Current Protocols in Protein Science

December 2, 2013

Dr. Howard S. Young
University of Alberta
Edmonton, Alberta, Canada

Tel: 780-492-3931
Email: hyoung@ualberta.ca

RE: CPPS Unit 29.8 Page Proofs

1. Your manuscript has been edited to Current Protocols style and structure. Please **carefully read** your page proofs and mark any changes right in the text.

2. You may submit your corrections via email, annotated or scanned pdf in email (**preferred**), fax, or overnight express service (not regular postal mail).

3. If you have large inserts to add, please supply us with an electronic file and hard copy of the insert(s) and indicate where they should go.

4. Read and answer all the queries on the Author/Editor Query Sheet and return this sheet with your proofs.

5. If you have not returned your signed contract yet, you must return it at this time.

6. Please forward all materials by **December 10th** to:

Allen Ranz
Current Protocols
John Wiley & Sons, Inc.
111 River Street, MSC 8-02
Hoboken, NJ 07030-5774

Tel: (201) 748-6278
Fax: (201) 748-6207
aranz@wiley.com

cc: Allen Ranz

Author/Editor Queries

1. Unit Introduction, 1st paragraph, 5th sentence: Is there a more recent literature reference for Membrane Protein Data Bank than Raman et al. (2006)?

2. Basic Protocol 1, (a) step 8: Please specify the concentrations of the solutions whose volumes are itemized here.

   (b) step 13: Do you really mean "boil" the resuspended colonies (i.e., such that ebullition is observed) or do you really mean place it in a 100°C heat block for 5 min (whereby the temperature in the tube itself will remain slightly below the boiling point)?

3. Basic Protocol 2 materials list: Correct that for LB+amp medium you add 1 ml of 100 mg/ml ampicillin stock per liter LB medium (100 μg/ml final; equivalent to the LB+amp plates minus the agar)?

4. Basic Protocol 2, step 2: Correct that a 50-ml Erlenmeyer flask is used?

5. Basic Protocol 3 title: Please expand on this title to be more topical and descriptive.

6. Basic Protocol 3, (a) steps 10 and 12: What size separatory funnel is used in each of these steps?

   (b) step 11: At what temperature are the layers left to separate?

7. Alternate Protocol 1 title: Please expand on this title to be more topical and descriptive.

8. Alternate Protocol 1 materials list: Is the protein pellet obtained from one of the preceding protocols? Which protocol and which step?

9. Reagents and Solutions, (a) We interpreted your original note at the top of Reagents and Solutions to mean that the storage temperature is room temperature unless otherwise stated and that the maximum length of storage is 1 year unless otherwise stated, and specified storage conditions accordingly. Please check the storage conditions and forgive (and of course correct) any errors.

   (b) LB+amp plates: Correct that you add 1 ml of 100 mg/ml ampicillin stock per liter (100 μg/ml final)?

10. Critical Parameters and Troubleshooting, "Expression and purification of fusion protein": Note that the LB medium that you describe in Reagents and Solutions is

defined throughout Current Protocols as Luria-Bertani medium (not “Lysogeny Broth”), so we have changed it to the former designation.

11. Literature Cited: Bertani (2004) and Marblestone et al. (2006) are not cited in the text; please find appropriate locations at which to cite.

Membrane Transport Piece by Piece: Production of Transmembrane Peptides for Structural and Functional Studies

Grant Kemp,[1] Larry Fliegel,[1] and Howard S. Young[1,2]

[1]Department of Biochemistry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada
[2]National Institute for Nanotechnology, University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Membrane proteins are involved in all cellular processes from signaling cascades to nutrient uptake and waste disposal. Because of these essential functions, many membrane proteins are recognized as important, yet elusive, clinical targets. Recent advances in structural biology have answered many questions about how membrane proteins function, yet one of the major bottlenecks remains the ability to obtain sufficient quantities of pure and homogeneous protein. This is particularly true for human membrane proteins, where novel expression strategies and structural techniques are needed to better characterize their function and therapeutic potential. One way to approach this challenge is to determine the structure of smaller pieces of membrane proteins that can be assembled into models of the complete protein. This unit describes the rationale for working with single or multiple transmembrane segments and provides a description of strategies and methods to express and purify them for structural and functional studies using a maltose binding protein (MBP) fusion. The bulk of the unit outlines a detailed methodology and justification for producing these peptides under native-like conditions. *Curr. Protoc. Protein Sci*. 75:29.8.1-29.8.28 © 2013 by John Wiley & Sons, Inc.

Keywords: membrane proteins • hydrophobic peptides • bacterial expression • maltose binding protein • organic extraction • molecular structure

INTRODUCTION

Structural biology is an important tool in the progression of modern medicine. From understanding the chemistry of relatively simple molecules, such as drugs and antibiotics, to being able to visualize large proteins and protein complexes, structural biology has changed the way we understand the human body and its environment. However, our understanding of membrane proteins lags behind that of soluble proteins, and determining the structure of a membrane protein remains a challenging endeavor. Approximately one-third of the human genome encodes membrane proteins (Wallin and Von Heijne, 1998), and many of these are potential drug targets. However, membrane proteins only account for ~1.6% of structures deposited in the Protein Data Bank (calculated using the Membrane Protein Data Bank; *http://www.mpdb.tcd.ie*; Raman et al., 2006). One reason for this lies in the relative difficulty in expressing, purifying, and studying membrane proteins in vitro. Not only must all the considerations for working with soluble proteins be optimized, but there are unique challenges associated with maintaining the appropriate membrane or membrane-like environment. Before this optimization can begin, the first hurdle is the fact that most membrane proteins are not naturally abundant, necessitating their heterologous expression. It is also typically more difficult to express and purify membrane proteins than soluble proteins, especially human or mammalian variants (Junge et al., 2008). This is compounded when producing large quantities of polytopic membrane proteins for structural studies.

Current Protocols in Protein Science 29.8.1-29.8.28, February 2014
Published online February 2014 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/0471140864.ps2908s75

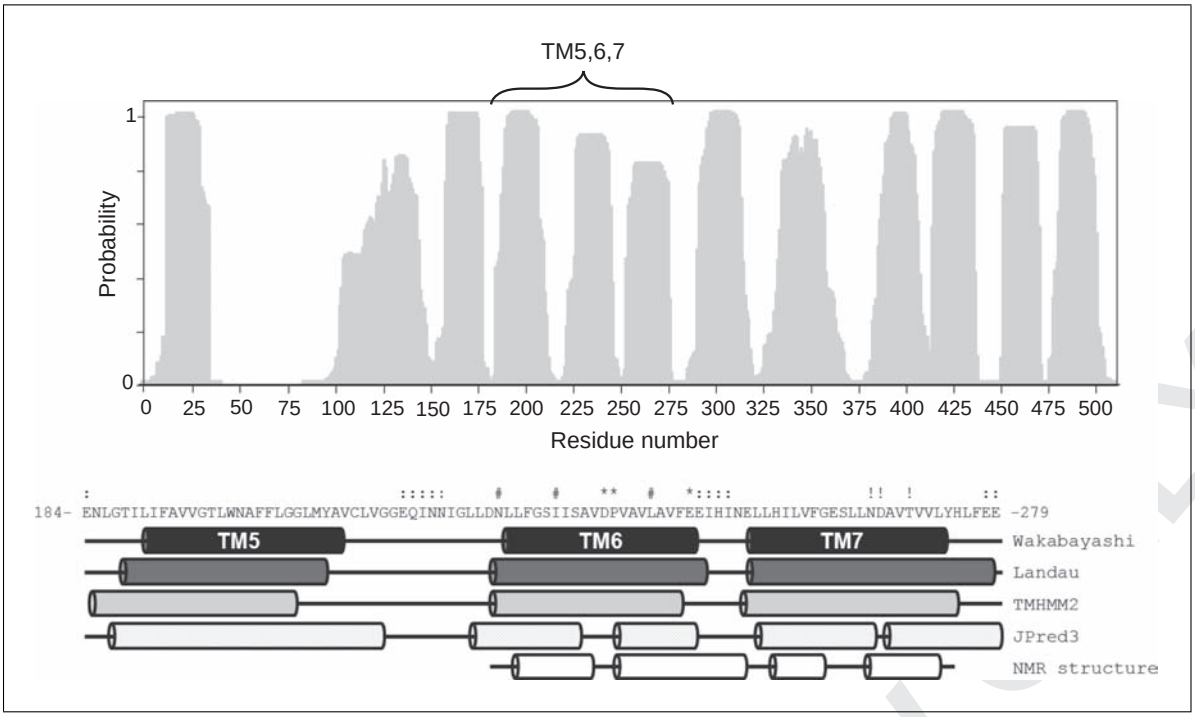

Figure 29.8.1 TM peptide design. The upper panel depicts the output of TMHMM2 prediction of membrane topology (Krogh et al., 2001) for the human sodium proton exchanger isoform 1 (NHE1) residues 1 to 510. The probability of being located in the membrane is plotted as a function of residue number. The lower panel displays the sequence of NHE1 between residues 184 and 279 with various TM segment predictions, the regions of helical content determined by NMR, and known functional mutations. The model of Wakabayashi et al. (2000) is shown as black cylinders; Landau et al. (2007) is in dark gray; TMHMM2 prediction is in light gray; helices predicted by the JPred3 algorithm (Cuff and Barton, 1999) are patterned; and helices indicated by the NMR structures [Tzeng et al. (2010) and Ding et al. (2006)] are in white. Above the sequence functionally important residues are labeled: (*) represents Cys mutants that inactivate NHE1 (Ding et al., 2006); (#) represents Cys mutants that are affected by cysteine modification reagents (Ding et al., 2006); (!) represents Ala mutants that greatly reduce NHE1 activity (Tzeng et al., 2010); and (:) represents Cys mutants that are accessible from the extra- or intracellular space (Wakabayashi et al., 2000).

This unit discusses one strategy in the 'divide-and-conquer' approach (Yeagle et al., 1995a; Lee et al., 2011) to studying membrane proteins. It includes information on the design, heterologous expression, and purification to homogeneity of transmembrane (TM) peptides that are representative fragments of larger membrane proteins, and which can be used for structural biology.

The method uses a maltose binding protein (MBP) fusion that allows otherwise insoluble or toxic hydrophobic peptides to be expressed and purified as a soluble protein construct in *E. coli* (Kapust and Waugh, 1999). To illustrate, the human Na^+/H^+ exchanger isoform 1 (NHE1) is presented as an example (Fig. 29.8.1). The biological significance and the role of this protein in disease have been well established, yet determining the structure of the full-length protein has remained a challenge. There have been many studies of this class of transport proteins, revealing aspects of function, sites critical for ion binding and transport, and mechanisms of regulation (Slepkov and Fliegel, 2002). To bring the understanding of NHE1 to the molecular level, our group and others have carried out structural studies. These include a low-resolution molecular envelope by single-particle electron microscopy (Moncoq et al., 2008), the structure of five individual transmembrane segments (Slepkov et al., 2005; Ding et al., 2006; Reddy et al., 2008; Lee et al., 2009b; Tzeng et al., 2010), an extracellular loop (Lee et al., 2009a), and the region corresponding to a regulatory protein binding site (Mishima et al., 2007). Additionally, two three-dimensional structural homology models based on the crystal structure of the related

bacterial transporter NhaA have been published using evolutionary conservation and fold alignment (Landau et al., 2007), and electron paramagnetic resonance spectroscopy (Nygaard et al., 2011). However, further structural work on either the full-length protein or a truncated version containing only the transmembrane domain has not been successful. Herein we describe an approach that provides structural knowledge on fragments of a membrane transport protein such as human NHE1. One or more transmembrane segments are fused to maltose binding protein, which facilitates high-level expression as well as ease of purification and structure determination. We use a region of NHE1 that includes transmembrane segments 5 to 7 as an example of the successful application of the technique. The basic protocols offer guidance to researchers interested in the structure of other relevant membrane protein targets. The challenges that may be encountered, alternative approaches, and the physiological relevance are discussed below.

The Strategic Planning section provides guidance on the design of a membrane protein fragment that contains one or more TM segments of a polytopic membrane transport protein. In Basic Protocol 1, the DNA sequence that encodes the peptide fragment is cloned into an expression vector as a MBP fusion protein. As an example, the design of a vector for the expression of human Na^+/H^+ exchanger TM segments is described, including the design of a protease cleavage site and restriction sites for ease of cloning. Basic Protocol 2 describes the expression and purification of a fusion protein consisting of MBP and a peptide fragment containing one or more TM segments. This is a commonly used recombinant approach for generating a target peptide as an in-frame fusion with a protein that assists in high-level expression and proper folding. Conditions are described for the overexpression of the fusion protein as a soluble protein and one-step purification by affinity chromatography, followed by protease cleavage to liberate the target TM peptide. Next, the difficult task of TM peptide purification is described in Basic Protocol 3. This protocol and Alternate Protocol 1 outline several methods for the purification of hydrophobic TM peptides using selective solubilization with denaturants, organic extraction, and reversed-phase high-performance liquid chromatography (HPLC). Since the purification of hydrophobic peptides is largely empirical, a workflow is described for the analytical-scale evaluation of different purification strategies, as well as the preparative-scale generation of large amounts of pure peptide. Finally, the Commentary provides insights into the major challenges and steps that can be taken for the successful design, expression, and purification of a piece of a membrane transport protein.

