METHODS AND APPLICATIONS

Fluorescent probe for high-throughput screening of membrane protein expression

A. E. Backmark,¹ N. Olivier,² A. Snijder,¹ E. Gordon,¹ N. Dekker,¹ and A. D. Ferguson^{2*}

¹Discovery Sciences, Reagents and Assay Development, AstraZeneca Pharmaceuticals, Mölndal, Sweden

²Discovery Sciences, Structure and Biophysics, AstraZeneca Pharmaceuticals, Waltham, Massachusetts 02451

Received 27 March 2013; Revised 2 June 2013; Accepted 6 June 2013

DOI: 10.1002/pro.2297

Published online 15 June 2013 proteinscience.org

Abstract: Screening of protein variants requires specific detection methods to assay protein levels and stability in crude mixtures. Many strategies apply fluorescence-detection size-exclusion chromatography (FSEC) using green fluorescent protein (GFP) fusion proteins to qualitatively monitor expression, stability, and monodispersity. However, GFP fusion proteins have several important disadvantages; including false-positives, protein aggregation after proteolytic removal of GFP, and reductions in protein yields without the GFP fusion. Here we describe a FSEC screening strategy based on a fluorescent multivalent NTA probe that interacts with polyhistidine-tags on target proteins. This method overcomes the limitations of GFP fusion proteins, and can be used to rank protein production based on qualitative and quantitative parameters. Domain boundaries of the human G-protein coupled adenosine A2a receptor were readily identified from crude detergent-extracts of a library of construct variants transiently produced in suspension-adapted HEK293-6E cells. Well expressing clones of MraY, an important bacterial infection target, could be identified from a library of 24 orthologs. This probe provides a highly sensitive tool to detect target proteins to expression levels down to 0.02 mg/L in crude lysate, and requires minimal amounts of cell culture.

Keywords: membrane proteins; FSEC; protein expression and purification; construct design

Introduction

Structure-based drug discovery requires the overproduction and isolation of recombinant protein for various biophysical applications. Although validated drug targets are selected based upon their correlation to human health and disease, many are difficult to produce and isolate in a functional form. This is

the case for membrane proteins, such as G-protein coupled receptors (GPCRs), which are of intense pharmaceutical interest as they regulate essential signal transduction pathways that control diverse cellular responses. Disrupting these pathways can be correlated with disease states ranging from diabetes, heart failure and cancer. Unfortunately, these, and many other membrane proteins, can only be recombinantly produced at minimal levels and have limited stability when extracted from the membrane environment. Consequently, obtaining sufficient quantities of high quality purified protein for target

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: A. D. Ferguson, Discovery Sciences, Structure and Biophysics, AstraZeneca Pharmaceuticals, Waltham, MA. E-mail: andrew.ferguson@astrazeneca.com

Figure 1. Chemical structure of the peptide-based fluorescent FSEC probe P3NTA.

validation, high-throughput screening and structural biology remains challenging.

Successful strategies to overcome protein production issues involve screening large numbers of protein variants, mutants or functional orthologs, to identify constructs with improved expression levels and stability.^{1–3} Efficient screening of multiple constructs, often under varying buffer conditions, requires a rapid and highly sensitive platform with minimal sample handling. Fluorescence-based protein detection methodologies offer the required level of sensitivity and specificity to detect protein production in crude cell preparations, including clarified cell lysates, solubilized membranes or secretion media. Fluorescence-detection size-exclusion chromatography (FSEC) can be used to rank protein constructs based on stability, monodispersity and yield.⁴ FSEC has been successful as a tool to select optimal protein constructs for the determination of crystallographic structures of several novel classes of membrane proteins.^{5–7}

The most widely applied method for performing FSEC analyses on recombinant proteins involves the use of green fluorescent protein (GFP) fusions or fragments thereof.^{4,8,9} However, there are several important disadvantages associated with using GFP as a reporter of protein production and stability. First, false-positives can be observed as GFP is a highly soluble 240 residue protein.¹⁰ Second, difficulties in the proteolytic removal of GFP from the fusion protein are common,¹¹ as are protein aggregation issues following cleavage.¹² These issues often preclude downstream applications that require the removal of the GFP fusion partner. Third, significant reductions in yield are frequently observed when the GFP is removed from the target protein, often requiring re-optimization of overproduction conditions.^{9,11} Fourth, GFP fusion proteins can alter the function of the target protein, its cellular localization and processing.¹³ Finally, within the industrial setting, licensing expenses and restrictive applications often preclude the use of GFP related technologies.

