Quality Control in Eukaryotic Membrane Protein Overproduction

Jennifer A. Thomas and Christopher G. Tate

MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK

Correspondence to Christopher G. Tate: cgt@mrc-lmb.cam.ac.uk.
http://dx.doi.org/10.1016/j.jmb.2014.10.012
Edited by B. Poolman

Abstract

The overexpression of authentically folded eukaryotic membrane proteins in milligramme quantities is a fundamental prerequisite for structural studies. One of the most commonly used expression systems for the production of mammalian membrane proteins is the baculovirus expression system in insect cells. However, a detailed analysis by radioligand binding and comparative Western blotting of G protein-coupled receptors and a transporter produced in insect cells showed that a considerable proportion of the expressed protein was misfolded and incapable of ligand binding. In contrast, production of the same membrane proteins in stable inducible mammalian cell lines suggested that the majority was folded correctly. It was noted that detergent solubilisation of the misfolded membrane proteins using either digitonin or dodecylmaltoside was considerably less efficient than using sodium dodecyl sulfate or foscholine-12, whilst these detergents were equally efficient at solubilising correctly folded membrane proteins. This provides a simple and rapid test to suggest whether heterologously expressed mammalian membrane proteins are indeed correctly folded, without requiring radioligand binding assays. This will greatly facilitate the high-throughput production of fully functional membrane proteins for structural studies.

Introduction

Structure determination of integral membrane proteins requires the production of milligrammes of pure, authentically folded protein for crystallisation [1]. As a natural prerequisite, the protein needs to be expressed in one of a number of heterologous expression systems, such as Escherichia coli, yeasts, insect cells or mammalian cells [2]. A number of expression strategies have been developed for each host system and many are now efficient for expression trials of hundreds of proteins in parallel [3]. A popular strategy for the expression of membrane proteins in E. coli is to generate fusion proteins with green fluorescent protein (GFP), which can be used as an indicator for both the quantity of protein expressed [4] and its relative stability upon detergent solubilisation by fluorescence-detection size-exclusion chromatography (FSEC) [5]. The utility of this strategy is that fluorescence of the fusion protein expressed in bacteria discriminates between correctly folded membrane protein (the GFP tag is fluorescent) and misfolded, aggregated membrane protein (the GFP tag is not fluorescent) [6,7]. In E. coli, it appears that the misfolded membrane protein promotes the formation of inclusion bodies and, once in an aggregate, the GFP is unable to fold and attain fluorescence. However, in eukaryotic cells, such as yeasts, insect cells used in the baculovirus expression system and in mammalian cells, GFP tagged to a membrane protein remains fluorescent regardless of whether the membrane protein is misfolded in the endoplasmic reticulum or correctly folded in the plasma membrane [8–11]. Higher eukaryotes have an efficient quality control system in the endoplasmic reticulum so that only folded proteins exit the endoplasmic reticulum, whilst misfolded proteins are retained for degradation [12]. Thus, GFP is not an appropriate marker for the folding status of membrane proteins produced using either mammalian cells or the baculovirus expression system, although it is still
useful in analysing the stability of a membrane protein in different detergents by FSEC.

The baculovirus expression system has proven efficient for the production of many eukaryotic membrane proteins, such as G protein-coupled receptors (GPCRs) [13], some of which have been crystallised and their structures determined [14]. However, recombinant baculovirus is not a panacea and there are many proteins that are poorly expressed and there have also been reports that some membrane proteins are expressed predominantly in a misfolded state [2,15]. This is a serious problem for structural biology, as it is not obvious from current methodology whether an overexpressed membrane protein is predominantly folded or misfolded. If misfolded material is inadvertently purified, then this will likely have a detrimental effect on the ability of the sample to crystallise and may also adversely affect the quality of any crystals obtained. The best way to determine whether misfolded material is present is to perform quantitative Western blotting to assess the total amount of membrane protein expressed in conjunction with radioligand binding assays to determine how much is functional [16,17]. However, this is expensive, difficult to perform and is also impossible for the majority of membrane proteins that do not possess high-affinity radioligands. It is also unclear whether the presence of misfolded overexpressed membrane protein is a rare event or whether it is commonly observed. We have therefore studied a number of membrane proteins produced both in stable mammalian cell lines and using the baculovirus expression system. The data show that all the four membrane proteins analysed are expressed in the baculovirus system as a mixture of folded and misfolded proteins, whereas mammalian cell lines are much more efficient at producing only correctly folded membrane proteins. A simple comparative detergent solubilisation assay is described, which is an excellent indicator for the presence of misfolded membrane proteins.

Results

Comparative expression of the angiotensin II type 1 receptor in insect cells and stable mammalian cell lines

