

Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host

Florent Bernaudat^{1,2,3,*}, Annie Frelet-Barrand^{4,5,6,7,8a,b}, Nathalie Pochon^{8,9,10}, Sébastien Dementin^{11,12,13,c}, Patrick Hivin^{14,d}, Sylvain Boutigny^{4,5,6,7}, Jean-Baptiste Rioux^{11,12,13}, Daniel Salvi^{4,5,6,7}, Daphné Seigneurin-Berny^{4,5,6,7}, Pierre Richaud^{15,16,17}, Jacques Joyard^{4,5,6,7}, David Pignol^{11,12,13}, Monique Sabaty^{11,12,13}, Thierry Desnos^{8,9,10}, Eva Pebay-Peyroula^{1,2,3}, Elisabeth Darrouzet¹⁸, Thierry Vernet^{1,2,3}, Norbert Rolland^{4,5,6,7}

1 Institut de Biologie Structurale Jean-Pierre Ebel, CEA, Grenoble, France, **2** Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075 CNRS, Grenoble, France, **3** Institut de Biologie Structurale Jean-Pierre Ebel, Université Joseph Fourier Grenoble I, Grenoble, France, **4** Laboratoire de Physiologie Cellulaire & Végétale, CEA, DSV, iRTSV, Grenoble, France, **5** Laboratoire de Physiologie Cellulaire & Végétale, CNRS, UMR 5168, Grenoble, France, **6** Laboratoire de Physiologie Cellulaire & Végétale, INRA, UMR 1200, Grenoble, France, **7** Laboratoire de Physiologie Cellulaire & Végétale, Université Joseph Fourier Grenoble I, Grenoble, France, **8** Laboratoire de Biologie du Développement des Plantes, CEA, DSV, iBEB, SBVME, St Paul les Durance, France, **9** Laboratoire de Biologie du Développement des Plantes, UMR 6191 CNRS, St Paul les Durance, France, **10** Laboratoire de Biologie du Développement des Plantes, Université Aix-Marseille, St Paul les Durance, France, **11** Laboratoire de Bioénergétique Cellulaire, CEA, DSV, iBEB, SBVME, St Paul les Durance, France, **12** Laboratoire de Bioénergétique Cellulaire, UMR 6191 CNRS, St Paul les Durance, France, **13** Laboratoire de Bioénergétique Cellulaire, Université Aix-Marseille, St Paul les Durance, France, **14** Laboratoire d'Ingénierie Cellulaire et Biotechnologie, CEA, DSV, iBEB, SBTN, Bagnols-sur-Cèze, France, **15** Laboratoire des Echanges Membranaires et Signalisation, CEA, DSV, iBEB, SBVME, St Paul les Durance, France, **16** Laboratoire des Echanges Membranaires et Signalisation, UMR 6191 CNRS, St Paul les Durance, France, **17** Laboratoire des Echanges Membranaires et Signalisation, Université Aix-Marseille, St Paul les Durance, France, **18** Laboratoire des Transporteurs en Imagerie et Radiothérapie en Oncologie, CEA, DSV, iBEB, SBTN, Bagnols-sur-Cèze, France

Abstract

Background: Membrane proteins are the targets of 50% of drugs, although they only represent 1% of total cellular proteins. The first major bottleneck on the route to their functional and structural characterisation is their overexpression; and simply choosing the right system can involve many months of trial and error. This work is intended as a guide to where to start when faced with heterologous expression of a membrane protein.

Methodology/Principal Findings: The expression of 20 membrane proteins, both peripheral and integral, in three prokaryotic (*E. coli*, *L. lactis*, *R. sphaeroides*) and three eukaryotic (*A. thaliana*, *N. benthamiana*, Sf9 insect cells) hosts was tested. The proteins tested were of various origins (bacteria, plants and mammals), functions (transporters, receptors, enzymes) and topologies (between 0 and 13 transmembrane segments). The Gateway system was used to clone all 20 genes into appropriate vectors for the hosts to be tested. Culture conditions were optimised for each host, and specific strategies were tested, such as the use of Mystic fusions in *E. coli*. 17 of the 20 proteins were produced at adequate yields for functional and, in some cases, structural studies. We have formulated general recommendations to assist with choosing an appropriate system based on our observations of protein behaviour in the different hosts.

Conclusions/Significance: Most of the methods presented here can be quite easily implemented in other laboratories. The results highlight certain factors that should be considered when selecting an expression host. The decision aide provided should help both newcomers and old-hands to select the best system for their favourite membrane protein.

Citation: Bernaudat F, Frelet-Barrand A, Pochon N, Dementin S, Hivin P, et al. (2011) Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host. PLoS ONE 6(12): e29191. doi:10.1371/journal.pone.0029191

Editor: Hendrik W. van Veen, University of Cambridge, United Kingdom

Received: August 25, 2011; **Accepted:** November 22, 2011; **Published:** December 21, 2011

Copyright: © 2011 Bernaudat et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work has been supported by the CEA (Commissariat à l'Energie Atomique et aux Energies Alternatives, funding of the CEA-PM project and Post-Doc fellowships to F. Bernaudat, A. Frelet-Barrand, P. Hivin and S. Dementin), by the CNRS (Centre National de la Recherche Scientifique, Post-Doc fellowship to A. Frelet-Barrand) and by the University of Grenoble 1. This work was partly supported by the European grant LSMH-CT-EUR-INTAFAR 2004-512138. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: florent.bernaudat@ibs.fr

These authors contributed equally to this work.

^a Current address: CEA Saclay, iBiTec-S, Service de Bioénergétique Biologie Structurale et Mécanismes (SB2SM), Gif-Sur-Yvette, France

^b Current address: CNRS-URA 2096, Gif-Sur-Yvette, France

^c Current address: Laboratoire de Bioénergétique et Ingénierie des Protéines, UPR 9036 CNRS, Marseille, France

^d Current address: Deinove, Montpellier, France

Introduction

Membrane proteins (MPs) perform a wide range of essential biological functions and represent the largest class of protein drug targets (for reviews, see [1–3]). Approximately 25% of all genes in both prokaryotes and eukaryotes code for MPs [4] and in humans 15% of these are G protein-coupled receptors (GPCRs) [2]. However, the vast majority of MPs still have no assigned function and only a little over 300 unique high-resolution 3D structures have been obtained for transmembrane proteins so far. Most of these structures are for bacterial and archaeal proteins, with only very few from eukaryotic systems [1,2,5] (<http://blanco.biomol.uci.edu/mpstruc>). This does not reflect the efforts deployed for the study of MPs in laboratories worldwide, but is an indication of the technical challenges posed by the hydrophobic nature, generally low natural abundance and intrinsic instability of these proteins. Obtaining sufficient amounts of MPs for functional and structural studies is the first major bottleneck in their study [6–12]; and when expressed in heterologous systems, the proteins are frequently *i*) toxic for the host, *ii*) expressed at a very low level in a spatially-delimited membranous environment and *iii*) mis- or unfolded (and thus inactive) [13]. Protein overexpression involves three elements: a gene, a vector and an expression host. The appropriate combination of these elements maximises the amount and quality of protein produced. However, since proteins are very diverse in structure and physico-chemical properties, it is impossible to predict whether a protein of interest will express well, be easy to purify, be active or crystallise in any given experimental setup [14]. Consequently, it is often necessary to test various constructions in diverse expression hosts. Traditional cloning methods with REaL (Restriction Enzyme and Ligase) steps to generate multiple expression plasmids (and constructs) are both labour-intensive and time-consuming. This makes them incompatible with a massively parallel strategy of expression screening. However, over the past few years, several recombinatorial cloning systems have been developed to allow rapid cloning of hundreds of genes and constructs simultaneously [14–17]. Among these, the Gateway technology [18], Creator [19] and the fragment exchange (FX) cloning [20] present the advantage of enabling subcloning of an open reading frame (ORF) into multiple expression vectors. Even if often adding extra-sequences to the proteins, Gateway is the most widely used and this technique has already been successfully exploited for high-throughput cloning of MPs [21], and several libraries from various ORFome projects have been constructed using Gateway vectors [22–27]. Gateway technology uses bacteriophage lambda Int/Xis/IHF recombination at *att* sites to transfer ORFs into vectors [28]. This divides the cloning procedure into three steps, as illustrated in Figure S1. In addition, most of the expression vectors available can be made Gateway-compatible by inserting an adapter cassette containing Gateway-specific recombination sites.

Once the expression vectors are obtained, production of the target proteins can be tested in different prokaryotic and eukaryotic expression systems suitable for overexpression of MPs (for reviews, see [12,15,29–34]). However, each of these systems has pros and cons, and the choice of the appropriate expression system often remains empirical, particularly with regard to the levels of functional protein expression. In the following paragraphs, we will briefly present the host systems tested in this study.

Escherichia coli

E. coli is by far the most widely used expression host for the production of recombinant proteins. Its short generation time, low cost and ease of use, as well as its extensive characterisation make

it an ideal candidate (for a review see [35]). However, *E. coli* presents some disadvantages for protein overexpression. In particular, many MPs do not fold properly and form aggregates that are then stored in inclusion bodies. Several recent developments have improved the expression of recombinant MPs in *E. coli* [36]. Strains like C41, C43 [13] or Lemo21 [37], which are more tolerant to toxic MPs, or the introduction of tags like GFP [38], MBP, GST, NusA [30] or Mistic [39] can facilitate and improve MP production. Mistic is a 13 kDa protein from *Bacillus subtilis*, which, when produced in *E. coli*, spontaneously associates with the inner membrane, without requiring recognition by the Sec translocon machinery. Due to this spontaneous association with the membrane, Mistic has been successfully used as an *N*-terminal fusion tag to target and facilitate membrane insertion of various cargo MPs in *E. coli* [39–45].

Lactococcus lactis

L. lactis, like other food-grade lactic acid bacteria, is a non-pathogenic, non-invasive Gram-positive bacterium. These properties have made it a popular candidate for the oral administration of mucosal vaccines (for recent reviews, see [46–50]). Thanks to the development of a wide range of genetic engineering tools (for a review see [51]), it is also widely used today for large-scale production of heterologous proteins [29,30,46].

Recombinant protein production in *L. lactis* can be performed using the Nisin-Inducible Controlled gene Expression (NICE) system, in which nisin, an antimicrobial peptide, is used to promote the expression of genes positioned in plasmids under the control of the nisin-inducible promoter P_{nisA} (see review [47]). This system has been used to produce various eukaryotic MPs in *L. lactis* [9,30,46,52–54]. GFP has also been used to monitor the state of protein folding, in order to select evolved hosts with enhanced functional expression of membrane proteins [55]. One of the major advantages of *L. lactis* over *E. coli* is that inclusion bodies have (so far) not been observed in this host [9]. In addition, it only has a single cell membrane, making the direct use of ligands or inhibitors for activity studies of membrane proteins in whole cells possible. Until recently, expression screening of multiple constructs in *L. lactis* was limited by the absence of efficient cloning procedures, but recent developments based on ligation-independent cloning (LIC) and Gateway technology have made it possible to clone many genes in parallel [54,56,57].

Rhodobacter sphaeroides

R. sphaeroides is a purple non-sulphur photosynthetic bacterium. The pigment-protein complexes of the photosynthetic apparatus (reaction centres, light-harvesting complexes) are located in invaginations of the cytoplasmic membrane, known as chromatophores. In response to light and/or lowered oxygen tension, the bacteria synthesises large amounts of photosystems [58], and the increasing number of chromatophores causes the membrane surface area to increase vastly. This increase in the intracytoplasmic membrane surface could be very useful for the production of MPs. Indeed, one of the major limitations for MP production in many hosts is the limited membranous space available. In *R. sphaeroides*, foreign MP synthesis can be coordinated with the synthesis of new membranes to accommodate them. This property has already been used to produce heterologous MPs for structural studies [59].

Arabidopsis thaliana

A small flowering plant with a relatively short life cycle of two months, *A. thaliana* is a popular model organism in plant biology and genetics. Its small genome was fully sequenced in 2000 [60]. *A.*

thaliana is not regarded as a classical overexpression system since most plant MPs are overexpressed in plants to test their *in vivo* function rather than to obtain sufficient amounts for crystallisation trials. *A. thaliana* can be both stably transformed (by floral dipping [61]) and transiently transformed (by agro-infiltration with *Agrobacterium tumefaciens* [62]). When overexpressing MPs in this organism using stable transformants, the main limitation is the long culture cycle, lasting two months between generations of plant seeds, as compared to only 30 to 50 min for bacteria.

Nicotiana benthamiana

Widely used as an experimental host in plant virology, *N. benthamiana* can be efficiently genetically transformed and regenerated. It is therefore amenable to transient protein expression [63]. This host is rapidly gaining popularity in plant biology, particularly in studies requiring protein localisation, interaction, or plant-based systems for protein expression and purification. Transient *Agrobacterium*-mediated transformation of *N. benthamiana* using leaf disks has provided the plant community with a valuable tool to rapidly evaluate transgenes in higher plants [64] and to produce gram quantities of recombinant proteins [65]. This protocol has a number of significant advantages: readily available explant material, high efficiency, and a relatively quick turnaround time.

Insect cells and the baculovirus system

The baculovirus system is widely used for eukaryotic protein expression in insect cells [66,67] as a compromise between bacterial expression and expression in mammalian (stably or transiently transfected) cells. Indeed, although more expensive and time-consuming than expression in *E. coli*, this system is more compatible with eukaryotic proteins because of similar codon usage rules, providing better expression levels and fewer truncated proteins than in bacteria. In addition, this system allows for post-translational modifications. Some of the post-translational modifications produced are not identical to those found in mammals (glycosylations for example), but they are nevertheless closer than those produced by bacteria, or even yeast [68]. Insect cells are easier and cheaper to handle than adherent cells like HEK 293, COS or CHO cells, especially for scale-up. Thus, these cells used with the baculovirus system have a real potential for the heterologous production of MPs. Briefly, the baculovirus system relies on the infection of insect cell lines (usually Sf9, Sf21 or High Five®) by recombinant viruses encoding the gene(s) of interest. Many improvements to recombinant baculovirus generation have been implemented over the last twenty years [69], including the Bac-to-Bac system (Invitrogen), which uses site-specific transposition in *E. coli* rather than homologous recombination in insect cells. Gene expression is generally driven by the polyhedrin or p10 late promoter. A similar system (BacMam, Invitrogen) has recently been developed to allow baculovirus-based expression in mammalian cells.

