Protocol

Journal of PeptideScience

Received: 18 November 2009 Revised: 21 January 2010 Accepted: 1 February 2010 Published online in Wiley Interscience: 1 April 2010

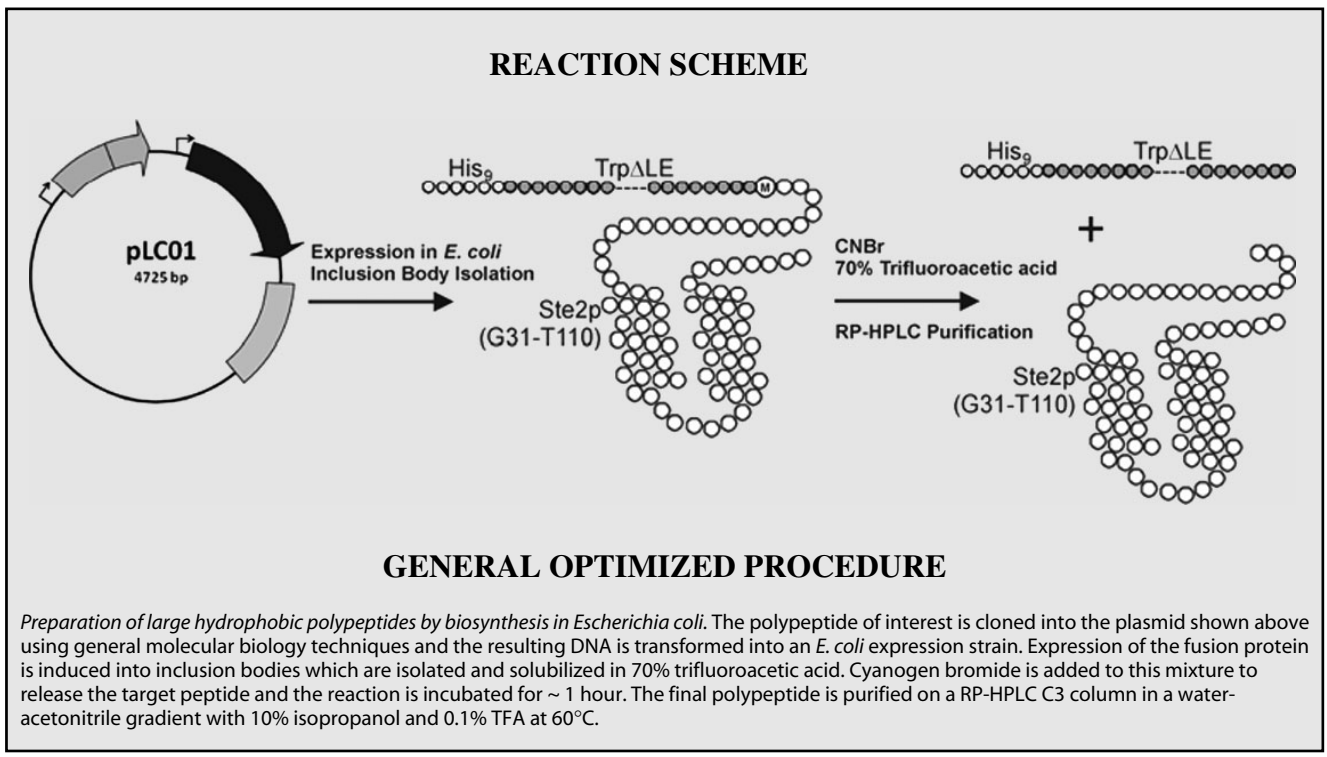
(www.interscience.com) DOI 10.1002/psc.1222

Biosynthesis of peptide fragments of eukaryotic GPCRs in *Escherichia coli* by directing expression into inclusion bodies

Leah S. Cohen,[a,b] Jeffrey M. Becker[c] and Fred Naider[a,b]*

Biosynthesis of peptides in heterologous systems is often a prerequisite to biophysical analyses. Large amounts of peptides, in particular portions of membrane proteins, are needed to optimize conditions for secondary and tertiary structure analysis. Chemical synthesis of these peptides is limited by their high hydrophobicity and also due to the need to incorporate isotopic labels for high resolution NMR analysis. In this protocol, we describe a method for the heterologous expression and purification of unlabeled and isotopically labeled peptide fragments of Ste2p, an integral membrane G protein-coupled receptor. Copyright © 2010 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: biosynthesis; G protein-coupled receptor; GPCR; peptide fragments; isotopic labeling; membrane peptides