The human angiotensin II type 1 receptor (AT₁R) is a GPCR with the typical predicted architecture of seven transmembrane regions with the N-terminus on the extracellular surface of the cell. The receptor contains three N-linked glycosylation sites with one in the N-terminal region (Asn4) and two in extracellular loop 2 (Asn176 and Asn188). Two expression systems were used for the production of AT₁R, the baculovirus expression system and stable tetracycline-inducible mammalian cell lines (the T-Rex system). AT₁R was expressed with a C-terminal decahistidine tag (H₁₀) from the polyhedrin promoter in the recombinant baculovirus bVA T₁R-H₁0. In inducible mammalian HEK293 cells (iHEK), AT₁R was expressed with a C-terminal GFP-H₁₀ tag from the CMV promoter after induction with tetracycline; the stable cell line iHEK(AT₁R-GFP-H₁₀) was generated through random integration of the plasmid in the genome followed by fluorescence-activated cell sorting to isolate a high-expressing clonal cell line. Initial analysis of expression was performed by Western blotting using an anti-penta-His horseradish peroxidase (HRP) conjugated antibody for detection (Fig. 1). AT₁R was extensively N-glycosylated in mammalian cells, which could be removed by treatment with PNGase F to yield a major product AT₁R-GFP-H₁₀ of apparent molecular mass of 60 kDa; no unglycosylated AT₁R-GFP-H₁₀ was visible in untreated cells. In Sf9 cells, AT₁R-H₁₀ was expressed as a mixture of glycosylated and unglycosylated receptor, which yielded a single major product (apparent molecular mass of 36 kDa) after treatment with PNGase F. The blot in Fig. 1 contained the same number of cells per lane; thus, an assessment of band intensities by eye suggested that there were similar levels of AT₁R expressed from the baculovirus expression system and from the stable mammalian cell line. However, despite the apparently similar levels of AT₁R polypeptide expressed in Sf9 and iHEK cells, radioligand binding assays showed that there was 20× more functional AT₁R expressed per cell in mammalian cells compared to the best baculoviral expression observed (Fig. 1). This implied that a large proportion of AT₁R expressed in insect cells was misfolded and incapable of binding antagonist.

Detergent solubilisation is the first step in the purification of a membrane protein; thus, the ability of AT₁R expressed stably in the iHEK cell line iHEK(AT₁R-GFP-H₁₀) to be solubilised by four different detergents was tested. The four detergents used in order of decreasing “harshness” [18] were sodium dodecyl sulfate (SDS), foscholine-12 (FC12), dodecylmaltoside (DDM) and digitonin. Digitonin was the mildest detergent used and it is very effective in maintaining membrane proteins in a functional state. DDM is one of the most popular mild detergents used for membrane protein purification, but it is a little harsher than digitonin. Only very few bacterial membrane proteins are sufficiently stable to maintain their integrity in either FC12 or SDS; thus, no ligand binding would be expected for AT₁R solubilised in either SDS or FC12. All four detergents were equally effective at solubilising AT₁R polypeptide expressed in iHEK cells (Fig. 2). However, as expected from the differing “harshness” of the detergents, only DDM and digitonin maintained the
integrity of $^{125}\text{I-Sar}^1$-bound AT$_1$R so that receptor-bound radioligand could be detected (Fig. 2). In contrast, bound radioligand was not detected when SDS was used to solubilise $^{125}\text{I-Sar}^1$-bound AT$_1$R, and only a small amount was detected when FC12 was used. Assays on DDM-solubilised AT$_1$R measured nearly twice as much receptor as detected in membranes (Fig. 2), which could be due to freeze-thawed membranes being a mixture of both right-side-out vesicles and inside-out vesicles, and the membrane-impermeant peptide $^{125}\text{I-Sar}^1$ could bind only to AT$_1$R in the right-side-out vesicles.

Detergent solubilisation of $^{125}\text{I-Sar}^1$-bound AT$_1$R from Sf9 cell membranes after expression from the recombinant baculovirus bvAT$_1$R-H$_{10}$ followed a similar pattern to that observed from the stable mammalian iHEK cell line; that is, double the amount of radioligand was observed in DDM-solubilised receptor compared to membranes and no binding was detected when SDS was used. Note that the binding data in Fig. 2 are normalised for ease of comparison, whereas in actuality, there is 20-fold less functional AT$_1$R per cell in Sf9 cells compared to the stable mammalian cell line. However, the Western blotting data of AT$_1$R produced in Sf9 cells are different from the analogous data from the iHEK cell line. Orders of magnitude more AT$_1$R polypeptide was solubilised from Sf9 cell membranes by SDS or FC12 compared to either DDM or digitonin (Fig. 2). It is reasonable to assume from the $^{125}\text{I-Sar}^1$ binding data that DDM solubilised all the functional AT$_1$R and therefore the difference between the signal on the Western blot for DDM-solubilised AT$_1$R and SDS-solubilised AT$_1$R represents misfolded AT$_1$R.

There is a significant difference between the Western blotting signal for SDS-solubilised AT$_1$R and DDM-solubilised AT$_1$R from Sf9 cells and that difference represents an amount of misfolded AT$_1$R that can be solubilised by SDS but not by DDM. However, there may be more misfolded AT$_1$R present in Sf9 cell membranes than suggested from the differential solubility in SDS versus DDM because it is plausible that DDM can also solubilise some AT$_1$R that cannot bind antagonist. To assess this possibility, we diluted membranes from the stable mammalian cell line iHEK(AT$_1$R-GFP-H$_{10}$) and insect cell membranes expressing AT$_1$R-H$_{10}$ to give the same amount of functional AT$_1$R per millilitre, solubilised in DDM and then analysed by Western blotting. The data (Fig. 3) showed clearly that there was considerably more AT$_1$R polypeptide solubilised from Sf9 cells than from the mammalian cell line and that this difference is due to misfolded receptor given that there was the same amount of functional receptor per lane. Efforts to decrease the amount of misfolded AT$_1$R in the baculovirus

![Western blot](image_url)