Rationale for the current study

Several studies comparing different expression systems for MP production have already been performed. However, these studies focused either on the expression of a given protein [7] or a family of proteins such as GPCRs [12,70]. Other laboratories have tried to express MPs only in *E. coli* [21,71,72] or *L. lactis* [73]. Moreover, except for GPCRs [12,70], the expression of eukaryotic MPs has only been compared in either prokaryotic [74] or eukaryotic [75] hosts. To our knowledge, our study is the first to compare the overexpression of 20 prokaryotic and eukaryotic MPs in both prokaryotic (*E. coli*, *L. lactis* and *R. sphaeroides*) and eukaryotic (*A.*

thaliana, *N. benthamiana* and Sf9 insect cells) expression hosts. This study is also original as we evaluate commonly used hosts such as *E. coli* and, to a lesser extent, insect cells together with more unusual systems, to test their ability to be used as alternative expression hosts. As overexpression of membrane proteins is a challenge in itself, we focused our attention on the production step, and on the yields obtained in the various expression hosts tested. However, in extensions of the present study, we were able to show that some of the proteins produced here could be purified to homogeneity and were active [54,76,77]. The present article highlights several successful strategies for the heterologous expression of the MPs studied (from different protein families and with large variations in topology and origin) and discusses possible further improvements to MP expression. But, most importantly, it provides a first-stop analysis of the pitfalls and advantages of the various systems tested depending on the nature of the MP to be expressed. This should be of use to all who are about to venture into this exciting, and sometimes frustrating, field of biology.

Materials and Methods

Cloning using the Gateway technology

The cloning steps were performed using Gateway technology (Invitrogen) according to the manufacturer's instructions, but by reducing the volume and quantities of all components (clonase enzyme, buffer, PCR products and vectors) to 1/8th during the recombination steps (BP and LR reactions), to yield a total reaction volume of 2.5 µl that was entirely used for transformation. Briefly, the ORFs coding for the selected proteins (Table 1) were amplified by PCR and flanked with *attB* specific recombination sites. All the genes were also extended with a sequence coding for a *Strep*-tag II at either the *N*- or *C*- terminal end of the constructs. The PCR products were purified and either recombined with a pDONR221 donor vector (Invitrogen) through a BP reaction or cloned into pENTR-D-TOPO vectors through directional topoisomerase-mediated cloning (TOPO, Invitrogen) to yield the "entry" clones. The entry clones were first sequenced to check the integrity of the cloned genes and then used in an LR Gateway reaction together with various destination vectors to yield expression vectors specific to each expression system tested in this study (Table 2).

***E. coli* expression vectors.** To test the expression of the proteins in *E. coli*, the genes were transferred into the destination vectors pDEST17 (Invitrogen) and pDEST-Mistic. pDEST-Mistic was obtained by modifying the vector pDEST17 by introducing the sequence coding for Mistic (Accession n° AAX20121) between the coding sequences of the hexa-histidine tag and the *attB1* site through RF cloning as described by van den Ent and Löwe [78].

***L. lactis* expression vector.** The vector pNZ8148 containing the NICE system was used for expression in *L. lactis*. This vector wasn't converted into a Gateway destination vector, because it is known to be very unstable in *E. coli* and because of the lack of *Lactococcus* strains able to propagate Gateway vectors. Therefore, the cDNAs were first transferred into the vector pBS-RfA using the Gateway system and subsequently cloned into pNZ8148 through digestion of pBS-RfA vectors by EcoRV and re-ligation (for details, see [54]). For some proteins (MraY, AtHMA3, AtHMA4 and $\alpha 2\delta$ subunit), with one or several EcoRV restriction sites within the ORF sequence, a partial restriction of the donor plasmids with this restriction enzyme led to a correct excision of the cassette containing the entire gene. Afterwards, *Lactococcus* strain NZ9000 was transformed with the recombinant plasmids as previously described [79] and the

Table 1. List of selected target proteins.

| Acc n° UNIPROT | Protein name | Function | Organism | Size (kDa) | Topology ^a | Reference |
|----------------|--------------------------|--|-----------------------|------------|-----------------------|-----------|
| Q6NCP8 | P450 | Cytochrome - mono-oxygenase | <i>R. palustris</i> | 49.7 | Peripheral | [87] |
| O88116 | NapC | Cytochrome - electron transfer | <i>R. sphaeroides</i> | 24.2 | 1 TM | [110] |
| Q8DMY2 | MreC | Peptidoglycan synthesis | <i>S. pneumoniae</i> | 29.7 | 1 TM | [111] |
| Q8DQH3 | FtsX | Cell division | <i>S. pneumoniae</i> | 34.2 | 4 TM | [112] |
| Q8DR69 | MraY | Peptidoglycan synthesis | <i>S. pneumoniae</i> | 36.0 | 10 TM | [113] |
| A5X8Y8 | LPR1 | Multi-copper oxydase | <i>A. thaliana</i> | 60.5 | Peripheral | [114] |
| Q9SV68 | ceQORH | Quinone oxydoreductase - electron transfer | <i>A. thaliana</i> | 33.1 | Peripheral | [100,101] |
| Q8GYE0 | PHF | Phosphate transport regulation | <i>A. thaliana</i> | 42.4 | 1 TM | [115] |
| Q9M3H5 | AtHMA1 | Heavy metal transporter | <i>A. thaliana</i> | 80.1 | 6 TM | [116] |
| P31167 | AAC | Mitochondria ADP/ATP transporter | <i>A. thaliana</i> | 33.2 | 6 TM | [117] |
| Q66474 | AtHMA4 | Heavy metal transporter | <i>A. thaliana</i> | 126.7 | 8 TM | [118] |
| Q9SZW5 | AtHMA3 | Heavy metal transporter | <i>A. thaliana</i> | 81.4 | 8 TM | [119] |
| Q96303 | PHT1;4 | Phosphate transporter | <i>A. thaliana</i> | 57.2 | 12 TM | [120] |
| Q39002 | NTT1 | Chloroplast ADP/ATP transporter | <i>A. thaliana</i> | 57.5 | 12 TM | [121] |
| P54290 | $\alpha 2\delta$ subunit | Calcium channel regulation | <i>R. norvegicus</i> | 122.2 | 1 TM | [122] |
| P04633 | UCP1 | Uncoupling protein | <i>R. norvegicus</i> | 31.3 | 6 TM | [123] |
| Q07817 | Bcl-xL | Apoptosis regulation | <i>H. sapiens</i> | 24.7 | 1 TM | [124,125] |
| P61073 | CXCR4 | GPCR | <i>H. sapiens</i> | 37.9 | 7 TM | [126] |
| P51681 | CCR5 | GPCR | <i>H. sapiens</i> | 38.7 | 7 TM | [127,128] |
| Q92911 | NIS | Iodide transporter | <i>H. sapiens</i> | 67.6 | 13 TM | [129] |

^aFor some of the proteins, the topology is still unclear and the number of TMs given here corresponds to the predicted topology.
doi:10.1371/journal.pone.0029191.t001

presence of the insert in the right orientation was confirmed using restriction analyses, PCR amplification and subsequent sequencing [80].

***R. sphaeroides* expression vector.** For expression in *R. sphaeroides*, the broad-host-range plasmid pBBR1MCS-2 [81] was modified to convert it into a Gateway destination vector and to change the antibiotic resistance. An omega cartridge encoding resistance to streptomycin and spectinomycin was obtained

through BamHI digestion of pHP45 Ω plasmid [82] and cloned into pBBR1MCS-2 previously digested with BglII. The *aph* gene encoding resistance to kanamycin was inactivated by the excision of a 400 bp NcoI fragment. To enhance protein expression, the strong promoter and the RBS of the *pu* operon (encoding light harvesting complexes II) were cloned into the resulting plasmid, pBBR1MCS Ω . This was done by amplifying *R. sphaeroides* genomic DNA by PCR, using the primers 5'-AAGGTACCTGC-

Table 2. Protein constructs obtained from the different expression vectors.

| Expression host | Expression vector | Expressed protein construct* |
|-----------------------------------|-------------------|---|
| <i>E. coli</i> | pDEST17 | MSYY(H ₆)LE-attB1-MP-Strep |
| <i>E. coli</i> | pDEST-Mistic | MSYY(H ₆)LE-Mistic-attB1-MP-Strep |
| <i>L. lactis</i> | pNZ8148 | MI-attB1-MP-Strep |
| <i>R. sphaeroides</i> | pDEST-E | VDI-attB1-MP-Strep |
| <i>A. thaliana/N. benthamiana</i> | pAlligator-3 | M-MP-Strep |
| Insect cells | pDEST8 | M-MP-Strep |

Sequences are presented using one-letter code for amino acids. attB1: amino acid sequence encoded by the attB1 cloning site corresponding to TSLYKKAGS when the entry clone was prepared through BP cloning and TSLYKKAGSAAAPFT when the entry clone was prepared through TOPO recombination (NapC, P450, LPR1, PHF, PHT1;4, ceQORH, AtHMA1, Bcl-xL). MP: amino acid sequence of the different membrane proteins. Mistic: amino acid sequence of the fusion tag Mistic. Strep: amino acid sequence of the Strep-tag II fusion tag corresponding to WSHQPQFEK.

*The Strep-tag II was fused to the C-terminus of most proteins, except for proteins AthMA3, AthMA4 and Bcl-xL for which the Strep-tag II was located at the N-terminus of the MP sequence.

doi:10.1371/journal.pone.0029191.t002

AGGCCACGCCCTGAA-3' and 5'-AAGATATCCACTGT-GTCGTCTCCCAACT-3'. The 0.7 kbp PCR product was then digested with KpnI and EcoRV and cloned into pBBR1MCSΩ. Finally, the resulting plasmid was linearised with EcoRV and a Reading Frame Cassette A (RfA) (Invitrogen) was introduced to convert it into a Gateway destination vector.

A. thaliana and N. benthamiana expression vector. The expression vector used for plant transformation was the pAlligator3 vector [83] containing the spectinomycin resistance marker gene and the CaMV 35S promoter (Cauliflower Mosaic Virus). This vector also includes a gene coding for GFP, driven by the At2S3 seed-specific promoter and used as a selectable marker for transformed seeds, as well as the Gateway cloning cassette [83]. *A. tumefaciens* strain (C58) was transformed with the different expression vectors as previously described [84] and the presence of recombinant vectors was verified by plasmid isolation and restriction analysis.

Insect cell expression vectors and bacmids. The entry clones were recombined (LR reaction) with the commercial destination vector pDEST8 (Invitrogen) to generate the expression plasmids, which were checked by restriction digest. According to the Invitrogen manual, the only requirement needed to use pDEST8 when designing the "Entry" clone, is the insertion of an ATG start codon for proper initiation of translation. These plasmids were subsequently transformed into DH10Bac™ (Invitrogen) for transposition with the parent bacmid. After the blue/white screening of positive recombinants (LacZα complementation system on the bacmid), the various recombinant bacmids thus obtained were further checked by PCR for the presence of the genes of interest.

Protein expression in the different systems

E. coli based expression. Expression vectors were used to transform C43(DE3) (Avidis) and BL21-AI (Invitrogen) competent cells. Expression tests were performed in 24-Deep well plates containing 3 mL of TB medium (100 µg/mL Ampicillin). The cultures were inoculated with overnight pre-cultures at a 1/40th dilution and grown for 2 h at 37°C under agitation (250 rpm). Protein expression was then induced by addition of either 1 mM IPTG for C43 cells or 0.005% (w/v) arabinose for BL21-AI cells and the cultures were incubated for another 16 h at 20°C under agitation (250 rpm). The cells were harvested by centrifugation (3200 g, 10 min, 4°C) and the cell pellet resuspended in 250 µL of PBS buffer containing lysozyme (Novagen), benzonase (Novagen) and Complete antiprotease cocktail (Roche). Cells were disrupted using a water bath sonicator and debris were removed by centrifugation (20,000 g, 20 min, 4°C). Membranes present in the supernatant were separated by ultracentrifugation (100,000 g, 1 h, 4°C). Finally, the membrane pellet was resuspended in 250 µL PBS buffer and 10 µL aliquots were analysed on gradient 4–20% SDS-PAGE gels (Bio-Rad) and by western blots (WB). Total MP content was determined using the BCA protein assay (Pierce).

L. lactis based expression. Expression tests were performed in 1 L-cultures and crude bacterial membranes were purified as previously described [80]. Briefly, transformed NZ9000 *Lactococcus* cells were grown in 1L of M17 medium (Difco) supplemented with 1% (w/v) glucose and 10 µg/mL chloramphenicol. Cultures were inoculated with 25 mL of overnight pre-cultures and grown at 30°C under gentle shaking (90 rpm). Protein expression was induced when the OD₆₀₀ reached 0.8, with a 0.005 volume of the nisin A-containing supernatant obtained from a culture of the *L. lactis* NZ97000 nisin-producing strain (NIZO). After induction the cells were grown for an additional 4 h at 30°C, under gentle

shaking (90 rpm). The cells were then harvested (5000 g, 15 min, 4°C) and resuspended in 40 mL of Tris buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl). Bacteria were lysed using a cell disruptor (One shot, Constant Systems) by 2-fold passages at 35,000 p.s.i. (= 2.3 kbars) and the lysates clarified by centrifugation (10,000 g, 10 min, 4°C). The supernatant containing the membranes was then ultracentrifuged (150,000 g, 1 h, 4°C) and the membranes were resuspended in 2 mL of PBS-Glycerol (10% (v/v)). Total MP content was measured using the Bio-Rad protein assay (Bio-Rad) and 20 µg of proteins were analysed on 10% SDS-PAGE and by western blots. Bacteria containing the empty pNZ8148 vector were systematically grown in parallel and used as negative control to validate the nature of the detected signals.