In this study, we present the evaluation of a multivalent nitrilotriacetic acid (NTA) fluorescent

probe that permits FSEC analyses of polyhistidine-tagged membrane proteins in crude cell lysates. We also compare the utility of this probe to traditional GFP fusions and demonstrate that they are comparable in sensitivity, with the added advantages of eliminating false-positives, the need for subsequent subcloning, and the re-optimization of overproduction conditions. Moreover, this probe can be used to enhance the utility of established FSEC-based thermostability assays¹⁴ and thermofluor techniques to screen for optimal buffer and detergents for purified membrane proteins.¹⁵

Results

Peptide-based probe for FSEC

Covalent fusion of GFP to target proteins is the conventional method for high-throughput precrystallization screening of recombinant proteins by FSEC.^{4,8} In this study, we sought to completely eliminate the need to generate GFP fusion proteins by capitalizing on the ubiquitous use of polyhistidine-tags to purify recombinant proteins by immobilized metal-affinity chromatography. A peptide-based fluorescent probe, designated as P3NTA was evaluated (Fig. 1). The chemical synthesis of P3NTA involves solid-phase peptide synthesis to introduce a fluorescein label. The three adjacent thiol groups in the reduced, purified peptide were modified by a reaction with maleimido-C3-NTA. A short ten residue affinity-tag was added to the probe to facilitate its separation on Streptactin resin. The synthetic reaction was monitored by mass spectrometry.

In contrast to GFP fusions, the central design principle for the P3NTA probe was to avoid any potential impact that a covalently attached GFP fusion partner could have on the function, cellular localization and processing of target proteins, overproduction (artificially decreased or increased yield), and then eliminate the need for proteolytic removal of the GFP fusion partner following purification. We believe that because the P3NTA probe is added to crude cell preparations following protein overproduction and that it is noncovalently linked to the target protein, these features minimize false-positive rates that have been previously observed with GFP fusion proteins.

The physical interaction between the P3NTA probe and the target protein is driven by the same principles that are routinely used to purify and capture polyhistidine-tagged proteins using NTA-based resins and matrices. Given that the P3NTA probe has three NTA groups (Fig. 1), it is possible that the probe-to-target binding stoichiometry could be greater than one. However, previous studies have thoroughly characterized the interaction between hexahistidine and decahistidine-tagged proteins with spectroscopic probes containing two-to-four NTA moieties with a single covalently linked

Figure 2. Validating A2a receptor production by FSEC. Suspension-adapted HEK293-6E cells expressing an A2a (residues 1-316)-eGFP fusion protein was detergent solubilized, and analyzed by FSEC (blue). Addition of the P3NTA probe to saturation resulted in a two-fold increase in the fluorescence of the 2.6 mL elution volume peak, corresponding to the expected position of the fusion protein (red). The void volume of the column is 1.9 mL.

fluorescein label by isothermal titration calorimetry and analytical size exclusion chromatography.^{16,17} These studies demonstrate that the binding affinity and stability of the probe increases with the accessibility and the number of available NTA moieties, and that the binding stoichiometry is approximately 1:1 (probe-to-target) when working at saturating probe concentrations. We observed stable fluorescence signals for replicate samples that were analyzed across 16 hour automated HPLC sequences following a four hour incubation at 4°C with excess P3NTA probe prior to FSEC. As with related spectroscopic probes, we conclude that saturatable and constant binding with the P3NTA probe has been established.

Validating GPCR expression

The utility of the P3NTA probe for FSEC was evaluated using the human adenosine 2a receptor as a model membrane protein, which has well-defined production and purification protocols.^{18–21} In a series of experiments, we compared the use of the peptide-based P3NTA probe to detect the production of an A2a-eGFP fusion protein (residues 1-316 of the GPCR) (Fig. 2). Suspension-adapted HEK293-6E cells transiently expressing this A2a-eGFP fusion protein with a 10xHis-tag were analyzed by FSEC. Detergent-solubilized membranes were analyzed in the absence and presence of saturating concentrations of the P3NTA probe. In determining the amount of P3NTA probe that was needed for saturation, we found no evidence to suggest that this probe induced protein aggregation. The void volume of the column used in this study is 1.9 mL, and we

observed an equivalent small peak in the absence (blue trace) and presence (red trace) of the P3NTA probe containing detergent-solubilized protein aggregates (Fig. 2). Given that the fluorescence emission wavelengths of eGFP and the P3NTA probe overlap, the observed fluorescent peak is the sum of both components (Fig. 2). By subtracting the fluorescent signal originating from the eGFP fusion protein in the absence of P3NTA, the fluorescent signal from the probe can be calculated. As shown in Figure 2, the addition of the P3NTA probe at a saturating concentration of 1 μ M, results in a two-fold increase in fluorescence at an elution volume of 2.6 mL, corresponding to the expected elution volume for this A2a-eGFP fusion protein. We conclude that the fluorescence originating from the P3NTA probe is equivalent to that of the eGFP fusion, further demonstrating the utility of the P3NTA probe.