STRATEGIC PLANNING: DESIGNING A PEPTIDE FOR EXPRESSION

Several issues need to be considered when designing a transmembrane peptide for structural or functional studies. Since this protocol is designed for expression of pieces of a polytopic membrane protein, there is no guarantee that the peptide of interest will be properly folded in isolation. This is especially true for multiple TM segments. Below are some strategic points, with relevant references for NHE1, that will help maximize the likelihood of the chosen peptide folding successfully. Below this detailed description, the key points are listed stepwise. Using all available experimental data, identify transmembrane regions with critical amino acids—e.g., those involved in substrate affinity, ion transport or drug binding (Slepkov and Fliegel, 2002). Once a region of interest has been identified, gather as much information as possible about its potential topological orientation. Although the final focus will be on a shorter transmembrane peptide, using the full-length protein sequence (or at least the full transmembrane domain) in the following steps will provide the most accurate results. First, collect empirical biochemical data such as that afforded by the Substituted Cysteine Accessibility Method, which gives topological information (Wakabayashi et al., 2000), and biophysical data indicating secondary structure, such as circular dichroism (Moncoq et al., 2008) or Fourier transform infrared spectroscopy, and limited tertiary structure, such as electron

paramagnetic resonance (Nygaard et al., 2011). This will be useful in determining the accuracy of the subsequent hypothetical topology or secondary structure predictions. Published data collected purely in silico such as topology based on hydrophobicity and secondary-structure predictions may be included, as well as nonempirical information such as evolutionary conservation and fold alignment, if available (Landau et al., 2007). Once all the published data have been collected, predictive algorithms are used to knit together the likely topology and secondary structure for the target protein. Several good algorithms exist for predicting membrane topology. Two of the most accurate algorithms, TMHMM2 and HMMTop2, use a hidden Markov model to determine the probability of residues being in a transmembrane helix. TMHMM2 calculates the probability based on charge, polarity, and hydrophobicity of a residue being in either a loop, helix cap, or helix center (Krogh et al., 2001). HMMTop2 calculates probability based on a set of transmembrane proteins of known structure (Tusnády and Simon, 2001). Both algorithms consider the relative probabilities of contiguous series of residues to determine the most likely placement of a transmembrane helix. This is distinctly advantageous over traditional methods based solely on hydropathy scales for predicting transmembrane helices of transport proteins that are likely to have charged and polar residues in the center of the membrane. Another useful algorithm is JPred3 (Cole et al., 2008). This algorithm uses a BLAST search to find proteins that have a similar sequence and compares predicted secondary structure and known structures to the target protein. This data is used to generate a consensus secondary structure prediction. Once all the experimental and predictive data has been collected, a simple alignment can be used to identify appropriate peptide sequences. An appropriate sequence should contain the transmembrane region of interest and at least three extra-membrane residues on either side. Further capping of the sequence with added lysine residues to increase solubility may also be considered. Agreement of the experimental data, topology prediction, and secondary structure prediction is taken to indicate a transmembrane helix. Figure 29.8.1 demonstrates how the data for a region of NHE1 (TM 5 to 7) was compared and a peptide was chosen.

Design of the peptide (Fig. 29.8.1)

(1) Locate the protein sequence of interest.
   (a) UniProt (*http://www.uniprot.org*) and NCBI (*http://www.ncbi.nlm.nih.gov*) are good online sources.
(2) Survey the literature for relevant biochemical and biophysical data (see above).
(3) Perform topology and secondary structure predictions.
   (a) For TM prediction we recommend TMHMM2 (*http://www.cbs.dtu.dk/services/TMHMM/*) and HMMTop2 (*http://www.enzim.hu/hmmtop/*).
   (b) For secondary structure prediction we recommend Jpred3 (*http://www.compbio.dundee.ac.uk/www-jpred/*), and many others can be found on ExPASy (*http://www.expasy.org/tools/#secondary*).
(4) Use the data collected above to choose a peptide sequence.
   (a) Include ≥ 3 extra-membrane residues on either side of a predicted transmembrane segment.

If solubility problems are encountered in the initial stages of peptide purification (see below), it may be desirable to include either additional residues from extra-membranous loops or capping lysine residues that flank the TM segment.

BASIC PROTOCOL 1

VECTOR CONSTRUCTION AND CLONING A TRANSMEMBRANE PEPTIDE FOR EXPRESSION

Once a transmembrane construct has been designed, the TM peptide is cloned into an MBP expression vector. Expression vectors for MBP fusion proteins are commercially available (New England Biolabs). We have previously described how pMal-c2x was

adapted to allow more efficient screening of multiple constructs of phospholamban and sarcolipin (Douglas et al., 2005). Below we adapt this procedure for the currently available pMal-c5X. Briefly, a polymerase chain reaction (PCR) product containing a tobacco etch virus (TEV) protease site and the restriction sites *Bam*HI and *Eco*RI is created, allowing a single protocol for the cloning, expression, purification, and characterization of various peptide constructs. Selection of the TEV protease was based on its specificity and efficiency in our preparation conditions, and because recombinant protein can be readily obtained. Factor Xa, included in the pMal-c5X vector, is a popular alternative (Nagai et al., 1985).

The pMal-c5X+TEV vector is only created for the first construct. Subsequent cloning uses the *Bam*HI and *Eco*RI sites incorporated during this protocol. This protocol will result in a pMal-c5X vector containing TEV–*Bam*HI–*gene of interest*–STOP–*Eco*RI–*Sal*I- following MBP. Remember that an initiating methionine is not included after the *Bam*HI, so a C-terminal fusion protein is created. If deviating from the protocol below, ensure that the TEV and the 5′ restriction site are in frame with MBP to create the correct fusion protein.

Primer design for vector construction

PCR primers should contain 18 to 30 annealing base pairs (bp) with a melting temperature (T_m) of 50° to 60°C. Longer primers may have overly high T_m or lead to unfavorable DNA secondary structure (i.e., hairpins). These guidelines do not apply to the 5′ non-annealing region of the primer (e.g., tags or restriction sites), so the final primer may be significantly longer than 30 bp. If restriction sites are included at the end of the primer, additional bases should be added (two or more) to ensure efficient restriction digestion of the PCR product. Enzyme-specific information about "Cleavage Close to the End of DNA Fragments" should be available online from the manufacturer. Many online tools exist to determine the information listed above (e.g., *http://www.idtdna.com/scitools/*). Remember that T_m and annealing temperature (T_a) should be calculated using only the annealing sequence, but DNA secondary structure (e.g., hairpins) analysis should be done using the entire primer sequence. If following the protocol below directly, the primers should be as follows: forward, 5′-GAGAACCTGTACTTCCAG*GGATCC***(**$(NNN)_{6\text{-}10}$-3′—TEV site (protein sequence, ENLYFQ) is underlined, *Bam*HI site is italicized, 18 to 30 annealing base pairs from the gene of interest (without an initiating methionine) are shown as $(NNN)_{6\text{-}10}$; reverse: 5′-ACTG*GAATTC***TCA(**$(NNN)_{6\text{-}10}$-3′, STOP codon is bold, *EcoRI* site is italicized, 18 to 30 annealing base pairs from the gene of interest are shown as $(NNN)_{6\text{-}10}$. ACTG at the 5′ end of the reverse primer allows efficient restriction digestion of the PCR product (primers based on Douglas et al., 2005). These primers are ordered from an oligonucleotide synthesis company (e.g., Integrated DNA Technologies, *http://www.idtdna.com*).

Materials

10 μM specific PCR forward primer (Integrated DNA Technologies (*http://www.idtdna.com*; see note below for suggestions on primer design)
T4 kinase kit (New England Biolabs, cat. no. M0201S)
10 μM specific PCR reverse primer (Integrated DNA Technologies, *http://www.idtdna.com*)
cDNA template containing the gene of interest
DNA polymerase (high fidelity, such as Pfx, preferred; Invitrogen, cat. no. 11708-013) and its specific buffers (also see *APPENDIX 4J*)
10 mM dNTP mix (Fermentas, cat. no. R0181, or see recipe in *APPENDIX 4J*)
QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
QIAquick Gel Extraction Kit (Qiagen, cat. no. 28704)

Restriction enzymes: *Pdm*I (Thermo Scientific, cat. no. FD1534), *Eco*RI (Thermo Scientific, cat. no. FD0274), *Bam*HI (Thermo Scientific, cat. no. cat. no. FD0054), and respective buffers
pMal-c5X plasmid (New England Biolabs, cat. no. N8108)
T4 DNA ligation kit (New England Biolabs, cat. no. M0202S)
Competent DH5α *E. coli* (Invitrogen, cat. no. 18258-012; stored at −80°C)
LB+amp plates (see recipe)
80% (v/v) glycerol (autoclaved)
Qiagen Plasmid MidiPrep Kit (Qiagen, cat. no. 12143)
pMal sequencing primers (New England Biolabs)

200-μl PCR tubes
Thermal cycler
Sterile loop
Heat block for boiling resuspended colonies
37°C shaking incubator
1.5-ml cryotubes (Nunc, cat. no. 114858) for glycerol stocks

Additional reagents and equipment for PCR (*APPENDIX 4J*), agarose gel electrophoresis (*APPENDIX 4F*), and DNA sequencing (Shendure et al., 2011)

NOTE: At any stage, purified DNA (without added enzymes) can be stored at 4°C for a few days or at −20°C for several months creating convenient stopping points.

Prepare vector and perform PCR (Day 1)

1. Phosphorylate the forward primer using a T4 kinase kit.

   Blunt-end cloning requires a phosphorylated 5′ end for ligation.

   Performing this step on the primer rather than the PCR product reduces PCR product loss.

2. In a 200-μl PCR tube, prepare the PCR reaction on ice using the phosphorylated forward primer and the cDNA of interest as template, with a high-fidelity DNA polymerase, reverse primer, and 10 mM dNTP mix (see *APPENDIX 4J*).

   Specific PCR reaction conditions are variable and should be obtained from the polymerase manufacturer. See APPENDIX 4J.

3. Run the PCR reaction in a thermal cycler (see *APPENDIX 4J*).

   Typical cycling conditions should be obtained from the polymerase manufacturer.

   Annealing temperature (T_a) should be the melting temperature of the primers minus 5°C (see note on primer design above).

   At this stage, the PCR products can be stored at 4°C for at least 3 days or frozen at −20°C for several months.

4. Confirm the success of the PCR reaction by analyzing 2 to 5 μl of the PCR reaction by agarose gel electrophoresis (*APPENDIX 4F*).

5. If the correct-sized product is observed, purify the PCR product from the reaction (e.g., using a QIAquick Gel Extraction Kit).

6. Digest the purified PCR product with *Eco*RI, following the manufacturer's instructions, and re-purify as in step 5.

7. Simultaneously digest 4 μg of pMal-c5X vector with *Pdm*I and *Eco*RI and purify (e.g., using a QIAquick method).

Not all restriction enzymes are functional under the same conditions. Consult the double-digest enzyme compatibility provided by the manufacturer. This is done by comparing each enzyme's compatibility in each restriction buffer and choosing the one that provides the highest activity for both enzymes while avoiding 'star activity' (nonspecific digestion). If one buffer cannot be found, then digest with one enzyme, purify the plasmid (e.g., using a QIAquick method), and then digest with the second enzyme.

If restriction enzymes are being purchased for this protocol, the authors recommend FastDigest enzymes from Thermo Scientific. All FastDigest enzymes are compatible in one universal buffer at the same temperature, and digestions are complete within 15 min (although longer digestions may be required for large amounts of DNA). This saves time required for multiple digestions and subsequent purifications.

Ligate the PCR product into pMal-c5X

8. Ligate the cut PCR product and cut vector using a DNA ligation kit; typical conditions are: 13 μl μg/μl cut PCR product, 2 μl ?? μg/μl cut vector, 4 μl 5× ligation buffer, and 1 μl ?? U/μl T4 ligase. Incubate at room temperature for 2 hr or overnight at 4°C.

Transform the pMal-c5X-peptide (ligation reaction) into E. coli (Day 2)

9. Defrost a 100-μl aliquot of chemically competent DH5α cells on ice. Transform them by the sterile addition of 5 or 10 μl of the ligation reaction followed by incubation on ice for 40 to 60 min. Spread the entire aliquot of cells onto an LB+amp plate and place at 37°C overnight.

   Another cloning strain of E. coli or a different transformation protocol is also suitable (see Commentary).