**Fig. 1.** Functional expression of AT$_1$R in mammalian cells that is 5-fold higher compared to insect cells. (a) Western blot of whole cells expressing AT$_1$R solubilised in SDS. Lanes 1 and 3, iHEK parental cells; lanes 2 and 4, iHEK(AT$_1$R-GFP-H$_{10}$) stable clonal cell line; lanes 5 and 7, uninfected Sf9 cells; lanes 6 and 8, bvAT$_1$R-H$_{10}$ infected Sf9 cells. N-Linked glycosylation was removed using PNGase F where indicated (+). Bands corresponding to AT$_1$R-GFP-H$_{10}$ in mammalian cells are indicated with a yellow asterisk (*). iHEK cell lines were induced with 1 μg/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 48 h. The blot was probed with an anti-pentaHis-HRP conjugated antibody. (b) The amount of functional AT$_1$R in each expression system was determined by measuring specific binding of the antagonist $[^{125}\text{I-Sar}]^1$. Baculoviral expression was performed in Sf9, Sf21 or Hi5 cells. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting. $[^{125}\text{I-Sar}]^1$-bound AT$_1$R is stable in DDM, but not in SDS. The amount of functional AT$_1$R was most accurately determined after solubilisation with DDM to ensure that all the receptor was accessible to ligand (see Fig. 2, where twice as much receptor could be measured upon solubilisation in DDM compared to in membranes). Each data point was determined in triplicate from two independent experiments and was plotted as mean ± SEM (standard error of the mean).
Misfolded AT₃R produced by the baculovirus expression system is poorly solubilised either by DDM or digitonin. (a) Western blot of AT₃R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGase F prior to SDS-PAGE. AT₃R was expressed either in the stable mammalian cell line iHEK(AT₃R-GFP-H10) or by using the recombinant baculovirus bvAT₃R-H₁₀ to infect Sf9 cells. The iHEK cell line was induced with 1 μg/ml tetracycline for 24 h and Sf9 cells were infected for 48 h. The Western blot insert is a 7× longer exposure. (b) The amount of functional detergent-solubilised AT₃R was determined by measuring specific binding of the antagonist [¹²⁵I]Sar1. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, AT₃R expressed in Sf9 cells; blue-filled bars, AT₃R expressed in iHEK cells. The amount of AT₃R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, AT₃R expressed in Sf9 cells; blue hatched bars, AT₃R expressed in iHEK cells. For ease of comparison, binding data have been normalised with respect to AT₃R in membranes (100%), which is equivalent to 1400 ± 240 dpm (n = 2; 380 fmol per million cells) for baculovirus-infected Sf9 cells and 12,000 ± 300 dpm (n = 2; 8.8 pmol per million cells) for iHEK(AT₃R-GFP-H10) cells. Absolute levels of AT₃R therefore cannot be compared meaningfully between the two expression systems using this bar graph (see Fig. 1). Binding assays for AT₃R contained either 150,000 Sf9 cells or 55,000 iHEK cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.
expression system either by using different cell lines (Sf21, Hi5) or by including an N-terminal signal sequence on AT1R were ineffective (Fig. 3).

In order to ascertain the quality of AT1R expressed in either mammalian cells or insect cells, we analysed two biophysical parameters of the detergent-solubilised receptor. Firstly, the thermostability of DDM-solubilised AT1R was determined and the apparent \( T_m \) values of \( 125^\text{I}-\text{Sar}^1 \)-bound AT1R expressed in either Sf9 cells or mammalian cells were found to be identical (Sf9 cells, 46 ± 0.8 °C; iHEK(AT1R-GFP-H10), 46 ± 0.7 °C). Secondly, the mobility on size-exclusion chromatography (SEC) of AT1R expressed using the baculovirus expression system in Sf9 cells or from the stable mammalian cell line was compared and also found to be very similar (Fig. 4). These data, coupled with the similarity in pharmacology between AT1R expressed in the two cell types [19], suggest that there is no significant difference between correctly folded AT1R produced in the baculovirus expression system and AT1R produced in the stable mammalian cell line.

The presence of misfolded protein upon overexpression from recombinant baculovirus is not uncommon

The presence of substantial amounts of misfolded AT1R upon production in the baculovirus expression system raised the question of whether this is specific for AT1R or whether other membrane proteins also exhibited this property. As it is not possible to test rigorously all membrane proteins, a careful selection was made of interesting test cases. The avian \( \beta_1 \)-adrenergic receptor (\( \beta_1\text{AR} \)) is a well-characterised GPCR and its structure has been determined bound to many different ligands of different efficacy [20–23]. All of the \( \beta_1\text{AR} \) crystals were grown from protein expressed in either Sf9 or Hi5 cells using recombinant

---

**Fig. 3.** DDM solubilises considerable amounts of inactive AT1R produced in the baculovirus expression system. (a) Western blot of DDM-solubilised AT1R, with equal amounts of active receptor per sample (lanes 2, 3 and 5–10). The blot was probed with an anti-penta-His-HRP conjugated antibody. Lane 1, iHEK parental cells; lanes 2 and 3, iHEK(AT1R-GFP-H10) stable clonal cell line; lane 4, uninfected Sf9 cells; lanes 5–10, bvAT1R-H10 infected insect cells. N-Linked glycosylation was removed using PNGase F where indicated (+). AT1R was expressed either in the stable mammalian cell line iHEK(AT1R-GFP-H10) or by using the recombinant baculoviruses bvAT1R-H10 and bvAT1R-LS-H10 to infect Sf9, Sf21 and Hi5 cells as indicated. iHEK cell lines were induced with 1 \( \mu \)g/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 48 h. The amount of functional AT1R was determined by measuring specific binding of the antagonist \( 125^\text{I}\)-Sar1.