R. sphaeroides based expression. The expression vectors were mobilised from *E. coli* to *R. sphaeroides* f. sp. *denitrificans* IL106 by conjugation. Cells were grown for 24 h at 30°C in Hutner modified medium [85] under aerobic conditions (100 mL medium in 250-mL erlenmeyer flasks, 150 rpm) or phototrophic conditions (180 µmol of photons.m⁻².s⁻¹) with 25 µg/mL kanamycin. Cells were harvested (7000 g, 10 min, 4°C) and resuspended in 8 mL of Tris buffer (50 mM Tris-HCl pH 8.0). The bacteria were lysed using a cell disruptor (One shot, Constant Systems) and the lysates clarified by centrifugation (7000 g, 10 min, 4°C). The supernatant containing the membranes was then ultracentrifuged (200,000 g, 1 h, 4°C) and the membranes were resuspended in 1 mL of Tris buffer. The protein content was measured with the BC assay (Interchim) in 2% SDS, and 25 µg of proteins were analysed by 10% SDS-PAGE and by western blots.

A. thaliana based expression. Plants were grown in culture chambers at 23°C (8-h light cycle) with a light intensity of 150 µmol.m⁻².s⁻¹ in standard conditions. Wild-type *Arabidopsis* plants ecotype Wassilevskija background were transformed by dipping the floral buds of 4–5-weeks-old plants into an *A. tumefaciens* (C58 strain) solution containing a surfactant (Silwett L-77) according to Clough and Bent [86]. Primary transformant seeds were selected on the basis of GFP fluorescence [83] and germinated in Petri dishes containing solidified medium (Murashige and Skoog, 0.5% (w/v) sucrose, and 0.8% (w/v) agarose) for 2 weeks before transfer to soil. After 3–4 weeks, total MPs were extracted from 1–2 leaves. Finally, membrane proteins were diluted in 200 µL of Tris buffer (50 mM Tris-HCl pH 6.8, 1% Triton X-100) and 25 µg aliquots were analysed on 12% SDS-PAGE and by western blots.

N. benthamiana based expression. Plants were grown in culture chambers at 20°C (14 h light cycle) with a light intensity of 60–120 µmol.m⁻².s⁻¹ in standard conditions. Three or four week-old wild-type *Nicotiana benthamiana* plants were infiltrated with a solution of *A. tumefaciens* (C58 strain) according to Witte et al. [64]. Total MPs were extracted from 2 leaf discs harvested after 4 days [59]. Finally, membrane proteins were resuspended in 70 µL of buffer (100 mM Tris-HCl pH 8, 5 mM EDTA, 150 mM NaCl, 5 mM DTT, anti protease inhibitors, 1% Triton X-100) and 10 µL aliquots were analysed on 10% SDS-PAGE and by western blots.

Sf9 insect cells based expression. The bacmids were amplified in DH10Bac and purified using the S.N.A.P.™ MidiPrep Kit (Invitrogen). Sf9 cells were transfected with cellfectin according to Invitrogen's protocol (Bac to Bac baculovirus expression system) and incubated for 72 h to get the P1 viral stock. This P1 viral stock was then amplified by infecting Sf9 cells and the P2 viral stock thus obtained was subsequently used for expression experiments. The precise titers of these viral stocks have not been determined and after preliminary experiments to determine the best conditions for protein

expression, 10% of viral inoculum was used for all the experiments. After infection, cells (approximately 10^6 per well) were incubated at 27°C for 48 h, and centrifuged for 5 minutes at 1000 *g*. For analyses on whole cell extracts, the cells were then washed in PBS and resuspended in 300 μ L of 10 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS plus protease inhibitors and kept on ice for 20 min. The lysate was centrifuged at 16,000 *g* for 15 min to remove the non-solubilised material. For analyses on membrane fractions, the cell pellet was suspended in 1 mL of cold 20 mM Tris pH 7.5, 250 mM sucrose, plus anti-proteases (Complete, Roche) buffer. After breaking the cells with a Dounce homogeniser (10–15 passages on ice), the lysate was centrifuged at 1000 *g* for 10 min. The supernatant was transferred and centrifuged at 10,000 *g* for 10 min. At last, from this supernatant, membranes were concentrated as a pellet at 100,000 *g* for 1 h. All steps were performed at 4°C or on ice. MPs were diluted in 300 μ L of 25 mM Tris, pH 7.5, 100 mM mannitol plus anti-proteases (Complete, Roche). Total MPs were determined using the BCA protein assay (Pierce). For western blot analysis 20 μ g of proteins were loaded onto a NOVEX NuPAGE 4–12% Bis-Tris gel (Invitrogen) with MES/SDS running buffer. Non-infected cells were used as a control.

Western blot analysis

The membrane fraction extracted from cells from each expression system was analysed by western blotting using the *Strep*-tag II sequence as the antigenic epitope, unless specified otherwise. Western blots were performed using the *Strep*-tag HRP Detection Kit (IBA) according to the manufacturer's instructions, unless otherwise stated. The amounts of target proteins present in isolated membrane samples were quantified by densitometry with background correction and comparison to known amounts of a control *Strep*-tagged protein loaded on the same blot. For both plant expression systems we followed the protocol described by Witte et al. [64], with some modifications for *Arabidopsis* by adding a blocking step with the biotin blocking buffer because of the presence of several biotinylated proteins in *Arabidopsis* crude membrane extracts. For *L. lactis*, two different methods were applied depending on the expression level of the protein as previously described [54]. Total membrane protein (TMP) concentrations in isolated membrane samples were also determined using conventional colorimetric methods as stated above.

Results

Generation of expression plasmids and cell lines

Our aim was to test the overexpression of 20 MPs (Table 1) in six host organisms, this required engineering 120 expression vectors. Gateway technology was used to optimise and streamline cloning, providing a success-rate over 99% for plasmid generation. The only expression plasmid not produced at all was the *L. lactis* expression vector for the $\alpha 2\delta$ subunit, which was lost in the cloning step after the Gateway step. This was probably due to the large (>4 kbp) size of the cDNA, or to the presence of several *EcoRV* restriction sites within the gene sequence. In the baculovirus system, all 20 pDEST8 recombinant plasmids were obtained. However, the corresponding bacmids could not be produced for P450 and NIS. In all other organisms, all 20 cell lines were successfully produced.

Expression results

The proteins in this study belong to diverse protein families, are of both prokaryotic and eukaryotic origin, and their topology

ranges from peripheral MPs to integral MPs (IMPs) containing between one and thirteen predicted transmembrane (TM) regions (Table 1). To evaluate the efficiency of the different expression systems, after protein expression, the membranes were isolated as described in Material and Methods. The amount of target protein in the membranes was determined by western-blot, using the *Strep*-tag II sequence (if not otherwise stated) to reveal the presence of target protein on the membrane (See Figure 1 (A) to (G) and Table 3). Expression levels are generally given in mg/L of culture for bacteria and Sf9 cells. However, because we also used plant systems, we also considered the production levels as a percentage, target MP within the total pool of membrane proteins (TMP) (Table 3). This made it possible to compare all the different expression systems used here.

Expression in *E. coli*. Prior to the screen of the 20 proteins, several expression conditions (concentration of inducing agent and temperature) were tested for the production of a few proteins in BL21-AI. A concentration of 0.005% arabinose and an overnight induction at 20°C gave the best results. For C43 strain, a concentration of 1 mM IPTG was retained. These conditions were then applied in the expression screening that was performed in triplicate for all proteins in both strains. No significant differences were observed between the strains in terms of expression levels, therefore the results were averaged in Table 3. Two plasmids were used to transform *E. coli*: pDEST17 yielded a construct in which the amino acids encoded by the *attB1* recombination site formed a linker between an *N*-terminal His-tag and the proteins (Table 2); whereas with pDEST-Mistic, Mistic was located between the *N*-terminal His-tag and the *attB1* sequence, followed by the target protein (Table 2). A representative western blot of proteins produced in *E. coli* is shown in Figure 1A. Detection of western blots using *Strep*-Tactin HRP conjugates had a useful side-effect in *E. coli*, where a soluble endogenous biotinylated protein, biotin carboxyl carrier protein (BCCP; 22.5 kDa), was detected. This protein should be absent from the membrane fraction and was therefore used to control the purity of this fraction (Figure 2).

Detectable amounts of full-length protein were obtained for 15 out of the 20 MPs (the three peripheral proteins and 12 IMPs) in *E. coli* with or without fusion to Mistic. For the other five proteins, either no signal was detected, or the MW was too far removed from the expected value (e.g. for NIS, a signal was observed at one third the expected MW). Bands of this type could be the result of proteolytic degradation, internal initiation or premature termination. Since the proteins produced in *E. coli* had both a His-tag and the *Strep*-tag II sequence, western blots could also be probed using anti-His antibodies. This was done for a few proteins that were not detected with the *Strep*-tag, to check whether the lack of signal was due to the absence of the protein, or to a tag detection problem. For FtsX, a signal was indeed obtained with anti-His, indicating that, for this protein there was some problem with the *Strep*-tag II. This type of problem may also have occurred for some proteins in the other expression systems (see below).

Mistic fusion significantly increased the yields of the 12 IMPs produced in *E. coli*. In contrast, it had a negative effect when fused to peripheral proteins, drastically reducing the amount of target protein associated with the isolated membranes in all three cases. This should therefore be taken into account when selecting a vector for protein expression.

Functional studies, detailed elsewhere [54,76,77], showed several of these proteins to be active and readily purified.

Expression in *L. lactis*. Before screening for expression of all proteins in *L. lactis*, culture conditions were optimised (temperature, induction time and concentration of nisin) for two representative proteins, one peripheral (ceQORH) and one

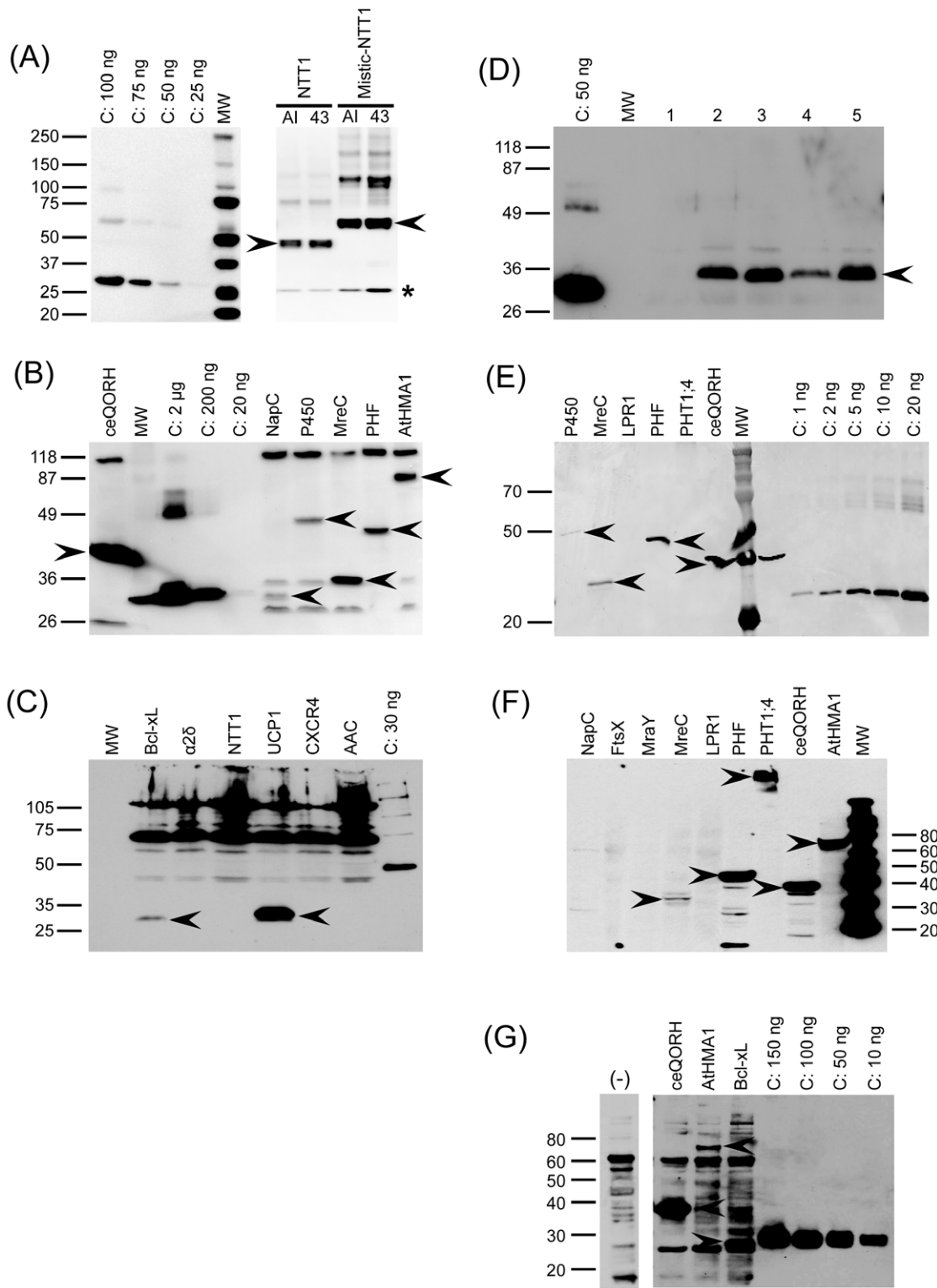


Figure 1. Examples of western blot analysis of cell extracts from the different hosts. (A) **Western blot analysis of membrane extracts of *E. coli*.** In this case, the native and Mistic-NTT1 fusion. C: *Strep*-tag II control protein loaded at 25, 50, 75 or 100 ng. AI: proteins produced in BL21-AI. 43: proteins produced in C43. MW: molecular weight standard. Arrows point out the different target proteins. *: Partly proteolysed NTT1 protein. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). (B) **Western Blot analysis of membrane extracts of *L. lactis*.** C: *Strep*-tag II control protein loaded at 2000, 200 or 20 ng as written above. MW: molecular weight standard. Arrows point out the different target proteins. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). (C) **Western blot analysis of membrane extracts of *R. sphaeroides*.** C: *Strep*-tag II control protein loaded at 30 ng as written above. MW: molecular weight standard. Arrows point out the different target proteins. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). (D) **Western blot analysis of membrane extracts of *A. thaliana*.** In this case, the expression of the protein AAC was tested in 5 different transformed plants. The membrane fraction was isolated and the extracts corresponding

