High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*

Simon Newstead*, Hyun Kim†, Gunnar von Heijne†, So Iwata*‡§¶, and David Drew*†‡

*Division of Molecular Biosciences, Membrane Protein Crystallography Group, Imperial College, London SW7 2AZ, United Kingdom; †Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-10691 Stockholm, Sweden; §ERATO Human Receptor Crystallography Project, Kawasaki, Kanagawa 210-0855, Japan; and ¶RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

Edited by Bert van den Berg, University of Massachusetts Medical School, Worcester, MA, and accepted by the Editorial Board July 20, 2007 (received for review May 15, 2007)

Eukaryotic membrane proteins are often difficult to produce in large quantities, which is a significant obstacle for further structural and biochemical investigation. Based on the analysis of 43 eukaryotic membrane proteins, we present a cost-effective high-throughput approach for rapidly screening membrane proteins that can be overproduced to levels of >1 mg per liter in *Saccharomyces cerevisiae*. We find that 70% of the well expressed membrane proteins tested in this system are stable, targeted to the correct organelle, and monodisperse in either Fos-choline 12 (FC-12) or *n*-dodecyl-β-D-maltoside. We illustrate the advantage of such an approach, with the purification of monodisperse human and yeast nucleotide-sugar transporters to unprecedented levels. We estimate that our approach should be able to provide milligram quantities for at least one-quarter of all membrane proteins from both yeast and higher eukaryotic organisms.

GFP-based fusion technology | structural genomics

Structural and functional analysis of integral membrane proteins lags far behind that of soluble proteins. Although more structures of bacterial membrane proteins are emerging, little success has so far been achieved for eukaryotic membrane proteins. To date, only the rat voltage-gated potassium channel (1), spinach aquaporin (2), and human 5-lipoxygenase-activating protein (3) structures have been obtained with material isolated from heterologous overexpression.

One way to improve the throughput of eukaryotic membrane protein structures is to develop methods that reliably facilitate the identification of well expressing proteins that are stable and functional. In *Escherichia coli*, C-terminal tagging of membrane proteins with GFP facilitates this process, in part because GFP folds and becomes fluorescent only if the upstream membrane protein integrates into the membrane (4). The fluorescence resulting from overexpression is a fast and accurate measure of expression in the membrane and is easy to measure both in liquid culture and directly in standard SDS gels (5, 6). Once expression is optimized, the fluorescence from the GFP tag considerably speeds up detergent screening and purification. Further, the stability and monodispersity of a small amount of fusion protein in different detergents for precrystallization screening can be assayed by using fluorescence size-exclusion chromatography (FSEC) by coupling a fluorescent detector to a size-exclusion column (7).

Unfortunately, most eukaryotic membrane proteins are not functional when overexpressed in *E. coli*. Even if functional, subtle differences in, e.g., copurifying lipids may produce material unsuitable for crystallization. In part because of these limitations, only the human 5-lipoxygenase-activating protein structure to 4-Å resolution has been obtained by using material overexpressed in *E. coli* (3). Instead, other recent high-resolution eukaryotic membrane protein structures were based on material isolated from heterologous overexpression in the yeast *Pichia pastoris* (1, 2).

Here, we describe a GFP overexpression and purification scheme specifically tailored to *Saccharomyces cerevisiae*, an organism that offers several advantages over *P. pastoris* such as the option of rapidly cloning many genes into a 2μ plasmid by homologous recombination for direct expression testing (no need to make constructs in *E. coli* first), the availability of numerous different strains and expression plasmids, and the possibility to carry out functional complementation with mutant or deletion strains *in vivo* (8). For this last reason, *S. cerevisiae* is often the first eukaryotic host used to characterize the function of membrane proteins from higher eukaryotic organisms. Although *P. pastoris* has the advantage that it generates a large biomass during fermentation on methanol, poorly produced membrane proteins in *P. pastoris* are often difficult to purify, because they represent only a small fraction of total protein. By high-throughput expression screening many membrane proteins in *S. cerevisiae*, we can quickly identify those that are highly overexpressed. Importantly, several membrane protein-GFP fusions have been functionally produced to high levels in both *S. cerevisae* and *P. pastoris* (9, 10), and well diffracting crystals of the rabbit Ca^{2+} ATPase using material overexpressed in *S. cerevisiae* (11) has shown that success in structure determination from yeast expression systems is not limited to proteins produced in *P. pastoris*.

Results and Discussion

Construction of GFP-Fusion Expression Plasmids. Two vectors, one carrying the constitutive translation elongation factor promoter (pRS426TEF) (12) and the other the inducible GAL1 promoter (pRS426GAL1) (13), were used for subcloning of the yeast-enhanced GFP gene (14). In addition to GFP, both vectors were constructed to harbor a yeast codon-optimized tobacco etch virus protease site upstream of GFP for removal of the tag and a His_8 sequence downstream of GFP for immobilized metal affinity chromatography (IMAC) purification. GFP fusions can be further subcloned from these plasmids directly into the pPIC9K (Invitro-

Author contributions: S.N. and H.K. contributed equally to this work; D.D. designed research; S.N., H.K., and D.D. performed research; S.N., H.K., and D.D. contributed new reagents/analytic tools; S.N., H.K., G.v.H., S.I., and D.D. analyzed data; and S.N., H.K., G.v.H., S.I., and D.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.v.d.B. is a guest editor invited by the Editorial Board.

Abbreviations: DDM, *n*-dodecyl-β-D-maltoside; FC-12, Fos-choline 12; SEC, size-exclusion chromatography; FSEC, fluorescence SEC; IMAC, immobilized metal affinity chromatography; LDAO, *N,N*-dimethyldodecylamine-*N*-oxide; *H.s*, *Homo sapiens*; ER, endoplasmic reticulum.

