Modular Integrated Secretory System Engineering in *Pichia pastoris* To Enhance G-Protein Coupled Receptor Expression

Katrien Claes,†‡∇ Kristof Vandewalle,†‡∇ Bram Laukens,†‡ Toon Laeremans,§∥ Olivier Vosters,† Ingrid Langer,† Marc Parmentier,†§# Jan Steyaert,§∥ and Nico Callewaert*†‡

†Unit for Medical Biotechnology, Medical Biotechnology Center, VIB, Technologiepark 927, 9052 Ghent, Belgium
‡Laboratory for Protein Biochemistry and Biomolecular Engineering, Department of Biochemistry and Microbiology, Ghent University, K.L.-Ledeganckstraat 35, 9000 Ghent, Belgium
§Structural Biology Research Center, VIB, Pleinlaan 2, 1050 Brussels, Belgium
∥Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
⊥Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM), Campus Erasme, 808 Route de Lennik, B-1070 Brussels, Belgium
#Welbio, Université Libre de Bruxelles (U.L.B.), Campus Erasme, 808 Route de Lennik, B-1070 Brussels, Belgium

**ABSTRACT:** Membrane protein research is still hampered by the generally very low levels at which these proteins are naturally expressed, necessitating heterologous expression. Protein degradation, folding problems, and undesired post-translational modifications often occur, together resulting in low expression levels of heterogeneous protein products that are unsuitable for structural studies. We here demonstrate how the integration of multiple engineering modules in *Pichia pastoris* can be used to increase both the quality and the quantity of overexpressed integral membrane proteins, with the human CXCR4 G-protein coupled receptor as an example. The combination of reduced proteolysis, enhanced ER folding capacity, GlycoDelete-based N-Glycan trimming, and nanobody-based fold stabilization improved the expression of this GPCR in *P. pastoris* from a low expression level of a heterogeneously glycosylated, proteolyzed product to substantial quantities (2−3 mg/L shake flask culture) of a nonproteolyzed, homogeneously glycosylated proteoform. We expect that this set of tools will contribute to successful expression of more membrane proteins in a quantity and quality suitable for functional and structural studies.

**KEYWORDS:** *Pichia pastoris*, membrane protein expression, CXCR4, nanobodies, N-glycosylation

Many proteins reside within the plasma membrane and intracellular membranes of the cell, establishing communication with the environment and subcellular compartments. This is achieved through transporters, receptors, enzymes, and anchoring proteins. Because of these essential functions and the accessibility of many of these proteins from the outside of the cell, they are targeted by over 50% of the currently prescribed drugs. Furthermore, for drug discovery research in the membrane protein field, the availability of expression host cells that can produce the drug target at high levels in a homogeneous fashion is of paramount importance. Such cells are used for direct drug screening. Many drugs have severe side effects or have to be administered in high doses because they are not selective, nor potent enough. One way to improve this is to perform structure-based drug design, but there are still much fewer membrane protein 3D structures as compared to soluble proteins. This is in part because the natural abundance of membrane proteins in native tissue is most often low and therefore, recombinant production is needed. Overexpression and purification of membrane proteins is very challenging and it is essential to express the protein in an as natural and homogeneous conformation as possible. Both the quality and quantity of the heterologously expressed membrane protein is crucial. Heterogeneity due to glycosylation, proteolysis, or different conformational states can severely hamper crystal growth. Furthermore, to test many different crystallization conditions, milligram amounts of membrane protein are generally required. While these amounts are slowly dropping due to robotics that can handle nano- to picoliter volumes, many of the more difficult membrane proteins are still way out of reach.

**Special Issue:** Synthetic Biology in Europe

**Received:** January 30, 2016
In this contribution, we used the human G-protein coupled receptor (GPCR) CXCR4 as a case study membrane protein to investigate the effects of several yeast secretory system engineering modules. The receptor is naturally expressed on CD4+ T cells and is, together with CCR5, especially known as one of the primary coreceptors for HIV viral infection. GPCRs in general consist of seven membrane spanning helices, connected by intracellular and extracellular loops. They can bind molecules on the outside of the cell and then change conformation to activate a signaling cascade inside the cell, resulting in a cellular response. GPCRs are involved in many key physiological functions and in multiple diseases, such as neurodegenerative disorders and cancer. They are of great clinical importance, and it is not surprising that GPCRs are still among the most researched proteins in the pharmaceutical industry. A major breakthrough in GPCR structure determination was the discovery that a T4 lysozyme insertion, in the place of the flexible third internal loop of several A-class GPCRs, protected the receptors from proteolytic degradation and provided a larger hydrophilic interface during crystallization. Except for the first characterized GPCR structure of bovine rhodopsin, almost all GPCR crystal structures that have been solved today were obtained using the T4 lysozyme insertion strategy. However, the vast majority of this class of receptors still remains intractable from a heterologous expression point of view. Even for those membrane proteins that express relatively well and could be successfully purified and crystallized, most often extensive protein engineering through, for example, N- and C-terminal truncations and stabilizing mutations was required. In this contribution, we focused on the modular engineering of the expression host rather than of the expressed membrane protein (Figure 1).
We codon-optimized the human CXCR4 receptor gene to facilitate heterologous expression in *Pichia pastoris*, and N-terminally fused it to the *Saccharomyces cerevisiae* alpha mating factor prepro-signal to direct the protein to the ER membrane.

