Downloaded from www.proteinscience.org on April 6, 2006

Expression, purification, and characterization of Thermotoga maritima membrane proteins for structure determination

Linda Columbus, Jan Lipfert, Heath Klock, Ian Millett, Sebastian Doniach and Scott A. Lesley

Protein Sci. published online Apr 5, 2006;
doi:10.1110/ps.051874706

P<P Published online April 5, 2006 in advance of the print journal.

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

Online First contains unedited articles in manuscript form that have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Online First articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Online First articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Protein Science* go to:
http://www.proteinscience.org/subscriptions/

© 2006 Cold Spring Harbor Laboratory Press

Downloaded from www.proteinscience.org on April 6, 2006

Expression, purification, and characterization of *Thermotoga maritima* membrane proteins for structure determination

LINDA COLUMBUS,[1] JAN LIPFERT,[2] HEATH KLOCK,[5] IAN MILLETT,[2] SEBASTIAN DONIACH,[2,3,4] AND SCOTT A. LESLEY[1,5]

[1]The Joint Center for Structural Genomics and The Scripps Research Institute, Department of Molecular Biology, La Jolla, California 92037, USA
Departments of [2]Physics and [3]Applied Physics and [4]Biophysics Program, Geballe Laboratory of Advanced Materials, Stanford University, Stanford, California 94305, USA
[5]The Joint Center for Structural Genomics and the Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA

(RECEIVED September 28, 2005; FINAL REVISION January 2, 2006; ACCEPTED January 27, 2006)

Abstract

Structural studies of integral membrane proteins typically rely upon detergent micelles as faithful mimics of the native lipid bilayer. Therefore, membrane protein structure determination would be greatly facilitated by biophysical techniques that are capable of evaluating and assessing the fold and oligomeric state of these proteins solubilized in detergent micelles. In this study, an approach to the characterization of detergent-solubilized integral membrane proteins is presented. Eight *Thermotoga maritima* membrane proteins were screened for solubility in 11 detergents, and the resulting soluble protein–detergent complexes were characterized with small angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) spectroscopy, circular dichroism (CD) spectroscopy, and chemical cross-linking to evaluate the homogeneity, oligomeric state, radius of gyration, and overall fold. A new application of SAXS is presented, which does not require density matching, and NMR methods, typically used to evaluate soluble proteins, are successfully applied to detergent-solubilized membrane proteins. Although detergents with longer alkyl chains solubilized the most proteins, further characterization indicates that some of these protein–detergent complexes are not well suited for NMR structure determination due to conformational exchange and protein oligomerization. These results emphasize the need to screen several different detergents and to characterize the protein–detergent complex in order to pursue structural studies. Finally, the physical characterization of the protein–detergent complexes indicates optimal solution conditions for further structural studies for three of the eight overexpressed membrane proteins.

Keywords: membrane proteins; NMR; SAXS; protein–detergent complexes; detergent micelles

Supplemental material: see www.proteinscience.org

Reprint requests to: Linda Columbus, The Scripps Research Institute, Department of Molecular Biology, 10550 N. Torrey Pines Road MB-44, La Jolla, CA 92037, USA; e-mail: columbus@scripps.edu; fax: (858) 784-8014.
Abbreviations: CMC, critical micelle concentration; PDC, protein–detergent complex; NMR, nuclear magnetic resonance; SAXS, small angle X-ray scattering; CD, circular dichroism; SDS, sodium n-dodecyl sulfate; OG, n-octyl-β-D-glucoside; NG, n-nonyl-β-D-glucoside; DG, n-decyl-β-D-glucoside; DM, n-decyl-β-D-maltoside; DoDM, n-dodecyl-β-D-maltoside; CHAPS, 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate; DHPC, 1,2-dihexanoyl-sn-glycero-phosphocholine; FC-10, n-decylphosphocholine; FC-12, n-dodecylphosphocholine; LDAO, lauryldimethylamine oxide; LPPG, 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)]; DSG, disuccinimidyl gluterate; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051874706.

Downloaded from www.proteinscience.org on April 6, 2006

Preparation of integral membrane proteins for structure elucidation is notoriously difficult due to two major barriers: expression yields of properly folded proteins, and the selection of appropriate (i.e., solubilizing) solution conditions. With the advent of high-throughput structural proteomics methods, expression of membrane proteins can be screened rapidly. In addition, different *Escherichia coli* strains (Miroux and Walker 1996), coexpression with helper proteins (Chen et al. 2003), and fusion tags (Weiss and Grisshammer 2002) have been developed to facilitate recombinant-expression of membrane proteins. However, difficulties still remain in optimizing the expression of membrane proteins. Limited empirical data have been acquired on membrane protein expression, and it is unclear why some overexpress and others do not.