to the different plants tested (lanes 1 to 5) were analysed. C: *Strep*-tag II control protein loaded at 50 ng as written above. MW: molecular weight standard. The arrow points out the protein AAC. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). (E) **Western blot analysis of membrane extracts of *N. benthamiana* leaf discs.** C: *Strep*-tag II control protein loaded at 1, 2, 5, 10 or 20 ng as written above. MW: molecular weight marker. Arrows point out the different target proteins. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). (F) **Western blot analysis of whole cell extracts of Sf9 insect cells.** MW: molecular weight standard. Arrows point out the different target proteins. The membrane was probed with the anti-*Strep*-Tag II (IBA) and a goat anti mouse-HRP secondary antibody. (G) **Western blot analysis of membrane extracts of Sf9 insect cells.** This figure is an example of a western-blot for the quantification of target proteins in Sf9 cells membrane vesicles. Here, membrane vesicles of Sf9 cells overproducing either no protein (–), ceQORH, AtHMA1 or Bcl-xL were deposited. C: *Strep*-tag II control protein loaded at 150, 100, 50, 10 ng as written above. Arrows point out the different target proteins. The membrane was probed with the Anti-*Strep*-Tag II (IBA) and a goat anti mouse-HRP secondary antibody.
doi:10.1371/journal.pone.0029191.g001

intrinsic MP (AtHMA1). The nisin used for induction was produced in-house as described in Material and Methods, and the same batch was used for all the tests performed in this study. Optimal production of both proteins was achieved by adding 0.005 volume nisin A-containing NZ9700 medium supernatant to a culture at OD600≈0.8. Production levels were two- to three-fold higher when the cells were grown at 30°C for 4 h after induction, rather than overnight at 20°C (data not shown). These culture and induction conditions were then applied to test the expression of all 20 proteins. Thirteen of the 20 proteins tested were successfully produced. The quantities of MPs obtained were about 1/10th those provided by *E. coli* (Table 3; for a representative western blot, see Figure 1B). However, the plant protein ceQORH was produced at 9 mg/L, which corresponded to 30% of the TMPs in these cells. This is comparable to the levels obtained with prokaryotic MPs. As in *E. coli*, some of the proteins that could not be identified through the *Strep*-tag II, could be detected using other antigenic epitopes. For example, AtHMA4, which contains an internal poly-histidine sequence, could be detected using anti-

His antibodies, while AtHMA3 and Bcl-xL were detected using protein-specific antibodies (Figure 3).

Further functional studies were performed on some of the proteins expressed in *L. lactis*, these are detailed elsewhere [54]. The specific activity of the protein ceQORH was significantly improved in this host compared to *E. coli*.

Expression in *R. sphaeroides*. In *R. sphaeroides*, intracytoplasmic membrane is synthesised in response to specific growth conditions. We tested the expression of target proteins under both phototrophic anaerobic conditions and semi-aerobic conditions. The different conditions did not have a significant impact on results, and only four proteins could be produced in this host (Table 3; for a representative western blot, see Figure 1C). Cytochrome P450 was found to be correctly folded and active, since it could fix CO [87]. However, we were quite surprised by the limited (20%) success rate of membrane protein expression using this system. Indeed, in other experiments, large amount of soluble proteins were produced using either a pRK415 [88] or a pBBR1MCS-2 derivative with the *puc* promoter. This vector also allows expression of MP, since we were able to express cytochrome

Table 3. Protein yields obtained in the different expression hosts.

| Protein name | <i>E. coli</i> (His) | <i>E. coli</i> (Mistic) | <i>L. lactis</i> | <i>R. sphaeroides</i> | Insect cells | <i>A. thaliana</i> | <i>N. benthamiana</i> |
|--------------|----------------------|-------------------------|------------------|-----------------------|---------------|--------------------|-----------------------|
| P450 | ◆◆◆◆ (5–10%) | – | ◆◆ (1–3%) | ◆◆◆◆ (1–3%) | n.c. | n.a. (<0.1%) | n.a. (<0.1%) |
| NapC | ◆◆◆ (3–5%) | ◆◆◆◆ (15–20%) | ◆◆ (0.5–1%) | – | ◆ (<0.1%) | n.a. (0.1–0.5%) | n.a. (<0.1%) |
| MreC | ◆◆◆◆ (10–15%) | ◆◆◆◆◆ (15–20%) | ◆◆◆ (0.5–1%) | – | ◆ (<0.1%) | n.a. (<0.1%) | n.a. (<0.1%) |
| FtsX | ◆◆◆ (3–5%) | ◆◆◆◆ (5–10%) | – | – | – | – | – |
| MraY | – | – | – | – | – | – | – |
| LPR1 | ◆◆◆◆ (1–3%) | – | ◆ (0.1–0.5%) | – | – | – | – |
| ceQORH | ◆◆◆◆◆ (15–20%) | ◆◆◆◆ (10–15%) | ◆◆◆◆◆ (30%) | ◆ (<0.1%) | ◆◆ (0.5–1%) | n.a. (0.1–0.5%) | n.a. (0.5–1%) |
| PHF | ◆◆◆◆ (5–10%) | ◆◆◆◆◆ (10–15%) | ◆◆◆◆ (1–3%) | – | ◆◆ (0.1–0.5%) | – | n.a. (0.1–0.5%) |
| AtHMA1 | – | ◆◆◆ (3–5%) | ◆◆ (1–3%) | – | ◆ (<0.1%) | n.a. (<0.1%) | – |
| AAC | ◆◆◆ (1–3%) | ◆◆◆◆ (3–5%) | – | – | ◆◆ (0.1–0.5%) | n.a. (0.1–0.5%) | – |
| AtHMA4 | – | – | ◆ (0.5–1%) | – | – | – | – |
| AtHMA3 | – | – | ◆◆ (0.5–1%) | – | – | n.a. (<0.1%) | – |
| PHT1,4 | ◆◆◆ (3–5%) | ◆◆◆◆ (5–10%) | – | – | – | – | – |
| NTT1 | ◆◆◆◆ (3–5%) | ◆◆◆◆◆ (5–10%) | ◆◆ (0.1–0.5%) | – | – | n.a. (<0.1%) | – |
| α2δ subunit | – | – | n.c. | – | – | – | – |
| UCP1 | ◆◆◆ (3–5%) | ◆◆◆◆ (5–10%) | – | ◆◆ (0.1–0.5%) | ◆◆ (0.1–0.5%) | n.a. (<0.1%) | n.a. (<0.1%) |
| Bcl-xL | ◆◆◆ (1–3%) | ◆◆◆◆ (3–5%) | ◆◆ (0.5–1%) | ◆ (<0.1%) | ◆◆ (0.1–0.5%) | n.a. (<0.1%) | n.a. (<0.1%) |
| CXCR4 | ◆◆◆ (1–3%) | ◆◆◆◆ (3–5%) | ◆ (<0.1%) | – | – | – | – |
| CCR5 | ◆◆◆ (1–3%) | ◆◆◆◆ (3–5%) | ◆ (<0.1%) | – | – | – | – |
| NIS | – | – | – | – | n.c. | – | – |

For each protein, the yield obtained is expressed in mg of target protein/liter of cell culture (black squares) and as percentage of total membrane protein (% TMP) (figures in brackets). –: protein not detected by western blot, ◆ = protein produced at a yield below 0.1 mg/L culture, ◆◆ = 0.1–0.5 mg/L, ◆◆◆ = 0.5–1 mg/L, ◆◆◆◆ = 1–4 mg/L, ◆◆◆◆◆ = 4–7 mg/L and ◆◆◆◆◆◆ >7 mg/L. n.a. = not applicable, as the two plant systems are not cultured in liquid media. n.c. = not cloned.
doi:10.1371/journal.pone.0029191.t003

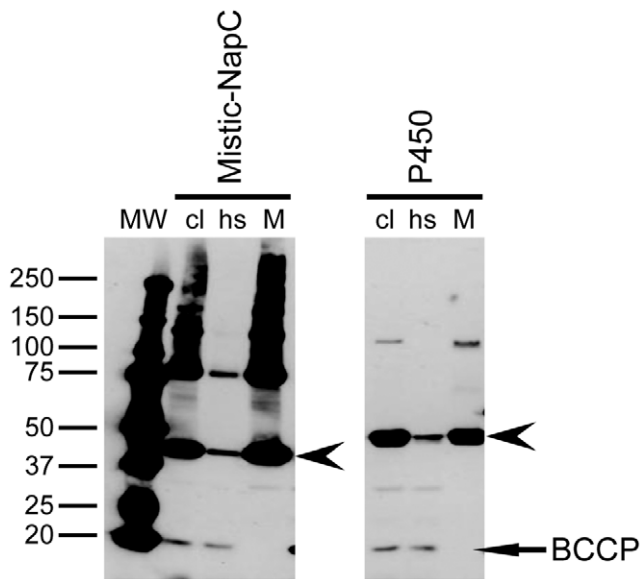


Figure 2. Isolation of the membrane fraction from *E. coli* cells. In this case extracts of *E. coli* cells overexpressing either the Mistic fusion of the protein NapC or the protein P450. cl stands for “cleared lysate” corresponding to the supernatant recovered after centrifugation, at 20,000 *g*, of the cell lysate. hs and M stand for “high-speed supernatant” and “membrane fraction”, respectively, corresponding to the supernatant and the resuspended pellet recovered after ultracentrifugation, at 100,000 *g*, of the “cleared lysate”. Arrows point out the different target proteins and the endogenous *E. coli* biotinylated protein BCCP. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA).

doi:10.1371/journal.pone.0029191.g002

bc1 and to complement a cytochrome *bc1* null mutant (data not shown). Even more surprising was the lack of production of homologous NapC, the tetraheme electron donor of the periplasmic nitrate reductase, NapAB. To test whether this was due to the use of the Gateway cloning approach, we cloned the *napC* gene in the pBBR1MCS2 vector with the *pu*c promoter using traditional cloning methods (Restriction enzyme and Ligase). This plasmid was able to restore the nitrate reductase activity in a *R. sphaeroides napC* null mutant, thus demonstrating effective expression of active NapC. This indicates that, with the Gateway

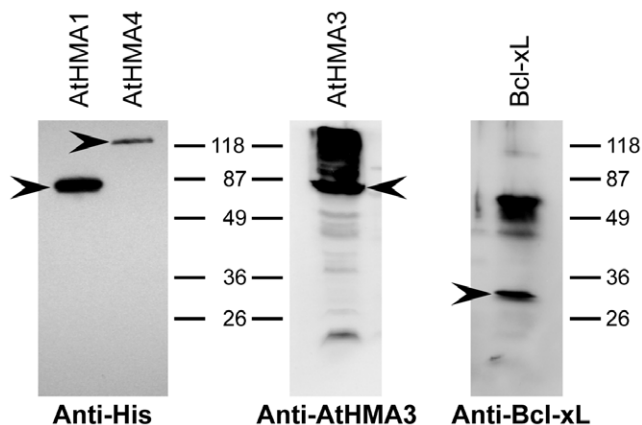


Figure 3. Particular cases of proteins detected in western blots using specific antibodies. For the detection of AthMA4 by anti-His antibodies, AthMA1 was also added on the blot as a positive control. Arrows point out the different target proteins.

doi:10.1371/journal.pone.0029191.g003

expression system, the additional amino acids encoded by the *attB1* sequence (Table 2) may be a source of problems in this host. Alternatively, as we did not perform functional studies with the Gateway clones, the absence of protein detected by western blot could simply be the result of problems with *Strep*-tag II detection, as mentioned above for other proteins, rather than low expression or absence of product in this system. Whatever the case, this host is not an ideal candidate for Gateway-based protein expression of MPs in the conditions used in this study.

Expression in *A. thaliana*. In *A. thaliana*, 50% of the proteins tested were produced (Table 3, for a representative western blot, see Figure 1D). Surprisingly, not all the proteins originating from Arabidopsis were successfully expressed. This may be due to a detection problem, as discussed above. Alternatively, it may stem from silencing of the transgene by the host cells. For example, all the plants carrying the *LPR1* transgene displayed a typical *lpr1* mutant phenotype, indicating that the endogenous *LPR1* gene was also silenced (data not shown). The ubiquitous presence of biotinylated proteins in *A. thaliana* made it necessary to adapt conditions for the western blot analysis, such as by the use of specific blocking buffer and/or avidin prior to conjugate incubation, as described by Witte et al. [64]. As several Arabidopsis proteins were tested in this study, in some cases we checked whether the recombinant protein was correctly targeted to its native location (membrane). For instance, AAC is an ADP/ATP transporter located in the inner mitochondrial membrane. Mitochondrial membranes were purified from the leaves of transgenic Arabidopsis plants overexpressing AAC using two different isolation procedures [89,90]. The presence of the recombinant protein was assessed by western blot using the *Strep*-Tactin HRP conjugate (Figure 4). While we found the procedure described by Brugière et al. [89] to be more efficient, in both cases the protein was enriched in the mitochondrial fraction compared to crude cell membranes, demonstrating efficient targeting of the recombinant protein to the organelle and indicating that the *Strep*-tag II did not interfere with its insertion into the membrane.