![Quality improvement of the expressed membrane protein](image1)

**Figure 2.** Quality improvement of the expressed membrane protein. By shifting the expression background for the CXCR4 receptor from wild type GS115 (lane 1) to the protease deficient *P. pastoris* strain, Δpep4 (lane 2), a higher amount of the receptor with an intact N-terminus, but still heterogeneous N-glycans, can be observed. To increase the homogeneity of the sample, the receptor was expressed in the GlycoDelete strain (lane 3) and in the GlycoDelete Δpep4 strain (lane 4). This results in membrane proteins that are decorated with a single N-acetyl glucosamine residue (GlcNAc). Expression of the GPCR in the GlycoDelete Δpep4 strain strongly improves the quality of the CXCR4 receptor. The presence of N-glycans on the receptor of the Δpep4 strain (lane 5) and the GlycoDelete Δpep4 strain (lane 6) was confirmed by treating the membrane protein samples with PNGaseF.

![Quantitative improvement of the expressed membrane protein](image2)

**Figure 3.** Quantitative improvement of the expressed membrane protein. (a) Western blot analysis showing the expression level increase of CXCR4, starting from the GlycoDelete Δpep4 strain (lane 1). Upon coexpression of HAC1 (lane 2) and Nanobody 39 (lane 3) an increase in expression levels can be observed, with an even higher increase upon combining both strategies (lane 4). Coexpression of the control nanobody directed against another GPCR did not increase the expression level of the CXCR4 receptor. (b) Expression levels of three independent biological replicates, quantified using the Odyssey system. A total 4- to 5-fold enhancement of membrane protein expression could be obtained.

We codon-optimized the human CXCR4 receptor gene to facilitate heterologous expression in *Pichia pastoris*, and N-terminally fused it to the *Saccharomyces cerevisiae* alpha mating factor prepro-signal to direct the protein to the ER membrane.
We added a C-terminal Rho1D4 tag for easy detection and put the gene under control of the inducible AOX1 promoter. Upon expression of the human CXCR4 receptor in the GS115 background, it was clear that the protein was sensitive to N-terminal degradation. The electrophoretic mobility of the receptor is somewhat faster than the mobility of soluble globular standard proteins of 40 kDa (Figure 2). As overexpressed membrane proteins in yeast likely will be distributed over multiple compartments of the cell (ER, Golgi, plasma membrane, endosomes and vacuole) for prolonged periods of time, we reasoned that exposure to vacuolar proteases would be the main reason for the observed degradation of yeast-expressed membrane proteins. Protease A is an aspartic protease that is produced as azymogen and can autoactivate itself under acidic pH. In turn, it activates a whole array of other vacuolar proteases and is therefore a good target for reducing vacuolar proteolysis. By changing to the SMD1168 host strain (his4 Δpep4), which lacks the vacuolar protease A, higher molecular weight species were produced, but in a much more heterogeneous fashion (Figure 1 and Figure 2, lane 1 vs lane 2), which is indicative of heterogeneous glycosylation. Indeed, many GPCRs, including CXCR4, are N-glycosylated on sites in the N-terminal luminal sequence. This glycosylation site (Asn15 for CXCR4) is expectedly deleted in the GlycoDelete strain, hence the absence of this smear. Indeed, this hypothesis was confirmed upon PNGaseF digestion of the Δpep4 expressed membrane proteins, which converted the smear to a tight band at the approximate expected molecular weight (Figure 2 lane 4 vs lanes 2 and 5). A PNGase F treatment Δpep4 strain also appears to be absent in the GlycoDelete strain (Figure 2 lanes 5 and 6), providing evidence that the GlycoDelete-mediated deglycosylation of the CXCR4 receptor is complete. Together, the Δpep4 and GlycoDelete engineering modules lead to expression of this membrane protein in a very homogeneous, un-degraded quality.