‡To whom correspondence may be addressed. E-mail: s.iwata@imperial.ac.uk or d.drew@imperial.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704546104/DC1.

gen, Carlsbad, CA) vector for expression in *P. pastoris* (data not shown).

Selection of Target Proteins and Strains. Previously, we estimated overexpression levels for 553 cloned *S. cerevisiae* polytopic membrane proteins by whole-cell Western blot analysis (15). From this collection, 20 proteins with expression levels ranging from 77 to 500 arbitrary units (the maximal expression recorded in the collection was 1,000 arbitrary units) were selected for the current study, the majority being transporters of various kinds [supporting information (SI) Table 1, top half].

Two different yeast strains were chosen for expression of the test proteins, STY50 (16) and FGY217 (17), in which the vacuolar endopeptidase Pep4p is deleted. Deletion of the *pep4* gene not only inhibits Pep4p protease activity but also reduces the levels of other vacuolar hydrolases (18). The 20 test proteins were cloned directly by homologous recombination into the two vectors in the STY50 and FGY217 strains. Thus, a total of 80 overproduction tests were carried out initially.

Optimizing Overexpression Conditions. Because expression was typically lower than seen with *E. coli* membrane proteins (6), 10-ml cell cultures were used in the screening to ensure a reliable GFP signal-over-background fluorescence in whole-cell measurements. Because in most cases aeration improved expression by 50%, cultures were grown in aerated 10-ml capped tubes (data not shown). GFP fluorescence measured for 10-ml cultures gave a reliable estimate of expression, as shown for the scale-up to 1-liter cultures grown in 2.5-liter baffled flasks (SI Fig. 5*a*).

Yeast cells carrying the 20 GFP fusions under the constitutive TEF promoter were cultured in selection medium (SC-Ura) to stationary phase. For cells harboring GFP fusions under the inducible GAL1 promoter, cultures were initially grown overnight in SC-Ura medium with 2% glucose and back-diluted the next day in SC-Ura medium containing 0.1% glucose to an $OD_{600} = 0.1$. By monitoring the fluorescence of many constructs, we concluded it was best to induce cells with 2% galactose at an $OD_{600} = 0.6$ and to culture them for 24 h (at which point fluorescent levels no longer increase over time; SI Fig. 5*b*). Cells were harvested, washed, and resuspended in buffer containing 0.12 M sorbitol and 10% glycerol (important to maintain a uniform cell suspension), and the fluorescence was measured (SI Fig. 6).

In 19 of 20 cases, overexpression in whole cells was higher with the GAL1 promoter than with the TEF promoter. Of these 19 proteins, 15 had higher expression levels in the FGY217 *pep4* deletion strain, against 4 in STY50. Nevertheless, the most highly expressing protein with the TEF promoter (Dur3p) gave 1.5-fold higher expression levels compared with GAL1 induction in either strain. In five cases, the FGY217 strain with the GAL1 promoter was >2-fold better than the next best. To summarize, most but not all proteins expressed best under the GAL1 promoter, especially in the FGY217 strain.

Estimating Levels of Membrane-Integrated Protein from Whole-Cell Measurements. To estimate the amount of membrane-integrated fusion protein, cell lysates were subjected to SDS/PAGE and analyzed by in-gel fluorescence. Little or no full-length protein was detectable for the poorly expressed proteins (<4,000 relative fluorescent units; data not shown). The presence of "free" GFP after overexpression of membrane protein-GFP fusions in yeast may be a result of degradation of the membrane protein, whereas GFP remains intact (19). To remove free GFP and isolate membrane-integrated fusions, crude membranes were fractionated from whole cells. This was accomplished in a high-throughput format by the use of a 48-well tissue lyser for cell breakage and a desktop centrifuge for isolation of crude membranes.

GFP activity for whole cells vs. extracted membranes correlated well for all strain-promoter pairs ($R^2 = 0.85$–0.97), indicating that the estimates from small whole-cell cultures are reliable (SI Fig. 7 *i–iv*). Apart from the outlier Tpo4p, the correlation was also in agreement with the recovery from larger 2-liter cultures ($R^2 = 0.81$; SI Fig. 8*a*). Half of the whole-cell GFP activity was recovered in membranes. Given that cell breakage efficiency is 80%, we estimate ≈65% of GFP activity in whole cells is membrane-integrated material. Consequently, from the slope of a purified GFP standard curve, the GFP activity in whole cells was converted to milligram of membrane protein produced per liter (SI Table 1, top half).

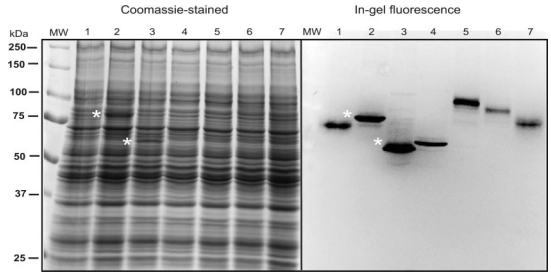

Fig. 1. Gel analysis of membrane protein-GFP fusions. Yeast membranes containing overexpressed fusions in strain FGY217 under the GAL1 promoter separated on a 10% SDS gel and analyzed by in-gel fluorescence (*Right*) and Coomassie staining (*Left*). Lane 1, Azr1p-GFP; lane 2, Isc1p-GFP; lane 3, Hsp30p-GFP; lane 4, Sec61p-GFP; lane 5, Mph3p-GFP; lane 6, Dal4p-GFP; and lane 7, Hxt1p-GFP. *, fusions that stained well with Coomassie.

To test whether the GAL1 promoter in combination with the FGY217 strain could overexpress membrane proteins from higher eukaryotic organisms, 14 metal and sugar transporters from *Arabidopsis thaliana*, *Homo sapiens* (*H.s*), *Mus musculus*, and *Caenorhabditis elegans* were amplified from cDNA, cloned, and overexpressed (for completeness, yeast homologs were also included; SI Table 1, lower half).

