FEBS Letters 507 (2001) 220–224 FEBS 25374

Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*

David E. Drew, Gunnar von Heijne, Pär Nordlund, Jan-Willem L. de Gier*

Department of Biochemistry and Biophysics, Arrhenius Laboratories, Stockholm University, S-106 91 Stockholm, Sweden

Received 10 May 2001; revised 24 September 2001; accepted 25 September 2001

First published online 11 October 2001

Edited by Matti Saraste†

Abstract ***Escherichia coli*** **is one of the most widely used vehicles to overexpress membrane proteins (MPs). Currently, it is not possible to predict if an overexpressed MP will end up in the cytoplasmic membrane or in inclusion bodies. Overexpression of MPs in the cytoplasmic membrane is strongly favoured to overexpression in inclusion bodies, since it is relatively easy to isolate MPs from membranes and usually impossible to isolate them from inclusion bodies. Here we show that green fluorescent protein (GFP), when fused to an overexpressed MP, can be used as an indicator to monitor membrane insertion versus inclusion body formation of overexpressed MPs in *E. coli*. Furthermore, we show that an overexpressed MP can be recovered from a MP–GFP fusion using a site specific protease. This makes GFP an excellent tool for large-scale MP target selection in structural genomics projects.**

Key words: Membrane protein; Green fluorescent protein; Protein overexpression; *Escherichia coli*

1. Introduction

With integral membrane proteins (MPs) accounting for 20–25% of all open reading frames in fully sequenced genomes, and their well-recognised importance as e.g. drug targets, they represent one of the most challenging classes of proteins in the areas of overexpression/purification and structural biochemistry [1,2]. The overexpression problems associated with MPs are a major bottleneck to overcome in studies of MP structure and function. *Escherichia coli* is one of the most widely used vehicles in attempting to overexpress both pro- and eukaryotic MPs. Overexpression of MPs in the cytoplasmic membrane is strongly favoured to overexpression in inclusion bodies, since it is relatively easy to isolate MPs from membranes, and usually impossible to isolate them from inclusion bodies. Unfortunately, it is not possible to predict whether an overexpressed MP will end up in the cytoplasmic membrane or in inclusion bodies.

Ideally, one would like to screen the expression of large numbers of MPs, but this is very time consuming using standard technology. For this reason we set out to develop an MP overexpression screen in *E. coli* using green fluorescent protein (GFP) as a folding indicator to distinguish between cytoplasmic membrane insertion versus inclusion body formation. GFP was chosen as it has been shown to be a sensitive folding indicator for the overexpression of globular proteins [3]. In short: when GFP is fused to the C-terminus of overexpressed globular proteins, then overexpression in a soluble form allows GFP to fold correctly and become fluorescent. However, if the globular protein-GFP fusion is overexpressed in inclusion bodies then GFP is not fluorescent.

In this report, we show that the use of GFP as a folding indicator can be expanded to encompass MP overexpression in *E. coli*; i.e. GFP can be used to distinguish between the overexpression of MPs in the cytoplasmic membrane and in inclusion bodies. Furthermore, we show that an overexpressed MP can be recovered from an MP–GFP fusion using a site specific protease. This makes GFP an excellent tool for large-scale MP target selection in structural genomics projects.

2. Materials and methods

2.1. Cloning and culture conditions

The genes coding for the *E. coli* MPs YidC [4], ProW [5], Lep-inv [6], the *E. coli* M13 bacteriophage MP procoat protein [7], the rat MP olfactory GPCR OR5 with an N-terminally fused glutathione-*S*-transferase (GST)-tag [8] and the human MP KDEL-receptor [9] were amplified by conventional PCR from plasmids available in house, and cloned into the C-terminal GFP fusion expression vector constructed by Waldo et al. [3]. This vector is a modified pET28(a+) vector, and it contains a GFP variant that is selected to fold well in *E. coli* and has the red-shifted mutation S65T and the folding mutation F64L [10]. The expression vectors harbouring the MP–GFP fusions were transformed freshly for each experiment into BL21(DE3)-pLysS and grown in 50 ml cultures on Luria broth (LB) medium containing 50 μg/ml kanamycin and 30 μg/ml chloroamphenicol. Cells were initially cultured at 37°C. When the cultures had reached an OD_{600} of 0.3–0.4, the temperature was switched to 25°C and MP–GFP fusion expression was induced for 10 h with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG).