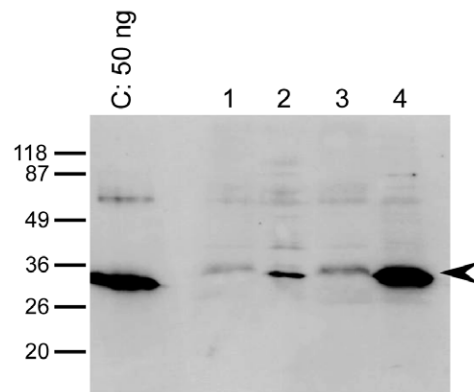


Figure 4. Homologous production of AAC in *A. thaliana* and presence of the recombinant protein in its original organelle, revealed by western blot. Mitochondria were isolated and enriched, from the leaves of 8 weeks old heterozygous *Arabidopsis* plants overexpressing the protein AAC, according to two isolation methods described by Keech et al. (lane 2) or by Brugière et al. (lane 4). Lanes 1 & 3 total membrane extracts before the mitochondria isolation treatments. C: *Strep*-tag II control protein loaded at 50 ng. The arrow points out the protein AAC. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA).

doi:10.1371/journal.pone.0029191.g004

Expression in *N. benthamiana*. Seven proteins were successfully expressed in *N. benthamiana* (Table 3), as shown for five of them in Figure 1E. These proteins are mainly peripheral proteins or proteins with only one predicted TM domain. A single protein with more than one predicted TM (UCP1 = 6 TM) was successfully produced. ceQORH was particularly well expressed in this system, to levels detectable on Coomassie gels (data not shown). All but one (PHF) of the proteins detected in *N. benthamiana* were also produced in *A. thaliana*, indicating that both hosts can be used almost equally successfully.

During optimisation of our experimental set-up for the overexpression of MPs in *N. benthamiana*, we observed that two conditions significantly influenced the levels and/or stability of several recombinant proteins: the growth stage of the plant, and the light intensity in the growth chamber. The expression levels of the recombinant MreC and ceQORH, in variations on the above-cited growth conditions, are shown in Figure 5 as examples of the effects of these parameters. Both MreC and ceQORH accumulate more in older plants than in young ones. Increased accumulation was also observed when young plants were grown under low light. Although we did not perform any further experiments to elucidate this phenomenon, it is possible that, since high light can induce oxidative stress in *N. benthamiana* [91], this may depress protein synthesis and/or accumulation.

Expression in Sf9 cells. As mentioned above, recombinant bacmids were obtained for 18 genes. A first set of expression tests was analysed by western blot on total extracts from cells infected with these bacmids (see an example in Figure 1F). Twelve proteins could be detected in whole-cell extracts, and their expression was also analysed on membrane vesicles (Figure 1G). The other 6 proteins (FtsX, MraY, AtHMA4, AtHMA3, $\alpha 2\delta$ subunit, and CXCR4) were undetectable. Among the proteins expressed, AAC showed variable results between expression experiments, and even over time with the same sample. This suggests significant protein instability. NapC and LPR1, although expressed were mainly present in the unsolubilised material (see Materials and Methods); while PHT1;4, NTT1 and CCR5 all migrated at very high molecular weights, suggesting that they were aggregated (see Figure 1F for PHT1;4). Some of the proteins produced in this host were difficult to quantify for reasons including: very high background staining on western blots with membrane vesicles (NapC); poor transfer of aggregated protein forms (PHT1;4, NTT1 and CCR5), and; extremely low presence in the purified

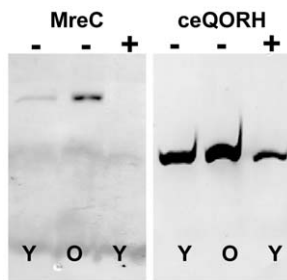


Figure 5. Effect of plant age and light intensity on the expression of MreC and ceQORH in *N. benthamiana*. The *N. benthamiana* plants were grown under low light (60–120 μ E) or high light (240 μ E) before the infiltration with *Agrobacterium*. The young plants had 4 to 6 leaves whereas the old plants started to blossom. The membranes were then extracted and 6.8 μ g of total proteins were loaded on a gel and western-blotted. Y: young plant; O: old plant; (–): light intensity of 60–120 μ E; (+): light intensity of 240 μ E. The membrane was probed with the *Strep*-Tactin HRP conjugate (IBA). doi:10.1371/journal.pone.0029191.g005

membrane fraction (LPR1). Taking all this into consideration, eight out of 18 proteins were correctly expressed at levels ranging from around 10–20 μ g/L of culture, for AtHMA1, NapC and MreC, to 330 μ g/L, for ceQORH; Bcl-xL, PHF, UCP1 and AAC were produced at intermediate levels (Table 3). Although these levels are quite low compared to those produced in bacterial cultures, as mentioned above insect cells have certain advantages when expressing eukaryotic proteins, such as the ability to insert post-translational modifications and disulfide bridges. Thus, in functional tests, these cells might be a better choice.

Discussion

Gateway vectors for cloning and expression

The cloning strategy chosen for this project, based on Gateway technology, enabled us to obtain expression vectors for the different systems in a convenient and very efficient manner (over 99% success). Adaptation of the manufacturer's protocol by an 8-fold reduction of the volumes and quantities of the components used in the BP and LR reactions also enabled us to significantly reduce the cost of the cloning with a maintained efficiency. However, Gateway does present some disadvantages; in particular, the specific *attB* recombination sites used for cloning introduce additional amino sequences at the *N*-terminus of the recombinant proteins. Because we decided to use the same “entry” clones for all the expression systems, for expression in bacterial systems, the Ribosome Binding Site (RBS) necessary for the translation as well as the initiation codon had to be present in the destination vector upstream of the *attR1* sequence. After the LR Gateway reaction, the expression vector codes for a protein that contains 12 to 18 additional residues at its *N*-terminus (Table 2). Although short, this additional sequence has a net charge which could interfere with membrane insertion of the target proteins. A previous study demonstrated that shortened *att* recombination sites increased the success rate for MP expression in *E. coli* [72]. However, the influence of these extensions appears to be variable, depending on the topology of the tested proteins [92]. Indeed, in this study expression in *E. coli* was highly successful using this strategy (15 proteins out of 20 detected in isolated membranes). In a previous study, we showed that the presence or absence of these sites did not affect the level of MP production in *L. lactis* [54]. However, the presence of these additional residues could affect expression in other bacterial hosts and perhaps explains the lower rate of success in *R. sphaeroides*. In mammalian cells, adding this extra sequence at the *N*-terminus of NIS protein has quite a negative effect, worse than the absence of a Kozak consensus sequence (data not shown). The addition of a Flag-tag epitope to the *N*-terminus of NIS also hampers its expression (yields, maturation) [93]. Because of these potential problems with protein expression, the constructs for expression in insect cells were designed not to contain the *att* sequence within the expressed protein. Among recombinatorial cloning methods, only MAGIC [94] and In-Fusion [95] enable seamless cloning, but these two methods require independent PCR products for every new construct and are thus not readily compatible with high-throughput approaches. A recent work by Geertsma and Dutzler [20] presents an elegant new system termed fragment exchange (FX) cloning, which enables subcloning into multiple expression vectors and introduces only a single amino acid to either side of the protein. FX cloning will most probably become very popular in a near future, but for the time being, no compatible vectors are yet commercially available and plasmids need to be constructed and adapted to the technique. To conclude on the cloning strategy, given the efficiency of the cloning step and the number of ORFeome projects (and thus the huge number of

readily available entry clones) for which Gateway technology has been successfully exploited (for some recent examples see [24–27]), we recommend its use when cloning a large number of target genes in various vectors.

Strep-tag II to reveal protein expression

In this study, the 20 proteins were labelled with a *Strep*-tag II sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys). This tag binds strongly to an engineered streptavidin derivative called *Strep*-Tactin. It enables fast and simple one-step purification, and is compatible with a wide range of detergents commonly used for the solubilisation of membrane proteins [96]. Comparison of a range of affinity tags to purify recombinant proteins from various cell types also revealed that the *Strep*-tag offered the best compromise in terms of purity and costs [97]. In our case, the tag was mainly used for target detection in western blotting, to determine the expression level of the proteins. The main reason why we chose *Strep*-tag II over the more commonly used His₆-tag is that a large number of plant proteins contain natural polyhistidine sequences that could lead to false positive responses with anti-His antibodies. This would also hinder the detection of poorly expressed proteins. In addition, commercial anti-*Strep* HRP-conjugates or antibodies were available, together with protein standards for blot calibration in all the laboratories involved in this project. Despite all these arguments in favour of *Strep*-tag II, analysis of overexpressed proteins in *A. thaliana* and Sf9 cells was difficult because a large number of endogenous biotinylated proteins were present and revealed by the *Strep*-Tactin conjugates used for detection in western blots. In these systems, additional blocking steps were required, to saturate the biotinylated proteins with avidin. In other systems, however, these endogenous biotinylated proteins were an advantage. For example, in *E. coli*, the biotinylated BCCP protein was used to control for membrane isolation (Figure 2) while, in *L. lactis*, several endogenous membrane proteins are biotinylated and could be used as protein loading controls for western blots (Figure 1B).

The topology of some proteins in this study is still unknown, and could be modified when they are produced in heterologous systems. We could not therefore predict which protein extremity would be cytoplasmic. This was another reason to avoid using polyhistidine tags, which can be positively charged at physiological pHs. A positive charge is not theoretically favourable to insertion in, or passage through, membranes. In contrast, the *Strep*-tag II is neutral [92]. Despite this, all the constructs expressed in *E. coli* had an *N*-terminal His₆-tag and this had no apparent deleterious effect given the high success rate and protein yields in this host. However, it is possible that these yields could have been further improved if the tag had not been included.

We do not know whether the failed detections in western blots were due to an absence of protein expression, or to loss or inaccessibility of the *Strep*-tag II. In several cases (FtsX, AtHMA3, AtHMA4, Bcl-xL) the protein was detected using other antigenic epitopes (anti-His or protein-specific antibodies) (Figure 3). Therefore, it should be kept in mind that failed detection does not necessarily indicate failed expression. To avoid this type of problem, it is advisable to use different methods of detection, or to modify the protein construct, for example, by moving the tag to the other extremity, or adding extra linkers between the tag and the protein. In another study, modified constructs of AtHMA3 and AtHMA4 which placed the *Strep*-tag II at the *C*-terminus were engineered for expression in *L. lactis*. These proteins were detected normally using *Strep*-Tactin [54]. However, in this parallel screening procedure, it was necessary to limit the number of constructs to be tested (already 120) by making

choices, these may cause our results to appear poorer than they are in reality.

Influence of protein properties on expression

Six expression systems, three prokaryotic (*E. coli*, *L. lactis* and *R. sphaeroides*) and three eukaryotic (*A. thaliana*, *N. benthamiana* and Sf9 cells), were evaluated for their ability to overexpress a set of 20 MPs in this study. Besides their scientific importance, the different proteins were selected to cover a broad range of protein families, source organisms, topologies and functions. Of the 20 MPs, 17 (85%) were produced in at least one of the expression hosts; at levels compatible with further functional and even structural studies in some cases (70% >1 mg/L). However, as shown in Table 3, the expression results were extremely variable. Proteins with a large number of predicted TMs or a large MW were generally less successful (Figure 6A). However, the trend in individual systems sometimes differed from the global picture (Figures S2 and S3). All the peripheral, and most IMPs containing between one and six predicted transmembrane helices, were successfully overexpressed. However, for IMPs with higher numbers of TMs, the outcome was more random, and these were often not expressed at all, which is consistent with previous studies [98,99]. The size of the proteins is expected to affect their expression, as smaller proteins necessarily contain fewer TMs, and in our selection all seven proteins of less than 35 kDa contain a maximum of six predicted TMs. These two parameters could explain why some proteins were (or were not) expressed in all the systems tested. For instance, ceQORH was produced in high quantities in all systems. It is a peripheral, rather hydrophilic protein, and, as previously demonstrated [54,100,101], only interacts with the membrane through electrostatic interactions. Bcl-xL, contains only one predicted TM, and was also well expressed. On the other hand, among the proteins not expressed, “ α 2 δ subunit” has only one predicted TM but a molecular weight over 120 kDa, while MraY and NIS are predicted to have 10 and 13 TMs respectively.

The organism of origin of the protein (including eukaryote versus prokaryote) did not appear to have a significant influence on the efficiency of expression (Figure 6B). Heterologous expression was often successful, and homologous expression sometimes failed, as in the cases of *R. sphaeroides* and *A. thaliana*.

When working with large and/or highly hydrophobic MPs, none of the hosts tested really stood out, with maybe a small exception for *L. lactis*, but the statistics are too small to really conclude. With these proteins one should thus expect that successful expression will require more effort and should focus on optimising the expression conditions (level and time of induction, temperature, additive in the growth medium like glycerol or sucrose...). For example, in the work by Wagner et al. [37] describing the *E. coli* strain Lemo21(DE3), the authors clearly demonstrated that variations of growth conditions could significantly impact on the levels of expression. In a parallel screening, choices need to be made and the number of growth conditions tested is limited. In this study, prior to the screen, several conditions were tested in the different hosts, with one or more proteins to define standard expression conditions that were finally used for all proteins. These conditions were therefore most probably not optimal in all cases. Then, after identification of the most suitable expression system for one candidate protein, it may also be required to further optimise specific expression conditions.