Preparation of large hydrophobic polypeptides by biosynthesis in Escherichia coli. The polypeptide of interest is cloned into the plasmid shown above using general molecular biology techniques and the resulting DNA is transformed into an *E. coli* expression strain. Expression of the fusion protein is induced into inclusion bodies which are isolated and solubilized in 70% trifluoroacetic acid. Cyanogen bromide is added to this mixture to release the target peptide and the reaction is incubated for ~ 1 hour. The final polypeptide is purified on a RP-HPLC C3 column in a water-acetonitrile gradient with 10% isopropanol and 0.1% TFA at 60°C.

Scope and Comments

Heterologous expression of integral membrane peptide fragments is often hampered by their toxicity to the host cell. When large amounts of these polypeptides are generated, there can be misfolding and/or mis-incorporation into the membrane resulting in early cell death and, therefore, low protein yield. Large-scale expression of integral membrane peptide fragments is extremely important for their structural studies by X-ray crystallography and NMR. Both of these structural methods are hampered by the hydro-

* *Correspondence to: Fred Naider, College of Staten Island, CUNY, 2800 Victory Blvd, Staten Island, NY 10314, USA. E-mail: fred.naider@csi.cuny.edu*

a *Department of Chemistry, College of Staten Island, CUNY, 2800 Victory Blvd, Staten Island, NY 10314, USA*

b *Department of Biochemistry, Graduate Center of the City University of New York, 365 Fifth Ave New York, NY 10016, USA*

c *Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA*

phobicity or the large size of the polypeptide : detergent complex. NMR analyses are greatly facilitated by isotopic labeling to incorporate ^{15}N, ^{13}C, and/or ^{2}H nuclei. Such labeling has been optimized in *E. coli* and the use of this organism is also beneficial because it does not perform some of the posttranslational modifications (i.e. glycosylation) that may result in a heterogeneous peptide preparation that is difficult to crystallize. One method that improves expression yields utilizes *N*-terminal fusion proteins to drive the expressed peptide into insoluble inclusion bodies (IBs) thereby eliminating the toxicity of the desired product peptide. These fusion proteins can later be removed from the chimeric protein to generate the target peptide. It is necessary to optimize the growth conditions as to the *E. coli* strain, medium, temperature, fusion protein, and inducer for each target peptide.

The use of the TrpΔLE fusion protein was first described by Kim and Staley in the expression of bovine pancreatic trypsin inhibitor [1]. The fusion protein is composed of the *E. coli* trp leader polypeptide L and the *E. coli* trp polypeptide E with the deletion LE 1413 [2,3] and it also contains the T7 promoter–operator sequence which allows for high levels of transcription into mRNA and a ribosome binding site (Shine–Dalgarno) sequence which allows the translation machinery to recognize the mRNA and translate it into protein [4]. We have reported the use of the TrpΔLE fusion protein to aid in the expression of multiple peptide fragments of Ste2p [5–7]. Other groups have also used this leader peptide to express an HIV-1 gene *vpu* [8,9], multiple transmembrane (TM) domain constructs of the cannabinoid CB2 GPCR [10–12], isotopically labeled charged multivesicular body protein peptide from residues 166 to 181 [13], and TM domains of caveolin [14] and of the p7 protein from the hepatitis C virus [15]. We have been able to express peptide fragments of Ste2p of one [5,6], two [7] and three (data not published) TMs. Our protocol generated sufficient peptide to produce high resolution structures in an organic aqueous solvent [6] and lyso-1-palmitoylphosphatidylglycerol micelles [16] of two of these constructs. Furthermore, the expressed 3TM fragment of Ste2p resulted in 15–20 mg per L of culture of [^{15}N]-labeled Ste2p(G31-R161) (data not published). In this protocol, the cloning, expression, and purification of an 80 amino acid peptide Ste2p(G31-T110) fragment that contains part of the *N*-terminus, TM1, intracellular loop 1, TM2, and part of extracellular loop 1 will be presented. The protocol resulted in 10–20 mg per L of culture of the purified peptide.