2.2. Cell fractionation

After expression, cells were harvested and re-suspended in 1 ml of buffer containing 50 mM Tris–HCl pH 8.0, 200 mM NaCl, 15 mM EDTA and 100 μM PMSF (hereafter referred to as sonication buffer). Cells were sonicated on ice using a Xl-gies Microson (Heat Systems Incorporated) sonicator. Cells were sonicated at 50% duty cycle for four 45 s bursts with 1 min intervals. The sonication mixture was subsequently centrifuged for 10 min at 13 000 rpm in a microfuge to isolate any inclusion bodies and unbroken cells (hereafter referred to as the low speed spin pellet fraction). The supernatant of the low speed spin was centrifuged in a Beckman TL-100 table ultracentrifuge in a TLA100.2 rotor for 1 h at 80 000 rpm to isolate the cytoplasmic membranes (hereafter referred to as the high speed spin pellet frac-

*Corresponding author. Fax: (46)-8-153679.
E-mail address: degier@dbb.su.se (J.-W.L. de Gier).

Abbreviations: MP, membrane protein; GFP, green fluorescent protein; GST, glutathione-*S*-transferase; TEV, tobacco etch virus

0014-5793/01/$20.00
PII: S0014-5793(01)02980-5

tion) and the soluble fraction (hereafter referred to as the high speed spin supernatant fraction).

2.3. Fluorescence measurements

The low speed spin and the high speed spin pellet fractions were re-suspended in 1 ml of sonication buffer. For whole cell fluorescence measurements, 1 ml of culture was taken from the 50 ml cell cultures before sonication. Cells were harvested and subsequently re-suspended in 1 ml of sonication buffer minus PMSF. GFP emission of all these samples and the high speed spin supernatant fraction was measured using a Perkin-Elmer LS50B spectrofluorimeter with the excitation wavelength set at 480 nm and the emission wavelength set at 508 nm (the bandwidth for each was set at 5 nm), essentially as described by Waldo et al. [3]. For comparison, all the GFP fluorescence data were normalised by the maximum fluorescence signal.

2.4. Immunoblotting

5% of the low/high speed spin pellet fractions, and 1 ml whole-cell suspension, that had previously been used for fluorescence measurements, were re-centrifuged and re-suspended in 20 μl of sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) solubilisation (SB) buffer. The high speed spin supernatant was adjusted to the same v/v with SB buffer. 17.5 μl of the high speed spin supernatant and 10 μl of the other samples was used for separation in standard 12% SDS–polyacrylamide gels. Proteins were subsequently transferred from the gel to a PVDF membrane by means of Western blotting. Blots were decorated with either specific antibodies to the different MPs used in this study (in house antibody collection) or with a GFP specific antibody (Novagen). Blots were developed using the alkaline phosphatase system (according to the instructions of the manufacturer, Sigma).

Blots of the cell fractionation experiments were scanned using a standard flatbed scanner (MICROTEK ScanMakerX12USI), and individual band densities were measured and compared using the NIH imaging software (http://rsb.info.nih.gov/nih-image/) to a known amount of YidC-GFP standard to calculate the expected fluorescence for all MP–GFP fusions.

2.5. Construction, purification and proteolytic cleavage of YidC-TEV-GFP-His

Using linker insertion the DNA sequence encoding the amino acid sequence for the recombinant tobacco etch virus (TEV) protease cleavage site (protease recognition site: ENLYFQ/G;/: cleavage site) was introduced between the polylinker and the gene encoding GFP in Waldo's expression vector (see Section 2.1) [11]. Subsequently, the C-terminal stop codon in GFP in Waldo's expression vector was removed by site directed mutagenesis to include transcription of six histidines, yielding the expression vector pWaldo-TEV-GFP-His. The gene encoding YidC was subcloned into pWaldo-TEV-GFP-His. YidC-TEV-GFP-His was expressed in 1 l cultures, essentially as described in Section 2.1. Cells were spun down and re-suspended in 5 ml of sonication buffer. Subsequently, 50 mg of lysozyme, 10 mM $MgCl_2$ and some grains of DNAseI were added. Cells were lysed by repeated freeze thawing, followed by sonication. Membranes were isolated as described in Section 2.2. Membranes were washed in 1 ml of EDTA-free sonication buffer, re-centrifuged at 13 000 rpm for 20 min, and solubilised in 1 ml of 2% (w/v) *n*-dodecyl-β-D-maltoside (DDM), 10 mM Tris–HCl pH 8.0, 100 mM NaCl, 10 mM imidazole, and 20% (v/v) glycerol. The YidC-TEV-GFP-His fusion was purified essentially as described by van der Laan et al. [12].