   Chemically competent DH5α cells can be purchased or prepared in house (Hanahan, 1983).

   A heat-shock step (60°C for 30 sec) can be performed before spreading the cells onto an LB+amp plate. This step seals the bacterial membrane, trapping the plasmid inside. Due to high transformation efficiency, this is typically not required for DH5a and pMal-c5X.

Confirm PCR product ligation into pMal-c5X (Day 2 or 3)

10. With a sterile loop or toothpick, select an individual transformant and suspend it in 100 μl of sterile water in a sterile microcentrifuge tube by vigorously stirring the water with the loop.

11. Immediately streak this loop onto a fresh LB+amp plate and place it at 37°C overnight.

    This maintains the colony for later glycerol stocks and plasmid purification.

12. Repeat steps 10 and 11 for at least three more colonies (up to 10 is recommended)

13. ?Boil? the resuspended colonies for 5 min.

    This is the PCR template.

14. Repeat PCR reaction (steps 2 to 3) using 2 μl of the ?boiled? colony resuspension as template.

    Use the same primers generated in step 1. The forward primer does not need to be phosphorylated.

15. Confirm successful ligation reactions by analyzing the PCR reaction by agarose gel electrophoresis (*APPENDIX 4F*).

    A successful ligation will yield the same PCR product that was visualized in step 4. Ligation products that are absent or are of incorrect size indicate that the PCR product (step 3) was not ligated properly into pMal-c5X.

Prepare glycerol stock and purify plasmid DNA (Day 3 and 4)

16. The next day, use a sterile loop or toothpick to select a single colony from the plate streaked in step 11, and inoculate it in 2 ml of LB+amp; incubate for 6 to 8 hr in a 37°C shaking incubator.

    Note that the single colony corresponds to the positive ligation as observed in step 15.

17. Transfer this culture to 50 ml of LB+amp and grow overnight in a 37°C shaking incubator.

18. The next day, make a glycerol stock by taking 800 μl of the overnight culture and mixing it with 200 μl of sterile 80% glycerol in a 1.5-ml cryotube. Store at −80°C. Use the remainder of the culture for plasmid purification with a Qiagen plasmid MidiPrep Kit.

19. Submit a sample of the plasmid DNA for sequencing (Shendure et al., 2011) to confirm construct insertion, and check for any mutations arising during PCR.

    pMal sequencing primers are available from New England Biolabs.

Modify vector for a new peptide construct

Once the vector has been constructed and the sequence has been confirmed, the peptide sequence can be replaced with a new target peptide sequence as follows:

20. Design primers for the cDNA of interest incorporating a forward primer *Bam*HI and a reverse primer *Eco*RI sites.

    Forward: 5′-ACTG*GGATCC*$(NNN)_{6\text{-}10}$-3′
    Reverse: 5′-ACTG*GAATTC***TCA**$(NNN)_{6\text{-}10}$-3′.

    Restriction sites are shown in italics and the STOP codon is in bold.

    Use the primer guidelines described above to help design the primers.

21. Perform PCR (*APPENDIX 4J*) using the primers from step 20 and the new cDNA template.

    Refer to PCR guidelines in step 3.

22. Purify the PCR product (e.g., using a QIAquick method).

23. Digest the purified PCR product with *Bam*HI and *Eco*RI and purify (e.g., using a QIAquick method).

    Refer to digestion guidelines in step 7.

24. Proceed with ligation beginning at step 8.

    The resultant plasmid contains the following elements—MBP-linker-FacXa-TEV-BamHI-gene-of-interest-EcoRI. The above procedure can be modified to include other compatible restriction sites, alternate protease cleavage sites, or additional affinity tags added before or after the target sequence (see Routzahn and Waugh, 2002).

BASIC PROTOCOL 2

EXPRESSION AND PURIFICATION OF FUSION PROTEIN

After successful cloning of a gene of interest into the modified pMal vector, confirmation by sequencing, and transformation into a bacterial strain of choice (Basic Protocol 1), the transformants are screened for expression. We have found that expression levels do not vary widely between transformants, yet occasionally a particular colony shows better

growth and expression. Thus, it is advisable to choose several colonies for expression testing.

The general steps for purification include centrifugation, chromatographic purification of the fusion protein, protease cleavage to liberate the target peptide, and peptide purification. The procedure is optimized for a soluble MBP fusion protein, which is the case for many of our peptide targets. However, the MBP fusion protein may precipitate during the purification procedure. In this case, the chromatography buffers should be supplemented with 10% to 20% glycerol, and the salt concentration (PSE base, see recipe for purification buffer) should be lowered to increase solubility.

Materials

Plate containing positive transformants (from Basic Protocol 1, step 11, or fresh LB+amp plate streaked using glycerol stocks from Basic Protocol 1, step 16)
LB+amp medium: add 1 ml of 100 mg/ml ampicillin stock (see recipe) per liter LB medium (see recipe)
0.5 M isopropyl β-D-1-thiogalactopyranoside (IPTG; filter sterilized, see recipe)
Lysis buffer (see recipe), cold
Amylose resin (maltose affinity resin; New England Biolabs, cat. no. E8021L)
Purification buffer (see recipe)
1× Bradford reagent (BioRad, cat. no. 500-0205)
Elution buffer (see recipe)
0.1% (w/v) sodium dodecyl sulfate (SDS)
0.02% sodium azide (NaN_3) or 20% (v/v) ethanol for column storage
Tobacco Etch Virus (TEV) protease (Sigma-Aldrich, cat. no. T4455)
1 M dithiothreitol (DTT; see recipe)

10-ml sterile culture tubes with caps (Simport, cat. no. T406-2A, *http://www.simport.com/*)
Sterile loop
37°C shaking incubator
50-ml Erlenmeyer flasks with caps or foil coverings, sterile
Spectrophotometer for reading optical density (600 nm)
Refrigerated centrifuge and ultracentrifuge with appropriate bottles/tubes
Cell lysis apparatus: e.g., sonicator, French press, high-pressure homogenizer (Emulsiflex); see Commentary for more information
Nutator or other apparatus for batch incubation of lysate with amylose resin
Gravity purification column and caps (49-ml glass Econo-Columns, BioRad, cat. no. 737-2512)
Tubes for collecting column flowthrough, washes, and eluates
Filter-driven concentrator apparatus (Amicon Stirred Cell, Millipore, cat. no. 5124) and ultrafiltration membranes (MWCO 10,000; Millipore, cat. no. PLGC07610)
16°C incubator for TEV cleavage

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*)

Screen transformants for expression

1. Pick a single transformant from a plate and add to 1.5 ml of LB+amp medium in a loosely capped 10-ml culture tube. Grow overnight at 37°C in a rotary shaker at high speed (≥200 rpm).

   Large scale: add a single colony to 10 ml LB+amp.

2. The following morning, inoculate 100 μl of the overnight culture into 10 ml of LB+amp medium in a 50-ml sterile Erlenmeyer flask and grow at 37°C in a rotary shaker (150 to 200 rpm) until an OD_{600} of 0.4 to 0.6 is reached (typically 3 to 6 hr).

   M9+amp medium [M9 medium (see recipe) containing 100 μg/ml ampicillin] can be used in place of LB (see Commentary for growth medium selection).

   For large scale: add the entire 10 ml overnight culture to 1 liter of LB+amp.

   Also for large scale: the culture can be moved to a lower expression temperature prior to induction (e.g., 18° to 22°C). It is recommended that the culture be placed at the lower temperature for 30 min prior to induction, to ensure equilibration. This is particularly important for poorly expressing constructs.

   A glycerol stock should also be made to preserve this overnight culture.

   To achieve proper aeration, use culture flasks with a medium-to-air ratio of 1:4 or 1:5 (e.g., 10 ml medium in a 50-ml flask)

3. Remove 1 ml of the non-induced sample for later gel analysis.

4. Add the appropriate amount of IPTG (0.1 to 1 mM) from 0.5 M stock.

   For the initial expression test, higher concentrations of IPTG should be used (0.2 mM for LB and 1 mM for M9). After expression has been observed (see below), further expression tests varying the IPTG concentration and the induction time are carried out to achieve the highest expression level. Note that the goal here is to achieve high-level expression of soluble protein and to limit the formation of inclusion bodies or the degradation of fusion protein.

5. Place the culture in a rotary shaker incubator (100 to 150 rpm) at the appropriate temperature for the appropriate time, as determined above.

   Higher protein expression is typically observed when cells are induced at 18° to 22°C for 12 to 48 hr. During the pilot experiment, remove samples at a wide range of time intervals (4, 8, 12, 24, 48 hr) and analyze by SDS-PAGE (UNIT 10.1).

Maltose affinity chromatography

Take samples at each step of the fusion protein purification for SDS-PAGE analysis (Fig. 29.8.2). Unless specifically noted, steps 6 to 14 can be performed at either room temperature or 4°C. 4°C is recommended if protein precipitation is observed. Steps 6 to 14 should also be carried out in 1 day.

6. Harvest bacteria by centrifugation for 15 min at 6000× *g*, 4°C, and resuspend the pellet completely in cold lysis buffer. For cell lysis by sonication, resuspend the cells at ~0.1 to 2 g/ml. For high-pressure cell-lysis methods, use ~0.5 to 0.6 g/ml.

   Using pre-chilled lysis buffer reduces the risk of protein degradation or precipitation following lysis.

   Cell yield depends on induction point (OD_{600}), induction temperature, and induction time. Fusion proteins that express poorly usually yield 3 to 4 g of cells per liter of cell culture (e.g., induce at $OD_{660} = 0.4$ at 18°C for 48 hr). Fusion proteins that express at high levels usually yield >10 g of cells per liter of cell culture (e.g., induce at $OD_{660} = 0.6$ at 37°C for 12 hr).

7. Lyse cells using standard procedures (sonication, Emulsiflex, etc.). Ensure the cells remain at 4°C during lysis.

8. Clarify the lysate by centrifugation for 25 min at 50,000 × *g*, 4°C, to remove unbroken cells and insoluble material.

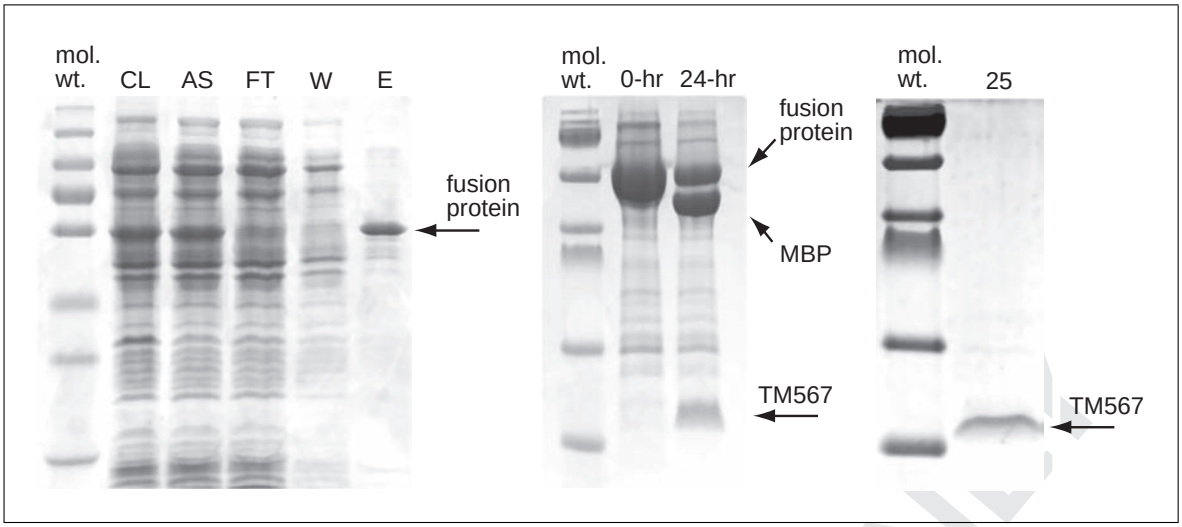

Figure 29.8.2 SDS-PAGE analysis of TM567 purification (peptide design in Fig. 29.8.1). Molecular weight ladder (mol. wt.) as marked, in kDa. Left panel is a 12% Tris-glycine SDS-PAGE gel showing fractions of the purification: CL, crude lysate (5 μl); AS, after lysate ultracentrifugation (5 μl); FT, amylose column flowthrough (5 μl); W, amylose column wash (10 μl); E, amylose column elution (2 μl, ~10 μg). The position of the fusion protein (MBP-TM567) is marked. Center panel is a 16% Tris-Tricine SDS-PAGE gel showing TEV protease digestion progress: 0 hr, concentrated elution before TEV addition; 24 hr, sample after 24-hr digestion at 16°C. The position of the fusion protein (MBP-TM567), cleaved MBP, and the peptide (TM567) are marked. Right panel is a 16% Tris-Tricine SDS-PAGE gel showing 25-μl sample of the organic phase of the organic extraction, dried and then resuspended in sample loading buffer. The position of the peptide (TM567) is marked.