Mistic, a boost in *E. coli* expression

Overall, the best expression results in terms of success rate and protein amounts were obtained with *E. coli* (protein yields >1 mg

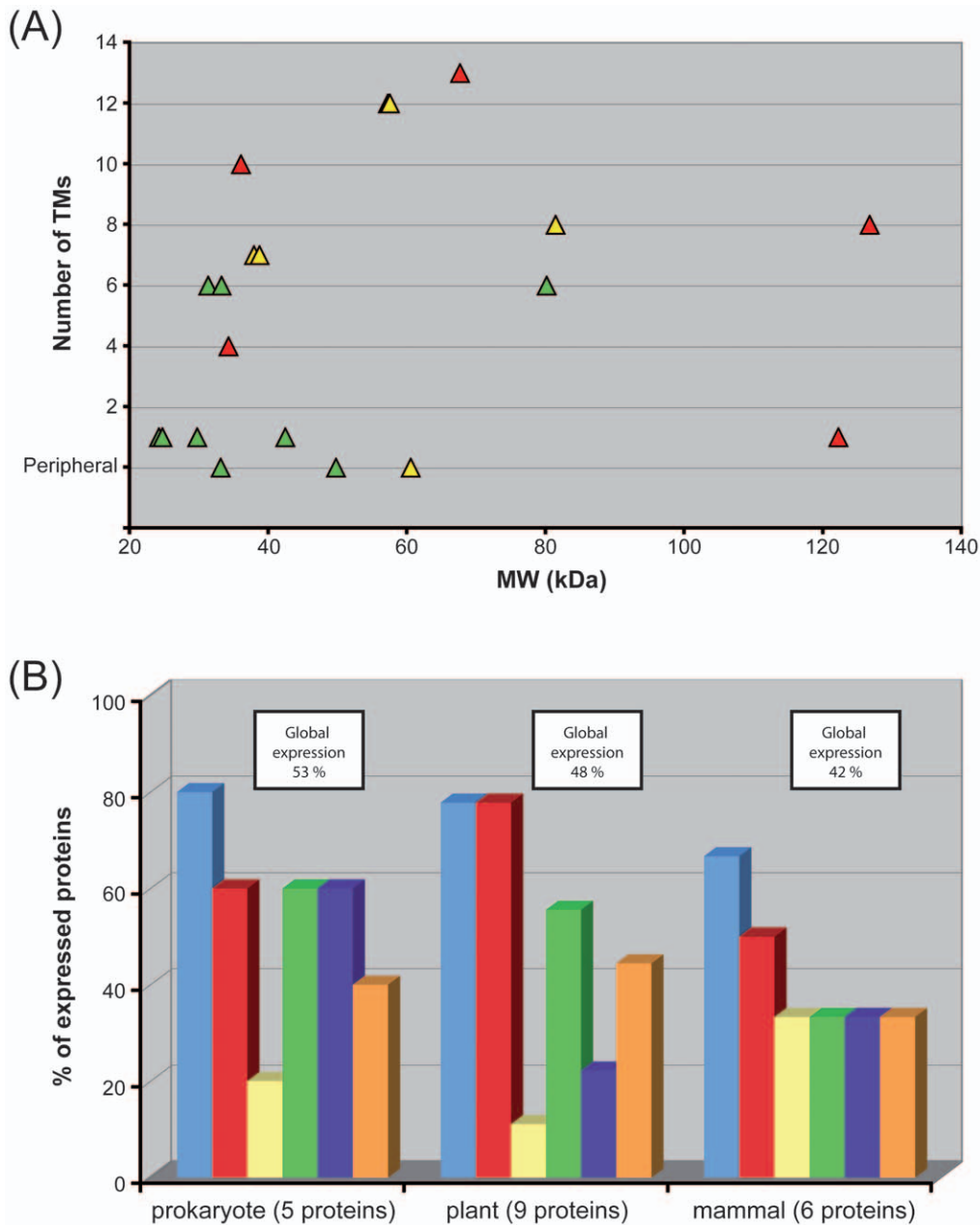


Figure 6. Influence of protein properties on expression. (A) Influence of the protein size and the number of TMs on the expression success rate. The triangles represent each proteins and their colour the success with which they were expressed in the different expression systems. Red=protein expressed in none or only one of the expression systems. Yellow=protein expressed in two or three of the expression systems. Green=protein expressed in four to six of the expression systems. (B) **Influence of the origin of proteins on the expression in the different systems.** The bars represent the percentage of positively expressed proteins in each expression host for a given category. Light blue: *E. coli*; Red: *L. lactis*; Yellow: *R. Sphaeroides*; Green: *A. thaliana*; Dark blue: *N. benthamiana*, Orange: insect cells. Global expression represents the percentage of positively expressed proteins in all expression hosts for a given category. doi:10.1371/journal.pone.0029191.g006

protein/L culture for 14 proteins). Many possible fusion tags or proteins such as GST, MBP, NusA or Mystic [30,39] have been described in the literature, however no system is perfect and can solve all the problems. Mystic seems to act like a signal sequence that targets the proteins to the inner membrane of *E. coli*. Its properties and the good results obtained by others with Mystic fusions [39,40,43,45] determined our choice to use it in this study. In our hands too, Mystic had a positive effect and significantly increased the yields obtained for all the 12 IMPs produced in *E.*

coli. In a recent report, Leviatan et al. [102] presented the use of two hydrophilic bacterial proteins, YaiN and YbeL, for membrane targeting of cargo proteins and compared these fusions tags with Mystic. The yields obtained were equivalent, or even better in one case than with Mystic, but the approach used was combinatorial, testing 8 different constructs for each target protein to find the best combination. This was not suitable for use here. In any case, given the good results obtained here and elsewhere with Mystic, we consider that the strategy applied here was more efficient and less

laborious. Interestingly, when Mystic was fused to the three peripheral MPs in this study, it had a negative effect, significantly reducing the yields of protein recovered with the plasma membrane. This leads us to hypothesise that Mystic might not just address the proteins to the membrane, but actually force them into the lipid bilayer. Since peripheral proteins are quite hydrophilic and interact with the membrane mainly through electrostatic interactions, being forced into the hydrophobic environment of the membrane, because of fusion to Mystic, could thus be unfavourable.

L. lactis, an efficient and valuable alternative to *E. coli*

In terms of success rate and protein yields, *L. lactis* also gave good results. Together with *E. coli*, it proved to be an adequate system for the expression of *A. thaliana* MPs (Figure 6B) but, in this case, without requiring fusion to Mystic-type tags. We believe that this is due to the fact that both *L. lactis* and *A. thaliana* have very similar GC-content in their genome, as well as similar amounts and types of glycolipids in their membranes [54]. Moreover, even produced at very low levels in *L. lactis* (around 0.2% of TMP), some recombinant MPs were active in this system [54]. The difference in protein yields obtained with the two bacteria could be explained by the limited capacity of *L. lactis* to accumulate branched-chain amino acids, thus limiting overexpression in this host [103]. *L. lactis* appears to be complementary to *E. coli*: the 17 MPs expressed (including all the MPs from *A. thaliana*), could be produced in at least one of these two bacteria. The three proteins that failed to express in these bacteria (MraY, $\alpha 2\delta$ subunit and NIS) also failed in all the other systems tested. For the less hydrophobic MPs, the two systems were equivalent: all the peripheral proteins and those containing a single predicted TM were produced in both bacteria. In contrast, MPs with higher numbers of TMs (UCP1, AAC and PHT1;4) were only produced in *E. coli*, whereas AtHMA3 and AtHMA4, as well as AtHMA1 (without the Mystic tag), were only detectable in *L. lactis*. Thus, *L. lactis* is an efficient expression system and it should be considered as an alternative when overexpression fails in *E. coli*.

The benefits of homologous and eukaryotic expression

With heterologous protein expression the recombinant protein produced does not always truly resemble the native protein. Conditions that produce the largest amount of protein do not necessarily generate functional proteins [7,29,104–106] and, in many cases, proteins are only functional after post-translational modification, such as through glycosylation and formation of disulfide bonds. Although several prokaryotic strains have been developed to overcome some of these hurdles (e.g. *E. coli* trxB mutants or *E. coli* glycosylation enabled mutants [107]), eukaryotic systems are sometimes necessary. Three eukaryotic hosts were selected in this project (*A. thaliana*, *N. benthamiana*, Sf9). *A. thaliana* enabled homologous expression of several proteins (9 out of 20 originate from this organism) and it allowed us to show that the protein AAC was correctly targeted to mitochondria (Figure 4). Different approaches to protein production were used in the two plant systems. In *A. thaliana* stable cell lines were generated, while in *N. benthamiana* transient agro-infiltration was used. Equivalent yields were obtained for proteins expressed in both systems, but more targets were produced in *A. thaliana*. However, the faster turnaround time with transient agro-infiltration is a great advantage, and facilitates screening for optimal production conditions (e.g. light intensity). In Sf9 cells, eight proteins were correctly overexpressed, and all five well-expressed proteins were of eukaryotic origin. Nonetheless the rat $\alpha 2\delta$ subunit and human CXCR4 proteins were not expressed at all, and human CCR5 was

produced in an aggregated form. This suggests that insect cells are not necessarily able to handle mammalian proteins properly.

Based on these results, and those discussed above for bacterial protein expression, prokaryotic and eukaryotic systems are complementary. Even though compared to *E. coli*, the expression levels in the other systems are generally lower, a significant number of targets could however be expressed, proving that *L. lactis*, *R. sphaeroides*, *A. thaliana* or *N. benthamiana* are valuable alternatives to more conventional expression hosts and can be considered for expression of membrane proteins. Most of these systems can be rather easily implemented in a laboratory. *L. lactis* and *R. sphaeroides* systems require similar handling procedures and instrumentations that are used for *E. coli*. Many commercial alternatives are available for expression in insect cells and the protocols are well established. However, insect cells culture medium is three to four times more expensive than *E. coli* medium. The agroinfiltration procedure required for expression in *N. benthamiana* is rather simple and well described in the literature [62,108]. Plant culturing requires an illuminated growth chamber, but doesn't cause major problems and a few days training in an expert laboratory should be sufficient to learn the necessary techniques. On the other hand, the procedures described here for expression in *A. thaliana* and the obtention of stable lines requires much more time and expertise, and one should consider a collaboration with an appropriate laboratory.

Protein activity also needs to be evaluated to choose the appropriate expression system for further functional studies.

Towards functional characterisation of the recombinant proteins produced

Overexpression of membrane proteins is a challenge in itself. Many groups aiming to characterise one of these difficult membrane proteins, must first test the efficiency of various expression systems before any functional characterisation can be performed. In this study, we intentionally focused on the expression yields of the 20 proteins tested, and on the ability of the six different hosts to produce our target proteins. Indeed, a good expression yield is a prerequisite for most biochemical and biophysical experiments as it more or less determines the final purity, concentration in solution, amount of protein available, cost of production, etc. In addition, in some cases, producing enough protein, whether functional or not, is a goal in itself. These proteins can be used to develop precious tools such as, for example, antibodies that usually work better than antibodies raised against shorter synthetic peptides. In this study, we also chose to study quite a diverse range of proteins (size, hydrophobicity, origin) with different known or predicted activities, even though, for many of them, no functional assays had yet been performed. For most of these proteins, functional characterisation represents a stand-alone project and it would not be possible to perform it for the 120 host/protein combinations described here.

In parallel studies, the functionality of some of the proteins produced during the present study has been analysed in one or more expression hosts. Thus, ceQORH proved to be active when produced in either *E. coli* or *L. lactis*, but its specific activity was higher in *L. lactis*, and the protein could be purified without difficulty using either the His-tag or the *Strep*-tag II [54]. The activity of AAC, produced in *E. coli*, was also tested, it was found to transport radioactive ATP, and was also sensitive to variations in NaCl concentration [77]. The functionality of NTT1 in the two bacterial systems (*E. coli* and *L. lactis*) and the influence of fusion to Mystic fusion (in *E. coli*) were also analysed. This protein was active in both in *E. coli* and *L. lactis* [54,76]. When produced in *E. coli*, the protein could be purified to homogeneity and, by optimising

growth conditions, the Mystic fusion led to a 16-fold increase in protein yield. Although the transport activity of the fusion protein was impaired in *E. coli* membranes, cleavage of the Mystic moiety *in vivo* delivered a functional transporter, proving that the Mystic strategy is a valuable approach [76]. Last but not least, of the four plant heavy metal ATPases tested in *L. lactis*, two were efficiently produced and purified from this host [54]. This system was also recently used to perform the first biochemical characterisation of a plant copper ATPase [109].

Conclusion

This study compared several approaches for the overexpression of a variety of recombinant membrane proteins in six expression hosts. Our success rate was high, with 17 out of 20 proteins tested expressed in at least one system, but large and very hydrophobic proteins remained however hard to express whatever the host used. It is therefore important to keep in mind that optimisation of expression conditions can greatly improve the yields of protein produced and it should be thoroughly undertaken after selecting the host. If a first screen fails to identify an appropriate expression system for a given target protein, optimisation of growth conditions could then be attempted in one of the most successful hosts presented here, e.g. *E. coli* or *L. lactis*.

The different systems present various advantages. Very good yields (several mg protein/L culture) could be obtained with *E. coli*, especially when Mystic fusions were used. But it is important to gather topological information on the target before fusing it to Mystic, as fusion with peripheral MPs was detrimental for the three proteins tested. Therefore, Mystic should only be considered as an aide to correct membrane targeting of IMPs. *L. lactis* was an appropriate host for the expression of plant MPs, as well as a good alternative to *E. coli* when expression fails in this system. We demonstrated homologous expression in *A. thaliana* to be beneficial, as it allows the investigation of subcellular targeting (as for ACC here). The baculovirus system was less efficient than *E. coli* or *L. lactis*, both in terms of number of expressed proteins and quantity of protein produced, but is nonetheless a good system (eight proteins expressed, five with a reasonable yield, out of 18 successfully cloned candidates). Moreover, insect cells appear to be more useful for the production of functional proteins with specific post-translational modifications, as are the other eukaryotic hosts: *A. thaliana* and *N. benthamiana*.

In this work, we have developed a certain number of methods to increase the throughput and rationalise the screening of MP overexpression in both prokaryotic and eukaryotic systems. This systematic approach was efficient since less conventional expression systems proved to be valuable alternatives and, as discussed above, besides *A. thaliana* all other systems can be rather easily implemented in other laboratories. We believe that the evaluation of expression systems presented here is a useful starting guide for

biologists aiming to produce their favourite membrane protein in amounts compatible with further biochemical and structural characterisation.

Supporting Information

Figure S1 The successive cloning steps in the Gateway technology. The Gateway Technology uses the λ recombination system to facilitate transfer of heterologous DNA sequences (flanked by modified *att* sites) between vectors. BP Reaction: Facilitates recombination of an *attB*-PCR product with an *attP*-containing donor vector to create an *attL*-containing entry clone. This reaction is catalysed by BP Clonase. LR Reaction: Facilitates recombination of an *attL*-containing entry clone with an *attR*-containing destination vector to create an *attB*-containing expression clone. This reaction is catalysed by LR Clonase. (TIF)

Figure S2 Influence of the number of TMs on the expression in the different systems. The bars represent the percentage of positively expressed proteins in each expression host for a given category. Light blue: *E. coli*; Red: *L. lactis*; Yellow: *R. Sphaeroides*; Green: *A. thaliana*; Dark blue: *N. benthamiana*, Orange: insect cells. Global expression represents the percentage of positively expressed proteins in all expression hosts for a given category. (TIF)

Figure S3 Influence of the protein size on the expression in the different systems. The bars represent the percentage of positively expressed proteins in each expression host for a given category. Light blue: *E. coli*; Red: *L. lactis*; Yellow: *R. Sphaeroides*; Green: *A. thaliana*; Dark blue: *N. benthamiana*, Orange: insect cells. Global expression represents the percentage of positively expressed proteins in all expression hosts for a given category. (TIF)

Acknowledgments

We thank Marjolaine Noireclerc-Savoie and Benoit Gallet (RoBioMol platform, Institut de Biologie Structurale J.P Ebel, Grenoble, France) for their technical advices, as well as Martin Kieffer (University of Leeds, UK) and François Parcy (iRTSV/LPCV, Grenoble, France) for the pAlligator3 vector. We thank Carlo Petosa for comments on the manuscript.