The second major barrier in membrane protein structure determination is the choice of solubilizing reagents. Typically the solubilizing reagent is a detergent, which, when above its CMC, forms a protein–detergent complex (PDC) in which micelles encapsulate the hydrophobic regions of the protein while exposing their hydrophilic head groups. Although the hydrophobic interactions between the detergent and the protein are similar regardless of which protein or detergent is considered, a protein is generally only soluble in a select few detergents, which vary depending on the protein. Therefore, for each protein, several solubilizing conditions need to be screened. Recent studies have utilized high-throughput efforts to screen expression (Eshaghi et al. 2005; Korepanova et al. 2005) and detergent conditions (Eshaghi et al. 2005) to generate a few conditions that yielded crystals (Dobrovetsky et al. 2005) or NMR spectra (Krueger-Koplin et al. 2004; Tian et al. 2005). However, screening solubility is not the only requirement for structural studies of membrane protein samples; the proteins need to be properly folded in the detergent conditions. Detergents can solubilize membrane proteins in inactive conformations (Vinogradova et al. 1998; Bowie 2001) and non-native folds (i.e., soluble aggregates). In the absence of a functional assay, methods for screening the overall fold and homogeneity of PDCs are needed in order to facilitate the structure determination of the ~50% of membrane proteins that have unknown function.

In this study, the high-throughput expression pipeline of the Joint Center for Structural Genomics (Lesley et al. 2002) was used to assay the expression of *Thermotoga maritima* α-helical membrane proteins in *E. coli*. *T. maritima* has ~446 putative α-helical membrane proteins that vary from one to 16 transmembrane helices (Fernando et al. 2004; http://144.16.71.10/thgs/). Forty-five of the smallest α-helical membrane proteins (MW < 16 kDa; shown in Supplemental Fig. 1) were selected as protein targets for solution NMR structure characterization and elucidation. Although TROSY-based NMR techniques have advanced the molecular weight limitations to ~100 kDa (Riek et al. 2000), smaller proteins were chosen due to the additional contribution that the detergent micelle and potential oligomers will have on the overall molecular tumbling.

The 45 targeted *T. maritima* membrane proteins were assayed for expression in *E. coli*. The expressing proteins were solubilized, purified, and subjected to a screen of 11 detergents (Table 1) in order to obtain soluble membrane proteins for structural studies. The detergent-soluble proteins were characterized with small angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) spectroscopy, circular dichroism (CD) spectroscopy, and chemical cross-linking to evaluate the homogeneity, the oligomeric state, the radius of gyration, and the overall fold.

Table 1. *Detergent properties*

Detergent (abbreviation)	MW (Da)	CMC (mM)	M_{mic} from literature	~Micelle size (Da)	$V_{monomer}$ (Å^3)	$\rho_{electron}$ (e/Å^3)	Ionic property
n-octyl-β-D-glucopyranoside (OG)	292	18[a]–23[b]	27–100[a]	20,000	418.6	0.382	Nonionic
n-nonyl-β-D-glucopyranoside (NG)	306	6.5[c]	NR	NR	445.5	0.377	Nonionic
n-decyl-β-D-glucopyranoside (DG)	320	2.2[d]	200–400[e]	96,000	472.4	0.373	Nonionic
n-decyl-β-D-maltoside (DM)	483	1.8[f]	69[f]	33,000	644.0	0.407	Nonionic
n-dodecyl-β-D-maltoside (DoDM)	511	0.17[f]	78–149[f]	58,000	697.8	0.398	Nonionic
1,2-dihexanoyl-sn-glycero-phosphocholine (DHPC)	453	14–15[g]	35[g]	16,000	731.0	0.363	Zwitterionic
3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS)	615	8.0[h]	10[i]	6000	830.3	0.405	Zwitterionic
Lauryldimethylyamine oxide (LDAO)	229	1–2[j]	76[k]	17,000	430.5	0.302	Zwitterionic
n-decylphosphocholine (FC-10)	323	11[f]	NR	NR	494.3	0.360	Zwitterionic
n-dodecylphosphocholine (FC-12)	351	1.5[f]	50–60[l]	19,000	548.1	0.354	Zwitterionic
1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LPPG)	507	0.018[m]	~125[n]	63,000	692.8	0.395	Ionic

[a] Lorber et al. 1990; [b] Chattopadhyay and London 1984; [c] DeGrip and Bovee-Geurts 1979; [d] Helenius et al. 1979; [e] Nilsson et al. 1998; [f] Anatrace, Inc. (2003 catalog); [g] Tausk et al. 1974, Burns et al. 1982; [h] Hjelmeland et al. 1983; [i] Womack et al. 1983; [j] Herrmann 1962; [k] Herrmann 1966; [l] le Maire et al. 2000; [m] Stafford 1989; [n] Chou et al. 2004.

