

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Isotopically labeled expression in *E. coli*, purification, and refolding of the full ectodomain of the influenza virus membrane fusion protein

Jaime Curtis-Fisk, Ryan M. Spencer, David P. Weliky *

Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:

Received 27 April 2008
and in revised form 17 June 2008
Available online 28 June 2008

Keywords:
Membrane proteins
Viral proteins
Circular dichroism
NMR
Solid-state NMR
Hemagglutinin
Protein refolding
Bacterial expression
Isotopic labeling
Influenza
Protein expression

ABSTRACT

This paper describes methods to produce an isotopically labeled 23 kDa viral membrane protein with purified yield of 20 mg/L of *Escherichia coli* shake flask culture. This yield is sufficient for NMR structural studies and the protein production methods are simple, straightforward, and rapid and likely applicable to other recombinant membrane proteins expressed in *E. coli*. The target FHA2 protein is the full ectodomain construct of the influenza virus hemagglutinin protein which catalyzes fusion between the viral and the cellular endosomal membranes during infection. The high yield of FHA2 was achieved by: (1) initial growth in rich medium to $A_{600} \sim 8$ followed by a switch to minimal medium and induction of protein expression; and (2) obtaining protein both from purification of the detergent-soluble lysate and from solubilization, purification, and refolding of inclusion bodies. The high cell density was achieved after optimization of pH, oxygenation, and carbon source and concentration, and the refolding protocol was optimized using circular dichroism spectroscopy. For a single residue of membrane-associated FHA2 that was obtained from purification and refolding of inclusion bodies, native conformation was verified by the ^{13}CO chemical shifts measured using solid-state nuclear magnetic resonance spectroscopy.

© 2008 Elsevier Inc. All rights reserved.

While there have been a number of high-resolution structures of bacterial membrane proteins in recent years, there have been fewer structures of viral and eukaryotic membrane proteins in part because of difficulties with production of milligram quantities of pure and folded protein by heterologous expression in *Escherichia coli* [1]. This paper describes methods to address this problem with a particular focus on production of isotopically labeled membrane protein in *E. coli* for nuclear magnetic resonance (NMR)¹ studies. Isotopic labeling of proteins is usually done in a minimal medium with consequent reduction in cell growth relative to rich medium and lower production of protein. For membrane proteins, the yield may be reduced further because of higher hydrophobicity and limited cell membrane space needed to maintain native structure.

The protein production in this study was accomplished with conventional shake flask fermentation rather than a commercial fermenter. Although cell density can often be increased in a fermenter by controlling growth parameters such as carbon concen-

tration, pH, and dissolved oxygen, the fermenter is expensive and not available in all laboratories. This paper describes the alternate but related approach of controlling growth parameters in shake flask fermentation. The culture was grown to maximum cell density in rich medium and then switched into fresh minimal medium prior to induction of protein expression [2]. Labeled amino acids were added to the minimal medium at induction and were incorporated into the expressed protein [3].

As stated earlier, one difficulty with membrane protein expression is the limited cell membrane space. Excess membrane protein is often stored as insoluble aggregates which are termed inclusion bodies. The amount of protein that can be recombinantly produced as inclusion bodies is up to 25% of the total cell mass, making inclusion bodies an attractive target for protein production [4]. However, the inclusion body protein has to be solubilized, purified, and properly folded in membranes or detergent. A typical protocol includes concomitant denaturation and solubilization using a solution containing a denaturant such as urea, purification, and then refolding with solutions containing detergent and possibly lipid vesicles [5–7]. In this paper, refolding was achieved by first mixing the solution of purified denatured fusion protein with a solution containing high concentration detergent and then dialyzing to reduce the denaturant concentration to a negligible level.

The target “FHA2” influenza fusion protein in this study is prototypical for membrane enveloped viruses which initiate cellular

* Corresponding author. Fax: +1 517 353 1793.

E-mail address: weliky@chemistry.msu.edu (D.P. Weliky).

¹ Abbreviations used: NMR, nuclear magnetic resonance; LB, Luria–Bertani; bTOG, *n*-octyl-β-D-thioglucopyranoside; DTPC, di-O-tetradecylphosphatidylcholine; DTPG, di-O-tetradecylphosphatidylglycerol; IPTG, isopropyl thiogalactoside; MAS, magic angle spinning; REDOR, rotational-echo double-resonance; CD, circular dichroism.

infection by fusion between the viral and host cell membranes. The influenza virus is first endocytosed into the cell and fusion occurs between the viral and endosomal membranes at the endosomal pH of ~5. The HA2 protein in the viral membrane catalyzes fusion and brings the two membranes together with consequent lipid mixing and formation of a fusion pore through which the viral nucleocapsid enters the cytoplasm [8]. HA2 has a 185 residue N-terminal “FHA2” ectodomain which lies outside the virus, a 25-residue transmembrane region, and a short 10 residue C-terminal endodomain [9]. The N-terminus of HA2 and FHA2 is the ~20 residue apolar “fusion peptide” region which is in the protein interior prior to endocytosis. After the pH in the endosome is reduced to ~5, HA2 undergoes a large conformational change which exposes the fusion peptide. The fusion peptide then binds to the endosomal membrane and fusion ensues. High-resolution structures exist for the pre-fusion ectodomain complex crystallized from aqueous solution at pH 8 and for a “soluble ectodomain” of HA2 (residues 34–178) crystallized from aqueous solution at pH 4.4 [9,10]. In addition, there are liquid-state NMR structures of the fusion peptide in detergent micelles and electron spin resonance measurements of membrane location of the fusion peptide and of an HA2 ectodomain fragment representing residues 1–127 [11,12]. There have also been solid-state NMR measurements of local conformation at three residues in membrane-associated FHA2 whose sequence is displayed in Fig. 1 [13]. Functional relevance of FHA2 was demonstrated by induction of rapid vesicle fusion at pH 5.0 and much less fusion at pH 7.4 which correlated with the low pH of influenza virus/endosome fusion.

In the earlier solid-state NMR study, the purified yield of the 23 kDa FHA2 protein expressed in *E. coli* was 3 mg/L culture. The present paper describes production of 20 mg/L quantities of isotopically labeled FHA2 which is a very good yield for a non-bacterial membrane protein expressed in *E. coli* [14,15].