9. Incubate the supernatant (clarified lysate) with prepared maltose affinity resin (amylose resin) in batch mode with gentle agitation for 1 hr.

   The resin is prepared by washing with at least 3 column volumes of purification buffer. If the resin is in 20% ethanol, wash with 1 to 2 column volumes of water followed by 3 column volumes of purification buffer to avoid salt precipitation.

   Gently agitate the resin on a rotary platform device (e.g., a gel rocker or Nutator is recommended; a stir bar may damage the resin).

   25 ml of amylose resin will bind at least 150 mg of fusion protein. This resin can be washed as recommended by the manufacturer and reused at least 10 times (until binding efficiency begins to decrease).

   Between 50 to >400 mg of fusion protein can be produced per liter of cell culture. Typical expression levels of single TM segment fusion proteins is 250 to 300 mg, requiring at least 50 ml of amylose resin for a pilot experiment. If more fusion protein is present than expected, the column flowthrough and wash (see below) can be collected and reapplied to the column.

10. Pour the resin into a 49-ml Econo-Column and allow it to pack by gravity while collecting the flowthrough.

    For ease of use, select a column that leaves enough space to add the entire lysate from step 8. For subsequent purifications, apply fresh lysate (step 8) to cleaned resin in the column (step 9), then cap the column and place it on a rotary platform device for batch mixing and incubation (step 9).

    The flowthrough can be stored at 4°C for a few hours, although it typically does not remain stable overnight (due to the presence of proteases in the lysate). Therefore, if SDS-PAGE analysis reveals that a significant amount of fusion protein remains in the flowthrough (insufficient resin), the flowthrough can be reapplied to clean resin. Since this can only be done on the same day as lysis, it is recommended that more resin be used initially to avoid storing the flowthrough for reapplication.

11. Wash the column with at least 3 column volumes of purification buffer to remove any nonspecifically bound protein.

    To ensure complete washing, a quick qualitative protein assay can be used. Using a micropipet, take 2 μl of liquid from a drop hanging from the tip of the column and mix it with 3 drops of 1× Bradford reagent (BioRad), initially brown in color. If protein is still washing off from the resin, the mixture will turn blue. When all the contaminants have been removed, the mixture will remain brown.

12. Add 0.2 to 0.5 column volumes of elution buffer to the column, mix well (cap the column at both ends and place on a rotary platform device), and allow it to incubate for 20 min.

    This step helps remove the protein in a smaller volume and reduces the later requirement for concentration.

13. Collect the eluent and continue to add elution buffer until all the protein has been eluted (typically a total of 1 to 2 column volumes).

    The same Bradford test described above can be used here, as well.

14. Clean the amylose resin by washing it as follows: 1 column volume water, 3 column volumes of 0.1% SDS, 1 column volume of water, and then 3 column volumes of 0.02% NaN_3 or 20% ethanol.

    Hydrolases present in E. coli cell lysates will lead to amylose resin degradation. Cleaning the column immediately following fusion protein elution maximizes the number of effective purifications one batch of resin will produce.

15. Analyze the purification steps by SDS-PAGE (*UNIT 10.1*) to determine areas of significant loss of fusion protein and quantify the total amount of fusion protein purified.

16. Concentrate the purified fusion protein to ≥5 mg/ml using a filter-driven concentrator (10,000 MWCO).

    This improves TEV cleavage as the enzyme is more effective at higher concentrations of fusion protein.

    The eluent may be stored up to one month at 4°C before or after concentration.

Perform protease digestion

17. Add 10 U of TEV per 100 mg of fusion protein and 1 mM (final) of DTT to the concentrated eluent and incubate at 16°C until cleavage is completed.

    Cleavage is typically complete after 24 to 72 hr of incubation, although it may take longer for some fusion protein constructs. If longer times are required, add 1 mM fresh DTT every 2 or 3 days. Alternately, add more TEV protease if poor digestion (<30%) is observed after 3 days.

    The effectiveness of digestion is observed by a mobility shift from fusion (>42 kDa) to free MBP (~42 kDa) by SDS-PAGE (UNIT 10.1). 8% acrylamide should be sufficient to observe a mobility shift of ≥2 kDa. See Figure 29.8.2.

BASIC PROTOCOL 3

PEPTIDE RECOVERY

At this stage (following Basic Protocol 2), there are many possible ways to remove MBP and purify the peptide. Here, we focus on methods that purify the peptide from a pellet of precipitated protein. However, the best method to use for purification needs to be determined empirically depending on how a particular peptide behaves during the purification protocol (e.g., the solubility of the peptide at different stages). If the target peptide precipitates following protease cleavage, high-speed centrifugation (45 min at 100,000 × *g*, 4°C) may be used to collect a pellet enriched in the peptide.

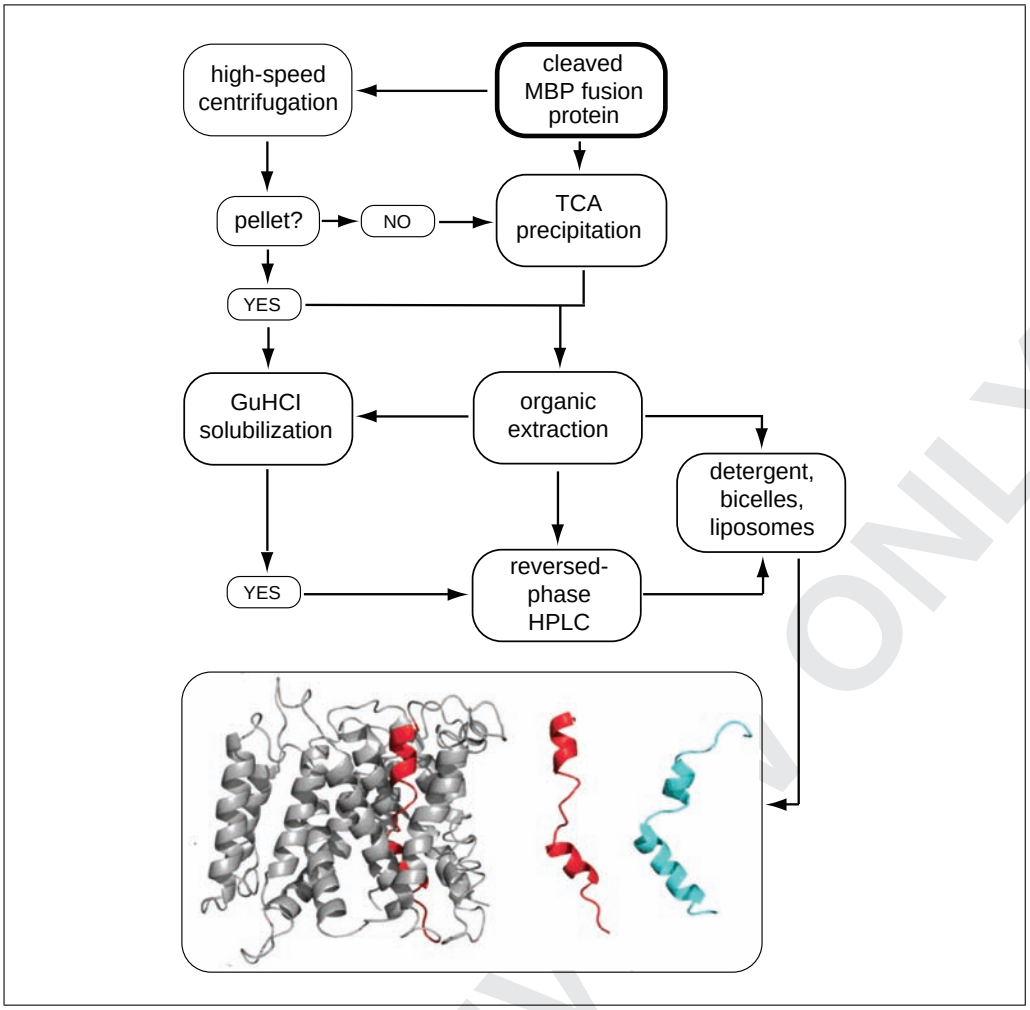

Figure 29.8.3 Workflow for peptide purification for structural studies. Begin with 'Cleaved MBP fusion protein' and follow the arrows. The final box of the flowchart shows, on the left, a cartoon representation of a three-dimensional model of the human sodium proton exchanger isoform 1 (Landau et al., 2007) with the sequence of TM VI highlighted in red. In the center is shown the predicted structure of TM VI (red) and on the right is shown the solution NMR structure of TM VI for comparison. For the color version of this figure, go to *http://www.currentprotocols.com/protocol/ps2908*.

This is ideal for quantities of fusion protein in excess of 200 mg without glycerol in the buffer. The peptide pellet may then be purified by organic extraction (as described in the protocol below) or by reversed-phase HPLC (Alternate Protocol 1). If the peptide does not precipitate on its own, trichloroacetic acid (TCA) is used to precipitate all protein. The peptide is then extracted from this pellet by differential solubilization with guanidine hydrochloride (GuHCl) or organic extraction. Additionally, a combination of techniques may be required to isolate the peptide, depending on the sample, and it may be necessary to try various protocols on aliquots of the sample. Figure 29.8.3 depicts a flowchart of these possible combinations.

Materials

Protease digestion reaction (output of Basic Protocol 2)
60% trichloroacetic acid (TCA)
Chloroform
Isopropanol

Centrifuge, rotor, and glass tubes capable of centrifugation at 9000 × *g*
Glass rod

Metal spatula to resuspend precipitated protein pellet
Glass Dounce homogenizer (Fisher Scientific, cat. no. FB56699)
??-ml separatory funnel
Teflon-lined screw-cap glass tubes for collecting organic extract
Rotary evaporator
Nitrogen gas tank
Heat block and/or lyophilizer

Additional reagents and equipment for Tris-Tricine SDS-PAGE (*UNIT 10.1*)

IMPORTANT NOTE: The following steps should be done using ONLY chloroform-insensitive materials like glass, metal, and Teflon. Many plastics dissolve in chloroform and will contaminate the sample. To simplify the following steps, all the volumes given are for 100 mg of fusion protein. Increase the volumes accordingly for different starting amounts.

1. After protease digestion is complete (output of Basic Protocol 2), transfer the reaction mixture to a glass centrifuge tube and slowly add 2.5 ml of 60% TCA while swirling to precipitate the protein (see Video 29.8.1, "TCA precipitation," at *http://www.currentprotocols.com/protocol/PS2908*).

   Slow addition of concentrated TCA helps prevent the formation of large aggregates of precipitate, aiding in the later extraction step.

2. Incubate this mixture on ice for 30 min to complete the precipitation.

3. Collect the precipitate by centrifuging 30 min at 6000 × *g*, 4°C.

   Glass centrifuge tubes may break at RCF values over 9000 × g; therefore, lower speeds (for longer times) are required.

4. Gently rinse the pellet three times with purified water, being careful not to disturb the pellet. If proceeding with HPLC purification, go to Alternate Protocol 1.

5. Prepare fresh organic solvent mixture by combining 16 ml of chloroform with 16 ml of isopropanol. Mix well with a glass rod.

   The final extraction solution will contain 5:5:1 (v/v/v) chloroform:isopropanol:water.

   Adding the denser chloroform to the isopropanol will make complete mixing easier.

   NMR experiments of the peptide directly in the organic extraction mixture may be possible for some samples. If this is desired, deuterated solvents must be used to prevent a strong solvent signal from overshadowing the peptide signal. Due to the expense of deuterated solvents, it is recommended that the volume of solvent mixture be reduced by up to 5-fold. Although this reduces the efficiency of extraction and decreases the total yield of peptide, it greatly increases the concentration of peptide in the organic solvent and consumes less solvent.

6. Add 25 ml of the solvent mixture to the precipitated protein and incubate it on ice for 15 min.

7. Using a metal spatula, mince the pellet and gently scrape the walls of the tube until the entire pellet has been resuspended (see Video 29.8.2, "TCA pellet resuspension," at *http://www.currentprotocols.com/protocol/PS2908*).

   Achieving a fine suspension aids in the later steps.

8. Pour the mixture into an appropriately sized all-glass Dounce homogenizer and homogenize until the solution is uniform in appearance (see Video 29.8.3, "Homogenization," at *http://www.currentprotocols.com/protocol/PS2908*).

CAUTION: Care should be taken to avoid splashing the solvent during homogenization. A fume hood is recommended to avoid chloroform fume inhalation.

Homogenization can be done over the course of 30 min or more, allowing the homogenizer to rest on ice in between homogenization sessions.

A fibrous pellet of insoluble denatured protein will appear at the bottom of the homogenizer (whitish in appearance).

9. Add 3.2 ml of pure water and continue to homogenize.

   The pellet may become more translucent and fibrous.