NR, Reference was not available; ND, no data were obtained for this detergent; NC, the values could not be calculated due to interparticle interference; M_{mic}, the aggregation number.

Downloaded from www.proteinscience.org on April 6, 2006

SAXS is a technique used to characterize the size and shape of macromolecules in solution, and has been widely applied to water-soluble proteins and nucleic acids (Doniach 2001; Svergun and Koch 2003). In addition, detergent micelles have been investigated with small-angle neutron scattering (Hayter and Zem 1982; Bendedouch et al. 1983; Lin et al. 1986; Bezzobotnov et al. 1988; Thiyagarajan and Tiede 1994) and SAXS (Dupuy et al. 1997; Zhang et al. 1999; He et al. 2002). A major challenge in the analysis of the scattering data of PDCs is that the signal has contributions from both the PDC and ''empty'' detergent micelles, which need to be separated in order to obtain information about the size and shape of the PDC. Engelman and coworkers (Bu and Engelman 1999) used SAXS to characterize the molecular weight and the radius of gyration of membrane proteins in detergent micelles by carefully matching the average electron density of the buffer to that of the detergent, such that the scattering contribution from the detergent molecules cancels. However, matching the electron density in this way is only possible for a select few detergents that have an electron density similar to that of the buffer. The scattering contrast for SAXS measurements is difficult to adjust, in contrast to neutron scattering experiments for which the scattering contrast of the buffer solution can be changed relatively easily by adjusting the ratio of D_2O to H_2O. An alternative approach, employed by Loll and coworkers (Loll et al. 2001) in light scattering studies on PDCs, is to exhaustively dialyze the sample against a buffer of known micelle concentration to ensure a fixed concentration of free detergent micelles and to subtract their scattering signal. This method is problematic for screens involving many different protein–detergent combinations, as for each PDC, dialysis conditions need to be optimized and several days of dialysis are required. In this study, a SAXS method is presented to estimate the oligomerization state of the protein and to measure the radius of gyration of the PDC without the need to match solvent or detergent properties or perform dialysis providing a method well suited for high-throughput screening.

One-dimensional (1D) ^{1}H (Rehm et al. 2002; Peti et al. 2004) and two-dimensional (2D) ^{15}N, ^{1}H-HSQC NMR (Folkers et al. 2004) spectroscopy has been implemented into several structural genomics centers in order to evaluate the quality of protein samples for NMR and crystallographic structure determination. In this study, the applicability of 1D ^{1}H NMR screening to membrane proteins is explored. A 1D ^{1}H NMR spectrum can be recorded in a few minutes, and a qualitative estimate of the amount of secondary structure is immediate. For PDCs, the 1D ^{1}H NMR spectrum is complicated by the additional signal from the detergent; however, most detergents do not have signal in the amide proton ''footprint'' region, which is often used to evaluate the secondary structure (Wüthrich 1986). A caveat is that α-helical proteins, in general, have less signal dispersion than β-sheet proteins due to the length and strength of the hydrogen bonds (Wishart and Sykes 1994), which can make the evaluation of the overall fold of α-helical proteins solely on proton dispersion difficult. In this study, the evaluation of PDCs based on 1D ^{1}H NMR spectra is compared with other biophysical techniques as well as 2D ^{15}N, ^{1}H-TROSY NMR spectra to determine the reliability of 1D ^{1}H NMR screening of α-helical membrane proteins.

In addition to SAXS and NMR, far-UV CD spectra and chemical cross-linking data were obtained for all PDCs to confirm the helical secondary structure and to probe the oligomeric states of the proteins, respectively. The data provide a comprehensive study of the solubility and physical properties of eight membrane proteins as well as establishing NMR and SAXS screening methods. From the screen of 45 membrane proteins, detergent conditions for three of the proteins were determined to be suitable for structural studies.

Results

Expression, localization, and purification

Large quantities of material are required for structural studies; therefore, only highly expressing targets are of interest. The expression of the 45 targets was assayed by electrophoresis of the cell lysates on an SDS-PAGE gel shown in Figure 1. Ten of the proteins overexpressed at substantial levels in *E. coli*, and are indicated in bold print (Fig. 1). The localization (membrane, soluble, or insoluble fraction) of each of the expressing proteins was determined. Examples of the two types of fractionation that were observed are shown in Figure 2A. TM1554 and TM0093 had a low overall expression and were localized predominately to the insoluble fraction. The other eight overexpressed proteins localized to the membrane as well as the insoluble fraction. Because of the difficulty in refolding α-helical membrane proteins (Kiefer 2003), the optimal localization for structural studies is to the membrane, and only proteins with detectable overexpression in the membrane were pursued; the molecular weight and the number of putative transmembrane α-helices are tabulated in Table 2. Each of the proteins that expressed and localized to the membrane was then solubilized in DM and purified with Co^{2+}-affinity chromatography. An SDS-PAGE gel for each fraction of the Co^{2+}-affinity purification of TM1514 is shown in Figure 2B as an example. All proteins were observed in the eluent and were substantially pure as assessed from the Coomassie-stained gel. After concentrating and dialyzing, the final yield of the proteins was ~2–5 mg/mL for detergent conditions in which the protein was soluble.