Materials and methods

Materials

The FHA2 plasmid was obtained from Dr. Yeon-Kyun Shin at Iowa State University and contained the Lac promoter and kanamycin resistance. The plasmid was transformed into *E. coli* BL21(DE3) cells. Unless noted, all chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Luria-Bertani (LB) medium was purchased from Acumedia (Lansing, MI). The detergents *n*-octyl- β -D-thioglucopyranoside (bTOG) and octyl pentaethylene glycol ether (C8E5) were purchased from Anatrace (Maumee, OH). The ether linked lipids di-O-tetradecylphosphatidylcholine (DTPC) and di-O-tetradecylphosphatidylglycerol (DTPG) were obtained from Avanti Polar Lipids (Alabaster, AL). Leucine with 1-¹³C, ¹⁵N labeling was purchased from Cambridge Isotope Labs (Andover, MA).

Culture growth

All cell cultures were grown in media containing 15 mg/L kanamycin. Bacterial growth was initiated from a glycerol stock of

GLFGAIAGFIENGWEGMIDGWYGFRHQNSEGT
GQAADLKSTQAAIDQINGKLNVRIEKTNEKFHQI
EKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLV
ALENQHTIDLTDSEMNLFEKTRQLRENAEEM
GNGSFKIYHKCDNACIESIRNGTYDHDVYRDEA
LNNRFQIKGVELKSGYKDWEHHHHH

Fig. 1. Protein sequence of the FHA2 construct from the X31 strain of the influenza virus. The last eight residues are non-native. Residue Leu-98 was probed by solid-state NMR spectroscopy and is highlighted.

the recombinant bacterial cells to 1 L of “enriched LB” which contained LB supplemented with 10 mL glycerol. The cell suspension was grown overnight to maximum cell density in a 2.8 L baffled Fernbach flask with a foam closure, shaking at 37 °C and 140 rpm. The cell suspension was then centrifuged at 10,000g for 10 min to produce a solid cell pellet. The pellet was resuspended into 1 L of minimal medium whose optimal composition included the commercial M9 minimal medium salts (6.8 g/L Na₂HPO₄, 3.0 g/L NaH₂PO₄, 0.50 g/L NaCl, 1.0 g/L NH₄Cl), 2.5 g/L MgSO₄, and 10 g/L glycerol at pH 8.0. Cell growth was continued by shaking at 37 °C at 140 rpm.

Many of the experiments in this paper examine the effect of a parameter such as glucose concentration on cell growth. Each growth optimization experiment was done using a 100 mL culture volume in a 250 mL baffled Erlenmeyer flask and all of the growths as a function of a single parameter were performed at the same time in the same shaker. All of the comparative experiments presented in this paper were repeated and for each parameter, the trends in cell growth were reproducible. The value of the parameter which yielded highest growth was also reproducible. Examination of all of the A₆₀₀ or ΔA₆₀₀ values between the different comparative growths showed that there was systematic variation by as much as 20%, e.g. for each parameter value, (A₆₀₀)_{first growth}/(A₆₀₀)_{second growth} would be ~1.2. It was therefore difficult to estimate error bars which would reflect data uncertainties within a single comparative growth.

Isotopically labeled FHA2 production

After 1 h of resuspended cell growth in optimized minimal medium, FHA2 expression was induced by addition of isopropyl thiogalactoside (IPTG) to a final concentration of 0.2 mM. For production of FHA2 with 1-¹³C, ¹⁵N Leu isotopic labeling, 100 mg/L of labeled amino acid was added at the time of protein induction. FHA2 production was continued for 3 h at 37 °C. The cell pellet was harvested by centrifugation at 10,000g for 10 min, and the pellet was then stored at -80 °C until purification.

Purification of native FHA2 from the soluble cell lysate

The FHA2 purification was based on the C-terminal His-tag. All buffers were refrigerated at 4 °C prior to use in the purification. The following protocol provided optimal purity and yield from 5 g cells. Purification of larger cell quantities may require further optimization. Cells (5 g) were suspended in 25 mL of wash buffer A (50 mM sodium phosphate at pH 8.0, 300 mM NaCl, 20 mM imidazole, and 0.5% (w/v) *N*-lauroylsarcosine (Sarkosyl) detergent and cell walls were lysed during four sonication periods of 1 min duration with a 1 min delay between sonication periods. Each period contained 0.8 s on/0.2 s off cycles with 80% amplitude during the on cycle. Cell debris was removed by centrifugation at 48,000g and 4 °C for 20 min. FHA2 in the centrifugation supernatant was bound to 0.5 mL of chelated cobalt His-Select resin with 1 h of mixing on a LabQuake shaker. A resin pellet was formed by centrifugation at 1000g for 1 min and the pellet was transferred to a column and then washed with 1 column volumes (0.5 mL) of wash buffer A. Detergent/buffer exchange was then completed by washing with 1 column volume each of wash buffer B (5 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes), 10 mM 2-(*N*-morpholino)ethanesulfonic acid, (MES), 20 mM imidazole, 0.5% Sarkosyl, 0.5% bTOG, and 0.4% C8E5 at pH 7.4) and wash buffer C (5 mM Hepes, 10 MES, 20 mM imidazole, 0.5% bTOG, and 0.4% C8E5 at pH 7.4). FHA2 was eluted from the resin with 5 column volumes (2.5 mL) of elution buffer (5 mM Hepes, 10 MES, 250 mM imidazole, 0.5% bTOG, and 0.4% C8E5 at pH 7.4). Protein concentrations were quantified using

A_{280} with the FHA2 $\epsilon_{280} = 34,000 \text{ M}^{-1} \text{ cm}^{-1}$ which was calculated from the numbers of tyrosines and tryptophans in the sequence and their extinction coefficients.