   This may be done over the course of 30 min or more, allowing the homogenizer to rest on ice in between homogenization sessions.

10. Pour the entire contents of the homogenizer into a ??-ml separatory funnel. Rinse the centrifuge tube and the homogenizer with the remaining solvent mixture (~7 ml) and pour into the funnel.

11. Shake and vent the funnel every 30 min for 2 hr or more and allow the layers to separate for >12 hr at ??ºC.

    If convenient, the mixture may be shaken and vented intermittently for the remainder of the day and left to separate overnight.

12. After the layers have separated completely, collect the organic (bottom) layer into a clean ??-ml separatory funnel. Add 3.2 ml of purified water and repeat step 11.

    The majority of remaining insoluble material and the aqueous layer should remain in the funnel but as the extraction is being repeated ensure that the entire organic layer is collected.

13. Collect the bottom layer, avoiding the collection of any of the upper aqueous phase, and store it in a glass container with a Teflon-lined or ground-glass closure.

    This sample may be used directly for some experiments such as mass spectrometry or NMR (see step 5).

    Step 12 can also be repeated to increase sample purity.

14. Recover the peptide by removing the organic solvent using a rotary evaporator, under a gentle stream of nitrogen (or argon) gas, or by lyophilization. Analyze the purity of the sample by Tris-Tricine SDS-PAGE (Schägger, 2006; Fig. 29.8.2).

    At this point, the sample may be used directly, or further purification may be required (proceed to Alternate Protocol 1). It is recommended that the solubility of the peptide be tested in detergent solution or more polar solvent systems. It may also be possible to reconstitute the dried peptide into liposomes by various methods (e.g., Reddy et al., 1995).

    To test for solubility in detergent, dry 100 μg of peptide to a thin film on the wall of a glass test tube. This can be done by gently blowing a stream of nitrogen gas over the solvent while vortexing.

    Add water (100 to 200 μl) and heat the solution (in the range of 37° to 50°C) for 30 min. The peptide should become flocculent.

    Bring the solution to room temperature and add buffer components of choice (buffer, pH, salts) followed by detergent (e.g., 0.7 mg DPC) and vortex vigorously for 3 min. At this point, the clarity of the solution is a good indicator of peptide solubility.

ALTERNATE PROTOCOL 1

HPLC

If proceeding to HPLC purification, the starting material for this process can be a 100,000 × *g* pellet, a TCA precipitated pellet, or the dried sample from organic extraction. The pellet is first solubilized in GuHCl as described below. However, peptides purified by any of the methods described above may become insoluble following solvent removal (e.g., lyophilized peptides may not be soluble in chloroform/isopropanol, GuHCl, SDS, etc.). However, solvent exchange can be achieved by partly drying the organic extraction followed by the addition of trifluoroethanol (TFE). This partial drying and addition of TFE is repeated until the sample is only in TFE. If the peptide is soluble in TFE, it can then be injected onto an HPLC column.

Materials

Protein pellet (?protocol?; step?)
1 M and 7 M buffered GuHCl (see recipe)
Reversed-phase solvents:

- Solvent A (H_2O-TFA; see recipe)
- Solvent B (isopropanol-TFA; see recipe)

Dounce homogenizer (Fisher Scientific, cat. no. FB56699)
Refrigerated centrifuge, rotor and bottles capable of 10,000 × *g* at 4°C
Reversed-phase semi-preparative HPLC column (Zorbax SB300 C8 semi-preparative column, Agilent Technologies, cat. no. SB300 C8)
HPLC equipped with column heater and fraction collector
Additional reagents and equipment for Tris-Tricine SDS-PAGE (*UNIT 10.1*) and reversed-phase separation of peptides (*UNIT 8.7*)

1. Resuspend the pellet in 20 ml of 1 M buffered GuHCl per 100 mg of fusion protein using a spatula and vigorous vortexing followed by Dounce homogenization.
2. Centrifuge the sample 20 min at 10,000 × *g*, 4°C, to pellet insoluble material. Set aside the supernatant for Tris-Tricine SDS-PAGE analysis (*UNIT 10.1*).

   This step should remove any remaining soluble impurities such as MBP and TEV from the sample.

   The supernatant should be analyzed by Tris-Tricine SDS-PAGE (UNIT 10.1) to ensure that no peptide loss is occurring (Schägger, 2006).
3. Resuspend the pellet in ~5 ml of 7 M buffered GuHCl per 100 mg of fusion protein.

   Keeping the sample volume low will aid in HPLC column loading.
4. Centrifuge the sample 20 min at 10,000 × *g*, 4°C. Retain the supernatant.
5. Purify the TM peptide using reversed-phase HPLC (also see *UNIT 8.7*).

   Hydrophobic peptide samples in GuHCl can be directly injected onto a reversed-phase HPLC semi-preparative column (e.g., a Zorbax SB300 C8 column, Agilent Technologies) heated at 60°C using 100% Solvent A. Absorbance is monitored at 220 and 280 nm wavelengths.

   Some peptides may be difficult or impossible to elute from the column, thereby reducing the life of the HPLC column. This is difficult to predict a priori (Table 29.8.1), although it may depend on the length, hydrophobicity, and secondary structure of a particular peptide. We find water-isopropanol-TFA gradients to be far superior to water-acetonitrile-TFA.

   Extensive washing of the column with high concentrations of isopropanol-TFA is required to remove unwanted contamination from peptides and GuHCl between runs. It is not

Table 29.8.1 Summary Data for Some of the Peptide Constructs Studied

TM segment[a]	No. of amino acids[b]	Sequence[c]	Mol. wt.	Pure[d]	Hydrophobicity[e]
Human phospholamban (PLN) and sarcolipin (SLN) (Douglas et al., 2005***)***					
PLN	51 (2)	residues 2-52 GSEKVQYLTRSAIRRASTIEMPQQARQK LQNLFINFCLILICLLLICIIVMLL	5977	Yes	51.1 (0.96)
SLN	31 (12)	residues 1-31 MGINTRELFLNFTIVLITVILMWLLVR SYQY	3761	Yes	0.7 (0.02)
Human Na^+/H^+ exchanger (NHE1)					
TM4	26	residues 155-180 FLQSDVFFLFLLPPIILDAGYFLPLR (Slepkov et al., 2005)	3056	Yes	40.9 (1.57)
TM6	25 (6)	residues 226-250 KKKDNLLFGSIISAVDPVAVLAVFEEIH KKK (Tzeng et al., 2010)	3438	CS	–10.2 (–0.33)
TM7	25 (2)	residues 250-274 HINELLHILVFGESLLNDAVTVVLYKK (Ding et al., 2006)	3079	CS	8.2 (0.30)
TM9	28 (3)	residues 338-365 KSYMAYLSAELFHLSG IMALIASGVVMRPKK (Reddy et al. 2008)	3413	CS	6.5 (0.21)
TM11	25	residues 447-472 KDQFIIAYGGLRGAIAFSLGYLLDKK (Lee et al., 2009)	2858	Yes & CS	5.2 (0.21)
TM3,4	55 (8)	residues 126-180 GSKKKSSIVPESCLLIVVGLLVGGLIK GVGETPPFLQSDVFFLFLLPPI ILDAGYFLPLRKKK	6798	No	10.5 (0.17)
TM6,7	49 (8)	residues 226-274 GSKKKDNLLFGSIISAVDPVAVLAVFEEI HINELLHILVFGESLLNDAVTVV LYKKK	6249	Yes	15.4 (0.27)
TM7-8	64 (8)	residues 250-313 GSKKKHINELLHILVFGESLLNDAVTVVL YHLFEEFANYEHVGIVDIFLGFLSFFVVAL GGVLVGVVYGKKK	7992	No	39.4 (0.55)
TM5,6,7 (Figs. 29.8.1 and 29.8.2)	96 (9)	residues 184-279 GSGGGKKENLGTILIFAVVGTLWNAFFLG GLMYAVCLVGGEQINNIGLLDNLLFGSII SAVDPVAVLAVFEEIHINELLHILVFGE SLLNDAVTVVLYHLFEEKK	10429	Yes	65.6 (0.62)
TM6,7,8	88 (8)	residues 226-313 GSKKKDNLLFGSIISAVDPVAVLAVFEEI HINELLHILVFGESLLNDAVTVVLYHL FEEFANYEHVGIVDIFLGFLSFFVVALG GVLVGVVYGKKK	10506	No	63.8 (0.66)

continued

Table 29.8.1 Summary Data for Some of the Peptide Constructs Studied *continued*

TM segment[a]	No. of amino acids[b]	Sequence[c]	Mol. wt.	Pure[d]	Hydrophobicity[e]
Lan8,9,10	102 (8)	residues 316-417 GSKKKAAFTSRFTSHIRVIEPLFVFLYSY MAYLSAELFHLSGIMALIASGVVMRPY VEANISHKSHTTIKYFLKMWSSVSETLI FIFLGVSTVAGSHHWNWTFVISTKKK	12523	No	21.4 (0.19)

[a]TM predictions for NHE1 are based on the model of Wakabayashi et al. (2000), except for Lan8,9,10 based on the model of Landau et al. (2007).
[b]Number of residues from the primary structure of NHE1 that were included in peptide design. Parentheses indicate that additional residues were included (capping lysine residues, sequence changes associated with the TEV site, and/or a His tag).
[c]Underlined residues indicate amino acids not in the native sequence that were added to increase solubility or remaining after TEV cleavage.
[d]Indicates whether or not the construct could be expressed and purified. CS indicates chemical synthesis was used.
[e]Calculated using the hydrophobicity scale of Liu and Deber (1998). A larger numerical value indicates a more hydrophobic construct. The values indicated were calculated for the sequence shown. In parentheses are the values normalized for the number of residues.

unusual to have multiple HPLC columns in use, where each column is dedicated to the purification of a single peptide or a set of related peptides.

Heating the column to 60°C during the purification assists in peptide elution and prevents excessively high column pressures caused by the viscosity of isopropanol.

20 to 40 mg of peptide (≤7 ml at 2-8 mg/ml) in 7 M GuHCl can be injected onto the SB300-C8 column using multiple injections depending on the sample loop volume.

6. Run 20% Solvent B for 10 min at 1 ml/min and 60°C.
7. Run a 200-min gradient from 20% to 80% Solvent B at 1 ml/min and 60°C, collecting 2-min (2-ml) fractions.

   Typically, transmembrane peptides of interest will elute between 30% and 60% Solvent B.

8. Clean the column by running four alternating 5-min washes each with 100% Solution A and 90% Solution B. To ensure that all residual peptide and GuHCl are removed before subsequent purifications, perform a mock injection and blank run (with no peptide).

   Reversed-phase HPLC is a good "polishing" method for getting a highly pure peptide that can be solubilized in detergent or reconstituted into liposomes. However, some peptides may not separate or elute efficiently using this method, and organic extraction may serve as an alternative purification option. If both methods fail, it may be necessary to modify the peptide design and begin anew (see Commentary).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Ampicillin stock, 100 mg/ml

Dissolve 1 g of ampicillin in 10 ml of pure water, filter sterilize into a sterile container, and aliquot at 1-ml volumes in sterile microcentrifuge tubes. Store up to 1 year at −20°C.

Buffered guanidine hydrochloride (GuHCl) solutions

First prepare 8 M GuHCl by dissolving 76.424 g of guanidine hydrochloride in 100 ml (final volume) of H_2O

For 1 M buffered GuHCl:
3.252 ml 8 M GuHCl
12.5 ml 0.1 M sodium phosphate buffer, pH 8 (see recipe)
Dilute to 25 ml with H_2O
For 7 M buffered GuHCl:
21.875 ml 8 M GuHCl
2.5 ml 0.1 M Na_2HPO_4 (not buffered; the final pH will be around 7.15)
Store solutions up to 1 year at room temperature

IPTG, 0.5 M

Dissolve 1.19 g of isopropyl β-d-1-thiogalactopyranoside in 10 ml of pure water, filter sterilize into a sterile container, and aliquot into 1-ml volumes in sterile microcentrifuge tubes. Store up to 1 year at −20°C.

LB medium

In 800 ml of H_2O, dissolve:
10 g tryptone (BD Difco)
5 g yeast extract (BD Difco)
10 g NaCl
Adjust pH to 7.5
Dilute to a final volume of 1000 ml with pure H_2O
Store up to 1 year at room temperature

LB+amp plates

Prepare LB medium (see recipe), except add 1 g agar per 100 ml LB before autoclaving. After autoclaving, allow liquid to cool to ~50°C. Add 1 ml of 100 mg/ml ampicillin stock (see recipe) per liter of agar-containing medium (100 μg/ml final), mix gently (to prevent bubble formation), then pour into sterile Petri plates. Allow agar to completely set for at least 60 min. For storage, turn the plates upside down and allow them to dry overnight on the bench. The following day, stack the plates upside down into their original bag and store at 4°C. The plates are stable for at least 30 days.