The addition of the P3NTA probe (molecular weight of less than 3 kDa) is not expected to impact the elution profile of the detergent-solubilized membranes containing the A2a-eGFP fusion protein (expected to be greater than 100 kDa), as the difference in molecular weight is not sufficient to alter the elution volume (Fig. 2). As with all integral membrane proteins, the elution volume, and therefore, observed molecular weight, is dependent upon the physical properties of the membrane protein in complex with the detergent micelle that is used for membrane solubilisation.^{22,23} For example, the protein-detergent complex of the neurotensin receptor NTS1 shows an observed molecular weight of 210 kDa by size exclusion chromatography, whereas MALS analysis demonstrates that molecular weight of the dimeric protein to be 77 kDa.²⁴

Finally, a broad elution peak containing the free P3NTA and the P3NTA probe mixed with cellular debris is also observed at an elution volume greater than 3.3 mL, and is not shown in this figure.

For difficult to produce membrane protein targets, it is common practice to screen large libraries of protein constructs to identify suitable variants with improved production and stability properties. Efficient parallel screening of these variants is reliant upon a high-throughput miniaturized plate-based screening platform to evaluate protein production levels and to perform biophysical characterization. We validated the utility of the P3NTA probe for analysis of protein levels, in a high-throughput manner, using a deletion library of this GPCR. The construct library was transiently expressed in mammalian cells using a plate-based 24-well platform for suspension-adapted HEK293-6E cells. This platform allowed us to complete the analysis of triplicate cultures of ten constructs plus controls within 4 days. FSEC analysis only required approximately 100,000 cells per culture, corresponding to about 100 μ L of suspension culture.

Figure 3. High-throughput FSEC screening of the A2a receptor. A: High-throughput FSEC screening of full-length and truncated variants of the human A2a receptor transiently expressed in HEK293-6E cells. The A2a receptor was detergent solubilized from whole cells using dodecylmaltoside, and analyzed by FSEC by incubating the clarified extract with the P3NTA probe. FSEC peak profiles for full-length and five C-terminal truncation variants are shown. Cells transfected with pDEST12.2 oriP containing non-tagged protein is included as the negative control. B: Western blot analysis of transient HEK293-6E cell expression of full-length and five C-terminal truncations. Two independent transfections were analyzed in adjacent lanes for each construct. Target protein was detected using an anti-His antibody that recognizes the C-terminal His-tag. Prestained protein markers are shown in the left lane, and an empty vector control (showing the detergent-solubilized HEK293-6E whole cell membrane background) is shown in the right lane.

Figure 3(A) shows the FSEC results derived from the P3NTA probe for full-length A2a, five C-terminal truncations and a nontagged negative control. Using triplicate cultures for each construct, FSEC peak profiles were consistent with %CV for the peak area below 20%. Replicate samples were analyzed using automated HPLC sequences such that equivalent detergent-solubilized membrane samples were analyzed at six hour time intervals, and our results showed no significant differences in fluorescence sig-

nal. These observations were also consistent with the associated negative controls reflecting signal background. Taken together, we conclude that these experiments were performed under conditions where P3NTA binding was both saturated and constant.

Relative elution times were found to correlate with the expected molecular weights for these variants [Fig. 3(A)], and the FSEC peak profiles are in agreement with the published literature.¹⁸ For example, the FSEC peak profiles for truncation 2 (residues 1–334) and truncation 3 (residues 1–316) show single monodisperse peaks [Fig. 3(A)]. Both constructs have previously been demonstrated to be stable in detergent, and truncation 3 with additional directed point mutations was used to determine crystallographic structures of this GPCR. Shorter variants, truncation 4 (residues 1–294) and truncation 5 (residues 1–279), showed no overproduction and had FSEC peak profiles that were identical to the nontagged negative control. Truncation 5 removes a portion of the seventh transmembrane helix, and would expectedly destabilize the GPCR and negatively impact the FSEC profile.