Downloaded from www.proteinscience.org on April 6, 2006

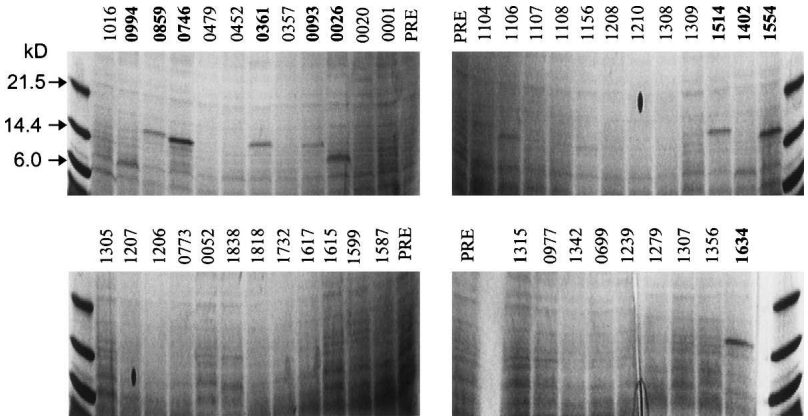

Figure 1. Expression of *T. maritima* membrane proteins in *E. coli*. A Coomassie-stained denaturing SDS-PAGE gel of the cell lysate of the 45 proteins after induction. Each target is labeled with the TM (*Thermotoga maritima*) protein identification number, the overexpressing proteins are labeled in bold, and the cell lysate pre-induction control is labeled PRE.

Solubility screen

Eleven detergents were chosen based on physical properties, such as head group, chain length, and charge, and success with previous studies; some of their physical properties are listed in Table 1. Table 2 summarizes the solubility of each protein in the 11 different detergents screened. An example of the results from the detergent screen for TM0361 is shown in Figure 2C. TM0361 is only significantly soluble in four out of the 11 detergents: CHAPS, LDAO, FC-10, and LPPG (assayed on a different

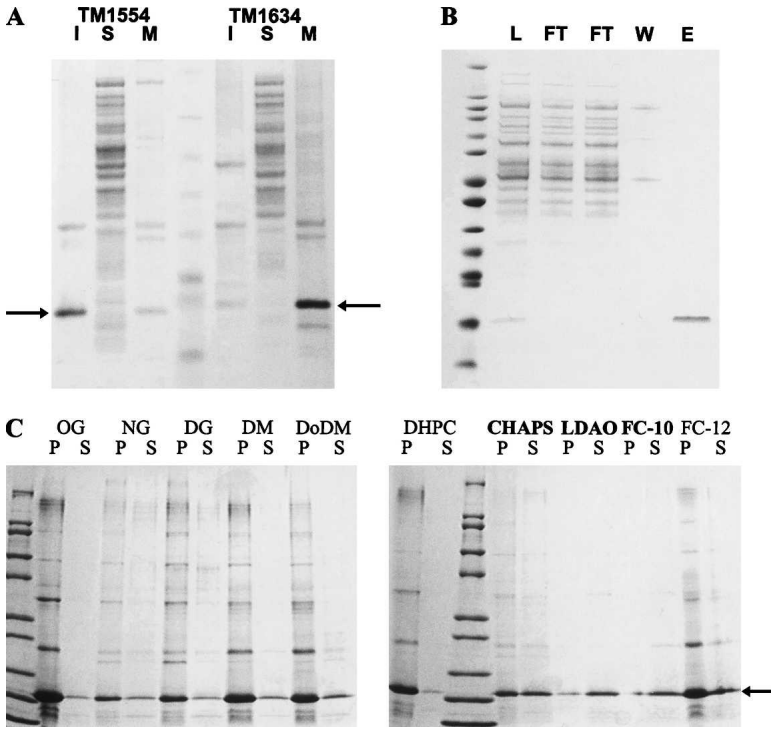

Figure 2. Localization, purification, and solubility of *T. maritima* membrane proteins. (*A*) A Coomassie-stained SDS-PAGE gel of the insoluble (I), soluble (S), and membrane (M) fractions for TM1514 and TM1634. The arrow indicates the protein band of each protein at the appropriate molecular weight. (*B*) Coomassie-stained denaturing SDS-PAGE gel of the loaded sample (L), the flow through (FT), the wash (W), and the elution (E) fractions of the Co^{2+}-affinity purification of TM1514. (*C*) The solubility of TM0361 in various detergents. A Coomassie-stained SDS-PAGE gel of the supernatant (S) and precipitant (P) of TM0361 in each detergent. LPPG is not shown because it was assayed on a different gel. The detergents that yielded soluble protein are in bold, and the arrow indicates the protein band corresponding to TM0361.