Author Contributions

Conceived and designed the experiments: JJ DP ED MS TD EPP TV NR. Performed the experiments: FB AFB NP SD PH SB JBR DS DSB. Analyzed the data: FB AFB NP SD PH SB JBR DS DSB PR DP ED MS TD EPP TV NR. Contributed reagents/materials/analysis tools: DSB PR DP ED MS TD EPP TV NR. Wrote the paper: FB AFB ED MS TD NR.

References

- Lacapère J-J, Pebay-Peyroula E, Neumann J-M, Etchebest C (2007) Determining membrane protein structures: still a challenge! Trends Biochem Sci 32: 259–270. doi:10.1016/j.tibs.2007.04.001.
- von Heijne G (2007) The membrane protein universe: what's out there and why bother? J Intern Med 261: 543–557. doi:10.1111/j.1365-2796.2007.01792.x.
- McLuskey K, Roszak AW, Zhu Y, Isaacs NW (2010) Crystal structures of all-alpha type membrane proteins. Eur Biophys J 39: 723–755. doi:10.1007/s00249-009-0546-6.
- Wagner S, Bader ML, Drew D, de Gier J-W (2006) Rationalizing membrane protein overexpression. Trends Biotechnol 24: 364–371. doi:10.1016/j.tibtech.2006.06.008.
- White SH (2009) Biophysical dissection of membrane proteins. Nature 459: 344–346. doi:10.1038/nature08142.
- Tate CG (2001) Overexpression of mammalian integral membrane proteins for structural studies. FEBS Lett 504: 94–98.
- Tate CG, Haase J, Baker C, Boorsma M, Magnani F, et al. (2003) Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. Biochim Biophys Acta 1610: 141–153.
- Reeves PJ, Kim J-M, Khorana HG (2002) Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. Proc Natl Acad Sci USA 99: 13413–13418. doi:10.1073/pnas.212519199.
- Kunji ERS, Slotboom D-J, Poolman B (2003) Lactococcus lactis as host for overproduction of functional membrane proteins. Biochim Biophys Acta 1610: 97–108.
- Kunji ERS, Chan KW, Slotboom DJ, Floyd S, O'Connor R, et al. (2005) Eukaryotic membrane protein overproduction in Lactococcus lactis. Curr Opin Biotechnol 16: 546–551. doi:10.1016/j.copbio.2005.08.006.
- Lundstrom K (2003) Semliki Forest virus vectors for rapid and high-level expression of integral membrane proteins. Biochim Biophys Acta 1610: 90–96.

12. Lundstrom K, Wagner R, Reinhart C, Desmyter A, Cherouati N, et al. (2006) Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. *J Struct Funct Genomics* 7: 77–91. doi:10.1007/s10969-006-9011-2.
13. Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260: 289–298. doi:10.1006/jmbi.1996.0399.
14. Hartley JL (2006) Cloning technologies for protein expression and purification. *Curr Opin Biotechnol* 17: 359–366. doi:10.1016/j.copbio.2006.06.011.
15. Hunt I (2005) From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expr Purif* 40: 1–22. doi:10.1016/j.pep.2004.10.018.
16. Koehn J, Hunt I (2009) High-Throughput Protein Production (HTPP): a review of enabling technologies to expedite protein production. *Methods Mol Biol* 498: 1–18. doi:10.1007/978-1-59745-196-3_1.
17. Lesley SA (2009) Parallel methods for expression and purification. *Meth Enzymol* 463: 767–785. doi:10.1016/S0076-6879(09)63041-X.
18. Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorton MA, et al. (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Meth Enzymol* 328: 575–592.
19. Colwill K, Wells CD, Elder K, Goudreault M, Hersi K, et al. (2006) Modification of the Creator recombination system for proteomics applications—improved expression by addition of splice sites. *BMC Biotechnol* 6: 13. doi:10.1186/1472-6750-6-13.
20. Geertsma ER, Dutzler R (2011) A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* 50: 3272–3278. doi:10.1021/bi200178z.
21. Eshaghi S, Hedrén M, Nasser MIA, Hammarberg T, Thornell A, et al. (2005) An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. *Protein Sci* 14: 676–683. doi:10.1110/ps.041127005.
22. Bechtel S, Rosenfelder H, Duda A, Schmidt CP, Ernst U, et al. (2007) The full-ORF clone resource of the German cDNA Consortium. *BMC Genomics* 8: 399. doi:10.1186/1471-2164-8-399.
23. Yashiroda Y, Matsuyama A, Yoshida M (2008) New insights into chemical biology from ORFeome libraries. *Curr Opin Chem Biol* 12: 55–59. doi:10.1016/j.cbpa.2008.01.024.
24. Dahlroth S-L, Lieu V, Haas J, Nordlund P (2009) Screening colonies of pooled ORFeomes (SCOOP): a rapid and efficient strategy for expression screening ORFeomes in *Escherichia coli*. *Protein Expr Purif* 68: 121–127. doi:10.1016/j.pep.2009.07.010.
25. Matsuyama A, Yoshida M (2009) Systematic cloning of an ORFeome using the Gateway system. *Methods Mol Biol* 577: 11–24. doi:10.1007/978-1-60761-232-2_2.
26. Pellet J, Tafforeau L, Lucas-Hourani M, Navratil V, Meyniel L, et al. (2010) ViralORFeome: an integrated database to generate a versatile collection of viral ORFs. *Nucleic Acids Res* 38: D371–D378. doi:10.1093/nar/gkp1000.
27. Rajagopala SV, Yamamoto N, Zweifel AE, Nakamichi T, Huang H-K, et al. (2010) The *Escherichia coli* K-12 ORFeome: a resource for comparative molecular microbiology. *BMC Genomics* 11: 470. doi:10.1186/1471-2164-11-470.
28. Hartley JL, Temple GF, Brasch MA (2000) DNA cloning using in vitro site-specific recombination. *Genome Res* 10: 1788–1795.
29. Midgett CR, Madden DR (2007) Breaking the bottleneck: eukaryotic membrane protein expression for high-resolution structural studies. *J Struct Biol* 160: 265–274. doi:10.1016/j.jsb.2007.07.001.
30. Junge F, Schneider B, Reckel S, Schwarz D, Dötsch V, et al. (2008) Large-scale production of functional membrane proteins. *Cell Mol Life Sci* 65: 1729–1755. doi:10.1007/s00018-008-8067-5.
31. Freigassner M, Pichler H, Glieder A (2009) Tuning microbial hosts for membrane protein production. *Microb Cell Fact* 8: 69. doi:10.1186/1475-2859-8-69.
32. Sørensen HP (2010) Towards universal systems for recombinant gene expression. *Microb Cell Fact* 9: 27. doi:10.1186/1475-2859-9-27.
33. Pannels V, Kock I, Krijnse-Locker J, Rezagou M, Sinning I (2011) *Drosophila* photoreceptor cells exploited for the production of eukaryotic membrane proteins: receptors, transporters and channels. *PLoS ONE* 6: e18478. doi:10.1371/journal.pone.0018478.
34. Mizutani K, Yoshioka S, Mizutani Y, Iwata S, Mikami B (2011) High-throughput construction of expression system using yeast *Pichia pastoris*, and its application to membrane proteins. *Protein Expr Purif* 77: 1–8. doi:10.1016/j.pep.2010.12.009.
35. Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* 307: 249–264. doi:10.1007/s11010-007-9603-6.
36. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* 72: 211–222. doi:10.1007/s00253-006-0465-8.
37. Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, et al. (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci USA* 105: 14371–14376. doi:10.1073/pnas.0804090105.
38. Drew D, Lerch M, Kunji E, Slotboom D-J, de Gier J-W (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods* 3: 303–313. doi:10.1038/nmeth0406-303.
39. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, et al. (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* 307: 1317–1321. doi:10.1126/science.1106392.
40. Kefala G, Kwiatkowski W, Esquivies L, Maslennikov I, Choe S (2007) Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from *Escherichia coli*. *J Struct Funct Genomics* 8: 167–172. doi:10.1007/s10969-007-9033-4.
41. Liu X, Wu L, Deng G, Li N, Chu X, et al. (2008) Characterization of mitochondrial trifunctional protein and its inactivation study for medicine development. *Biochim Biophys Acta* 1784: 1742–1749. doi:10.1016/j.bbapap.2008.06.018.
42. Dvir H, Choe S (2009) Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs. *Protein Expr Purif* 68: 28–33. doi:10.1016/j.pep.2009.06.007.
43. Nekrasova OV, Wulfson AN, Tikhonov RV, Yakimov SA, Simonova TN, et al. (2010) A new hybrid protein for production of recombinant bacteriorhodopsin in *Escherichia coli*. *J Biotechnol*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20363267>.
44. Blain KY, Kwiatkowski W, Choe S (2010) The Functionally Active Mistic-Fused Histidine Kinase Receptor, EnvZ. *Biochemistry*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20849081>.
45. Petrovskaya LE, Shulga AA, Bocharova OV, Ermolyuk YS, Kryukova EA, et al. (2010) Expression of G-protein coupled receptors in *Escherichia coli* for structural studies. *Biochemistry Mosc* 75: 881–891.
46. Zhou XX, Li WF, Ma GX, Pan YJ (2006) The nisin-controlled gene expression system: construction, application and improvements. *Biotechnol Adv* 24: 285–295. doi:10.1016/j.biotechadv.2005.11.001.
47. Mierau I, Kleerebezem M (2005) 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl Microbiol Biotechnol* 68: 705–717. doi:10.1007/s00253-005-0107-6.
48. Le Loir Y, Azevedo V, Oliveira SC, Freitas DA, Miyoshi A, et al. (2005) Protein secretion in *Lactococcus lactis*: an efficient way to increase the overall heterologous protein production. *Microb Cell Fact* 4: 2. doi:10.1186/1475-2859-4-2.
49. Wells JM, Mercenier A (2008) Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nat Rev Microbiol* 6: 349–362. doi:10.1038/nrmicro1840.
50. del Rio B, Fuente JL, Neves V, Dattwyler R, Seegers JFML, et al. (2010) Platform technology to deliver prophylactic molecules orally: an example using the Class A select agent *Yersinia pestis*. *Vaccine* 28: 6714–6722. doi:10.1016/j.vaccine.2010.07.084.
51. Morello E, Bermúdez-Humarán LG, Llull D, Solé V, Miraglio N, et al. (2008) *Lactococcus lactis*, an efficient cell factory for recombinant protein production and secretion. *J Mol Microbiol Biotechnol* 14: 48–58. doi:10.1159/000106082.
52. Mierau I, Olieman K, Mond J, Smid EJ (2005) Optimization of the *Lactococcus lactis* nisin-controlled gene expression system NICE for industrial applications. *Microb Cell Fact* 4: 16. doi:10.1186/1475-2859-4-16.
53. Steen A, Wiederhold E, Gandhi T, Breiting R, Slotboom DJ (2010) Physiological adaptation of the bacterium *Lactococcus lactis* in response to the production of human CFTR. *Mol Cell Proteomics*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20421633>.
54. Frelet-Barrand A, Boutigny S, Moyet L, Deniaud A, Seigneurin-Berny D, et al. (2010) *Lactococcus lactis*, an alternative system for functional expression of peripheral and intrinsic Arabidopsis membrane proteins. *PLoS ONE* 5: e8746. doi:10.1371/journal.pone.0008746.
55. Linares DM, Geertsma ER, Poolman B (2010) Evolved *Lactococcus lactis* strains for enhanced expression of recombinant membrane proteins. *J Mol Biol* 401: 45–55. doi:10.1016/j.jmb.2010.06.002.
56. Geertsma ER, Poolman B (2007) High-throughput cloning and expression in recalcitrant bacteria. *Nat Methods* 4: 705–707. doi:10.1038/nmeth1073.
57. Douillard FP, Mahony J, Campanacci V, Cambillau C, van Sinderen D (2011) Construction of two *Lactococcus lactis* expression vectors combining the Gateway and the Nisin Controlled Expression systems. *Plasmid*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21807023>. Accessed 24 October 2011.
58. Kiley PJ, Kaplan S (1988) Molecular genetics of photosynthetic membrane biosynthesis in *Rhodobacter sphaeroides*. *Microbiol Rev* 52: 50–69.
59. Laible PD, Scott HN, Henry L, Hanson DK (2004) Towards higher-throughput membrane protein production for structural genomics initiatives. *J Struct Funct Genomics* 5: 167–172. doi:10.1023/B:JSFG.0000029201.33710.46.
60. The Arabidopsis Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815. doi:10.1038/35048692.
61. Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H (2006) Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat Protoc* 1: 641–646. doi:10.1038/nprot.2006.97.
62. Lee MW, Yang Y (2006) Transient expression assay by agroinfiltration of leaves. *Methods Mol Biol* 323: 225–229. doi:10.1385/1-59745-003-0-225.
63. Voimnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J* 33: 949–956.