Although the boundaries for A2a constructs 2 (residues 1–334) and 3 (residue 1–316) are equivalent to those published previously,¹⁸ which have been demonstrated to be stable in detergent, the constructs used in this study are not precisely equivalent as they contain a N-terminal decahistidine-tag, a TEV cleavage site and a short four residue linker prior to the A2a coding sequence. Moreover, the literature construct that was used to determine the A2a crystal structure also contains a cluster of point mutations to facilitate crystallization. However, qualitative appearance of UV size exclusion chromatography profile for the detergent stable construct (residues 1–334) of A2a¹⁸ precisely matches the FSEC profile observed in this study using the P3NTA probe. This data demonstrate that when coupled with FSEC, the P3NTA probe can be used to identify appropriate construct boundaries using protein transiently expressed in suspension-adapted HEK293-6E cells with a plate-based miniaturized platform without the need for purification.

The overproduction levels of monodispersity A2a protein for these constructs were estimated by integrating the elution peak volume. Several assumptions were made in calculating these values: (1) calibrating the concentration of the free P3NTA probe under identical chromatographic conditions; (2) 1:1 binding stoichiometry between the P3NTA probe and the decahistidine-tag of the A2a protein; and (3) an average cell density of 1×10^9 cells/L. Using these measurements, the overproduction level of the full-length A2a fusion protein was estimated to be 0.7 mg/L. Constructs ending at residues V334 (0.4 mg/L), A316 (0.45 mg/L) or L267 (0.1 mg/L) had lower overproduction levels, whereas those ending at

residues T279 or E294 could not be detected. Taken together, we estimate the detection limit of the P3NTA probe to be better than 0.02 mg/L.

The advantages of FSEC are retained when used in conjunction with the P3NTA probe in ranking of protein constructs as demonstrated by Western blot analysis [Fig. 3(B)]. Not only does FSEC provide information on the quality of the protein produced (monodispersity) rather than the total protein yield, but it also facilitates the ranking of proteins with complex banding patterns when resolved using SDS-PAGE. For membrane proteins, oligomeric bands detected by SDS-PAGE can often indicate SDS-induced protein aggregation. However, as demonstrated with the human M2 muscarinic acetylcholine receptor,²⁵ the T4 lysozyme fusion construct used to determine the crystal structure of the receptor appeared to be aggregated in the presence of SDS under reducing conditions (accounting for higher molecular weight bands on the gel), whereas the receptor appeared to be monodisperse by size exclusion chromatography under nondenaturing conditions.

Although all truncation variants of A2a gave multiple strong bands on Western blots at the expected molecular weights, no detergent-solubilized protein was observed in the FSEC profiles for truncations 1–279 and 1–294 [Fig. 3(A)]. Moreover, none of these truncated A2a constructs produced monodisperse FSEC profiles. Further complexities arise in interpreting Western blotting data from the heterogeneous glycosylation patterns of recombinantly produced GPCRs, including the human A2a receptor. Although directed mutants, designed to eliminate specific glycosylation sites on the A2a receptor, displayed simplified banding patterns on Western blots, understanding the impact on protein yield and stability remains a significant challenge. The FSEC peak profiles of variants of truncation construct 3 (residues 1–316) with a single point mutation at residue N154 showed increased homogeneity, reduced molecular weights, and displayed small but consistent differences in the appearance of their FSEC peaks, and the FSEC profile of the N154D mutant can clearly be distinguished from the N154A, N154S, or N154Q mutants (Supporting Information Fig. S1). The accompanying Western blot is shown in Supporting Information Figure S2, which further emphasizes the complexity of interpreting Western blots data generated from crude whole cell lysates. Taken together these results further demonstrate the sensitivity and utility of the P3NTA probe when coupled to FSEC.

Bacterial ortholog screening

We next evaluated the utility of the P3NTA probe on a bacterial membrane protein target *MraY*, which is a member of the peptidoglycan biosynthetic pathway that catalyzes the translocation of a UDP-linked phospho-N-acetylmuramoyl-pentapeptide from the cytosol

to the cytosolic leaflet of the cytoplasmic membrane through the formation of a covalent linkage to membrane-embedded undecaprenyl phosphate.²⁶ Secondary structure predictions suggest that *MraY* consists of ten transmembrane helices that are connected by five cytoplasmic and four periplasmic loops.²⁷

The conceptual basis of ortholog screening is that primary sequence differences between functionally and structurally related membrane proteins can be used to identify an ortholog with improved recombinant production levels and stability for biophysical studies. A panel of twenty-four *MraY* orthologs, selected from a diverse dendrogram composed of the primary sequences of 110 *MraY* proteins, identified by a PSI-BLAST search using the primary sequence of *E. coli* *MraY*, was generated. A high-throughput plate-based screening platform for expression in *E. coli*, was used to evaluate *MraY* recombinant protein production levels.