In all, 25 of the 29 yeast membrane proteins and 4 of the 14 membrane proteins from higher eukaryotic organisms were produced to levels >1 mg/liter.

Confirming the Integrity of Membrane-Integrated Fusions. To confirm the integrity of membrane-integrated fusions, samples were analyzed by in-gel fluorescence after standard SDS/PAGE (SI Fig. 9), as illustrated for seven proteins in Fig. 1 *Right*.

For six membrane proteins from higher eukaryotic organisms with poor levels of overexpression (≈0.1 mg/liter), no band was detectable. For the remaining samples, membrane-integrated material was detected as predominantly full-length GFP fusions. As is typical for integral membrane proteins, almost all migrated ≈10–20% faster than their predicted molecular mass (20), with the exception of Ctr1p that migrates more slowly than predicted because of *O*-linked glycosylation (21). These results indicate that, although fewer membrane proteins from higher eukaryotes express well, those that do are stable. When the gels were further stained with Coomassie blue, only ≈20% of the membrane protein-GFP fusions could be detected (data not shown). Although GFP-fusion intensities measured from in-gel fluorescence bands correlate well with fluorescent levels in yeast membrane fractions (SI Fig. 8*b*), the Coomassie staining levels matched in-gel fluorescent intensities poorly, as illustrated by differential staining of Isc1p and Hsp30p in Fig. 1 *Left*, lanes 2 and 3. Therefore, although GFP intensities from in-gel fluorescent bands remain consistent, Coomassie staining does not provide a reliable estimate for membrane protein expression, because staining is poor and inconsistent for many membrane protein samples. We estimate a 20-fold working range for the amount of membrane protein-GFP fusion that can be loaded on a gel for a linear in-gel fluorescent signal at a single exposure time (SI Fig. 8*c*).

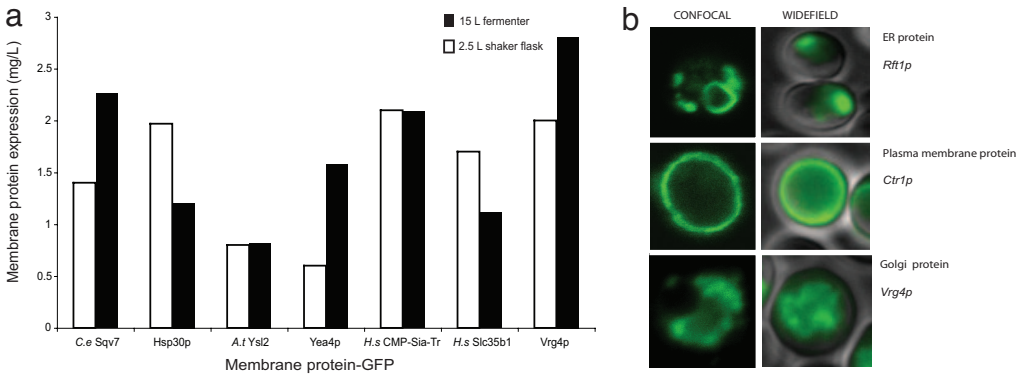

Fig. 2. Comparison of the upscale expression in 2.5-liter shaker flask vs. 15-liter fermenter and localization of fusion proteins by microscopy. (*a*) Bars represent overexpression levels of yeast and higher eukaryotic membrane protein fusions in strain FGY217 under GAL1 promoter. White bars, growth in 2.5-liter baffled shaker flasks; black bars, growth in 15-liter fermenter. (*b*) Confocal and widefield images of differently localized membrane proteins.

Optimizing Expression with Temperature and Chemical Chaperones. Because GFP fluorescence in whole cells reliably estimates membrane-integrated expression, it is straightforward to screen for optimal overexpression conditions. In previous literature, it has been reported that the addition of chemical chaperones such as DMSO (2.5% vol/vol), glycerol (10% vol/vol), or histidine (0.04 mg/ml) to yeast cultures can improve membrane protein folding (22, 23). To test these modifications, 20 fusions (12 yeast and 8 from higher eukaryotic organisms) were randomly selected, and the chemical chaperone was added together with galactose. Because a lower temperature can also increase yields, expression at 20°C for 24 and 36 h after GAL1 induction was also tested (SI Table 2).

In 15 of 20 cases, the modifications gave significantly better expression (≥2 SD = 0.2 mg/liter) than the standard culturing condition. DMSO or histidine addition was often the most effective, accounting for 11 of the 15 cases. The mean level of improvement was 30%, meaning that proteins with poor expression cannot be rescued in this way. Glycerol had a negative effect in most cases, as did culturing at 20°C for 24 h after induction. Culturing at 20°C for 36 h after GAL1 induction produced expression levels similar to standard conditions in most cases.

To test the scalability of the expression system, the overexpression of *C. elegans* Sqv7, Hsp30p, *A. thaliana* Ysl2, Yea4p, *H.s* CMP-Sia-Tr, *H.s* Slc35b1, and Vrg4p was compared in a 15-liter reaction vessel. On average, fermentation expression levels were consistent with expression levels from 2.5-liter cultures (Fig. 2*a*). Vrg4p overexpression levels in a 15-liter fermenter were also the same as that in a 50-liter fermenter (data not shown). Thus, the system is scalable to very large culture volumes.

Assessing the Quality of Overproduced Fusions by Confocal Microscopy. Although we can rapidly select well expressing membrane proteins and improve their expression levels using GFP fluorescence, it is difficult to assay the quality of the overexpressed proteins in the same way. Functional assays are the gold standard, yet, for transport proteins in particular, this is time consuming and not amenable to rapid parallelization. Moreover, for the majority of membrane proteins, functional assays are not at hand, and alternative strategies are needed to assess protein quality.

