indicates α-helical secondary structure. SDS-PAGE (16% acrylamide) shows that purified and dialyzed hA2aR TM 7 peptide forms dimers through a disulfide bond, where addition of DTT reduces these intermolecular bonds (A). Purified hA2aR TM 7 was diluted to 41.4 μM in 20mM phosphate buffer, pH 7.4, 0.1% Fos-choline 16 with (closed circles) and without (open circles) 5 mM TCEP and characterized by CD (B). Fusion protein (FP); monomeric hA2aR TM 7 peptide (P-M); putative dimeric hA2aR TM 7 peptide (P-D) are indicated with arrows.
Expression of hA₂aR TM peptides in E. coli is TM-specific. SDS–PAGE (12% acrylamide) of cellular proteins corresponding to 1 OD cells analyzed by Coomassie staining (A) and anti-His₆ Western blot (B) shows samples obtained from strains expressing Strep-KSI-G(GSG)₈-G-Strep-Thrombin-G(GSG)₈G fusion protein with peptides corresponding to hA₂aR TM 7, TM 6–7, TM 5–7 and TM 6. A previously identified mutation in hA₂aR TM 6 that stabilizes the antagonist-bound form of hA₂aR (Rant 5) [65] had little effect on overall expression of the hA₂aR TM 6–7 fusion protein. Arrows indicate the fusion protein.
Fig. 7.
Reduced expression yields of hA2aR TM 6–7 are also observed in A1R TM 6–7 but not for A2bR TM 6–7. SDS–PAGE (12% acrylamide) of cellular proteins corresponding to 1 OD cells analyzed by Coomassie staining (A) and anti-His6 Western blot (B) shows samples obtained from strains expressing Strep-KSI-G(GSG)6-G-Strep-Thrombin-G(GSG)6-G fusion protein with peptides corresponding to hA1R, hA2aR and hA2bR TM 7 as well as hA1R, hA2aR and hA2bR TM 6–7 as indicated. All TM 7 peptides analyzed express at a high level, whereas only hA2bR TM 6–7 express at a high level. Arrows indicate the fusion protein.
Table 1

Primer sequences used in developing the expression constructs.

<table>
<thead>
<tr>
<th>Primer type</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward primers</strong></td>
<td></td>
</tr>
<tr>
<td>FPa KSI</td>
<td>GATATACATATGTGGAGCCACCCGCAGTTGCCGAAAAGATGCAATACCCCCAGACACATCAC</td>
</tr>
<tr>
<td>FP A2aR 259</td>
<td>GCAGGCTGCACTGCCCAGACTGCAG</td>
</tr>
<tr>
<td><strong>Reverse primers</strong></td>
<td></td>
</tr>
<tr>
<td>RP KSI</td>
<td>GAATTCGATCCGCATGCGTAATATTCCTTCGC</td>
</tr>
<tr>
<td>RP A2aR 316</td>
<td>GTGTAACCTCGAGTCACCTGAAGAACTGGTTGCTGAC</td>
</tr>
<tr>
<td><strong>Forward oligonucleotide</strong></td>
<td></td>
</tr>
<tr>
<td>FO Steap-Thrombin</td>
<td>GATCCTAGATCCGCGGCCGCCGACAGCTGTGGTGCGCAGAGCTTAAGGCGAG</td>
</tr>
<tr>
<td>FO B (GSG)8</td>
<td>GATCCGGTGGGTCTGGCGGATCGGGTGGGGAGTGGCGGATCGGATGCCGCTTCGGGGTTGGGAGG</td>
</tr>
<tr>
<td>FO A (GSG)8</td>
<td>TTAAGGCTTTTCGCCGCTGGGGAGTGGCGGATCGGATGCCGCTTCGGGGTTGGGAGG</td>
</tr>
<tr>
<td>FO A (GSG)4</td>
<td>TTAAGGCTTTTCGCCGCTGGGGAGTGGCGGATCGGATGCCGCTTCGGGGTTGGGAGG</td>
</tr>
<tr>
<td>FO A (GSG)2</td>
<td>TTAAGGCTTTTCGCCGCTGGGGAGTGGCGGATCGGATGCCGCTTCGGGGTTGGGAGG</td>
</tr>
<tr>
<td>FO A (GSG)0</td>
<td>TTAAGGCTTTTCGCCGCTGGGGAGTGGCGGATCGGATGCCGCTTCGGGGTTGGGAGG</td>
</tr>
<tr>
<td><strong>Reverse oligonucleotide</strong></td>
<td></td>
</tr>
<tr>
<td>RO Steap-Thrombin</td>
<td>TCGACTGCTTTAGGGCTCCTGGCGCCTGGAGTGGCTGCTACAGCTTAAGGCGAG</td>
</tr>
<tr>
<td>RO K (GSG)8</td>
<td>CCCACCCAGACCCACGGACCACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO S (GSG)8</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO S (GSG)4</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO X (GSG)8</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO X (GSG)4</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO X (GSG)2</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
<tr>
<td>RO X (GSG)0</td>
<td>TCGACCTACCCAGACCCAGATCGGACCTGCCCTACCTGATCCGCCACCTCCACAGATA</td>
</tr>
</tbody>
</table>