Although the *E. coli* ortholog of *MraY* has proven difficult to overproduce and purify, the *B. subtilis* ortholog has been produced recombinantly for a variety of biophysical studies.^{27–29} This open reading frame was used to evaluate bacterial strains for protein production, growth media formulations, expression plasmids, and the position of the polyhistidine-tag prior to evaluating the 24 member *MraY* ortholog panel. Using the P3NTA probe, FSEC was used to rank the impact of these changes to the overproduction conditions. Optical density readings were used to normalize cell density prior to cell lysis for each condition tested. These results indicated that *E. coli* BL21 (DE3) cultured in Superbroth supplemented with glycerol was optimal for *MraY* overproduction from plasmid pET26b+ containing a cleavable C-terminal 10xHis-tag. For the 24 member ortholog panel, each open reading frame was then codon optimized for *E. coli*, and subcloned into this plasmid. The entire process, from bacterial transformation to FSEC analysis took 3 days, including triplicate clones and all relevant controls. As shown in Figure 4, only three *MraY* orthologs showed symmetrical FSEC peak profiles that were superior to the *E. coli* enzyme, they are *B. subtilis* (blue trace #25), *C. bolteae* (green trace #24) and *T. thermophilus* (red trace #23). The vast majority of the orthologs were either not overproduced or showed only minimal levels.

On the basis of these results, *MraY* from *C. bolteae* was overproduced at a preparative scale. This *MraY* ortholog could be readily purified in milligram quantities (0.3 mg/L) to greater than 90% purity (Supporting Information Fig. S3). The anomalous migration pattern of *MraY* from *C. bolteae* on SDS-PAGE is consistent with previously reported results for the *B. subtilis* ortholog.²⁸ Moreover, using a recently developed assay,³⁰ purified *MraY* from *C. bolteae* demonstrated the expected enzymatic activity (unpublished results). Taken together these results exemplify how the P3NTA probe can be used in

Figure 4. Ortholog screening of a bacterial membrane protein using FSEC. Orthologs were expressed in *E. coli*, and whole cell lysates were solubilized with dodecylmaltoside prior to addition of probe. Monodisperse protein elutes at 2.9 mL, whereas the free probe elutes at 3.9 mL. Cells transformed with empty expression vector were used as a negative control. To ensure that the full expression range was observed, detergent-solubilized samples were analyzed with increasing probe concentrations until a significant peak of free probe was observed at 3.9 mL. Samples: 1 - empty vector control; 2 - *C. diptheriae*; 3 - *C. testosteroni*; 4 - *C. trachomatis*; 5 - *E. coli*; 6 - *E. faecalis*; 7 - *F. nodosum*; 8 - *F. placidus*; 9 - *F. prausnitzii*; 10 - *H. neptunium*; 11 - *H. pylori*; 12 - *M. thermophila*; 13 - *M. voltae*; 14 - *P. gingivalis*; 15 - *P. horikoshii*; 16 - *R. rickettsii*; 17 - *S. pneumoniae*; 18 - *T. mathranii*; 19 - *T. neutrophilus*; 20 - *M. thermotrophicus*; 21 - *M. thermoacetica*; 22 - *L. gasicomitatum*; 23 - *T. thermophilus*; 24 - *C. bolteae*; and 25 - *B. subtilis*.

conjunction with FSEC to support the identification of well-behaving clones that can be overproduced at milligram levels without the need for subcloning.

Discussion

Fluorescent detection of recombinantly produced proteins can now be achieved with a small peptide-based probe that utilizes the high-affinity, specificity and reversibility of Ni-NTA interactions for polyhistidine stretches. In contrast to previous studies using multivalent probes that require chemical synthesis of complex Tris-NTA derivatives,^{16,31,32} the synthesis of the peptide-based FSEC probe P3NTA is straightforward and can be performed by any protein biochemistry laboratory. The equipment required to utilize the P3NTA probe is equivalent to that for traditional fluorescence-based protein detection methodologies, simply a chromatography system with an online fluorescence detector. Probe detection is compatible with versatile sample and running buffers, samples containing high salt, low concentrations of Mg^{2+} , TCEP, detergents and lipids can be successfully analyzed. As probe binding is based on Ni-NTA interactions with the polyhistidine-tag, restrictions on sample and buffer components are similar with immobilized metal-affinity chromatography (i.e. diva-

lent cations, chelating and reducing agents, and imidazole may interfere with probe detection).

For FSEC analysis, the performance of this probe is comparable to conventional GFP fusion proteins. Targets with low production levels, monodisperse protein of 0.02 mg/L and above, can be detected from crude cell preparations with fluorescence signals on the same order of magnitude as traditional GFP fluorescence. The main advantage of this probe over GFP fusion proteins is that it offers generic detection of recombinant proteins with polyhistidine-tags, completely eliminating the need for subcloning to remove the GFP fusion and re-optimization of overproduction conditions. This advantage reduces the risk of altering construct expression, solubility or activity following the removal of the GFP fusion, which improves the validity of production screening results.