To address this issue, we took advantage of the fact that *S. cerevisiae* has a highly regulated quality control system in the endoplasmic reticulum (ER), ER-associated degradation, and only correctly folded proteins exit the ER. Misfolded proteins are either retained in the ER until they are folded correctly or dislocated from the ER and degraded by the proteasome in the cytosol (24). Proper subcellular targeting of overexpressed membrane proteins therefore provides a good indication of correct folding. Global fluorescence-microscopy localization studies on C-terminally GFP-tagged proteins have shown that for the majority of proteins, GFP does not interfere with trafficking (25). As a quality indicator, we therefore determined the localization of 25 fusions with different levels of expression (a cytoplasmic control protein was also included) after standard overexpression with the GAL1 promoter in the FGY217 strain.

Twenty-one of the 25 membrane proteins tested (from both yeast and higher eukaryotic organisms) were trafficked to their correct organelle (Fig. 2*b* and SI Fig. 10 *i–v*). Four plasma membrane proteins, Tpo4p, Hmn1p, Dur3p, and Pho87p, were found in vacuoles also, which suggests some degree of mistargeting, and has been described as such for Tpo4p and Dur3p previously (26). There was no obvious correlation between the level of overexpression and mistargeting. Indeed, a functional N-terminal deletion mutant of a copper transporter Ctr1pΔM1–8-GFP (27), which has been constructed to remove *O*-linked glycans for structural work, is overexpressed to ≈5% of total membrane protein and is found exclusively in the plasma membrane (SI Figs. 10*i* and 11*a*). Furthermore, this deletion mutant is also trimeric in *n*-dodecyl-β-D-maltoside (DDM) (SI Fig. 11 *b* and *c*), which is important given that homooligomerization of Ctr1 is essential for function (28).

To test our system with a membrane protein known to be nonfunctional as a C-terminal GFP fusion, the Golgi retrieval-receptor Rer1p was cloned and overexpressed. Rer1p-GFP is degraded upon overexpression in *S. cerevisiae*, because the C-terminal tail of Rer1p is essential for interaction with COP I subunits (29). Whole-cell fluorescent counts for Rer1p-GFP were very low [≈1,500 relative fluorescent units (RFU)], and the fusion could not be detected by in-gel fluorescence (SI Fig. 12*a*, lane 1). To confirm this was not because of difficulties in the overexpression of Rer1p *per se*, an N-terminal GFP fusion to Rer1p was constructed, which in contrast to Rer1p-GFP is functional (29). GFP-Rer1p was recoverable in membranes (≈21,000 RFU; SI Fig. 12*a*, lane 2), and monodisperse Rer1p in DDM could be purified from this fusion (SI Fig. 12*b*).

Assessing the Quality of Overproduced Fusions by FSEC. Recently, the solubilization efficiency of six detergents on ≈120 *S. cerevisiae* membrane proteins overexpressed in *S. cerevisiae* was determined by Western blotting (30). The zwitteronic detergents FC-12 and *N*,*N*-dimethyldodecylamine-*N*-oxide (LDAO) were proposed to be the most effective, followed by the nonionic detergent DDM.

We performed detergent extraction assays for 17 fusions for which the localization was previously monitored (SI Table 3). The median solubilization efficiency for FC-12 (75 ± 16%) was ≈20% higher than that for LDAO (57 ± 18%) or DDM (51 ± 20%) (Fig. 3*a*). No difference was observed between the detergent extraction efficiency of the yeast membrane proteins compared with those from higher eukaryotic organisms.

To determine whether detergent extraction efficiency was a good indicator of protein stability, we used FSEC. Recently, Kawate and Gouaux (7) showed that detergent-solubilized fusions that are monodisperse, as judged by FSEC, are typically stable after purification. This correlation removes the need to test each protein–

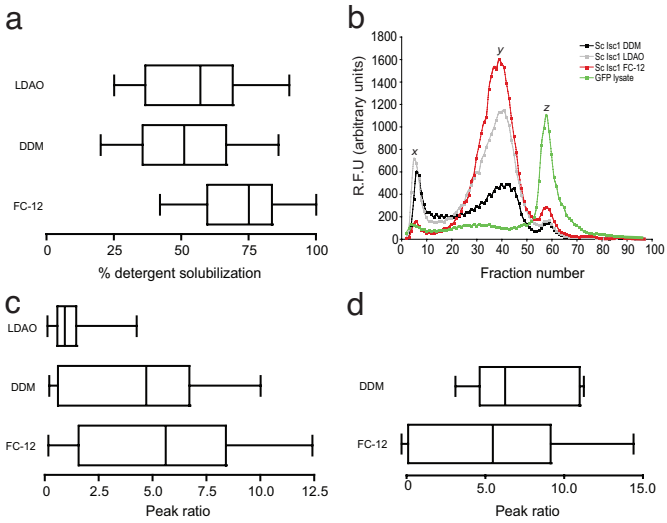

Fig. 3. Detergent extraction efficiency and FSEC profiles for membrane protein-GFP fusions in FC-12, DDM, and LDAO. Membrane protein-GFP fusions were assayed for detergent solubilization efficiency and monodispersity in FC-12, DDM, and LDAO. (*a*) Box-and-whiskers plots showing the distribution of the detergent solubilization efficiency. (*b*) Example of a typical FSEC profile, as shown here for Isc1p-GFP (for comparison, the GFP lysate is shown in green). Aggregation peak, *x*; membrane protein-GFP peak, *y*; and free GFP peak, *z*. Detergent FC-12, red; DDM, black; and LDAO, gray. (*c*) Box-and-whiskers plots showing the distribution of the protein-to-aggregation peak ratio (*y* divided by *x*). (*d*) Box-and-whiskers plots showing the distribution of the protein-to-aggregation peak ratio for fusions monodisperse in both DDM and FC-12 only.

detergent sample separately by purification. We injected FC-12, DDM, and LDAO solubilized samples onto a size-exclusion column equilibrated in buffer containing a low percentage of DDM (0.03%), which was found not to rescue aggregation of a protein solubilized in a different detergent (7). Instead of using an on-line fluorescence detector, 0.2-ml samples from the column were distributed on a 96-well plate for the fluorescence measurements. To improve the signal-to-noise ratio, GFP was excited at 470 nm instead of the commonly used 488 nm (SI Fig. 13). Our criterion for "monodispersity" was that the fusion peak was symmetric and equal to or larger than either the aggregate or free-GFP peaks.