---

**Fig. 4.** Comparison of AT1R expressed in mammalian cells and insect cells. (a) Stability of DDM-solubilised AT1R bound to the antagonist \( 125^\text{I}\)-Sar1. AT1R was expressed using three different expression systems: blue circles, baculovirus bvAT1R in Sf9 cells; red diamonds, stable clonal cell line iHEK(AT1R-GFP-H10). The apparent \( T_m \) values of AT1R expressed in each system are as follows: Sf9 cells, 46 ± 0.8 °C; iHEK(AT1R-GFP-H10), 46 ± 0.7 °C. Each data point was determined in triplicate and was plotted as a mean value ± SEM. (b) SEC was carried out using a Superdex 200 10/300 (24 ml) column. The elution of iHEK(AT1R-GFP-H10) was detected using GFP fluorescence (mV). The elution of bvAT1R-H10 was detected by Western blotting and band quantification (ImageJ value). iHEK(AT1R-GFP-H10) shows a symmetrical peak whereas bvAT1R-H10 shows two peaks; however, both systems show elution of a protein of a similar size. The void \( (V_o) \) and total \( (V_T) \) column volumes are indicated.
baculoviruses [24,25]. The assays described for AT1R were therefore repeated using wild-type β1AR with truncations at the N-terminus and at the C-terminus (bvβ1AR-H10), which facilitates expression of a homogenous protein. Comparison of the amount of β1AR-H10 solubilised by the different detergents clearly indicates that a large proportion of the receptor is indeed misfolded, as suggested by the higher proportion of receptor solubilised by either SDS or FC12 compared to either DDM or digitonin (Fig. 5). Efforts to improve the proportion of folded protein by using a thermostable β1AR mutant fused at the N-terminus to a leader sequence and a well-folded soluble protein (thioredoxin) did not increase the proportion of correctly folded β1AR (Fig. 5). However, it is interesting to note that receptor containing an uncleaved leader sequence was only extracted by SDS or FC12, suggesting that this sub-population of receptor was probably mainly

![Figure 5](image)

Fig. 5. Misfolded β1AR is poorly solubilised either by DDM or digitonin. (a) Western blot of β1AR solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. β1AR was expressed by using the recombinant baculovirus bvβ1AR-H10 to infect Hi5 cells. tsβ1AR was expressed by using the recombinant baculovirus bvtsβ1AR-H10 to infect Sf9 cells. Hi5 and Sf9 cells were infected for 48 h. The broken line indicates two separate blots. (b) The amount of functional detergent-solubilised β1AR and tsβ1AR was determined by measuring specific binding of the antagonist [3H]DHA. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, β1AR expressed in Hi5 cells; blue-filled bars, tsβ1AR expressed in Sf9 cells. The amount of β1AR in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, β1AR expressed in Hi5 cells; blue hatched bars, tsβ1AR expressed in Sf9 cells. For ease of comparison, binding data have been normalised with respect to β1AR in membranes (100%), which is equivalent to 11,000 ± 550 dpm (n = 3; 6.1 pmol per million cells) for baculovirus-infected Hi5 cells and 2600 ± 190 (n = 3; 1.4 pmol/l) for bvtsβ1AR-H10 infected Sf9 cells. All binding assays for β1AR and tsβ1AR contained 8300 cells, and therefore, comparison on absolute levels of receptor can be directly compared. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.
misfolded. In addition, it is unlikely that the fusion protein was efficiently trafficked to the cell surface given that FC12 extraction resulted in a 3-fold increase in the amount of receptor binding obtained compared to when membranes were used.

In a second example, we compared the expression of the adenosine A1 receptor (A1R) in both the stable mammalian cell line iGnTI − (A1R-GFP-H10) and the insect cells using the baculovirus expression system (bvA1R-H10). N-Linked glycosylated sites are found in extracellular regions of A1R (extracellular loop 2; Asn148 and Asn159), which produces a glycosylated form in mammalian cells that can be reduced in molecular weight by treatment with PNGase F, whereas the majority of the receptor is unglycosylated in Sf9 cells (Fig. 6). Expression of A1R-H10 in insect cells gave comparative Western blots analogous to those observed for AT1R-H10, with SDS and FC12 extracting orders of magnitude more polypeptide from insect cell membranes compared to either DDM or digitonin, consistent with a large excess of misfolded receptor produced in the baculovirus expression system (Fig. 7). In contrast, all the detergents used to solubilise A1R-GFP-H10 from a stable mammalian cell line were equally efficacious, indicating that there is minimal misfolded receptor in these cells (Fig. 7). The low levels of antagonist binding activity observed for A1R is a consequence of the poor stability of this receptor in detergent solutions. A1R also provided a nice example of the usefulness of confocal microscopy in defining whether a receptor is likely to be correctly folded. A1R-GFP-H10 is expressed in the stable cell line predominantly at the cell surface whereas a mutant of A1R that contained multiple changes introduced to try and facilitate crystallisation (A1R-GL26-GFP-H10; see Methods) was expressed predominantly in intracellular membranes (Fig. 8). The confocal data correlated well with the Western blotting data. The misfolded mutant A1R-GL26-GFP-H10 was only efficiently extracted from mammalian cells with SDS (Fig. 9), whereas the wild-type receptor was extracted equally efficiently using either digitonin or SDS (Fig. 7).