Purification buffer

First, prepare the following stock solutions:
10× PSE (phosphate-sodium-EDTA)—dissolve the following in 1 liter (final volume) of H_2O:
1.462 g Na_2HPO_4
26.88 g NaH_2PO_4
70.2 g NaCl
0.372 g EDTA
50% (v/v) glycerol: Dilute 500 ml glycerol to 1 liter with H_2O
5% (w/v) NaN_3: Dissolve 5 g of NaN_3 in 100 ml (final volume) of H_2O
500 mM maltose: Dissolve 90 g of maltose in 500 ml (final volume) of H_2O, autoclave
To prepare the purification buffer, mix together the following and dilute to 1 liter with H_2O:
100 or 50 ml 10× PSE
4 ml 5% NaN_3
400 ml 50% (v/v) glycerol (optional, increases protein stability)
Store up to 1 year at room temperature

Using less PSE may increase protein solubility.

Lysis buffer

Add 25 μl Halt protease inhibitor cocktail (Thermo Scientific, cat. no. 78425) and 100 μl of 1 M DTT (see recipe) to 100 ml of purification buffer (see recipe). Make fresh daily.

Elution buffer

Dilute 12 ml 500 mM maltose to 100 ml with purification buffer (see recipe; 60 mM maltose final). Make fresh daily.

DTT, 1 M

Dissolve 1.54 g of dithiothreitol (DTT) in 10 ml of pure water. Aliquot into 1-ml volumes in microcentrifuge tubes. Store up to 1 year at −20°C.

M9 minimal medium

For 1 liter mix together:
1 g $(NH_4)_2SO_4$
100 ml 10× M9 salts (see recipe)
673 ml H_2O
Autoclave and allow to cool. Using sterile technique add:
200 ml 5× phosphate solution (see recipe)
1 ml 1000× metal mix (see recipe)
1 ml 1% (w/v) thiamine
25 ml 20% (w/v) glucose
Store up to 1 year at room temperature

Casamino acids (0.7% (w/v) final) can also be added for non-labeled protein preparations. Make a 10% (w/v) stock solution, autoclave and add sterilely to M9 medium.

M9 salts, 10×

Dissolve the following in 1 liter (final volume) of H_2O:
128 g $Na_2HPO_4 \cdot 7H_2O$ or 67.8 g anhydrous Na_2HPO_4
30 g KH_2PO_4
5 g NaCl
Filter through an 0.45-μm filter
Store up to 1 year at room temperature

Metal mix, 1000×

Dissolve the following in 100 ml (final volume):
500 mg $MnSO_4$
92.5 mg $FeSO_4 \cdot 7H_2O$
5 g $MgSO_4 \cdot 7H_2O$
50 mg $CaCl \cdot 2H_2O$
Autoclave
Store up to 1 year at room temperature

Phosphates, 5×

Dissolve the following in 800 ml of H_2O:
53 g K_2HPO_4
24.7 g KH_2PO_4
Adjust pH to 7.5
Dilute to 1 liter
Autoclave
Store up to 1 year at room temperature

Sodium phosphate buffer, 0.1 M, pH 8

Prepare the following stock solutions:
0.2 M NaH_2PO_4, dissolve 5.52 g of NaH_2PO_4 in 1 liter (final volume) of H_2O
0.2 M Na_2HPO_4, dissolve 5.68 g of Na_2HPO_4 in 1 liter (final volume) of H_2O
Mix 8.5 ml of 0.2 M NaH_2PO_4 and 91.5 ml of 0.2 M Na_2HPO_4
Store up to 1 year at room temperature

Reversed-phase HPLC solvents

Solvent A (water-TFA): Add 2 ml of trifluoroacetic acid (TFA) to 4 liters of HPLC-grade water and thoroughly degas under vacuum. Store up to 1 year at room temperature.

Solvent B (isopropanol-TFA): Add 2 ml of trifluoroacetic acid (TFA) to 4 liters of HPLC-grade isopropanol and thoroughly degas under vacuum. Store up to 1 year at room temperature.

COMMENTARY

Background Information

Many useful resources for membrane protein expression in heterologous systems are available, including bacteria (Miroux and Walker, 1996; Newby et al., 2009), yeast (Daly and Hearn, 2005; Jidenko et al., 2005; Cregg et al., 2009), insect cells (Trometer and Falson, 2010), mammalian cells (Lundstrom, 2010), and cell-free systems (Klammt et al., 2007). Herein, we present an application of the 'divide and conquer' approach using, as an example, a human membrane transport protein that is a recognized clinical target in heart disease and cancer, the human Na^+/H^+ exchanger isoform 1 (Karmazyn et al., 2003). The basic tenet of the approach is that well defined pieces of a membrane protein retain their native structure, yet they are easier to express, purify, and analyze. Several groups including ours have successfully applied this method (Chopra et al., 2000; Katragadda et al., 2001; Hu et al., 2007; Bordag and Keller, 2010; Lee et al., 2011; Kocherla et al., 2012).

One of the central weaknesses of this approach is the correct selection of a TM segment or subdomain in the absence of a validated topology model for the target membrane protein (Fig. 29.8.1). Fortunately, there have been huge improvements in predictive algorithms that use both empirical and bioinformatic information to predict protein topology, secondary structure, and even tertiary fold (many are listed at *http://www.expasy.org*). These tools, combined with the wealth of biochemical data on many important membrane protein targets, can be used as a basis for the design of hypothetical TM segments and subdomains that can be more readily studied in vitro (Rost et al., 1995). Of course, dividing any protein into pieces brings to question whether the structure of this piece in isolation is relevant to the full-length physiological structure. While the inability to collect functional data on a membrane protein fragment prevents any guarantee of proper folding, there is ample support for the idea that isolated membrane segments and loops can remain properly folded (Hunt et al., 1997). This is perhaps best exemplified by the numerous pioneering studies of bovine rhodopsin, a G protein–coupled receptor (Albert and Yeagle, 2002). Using chemically synthesized peptides and nuclear magnetic resonance (NMR) spectroscopy, the structure of the carboxy-terminal domain of rhodopsin was determined first (Yeagle et al., 1995a). This success was followed by the structures of all four individual cytoplasmic loops (Yeagle et al., 1995b, 1996, 1997) and individual TM segments (Chopra et al., 2000; Yeagle et al., 2000; Katragadda et al., 2001). Remarkably, the structure derived from the NMR data was very similar to the crystal structure (Palczewski et al., 2000; Albert and Yeagle, 2002). Indeed, as more membrane protein crystal structures have been determined, the structures of individual TM segments appear to retain their native structures (Duff and Ashley, 1992; Reddy et al., 1993; Katragadda et al., 2000, 2001). The demonstration that individual TM segments in isolation may provide structural information reflecting the intact protein has led many researchers to begin studying individual TM segments.

Critical Parameters and Troubleshooting

Peptide design

There are many considerations and pitfalls in designing a peptide fragment when the membrane protein of interest has a complex topology. Unless the membrane protein target is sufficiently small [e.g., phospholamban (Simmerman et al., 1986) or sarcolipin (Wawrzynow et al., 1992)], or has a well defined topology [e.g., GPCRs (Bhave et al., 2003)], the ends of the putative membrane-spanning regions are ambiguous. TM α-helices are generally considered to be 20 residues in length, yet the available membrane protein structures reveal that TM regions can be tilted, kinked, unwound, or buried shorter or elongated helices. They can also cause local compaction or stretching of the membrane bilayer, leading to a shorter or longer than expected TM segment (De Planque and Killian, 2003; Cybulski and de Mendoza, 2011). Nonetheless, many accurate predictive tools exist for the in silico study of membrane proteins owing to quickly expanding bioinformatic databases and improved sequence and structural alignment algorithms (Punta et al., 2007). No algorithm is able to correctly predict every TM segment, and it is therefore recommended to run a few different algorithms and then manually analyze the sequence to pick out the most likely candidate (Fig. 29.8.1). At this point, one needs to consider all biochemical information available for the membrane protein target. For example, cysteine-scanning mutagenesis combined with in vivo labeling by membrane-permeant and -impermeant reactive compounds can help to define the limits of TM segments and extra-membranous loops (Akabas et al., 1992; Tang et al., 1998; Wakabayashi et al., 2000). In addition, TM segments containing highly conserved residues that are critical for membrane protein function are often chosen first (Slepkov et al., 2005). Based on these analyses, a putative TM segment is chosen including at least three extramembrane residues on the N- and C-termini. Often, limited solubility of the TM peptide in detergents or organic solvents is encountered, in which case it is advantageous to either lengthen the extra-membranous portions of the peptide or include lysine residues on the N- and C-termini of the peptide. In the latter case, this has been shown not to affect function (Afara et al., 2006, 2008), and to increase peptide solubility and the likelihood of insertion in membrane mimetics such as detergent micelles, bicelles, and liposomes (Melnyk et al., 2003). Many other considerations of peptide design reviewed by others (e.g., Cunningham and Deber, 2007) may also be applicable.

Figure 29.8.1 depicts an example of peptide design using the human Na^+/H^+ exchanger isoform 1. Functional and structural data from mutagenesis and NMR studies (Wakabayashi et al., 2000; Ding et al., 2006; Tzeng et al., 2010), as well as two conflicting three-dimensional models of the membrane domain of the protein (Landau et al., 2007; Nygaard et al., 2011), sparked interest in the region around amino acids 180 to 280. In this region, TMHMM analysis (Krogh et al., 2001) of the sequence predicts 3 TM segments (12 TM segments in total). In Figure 29.8.1, the sequence of residues 184 to 279 is displayed along with known functional and structural data, and three separate TM predictions: TMHMM2, JPred3 (Cuff and Barton, 1999), and evolutionary conservation/fold alignment (Landau et al., 2007). The variability in predictions illustrates the need to use the available biochemical data combined with multiple prediction methods when designing a TM construct.

Peptide production

Once the peptide is designed, there are many choices on how to obtain suitable quantities for structural analyses. Peptide synthesis is the simplest approach, though the strong hydrophobicity of many TM peptides presents challenges for their chemical synthesis and subsequent purification. Consequently, our experience has been that commercial peptide synthesis sometimes fails, and some companies may refuse to synthesize very hydrophobic peptides. Additionally, synthesis of peptides longer than 20 to 30 amino acids is costly and more problematic than shorter ones. If one chooses to use an expression system, as is our preference, there are different strategies for the expression of small hydrophobic proteins. One choice is to use a simple affinity tag such as hexahistidine (His_6), where the hydrophobic peptides tend to aggregate and accumulate in inclusion bodies (Kane and Hartley, 1988), necessitating purification under denaturing conditions followed by refolding into detergent or lipid suspensions (reviewed by Singh and Panda, 2005). Although many peptides can be purified and properly refolded in this manner **(*UNIT 4.8* & *UNIT 28.5*), their hydrophobic nature can promote aggregation

and low peptide yields (De Bernardez Clark, 1998). Compounding this issue, hydrophobic peptides are generally toxic to the cell, which leads to down-regulation of expression or targeted degradation. An attractive alternative is to use a large soluble tag linked to the hydrophobic TM peptide. The advantage is that a naturally abundant and highly soluble protein can be used both to increase expression and to maintain the TM segment in a state suitable for purification. Ideally, when fused to the hydrophobic peptide, the large soluble protein tag would allow purification as a soluble protein, avoiding all of the caveats and complications associated with inclusion bodies.

Several solubility-enhancing fusion protein systems are commercially available, with the most popular being MBP (Kapust and Waugh, 1999), glutathione *S*-transferase (GST; Bichet et al., 2000), Mistic (Roosild et al., 2005), NusA, and small ubiquitin-like modifier (SUMO; Zuo et al., 2005). GST is a small tag—about 26 kDa—that modestly increases the solubility of the target protein but allows facile and specific affinity purification. SUMO and Mistic are recently characterized tags that have both been shown to be effective in increasing solubility of membrane proteins for expression and purification. MBP (42 kDa) has proven to be a robust system for TM peptide expression (Douglas et al., 2005; Hu et al., 2007; Lee et al., 2011; Ceholski et al., 2012). This is because MBP has the remarkable ability to resist the denaturation forces imposed by fusion to an insoluble peptide sequence, thereby allowing high cellular expression levels of a soluble protein construct (Kapust and Waugh, 1999; Nallamsetty and Waugh, 2006). In our early efforts to purify phospholamban, a short single TM protein, an MBP fusion protein maintained a soluble state (Douglas et al., 2005), while a GST fusion was found in inclusion bodies.

