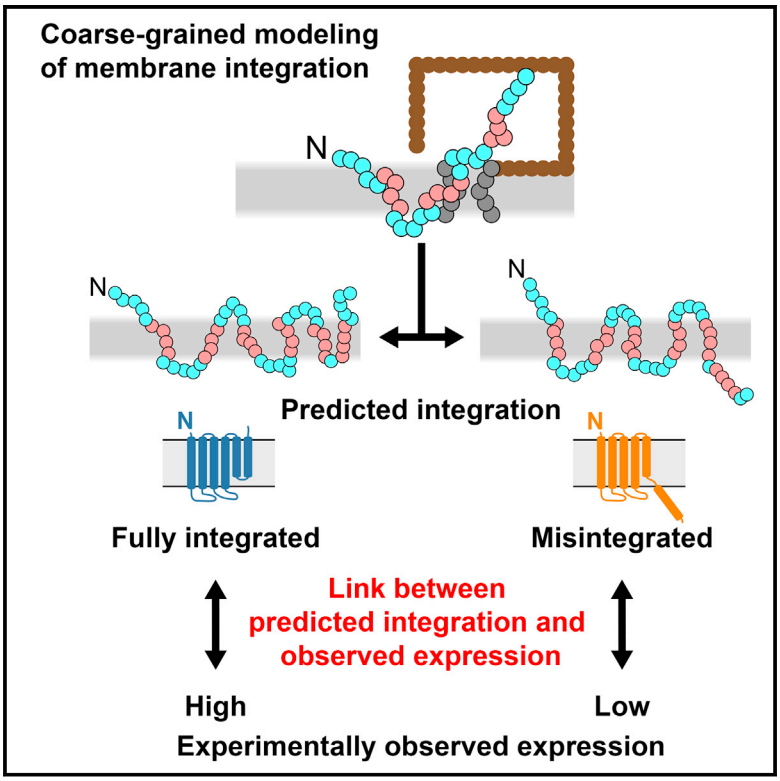
Article

Cell Reports

A Link between Integral Membrane Protein Expression and Simulated Integration Efficiency

Graphical Abstract



Highlights

- Closely related IMP sequences exhibit substantial differences in expression levels
- IMP expression levels are dependent upon the efficiency of membrane integration
- In distinct systems, mutations that improve IMP integration also improve expression
- Simulated IMP integration is used to design IMPs with enhanced expression

Authors

Stephen S. Marshall, Michiel J.M. Niesen, Axel Müller, ..., Bin Zhang, William M. Clemons, Jr., Thomas F. Miller III

Correspondence

clemons@caltech.edu (W.M.C.), tfm@caltech.edu (T.F.M.)

In Brief

Marshall et al. demonstrate that an important bottleneck for integral membrane protein (IMP) expression is integration into the membrane. A recently developed computational model for predicting IMP integration efficiency is used to understand, predict, and enhance experimentally observed IMP expression levels.

Marshall et al., 2016, Cell Reports *16*, 2169–2177
August 23, 2016
http://dx.doi.org/10.1016/j.celrep.2016.07.042

CellPress

Cell Reports
Article

A Link between Integral Membrane Protein Expression and Simulated Integration Efficiency

Stephen S. Marshall,[1,2] Michiel J.M. Niesen,[1,2] Axel Müller,[1] Katrin Tiemann,[1] Shyam M. Saladi,[1] Rachel P. Galimidi,[1] Bin Zhang,[1] William M. Clemons, Jr.,[1,3,*] and Thomas F. Miller III[1,*]
[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[2]Co-first author
[3]Lead Contact
*Correspondence: clemons@caltech.edu (W.M.C.), tfm@caltech.edu (T.F.M.)
http://dx.doi.org/10.1016/j.celrep.2016.07.042

SUMMARY

Integral membrane proteins (IMPs) control the flow of information and nutrients across cell membranes, yet IMP mechanistic studies are hindered by difficulties in expression. We investigate this issue by addressing the connection between IMP sequence and observed expression levels. For homologs of the IMP TatC, observed expression levels vary widely and are affected by small changes in protein sequence. The effect of sequence changes on experimentally observed expression levels strongly correlates with the simulated integration efficiency obtained from coarse-grained modeling, which is directly confirmed using an in vivo assay. Furthermore, mutations that improve the simulated integration efficiency likewise increase the experimentally observed expression levels. Demonstration of these trends in both *Escherichia coli* and *Mycobacterium smegmatis* suggests that the results are general to other expression systems. This work suggests that IMP integration is a determinant for successful expression, raising the possibility of controlling IMP expression via rational design.

INTRODUCTION

The central role of integral membrane proteins (IMPs) in many biological functions motivates structural and biophysical studies that require large amounts of purified protein, often at considerable costs in terms of both materials and labor. A key obstacle is that only a small percentage of IMPs can be overexpressed (i.e., heterologously produced at levels conducive to further study) (Lewinson et al., 2008). While extensive efforts have shown promising results for individual IMPs, including those focusing on expression conditions, host modification, and directed evolution (reviewed in Schlegel et al., 2010; Wagner et al., 2006; and Scott et al., 2013), none of these has proven broadly applicable, even among homologs of a given IMP. In general, the determinants for IMP expression are poorly understood, leading to the prevailing opinion that problems in membrane protein expression must be addressed on a case-by-case basis.

Closely related IMP homologs can vary dramatically in the amount of protein available after expression (Lewinson et al., 2008), which raises a fundamental question: what differentiates the expression of IMP homologs? The hypothesis raised here is that the efficiency with which an IMP is integrated into the membrane is a key determinant in the degree of observed IMP expression.

A fundamental step in the biosynthesis of most IMPs involves their targeting to and integration into the membrane via the Sec protein translocation channel (Rapoport, 2007). Integration of IMP transmembrane domains (TMDs) into the membrane is facilitated primarily through interaction between the nascent chain and SecY, which forms the core of the protein translocation complex, or translocon. Following the co-translational or post-translational insertion of nascent protein sequences into the translocon channel, hydrophobic segments pass through the lateral gate of SecY into the membrane to form TMDs. Factors such as TMD hydrophobicity (Harley et al., 1998; Hessa et al., 2005) and loop charge (Heijne, 1986; Goder and Spiess, 2003) have been shown to affect the efficiency of TMD integration and topogenesis. For example, TMD hydrophobicity is directly related to the probability with which TMDs partition into the lipid bilayer, while positively charged residues in the loop alter TMD orientation by preferentially occupying the cytosol (Goder and Spiess, 2003; Hessa et al., 2005; Heijne, 1986).

In this study, we investigated the connection between observed IMP expression levels and Sec-facilitated IMP integration efficiency (i.e., the probability of membrane integration with the correct multi-spanning topology). Systematic investigation of chimeras within an IMP family led to the identification of sequence elements that modulate expression levels. In silico modeling of IMP integration at the Sec translocation channel found that the sequence modifications that increase the calculated IMP integration efficiency correlate with in vivo overexpression improvements, suggesting that IMP integration efficiency is a determinant for successful expression. The result was found to be general across distinct expression systems (*E. coli* and *M. smegmatis*). Furthermore, an in vivo assay based on antibiotic resistance in *E. coli* experimentally confirmed the model that the integration efficiency of an individual TMD correlates with the observed IMP expression levels. The strong link between