Downloaded from www.proteinscience.org on April 6, 2006

Table 2. *Solubility of protein/detergent complexes and assessment of secondary structure and overall fold with 1D ^{1}H NMR of T. maritima α-helical membrane proteins*

Protein	MW	#TM	OG	NG	DG	DM	DoDM	DHPC	CHAPS[a]	LDAO	FC-10	FC-12	LPPG
TM0026	9.6	2				++	+			+	–	++	–
TM0361	15.1	1								–	–		+
TM0746	16.5	1		A/U						+			–
TM0859	14.3	1					–			+	–	–	–
TM0994	5.5	1	+				–						–
TM1402	7.6	1				–	+					+	–
TM1514	16.9	4					–	–			–		–
TM1634	16.6	1	A/U	A/U		+	–	+		++	–		–

The gray shaded areas correspond to conditions in which the proteins were insoluble. The white areas indicate a condition in which the protein was soluble. Based on the evaluation of the 1D ^{1}H NMR spectra as shown in Figure 5, each soluble condition is classified as: A/U, aggregated and/or unfolded; ++, high quality; +, good quality; –, poor quality.

[a]CHAPS has a proton resonance in the amide region that interferes with the evaluation of the PDC.

gel). There were a few noteworthy trends observed in the solubility of the eight membrane proteins in the different detergents. LPPG solubilized all of the proteins, and DoDM solubilized six; therefore, in terms of protein solubilization these detergents are superior. In contrast, the glucosides (especially DG) and DHPC performed poorly in solubilizing any of the proteins. The number of detergents in which each protein was soluble varied from three (TM0994) to eight (TM1634). The soluble PDCs were analyzed with 1D ^{1}H NMR spectroscopy, SAXS, CD spectroscopy, and chemical cross-linking.

1D ^{1}H NMR spectroscopy

One-dimensional ^{1}H NMR spectra of each PDC were analyzed for signal intensity, line widths, and dispersion in the indole, aromatic, and amide proton region (6–11 ppm). Examples of 1D spectra of various qualities are shown in Figure 3. Chemical shifts >8.5 ppm indicate the presence of protein secondary structure (arrows in Fig. 3A for TM0026 in DM and FC-12). The superposition of sharp lines and a very broad signal intensity suggest that the protein is partially unfolded and aggregated (Fig. 3B, TM0746 in NG). Broad lines, such as those observed for TM0746 in LPPG, could be a result of conformational exchange, oligomerization, and/or aggregation. Based on these evaluations, all of the samples were ranked for NMR structure determination as A/U, unfolded or aggregated; poor (–), probably is an oligomer or has chemical exchange; promising (+), possibly folded; or very promising (++), folded. The results are shown in Table 2.

Chemical cross-linking

The oligomeric states of the protein targets are unknown. In order to investigate the oligomeric state of the proteins, each PDC was chemically cross-linked and evaluated on a denaturing SDS-PAGE gel. Without cross-linking reagents, some of the proteins migrate on a denaturing SDS-PAGE gel as a monomer and/or oligomer depending on the initial detergent before adding the SDS loading buffer. For example, the migration of TM1514 (without cross-linking reagent) in four different initial detergents is shown in Figure 4A. Higher oligomeric states, as well as monomers, are observed in DHPC and FC-10, indicating the protein may be a higher-order oligomer. To probe the oligomeric species, a condition for which TM1514 migrates predominantly as a monomer (DoDM) was subjected to cross-linking with DSG, and the result is shown in Figure 4B. Only dimeric and monomeric species were observed, indicating that TM1514 in DoDM is likely a dimer. Monomers are observed because the cross-linking efficiency of DSG is low due to the competing hydrolysis reaction of the N-hyrdoxy succimide-ester. All soluble PDCs were chemically cross-linked with DSG and/or EDC to investigate the oligomeric state, and the results are summarized in Table 3. In general, the oligomeric states of the proteins were the same in each soluble detergent condition. These data should be taken as estimates due to the problems associated with cross-linking, some of which are the detection of collision complexes and the accessibility of the cross-linker to the necessary functional groups. Although concentrations and reaction times were optimized for each protein to eliminate collision complexes, these are still issues in evaluating the data.