In 13 of 17 cases, fusions were monodisperse in at least one detergent tested (example Fig. 3*b* and SI Fig. 14 *i–xvi*). For the four negative cases, Mrh1p and Hut1p were aggregated, the peak for Pho87p was diffuse, and Tpo4p had a free-GFP peak larger than the fusion peak itself. Interestingly, Tpo4p had already been picked up as an unstable protein, an outlier in our test set for comparing the average recovery of membrane-integrated material from whole cells (see SI Fig. 8*a*). In only four cases was a fusion monodisperse in all three detergents. FC-12 was more likely to give a monodisperse sample (13/17), followed by DDM (10/17) and LDAO (5/17) (SI Table 3). LDAO never gave a monodisperse profile for any of the samples that were not monodisperse in either FC-12 or DDM. FC-12 gave a monodisperse sample in three cases where the samples were aggregated in DDM. For the monodisperse samples, all apart from Dur3p were localized to the correct organelle.

To assign some relative "quality indicator" to the FSEC profiles, we calculated the ratio of the fusion peak height to aggregation peak height (Fig. 3*c*). The median ratio was highest for FC-12 (5.6), followed by DDM (4.7) and LDAO (0.9), reflecting the decreasing number of monodisperse samples in these detergents. However, if we compared membrane proteins that are monodisperse only in both FC-12 and DDM, the size of the upper quartile shifted clearly in favor of DDM (Fig. 3*d*).

How does the solubilization efficiency of a detergent correlate with the fraction of nonaggregated membrane protein? The membrane protein-to-aggregation peak ratio was plotted against the solubilization efficiency for each detergent (SI Fig. 15). For FC-12, there is no correlation. For both LDAO and DDM, there is a weak correlation, $R^2 = 0.35$. For proteins with solubilization efficiency ≈40% or less, the amount of aggregation was proportionally higher, such that the peak ratio was never >2 (this also applied to FC-12 in the two cases where the solubilization efficiency was ≈40%). Hence, by itself good detergent solubilization efficiency is a poor criterion for selecting which detergent to work with.

Purification of Eukaryotic Membrane Protein-GFP Fusions. We purified 10 eukaryotic membrane protein-GFP fusions by IMAC and reverse IMAC after cleavage of the GFP-His_8 tag (during overnight dialysis to remove imidazole) with His-tagged tobacco etch virus (TEV) protease. N-terminal sequencing of two human proteins, the CMP-sialic acid transporter and Slc35b1, confirmed there was no unspecific cleavage by the TEV protease (data not shown). As illustrated for the yeast GDP-mannose transporter Vrg4p, the putative human UDP-galactose transporter Slc35b1, and the human CMP-sialic acid transporter, this purification scheme yields pure protein (Fig. 4 *Insets*). Gel filtration profiles fit well with the profile for each fusion in the detergent solubilized membranes (Fig. 4 *Left*) and for the purified protein with or without the GFP-His_8 tag (Fig. 4 *Center* and *Right*). Average recovery of pure protein from membranes is ≈25%, that is, 13, 4, and 7 mg from a 15-liter fermenter, respectively. With the exception of refolding small amounts of a *M. musculus* CMP-sialic acid transporter from *E. coli* inclusion bodies (31), no transporter from the nucleotide-sugar transporter family has been purified to date.

Last, as a control, we attempted to purify Rft1p-GFP, which aggregated in DDM as shown by FSEC (SI Fig. 14*xii*). As anticipated, after IMAC-purified Rft1p-GFP was aggregated and could not be eluted from the column under standard conditions (results not shown).

Conclusions

GFP-fusion technology was originally designed for soluble proteins (32) but has also been used for rapid tests of expressibility, stability, and monodispersity of membrane proteins in *E. coli* (5–7). Here, using a test set of 43 eukaryotic membrane proteins tagged with GFP at the C terminus, we show that GFP fusions can also be used successfully to optimize membrane protein expression and purification in *S. cerevisiae*.

Initially, to find "optimal" expression conditions, 20 yeast membrane protein-GFP fusions were overproduced from two different promoters (the constitutive TEF and the inducible GAL1 promoters) and in two different strains. GFP fluorescence was measured for whole cells and isolated membrane fractions. Even though cleaved GFP products were detected in whole cells, a good correlation between the fluorescence measured in whole cells and membrane fractions means that expression screening in whole cells is reliable. We observed that most yeast membrane protein-GFP fusions expressed to higher levels under the GAL1 promoter than under the TEF promoter. Because the majority of membrane proteins when constitutively expressed retard cell growth (15), an inducible promoter such as GAL1 may be a better choice for overproduction in yeast. Moreover, most membrane protein-GFP fusions expressed better in the *pep4* deletion strain FGY217, suggesting that some overexpressed proteins may be susceptible to vacuolar proteolysis. Thus, we concluded that the combination of the GAL1 promoter in the FGY217 strain was the best choice for overproduction of membrane proteins and continued to screen the overexpression of proteins from higher eukaryotic organisms. The main difference between the overexpression of yeast membrane proteins compared with those from higher eukaryotic organisms was that only one-quarter of the latter could be overproduced to ≥1 mg/liter, compared with the majority of the yeast membrane