Solubilization, purification, and refolding of FHA2 from inclusion bodies

The inclusion body fraction of FHA2 was defined as the component which pelleted during the centrifugation of the cell lysate. The pellet was sonicated in denaturing lysis buffer (8 M urea, 100 mM NaH₂PO₄, 10 mM Tris-Cl), the suspension centrifuged, and the supernatant purified with chelated cobalt His-Select resin using methods similar to those described above. After binding the denatured protein to the resin, the column was washed with 3 column volumes of denaturing wash buffer (8 M urea, 100 mM NaH₂PO₄, 10 mM Tris-Cl, 20 mM imidazole) and the FHA2 was eluted with 5 column volumes of denaturing elution buffer (8 M urea, 100 mM NaH₂PO₄, 10 mM Tris-Cl, 250 mM imidazole). The denatured FHA2 solution was rapidly diluted into twice the volume of ice cold refolding buffer (1 M arginine, 10 mM Tris-Cl, 0.17% decyl maltoside, 2 mM EDTA at pH 8) and stored at 4 °C overnight. Removal of urea and refolding of FHA2 was achieved with dialysis at 4 °C for two days using 10,000 MWCO tubing and a dialysis buffer [7].

Circular dichroism spectroscopy

Spectra were obtained at 4 °C using a CD instrument (Chirascan, Applied Photophysics, Surrey, United Kingdom), a cuvette with 1 mm pathlength, a 260–200 nm spectral window, wavelength points separated by 0.5 nm, and 0.5 s signal averaging per point. For each sample, a difference spectrum was obtained by subtracting the sample buffer spectrum from the FHA2 spectrum. Refolded FHA2 was analyzed in dialysis buffer, and FHA2 from the native purification was analyzed in elution buffer without imidazole. The circular dichroism signal is reported in units of mean residue molar ellipticity which is a quantity normalized to the concentration of protein residues. The samples contained purified FHA2 and the mean residue molar ellipticity was therefore normalized to the FHA2 concentration.

Membrane reconstitution

The membrane composition was a 4:1 molar ratio of the ether linked lipids di-O-tetradecylphosphatidyl-choline (DTPC) and di-O-tetradecylphosphatidylglycerol (DTPG) and was chosen because: (1) choline is a predominant headgroup of lipids of membranes of respiratory epithelial host cells of the influenza virus; (2) the headgroup of DTPG is negatively charged like the headgroups of a minor fraction of the host cell lipids; and (3) DTPC and DTPG are ether- rather than ester-linked lipids and do not have a natural abundance ¹³C contribution to the carbonyl region probed in the NMR experiments [13,25]. The lipids (~40 mg total) and the detergent bTOG (~160 mg) were dissolved in chloroform. The solvent was removed by a stream of nitrogen gas and subsequent overnight pumping in a vacuum chamber. The lipid/detergent mixture was then dissolved in ~5 mL of 5 mM Hepes/10 mM MES buffer at pH 7.4. The FHA2 solution was added to the detergent/lipid solution to form a co-micelle solution of ~8 mg FHA2, detergent, and lipid. The solution was transferred to 10,000 MWCO tubing and dialyzed against 2 L of Hepes/MES buffer at pH 5.0. This pH is comparable to the one for fusion between influenza and endosomal membranes. The dialysis was done at 4 °C for three days with one buffer change. The FHA2 reconstituted in membranes was then harvested by centrifugation at 50,000g for 3 h.

Solid-state NMR spectroscopy

Data were obtained with a 9.4 T instrument (Varian Infinity Plus, Palo Alto, CA), a triple resonance magic angle spinning (MAS) probe, and a 4.0 mm diameter rotor with ~40 µL sample volume. It is estimated that the sample volume contained ~4 mg FHA2 and ~20 mg total lipid. Typical parameters of the rotational-echo double-resonance (REDOR) pulse sequence were: (1) 8.0 kHz MAS frequency; (2) a 6 µs ¹H π/2 pulse; (3) a 1.6 ms cross-polarization period with 63 kHz ¹H Rabi frequency and 80 kHz ¹³C Rabi frequency; (4) a 2 ms dephasing period with alternating 19 µs ¹⁵N π pulses and 8 µs ¹³C π pulses and 88 kHz two-pulse phase modulation (TPPM) ¹H decoupling; (5) ¹³C detection with 88 kHz TPPM ¹H decoupling; and (6) 1 s delay [13]. Data were acquired without (S_0) and with (S_1) the ¹⁵N π pulses during the dephasing period and, respectively, represented the full ¹³C signal and the ¹³C signal minus ¹³Cs directly bonded to ¹⁵N nuclei. The sample was cooled with nitrogen gas at -10 °C to counteract radiofrequency heating. Spectra were externally referenced to the methylene carbon of adamantane at 40.5 ppm, which corresponds to the ¹³C referencing used in liquid-state NMR of soluble proteins [21,26].

Results

Optimization of FHA2 expression in *E. coli*

Expression of amino acid-type isotopically labeled protein requires induction in minimal medium. However, when minimal medium was used for all stages of bacterial growth, the yield of FHA2 protein was less than 0.5 mg FHA2 per liter of culture. High-resolution structural methods such as NMR typically require at least 5 mg of protein, so preparation using growth solely in minimal medium would require large volumes of bacterial culture as well as large quantities of labeled amino acids. There are at least two general approaches to improving protein yield: increasing cell density and increasing protein expression for each cell. This paper describes progress with both approaches. Because the quantity of cell membrane is limited in a bacterial cell, excess FHA2 was sequestered in inclusion bodies. This paper also describes development of the solubilization, purification, and refolding protocols needed to obtain folded FHA2 from inclusion bodies.

The effort to increase cell density was focused on an initial growth of the cells in rich medium, followed by centrifugation and resuspension of the cells in minimal medium. Fig. 2A shows a comparison of the cell densities after growth in different media: minimal medium, LB, and LB supplemented with 10 g/L glycerol. The highest cell density was achieved with enriched LB and these cells also had the best subsequent growth in minimal medium. Fig. 2B demonstrates that the final cell density obtained with cell growth in rich medium and then minimal medium is ~5 times greater than the density obtained with growth only in minimal medium.

Because the cell densities for our protein expression are well above the normal range for shake flask fermentation, oxygenation may be a limiting factor on cell growth. The effect of oxygenation was first investigated by comparing growth with different types of flasks and closures, see Fig. 3. Baffled flasks are designed to increase oxygenation of the medium, and in our case significantly improved cell growth. The type of closure also had an effect on the oxygenation of the Parafilm (used as a control), foil, or a foam plug were compared, and the foam plug allowed for better oxygenation and increased growth. These improvements to the traditional method of using flat bottom Erlenmeyer flasks with foil closures increased the final cell density from 5 to 9.