SAXS: Protein–detergent complexes

SAXS provides a model-free measurement of the radius of gyration and total excess electron density of the PDC through Guinier analysis of the forward scattering intensity. However, the measurement is complicated by the fact that both the PDC and the ''empty'' micelle contribute to the scattering signal and their relative contributions are, a priori, unknown. In this study, two different estimates of the PDC scattering are obtained, which bracket the

Downloaded from www.proteinscience.org on April 6, 2006

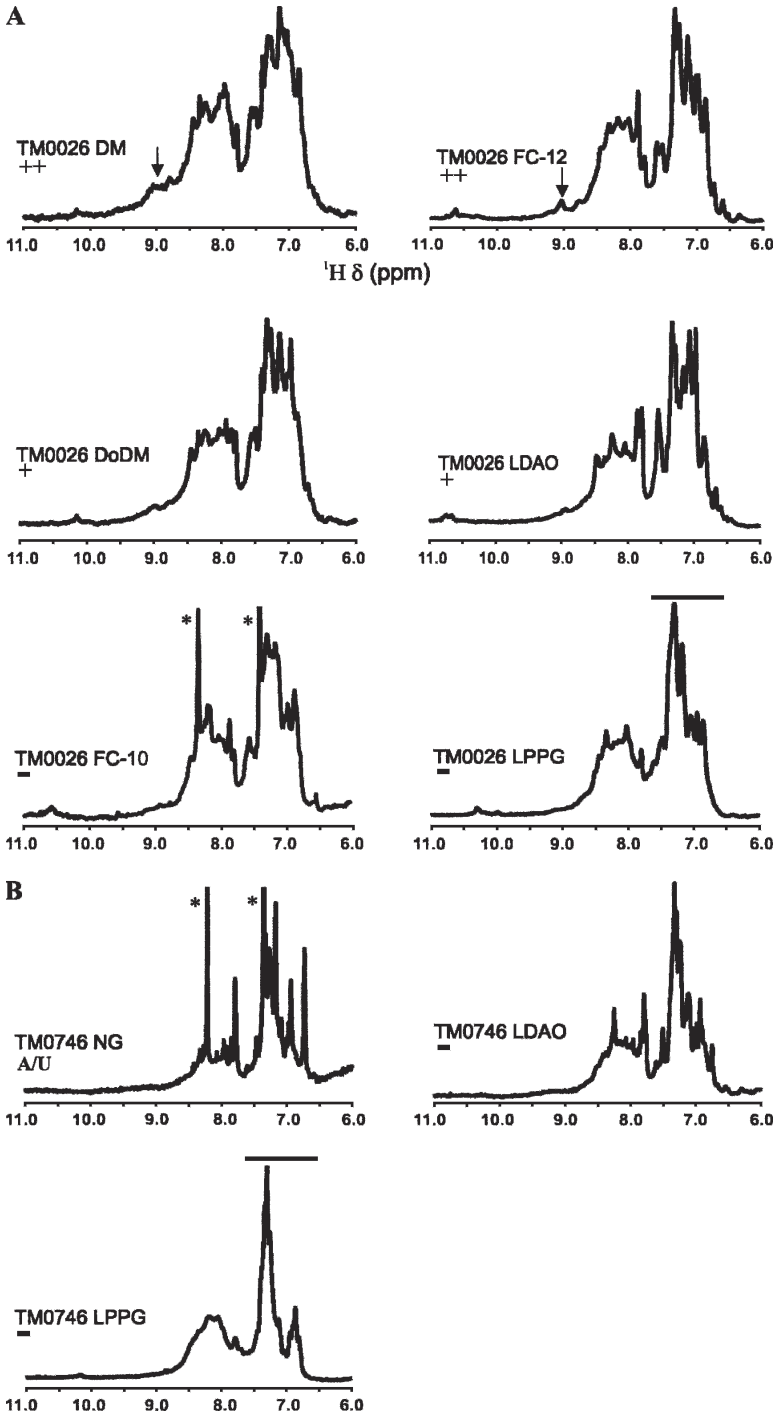

Figure 3. One-dimensional ^{1}H NMR spectra for the soluble conditions for TM0026 (*A*) and TM0746 (*B*). Solubilizing detergents are indicated. The asterisks (*) indicate signal from trace amounts of imidizole. The arrow points to amide proton dispersion indicative of secondary structure, and the bar indicates the aromatic region of the LPPG PDCs. The scores (++, +, −, A/U) correspond to those in Table 1.

scattering signal from the PDC and can accurately determine the PDC scattering intensity at low angles (the Guinier region). As detailed in Supplemental Material, the first approximation, I(complex − buffer), uses the scattering signal from the protein–detergent solution after subtraction of the buffer (no detergent) scattering as an estimate of the PDC scattering. This is an overestimate of the true PDC scattering, as the ''empty micelle'' also contributes to the