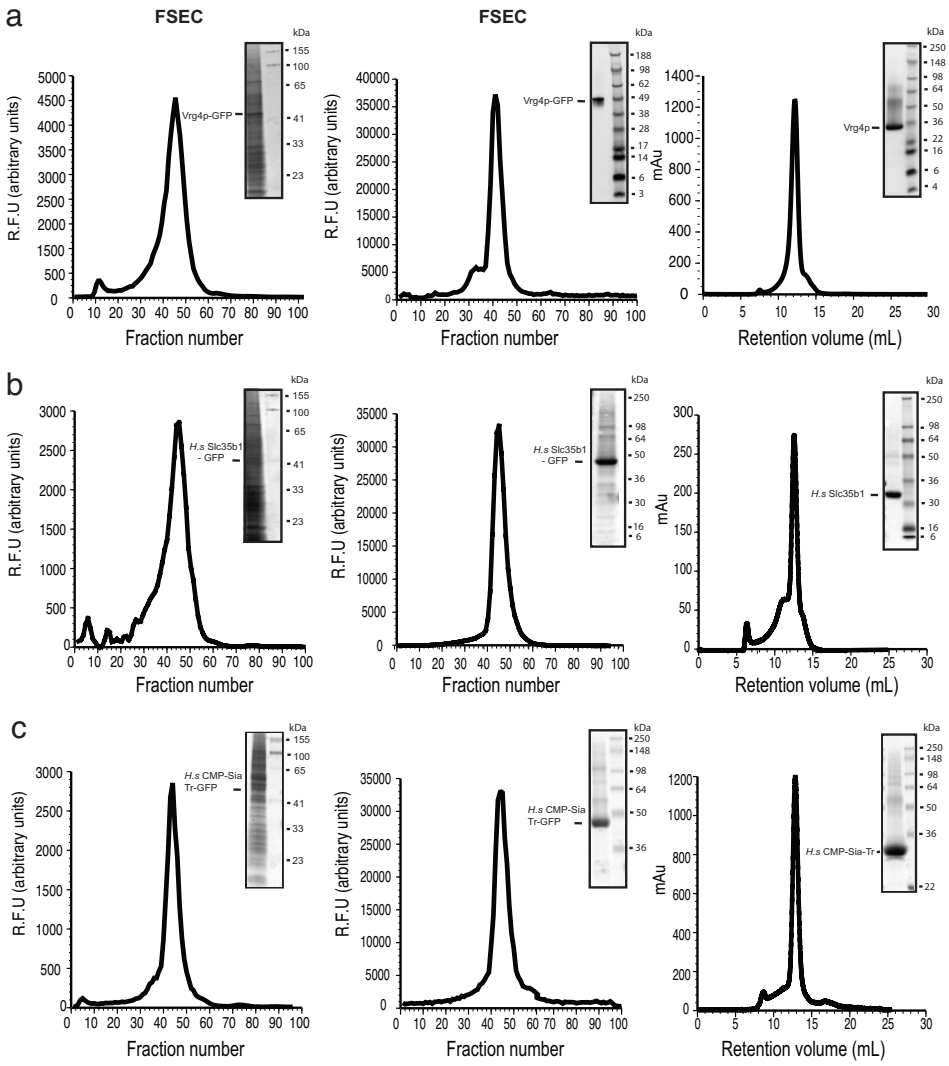

Fig. 4. Purification of nucleotide-sugar transporters Vrg4p, *H.s* Slc35b1 *H.s* CMP-Sia-Tr. Purification of Vrg4p (*a*), *H.s* Slc35b1 (*b*), and *H.s* CMP-Sia-Tr (*c*). (*Left*) FSEC profile for DDM-solubilized membranes. (*Center*) FSEC profile for Ni-NTA elution. (*Right*) SEC profile of membrane protein after removal of GFP-His_8 tag. The sample that was injected into either a Superose 6 or Superdex 200 gel-filtration column was analyzed by SDS/PAGE and is shown in the upper right corner of each chromatograph.

proteins in our test set. To confirm this estimate, we have since analyzed the expression and in-gel fluorescence levels of an additional 42 membrane proteins (8 ATP-binding cassette and 34 solute carrier transporters) from either human, mouse, or rat. Indeed, ≈25% of these proteins were expressed at ≥1 mg/liter (unpublished data). A recent study (33) comparing the overexpression of eukaryotic membrane proteins in *S. cerevisiae* to other expression hosts has reported that overexpression in this host is poor after 6-h induction with galactose; however, by systematic improvement of culturing conditions, we found that most protein is synthesized between 13 and 20 h after galactose addition to the medium.

Analyzing the quality of an overexpressed membrane protein is usually difficult to accomplish in a high-throughput manner. However, we could take advantage of the GFP fusion to confirm that the majority of the fusions tested were targeted to their correct organelle, which implies they are folded correctly. For a known nonfunctional C-terminal GFP fusion to Rer1p, no overexpression was detected. We could, however, purify monodisperse Rer1p from an N-terminal GFP-Rer1p fusion, which has been shown to be functional (29).

Because stability is important for functional studies and crystallization, we carried out FSEC on detergent-solubilized membranes to assess the monodispersity of fusions. We found that 14 of 17 fusions tested were monodisperse, and 13 of 14 of these samples localized to their correct organelle. This suggests there is some correlation between correct subcellular localization and monodispersity. In all, 70% of well expressed membrane proteins localize to their corresponding organelle and are monodisperse.

Last, we compared the level of membrane protein overexpression to the fraction of detergent-solubilized membrane protein and to the fraction of nonaggregated protein in the FSEC traces (data not shown) but found no correlation. These findings emphasize the necessity of rapid methods, like FSEC, to confirm that well expressed membrane proteins (especially those for which rapid functional assays are not available) are suitable for further studies. As shown in detail for three nucleotide–sugar transporters from yeast and human, SEC profiles obtained before and after removal of the GFP tag correlated well with each other.