Methods

Synthesis of the P3NTA probe

A peptide of sequence FAM-Ahx-SAWSHPQFEKCCC was synthesized, HPLC purified and supplied as lyophilized powder. 1 mg of synthetic peptide, was re-suspended in pH 7.4 PBS buffer containing 2 mM TCEP, and incubated at room temperature for 2 h. Ten milligrams of maleimido-C3-NTA was added, and the mixture was incubated in the dark for 24 h. The reaction mixture was applied to a 5 mL gravity flow column containing Streptactin agarose resin equilibrated in 20 mM Hepes (pH 7.5) and 200 mM NaCl. Uncoupled maleimido-C3-NTA reagent was removed by washing with two column volumes of 20 mM Hepes (pH 7.5) and 200 mM NaCl. The NTA groups were charged with nickel by washing with two column volumes of 20 mM Hepes (pH 7.5), 200 mM NaCl, and 10 mM $NiCl_2$, and excess nickel was removed by washing with an additional two column volumes of 20 mM Hepes (pH 7.5) and 200 mM NaCl. The completed P3NTA probe was eluted from the column with 2.5 mM desthiobiotin, and coupling efficiency was determined by electrospray ionization mass spectrometry. The probe was stored in small aliquots at $-80^{\circ}C$.

Cloning and expression of A2a mutants

Full-length human A2a, five C-terminal truncations variants and two enhanced GFP³³ fusion variants of construct 3 (N- and C-terminal) were subcloned into the pDEST12.2 oriP vector. All constructs share a common N-terminal sequence: METDTLLLWVLLLWVPGSTGDAPGHHHHHHHHHHENLYFQSGSGS containing an IgK light chain signal sequence, a 10xHis-tag, a TEV protease cleavage site and a short GSGS linker followed by the target coding sequence. All DNA was prepared using a HiSpeed Plasmid Maxi kit.

For mammalian expression, suspension-adapted HEK293-6E cells expressing EBNA-1 were

transfected at a cell density of $1.0 - 1.2 \times 10^6$ cells/mL in F17 media supplemented with 4 mM L-glutamine, 0.1% pluronic F-68 and 25 µg/mL geneticin. Transfections were performed in 24-deep well plates with each well containing 200 µL of F17 media, 4 µg of plasmid DNA, and 10 µg of PEI Max diluted with F17 media. Plates were gently agitated and incubated for 15 min at room temperature, prior to the addition of 3.6 mL of the transfection-ready HEK293-6E cell suspension to each well. Plates were sealed with a vented Capmat and incubated at 37°C, 5% CO₂ and agitated at 250 rpm. Constructs were set up in triplicate within the plate along with a negative control comprising the pDEST12.2 oriP vector with an insert lacking a 10xHis-tag. Cells were harvested 48 h post-transfection, cell densities were determined using a Cedex cell counter, and cell pellets were stored at -80°C.

Cloning and expression of *MraY* orthologs

Twenty-four orthologs of *MraY* covering diverse phylogenetic groups were selected. Synthetic genes codon optimized for *E. coli* were subcloned into the NcoI/XhoI cloning site of vector pET26b+, maintaining the *pelB* leader sequence. Coding sequences were engineered with the following N- and C-terminal sequences: TSRDHMVLHEYVNAAGITGGSGGSENLYFQS and GSGSENLYFQSHHHHHHHHHH. Both extensions contain a short GSGS linker and a TEV protease cleavage site, and the C-terminal extension includes a 10xHis-tag.

Protein production screening of the ortholog set was carried out in *E. coli* strain BL21 (DE3). 1 mL cultures of fresh transformants were grown in triplicate at 37°C in LB media supplemented with 20 mM D-glucose and 100 µg/mL kanamycin, in 24-deep well plates at 250 rpm using an orbital shaking incubator. Starter cultures were used to inoculate 3 mL cultures of Superbroth supplemented with 0.5% glycerol and 100 µg/mL kanamycin. When the OD₆₀₀ reached 0.6, the temperature was reduced to 19°C, protein expression was induced by 0.1 mM IPTG and cultures were grown overnight. Cells were harvested the following morning and cell pellets were stored at -80°C.