The final example we tested was the serotonin transporter (SERT). The expression of SERT has been studied intensively[16,17,26] and was the first example along with rhodopsin that showed the utility of mammalian cells for the overexpression of functional membrane protein using the tetracycline-inducible HEK293 cell line [27,28]. Here we demonstrated that the Western blot data show an identical pattern of results with constructs based on wild-type A1R and AT1R, namely, similar amounts of extractable SERT-SAH9-GFP-H10 from the mammalian cell line, regardless of the detergent used, whereas there are orders of magnitude more SDS-extractable SERT-H10 in Sf9 cells compared to the amount solubilised by digitonin or DDM (Fig. 10).

A simplified assay for the detection of misfolded membrane proteins

Analysis of the data in Figs. 1–10 suggests that the salient conclusions of this paper, that is, that the majority of AT1R, A1R and SERT constructs...
Fig. 7 (legend on next page)
expressed in insect cells were misfolded, whereas expression in mammalian cells produced correctly folded protein, could be deduced with a fraction of the work. Comparison of two lanes in each Western blot, namely, SDS-extracted protein and digitonin-extracted protein, is sufficient to draw the relevant conclusions. Importantly, this obviates the need for radioligand binding assays and a stable mammalian cell line for each membrane protein to be studied. Radioligands have been developed for only a small fraction of membrane proteins and not all radioligands are of sufficiently high affinity (100 nM or better) to make them suitable for assays on detergent-solubilised membrane proteins. In addition, construction of stable mammalian cell lines can take many months and sometimes the cell lines grow very poorly due to basal activity of the membrane protein. This was noticeable for the stable A1R cell line iGnTI−(A1R-GFP-H10) developed here, which grew very poorly compared to the stable cell line iGnTI+(A1R-GL26-GFP-H10) expressing the inactive A1R mutant, despite the use of an inducible promoter.

Using the methodology described in this paper, it would be relatively simple to test 50 or so different membrane protein expression trials in a day. However, if hundreds of samples are to be tested in 96-well plates, then the ultracentrifugation step will become limiting and will need to be replaced using filtration through low-protein-binding 0.2-μm filters. The use of a dot-blot apparatus and semi-quantification of the resulting signals in relation to a known standard would be sufficient to define how much functional membrane protein could be extracted and whether or not extra precautions may be required to remove potential misfolded protein during purification. It would also be possible to measure the fluorescence of a GFP-tagged membrane protein, rather than performing a Western blot, to improve further the high-throughput capabilities of this assay.

Discussion

A commonly held misconception, particularly amongst investigators new to the membrane protein field, is that if a membrane protein can be expressed into a membrane within a cell and can be extracted with detergent, then that membrane protein is folded authentically. Over the last 30 years, there have been sporadic reports of overexpressed membrane proteins in E. coli, yeast or the baculovirus expression system being predominantly misfolded and inactive [2]. The work presented here demonstrates that the baculovirus expression system is particularly prone to producing misfolded membrane proteins, even of apparently uncomplex GPCRs that were expressed over 20 years ago. However, the simple assay proposed here will rapidly demonstrate whether misfolded membrane protein is indeed present. A few words of caution are warranted with regard to the differential solubility assay. Firstly, we have tested the assay on membrane proteins expected to be expressed in the plasma membrane of mammalian cells, which is efficiently solubilised by DDM. This is evident from the similar levels of solubilisation between DDM and SDS of correctly folded membrane proteins in the plasma membrane.

Fig. 7. Misfolded A1R produced by the baculovirus expression system is poorly solubilised either by DDM or digitonin. (a) Western blot of A1R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGase F prior to SDS-PAGE. A1R was expressed either in the stable mammalian cell line iGnTI−(A1R-GFP-H10) or by using the recombinant baculovirus bvA1R-H10 to infect Sf9 cells. The iGnTI−cell line was induced with 1 μg/ml tetracycline for 24 h and Sf9 cells were infected for 72 h. (b) The amount of functional detergent-solubilised A1R was determined by measuring specific binding of the antagonist [3H]DPCPX. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, A1R expressed in Sf9 cells; blue-filled bars, A1R expressed in iGnTI−cells. The amount of A1R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, A1R expressed in Sf9 cells; blue hatched bars, A1R expressed in iGnTI−cells. For ease of comparison, binding data have been normalised with respect to A1R in membranes (100%), which is equivalent to 120,000 ± 2000 dpm (n = 3; 2.9 pmol per million cells) for baculovirus-infected Sf9 cells and 7500 ± 250 dpm (n = 3; 3.8 pmol per million cells) for iGnTI−(A1R-GFP-H10) cells. Absolute levels of A1R therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for A1R contained either 150,000 Sf9 cells or 7500 iGnTI−cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM. (c) Western blot of DDM-solubilised A1R, with equal amounts of active receptor per sample (lanes 2, 3, 5 and 6). The blot was probed with an anti-pentaHis-HRP conjugated antibody. Lane 1, iGnTI−parental cells; lanes 2 and 3, iGnTI−(A1R-GFP-H10) stable cell line; lane 4, uninfected Sf9 cells; lanes 5 and 6, bvAT1R-H10 infected Sf9 cells. N-Linked glycosylation was removed using PNGase F where indicated (+). A1R was expressed either in the stable mammalian cell line iGnTI+(A1R-GL26-GFP-H10) or by using the recombinant baculovirus bvAT1R-H10 to infect Sf9 cells. Bands corresponding to A1R-GFP-H10 in mammalian cells are indicated with a yellow asterisk (*). The iGnTI−cell line was induced with 1 μg/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 72 h. The amount of functional A1R was determined by measuring specific binding of the antagonist [3H]DPCPX binding.
Secondly, we are using the assay as a guide rather than as an exact measure for determining the number of molecules of the target membrane protein that are correctly folded compared to the number of molecules that are misfolded.