Taken together, our results establish the GFP-fusion technology in *S. cerevisiae* as a cost-effective high-throughput approach to obtain milligram quantities of stable and monodisperse eukaryotic membrane proteins for functional and structural studies.

Materials and Methods

Yeast Genetic Manipulations. Yeast membrane proteins were amplified from genomic DNA isolated from W303-a (*MATa*, *ade2*, *can1*, *his3*, *leu2*, *trp1*, and *ura3*) and other eukaryotic membrane proteins from their respective cDNA (all genes are listed in *SI Text*). Primers contained a 12-bp gene-specific region and a 30-bp homologous region on the forward primer 5′-TCG ACG GAT TCT AGA ACT AGT GGA TCC CCC -3′ and reverse primer 5′-AAA TTG ACC TTG AAA ATA TAA ATT TTC CCC-3′. PCR

product and SmaI linearized pDDGFP-1 or -2 plasmid [the construction of the plasmids pDDGFP-1 (pRS426TEF-GFP) and pDDGFP-2 (pRS426GAL1-GFP) is detailed in *SI Text*] were transformed into the strains STY50 or FGY217 (*MATa*, *ura3–52*, *lys2Δ201*, and *pep4*Δ) (17). Transformants were selected on -Ura plates, and positive clones were initially confirmed by colony PCR and/or whole-cell GFP fluorescence.

Membrane Protein Overexpression. Ten-milliliter cell cultures containing either pDDGFP-1 or -2 plasmid were grown in -Ura medium shaken at 280 rpm in 50-ml aerated capped tubes (TPP, Trasadingen, Switzerland) at 30°C overnight. For constitutive expression under the TEF promoter, the overnight cultures were grown to stationary phase. For GAL1 induction, cultures were back diluted to an OD_{600} = 0.1 and grown in the presence of 0.1% glucose to an OD_{600} = 0.6. At OD_{600} = 0.6, final 2% of galactose was added to culture. After 24 h, cells were harvested by centrifugation at 3,200 × *g* for 5 min, and supernatant was removed by aspiration. Cell pellets were resuspended in 200 μl of buffer containing 50 mM Tris·HCl, pH 7.6, 50 mM EDTA, 10% glycerol, and protease inhibitor mix (Roche, Indianapolis, IN). Cells were transferred to a clear 96-well optical plate (Nunc, Rochester, NY), and GFP emission was measured at 512 nm on bottom read, with an excitation wavelength of 488 nm, using a SpectraMax M2e (Molecular Devices, Sunnyvale, CA). Cell suspension in plate was transferred to 2-ml screw-capped tubes containing 250 μl of acid-washed 425- to 600-μm glass beads (Sigma, St. Louis, MO). Cells were lysed in automated biological sample lyser (Qiagen, Chatsworth, CA) and spun at high speed for 5 s at 4°C. Supernatant was collected and transferred to a new 1.5-ml tube. Two hundred microliters of buffer was added to the remaining pellet, and the previous step was repeated. Combined supernatants were spun at 21,040 × *g* by using a desktop centrifuge for 1 h at 4°C to collect crude membranes. Crude membranes were resuspended in 200 μl of buffer and transferred to a 96-well plate for measuring GFP emission, and expression levels were calculated as detailed in *SI Text*. To check consistency of cell breakage, total protein concentration was measured by using the BCA assay (Pierce, Rockford, IL).

Optimization of overexpression and upscale of overexpression and subsequent isolation of membranes from larger cultures are described in detail in *SI Text*.

Analysis of Fusions by In-Gel Fluorescence. Crude membranes (≈20 to −30 μg of total protein) or protein samples were added 1:1 with solubilization buffer containing 50 mM DTT; 50 mM Tris·HCl, pH 7.6; 5% glycerol; 5% SDS; 5 mM EDTA; 0.02% bromophenol blue; and protease inhibitors. Samples were analyzed by SDS/PAGE with 12% Tris-Glycine gels (Invitrogen). For in-gel fluorescence, GFP protein bands were imaged by using a CCD camera after exposure to blue light at 460 nm with a 515-nm filter cutoff for 30 sec (LAS-3000; Fuji, Tokyo, Japan). For Coomassie staining, brilliant blue R-250 was used as stated by the manufacturer's instructions.

Widefield and Confocal Microscopy. We assessed the localization of the membrane protein-GFP fusions by observing an aliquot of the cell suspension under a Leica (Wetzlar, Germany) TCS SP2 upright confocal microscope. Rft1p, Vrg4p, and Ctr1p GFP fusions were also monitored with a Zeiss (Jena, Germany) Axiovert 200 inverted fluorescent microscope.

Detergent Screening. In each screen, the membranes were diluted in PBS to a final concentration of 3 $mg \cdot ml^{-1}$ total protein in a final volume of 1 ml and mixed with the detergents at the final concentrations given: 1% (wt/vol) DDM, 1% (wt/vol) FC-12, and 1% (wt/vol) LDAO (Anatrace, Maumee, OH) The mixtures were incubated for 1 h at 4°C on a blood wheel, and the solubilized proteins were separated from the insolubilized proteins by a 1-h centrifugation at 120,000 × *g* at 4°C. To determine the amount of solubilized membrane protein-GFP fusion, the fluorescent counts in 100 μl of the sample before and after centrifugation were measured as previously described.