The next set of experiments focused on growth in minimal medium after the medium switch and were based on the hypothesis that there was a correlation between cell growth in this

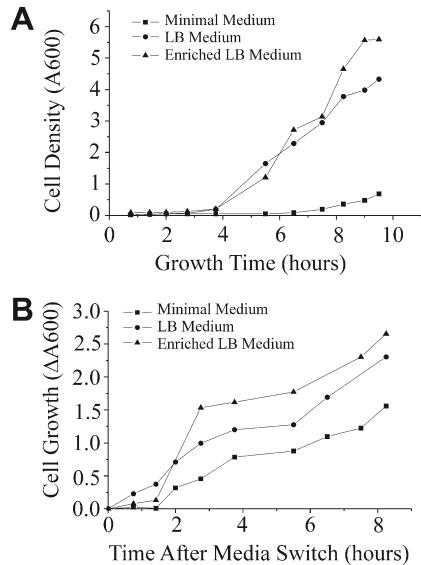


Fig. 2. (A) The initial cell growth varies with medium. Minimal medium shows the slowest cell growth and yields the lowest maximum density, while LB and LB enriched with 10 g/L of glycerol show more rapid initial cell growth, and enriched LB yields the highest final cell density. (B) Cell growth continues after the switch into minimal medium and the greatest growth is from enriched LB.

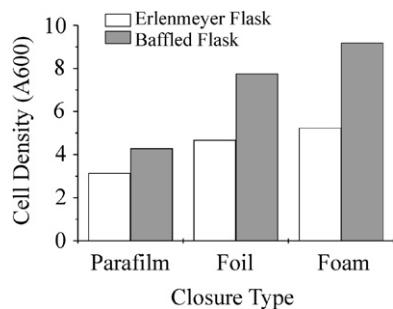


Fig. 3. The total cell density after overnight growth in enriched LB for different combinations of flask and closure types.

medium and expressed protein yield. The initial choice for the minimal medium was the commercially available M9 mix. Because the cell density was well above the typical density for this mixture, parameters such as pH and carbon source were varied to attain maximum cell growth.

As discussed earlier, growth of bacteria at high cell density may be impaired by a reduction in pH correlated with over-production of acetate and this was experimentally observed, see Fig. 4. The cell growth was then monitored as a function of initial pH of the minimal medium, see Fig. 5, and the highest growth was found for pH 8.0. Presumably, the pH reduction from an initial pH of 8.0 occurs in a pH range more amenable to bacterial growth. Increasing the buffer concentration had only minor effect on cell growth (not shown).

Glycerol may be a better carbon source than glucose because the uptake of glycerol into the cell occurs at a lower rate with less saturation of metabolic pathways and consequent lower production of acetate and reduction in pH [16,17]. Experiments were therefore carried out to study the effect of glycerol vs glucose on bacterial growth after the medium switch, see Fig. 6. For glycerol, the highest growth was obtained with 10 g/L and significantly less growth was observed with 5 or 20 g/L. For glucose, the highest growth was obtained with 5 g/L although reasonable growth was also obtained with 10 or 20 g/L. For 10 g/L of glucose, the pH of the medium was

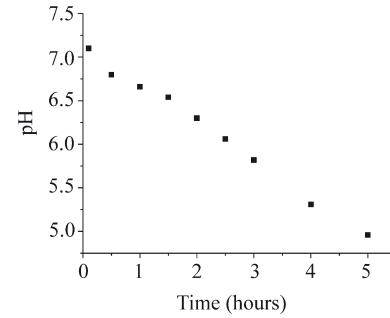


Fig. 4. There was a rapid drop in pH after resuspension of the cells in minimal medium containing the M9 mix of salts, 2.5 g/L MgSO₄, and 10 g/L glucose.

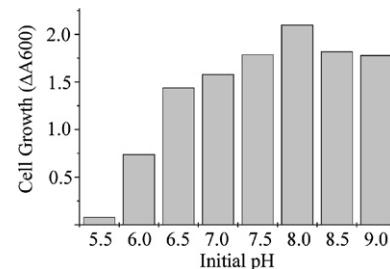


Fig. 5. The cell growth after the medium switch was monitored as a function of the initial pH of the minimal medium. Each bar represents the change in A₆₀₀ 3 h after the medium switch. The minimal medium contained the M9 mix of salts, 2.5 g/L MgSO₄, and 10 g/L glucose.

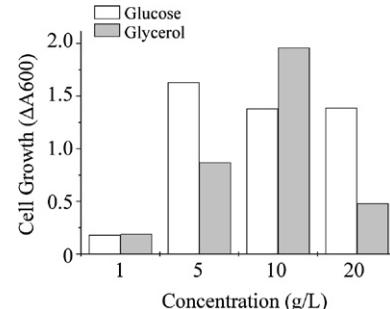


Fig. 6. Cell growth 3 h after the medium switch for different concentrations of glucose or glycerol in the minimal medium with initial pH of 8.0.

5.0 after 3 h of growth and for 10 g/L of glycerol, the pH was 6.4. This result correlated with the difference in cellular uptake rates of glucose and glycerol.

In summary, the highest cell density of A₆₀₀ ≈ 9 was achieved with: (1) overnight growth in LB enriched with 10 g/L glycerol; and (2) a switch to minimal medium containing M9 salts, 2.5 g/L MgSO₄, and 10 g/L glycerol at pH 8.0. For these conditions, there were ~10 g of wet cell mass per L fermentation culture.