The purified YidC-TEV-GFP-His fusion was incubated overnight at 30°C in the presence and absence of the recombinant TEV protease, as described in the manufacturer's instructions (Invitrogen). Fractions of the incubations were dissolved in SB buffer, and separated in 12% SDS–polyacrylamide gels. Proteins were subsequently transferred from the gel to a PVDF membrane by means of Western blotting. Blots were decorated with either an antibody to YidC or an antibody to GFP. Blots were developed as described in Section 2.4.

3. Results

3.1. MP–GFP fusions

A set of well-characterised pro- and eukaryotic MPs was

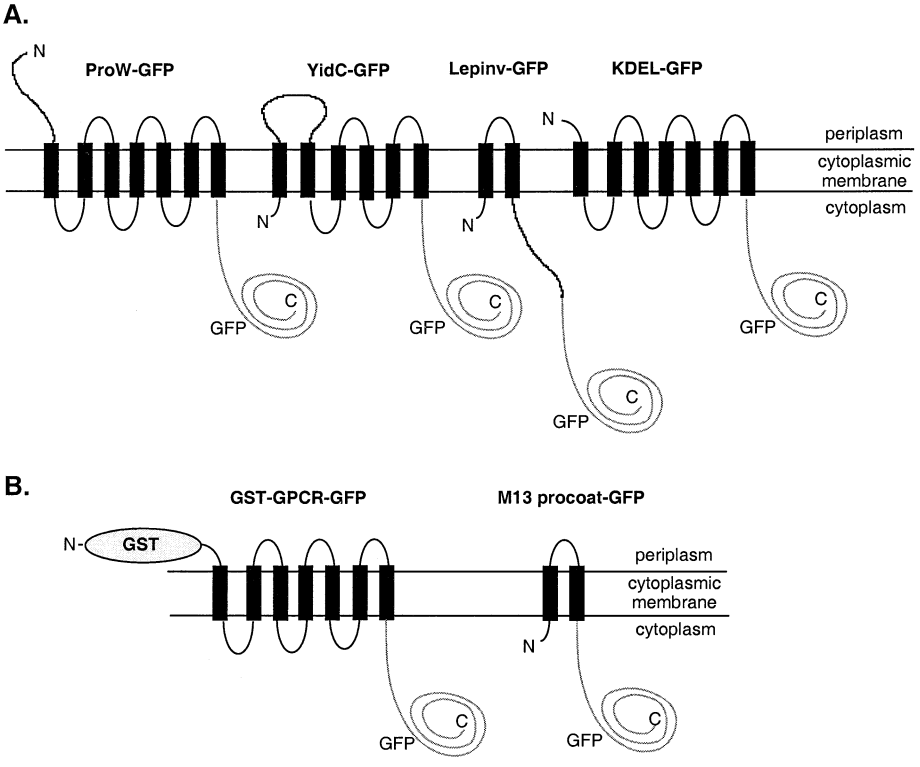

Fig. 1. Schematic representation of the MP–GFP fusions. A: MPs that express in *E. coli* predominantly in the cytoplasmic membrane. ProW, an *E. coli* MP that is part of the ProU osmoregulatory system [5]; YidC, an *E. coli* MP that is involved in MP assembly [4]; Lep-inv, a derivative of the *E. coli* MP leader peptidase that inserts with an inverted topology [6]; KDEL-receptor, human ERD2 receptor that is involved in the trafficking of proteins from the Golgi complex to the endoplasmic reticulum (ER) [17]. B: MPs that express in *E. coli* predominantly in inclusion bodies. M13 procoat, a bacteriophage coat protein [7]; GST-GPCR, a rat olfactory receptor with an N-terminal GST-tag [8].