Knowing that a proportion of an expressed membrane protein is misfolded is important. Many efforts have been made to parallelise expression of membrane proteins to facilitate high-throughput post-genomic approaches to determine rapidly membrane protein structures [3]. Although it has proven possible to do this for bacterial membrane proteins, it has proven harder to replicate these strategies for mammalian membrane proteins, partly because yields of membrane protein suggested from the quantification of polypeptide expressed have not reflected the yield of purified membrane protein. There are two factors that could explain this. Firstly, as described here, most of the membrane protein could be expressed in a misfolded state and therefore cannot be purified in mild detergents. Secondly, membrane proteins are often unstable in detergent and therefore they become...
inactive and aggregate during solubilisation and purification. The assay described here will define which is the problematic step, thus directing resources to solving the relevant problem. For example, knowing that the majority of membrane protein is misfolded in the baculovirus expression system suggests that using
stable inducible mammalian cell lines could improve yields [29]. In the work described here, the AT1R expressed in the baculovirus expression system would yield only 0.1 mg/l of functional receptor, whereas the stable clonal cell line iHEK(AT1R-GFP-H10) would yield 0.5 mg/l. However, the major advantage of using the mammalian expression system is that there is little or no misfolded AT1R expressed.

Is the misfolded membrane protein expressed in insect cells a potential problem for downstream purification and crystallisation? Even though DDM is a mild detergent and cannot solubilise misfolded
protein as well as SDS or FC12, misfolded AT₁R is the major component of DDM-solubilised insect cell membranes. In the initial stages of a project, this could be highly misleading, as it would appear that major losses were being incurred on, for example, the first Ni²⁺-affinity column, when in actual fact, it may be the case that the only protein lost was the misfolded material and that the yields of the correctly folded protein were around 80–90%. In the worst instance, researchers may note that FC12 extracts more of the target protein than DDM and then waste many years trying to purify and crystallise this material, not knowing that the target protein was likely to be totally inactive. Interestingly, the work here shows that β₁AR is expressed as a mixture of both folded and misfolded receptors, but β₁AR was purified and crystallised and its structure was determined without knowing this. Two effects may help in reducing the impact of misfolded membrane proteins on crystallisation trials. Firstly, SEC is a frequently used step in protein purification and will effectively remove any misfolded protein. Secondly, misfolded membrane proteins have a tendency to aggregate; thus, this portion may just “disappear”, either through retention on columns by non-specific effects or by being unable to pass through pre-filters that are normally present upstream of columns run on automated protein purification equipment. Thirdly, during crystallisation, any remaining inactive protein will precipitate more readily than the folded protein, hopefully allowing crystals to form later on.

Why are misfolded membrane proteins produced in the baculovirus expression system? Although there are many potential differences between insect cells and mammalian cells that may reduce the efficiency of membrane protein folding (e.g., potential specificity and amounts of molecular chaperones, different lipid composition, etc.), there are two overriding factors that have to be considered. Baculoviruses are lytic viruses, and one of the first effects of the virus is to impair the cells’ secretory pathway, which is precisely where membrane proteins are folded. Thus, during the infection cycle, the rate of secretion decreases and it is also observed that post-translational modifications (e.g., potential specificity and amounts of molecular chaperones, different lipid composition, etc.) are determined without knowing this. Two effects may help in reducing the impact of misfolded membrane proteins on crystallisation trials. Firstly, SEC is a frequently used step in protein purification and will effectively remove any misfolded protein. Secondly, misfolded membrane proteins have a tendency to aggregate; thus, this portion may just “disappear”, either through retention on columns by non-specific effects or by being unable to pass through pre-filters that are normally present upstream of columns run on automated protein purification equipment. Thirdly, during crystallisation, any remaining inactive protein will precipitate more readily than the folded protein, hopefully allowing crystals to form later on.

Materials and Methods

Materials

All radiolabelled ligands were purchased from PerkinElmer: [125I]Sar¹-Ile⁶-angiotensin II ([125I]Sar¹), [3H]dihydroalprenolol ([3H]DHA), [3H]dipropylcyclopentylxanthine ([3H]DPCPX) and [125I]2β-carbomethoxy-3β-(4-iodophenyl)tropane ([125I]RTI-55). The detergents n-dodecyl β-d-maltopyranoside (DDM) and fos-choline-12 (FC12) were purchased from Anatrace; SDS was purchased from Sigma and digitonin was purchased from Calbiochem. Anti-penta-histidine antibody conjugated to HRP (anti-pentaHis-HRP) was purchased from Qiagen. A tetracycline-inducible HEK293 cell line, T-Rex™, was purchased from Invitrogen. A tetracycline-inducible HEK293S cell line lacking N-acetylgalactosaminylationtransferase I (GnTI−) was kindly provided by Philip J. Reeves (Massachusetts Institute of Technology) [56].