Purification of native FHA2 from the soluble cell lysate

Purification of the FHA2 protein from *E. coli* cells takes advantage of the poly-histidine tag attached to the C-terminus of the protein, see Fig. 1. The cells were lysed by sonication, the insoluble components separated by centrifugation, and the soluble cell lysate bound to a chelated cobalt resin. Contaminants were removed by washing with a low concentration of imidazole buffer solution, and high purity protein was eluted with high concentration imidazole. N-Lauroyl sarcosine (Sarkosyl) was chosen to ini-

tially solubilize the protein during sonication but this detergent is difficult to remove after purification. Residual Sarkosyl can interfere with membrane reconstitution and has a large background signal in circular dichroism spectroscopy. The Sarkosyl detergent was therefore exchanged while the protein was bound to the resin. The first wash buffer contained both Sarkosyl and the target detergent mixture of bTOG and C8E5 and the second wash buffer and the elution buffer contained only the target detergent mixture. Relative to other detergents, FHA2 aggregation appeared smaller in the mixture of bTOG and C8E5 detergents. Evidence for removal of most of the Sarkosyl by this procedure included: (1) no obvious precipitation after addition of solutions containing divalent cations; (2) high signal-to-noise in the circular dichroism spectra without interference from strong far ultraviolet absorption; and (3) ability to reconstitute the FHA2 into membranes. Prior to exchange of the Sarkosyl, there was, respectively: (1) precipitation; (2) low signal-to-noise probably because of the far ultraviolet absorption of Sarkosyl; and (3) FHA2 did not reconstitute into membranes. Fig. 7 lanes 1, 2, 6, and 7 show the progress of the purification of native FHA2 from the soluble cell lysate and the high purity of the FHA2 in the eluent. The typical yield of FHA2 from this purification was 8–10 mg/L culture.

Circular dichroism (CD) spectroscopy

CD spectroscopy was used to assess the overall folding of the purified FHA2, see Fig. 8A. The shape of the spectrum is consistent with predominant helical secondary structure and the $\theta_{222\text{nm}} \approx -17,000 \text{ deg cm}^2 \text{dmol}^{-1}$ corresponds to ~50% of the residues in helical conformation [18]. There is a crystal structure for a soluble fragment representing residues 34–178 of FHA2 and a structure for a fragment in detergent representing residues 1–20 of FHA2 [9,12]. From these two fragments, ~60% of the residues are in helical conformation and the CD spectrum of FHA2 is therefore generally consistent with these previous structural studies.

Solubilization, purification, and refolding of FHA2 in inclusion bodies

The cell lysis was considered to be complete because increased lysis time did not increase the purified FHA2 yield. However, comparison of lanes 2 and 3 in Fig. 7 shows that a significant fraction of FHA2 is not solubilized in Sarkosyl detergent and is presumably in the form of inclusion bodies. An effort was therefore made to solubilize, purify, and refold FHA2 from this source. Different amounts

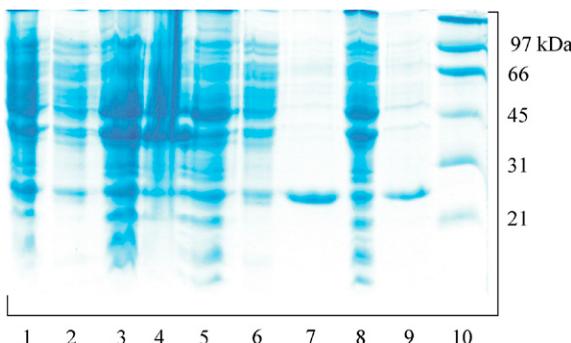


Fig. 7. SDS-PAGE gel electrophoresis at each step of the soluble lysate and inclusion body purifications. A “component” refers to the material used for the lane of the gel and components 2, 3, 4, and 5 are, respectively, the soluble cell lysate, the insoluble cell lysate, the insoluble cell lysate suspended in urea, and the portion of the insoluble cell lysate which is soluble in urea. Lane identification: (1) total cell lysate; (2) soluble portion of the cell lysate; (3) insoluble portion of the cell lysate; (4) component 3 suspended in urea; (5) portion of component 4 that is soluble in urea; (6) wash of component 2; (7) elution of component 2; (8) wash of component 5; (9) elution of component 5; (10) molecular weight standards.

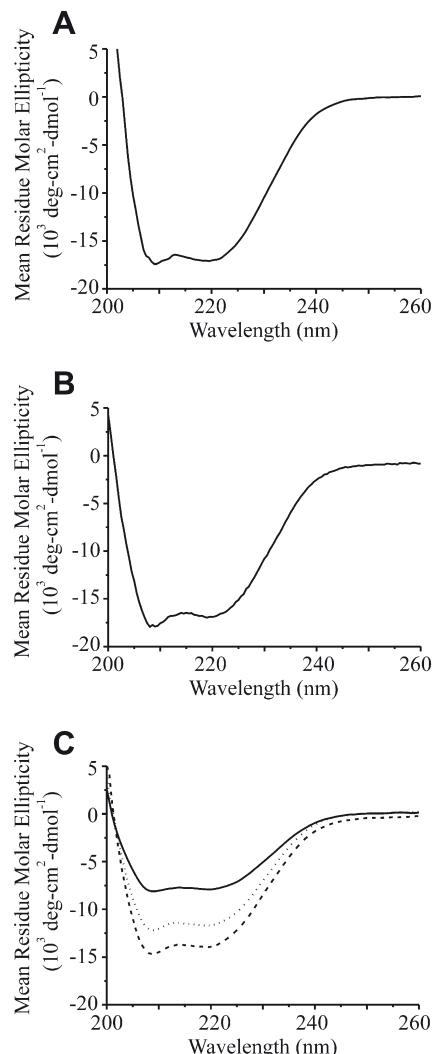


Fig. 8. Circular dichroism was used to determine the percentage helicity of FHA2. The temperature was 4 °C and the FHA2 concentrations were 30, 9, and 16 μM in A, B, and C, respectively. (A) FHA2 obtained from purification of native protein from the soluble cell lysate. The θ_{222} value corresponds to ~50% helical conformation. (B) FHA2 obtained from the denaturing purification of the inclusion body protein and refolded using the optimized protocol. The θ_{222} value corresponds to ~50% helical conformation. (C) Investigation of refolding with different dilution protocols: fast dilution at 4 °C (dash); slow dilution at 4 °C (dot); and fast dilution at room temperature (solid). These spectra were obtained after two days of dialysis.

of total protein were loaded in the lanes of the gel so the only meaningful comparison between the lanes is the purity of FHA2 relative to the other proteins. Using this metric, the purity of FHA2 is probably somewhat higher in the soluble cell lysate (lane 2) than in the insoluble cell lysate (lane 3).