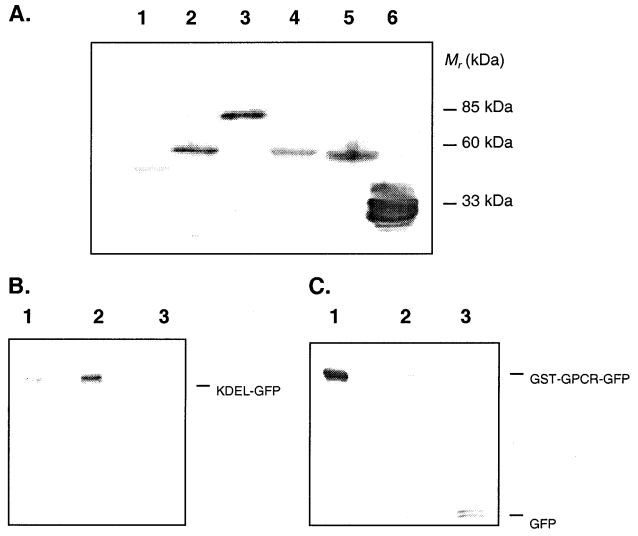

Fig. 2. MP–GFP expression. A: Western blots on whole cells after MP–GFP fusion overexpression. Gels were run and blots were developed using a GFP antibody as described in Section 2. Lane 1, KDEL-GFP; lane 2, GST-GPCR-GFP; lane 3, YidC-GFP; lane 4, Lep-inv-GFP; lane 5, ProW-GFP; and lane 6, M13-GFP. B: Western blots on subfractions of *E. coli* KDEL-GFP overexpressing cells. Gels were run and blots were developed using a GFP antibody as described in Section 2. Lane 1, low speed spin pellet fraction; lane 2, high speed spin pellet fraction; and lane 3, high speed spin supernatant fraction. C: Western blots on subfractions of *E. coli* GST-GPCR-GFP overexpressing cells. Gels were run and blots were developed using a GFP antibody as described in Section 2. Lane 1, low speed spin pellet fraction; lane 2, high speed spin pellet fraction; and lane 3, high speed spin supernatant fraction.

used to study the use of GFP as an MP folding indicator in *E. coli* (Fig. 1). All selected MPs were known to express well in *E. coli*, either predominantly in the cytoplasmic membrane (ProW, YidC, Lep-inv, KDEL-receptor), or predominantly in inclusion bodies (GST-GPCR, M13 procoat). As GFP is incorrectly folded as a C-terminal fusion that is translocated into the periplasm [13], we only selected MPs that have their C-termini located in the cytoplasm. Since at least 70% of all MPs are predicted to have this topology [14], the majority of MPs could be tested for overexpression in *E. coli* using this approach.

3.2. MP–GFP fusion expression analysis

All selected MPs were cloned into the C-terminal GFP fusion expression vector that has been constructed by Waldo et al. [3]. MP–GFP fusions were overexpressed in *E. coli* strain BL21(DE3)pLysS. After induction, expression of the MP–GFP fusions were detectable in whole cells by means of immunoblotting with MP specific antibodies (results not shown) and with a GFP antibody (Fig. 2A).

To monitor the expression status of the MP–GFP fusions (i.e. expression in the cytoplasmic membrane versus expression in inclusion bodies), MP–GFP fusion overexpressing cells were subfractionated. Cells were broken by means of sonication. The sonicated cells were subsequently subjected to a low speed centrifugation spin to collect unbroken cells/inclusion bodies (hereafter referred to as the low speed spin pellet). The supernatant of the low speed centrifugation spin was subsequently subjected to a high speed centrifugation spin to isolate the cytoplasmic membranes and the soluble fraction (hereafter referred to as the high speed spin pellet fraction and high speed spin supernatant fraction, respectively). The different subfractions were analysed by means of immunoblotting with a GFP antibody. An example of a subfractionation experiment, where most of MP–GFP (human KDEL-GFP) fusion appears to be in the cytoplasmic membrane is shown in Fig. 2B, and an example where most of MP–GFP (rat olfactory GST-GPCR-GFP) fusion appears to be in inclusion bodies is shown in Fig. 2C.