Downloaded from www.proteinscience.org on April 6, 2006

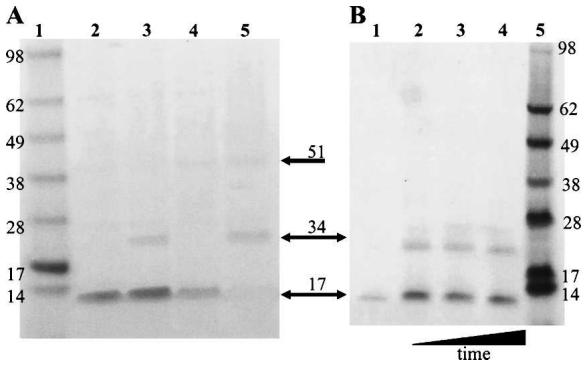

Figure 4. (*A*) Denaturing SDS-PAGE gel of TM1514 in DoDM, DHPC, LPPG, and FC-10 (lanes *2–4*, respectively). Lane *1* contains the molecular weight marker. Arrows point to observed oligomers. (*B*) Denaturing SDS-PAGE gel of TM1514 in DoDM after incubation with DSG for 0, 15, 30, and 60 min (lanes *1–4*, respectively). Lane *5* contains the molecular weight marker.

signal. The second approximation, I(complex − micelle), subtracts the micelle scattering signal recorded for the same concentration of detergent in the absence of protein from the scattering signal of the protein–detergent solution. This approximation is an underestimate of the true PDC scattering because in the presence of protein, a fraction of the detergent molecules associates into the PDCs, lowering the concentration of the ''empty'' micelles compared with the reference detergent solution. In practice, obtaining an upper and lower bound validates the obtained estimates, and, in most cases, the low angle scattering from the PDC can be determined with sufficient accuracy.

Scattering profiles for representative examples of TM0026 in FC-10, FC-12, and LPPG, as well as TM0994 in DoDM, are shown in Figure 5. One interesting feature noted for both the micelle and the PDC scattering is the observation of a strong second peak in the scattering profiles, which arises from the fact that the detergent alkyl chains have negative scattering contrast with respect to the solvent. The TM0026 data illustrate where the approximations work well. The micelle scattering is weakest for FC-12 and strongest for LPPG; however, the micelle scattering, even for LPPG, is well below the PDC estimates in the low s region used for the Guinier analysis. The insets show the data in Guinier representation (ln(I) as a function of s^2); black lines indicate the fitting range used for the R_g and I(0) estimates. The fourth example, TM0994 in DoDM, is representative of a case for which the approximate treatment fails. As can be seen from the scattering profiles (Fig. 6D), the buffer-subtracted scattering is stronger for the ''empty micelles'' as in the presence of the protein, which makes the I(complex − micelle) estimate unphysical (negative values, green solid curve). In this case, it is at best possible to check the PDC sample for the presence of large aggregates and to obtain a rough lower limit for the R_g of the PDC.

The R_g and I(0) values obtained for both approximations for the examples in Figure 5 are given in Table 4. For the weakest scattering micelle, FC-12, the two estimates give very similar values for both R_g and I(0); the R_g estimates deviate by ~1.5 Å, which is less than the experimental error. For FC-12 and LPPG, the R_g estimates exhibit a spread of ~5 Å, comparable with or slightly larger than the overall experimental error. The I(0) estimates show a similar trend. For FC-12, both approximations agree to within 5%, lower than the overall experimental error of ~10%. For FC-10 and LPPG, the spread is ~10% and 25%, respectively, still providing a meaningful estimate. For TM0994 in DoDM, the negative signal for I(complex − buffer) makes a physical estimate impossible; the R_g is at least as large as the apparent R_g from the I(complex − buffer) profile, but likely larger.

Scattering profiles were collected and analyzed for all 40 soluble PDCs. Similar to the case of the LDAO micelle, the five LDAO PDCs exhibited very strong interparticle interference, which prevented any estimation of the R_g and I(0). For TM0859 in DoDM and CHAPS, R_g and I(0) values were not obtained because the scattering signal was too low due to poor protein yields and/or the detergent micelle scattered too strongly. The R_g data for the remaining 33 soluble PDCs are presented in Table 3. The two different R_g estimates (using $I_{complex} - I_{buffer}$ and $I_{complex} - I_{micelle}$ methods) are tabulated, as well as a final R_g and error assessment. The two estimates typically bracket the true R_g value quite closely, with a spread comparable with (and in some cases even smaller than) the experimental error as determined from repeat measurements and Guinier fits with different fitting ranges. Similar to the example of TM0994 in DoDM in Figure 5 and Table 4, the strongly scattering micelles of DM, DoDM, and CHAPS caused fairly inaccurate R_g estimates with errors up to ±10 Å or in four instances only a lower bound of the R_g could be estimated.