Experimental Procedure

Cloning of membrane protein fragments

PCR was used to generate a DNA fragment encoding the Ste2p(G31-T110) sequence. Restriction enzyme sites were incorporated into the forward and reverse primers, respectively (5′-AGTACGCGCTAAGCTTTGGATGGGGAATGGATCTACCATCAC-3′ and 5′-CTCGGTACCCGGGGATCCTCAAGTCACTGAAGAGTAATTAG-3′) to facilitate cloning into the pMMHa plasmid [1]. The PCR product and plasmid were digested with *Hind*III and *Bam*HI, gel purified and ligated together using T4 DNA ligase. Following ligation, the product was transformed into DH5α cells and screened for the correct final construct by restriction enzyme digestion analysis of the amplified plasmid. Mutagenesis of the native methionine residues to other nonpolar amino acids (Leu, Val, and Ile) was carried out to facilitate CNBr cleavage to remove the fusion protein. This was performed by double-stranded mutagenesis based on a modified Quickchange mutagenesis protocol (Stratagene, LaJolla, CA, USA) [17]. Sequencing was performed to confirm the final DNA product which encodes Ste2p(G31-T110, M54L, C59S, M69V, M71I).

Expression of membrane protein fragments in fusion proteins

The resulting plasmid, pLC01 (Figure 1), was transformed into three different *E. coli* expression strains: BL21(DE3)pLysS, BL21-AI, and BL21Star(DE3)pLysS. Each of these strains was developed to express heterologous proteins that tend to be toxic to the cells. Different strains often result in optimized expression of different proteins. Initially, induction conditions were determined based on the manufacturer's recommendation. Generally, an overnight culture in Luria–Bertani (LB) broth was grown at 37 °C with shaking at 250 rpm. All cultures contained 200 μg/ml of ampicillin and the cultures for DE3 strains also contained 34 μg/ml of chloramphenicol. A 1 : 20 dilution of the overnight growth into fresh LB was made. As a general rule, the culture volume relative to the flask size was kept at 1 : 5 to allow for good aeration of the cells. The cells were grown until OD_{600} is 0.6 for the BL21(DE3)pLysS and BL21Star(DE3)pLysS strains and OD_{600} is 0.4 for the BL21-AI strain. When these ODs were reached, the expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) or 0.5% L-arabinose for the BL21(DE3)pLysS and BL21Star(DE3)pLysS strains, and the BL21-AI strain, respectively. Portions (1 ml) of each induction were taken, pelleted, and used to follow the expression over time. Small-scale growths (50-ml culture volume) were performed to choose an expression strain and then to optimize expression conditions as to temperature, time, and inducer concentration (Figure 2A). The optimal expression for the isotopically labeled ^{15}N and/or ^{13}C fusion protein TrpΔLE-Ste2p(G31-T110) was found in BL21-AI cells grown at 37 °C to OD_{600} 0.4 followed by induction with 0.5% L-arabinose, 1 mM IPTG, and expression at 22 °C for 22 h in minimal medium (M9: 1 g/l $^{15}NH_4Cl$, 20 mM KH_2PO_4, 48 mM Na_2HPO_4, 8.6 mM NaCl, 0.4% glucose or ^{13}C-glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 200 μg/ml ampicillin; Figure 2B). At the end of large-scale (1 l) growths, the cells were pelleted and frozen for further use. Yields of unlabeled, [^{15}N]- and [^{15}N, ^{13}C]-labeled Ste2p(G31-T110) were 8–10 mg per L of culture of the purified peptide fragments. Higher yields (up to 20 mg per L of culture) were observed when selective amino acid labeling was performed in minimal medium supplemented with 19 unlabeled amino acids. The desired [^{15}N]-amino acid was added just prior to induction.

Growth in deuterated media

Deuteration of peptide fragments for NMR analysis can be useful due to changes in relaxation parameters and also a reduction in the number of peaks visualized. Growth is slowed in deuterated media and therefore the expression of the fusion protein is reduced. Adaptation of the *E. coli* cells to the deuterated media allows for better protein yields. A 5-ml culture was prepared in 40% D_2O minimal medium and grown overnight at 37 °C. This culture was diluted into 50 ml of culture in 60% D_2O minimal medium grown overnight at 37 °C, and then inoculated into 1 l of minimal medium containing 1 g/l $^{15}NH_4Cl$ and/or 4 g/l ^{13}C-glucose and D_2O. This culture was then grown at 37 °C until OD_{600} reached approximately 0.4 and then the culture was induced with 0.5% L-arabinose and 1 mM IPTG and grown at 22 °C for 22 h. The level