The insoluble cell lysate pellet was sonicated in 8 M urea denaturing lysis buffer, centrifuged, and the FHA2 in the supernatant was purified using the chelated cobalt resin. Lanes 4 and 5 in Fig. 7 show that a large amount of inclusion body FHA2 was soluble in urea and lanes 8 and 9 show that the purification was effective. The FHA2 purity from this protocol is estimated to be ~90% based on relative band intensities in lane 9.

The FHA2 from inclusion bodies was purified in urea and therefore denatured and would only be useful in structural or functional studies if it could be refolded. The best refolding was accomplished by: (1) rapid dilution of the FHA2/urea solution into a solution containing decyl maltoside detergent, 10 mM Tris-Cl, 2 mM EDTA, and 1 M arginine at pH 8.0; (2) storage overnight at 4 °C; and (3)

two days of dialysis in buffer containing Tris-Cl and decyl maltoside to remove the urea and Arginine, with one buffer change [7]. Although arginine is required for refolding, its mode-of-action is not yet well-understood [19,20]. Very similar CD spectra were obtained for refolded FHA2 and for FHA2 from the purification, see Fig. 8A and B, which suggests that the refolding is quantitative. The typical yield of refolded FHA2 from inclusion bodies was 10 mg/L culture.

The effects of the rate and temperature of the dilution step in refolding were investigated using CD spectroscopy, see Fig. 8C. For the “rapid dilution” protocol, the FHA2/urea solution was pushed through a syringe with a narrow gauge needle into a stirring refolding solution. For the “slow dilution” protocol, the FHA2/urea solution was poured into the refolding solution and no attempt was made to quickly mix the solutions. For both protocols, the resultant mixture solution was dialyzed for two days to remove the urea. Both protocols were done at 4 °C and at room temperature and the most helical and presumably most folded FHA2 was obtained with rapid dilution at 4 °C. Lower helicity was observed for slow dilution at 4 °C, and even lower helicity was obtained for rapid or slow dilution at ambient temperature.

Membrane reconstitution

One long-term goal of this research is high-resolution structural characterization of FHA2 bound to membranes which is the most physiologically relevant state. Transfer of FHA2 from detergent to lipid was done by forming co-micelles of the lipid, detergent, and protein, and then removing the detergent via dialysis. The suspension was centrifuged and FHA2 was only detected in the lipid pellet and not in the supernatant, see lanes 2 and 3 in the gel of Fig. 9. This was the result both for FHA2 obtained from purification of the soluble cell lysate and for FHA2 obtained from denaturing purification of the inclusion bodies.

Solid-state NMR spectroscopy

Solid-state NMR spectroscopy in conjunction with isotopic labeling can provide information about conformation at specific residues in a large membrane-associated protein such as FHA2 [13]. In

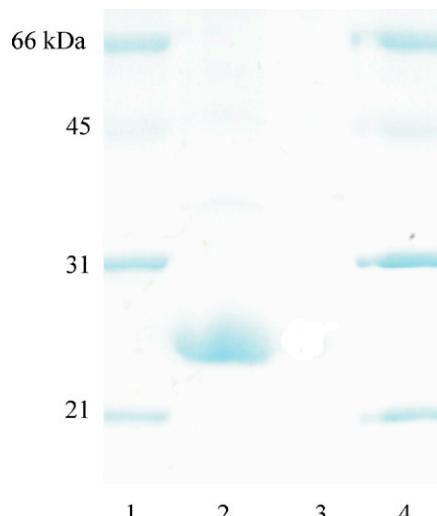


Fig. 9. SDS-PAGE gel of the pellet (lane 2) and supernatant (lane 3) from centrifugation of membrane-reconstituted FHA2. Lanes 1 and 4 are molecular weight standards. The protein had been purified from inclusion bodies and refolded. The absence of FHA2 in the supernatant suggests that most of the FHA2 is membrane-bound. The gel for membrane reconstitution of FHA2 purified from the soluble cell lysates also showed no protein in the supernatant.

this paper, solid-state NMR was applied to probe the conformation at Leu-98, see Fig. 10. The analysis relied on the well-known conformational dependences of ^{13}CO chemical shifts in proteins. For proteins with high-resolution structures, the distribution of Leu ^{13}CO shifts is 178.3 ± 1.3 ppm for Leu in helical conformation and 175.7 ± 1.5 ppm for Leu in β -strand conformation [21].

Fig. 10A displays the ^{13}C S_0 and S_1 REDOR spectra of a sample prepared with FHA2 obtained using purification of the soluble cell lysate. There appears to be substantial labeling as evidenced by the strong ^{13}CO carbonyl signal in the 170–180 ppm region. In addition, the S_1 ^{13}CO integrated signal intensity is ~6% smaller than the S_0 intensity and is consistent with ~80% incorporation of the $1\text{-}^{13}\text{C}$, ^{15}N Leu into FHA2 and with the loss of ^{13}CO signal from Leu-98 which is one of thirteen Leus in the sequence. This loss is expected because: (1) signals from ^{13}COs directly bonded to ^{15}N are highly attenuated in the S_1 spectrum; (2) Leu-98 residue is the only Leu in the sequence that is followed by a Leu; and (3) the natural abundance of ^{15}N is only 0.37%. The $S_0 - S_1$ difference spectrum is displayed in Fig. 10B and shows quantitative attenuation of natural abundance ^{13}C signals and a single sharp Leu-98 ^{13}CO signal with peak shift of 178.3 ppm and a full-width-at-half-maximum linewidth of ~2 ppm. The peak shift is more consistent with Leu-98 local helical conformation than with β -strand conformation. Fig. 10C displays a similar difference spectrum for a sample for which the FHA2 was purified and refolded from inclusion bodies. The peak shift of 178.4 ppm is also consistent with helical conformation. In the crystal structure of the soluble fragment representing residues 34–178 of FHA2, Leu-98 is part of a helix which extends from residue 38 to residue 105. The solid-state NMR data support retention of this conformation in membrane-associated FHA2. The solid-state NMR spectrum in Fig. 10C also provides residue-specific