\(^{a}\)Forward primer (FP) reverse primer (RP), forward oligonucleotide (FO) and reverse oligonucleotide (RO). NdeI (CATATG), BamHI (GTCGAC), KpnI (GGTACC), AflII (CTTAAG), SalI (GTCGAC), and XhoI (CTCGAG) restriction sites are underlined.
Table 2

<table>
<thead>
<tr>
<th>Protein Expr Punif. Author manuscript; available in PMC 2013 August 01.</th>
<th>Table 2</th>
<th>Table 2</th>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing the expression constructs. Plasmids purchased from UMR cDNA resource center. Forward primer (FP) reverse primer nucleotide (FO) and reverse oligonucleotide (RO) sequences are listed in Table 1.</td>
<td><strong>Table 2</strong></td>
<td><strong>Table 2</strong></td>
<td><strong>Table 2</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Primers/oligonucleotides</strong></td>
<td><strong>Template/parental vector</strong></td>
<td><strong>Destination vector</strong></td>
</tr>
<tr>
<td></td>
<td><strong>FP</strong> Strep-KSI, <strong>RP</strong> KSI</td>
<td><strong>pET31b</strong></td>
<td>pET-23b</td>
</tr>
<tr>
<td></td>
<td>FO Strep-Thrombin, RO Strep-Thrombin</td>
<td>--</td>
<td>pET-23b-Strep-KSI-His6</td>
</tr>
<tr>
<td></td>
<td>FO B G(GSG)3G, RO K G(GSG)3G</td>
<td>--</td>
<td>pET23b-Strep-KSI-Strep-Thrombin-His6</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>pET23b-Strep-KSI-G(GSG)3G-Strep-Thrombin-His6</td>
<td>pET31b</td>
</tr>
<tr>
<td></td>
<td>FO A G(GSG)3G, RO S G(GSG)3G</td>
<td>--</td>
<td>pET31b-Strep-KSI-G(GSG)3G-Strep-Thrombin-His6</td>
</tr>
<tr>
<td></td>
<td>FO A G(GSG)3G, RO X G(GSG)3G</td>
<td>--</td>
<td>pET31b-Strep-KSI-G(GSG)3G-Strep-Thrombin-His6</td>
</tr>
<tr>
<td></td>
<td>FO A G(GSG)3G, RO X G(GSG)3G</td>
<td>--</td>
<td>pET31b-Strep-KSI-G(GSG)3G-Strep-Thrombin-His6</td>
</tr>
</tbody>
</table>
Table 3

Summary of thrombin cleavage and overall GGSGG-hA₂₉R TM 7-His₆ purification yields for large-scale preparation from Strep-KSI-G(GSG)₆-Strep-Thrombin-GGSGG-hA₂₉R TM 7-His₆. Note that only a portion of the total fusion protein was treated during thrombin step (equivalent to 8 mg inclusion body extract).

<table>
<thead>
<tr>
<th>Step</th>
<th>Total protein (mg)</th>
<th>Fusion protein (mg)</th>
<th>Peptide (mg)</th>
<th>Step yield of peptide (%)</th>
<th>Overall peptide yield (%)</th>
<th>Purity</th>
<th>Relative purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion body extract</td>
<td>39</td>
<td>19</td>
<td>(6.3)²</td>
<td>N/A</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nickel purification</td>
<td>17</td>
<td>17</td>
<td>(5.7)</td>
<td>90</td>
<td>90</td>
<td>33.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Thrombin cleavage</td>
<td>4</td>
<td>4</td>
<td>(1.3)</td>
<td>43</td>
<td>39</td>
<td>32.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Nickel purification</td>
<td>2.6</td>
<td>0.52</td>
<td>90</td>
<td>35</td>
<td>20</td>
<td>20</td>
<td>1.2</td>
</tr>
<tr>
<td>Streptactin purification</td>
<td>0.33</td>
<td>0.33</td>
<td>65</td>
<td>23</td>
<td>&gt;99</td>
<td>6.2</td>
<td></td>
</tr>
</tbody>
</table>

²Parentheses indicate that peptide mass was a calculated value based on mass fraction of peptide in the fusion protein.