Journal of PeptideScience

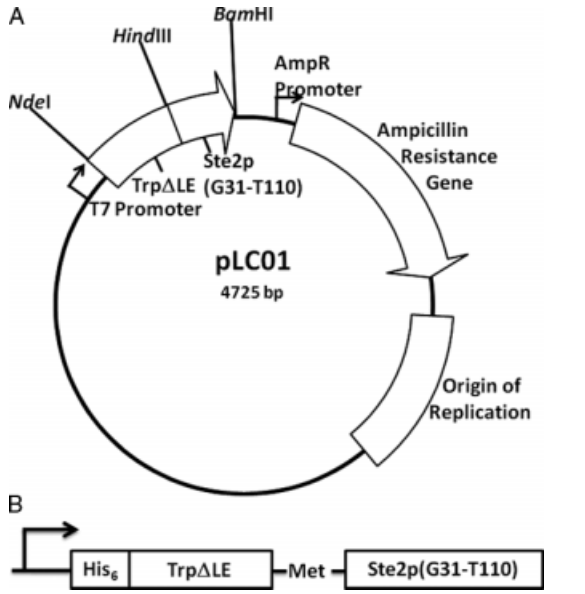


Figure 1. Plasmid that expresses the TrpΔLE–Ste2p(G31-T110) fusion protein. The plasmid pLC01 was based on pMMHa [1]. (A) The origin of replication allows for multiple copies to be made in an *E. coli* strain. The ampicillin resistance gene allows *E. coli* cells containing the plasmid to be selected in growth media. Promoter regions direct the expression of the genes using either *E. coli* RNA polymerase (AmpR promoter) or T7 RNA polymerase (T7 promoter). The T7 RNA polymerase has a higher rate of processivity, so more mRNA and, therefore, more protein can be made. The cloning sites that were used to create this plasmid are indicated above. (B) The fusion protein being expressed is blown up in a cartoon format indicating the Met residue where CNBr cleavage takes place. This cleavage releases the target receptor fragment.

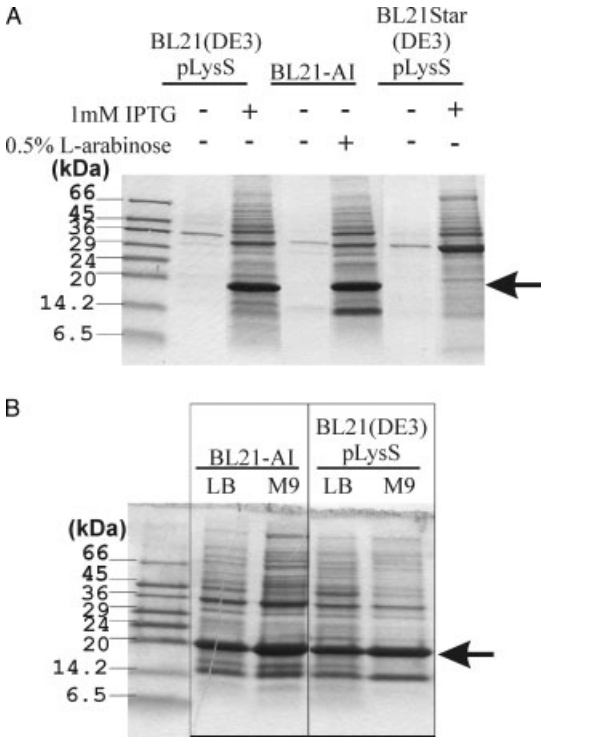


Figure 2. Analysis of TrpΔLE–Ste2p(G31-T110) expression. IBs were generated as described, and the solubilized IBs were separated on a 16% SDS-PAGE and stained by Coomassie Blue. The desired fusion protein [TrpΔLE–Ste2p(G31-T110)] (~22.3 kDa) is indicated by the arrow. (A) Expression analysis in three different *E. coli* strains grown in LB following the manufacturer's instructions. (B) Comparative expression of TrpΔLE–Ste2p(G31-T110) in LB and M9 minimal medium (used for isotopic labeling) in two different *E. coli* strains.

of D_2O in the final growth medium is dependent on the level of deuteration required and can be adjusted from 60 to 100% as required. Growth in deuterated medium resulted in 5–8 mg per L of culture of the purified peptide with the yield reduced at the higher percentages of D_2O.