Purification of *MraY*

Totally, 75 mL of Luria Broth was inoculated with *E. coli* BL21(DE3) containing *MraY* from *C. bolteae*, and incubated overnight at 37°C. The overnight culture was used to inoculate 7 L of Super Broth supplemented with 5% glycerol, and incubated at 30°C. When the OD₆₀₀ reached 0.5, the temperature was reduced to 19°C. Protein expression was induced by 0.1 mM IPTG, and grown for 16 hrs with vigorous shaking. All subsequent steps were performed at 4°C. Cells were harvested and re-suspended in Buffer A [25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM β-mercaptoethanol, 1 mM MgCl₂, Roche EDTA-free protease tablet and 10% glycerol] by adding 30 mL of Buffer A for each 5 g of cell paste. Cells

were lysed by French Press and clarified by centrifugation. Membranes were pelleted from the clarified cell lysate, washed by re-suspending the membranes in Buffer A, and pelleted again. Purified membranes were solubilized in 20 mL of Buffer A supplemented with 1% dodecylmaltoside and 5 mM imidazole using a glass homogenizer, incubated for 30 min with gentle rolling, centrifuged, and the detergent-soluble fraction was stored overnight at 4°C. The detergent-soluble fraction was mixed with 3 mL of Ni-NTA resin equilibrated in Buffer A, gently rolled for 1 hr, and poured into an empty gravity flow column. The column was washed using three 50 mL volumes of Buffer A supplemented with increasing imidazole concentrations (20, 40 and 80 mM), a 15 mL equilibration with Buffer A only, and eluted using 300 mM imidazole.

Western blotting

Filtered solubilized samples were prepared in reduced LDS sample buffer and separated on 4 - 12% NuPAGE Novex Bis-Tris gels. PVDF membranes were blocked with PBS containing 1% BSA. A primary anti-His6 antibody was diluted 1:15,000 in blocking buffer, and the secondary anti-mouse AP conjugate was diluted at 1:7500 in blocking buffer. A one-step NBT/BCIP solution was used to develop the blots.

FSEC screening

Analysis of samples containing His-tagged proteins was performed following the addition of the P3NTA probe. Harvested cells were lysed and solubilized directly in detergent-containing buffer using 1% dodecylmaltoside. Membranes were solubilized by gentle mixing at 4°C, and clarified by centrifugation at 100,000g for 1 h at 4°C or by filtration. FSEC samples were prepared by gently mixing 0.5 - 1.6 µM of the P3NTA probe with 100 µL sample aliquots, and arrayed into 96-well plates. Buffers that were compatible with the FSEC analysis included 50 mM HEPES, sodium phosphate or TRIS within the pH range of 7.3-9.0. Detergent-containing buffers with at least 0.5M NaCl, 1 mM TCEP and/or 10% glycerol had no effect on probe binding. Samples were incubated for four hours with the probe before analysis by size exclusion chromatography, which was run at a column temperature of 12°C.

An Agilent series 11 HPLC system equipped with online fluorescence detection was used. On this system, samples were resolved on a Tosoh TSKgel SW3000 column (4.6 × 300 mm). Similar results were obtained using a SUPELCO Discovery BIO GFC300 (4.6 × 300 mm) or Agilent SEC-3 column (4.6 × 300 mm). Maximum probe fluorescence was monitored using excitation and emission wavelengths of 482 and 520 nm, respectively. Chromatograms were processed using Agilent ChemStation software. For each FSEC run, 10-20 µL samples were resolved using appropriate buffers.

Acknowledgments

The authors thank C. Koth from Genentech for kindly sharing the original synthesis of the P3NTA probes for evaluation by AstraZeneca. They also thank Adam Shapiro for his expert analysis of enzymatic activity of MraY.