FSEC Analysis of the Membrane Protein-GFP Fusions. Five hundred microliters of the solubilized membranes containing the overexpressed GFP fusions in the three detergents DDM, FC-12, and LDAO was loaded onto a Superose 6 10/300 column (GE Healthcare, Uppsala, Sweden) preequilibrated in 50 mM Tris·HCl, pH 7.50/0.2 M NaCl/0.03% DDM connected to an AKTA FPLC system at 4°C. The flow rate was set at 0.4 $ml \cdot min^{-1}$, and 200-μl fractions were collected from 6 ml into a 96-well black clear bottom plate (Nunc) by using a Frac-950 fraction collector (GE Healthcare). After separation of the sample into the a 96-well plate, the GFP emission in each fraction was measured at 512 nm on bottom read, with an excitation wavelength of 470 nm on a SpectraMax $M2^e$ (Molecular Devices).

Purification of the Membrane Protein-GFP Fusions. Purification of membrane protein-GFP fusions was similar to that described in ref. 5 (for detailed description, see *SI Text*).

We thank Yo Sonoda and James Mansfield for technical assistance and Dr. Masato Kawasaki for work with Rer1p. We gratefully thank Dr. Per Ljungdahl (Department of Cell Biology, The Wennergren Institute, Stockholm University) for advice and the kind gift of the FGY217 yeast strain and Dr. Jan-Willem de Gier for critically reading the manuscript. This research was funded by grants from the Swedish Research Council (to G.v.H. and H.K.), the U.K. Biotechnology and Biological Sciences Research Council Membrane Protein Structure Initiative (to S.I.), the EU-PF6 E-MEP European Membrane Protein Consortium (to S.I.), and the European Molecular Biology Organization long-term fellowship (to D.D.).

1. Long SB, Campbell EB, Mackinnon R (2005) *Science* 309:897–903.
2. Tornroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, Tajkhorshid E, Neutze R, Kjellbom P (2006) *Nature* 439:688–694.
3. Ferguson AD, McKeever BM, Xu S, Wisniewski D, Miller DK, Yamin TT, Spencer RH, Chu L, Ujjainwalla F, Cunningham BR, *et al.* (2007) *Science* 317:510–512.
4. Drew DE, von Heijne G, Nordlund P, de Gier JW (2001) *FEBS Lett* 507:220–224.
5. Drew D, Lerch M, Kunji E, Slotboom DJ, de Gier JW (2006) *Nat Methods* 3:303–313.
6. Drew D, Slotboom DJ, Friso G, Reda T, Genevaux P, Rapp M, Meindl-Beinker NM, Lambert W, Lerch M, Daley DO, *et al.* (2005) *Protein Sci* 14:2011–2017.
7. Kawate T, Gouaux E (2006) *Structure (London)* 14:673–681.
8. Ton VK, Rao R (2004) *Am J Physiol* 287:C580–C589.
9. Niebauer RT, Robinson AS (2006) *Protein Expr Purif* 46:204–211.
10. Sarramegna V, Talmont F, Seree de Roch M, Milon A, Demange P (2002) *J Biotechnol* 99:23–39.
11. Jidenko M, Nielsen RC, Sorensen TL, Moller JV, le Maire M, Nissen P, Jaxel C (2005) *Proc Natl Acad Sci USA* 102:11687–11691.
12. Mumberg D, Muller R, Funk M (1995) *Gene* 156:119–122.
13. Mumberg D, Muller R, Funk M (1994) *Nucleic Acids Res* 22:5767–5768.
14. Cormack BP, Bertram G, Egerton M, Gow NA, Falkow S, Brown AJ (1997) *Microbiology* 143:303–311.
15. Osterberg M, Kim H, Warringer J, Melen K, Blomberg A, von Heijne G (2006) *Proc Natl Acad Sci USA* 103:11148–11153.
16. Strahl-Bolsinger S, Scheinost A (1999) *J Biol Chem* 274:9068–9075.
17. Kota J, Gilstring CF, Ljungdahl PO (2007) *J Cell Biol* 176:617–628.
18. Woolford CA, Daniels LB, Park FJ, Jones EW, Van Arsdell JN, Innis MA (1986) *Mol Cell Biol* 6:2500–2510.
19. Lagerstedt JO, Zvyagilskaya R, Pratt JR, Pattison-Granberg J, Kruckeberg AL, Berden JA, Persson BL (2002) *FEBS Lett* 526:31–37.
20. Rais I, Karas M, Schagger H (2004) *Proteomics* 4:2567–2571.
21. Dancis A, Haile D, Yuan DS, Klausner RD (1994) *J Biol Chem* 269:25660–25667.
22. Figler RA, Omote H, Nakamoto RK, Al-Shawi MK (2000) *Arch Biochem Biophys* 376:34–46.
23. Andre N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T, Pattus F, Michel H, Wagner R, Reinhart C (2006) *Protein Sci* 15:1115–1126.
24. Ellgaard L, Helenius A (2003) *Nat Rev Mol Cell Biol* 4:181–191.
25. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) *Nature* 425:686–691.
26. Uemura T, Kashiwagi K, Igarashi K (2007) *J Biol Chem* 282:7733–7741.
27. Puig S, Lee J, Lau M, Thiele DJ (2002) *J Biol Chem* 277:26021–26030.
28. Nose Y, Rees EM, Thiele DJ (2006) *Trends Biochem Sci* 31:604–607.
29. Sato K, Sato M, Nakano A (2001) *J Cell Biol* 152:935–944.
30. White MA, Clark KM, Grayhack EJ, Dumont ME (2007) *J Mol Biol* 365:621–636.
31. Tiralongo J, Ashikov A, Routier F, Eckhardt M, Bakker H, Gerardy-Schahn R, von Itzstein M (2006) *Glycobiology* 16:73–81.
32. Waldo GS, Standish BM, Berendzen J, Terwilliger TC (1999) *Nat Biotechnol* 17:691–695.
33. Eifler N, Duckely M, Sumanovsky LT, Egan TM, Oksche A, Konopka JB, Luthi A, Engel A, Werten PJ (2007) *J Struct Biol* 159:179–193.