Growth in media to selectively label methyl groups of aliphatic residues

Selective labeling of the aliphatic methyl groups (i.e. Val, Leu, Ile) can be very important in the determination of long-range connectivities in integral membrane peptide fragments. The following labeling protocol was adapted from Tugarinov and Kay [18]. Cells were streaked onto a LB agar plate containing ampicillin and incubated overnight at 37 °C. A 6-ml LBAmp culture was inoculated from a single colony and grown at 37 °C until OD_{600} ~ 0.75 and then the cells were pelleted gently by centrifugation at 1400*g* and resuspended in 20 ml M9 medium (as described above) to OD_{600} ~ 0.05–0.1. These cells were incubated with shaking at 37 °C until OD_{600} ~ 0.6 and then they were gently pelleted in sterile tubes and resuspended in 100 ml M9 medium in 100% D_2O with 4 g/l $^{13}C/^{2}H$-glucose and 1 g/l $^{15}NH_4Cl$. Incubation at 37 °C with shaking occurred until OD_{600} ~ 0.4–0.5. The cells were then diluted to 200 ml in the same medium and grown again to OD_{600} ~ 0.4–0.5 at which point the cells were diluted to 1 l in labeling medium [100% D_2O, 4 g/l $^{13}C/^{2}H$-glucose, 1 g/l $^{15}NH_4Cl$, 70 mg/l α-ketobutyric acid ($^{13}C_4$, 98%; 3,3-D_2, 98%) and 120 mg/l α-ketoisovaleric acid (1,2,3,4-$^{13}C_4$, 99%; 3,4′,4′,4′-D_4, 98%)]. The precursor α-ketobutyric acid is converted in the cells into isoleucine that is protonated only at the δ-methyl group, whereas the α-ketobutyric acid is converted into both leucine and valine in which one of the branch point methyl groups is protonated. These cells were grown to OD_{600} ~ 0.3–0.4 and induced with 0.5% L-arabinose and grown at 37 °C for 6–8 h and then pelleted by centrifugation at 4000*g* for IB preparation. Growth in methyl group labeling media resulted in about 6 mg per L of culture of the purified peptide.

IB preparation and analysis

As the TrpΔLE fusion protein directs the expression of the protein into IBs to reduce the toxicity to the *E. coli*, IB preparations were generated as described in [1,5,6]. Briefly, frozen cells were thawed slightly, resuspended in lysis mix which is lysis buffer (50 mM Tris–HCl, pH 8.7, 1 mM ethylenediaminetetracetic acid) with 300 μg/ml lysozyme and 1 mM phenylmethylsulfonyl fluoride and sonicated in a room temperature water bath until the cells appeared as a creamy uniform mixture. The mixture was then centrifuged at 39 000*g* for 20 min at 4 °C. The resulting supernatant was decanted and the pellet was resuspended by sonication in lysis buffer and pelleted by centrifugation as above, and then the supernatant was decanted. Next, the pellet was resuspended by sonication in mild detergent (1% igepal Ca-630 and 1% deoxycholic acid in lysis buffer) and pelleted by centrifugation. The pellet was resuspended by sonication in water and pelleted to remove any remaining detergent and the weight was determined. Finally, the IB pellet was resuspended in 8 ml of water per liter of original culture. This suspension was divided into eight-1 ml aliquots in nonstick