Methods

Constructs

Expression in mammalian cells was performed using derivatives of pcDNA4/TO (Invitrogen). The serotonin transporter cDNA was inserted into the EcoRV/NotI restriction sites in pcDNA4/TO, for expression from the tetracycline-inducible CMV promoter, and then a cassette encoding enhanced GFP, the StreptII tag and a decahistidine (H₁₀) tag was inserted after SERT in the NotI/AlaI sites (plasmid pJMA111, kindly provided by J. Andred, MRC Laboratory of Molecular Biology). The cDNA clone for human angiotensin II type 1 receptor (AT₁R) was obtained from the Missouri S&T cDNA Resource Center, amplified by polymerase chain reaction, flanked with EcoRV and NotI sites and cloned into the corresponding sites of pJMA111 to create plasmid pJAP2, which expressed AT₁R-GFP-H₁₀. Additionally, the cDNA for the human adenosine A₁R (Missouri S&T cDNA Resource Center) was cloned similarly into the EcoRV/NotI sites to create plasmid pJAP34, which expressed A₁R-GFP-H₁₀. In an effort to create a thermo-stable A₁R receptor, four mutations that stabilised the adenosine A₂A receptor in the active state (L48A, A54L, T65A, Q89A) [37] were transferred to A₁R (mutations L51A, A57L, L68A, Q92A). In addition, the mutations N148G and N159G were included to remove the putative N-linked expression systems that utilise viruses with strong promoters, such as the semiliki forest virus expression system, misfolded and inactive membrane protein has also been observed [19,34,35]. Thus, the current successes with the production of authentically folded membrane proteins in mammalian cells for structural studies are all about ensuring that there is a balance between the amount of mRNA produced and the ability of the membrane protein to fold [29]. This will be different for each membrane protein and will have to be optimised empirically on a case-by-case basis. However, the differential solubility assay described here will ensure that expression of only the correctly folded membrane protein will be optimised.
glycosylation sites. To remove flexible regions, we truncated the N-terminus between Pro2 and Ile5, truncated the C-terminus at Phe307 and also added the sequence VLRQEEPFKAA to the C-terminus, thus generating A1R-GL26. A synthetic cDNA encoding A1R-GL26 (Life Technologies) was cloned into the EcoRV/NotI sites in pJMA111 creating pJAP37, which expressed A1R-GL26-GFP-H10. For generating baculoviruses, A1R was cloned into the BamHI/EcoRI sites of the transfer vector pBacPAK8 (Clontech), A1R was cloned into the Xhol/EcoRI sites and A1R-GL26 was cloned into the EcoRI/EagI sites, creating plasmids pJAP15, pJAP44 and pJAP33, respectively. Additionally, AT1R was cloned into the BamHI/EcoRI sites in plasmid pAccGP67-B (BD Biosciences) in order to utilise the acidic glycoprotein gp67 signal sequence (LS) preceding the N-terminus of AT1R, creating plasmid pJAP16, which expressed AT1R-LS-H10. All baculovirus sequences were engineered to contain a C-terminal tobacco etch virus cleavage site and H10 tag. All constructs were verified by DNA sequencing (Source Biosciences, UK).

**Transient transfection, generation of stable cell lines and protein expression**

Mammalian expression plasmids for the expression of AT1R (pJAP2), A1R (pJAP43) and A1R-GL26 (pJAP37) were amplified in E. coli strain DH5α, purified using a Maxi-prep kit (Qiagen) and transiently transfected (GeneJuice, Novagen) into adherent mammalian iHEK cells or iGnTI cells following the manufacturer’s protocol. Cells were grown in Dulbecco’s modified Eagle’s media supplemented with 10% tetracycline-free foetal bovine serum (Invitrogen) and 5 μg/ml blasticidin (Invitrogen) and incubated at 37 °C in an atmosphere containing 5% CO2. Expression of plasmids was induced by addition of 1 μg/ml tetracycline and incubated at 37 °C for 24 h. Stable cell lines were generated by selection with media containing 200 μg/ml Zeocin (Invitrogen). An iGnTI stable cell line expressing a thermostable mutant of SERT, SERT-SA9 (J. Andréll and C. Tate, unpublished results; Ref.[38]) and (iGnTI- SERT-SA9-GFP-H10) was kindly provided by J. Andréll. A highly expressing clonal AT1R-GFP-H10 cell line (iGnTI (J. Andréll and C. Tate, unpublished results; Ref.[38]) and expressing a thermostable mutant of SERT, SERT-SAH9-GFP-H10 (J. Andréll and C. Tate, unpublished results; Ref.[38]) and (iGnTI- SERT-SA9-GFP-H10) was kindly provided by J. Andréll. A highly expressing clonal AT1R-GL26-GFP-H10 cell line was selected from a polyclonal cell line using fluorescence-activated cell sorting. After expression, cells were washed twice in phosphate-buffered saline (PBS), counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1200 g for 5 min) and resuspended at 10 million cells per millilitre in ice-cold cell buffer [50 mM Tris (pH 7.4) and 150 mM NaCl supplemented with Complete EDTA-Free Protease Inhibitor Cocktail (Roche)]. Cell suspensions were flash frozen in liquid nitrogen and stored at ~80 °C.

**Recombinant baculovirus generation and protein expression**

Recombinant baculoviruses that expressed AT1R, A1R or A1R-GL26 were generated using the BaculoGold Baculovirus Expression System according to manufacturer’s protocol (BD Biosciences). Viruses were isolated by plaque purification and screened for expression by Western blotting using an anti-pentaHis-HRP antibody. Recombinant baculovirus that expressed SERT with a H10 tag at its C-terminus was kindly provided by R. Nehme (MRC Laboratory of Molecular Biology) and a thermostable β1AR fused to thioredoxin (tsβ1AR) was kindly provided by T. Warne (MRC Laboratory of Molecular Biology). Recombinant baculoviruses were passaged twice in Sf9 cells to obtain high titre stocks. Viruses were used to infect Sf9, Sf21 or Hi5 cells for 48 or 72 h as indicated. After protein expression, cells were counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1200 g for 5 min) and washed twice in PBS, and the cell pellet was resuspended at 10 million cells per millilitre in ice-cold cell buffer [50 mM Tris (pH 7.4) and 150 mM NaCl supplemented with Complete EDTA-Free Protease Inhibitor Cocktail (Roche)]. Cell suspensions were flash frozen in liquid nitrogen and stored at ~80 °C.