Expression and purification of fusion protein

The next major consideration is bacterial strain and growth conditions. Due to its favorable expression, *Escherichia coli* strains like DH5α, normally reserved for cloning, are sometimes sufficient for expressing large amounts of fusion protein. Unfortunately, this step is largely empirical, depending on the construct. We have had success with DH5α, Rosetta DE3, XL1-Blue, BL21-PLysS, and TB1. TB1 (JM83) is recommended by New England Biolabs (the manufacturer of the pMal vector system) for expression. Another important consideration is the type of medium the cells are grown in. For a strain that expresses well, commonly used rich complex broths like Luria-Bertani (LB) are low cost and typically result in high protein yields. However, if structural studies are being pursued, M9 minimal medium is recommended because it can be conveniently modified to allow isotopic labeling. In addition, more difficult-to-express constructs often express better in M9 minimal medium than in LB. Other media considerations to account for biosynthetic deficiencies may also be required, e.g., TB1 requires supplementation with proline when grown in minimal medium. Generally, minimal medium is supplemented with building blocks, such as vitamins and amino acids, to improve bacterial growth. For specific labeling strategies, such as the use of N^{15} for collecting two-dimensional NMR spectra of a target peptide, minimal medium can be generated with N^{15}-labeled ammonium chloride as the sole source of nitrogen. Once fusion protein expression is achieved with a construct-strain-medium combination, several other variables can be optimized to improve the quality and quantity of the expressed material. Although bacteria are typically grown at 37°C, inducible expression of exogenous proteins at this temperature often results in little or no fusion protein. Lowering the temperature during induction changes the cell state to favor protein synthesis (Jones and Inouye, 1994). Our usual protocol is as follows. The cell culture is incubated at 37°C until the beginning of logarithmic growth, corresponding to an OD_{600} of ~0.6 to 0.8 (for difficult constructs, an earlier induction point, OD_{600} of 0.4 to 0.5, may be preferable). The culture is shifted to 22°C and equilibrated for 30 min. Once the culture reaches 22°C, the inducing agent is added [in our case isopropyl β-d-1-thiogalactopyranoside (IPTG)], and the cells are incubated for 12 to 72 hr. The concentration of inducing agent can also affect fusion protein quality and quantity. For IPTG, a range of 0.1 to 1 mM may be tested to optimize protein yield. Of course, the goal here is to optimize the quality and the level of protein expression (per cell) as well as the total cell mass, such that large quantities of material suitable for structural analyses are obtained. Induction of expression at lower cell density, lower temperatures, and lower IPTG concentrations will require longer induction times (up to 72 hr), while higher cell densities, temperatures, and IPTG concentrations will generally require shorter induction times

(e.g., >1 mM IPTG at 37°C usually maximizes expression within 2 to 4 hr). Some constructs are stable and express well with overnight induction at 37°C, whereas other constructs degrade at 37°C and require longer induction times at lower temperature. It is difficult to predict the expression conditions for a particular fusion protein construct, so a range of conditions should be empirically tested. Once a working expression protocol is determined, further optimization for large-scale cultures may be performed. It is recommended that this be done using 1-liter culture volumes, such that scale-up simply involves multiple 1-liter flasks, and variables like aeration and growth times remain consistent between cultures.

Following overexpression, the purification scheme invariably starts with cell lysis, regardless of whether the target protein is found in inclusion bodies, cell membranes, or the soluble cytosolic fraction. Various methods of lysis exist, with three of the most popular being sonication, mechanical, and high pressure. Sonication is the least delicate method and introduces a significant amount of heat into the sample. Nonetheless, the MBP-fusion protein is stable if care is taken to ensure that the cells remain cold. By limiting continuous sonication time and allowing the cells to cool on ice in between repetitive sonication steps, sonication is a very effective way to isolate MBP-fusion proteins from the cytosol. However, sonication becomes less effective for larger volumes of cell suspension. A more favorable and gentler method is high-pressure lysis with a French Press (*UNIT 6.2*) or Emulsiflex (Avestin, *http://www.avestin.com/*). These systems require a minimum suspension volume and are more effective at higher cell densities. Alternatively, several mechanical disruption apparatus that use various cell volumes exist such as bead-beaters and mills. These are also gentler than sonication, but they introduce heat into the sample, so care should be taken to keep the sample cool. Following lysis, the purification scheme of MBP-fusion proteins involves isolation of the cytosolic fraction, maltose affinity chromatography, protease cleavage, and peptide purification (via organic extraction, and/or HPLC purification). Since MBP is such a robust system, little optimization of the fusion protein purification should be required. However, some constructs may express well as soluble proteins in the cytosol, but become unstable during purification. If the fusion protein precipitates, supplement the affinity chromatography buffers with 10% to 20% glycerol and decrease the salt concentration by using 0.5× PSE (see recipe for purification buffer). Additionally, all purification steps can be carried out at 4°C. Interestingly, the addition of small amounts of Triton X-100 detergent (≤0.1% w/v) did not increase fusion protein solubility. This may be due to the fusion protein forming very large stable oligomers that sequester the hydrophobic peptide in the core of the oligomer (unpub. observ.). Adding glycerol and decreasing the salt concentration may stabilize this oligomer, whereas detergent destabilizes it. Following fusion protein purification, we have occasionally experienced difficulty achieving complete cleavage of some constructs. Since the fusion protein and peptide should be stable in the elution buffer, samples can be left to digest for a week if required. Additional enzyme and fresh DTT every 3 days may also help. If cleavage is still poor, the linker between MBP and the peptide may be redesigned and lengthened. We have found that insertion of two additional glycine residues in between the TEV site and the peptide often improves cleavage efficiency.

Peptide recovery

The recovery of pure peptide is the most challenging step. After removal of the MBP tag, the peptide is no longer stable in solution. If the peptide aggregates and becomes insoluble, this can be used to advantage. The peptide can be collected by centrifugation and partially purified with denaturants and organic solvents. Once optimized, the entire process of expression and purification can take as little as 4 to 10 days. However, if the peptide is too hydrophobic, it may become difficult to separate it from MBP, and the process of optimization may take longer. If denaturants are used, one must be concerned about peptide refolding, particularly for peptides that include multiple TM segments. Since most of the membrane proteins under study by these methods contain α-helical TM segments, they are more likely to retain their secondary structure and native fold. As might be expected, peptides containing a single TM are more likely to result in high-level expression, ease of purification, and solubility for structural studies (Fig. 29.8.3). Peptides containing multiple TM segments have a much lower success rate at all stages of the workflow described herein (Table 29.8.1). For a human membrane protein such as the Na^+/H^+ exchanger isoform 1, this may represent the increasing uncertainty in defining the correct limits for two or three sequential TM segments (Fig. 29.8.1). In this case, it may be necessary to design a multi-TM

peptide, test for high-level expression and solubility (e.g., in detergent or 70% isopropanol), and then redesign based on the outcome. This would be repeated until a highly expressed, soluble construct is obtained. The underlying assumption here is that expression and solubility may be useful indicators for a correctly designed multi-TM peptide. Additionally, the further experimental goals will determine the quality of peptide required. Indeed, sarcolipin purified by HPLC following organic extraction does not appear significantly more pure than the peptide after organic extraction alone, but without the HPLC "polishing" step sarcolipin cannot effectively be reconstituted into liposomes for functional experiments. Conversely, a single TM segment from a yeast Na^+/H^+ exchanger was sufficiently pure following organic extraction for NMR experiments.

Anticipated Results

The protocols described in this unit provide a conceptual framework and starting point for the design, expression, and purification of α-helical TM segments from any membrane protein of interest. To illustrate peptide design, we have used a region of the human Na^+/H^+ exchanger that contains three well defined TM segments and residues identified to be important for transport function. In previous work, two TM segments in this region, TM6 and TM7, were chemically synthesized, and their structures were determined by NMR spectroscopy (Table 29.8.1). The methods described herein were used to express and purify the tandem constructs, TM6-7 and TM5-7. The structure of TM6-7 has been determined (L.F., unpub. observ.). As mentioned, MBP is a facile system for expressing these hydrophobic TM peptides because the fusion protein can be handled as a stable, soluble protein. Using the methodologies and suggested conditions mentioned above, the authors have successfully expressed several TM segments of the human Na^+/H^+ exchanger, as well as dozens of mutants of human phospholamban and sarcolipin, single-TM regulatory proteins of the sarcoplasmic reticulum (e.g., see Ceholski et al., 2012). Many of these peptides are very hydrophobic, necessitating extreme purification strategies such as organic extraction and the use of denaturants. Phospholamban and the human Na^+/H^+ exchanger epitomize some of the difficult properties one can encounter in these peptide constructs. Phospholamban has a hydrophobic TM helix and a highly basic cytoplasmic domain (the pI for phospholamban is 9.5), while the limits of TM segments for the Na^+/H^+ exchanger are not well defined, making it challenging to design multi-TM constructs. Nonetheless, such TM constructs can be generated using the basic protocols described above. The anticipated outcome for a "well behaved" TM peptide is as follows. From 4 liters of cell culture, one can obtain 32 g of cell pellet, 1000 mg of fusion protein, and 10 to 20 mg of pure TM peptide. While the final yield depends on the peptide of interest (low of 1 mg and high of 30 mg), the method can provide suitable quantities for a variety of biophysical techniques aimed at characterizing the secondary and tertiary structure of membrane transport proteins piece by piece. Ultimately, this approach can provide valuable insights into the structure and transport mechanisms of membrane proteins that are often too problematic to achieve through expression of full-length proteins.

Time Considerations

Once optimized, the entire process of expression and purification can take 4 to 10 days. As indicated throughout this unit, there are many points where the sample can be stored and the process continued at a later time. While the time considerations vary widely for different constructs, a typical experience is as follows. The harvested bacterial cell pellet can be stored at −20°C for months; the amylose column eluent (before or after concentration) can be stored at 4°C for up to 3 days; the TEV digestion reaction is stable for at least 1 week at 16°C; the organic extraction, if sealed, is stable for months at room temperature; and a dried peptide from HPLC or organic extraction is stable for months under vacuum or at −80°C. The optimization process for expression may take several weeks or months depending on the number of bacterial strains and conditions screened. However, fusion protein purification and protease cleavage should require less optimization. Peptide purification can be the most time-consuming step, often requiring a few months to select and optimize a purification strategy. Unfortunately, it is not possible to purify some peptides (Table 29.8.1), and this behavior is difficult to predict during construct design. Nonetheless, a well behaved construct can be purified to homogeneity and analyzed by multidimensional solution NMR in approximately 2 weeks.

Acknowledgements

The authors would like to acknowledge several people who were instrumental in the optimization of the methods described above: J. L.

Douglas and C.A. Trieber for early design and optimization of the expression and purification system; J. P. Glaves and P. A. Gorski for the optimization of expression in M9 medium and the organic extraction technique; and D. K. Ceholski and P.A. Gorski for the optimization of HPLC preparation techniques.

Literature Cited

Afara, M.R., Trieber, C.A., Glaves, J.P., and Young, H.S. 2006. Rational design of peptide inhibitors of the sarcoplasmic reticulum calcium pump. *Biochemistry* 45:8617-8627.

Afara, M.R., Trieber, C.A., Ceholski, D.K., and Young, H.S. 2008. Peptide inhibitors use two related mechanisms to alter the apparent calcium affinity of the sarcoplasmic reticulum calcium pump. *Biochemistry* 47:9522-9530.

Akabas, M.H., Stauffer, D.A., Xu, M., and Karlin, A. 1992. Acetylcholine receptor channel structure probed in cysteine-substitution mutants. *Science* 258:307-310.

Albert, A.D. and Yeagle, P.L. 2002. Structural studies on rhodopsin. *Biochim. Biophys. Acta* 1565:183-195.

Bertani, G. 2004. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* 186:595-600.

Bhave, G., Nadin, B.M., Brasier, D.J., Glauner, K.S., Shah, R.D., Heinemann, S.F., Karim, F., and Gereau, R.W. 2003. Membrane topology of a metabotropic glutamate receptor. *J. Biol. Chem.* 278:30294-30301.

Bichet, P., Mollat, P., Capdevila, C., and Sarubbi, E. 2000. Endogenous glutathione-binding proteins of insect cell lines: Characterization and removal from glutathione S-transferase (GST) fusion proteins. *Protein Expr. Purif.* 19:197-201.

Bordag, N. and Keller, S. 2010. α-Helical transmembrane peptides: A "Divide and Conquer" approach to membrane proteins. *Chem. Phys. Lipids* 163:1-26.

Ceholski, D.K., Trieber, C.A., Holmes, C.F.B., and Young, H.S. 2012. Lethal, hereditary mutants of phospholamban elude phosphorylation by protein kinase a. *J. Biol. Chem.* 287:26596-26605.

Chopra, A., Yeagle, P.L., Alderfer, J.A., and Albert, A.D. 2000. Solution structure of the sixth transmembrane helix of the G-protein-coupled receptor, rhodopsin. *Biochim. Biophys. Acta* 1463: 1-5.

Cole, C., Barber, J.D., and Barton, G.J. 2008. The Jpred 3 secondary structure prediction server. *Nucl. Acids Res.* 36:W197-W201.

Cregg, J.M., Tolstorukov, I., Kusari, A., Sunga, J., Madden, K., and Chappell, T. 2009. Expression in the yeast *Pichia pastoris*. *Methods Enzymol.* 463:169-189.