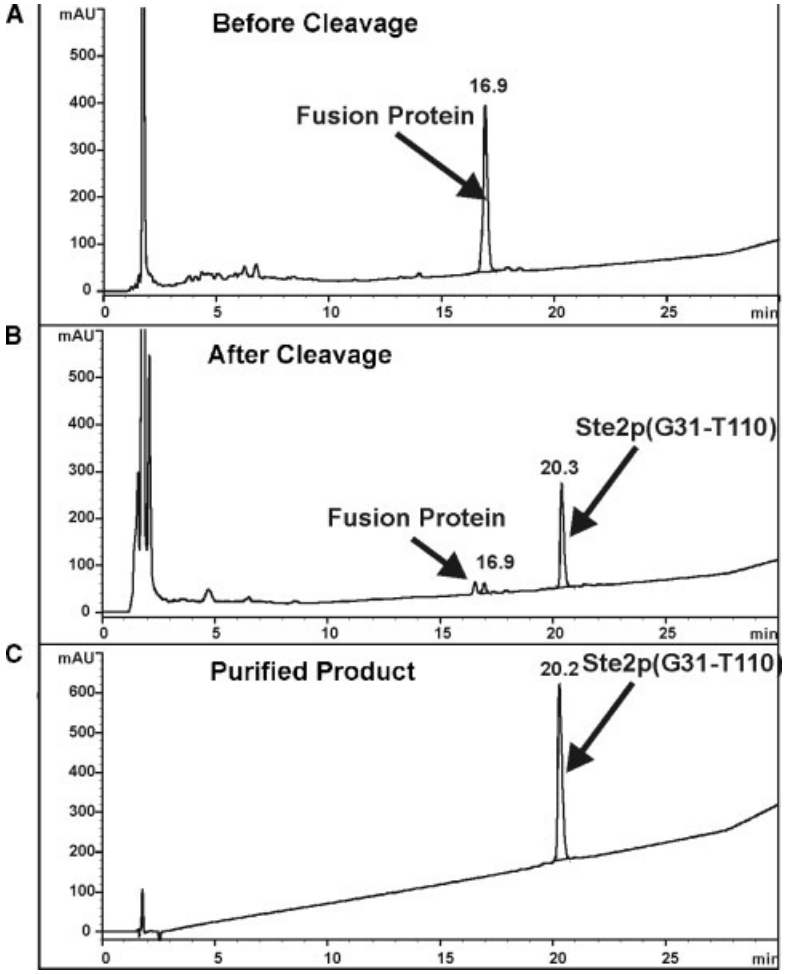


Figure 3. RP-HPLC analysis of TrpΔLE–Ste2p(G31-T110) CNBr cleavage. (A) IB pellet solubilized in 70% TFA and injected onto Zorbax 300SB-C3 analytical column. The fusion protein was determined to be eluted for 16.9 min by ESI-MS analysis. (B) Excess CNBr was added to the TFA solubilized IB pellet and the reaction mixture was incubated in the dark for 1 h. A small sample was injected onto the analytical column to determine if cleavage was complete. (C) The remainder of the reaction mixture was loaded onto a Zorbax 300SB-C3 Prep-HT column and the purified protein was visualized on the analytical column and observed to be >95% pure.

1.5 ml microcentrifuge tubes. The samples were pelleted at 4 °C using 21 100*g* for 1 h. For optimal protein yield, the supernatant at this last pelleting step should be clear. If it is a little cloudy then there will be fusion protein losses and longer centrifugation times are necessary. The 1-ml cell pellets that were removed during cell growth to observe protein expression were prepared similarly but only resuspended in 200 µl lysis mix, 200 µl mild detergent and finally in 100 µl SDS-PAGE loading buffer (4% SDS, 12% glycerol, 256 mM 2-mercaptoethanol, 50 mM Tris–HCl, pH 6.8, and a small amount of Coomassie Brilliant Blue R) and analyzed by SDS-PAGE.

CNBr cleavage and protein fragment purification

The aliquoted IBs were solubilized in 600 µl of 70% TFA by addition of approximately 420 µl 100% TFA followed by sonication and then the addition of approximately 180 µl water. To ensure complete cleavage of the fusion protein, excess amounts of CNBr was added to approximately 2 M final concentration and the reaction mixture was incubated at room temperature in the dark for 30 min. Cleavage was checked by analytical HPLC and when the fusion protein almost completely disappeared (compare the 16.9 min peaks in Figure 2A and B) the reaction mixture was injected onto a Zorbax 300SB-C3 column and purified using a 45–80% acetonitrile : water, 10% isopropanol, 0.1% TFA gradient (Figure 3). The desired peak, as verified by ESI-MS (Figure 4) was collected and lyophilized for use in further biophysical analyses.