bThrombin cleavage yield based on quantification of SDS–Page and Coomassie staining of fusion protein before and after thrombin digestion.
Cloning strategy for peptide expression constructs of the adenosine family of receptors. Plasmids purchased from UMR cDNA resource center. Forward primer (FP) and reverse primers (RP) listed in Table 5.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Primers/oligonucleotides</th>
<th>Template vector</th>
<th>Destination vector</th>
<th>Restriction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET-31b-Strep-KSI-G(GSG)G-Strep-Thrombin-G(GSG)G-hA2R TM 7-His&lt;sub&gt;6&lt;/sub&gt;</td>
<td>FP A&lt;sub&gt;1R&lt;/sub&gt; 260, RP A&lt;sub&gt;1R&lt;/sub&gt; 326</td>
<td>pCDNA3.1-hA&lt;sub&gt;1R&lt;/sub&gt;*</td>
<td>pET31b-Strep-KSI-G(GSG)G-Strep-Thrombin-G(GSG)G-hA&lt;sub&gt;2R&lt;/sub&gt;</td>
<td>SalI, XhoI</td>
</tr>
<tr>
<td>pET-31b-Strep-KSI-G(GSG)G-Strep-Thrombin-G(GSG)G-hA2R TM 6–7-His&lt;sub&gt;6&lt;/sub&gt;</td>
<td>FP A&lt;sub&gt;1R&lt;/sub&gt; 208, RP A&lt;sub&gt;1R&lt;/sub&gt; 326</td>
<td>pCDNA3.1-hA&lt;sub&gt;1R&lt;/sub&gt;*</td>
<td>pET31b-Strep-KSI-G(GSG)G-Strep-Thrombin-G(GSG)G-hA&lt;sub&gt;2R&lt;/sub&gt;</td>
<td>SalI, XhoI</td>
</tr>
</tbody>
</table>
Table 5

Description of peptides explored in this work. Forward primer (FP) and reverse primers (RP) listed in Table 5.

<table>
<thead>
<tr>
<th>GPCR Peptide</th>
<th>AA</th>
<th>Molecular weight</th>
<th>Oligonucleotides</th>
<th>PCR Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>hA2aR TM 7</td>
<td>259–316</td>
<td>6962.198</td>
<td>FP A2aR 259, RP A2aR 316</td>
<td>pCDNA 3.1 hA2aR</td>
</tr>
<tr>
<td>hA2aR TM 6–7</td>
<td>205–316</td>
<td>13152.621</td>
<td>FP A2aR 205, RP A2aR 316</td>
<td>pCDNA 3.1 hA2aR</td>
</tr>
<tr>
<td>hA2aR TM 5–7</td>
<td>170–316</td>
<td>17435.893</td>
<td>FP A2aR 170, RP A2aR 316</td>
<td>pCDNA 3.1 hA2aR</td>
</tr>
<tr>
<td>hA2aR TM 6</td>
<td>205–265</td>
<td>6922.217</td>
<td>FP A2aR 205, RP A2aR 265</td>
<td>pCDNA 3.1 hA2aR</td>
</tr>
<tr>
<td>Rant 5 TM 6–7</td>
<td>205–316</td>
<td>12968.379</td>
<td>FP A2aR 205, RP A2aR 316</td>
<td>pRG/III-hs-MBP-Rant 5 [65]</td>
</tr>
<tr>
<td>hA2bR TM 7</td>
<td>260–332</td>
<td>8087.404</td>
<td>FP A2bR 260, RP A2bR 332</td>
<td>pCDNA 3.1 hA2bR</td>
</tr>
<tr>
<td>hA2bR TM 6–7</td>
<td>211–332</td>
<td>13689.100</td>
<td>FP A2bR 211, RP A2bR 332</td>
<td>pCDNA 3.1 hA2bR</td>
</tr>
</tbody>
</table>

Molecular weights are based on the theoretical molecular weights calculated from the peptide primary sequence [70] and correspond to the peptide alone without contributions from the expression cassette.
Table 6

Primers used for generating peptides of the adenosine family of GPCRs. SalI to (GTCGAC) and XhoI (GTCGAC) restriction sites used for sub-cloning are underlined. A silent mutation was created in A2bR TM 6–7 to remove an XhoI restriction site, which is marked in bold.

<table>
<thead>
<tr>
<th>Forward primers</th>
<th>Reverse primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP A2aR 259</td>
<td>RP A2aR 316</td>
</tr>
<tr>
<td>GCAGC GTCGACTGCCCGACTGCAGC</td>
<td>GTGACTCGAAGTTTCTTGGTGCCTC</td>
</tr>
<tr>
<td>FP A2aR 205</td>
<td>RP A2aR 265</td>
</tr>
<tr>
<td>CAATAAGTCGAAGCTCAAGCTCAACAGAAG</td>
<td>GTCGAGGCCGAGCTGCCGCTGAG</td>
</tr>
<tr>
<td>FP A2aR 170</td>
<td>RP A2bR 326</td>
</tr>
<tr>
<td>GTACG GTCGAGGATGTTGAGTCCC CATGAACCTACA</td>
<td>TGATACGAGATTCCAGCGCAGCAG</td>
</tr>
<tr>
<td>FP A1R 260</td>
<td>RP A1R 326</td>
</tr>
<tr>
<td>CAATAATCGACTGCCGCTTGGCGCAC</td>
<td>TGATACGAGATTCCAGCGCAGCAG</td>
</tr>
<tr>
<td>FP A2bR 260</td>
<td>RP A2bR 332</td>
</tr>
<tr>
<td>GTACG GTCGACCAGCCATCAGGGTAAAAATAAG</td>
<td>GTGACTCGAAGTTTCTTGGTGCCTC</td>
</tr>
</tbody>
</table>

Protein Expr Purif. Author manuscript; available in PMC 2013 August 01.