Cuff, J.A. and Barton, G.J. 1999. Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins Struct. Funct. Bioinformat.* 34:508-519.

Cunningham, F. and Deber, C.M. 2007. Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods* 41:370-380.

Cybulski, L.E. and De Mendoza, D. 2011. Bilayer hydrophobic thickness and integral membrane protein function. *Curr. Protein Peptide Sci.* 12:760-766.

Daly, R. and Hearn, M.T. W. 2005. Expression of heterologous proteins in *Pichia pastoris*: A useful experimental tool in protein engineering and production. *J. Mol. Recogn.* 18:119-138.

De Bernardez Clark, E. 1998. Refolding of recombinant proteins. *Curr. Opin. Biotechnol.* 9:157.

De Planque, M.R.R. and Killian, J.A. 2003. Protein-lipid interactions studied with designed transmembrane peptides: Role of hydrophobic matching and interfacial anchoring. *Mol. Membrane Biol.* 20:271-284.

Ding, J., Rainey, J.K., Xu, C., Sykes, B.D., and Fliegel, L. 2006. Structural and functional characterization of transmembrane segment VII of the Na+/H+ exchanger isoform 1. *J. Biol. Chem.* 281:29817-29829.

Douglas, J.L., Trieber, C.A., Afara, M., and Young, H.S. 2005. Rapid, high-yield expression and purification of Ca2+-ATPase regulatory proteins for high-resolution structural studies. *Protein Expr. Purif.* 40:118-125.

Duff, K.C. and Ashley, R.H. 1992. The transmembrane domain of influenza A M2 protein forms amantadine-sensitive proton channels in planar lipid bilayers. *Virology* 190:485-489.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.

Hu, J., Qin, H., Li, C., Sharma, M., Cross, T.A., and Gao, F.P. 2007. Structural biology of transmembrane domains: Efficient production and characterization of transmembrane peptides by NMR. *Protein Sci.* 16:2153-2165.

Hunt, J.F., Earnest, T.N., Bousché, O., Kalghatgi, K., Reilly, K., Horváth, C., Rothschild, K.J., and Engelman, D.M. 1997. A biophysical study of integral membrane protein folding. *Biochemistry* 36:15156-15176.

Jidenko, M., Nielsen, R.C., Sørensen, T.L.-M., Møller, J.V., le Maire, M., Nissen, P., and Jaxel, C. 2005. Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 102:11687-11691.

Jones, P.G. and Inouye, M. 1994. The cold-shock response—a hot topic. *Mol. Microbiol.* 11:811-818.

Junge, F., Schneider, B., Reckel, S., Schwarz, D., Dötsch, V., and Bernhard, F. 2008. Large-scale production of functional membrane proteins. *Cell. Mol. Life Sci.* 65:1729-1755.

Kane, J.F. and Hartley, D.L. 1988. Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends Biotechnol.* 6:95-101.

Kapust, R.B. and Waugh, D.S. 1999. Escherichia coli maltose-binding protein is uncommonly

effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8:1668-1674.

Karmazyn, M., Avkiran, M., and Fliegel, L. (eds.). 2003. The Sodium-Hydrogen Exchanger: From Molecule to its Role in Disease. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Katragadda, M., Alderfer, J.L., and Yeagle, P.L. 2000. Solution structure of the loops of bacteriorhodopsin closely resembles the crystal structure. *Biochim. Biophys. Acta* 1466:1-6.

Katragadda, M., Alderfer, J.L., and Yeagle, P.L. 2001. Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin. *Biophys. J.* 81:1029-1036.

Klammt, C., Schwarz, D., Dötsch, V., and Bernhard, F. 2007. Cell-free production of integral membrane proteins on a preparative scale. *In* In Vitro Transcription and Translation Protocols Methods in Molecular Biology. (G. Grandi, ed.) pp. 57-78. Humana Press, Totowa, N.J. Available at: *http://www.springerlink.com/content/l3262163j2827804/abstract/* [Accessed October 2, 2012].

Kocherla, H., Marino, J., Shao, X., Graf, J., Zou, C., and Zerbe, O. 2012. Biosynthesis and spectroscopic characterization of 2-TM fragments encompassing the sequence of a human GPCR, the Y4 receptor. *Chembiochem* 13:818-828.

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. *J. Mol. Biol.* 305:567-580.

Landau, M., Herz, K., Padan, E., and Ben-Tal, N. 2007. Model structure of the Na+/H+ exchanger 1 (NHE1): Functional and clinical implications. *J. Biol. Chem.* 282:37854-37863.

Lee, B.L., Li, X., Liu, Y., Sykes, B.D., and Fliegel, L. 2009a. Structural and functional analysis of extracellular loop 2 of the Na(+)/H(+) exchanger. *Biochim. Biophys. Acta* 1788:2481-2488.

Lee, B.L., Li, X., Liu, Y., Sykes, B.D., and Fliegel, L. 2009b. Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na+/H+ exchanger. *J. Biol. Chem.* 284:11546-11556.

Lee, B.L., Sykes, B.D., and Fliegel, L. 2011. Structural analysis of the Na+/H+ exchanger isoform 1 (NHE1) using the divide and conquer approach. *Biochem. Cell Biol.* 89:189-199.

Liu, L.P. and Deber, C.M. 1998. Guidelines for membrane protein engineering derived from de novo designed model peptides. *Biopolymers* 47:41-62.

Lundstrom, K. 2010. Expression of mammalian membrane proteins in mammalian cells using Semliki Forest Virus vectors. *In* Heterologous Expression of Membrane Proteins Methods in Molecular Biology. (I. Mus-Veteau, ed.) pp. 149-163. Humana Press, Totowa, N.J. Available at: *http://www.springerlink.com/content/x18r02152686653r/abstract/* [Accessed October 2, 2012].

Marblestone, J.G., Edavettal, S.C., Lim, Y., Lim, P., Zuo, X., and Butt, T.R. 2006. Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. *Protein Sci.* 15:182-189.

Melnyk, R.A., Partridge, A.W., Yip, J., Wu, Y., Goto, N.K., and Deber, C.M. 2003. Polar residue tagging of transmembrane peptides. *Peptide Sci.* 71:675-685.

Miroux, B. and Walker, J.E. 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.

Mishima, M., Wakabayashi, S., and Kojima, C. 2007. Solution structure of the cytoplasmic region of Na+/H +exchanger 1 complexed with essential cofactor calcineurin B homologous protein 1. *J. Biol. Chem.* 282:2741-2751.

Moncoq, K., Kemp, G., Li, X., Fliegel, L., and Young, H.S. 2008. Dimeric structure of human Na+/H+ exchanger isoform 1 overproduced in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 283:4145-4154.

Nagai, K., Perutz, M.F., and Poyart, C. 1985. Oxygen binding properties of human mutant hemoglobins synthesized in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82:7252-7255.

Nallamsetty, S. and Waugh, D.S. 2006. Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* 45:175-182.

Newby, Z.E.R., O'Connell, J.D., Gruswitz, F., Hays, F.A., Harries, W.E.C., Harwood, I.M., Ho, J.D., Lee, J.K., Savage, D.F., Miercke, L.J.W., and Stroud, R.M. 2009. A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat. Protoc.* 4:619-637.

Nygaard, E.B., Lagerstedt, J.O., Bjerre, G., Shi, B., Budamagunta, M., Poulsen, K.A., Meinild, S., Rigor, R.R., Voss, J.C., Cala, P.M., and Pedersen, S.F. 2011. Structural modeling and electron paramagnetic resonance spectroscopy of the human Na+/H+ exchanger isoform 1, NHE1. *J. Biol. Chem.* 286:634-648.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M, and Miyano, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289:739-745.

Punta, M., Forrest, L.R., Bigelow, H., Kernytsky, A., Liu, J., and Rost, B. 2007. Membrane protein prediction methods. *Methods* 41:460-474.

Raman, P., Cherezov, V., and Caffrey, M. 2006. The membrane protein data bank. *Cell. Mol. Life Sci.* 63:36-51.

Reddy, G.L., Iwamoto, T., Tomich, J.M., and Montal, M. 1993. Synthetic peptides and four-helix bundle proteins as model systems for the pore-forming structure of channel proteins. II.

Transmembrane segment M2 of the brain glycine receptor is a plausible candidate for the pore-lining structure. *J. Biol. Chem.* 268:14608-14615.

Reddy, L.G., Jones, L.R., Cala, S.E., O'Brian, J.J., Tatulian, S.A., and Stokes, D.L. 1995. Functional reconstitution of recombinant phospholamban with rabbit skeletal Ca2+-ATPase. *J. Biol. Chem.* 270:9390-9397.

Reddy, T., Ding, J., Li, X., Sykes, B.D., Rainey, J.K., and Fliegel, L. 2008. Structural and functional characterization of transmembrane segment IX of the NHE1 isoform of the Na+/H +exchanger. *J. Biol. Chem.* 283:22018-22030.

Roosild, T.P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., and Choe, S. 2005. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* 307:1317-1321.

Rost, B., Sander, C., Casadio, R., and Fariselli, P. 1995. Transmembrane helices predicted at 95% accuracy. *Protein Sci.* 4:521-533.

Routzahn, K.M. and Waugh, D.S. 2002. Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *J. Struct. Funct. Genomics* 2:83-92.

Schägger, H. 2006. Tricine–SDS-PAGE. *Nat. Protoc.* 1:16-22.

Shendure, J.A., Porreca, G.J., Church, G.M., Gardner, A.F., Hendrickson, C.L., Kieleczawa, J., and Slatko, B.E. 2011. Overview of DNA sequencing strategies. *Curr. Protoc. Mol. Biol.* 96:7.1.1–7.1.23.

Simmerman, H.K., Collins, J.H., Theibert, J.L., Wegener, A.D., and Jones, L.R. 1986. Sequence analysis of phospholamban. Identification of phosphorylation sites and two major structural domains. *J. Biol. Chem.* 261:13333-13341.

Singh, S.M. and Panda, A.K. 2005. Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 99:303-310.

Slepkov, E. and Fliegel, L. 2002. Structure and function of the NHE1 isoform of the Na+/H+ exchanger. *Biochem. Cell Biol.* 80:499-508.

Slepkov, E.R., Rainey, J.K., Li, X., Liu, Y., Cheng, F.J., Lindhout, D.A., Sykes, B.D., and Fliegel, L. 2005. Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na+/H+ exchanger. *J. Biol. Chem.* 280:17863-17872.

Tang, X.-B., Fujinaga, J., Kopito, R., and Casey, J.R. 1998. Topology of the region surrounding Glu681 of human AE1 protein, the erythrocyte anion exchanger. *J. Biol. Chem.* 273:22545-22553.

Trometer, C. and Falson, P. 2010. Mammalian membrane protein expression in baculovirus-infected insect cells. *Methods Mol. Biol.* 601:105-117.

Tusnády, G.E. and Simon, I. 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17:849-850.

Tzeng, J., Lee, B.L., Sykes, B.D., and Fliegel, L. 2010. Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na+/H+ exchanger. *J. Biol. Chem.* 285:36656-36665.

Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. 2000. A novel topology model of the human Na(+)/H(+) exchanger isoform 1. *J. Biol. Chem.* 275:7942-7949.

Wallin, E. and Von Heijne, G. 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 7:1029-1038.

Wawrzynow, A., Theibert, J.L., Murphy, C., Jona, I., Martonosi, A., and Collins, J. H. 1992. Sarcolipin, the "proteolipid" of skeletal muscle sarcoplasmic reticulum, is a unique, amphipathic, 31-residue peptide. *Arch. Biochem. Biophys.* 298:620-623.

Yeagle, P.L., Alderfer, J.L., and Albert, A.D. 1995a. Structure of the carboxy-terminal domain of bovine rhodopsin. *Nat. Struct. Biol.* 2:832-834.

Yeagle, P.L., Alderfer, J.L., and Albert, A.D. 1995b. Structure of the third cytoplasmic loop of bovine rhodopsin. *Biochemistry* 34:14621-14625.

Yeagle, P.L., Alderfer, J.L., and Albert, A.D. 1996. Structure determination of the fourth cytoplasmic loop and carboxyl terminal domain of bovine rhodopsin. *Mol. Vision* 2:12.

Yeagle, P.L., Alderfer, J.L., Salloum, A.C., Ali, L., and Albert, A.D. 1997. The first and second cytoplasmic loops of the G-protein receptor, rhodopsin, independently form beta-turns. *Biochemistry* 36:3864-3869.

Yeagle, P.L., Danis, C., Choi, G., Alderfer, J.L., and Albert, A.D. 2000. Three dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin. *Mol. Vision* 6:125-131.

Zuo, X., Li, S., Hall, J., Mattern, M.R., Tran, H., Shoo, J., Tan, R., Weiss, S.R., and Butt, T.R. 2005. Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. *J. Struct. Funct. Genomics* 6:103-111.