Limitations

While this protocol has been found to be a useful general method to biosynthesize and purify isotopically labeled integral membrane peptide fragments, there are some peptides that have proven difficult to obtain. The same fusion protein expressed above but mutated to contain 1 or 2 cysteine residues cannot be purified as described due to difficulties that arise during purification. After the cleavage is performed on either IB preps or purified fusion protein, the desired protein fragment was not observed in the HPLC chromatogram, although the CNBr cleavage seems to occur based on the disappearance of the fusion protein peak. In the past, we have been able to cleave and purify an *N*-terminal Cys 73-residue peptide using conditions similar to those described earlier but the yield was low [19]. The expression, CNBr cleavage to remove the TrpΔLE carrier, and purification of α-defensins containing six

Journal of PeptideScience

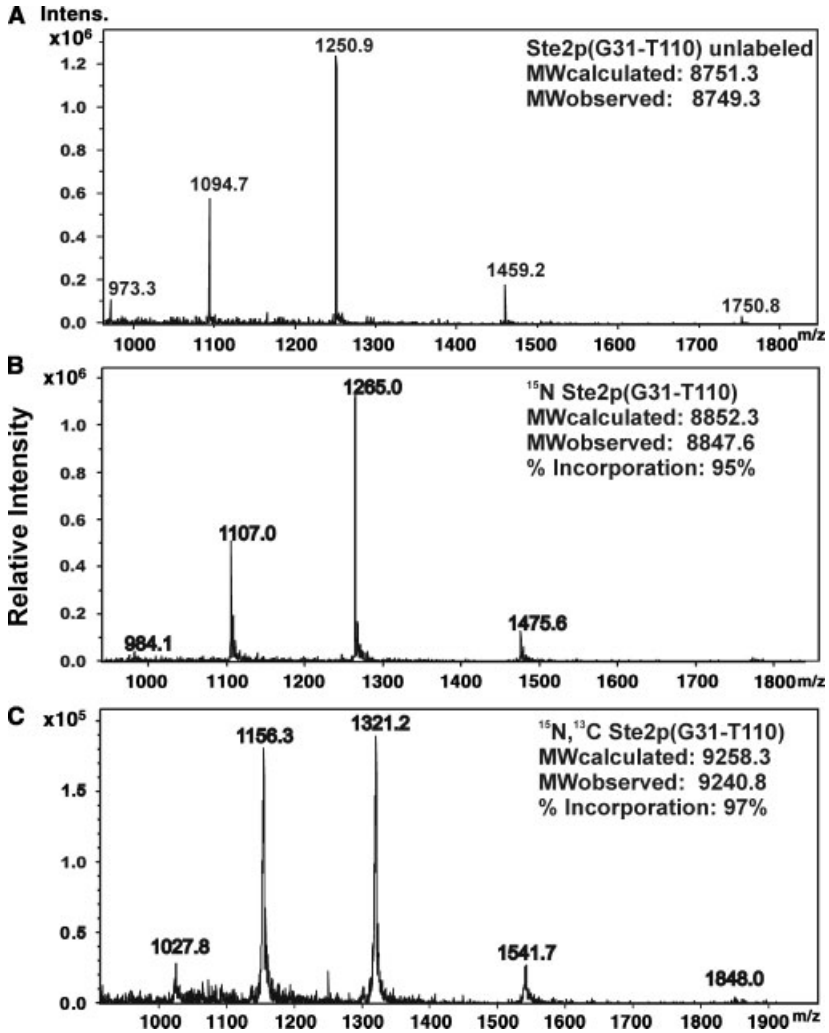


Figure 4. ESI-MS analysis of the purified protein fragments. (A) Unlabeled Ste2p(G31-T110) was determined to be the correct protein fragment based on molecular weight (MW). (B) The ^{15}N moiety was determined to be incorporated into 95% of the peptide nitrogens. The percent incorporation was determined by the following formula:

$$\% \text{ incorporation} = \frac{MW_{labeled/observed} - MW_{unlabeled}}{MW_{labeled/expected} - MW_{unlabeled}} \times 100 \quad (1)$$

(C) The ^{13}C and ^{15}N moieties were determined to be incorporated into 97% of the total carbons and nitrogens of the protein fragment.

Cys residues at 1–3 mg per L of culture has also been reported [20]. It might be possible to eliminate some of these difficulties by alkylating the cysteine residues prior to CNBr cleavage but this would result in a modified peptide. In addition to our experiences with Cys-containing peptides, the harsh conditions of the CNBr cleavage (70% TFA), and of the purification (acetonitrile gradients in the presence of isopropanol at 60 °C) probably cause the denaturation which would require a refolding protocol to achieve the native structure of the peptide fragment. The use of enzymatic cleavage (i.e. tobacco etch virus (TEV) protease) to remove the fusion tag while the peptide is in a more membrane-like micellar environment may help to maintain the native conformation of the peptide.