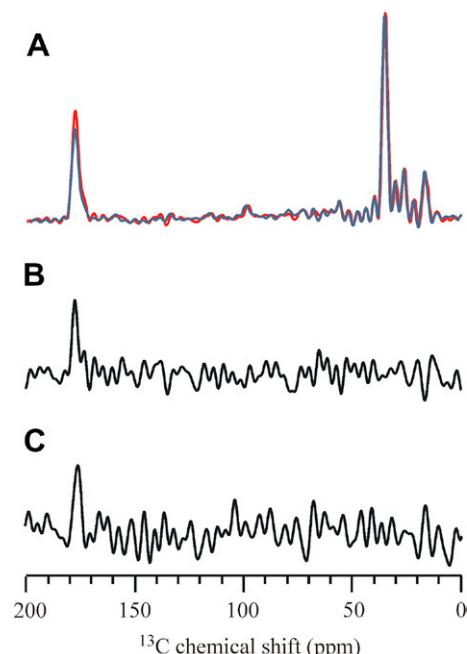


Fig. 10. ^{13}C solid-state NMR spectra of membrane-associated FHA2. (A) REDOR S_0 (red) and S_1 (blue) spectra of a sample with FHA2 labeled with $1\text{-}^{13}\text{C}$, ^{15}N leucine. The FHA2 was obtained from purification of the soluble lysate. (B) $S_0 - S_1$ difference of the two spectra from A. This difference is predominantly the Leu-98 ^{13}CO signal. (C) $S_0 - S_1$ difference spectrum of a sample containing FHA2 purified and refolded from inclusion bodies. The peak chemical shifts in spectra B and C are 178.3 and 178.4 ppm, respectively, and indicate helical conformation at Leu-98. The B and C spectra represent 16,000 and 90,000 total scans, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

support that the refolding of FHA2 was successful and are complementary to the CD analysis of refolding, Fig. 8B.

Discussion

This paper describes methods to express, purify, and refold the 23 kDa FHA2 membrane protein which comprises the full ectodomain of the influenza HA2 fusion protein. A yield of ~20 mg isotopically labeled FHA2/L cell culture was achieved by: (1) initially growing cells in rich medium followed by a switch to minimal medium prior to induction of expression; and (2) solubilization, purification, and refolding of FHA2 from inclusion bodies. The total process time from initializing cell growth to obtaining pure, folded protein is less than a week and most of that time is devoted to cell growth and dialysis, which require little attention from the scientist. The total yield is at least 40 times higher than the best yield obtained from growth solely in minimal medium. One reason for the larger yield was the increase in cell density A_{600} from ~2 in minimal medium growth to ~9 in rich medium/minimal medium growth. This increase relied on optimization of oxygenation and carbon source in the rich and minimal medium and on optimization of the pH and carbon source concentration in minimal medium.

The growth in rich medium is also rapid and should lead to a higher concentration of ribosomes within the cell and therefore greater production of expressed protein [22,23]. This supposition appeared valid for FHA2 as evidenced by the following analysis of yields of FHA2 from the soluble lysate. Relative to growth solely in minimal medium, there was a ~20-fold increase in the yield of FHA2 from growth in rich medium/minimal medium. This increase can be compared to the smaller ~5-fold increase in cell density. This difference suggests that relative to minimal medium growth, there was a ~4-fold increase in FHA2 per cell from growth in rich medium/minimal medium. This finding is consistent with a higher concentration of ribosomes in the cells.

Any inclusion body FHA2 solubilized by the 0.5% Sarkosyl detergent was likely folded after exchange into more benign detergent and purification as evidenced by the following properties of the final FHA2 product: (1) no precipitation; (2) high helical content in the CD spectrum; (3) induction of vesicle fusion with much greater activity at low pH than at neutral pH which correlated with the low pH of intact influenza viral fusion; and (4) quantitative reconstitution into membranes at least at the level of comparison of band intensities in a gel similar to the one displayed in Fig. 9 [13].

The FHA2 yield was doubled by solubilization, purification, and refolding of protein from inclusion bodies in the insoluble cell lysate. The refolded inclusion body protein could be isotopically labeled and was quantitatively incorporated into membranes, see Figs. 9 and 10C. Refolding was evidenced by the strong similarity of the CD spectrum of the protein that had undergone the refolding protocol to the spectrum of the protein that had never been unfolded in urea, see Fig. 8A and B. Most of the FHA2 structure is outside the membrane and folding is therefore driven in large part by the hydrophobic effect. The high helical content of folded FHA2 is primarily due to formation of a leucine zipper-like trimeric coiled coil where each FHA2 molecule in the trimer contributes one helix and the leucine zipper forms because of the hydrophobic effect [9]. From the point of view of the hydrophobic effect, it is difficult to understand how the protein would be both soluble and retain high helical content without formation of folded trimers.

Quantitative reconstitution (from Fig. 9) of FHA2 into membranes provided some evidence that the FHA2 was not extensively aggregated as both membrane binding and aggregation would likely occur through the hydrophobic fusion peptide and aggregated fusion peptide would be less likely to bind to membranes than non-aggregated peptide. Poorer membrane binding was observed for FHA2 in detergents other than bTOG detergent and

the poorer binding was interpreted as due to more extensive aggregation in these other detergents. In Fig. 10B and C, detection of helical conformation for Leu-98 in membrane-reconstituted FHA2 by solid-state NMR does provide some evidence that the ectodomain is correctly folded. Leu-98 has helical conformation in the native fold and is not in the membrane binding N-terminal domain of FHA2.

Comparison of CD spectra under different conditions showed that the best refolding was obtained from rapid dilution of the FHA2/urea solution at 4 °C, see Fig. 8C. The refolding literature proposes that there is competition between aggregation and refolding and this model can be applied to interpret of some of the FHA2 refolding results [19,24]. For example, in the slow dilution protocol, diffusion of urea will likely be faster than diffusion of FHA2 and the locally high concentrations of unfolded FHA2 may result in aggregation of FHA2. Although aggregation was not visually observed, there was lower helical content with slow dilution. The successful refolding of the inclusion body protein suggests that even higher yields of FHA2 could be obtained with longer induction times and the consequent production of greater quantities of FHA2-containing inclusion bodies.