After establishing enhanced protein homogeneity, we further implemented modules to increase the amount of produced membrane protein. First, we coexpressed the spliced form of HAC1 in the GlycoDelete Δpep4 CXCR4 strain to induce the unfolded protein response (UPR) upon methanol induction. This resulted in a modest 59% increase of expressed membrane protein (Figure 3a, lane 1 vs lane 2). In a previous contribution, we also showed that HAC1 overexpression did not dramatically increase expression levels of another GPCR (adenosine A3 receptor), but strongly improved receptor ligand binding activity, most likely through enhanced folding caused by improved chaperone activity in combination with effective degradation of incorrectly folded molecules. We have tried to perform radio-ligand binding assays using the tritiated CXCR4 agonist chemokine SDF1α on crude P. pastoris membrane extracts. However, we encountered an excessively high background signal (aspecific binding of the ligand) in strain extracts not containing the receptor, and therefore were unable to proceed with these experiments. However, the enhanced expression of this receptor in the face of enhanced ER-associated degradation (ERAD) capacity for misfolded proteins is evidence for the natively folded state of the receptor in our strains.

In addition to UPR induction, we wanted to test whether coexpression of CXCR4-specific nanobodies as protein-specific chaperones would further enhance expression levels by stabilizing the receptor. Because nanobodies express at very high levels in Pichia and the bulk of these molecules get secreted, a high concentration should be continuously present in the yeast’s secretory system, available for binding and stabilizing the CXCR4 receptor. To this end, two llamas were immunized with either intact Chinese hamster ovary cells (CHO-K1 cells) overexpressing the CXCR4 receptor, or with membranes derived from these cells. Nanobody-encoding open reading frames were then amplified from the cDNA of the peripheral blood lymphocytes and cloned into phage display libraries. The resulting libraries were enriched for CXCR4-specific nanobodies by several rounds of panning against CXCR4-expressing CHO-K1 cells and against CHO-K1 membrane extracts. Finally, the binding of eight individually purified nanobodies to the extracellular domain of the CXCR4 receptor was analyzed by FACS analysis. To this end, CHO cells with and without CXCR4 were incubated with His6-tagged nanobodies, and the cells were stained with Alexa Fluor 647-conjugated anti-His6 mAb. The fluorescence intensity shift between CXCR4 expressing and nonexpressing CHO cells, as determined via flow cytometry was used as a measure of specific binding of the nanobody to the extracellular side of the receptor (Supporting Information, Table S2).

We tested these eight nanobodies, binding to the extracellular region of CXCR4, and one control nanobody, recognizing another GPCR (ChemR23) for their capacity to enhance CXCR4 expression (Figure S1). From this set of nanobodies, coexpression of NbCA4139 with the receptor consistently showed the highest increase in expression yield: we observe a 2- to 3-fold increase in CXCR4 expression levels (Figure 3a, lane 1 vs lane 3). Coexpression of the control nanobody induced no increase in CXCR4 receptor expression levels (Figure 3a lane 1 vs lane 5).

In two functional assays, Nanobody NbCA4139 inhibited receptor binding of the CXCL12 ligand (Supplementary Figure S2 and Table S2). This nanobody was raised against the human CXCR4 receptor functionally expressed in mammalian cells and inhibits ligand binding in functional assays in mammalian cells. Its coexpression with the receptor in yeast strongly enhances the production level of the receptor. This again provides evidence that the CXCR4 receptor in yeast is natively folded, as no such effects would be expected if the conformational epitope for the nanobody were not present.

Finally, as nanobodies are massively expressed in Pichia pastoris, we wanted to exclude that the observed expression-improvement is due to unfolded protein response secretory system overload. To this end, we performed a quantitative reverse transcription polymerase chain reaction (RT-PCR) experiment on the KAR2 UPR sentinel gene. KAR2 encodes a
molecular chaperone responsible for assistance in protein folding. We tested UPR-induction for the different engineered strains in four independent experiments. From Figure 4 it is clear that upon HAC1 expression (i.e., the driving transcription factor for UPR), the UPR is induced, as expected. However, neither the GlycoDelete module nor expression of the anti-CXCR4 nanobody induced UPR. We therefore conclude that the nanobody-induced enhanced receptor expression level is not caused by an induction of the unfolded protein response, but suggest that it is due to a specific stabilization of the receptor, during or after protein folding at the ER membrane-lumen interface. These data suggested that nanobody coexpression and HAC1-mediated UPR induction could act synergistically on the expression level. Indeed, upon combining the HAC1 and receptor-specific nanobody coexpression modules, a 4- to 5-fold increase in CXCR4 expression could be obtained as compared to the GlycoDelete Δpep4 CXCR4 strain (Figure 3a: lane 1 vs lane 4, and Figure 3b). By using a Rho1D4-tagged IL-22 internal standard on Western blot, we estimated the total amount of CXCR4 in this final strain (Δpep4 CXCR4 Hac1 NbCA4139) at 2−3 mg per 125 mg (±2%) of total membrane protein per liter of Pichia shake culture (Figure S3). This is a level that is a feasible starting point for extraction and purification purposes. For comparison, the most easily expressed GPCR, bovine rhodopsin, which was also the first one to be crystallized, can be produced at ±2−5 mg per liter of highly optimized HEK293 suspension cell culture.21