Acknowledgements

Fred Naider is the Leonard and Esther Kurtz Term Professor at the College of Staten Island. This work was supported by the National Institute of General Medical Sciences grants GM22086 and GM22087.

References

1 Staley JP, Kim PS. Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci.* 1994; **3**: 1822–1832.
2 Miozzari GF, Yanofsky C. Translation of the leader region of the *Escherichia coli* tryptophan operon. *J. Bacteriol.* 1978; **133**: 1457–1466.
3 Kleid DG, Yansura D, Small B, Dowbenko D, Moore DM, Grubman MJ, McKercher PD, Morgan DO, Robertson BH, Bachrach HL. Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine. *Science* 1981; **214**: 1125–1129.
4 Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 1990; **185**: 60–89.
5 Arevalo E, Estephan R, Madeo J, Arshava B, Dumont M, Becker JM, Naider F. Biosynthesis and biophysical analysis of domains of a yeast G protein-coupled receptor. *Biopolymers* 2003; **71**: 516–531.

6 Estephan R, Englander J, Arshava B, Samples KL, Becker JM, Naider F. Biosynthesis and NMR analysis of a 73-residue domain of a Saccharomyces cerevisiae G protein-coupled receptor. *Biochemistry* 2005; **44**: 11795–11810.
7 Cohen LS, Arshava B, Estephan R, Englander J, Kim H, Hauser M, Zerbe O, Ceruso M, Becker JM, Naider F. Expression and biophysical analysis of two double-transmembrane domain-containing fragments from a yeast G protein-coupled receptor. *Biopolymers* 2008; **90**: 117–130.
8 Marassi FM, Ma C, Gratkowski H, Straus SK, Strebel K, Oblatt-Montal M, Montal M, Opella SJ. Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc. Natl. Acad. Sci. U S A* 1999; **96**: 14336–14341.
9 Kochendoerfer GG, Jones DH, Lee S, Oblatt-Montal M, Opella SJ, Montal M. Functional characterization and NMR spectroscopy on full-length Vpu from HIV-1 prepared by total chemical synthesis. *J. Am. Chem. Soc.* 2004; **126**: 2439–2446.
10 Zheng H, Zhao J, Wang S, Lin CM, Chen T, Jones DH, Ma C, Opella S, Xie XQ. Biosynthesis and purification of a hydrophobic peptide from transmembrane domains of G-protein-coupled CB2 receptor. *J. Pept. Res.* 2005; **65**: 450–458.
11 Zhao J, Zheng H, Xie XQ. NMR characterization of recombinant transmembrane protein CB2 fragment CB2(180–233). *Protein Pept. Lett.* 2006; **13**: 335–342.
12 Zheng H, Zhao J, Sheng W, Xie XQ. A transmembrane helix-bundle from G-protein coupled receptor CB2: biosynthesis, purification, and NMR characterization. *Biopolymers* 2006; **83**: 46–61.
13 Kieffer C, Skalicky JJ, Morita E, De Domenico I, Ward DM, Kaplan J, Sundquist WI. Two distinct modes of ESCRT-III recognition are required for VPS4 functions in lysosomal protein targeting and HIV-1 budding. *Dev. Cell* 2008; **15**: 62–73.
14 Diefenderfer C, Lee J, Mlyanarski S, Guo Y, Glover KJ. Reliable expression and purification of highly insoluble transmembrane domains. *Anal. Biochem.* 2009; **384**: 274–278.
15 Cook GA, Opella SJ. NMR studies of p7 protein from hepatitis C virus. *Eur. Biophys. J.* 2009.
16 Neumoin A, Cohen LS, Arshava B, Tantry S, Becker JM, Zerbe O, Naider F. Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles. *Biophys. J.* 2009; **96**: 3187–3196.
17 Kunkel TA. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U S A* 1985; **82**: 488–492.
18 Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc.* 2006; **1**: 749–754.
19 Balambika R, Inui T, Sargsyan H, Arshava B, Cohen LS, Ding FX, Becker JM, Naider F. Synthesis of a Double Transmembrane Domain Fragment of Ste2p by Native Chemical Ligation. *Int. J. Pept. Res. Ther.* 2007; **13**: 251–263.
20 Pazgier M, Lubkowski J. Expression and purification of recombinant human alpha-defensins in *Escherichia coli*. *Protein Expr Purif* 2006; **49**: 1–8.