3.3. Comparing GFP emission from whole cells to isolated membranes

Breakage of *E. coli* cells by sonication is never complete (see e.g. [15]). This made it necessary to study if the GFP emission from the low speed spin pellet fraction was produced by unbroken cells rather than fluorescent inclusion bodies. Using a GFP antibody we quantified on Western blots the band intensity of the MP–GFP fusions present in equivalent amounts of the low speed spin pellet fraction and the high speed spin pellet fractions. Subsequently, the GFP emission from these fractions was estimated and plotted against the GFP emission measured (Fig. 3A and B). It was clear that for the MPs which are known to be expressed predominantly in inclusion bodies (M13 procoat protein, GST-GPCR), expected GFP emission present in the low speed spin pellet fraction did not correlate to the GFP emission measured. We estimated that 53% of M13-GFP and 86% of GST-GPCR-GFP fusion proteins formed inclusion bodies that did not contribute to any GFP emission. These inclusion bodies were isolated using a sucrose step gradient [16], and no GFP emission could be detected (results not shown).

Contrary, the GFP emission present in the low speed spin pellet fraction from MPs expected to be expressed predominantly in the cytoplasmic membrane (YidC, ProW, Lep-inv, KDEL-receptor), corresponded very nicely to observed GFP emission, indicating the presence of mostly unbroken cells rather than inclusion bodies in the low speed spin pellet fraction.

In the high speed spin pellet fractions, all MP–GFP fusions produced an estimated GFP emission that correlated well to the GFP emission measured. This indicated no bias when using GFP emission from isolated membranes to estimate expression levels. To re-confirm the amount of non-fluorescent inclusion bodies, the low speed spin fraction was adjusted to reach the same GFP emission levels measured in the high speed spin fraction and analysed by means of Western blotting using a GFP antibody (Fig. 3C).

In our set of MP–GFP fusions the amount of GFP emission measurable from isolated membranes compared well to the amount of MP expression in the cytoplasmic membrane. Because the GFP emission measurable in the low speed spin pellet fraction was produced by unbroken cells only, whole cell GFP emission was used to estimate MP expression levels in the cytoplasmic membrane (Fig. 3D).

3.4. MP recovery from an overexpressed MP–GFP fusion

We have explored the possibility to recover MPs from MP–GFP fusions using YidC-GFP as an example. In this respect it is worth mentioning that, so far, we have not observed any notable differences between the overexpression of MPs fused and not fused to GFP (our unpublished observations). To facilitate the purification of YidC-GFP, a 6His-tag was engi-