With this synthetic biology study we showcase the development and integration of a panel of engineering modules for Pichia pastoris that increase the chances for difficult-to-express membrane proteins, using the model CXCR4 G-protein coupled receptor, to reach the quantity and quality needed to allow for use of the cells in drug screening projects and/or for starting up structure determination projects. The application of this set of engineering modules may provide powerful tools to explore the expression in sufficient quantity and quality of eukaryotic membrane proteins in Pichia pastoris, an expression system that in and by itself is already one of the more successful ones for the expression of membrane proteins.22 As GPCR stabilization can also be achieved by T4 lysozyme insertion in the third intracellular loop, it would be useful in future work to explore whether this could replace the nanobody coexpression in our approach. However, for other categories of membrane proteins, nanobody coexpression provides for a potentially more broadly applicable stabilization strategy. Contrary to other methods for enhancing membrane protein expression, it is of note that for our methodology, nothing needs to be changed in the protein’s native sequence, which should be useful to pharmacologists.

■ MATERIALS AND METHODS

Full details of all materials and methods are provided in online Supporting Information.

■ ASSOCIATED CONTENT

* Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.6b00032.

Figure S1, effect of CXCR4 expression level upon specific nanobody coexpression; Figure S2, functional characterization of the CXCR4-specific nanobody and the control nanobody; Figure S3, quantification of the CXCR4 expression level; Table S1, coding sequences of the CXCR4-specific and control nanobody; Table S2, summary of the nanobody–CXCR4 binding data; Table S3, primers used for UPR-detection in quantitative PCR experiment. Material and methods. Supplementary references (PDF)

**Figure 4.** Quantitative PCR analysis of the unfolded protein response. Four biological replicates in the early exponential phase were analyzed for induction of the unfolded protein response. The strains expressing HAC1 show an increase in mRNA levels of HAC1 and KAR2, the signature gene of the unfolded protein response. Because KAR2 transcription is not increased in the nanobody coexpressing strain, the enhanced CXCR4 (CR) expression levels seen when coexpressing the nanobody are not due to UPR induction. Statistical analysis is performed using the Prism6 software.
Author Information

Corresponding Author
*E-mail: Nico.Callewaert@vib-ugent.be.

Author Contributions
V.K.C. and K.V. contributed equally to this work, designed and performed experiments, analyzed and interpreted data, and drafted the manuscript. B.L. developed the GlycoDelete system in Pichia pastoris. T.L. generated the CXCR4 nanobodies under scientific supervision of J.S. O.V. and I.L. characterized the nanobodies under supervision of M.P. N.C. initiated and designed the study, coordinated the project, interpreted data, and cowrote the manuscript. All authors read and approved the final version of the manuscript.

Notes
The authors declare no competing financial interest.

Acknowledgments
K.C. and B.L. are supported by a Ph.D. grant from the Institute for the Advancement of Scientific and Technological Research in Industry (IWV—Vlaanderen). K.V. is supported by the Fund for Scientific Research Flanders (FWO). J.S. and N.C. thank INSTRUCT, part of the European Strategy Forum on Research Infrastructures (ESFRI) and the Hercules Foundation Flanders for their support to nanobody discovery. Research in the lab of J.S. is also supported by G049512N of the Fund for Scientific Research Flanders (FWO). Research in the lab of M.P. is supported by the Brussels Region (3D4Health program), the Fonds de la Recherche Scientifique de Belgique, and the Interuniversity Attraction Poles Programme (P7–14), Belgian State, Belgian Science Policy. Research in the lab of N.C. is supported by the Marie Curie Excellence Grant MEXT-14292 under EU Framework Program 6 and Research Grant G.0.541.08.N.10 of the Fund for Scientific Research Flanders (FWO).

References