64. Witte C-P, Noël LD, Gielbert J, Parker JE, Romeis T (2004) Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Mol Biol* 55: 135–147. doi:10.1007/s11103-004-0501-y.
65. Vézina L-P, Faye L, Lerouge P, D'Aoust M-A, Marquet-Blouin E, et al. (2009) Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol J* 7: 442–455. doi:10.1111/j.1467-7652.2009.00414.x.
66. O'Reilly D (1994) *Baculovirus expression vectors: a laboratory manual*. New York: Oxford University Press.
67. Possee RD (1997) Baculoviruses as expression vectors. *Curr Opin Biotechnol* 8: 569–572.
68. Jarvis DL, Finn EE (1995) Biochemical analysis of the N-glycosylation pathway in baculovirus-infected lepidopteran insect cells. *Virology* 212: 500–511. doi:10.1006/viro.1995.1508.
69. Condreay JP, Kost TA (2007) Baculovirus expression vectors for insect and mammalian cells. *Curr Drug Targets* 8: 1126–1131.
70. Sarraemna V, Talmont F, Demange P, Milon A (2003) Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification. *Cell Mol Life Sci* 60: 1529–1546.
71. Dobrovetsky E, Lu ML, Andorn-Broza R, Khutoreskaya G, Bray JE, et al. (2005) High-throughput production of prokaryotic membrane proteins. *J Struct Funct Genomics* 6: 33–50. doi:10.1007/s10969-005-1363-5.
72. Gordon E, Horsefield R, Swarts HGP, de Pont JHHM, Neutze R, et al. (2008) Effective high-throughput overproduction of membrane proteins in *Escherichia coli*. *Protein Expr Purif* 62: 1–8. doi:10.1016/j.pep.2008.07.005.
73. Surade S, Klein M, Stolt-Bergner PC, Muenke C, Roy A, et al. (2006) Comparative analysis and “expression space” coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci* 15: 2178–2189. doi:10.1110/ps.062312706.
74. Laage R, Langosch D (2001) Strategies for prokaryotic expression of eukaryotic membrane proteins. *Traffic* 2: 99–104.
75. Eifler N, Duckely M, Sumanovski LT, Egan TM, Oksche A, et al. (2007) Functional expression of mammalian receptors and membrane channels in different cells. *J Struct Biol* 159: 179–193. doi:10.1016/j.jsb.2007.01.014.
76. Deniaud A, Bernaudat F, Frelet-Barrand A, Juillan-Binard C, Vernet T, et al. (2011) Expression of a chloroplast ATP/ADP transporter in *E. coli* membranes: behind the Mistic strategy. *Biochim Biophys Acta* 1808: 2059–2066. doi:10.1016/j.bbame.2011.04.011.
77. Krammer E-M, Ravaut S, Dehez F, Frelet-Barrand A, Pebay-Peyroula E, et al. (2009) High-chloride concentrations abolish the binding of adenine nucleotides in the mitochondrial ADP/ATP carrier family. *Biophys J* 97: L25–L27. doi:10.1016/j.bpj.2009.08.047.
78. van den Ent F, Löwe J (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods* 67: 67–74. doi:10.1016/j.jbbm.2005.12.008.
79. Holo H, Nes IF (1989) High-Frequency Transformation, by Electroporation, of *Lactococcus lactis* subsp. *cremoris* Grown with Glycine in Osmotically Stabilized Media. *Appl Environ Microbiol* 55: 3119–3123.
80. Frelet-Barrand A, Boutigny S, Kunji ERS, Rolland N (2010) Membrane protein expression in *Lactococcus lactis*. *Methods Mol Biol* 601: 67–85. doi:10.1007/978-1-60761-344-2_5.
81. Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, et al. (1995) Four new derivatives of the broad-host-range cloning vector pBRR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166: 175–176.
82. Prentki P, Krusch HM (1984) In vitro insertional mutagenesis with a selectable DNA fragment. *Gene* 29: 303–313.
83. Benschmih S, To A, Lambert G, Kroj T, Giraudat J, et al. (2004) Analysis of an activated ABI5 allele using a new selection method for transgenic *Arabidopsis* seeds. *FEBS Lett* 561: 127–131. doi:10.1016/S0014-5793(04)00148-6.
84. Wise AA, Liu Z, Binns AN (2006) Three methods for the introduction of foreign DNA into *Agrobacterium*. *Methods Mol Biol* 343: 43–53. doi:10.1385/1-59745-130-4_43.
85. Clayton RK (1960) The induced synthesis of catalase in *Rhodospseudomonas sphaeroides*. *Biochim Biophys Acta* 37: 503–512.
86. Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16: 735–743.
87. Berne C, Pignol D, Lavergne J, Garcia D (2007) CYP201A2, a cytochrome P450 from *Rhodospseudomonas palustris*, plays a key role in the biodegradation of tributyl phosphate. *Appl Microbiol Biotechnol* 77: 135–144. doi:10.1007/s00253-007-1140-4.
88. Sabaty M, Avazeri C, Pignol D, Vermeglio A (2001) Characterization of the reduction of selenate and tellurite by nitrate reductases. *Appl Environ Microbiol* 67: 5122–5126. doi:10.1128/AEM.67.11.5122-5126.2001.
89. Brugière S, Kowalski S, Ferro M, Seigneurin-Berny D, Miras S, et al. (2004) The hydrophobic proteome of mitochondrial membranes from *Arabidopsis* cell suspensions. *Phytochemistry* 65: 1693–1707. doi:10.1016/j.phytochem.2004.03.028.
90. Keech O, Dizengremel P, Gardeström P (2005) Preparation of leaf mitochondria from *Arabidopsis thaliana*. *Physiologia Plantarum* 124: 403–409. doi:10.1111/j.1399-3054.2005.00521.x.
91. Carvalho LC, Santos S, Vilela BJ, Amâncio S (2008) *Solanum lycopersicon* Mill. and *Nicotiana benthamiana* L. under high light show distinct responses to anti-oxidative stress. *J Plant Physiol* 165: 1300–1312. doi:10.1016/j.jplph.2007.04.009.
92. Rahman M, Ismat F, McPherson MJJ, Baldwin SA (2007) Topology-informed strategies for the overexpression and purification of membrane proteins. *Mol Membr Biol* 24: 407–418. doi:10.1080/09687860701243998.
93. Huc-Brandt S, Marcellin D, Graslin F, Averseng O, Bellanger L, et al. (2011) Characterisation of the purified human sodium/iodide symporter reveals that the protein is mainly present in a dimeric form and permits the detailed study of a native C-terminal fragment. *Biochim Biophys Acta* 1808: 65–77. doi:10.1016/j.bbame.2010.08.013.
94. Li MZ, Elledge SJ (2005) MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* 37: 311–319. doi:10.1038/ng1505.
95. Gräslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, et al. (2008) Protein production and purification. *Nat Methods* 5: 135–146. doi:10.1038/nmeth.f.202.
96. Schmidt TGM, Skerra A (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* 2: 1528–1535. doi:10.1038/nprot.2007.209.
97. Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. *Protein Expr Purif* 41: 98–105. doi:10.1016/j.pep.2005.01.019.
98. Korepanova A, Gao FP, Hua Y, Qin H, Nakamoto RK, et al. (2005) Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli*. *Protein Sci* 14: 148–158. doi:10.1110/tp.041022305.
99. Psakis G, Nitschkowski S, Holz C, Kress D, Maestre-Reyna M, et al. (2007) Expression screening of integral membrane proteins from *Helicobacter pylori* 26695. *Protein Sci* 16: 2667–2676. doi:10.1110/ps.073104707.
100. Miras S, Salvi D, Ferro M, Grunwald D, Garin J, et al. (2002) Non-canonical transit peptide for import into the chloroplast. *J Biol Chem* 277: 47770–47778. doi:10.1074/jbc.M207477200.
101. Miras S, Salvi D, Piette L, Seigneurin-Berny D, Grunwald D, et al. (2007) Toc159- and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. *J Biol Chem* 282: 29482–29492. doi:10.1074/jbc.M61112200.
102. Leviatan S, Sawada K, Moriyama Y, Nelson N (2010) Combinatorial method for overexpression of membrane proteins in *Escherichia coli*. *J Biol Chem* 285: 23548–23556. doi:10.1074/jbc.M110.125492.
103. Marreddy RKR, Geertsma ER, Permentier HP, Pinto JPC, Kok J, et al. (2010) Amino acid accumulation limits the overexpression of proteins in *Lactococcus lactis*. *PLoS ONE* 5: e10317. doi:10.1371/journal.pone.0010317.
104. Lenoir G, Menguy T, Corre F, Montigny C, Pedersen PA, et al. (2002) Overproduction in yeast and rapid and efficient purification of the rabbit SERCA1a Ca(2+)-ATPase. *Biochim Biophys Acta* 1560: 67–83.
105. Griffith DA, Delipala C, Leadsham J, Jarvis SM, Oesterheld D (2003) A novel yeast expression system for the overproduction of quality-controlled membrane proteins. *FEBS Lett* 553: 45–50.
106. Bonander N, Hedfalk K, Larsson C, Mostad P, Chang C, et al. (2005) Design of improved membrane protein production experiments: quantitation of the host response. *Protein Sci* 14: 1729–1740. doi:10.1110/ps.051435705.
107. Abu-Qam M, Eichler J, Sharon N (2008) Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr Opin Struct Biol* 18: 544–550. doi:10.1016/j.sbi.2008.06.010.
108. Goodin MM, Zaitlin D, Naidu RA, Lommel SA (2008) *Nicotiana benthamiana*: its history and future as a model for plant-pathogen interactions. *Mol Plant Microbe Interact* 21: 1015–1026. doi:10.1094/MPMI-21-8-1015.
109. Catty P, Boutigny S, Miras R, Joyard J, Rolland N, et al. (2011) Biochemical Characterization of AtHMA6/PAA1, a Chloroplast Envelope Cu(I)-ATPase. *J Biol Chem* 286: 36188–36197. doi:10.1074/jbc.M111.241034.
110. Liu HP, Takio S, Satoh T, Yamamoto I (1999) Involvement in denitrification of the napKEFDABC genes encoding the periplasmic nitrate reductase system in the denitrifying phototrophic bacterium *Rhodobacter sphaeroides* f. sp. *denitrificans*. *Biosci Biotechnol Biochem* 63: 530–536.
111. Leaver M, Errington J (2005) Roles for MreC and MreD proteins in helical growth of the cylindrical cell wall in *Bacillus subtilis*. *Mol Microbiol* 57: 1196–1209. doi:10.1111/j.1365-2958.2005.04736.x.
112. Schmidt KL, Peterson ND, Kustusch RJ, Wissel MC, Graham B, et al. (2004) A predicted ABC transporter, FtsEX, is needed for cell division in *Escherichia coli*. *J Bacteriol* 186: 785–793.
113. Bouhss A, Crouvoisier M, Blanot D, Mengin-Lecreux D (2004) Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. *J Biol Chem* 279: 29974–29980. doi:10.1074/jbc.M314165200.
114. Raymond M, Svistoonoff S, Loudet O, Nussaume L, Desnos T (2006) Identification of QTL controlling root growth response to phosphate starvation in *Arabidopsis thaliana*. *Plant Cell Environ* 29: 115–125.
115. González E, Solano R, Rubio V, Leyva A, Paz-Ares J (2005) PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 is a plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in *Arabidopsis*. *Plant Cell* 17: 3500–3512. doi:10.1105/tpc.105.036640.
116. Seigneurin-Berny D, Gravot A, Auroy P, Mazard C, Kraut A, et al. (2006) HMA1, a new Cu-ATPase of the chloroplast envelope, is essential for growth

- under adverse light conditions. *J Biol Chem* 281: 2882–2892. doi:10.1074/jbc.M508333200.
117. Haferkamp I, Hackstein JHP, Voncken FGJ, Schmit G, Tjaden J (2002) Functional integration of mitochondrial and hydrogenosomal ADP/ATP carriers in the *Escherichia coli* membrane reveals different biochemical characteristics for plants, mammals and anaerobic chytrids. *Eur J Biochem* 269: 3172–3181.
 118. Verret F, Gravot A, Auroy P, Preveral S, Forestier C, et al. (2005) Heavy metal transport by AtHMA4 involves the N-terminal degenerated metal binding domain and the C-terminal His11 stretch. *FEBS Lett* 579: 1515–1522. doi:10.1016/j.febslet.2005.01.065.
 119. Gravot A, Lieutaud A, Verret F, Auroy P, Vavasseur A, et al. (2004) AtHMA3, a plant PIB-ATPase, functions as a Cd/Pb transporter in yeast. *FEBS Lett* 561: 22–28. doi:10.1016/S0014-5793(04)00072-9.
 120. Misson J, Thibaud M-C, Bechtold N, Raghothama K, Nussaume L (2004) Transcriptional regulation and functional properties of Arabidopsis Pht1;4, a high affinity transporter contributing greatly to phosphate uptake in phosphate deprived plants. *Plant Mol Biol* 55: 727–741. doi:10.1007/s11103-004-1965-5.
 121. Neuhaus HE, Thom E, Möhlmann T, Steup M, Kampfenkel K (1997) Characterization of a novel eukaryotic ATP/ADP translocator located in the plastid envelope of *Arabidopsis thaliana* L. *Plant J* 11: 73–82.
 122. Davies A, Hendrich J, Van Minh AT, Wratten J, Douglas L, et al. (2007) Functional biology of the alpha(2)delta subunits of voltage-gated calcium channels. *Trends Pharmacol Sci* 28: 220–228. doi:10.1016/j.tips.2007.03.005.
 123. Porter RK (2008) Uncoupling protein 1: a short-circuit in the chemiosmotic process. *J Bioenerg Biomembr* 40: 457–461. doi:10.1007/s10863-008-9172-8.
 124. Boise LH, González-García M, Postema CE, Ding L, Lindsten T, et al. (1993) bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell* 74: 597–608.
 125. Borner C (2003) The Bcl-2 protein family: sensors and checkpoints for life-or-death decisions. *Mol Immunol* 39: 615–647.
 126. Feng Y, Broder CC, Kennedy PE, Berger EA (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272: 872–877.
 127. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, et al. (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381: 661–666. doi:10.1038/381661a0.
 128. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, et al. (1996) HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 381: 667–673. doi:10.1038/381667a0.
 129. Riedel C, Dohán O, De la Vieja A, Ginter CS, Carrasco N (2001) Journey of the iodide transporter NIS: from its molecular identification to its clinical role in cancer. *Trends Biochem Sci* 26: 490–496.