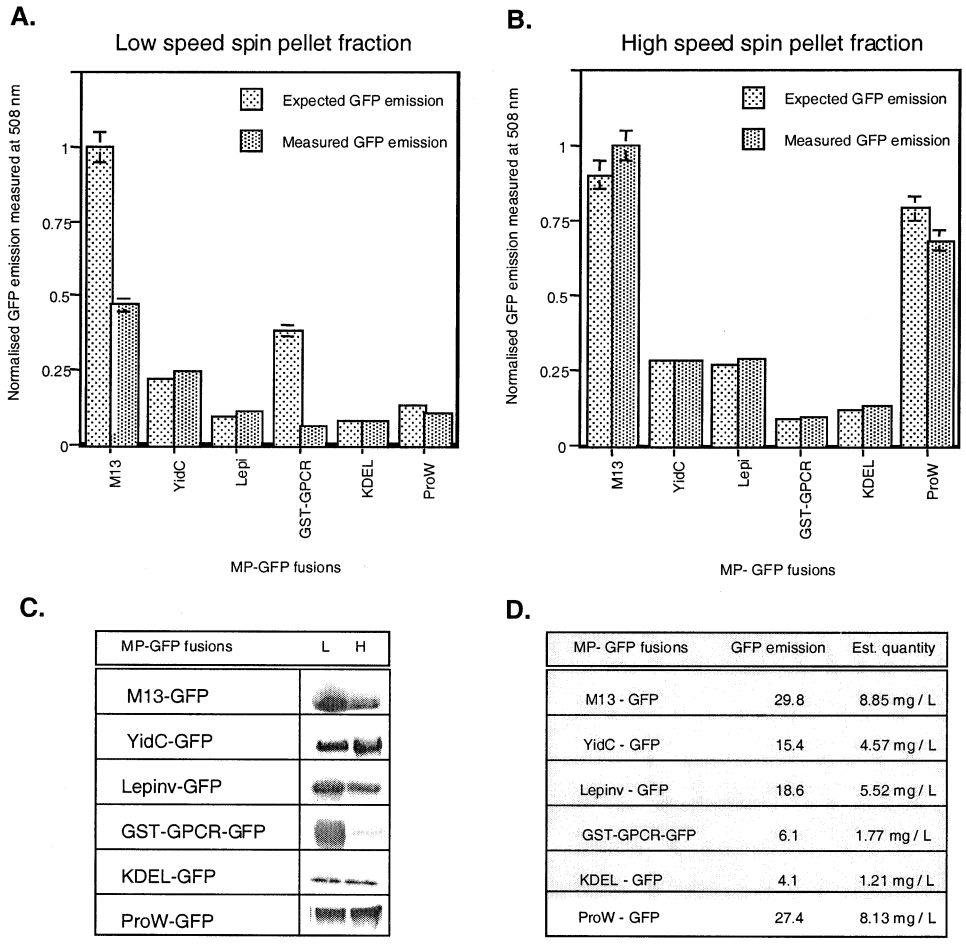

Fig. 3. GFP as an indicator to monitor MP overexpression. A: Comparison of the GFP emission estimated from the low speed spin pellet fraction by means of Western blotting using a GFP antibody to GFP emission measured by fluorescence spectrofluorimetry as described in Section 2. B: Comparison of the GFP emission estimated from the high speed spin pellet fraction by means of Western blotting using a GFP antibody to GFP emission measured by fluorescence spectrofluorimetry as described in Section 2. C: Western blots of MP–GFP low speed spin pellet (L) and high speed spin pellet (H), adjusted to the same GFP emission levels. Gels were run and blots were developed using a GFP antibody as described in Section 2. D: Estimated MP–GFP expression levels from whole cell GFP emission. As a standard purified GFP was used.

neered to the C-terminus of the fusion. To separate YidC and GFP, the site specific TEV protease cleavage site was introduced between YidC and GFP. The resultant YidC-TEV-GFP-His fusion could be purified using Ni^{2+}-NTA chromatography under similar conditions to YidC-His (Fig. 4A) [12]. Subsequently, at 50 mM imidazole purified YidC-TEV-GFP-His was incubated with the TEV protease, and protease digests were analysed by means of immunoblotting with antibodies to YidC and GFP (Fig. 4B). Strikingly, separation of YidC from the GFP moiety was almost complete.

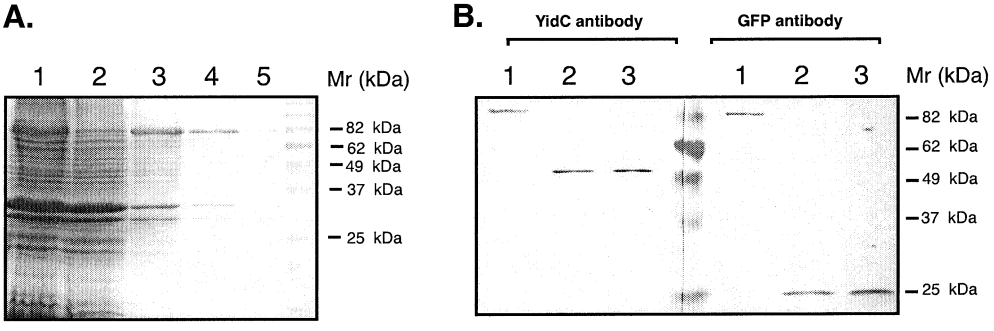

Fig. 4. MP recovery from an overexpressed MP–GFP fusion. A: Purification of YidC-TEV-GFP-His. Coomassie staining of SDS–PAGE gel representing purification of YidC-TEV-GFP-His by Ni^{2+}-NTA chromatography. Lane 1, solubilised inner membrane vesicles; lane 2, unbound material; lane 3, elution fraction in 30 mM imidazole; lane 4, elution fraction in 50 mM imidazole; and lane 5, elution fraction in 150 mM imidazole. B: Analysis of TEV protease digests. Western blots were decorated with antibodies to either YidC or GFP. Lane 1, YidC-TEV-GFP-His without TEV protease (control); lane 2, YidC-TEV-GFP-His with TEV protease; and lane 3, YidC-TEV-GFP-His with an excess of TEV protease.