The cell densities and FHA2 yields described in this paper are much larger than what is typical in the current literature for non-bacterial recombinant membrane proteins and the FHA2 methods should be applicable to some of these proteins. Strengths of the methods include use of shake flasks rather than fermenters and having half the protein folded in the bacterial membrane and subsequently purified using native methods. The short histidine tag did not interfere with folding and may be compared to larger chaperone proteins such as the maltose binding protein and GST. It is usually desirable to cleave the chaperone protein after purification and then refold the target protein. Each of these steps may be difficult and neither was required for the purification of native FHA2 with the His-tag.

Acknowledgments

The FHA2 plasmid was provided by Yeon-Kyun Shin, and financial support was provided by NIH AI47153 and the Michigan State University Quantitative Biology Initiative. Advice and assistance were also provided by Dr. R. Michael Garavito and by the REF Center for Structural Biology of Membrane Proteins (REF03-016) at Michigan State.

References

- [1] A.J. Link, G. Georgiou, Advances and challenges in membrane protein expression, *AIChE J.* 53 (2007) 752–756.
- [2] M.L. Cai, Y. Huang, K. Sakaguchi, G.M. Clore, A.M. Gronenborn, R. Craigie, An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*, *J. Biomol. NMR* 11 (1998) 97–102.
- [3] O.J. Murphy 3rd, F.A. Kovacs, E.L. Sicard, L.K. Thompson, Site-directed solid-state NMR measurement of a ligand-induced conformational change in the serine bacterial chemoreceptor, *Biochemistry* 40 (2001) 1358–1366.
- [4] S.M. Singh, A.K. Panda, Solubilization and refolding of bacterial inclusion body proteins, *J. Biosci. Bioeng.* 99 (2005) 303–310.
- [5] H. Lilie, E. Schwarz, R. Rudolph, Advances in refolding of proteins produced in *E. coli*, *Curr. Opin. Biotechnol.* 9 (1998) 497–501.
- [6] B.M. Gorzelle, J.K. Nagy, K. Oxenoid, W.L. Lonzer, D.S. Cafiso, C.R. Sanders, Constitutive refolding of diacylglycerol kinase, an integral membrane protein, *Biochemistry* 38 (1996) 16373–16382.
- [7] S.E. Swalley, B.M. Baker, L.J. Calder, S.C. Harrison, J.J. Skehel, D.C. Wiley, Full-length influenza hemagglutinin HA₂ refolds into the trimeric low-pH-induced conformation, *Biochemistry* 43 (2004) 5902–5911.
- [8] J.J. Skehel, D.C. Wiley, Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin, *Annu. Rev. Biochem.* 69 (2000) 531–569.
- [9] J. Chen, J.J. Skehel, D.C. Wiley, N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA₂ subunit to form an N cap that terminates the triple-stranded coiled coil, *Proc. Natl. Acad. Sci. USA* 96 (1999) 8967–8972.
- [10] I.A. Wilson, J.J. Skehel, D.C. Wiley, Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution, *Nature* 289 (1981) 366–373.

- [11] J.C. Macosko, C.H. Kim, Y.K. Shin, The membrane topology of the fusion peptide region of influenza hemagglutinin determined by spin-labeling EPR, *J. Mol. Biol.* 267 (1997) 1139–1148.
- [12] X. Han, J.H. Bushweller, D.S. Cafiso, L.K. Tamm, Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin, *Nat. Struct. Biol.* 8 (2001) 715–720.
- [13] J. Curtis-Fisk, C. Preston, Z. Zheng, R.M. Worden, D.P. Weliky, Solid-state NMR structural measurements on the membrane-associated influenza fusion protein ectodomain, *J. Am. Chem. Soc.* 129 (2007) 11320–11321.
- [14] B. Buck, J. Zamoon, T.L. Kirby, T.M. DeSilva, C. Karim, D. Thomas, G. Veglia, Over-expression, purification, and characterization of recombinant Ca-ATPase regulators for high-resolution solution and solid-state NMR studies, *Protein Expr. Purif.* 30 (2003) 253–261.
- [15] C.L. Tian, R.M. Breyer, H.J. Kim, M.D. Karra, D.B. Friedman, A. Karpay, C.R. Sanders, Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor, *J. Am. Chem. Soc.* 127 (2005) 8010–8011.
- [16] G.W. Luli, W.R. Strohl, Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations, *Appl. Environ. Microbiol.* 56 (1990) 1004–1011.
- [17] M.A. Eiteman, E. Altman, Overcoming acetate in *Escherichia coli* recombinant protein fermentations, *Trends Biotechnol.* 24 (2006) 530–536.
- [18] E.K. Oshea, R. Rutkowski, P.S. Kim, Evidence that the leucine zipper is a coiled coil, *Science* 243 (1989) 538–542.
- [19] B.M. Baynes, D.I.C. Wang, B.L. Trout, Role of arginine in the stabilization of proteins against aggregation, *Biochemistry* 44 (2005) 4919–4925.
- [20] Y.D. Liu, J.J. Li, F.W. Wang, J. Chen, P. Li, Z.G. Su, A newly proposed mechanism for arginine-assisted protein refolding—not inhibiting soluble oligomers although promoting a correct structure, *Protein Expr. Purif.* 51 (2007) 235–242.
- [21] H.Y. Zhang, S. Neal, D.S. Wishart, RefDB: a database of uniformly referenced protein chemical shifts, *J. Biomol. NMR* 25 (2003) 173–195.
- [22] R.J. Harvey, Regulation of ribosomal protein synthesis in *Escherichia coli*, *J. Bacteriol.* 101 (1970) 574–583.
- [23] R.J. Harvey, Fraction of ribosomes synthesizing protein as a function of specific growth-rate, *J. Bacteriol.* 114 (1973) 287–293.
- [24] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J.S. Philo, T. Arakawa, Role of arginine in protein refolding, solubilization, and purification, *Biotechnol. Prog.* 20 (2004) 1301–1308.
- [25] S.A. Rooney, L.L. Nardone, D.L. Shapiro, E.K. Motoyama, L. Gobran, N. Zaehringer, Phospholipids of rabbit type II alveolar epithelial cells: comparison with lung lavage, lung tissue, alveolar macrophages, and a human alveolar tumor cell line, *Lipids* 12 (1977) 438–442.
- [26] C.R. Morcombe, K.W. Zilm, Chemical shift referencing in MAS solid state NMR, *J. Magn. Reson.* 162 (2003) 479–486.