References

1. Savchenko A, Yee A, Khachatryan A, Skarina T, Evdokimova E, Pavlova M, Semesi A, Northey J, Beasley S, Lan N, Das R, Gerstein M, Arrowsmith CH, Edwards AM (2003) Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches. *Proteins*50: 392–399.
2. Surade S, Klein M, Stolt-Bergner PC, Muenke C, Roy A, Michel H (2006) Comparative analysis and expression space coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci*15:2178–2189.
3. Warne T, Serrano-Vega MJ, Tate CG, Schertler GF (2009) Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr Purif*65:204–213.
4. Kawate T, Gouaux E (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure*14: 673–681.
5. Jasti J, Furukawa H, Gonzales EB, Gouaux E (2007) Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature*449:316–323.
6. Kawate T, Michel JC, Birdsong WT, Gouaux E (2009) Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. *Nature*460:592–598.
7. Sobolevsky AI, Rosconi MP, Gouaux E (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature*462:745–756.
8. Drew D, Slotboom DJ, Friso G, Reda T, Genevaux P, Rapp M, Meindl-Beinker NM, Lambert W, Lerch M, Daley DO, Van Wijk KJ, Hirst J, Kunji E, de Gier JW (2005) A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci*14:2011–2017.
9. Drew D, Lerch M, Kunji E, Slotboom DJ, de Gier JW (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods*3:303–313.
10. Drew DE, von Heijne G, Nordlund P, de Gier JW (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Lett*507:220–224.
11. Hsieh JM, Besserer GM, Madej MG, Bui HQ, Kwon S, Abramson J (2010) Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. *Protein Sci*19:868–880.
12. Drew D, Newstead S, Sonoda Y, Kim H, von Heijne G, Iwata S (2008) GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in *Saccharomyces cerevisiae*. *Nat Protocols*3:784–798.
13. Thomas CL, Maule AJ (2000) Limitations on the use of fused green fluorescent protein to investigate structure-function relationships for the cauliflower mosaic virus movement protein. *Gen Virol*81: 851–855.
14. Hattori M, Hibbs RE, Gouaux E (2012) A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure*20:1293–1299.
15. Fan J, Heng J, Dai S, Shaw N, Zhou B, Huang B, He Z, Wang Y, Jiang T, Li X, Liu Z, Wang X, Zhang XC (2011) An efficient strategy for high throughput screening of recombinant integral membrane protein expression and stability. *Protein Expr Purif*78:6–13.
16. Lata S, Reichel A, Tampé R, Piehler J (2005) High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J Am Chem Soc*127: 10205–10215.
17. Lata S, Gavutis M, Tampé R, Piehler J (2006) Specific and stable fluorescence labeling of histidine-tagged proteins for dissecting multi-protein complex formation. *J Am Chem Soc*128:2365–2372.
18. Singh S, Hedley D, Kara E, Gras A, Iwata S, Ruprecht J, Strange PG, Byrne B (2010) A purified C-terminally truncated human adenosine A(2A) receptor construct is functionally stable and degradation resistant. *Protein Expr Purif*74:80–87.
19. Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature*474:521–525.
20. Doré AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, Hurrell E, Bennett K, Congreve M, Magnani F, Tate CG, Weir M, Marshall FH (2011) Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure*19:1283–1293.
21. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC (2011) Structure of an agonist-bound human A2A adenosine receptor. *Science*332:322–327.
22. Möller JV, le Maire M (1993) Detergent binding as a measure of hydrophobic surface area of integral membrane proteins. *J Biol Chem*268:18659–18672.
23. Kunji ER, Harding M, Butler PJ, Akamine P (2008) Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography. *Methods*46:62–72.
24. White JF, Grodnitzky J, Louis JM, Trinh LB, Shiloach J, Gutierrez J, Northup JK, Grisshammer R (2007) Dimerization of the class A G-protein-coupled neurotensin receptor NTS1 alters G protein interaction. *Proc Natl Acad Sci USA*104:121999–12204.
25. Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature*482:547–551.
26. Al-Dabbagh B, Henry X, El Ghachi M, Auger G, Blanot D, Parquet C, Mengin-Lecreulx D, Bouhss A (2008) Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. *Biochemistry*47: 8919–8928.
27. Bouhss A, Mengin-Lecreulx D, Le Beller D, Van Heijenoort J (1999) Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol*34:576–585.
28. Bouhss A, Crouvoisier M, Blanot D, Mengin-Lecreulx D (2004) Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. *J Biol Chem*279: 29974–29980.

29. Ma Y, Münch D, Schneider T, Sahl HG, Bouhss A, Ghoshdastider U, Wang J, Dötsch V, Wang X, Bernhard F (2011) Preparative scale cell-free production and quality optimization of MraY homologues in different expression modes. *J Biol Chem*286: 38844–38853.
30. Shapiro AB, Jahić H, Gao N, Hajec L, Rivin O (2012) A high-throughput, homogeneous, fluorescence resonance energy transfer-based assay for phospho-N-acetylmuramoyl-pentapeptide translocase (MraY). *J Biomol Screen*17:662–672.
31. Huang Z, Park JI, Watson DS, Hwang P, Szoka FC (2006) Facile synthesis of multivalent nitrilotriacetic acid (NTA) and NTA conjugates for analytical and drug delivery applications. *Bioconjug Chem*17: 1592–1600.
32. Shiroishi M, Tsujimoto H, Makyio H, Asada H, Yurugi-Kobayashi T, Shimamura T, Murata T, Nomura N, Haga T, Iwata S, Kobayashi T (2012) Platform for the rapid construction and evaluation of GPCRs for crystallography in *Saccharomyces cerevisiae*. *Microb Cell Fact*11:8.
33. Pedélaq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*24:79–88.