4. Discussion

E. coli is one of the most widely used vehicles to overexpress both pro- and eukaryotic MPs. Overexpression of MPs in the *E. coli* cytoplasmic membrane is strongly favoured to overexpression in inclusion bodies, since it is relatively easy to isolate MPs from membranes, and usually impossible to isolate them from inclusion bodies. In this report, we show using a well-characterised set of MPs that GFP can be used as an indicator for MP overexpression in *E. coli*; i.e. GFP can be used to distinguish between overexpression of MPs in the cytoplasmic membrane or in inclusion bodies. With the set of model MPs used in this study, we show that when the overexpressed MP–GFP fusion is expressed as inclusion bodies, GFP is not fluorescent. Contrary, if the MP–GFP is inserted into the cytoplasmic membrane, GFP can fold properly and becomes fluorescent. In addition, we show that GFP is not only an indicator for the overexpression status of MPs, but can also be used to estimate the overexpression levels of MPs in the cytoplasmic membrane of whole cells.

Though we have not observed the formation of MP–GFP inclusion bodies that are fluorescent, it cannot be excluded that in some rare cases MP–GFP inclusion bodies are fluorescent. Another drawback with using GFP as an indicator to monitor MP overexpression, is that GFP may be clipped off from the MP–GFP fusions by proteases in vivo. In our study, we have observed only some proteolysis (see Fig. 2B), which is consistent with marginal or absent GFP emission from the high speed spin supernatant fraction (data not shown). This still means that there are potential risks to overestimate the amount of expression in the cytoplasmic membrane when GFP emission is measured from whole cells. Fortuitously, any gross overestimate of MP overexpression in the cytoplasmic membrane in whole cells can easily be filtered out by monitoring the GFP emission in isolated membranes. In conclusion, we have shown that GFP is an excellent indicator to monitor MP overexpression in *E. coli*. Preliminary results indicate that whole cell MP–GFP expression can easily be monitored on a microtitre plate dish format (our unpublished observations). This will significantly help in screening the expression of a large number of MPs, and to select the MPs that are overexpressed in the cytoplasmic membrane in sufficiently high amounts for functional and structural studies. Furthermore, we have shown that MP–GFP fusions can be used to recover overexpressed MPs. Therefore, GFP is an excellent tool for the large-scale MP target selection in structural genomics projects.

Acknowledgements: We gratefully thank Chen-ni Chin for critically reading the manuscript, and Edmund Kunji and Geoffrey Waldo for generously donating plasmids. This research was supported by the Swedish Foundation for Strategic Research (SSF).

References

[1] Wallin, E. and von Heijne, G. (1998) Protein Sci. 7, 1029–1038.
[2] Miroux, B. and Walker, J.E. (1996) J. Mol. Biol. 260, 289–298.
[3] Waldo, G.S., Standish, B.M., Berendzen, J. and Terwilliger, T.C. (1999) Nat. Biotechnol. 17, 691–695.
[4] Saaf, A., Monne, M., de Gier, J.W. and von Heijne, G. (1998) J. Biol. Chem. 273, 30415–30418.
[5] Csonka, L.N. and Hanson, A.D. (1991) Annu. Rev. Microbiol. 45, 569–606.
[6] von Heijne, G. (1989) Nature 341, 456–458.
[7] Wickner, W. (1988) Biochemistry 27, 1081–1086.
[8] Kiefer, H., Krieger, J., Olszewski, J.D., von Heijne, G., Prestwich, G.D. and Breer, H. (1996) Biochemistry 35, 16077–16084.
[9] Scheel, A.A. and Pelham, H.R. (1996) Biochemistry 35, 10203–10209.
[10] Crameri, A., Whitehorn, E.A., Tate, E. and Stemmer, W.P. (1996) Nat. Biotechnol. 14, 315–319.
[11] Parks, T.D., Howard, E.D., Wolpert, T.J., Arp, D.J. and Dougherty, W.G. (1995) Virology 210, 194–201.
[12] van der Laan, M., Houben, E.N., Nouwen, N., Luirink, J. and Driessen, A.J. (2001) EMBO Rep. 2, 519–523.
[13] Feilmeier, B.J., Iseminger, G., Schroeder, D., Webber, H. and Phillips, G.J. (2000) J. Bacteriol. 182, 4068–4076.
[14] Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) J. Mol. Biol. 305, 567–580.
[15] Hill, R.A. and Sillence, M.N. (1997) Protein Expr. Purif. 10, 162–167.
[16] Fiermonte, G., Walker, J.E. and Palmieri, F. (1993) Biochem. J. 294, 293–299.
[17] Townsley, F.M., Wilson, D.W. and Pelham, H.R. (1993) EMBO J. 12, 2821–2829.