**Western blotting**

Cell suspensions were sonicated briefly and the total protein concentration was determined using the Bradford assay [39]. Samples were then solubilised in the detergent indicated [SDS, FC12, DDM or digitonin; all at 1% (w/v) final concentration] at either 4 °C (FC12, DDM, digitonin) or 20 °C (SDS) for 1 h. For blots corresponding to the differential solubility assay, the solubilisate was centrifuged at 280,000 g for 30 min at 4 °C to remove the insoluble fraction. SDS-loading buffer was added to the supernatant (corresponding to approximately 150,000 cells), and samples were separated on a 4–20% Tris glycine gel and transferred to nitrocellulose using standard techniques. Membranes were probed with anti-pentaHis-HRP at a dilution of 1:1000 and developed using enhanced chemiluminescence (GE Healthcare). Where indicated, 2 μl of PNGase F (New England Biolabs) was added to 15 μl of the supernatant and incubated at 37 °C for 1 h prior to SDS-PAGE to remove N-linked glycosylation.

**Thermostability assay of detergent-solubilised AT1R**

The cell suspension containing unpurified AT1R was sonicated briefly and diluted into buffer [50 mM Tris (pH 7.4), 5 mM MgCl2, 1 mM EDTA, 0.1% (w/v) bovine serum albumin, 150 mM NaCl and 40 μg/ml bacitracin, [125I]Sar1-labeled Sar1 and unlabelled Sar1 were added to give final concentrations of 0.5 nM and 100 nM, respectively, and incubated for 1 h at room temperature before chilling on ice and solubilising in 1% DDM (w/v, final concentration). The samples were then heated at varying temperatures for 30 min and the [125I]Sar1-bound receptor was separated from the free radioligand by gel-filtration spin columns as described previously [40–43]. Background was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells. Each reaction was performed in triplicate. Results were evaluated by nonlinear regression using GraphPad Prism.

**Detergent-solubilised and membrane-bound radioligand binding assays**

Cell suspensions were sonicated briefly and the total protein concentration was determined using the Bradford assay [39]. Cells were then diluted into buffer [150 mM NaCl and 50 mM Tris (pH 7.4)], incubated with the
respective radioligand (1 h, 4 °C) and solubilised in a final concentration of 1% detergent (DDM, FC12, digitonin, SDS) for 1 h at 4 °C. [3H]DHA was used at a final concentration of 200 nM and [3H]DPCPX was used at a final concentration of 39 nM in 150 mM NaCl and 50 mM Tris (pH 7.4). [125I]RTI-55 was used at a concentration of 1 nM in PBS. [125I]Sar1 was used as per the thermostability assay mentioned above. Bound and free radioligands were separated on gel-filtration spin columns as above.

To determine the amount of SERT, AT1R, β1AR or A1R present in cell membranes, we performed binding assays as mentioned above but without the samples being solubilised with detergent. Separation of receptor-bound and free radioligands was achieved by filtration through a 96-well glass fibre filter plates (Millipore) pre-treated with 0.1% polyethylenimine [38] except for [125I]Sar1 where no polyethyleneimine was used. Background for both assays was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells.

**FSEC and SEC analysed by Western blotting**

The void volume (8.16 ml) of the Superdex 200 10/300 (24 ml) (GE healthcare) was determined by running blue dextran through the column and observing where it eluted using A280. For FSEC, approximately 5 million iHEK(AT1R-GFP-H10) cells were thawed on ice and sonicated briefly. Cells were incubated at room temperature. After washing with PBS, we selectively stained membranes using a solution (10 μg/ml) of concanavalin A (ConA)–Alexa Fluor 647 conjugate (Invitrogen) in PBS for 10 min at room temperature. After washing with PBS, we stored icells n fresh PBS with 0.02% Na azide at 4 °C protected from light. Cells were visualised on a Leica TCS SP8 STED inverted laser-scanning microscope with 63x oil-immersion objective and a 1.4 numerical aperture. The white light laser was set to a wavelength of 488 nm to excite GFP and to 633 nm for Alexa Fluor 647 with the pinhole emission wavelength set to 580 nm.

**Fixing and staining cells for analysis by confocal laser-scanning microscopy**

Cells were grown on 35-mm glass bottom culture dishes, induced for 24 h under standard conditions and fixed using 2% paraformaldehyde [28]. After washing with PBS, we selectively stained membranes using a solution (10 μg/ml) of concanavalin A (ConA)–Alexa Fluor 647 conjugate (Invitrogen) in PBS for 10 min at room temperature. After washing with PBS, we stored icells n fresh PBS with 0.02% Na azide at 4 °C protected from light. Cells were visualised on a Leica TCS SP8 STED inverted laser-scanning microscope with 63x oil-immersion objective and a 1.4 numerical aperture. The white light laser was set to a wavelength of 488 nm to excite GFP and to 633 nm for Alexa Fluor 647 with the pinhole emission wavelength set to 580 nm.

**References**