The forward scattering intensity provides a model-free measurement of the total excess electron density of the PDC. The excess electron density can in principle be due to protein or detergent molecules, and a method for estimating the contributions of each is presented in the Supplemental Material. As a result, solutions satisfying the constraints imposed on the number of the detergent molecules in the PDC provide estimates for the oligomeric states of some of the PDCs, and the results are shown in Table 3.

2D ^{15}N, ^{1}H-correlation spectroscopy

The PDCs with high-quality ^1H NMR spectra (Table 2) are possible candidates for solution NMR structure determination. To further investigate the possibility of NMR structure determination of these PDCs, TM1634

Downloaded from www.proteinscience.org on April 6, 2006

Table 3. *Oligomeric states and radii of gyration of protein–detergent complexes*

		Oligomeric state[a]		R_g (Å)		
Protein	Detergent	Cross-linking	SAXS	I: Complex − Buffer	II: Complex − Micelle	Complex (error)
TM0026	DM	1	1	24	47	40 (±10)
TM0026	DoDM	1	1	46	46	46 (±3)
TM0026	FC-10	1	1	40	45	43 (±5)
TM0026	FC-12	1	1	46	47	46 (±3)
TM0026	LPPG	1	1	44	49	47 (±5)
TM0361	CHAPS	1	1	18	19	18 (±3)
TM0361	FC-10	1	1	43	44	43 (±3)
TM0361	LPPG	1	1	45	50	47 (±5)
TM0746	NG	1	1	51	51	51 (±3)
TM0746	CHAPS	1	1	28	48	45 (±10)
TM0746	LPPG	1	1	42	48	46 (±5)
TM0859	FC-10	1	1	51	53	52 (±5)
TM0859	FC-12	1	1	49	50	49 (±5)
TM0859	LPPG	1	1	53	58	55 (±5)
TM0994	OG	3	NR	37	34	35 (±3)
TM0994	DoDM	ND	NR	45	51	48 (±5)[b]
TM0994	LPPG	ND	NR	45	50	47 (±5)
TM1402	DM	ND	NR	20	[c]	>20
TM1402	DoDM	ND	NR	27	[c]	>28
TM1402	CHAPS	ND	NR	32	[c]	>33
TM1402	FC-12	ND	2	48	51	48 (±5)
TM1402	LPPG	ND	1	46	59	52 (±8)
TM1514	DoDM	2	2	32	86	>32
TM1514	DHPC	2	1/2	41	49	46 (±5)
TM1514	FC-10	2	1/2	38	44	42 (±5)
TM1514	LPPG	2	1/2	50	54	52 (±5)
TM1634	OG	2	2	37	38	38 (±3)
TM1634	NG	2	2	51	48	49 (±5)
TM1634	DM	2	2	30	47	40 (±10)
TM1634	DoDM	2	2	65	65	65 (±5)
TM1634	FC-10	1	1/2	44	48	47 (±5)
TM1634	DHPC	1	1/2	40	49	46 (±5)
TM1634	LPPG	ND	1	50	53	52 (±3)

[a] 1, Monomer; 2, dimer; ½, monomer or dimer; ND, data were not determined because optimal cross-linking conditions could not be found (i.e., a predominant species could not be determined depending on the reagent, the pH, and the time of incubation); NR, the oligomeric state was not determined.
[b] These values were obtained with a lower DoDM concentration (5 mM) compared with the experiment in Figure 7 (120 mM).
[c] $I(0)_{micelle} > I(0)_{complex}$ and the R_g could not be estimated.

was ^{15}N-labeled and TM0026 was ^{15}N, ^{2}H-labeled and ^{15}N, ^{1}H-TROSY spectra were recorded (Fig. 6). Each spectrum should have a cross-peak for each amide nitrogen–proton pair (except for proline residues) and each tryptophan indole nitrogen–proton pair. In addition to the number of peaks, the spectra should lack significant line broadening indicative of conformational exchange, and the peak intensities should be homogenous.

Based on the evaluation in Table 2, TM0026 in DM and FC-12 yielded very promising results for further structural studies. The ^{15}N, ^{1}H-TROSY spectra support this conclusion. For TM0026 in DM, 76 of the 78 expected cross-peaks were observed, the peak intensities were homogenous, and line broadening was observed only for a select few resonances. For FC-12, 74 of the 78 expected peaks were observed; however, the peak intensities are not as homogenous as for the TM0026 DM PDC. The overall pattern of cross-peaks were similar between DM and FC-12; however, some chemical shift differences were evident. For a comparison the ^{15}N, ^{1}H-TROSY spectra were recorded for the less favorable conditions FC-10 and DoDM. In both detergents, there were far less than the expected number of cross-peaks and significant line broadening was observed, consistent with the evaluations based on the 1D ^{1}H NMR spectra.

TM1634 had the most soluble detergent conditions; however, only the most promising detergent condition was investigated for comparison with the 1D ^{1}H NMR spectrum. TM1634 in LDAO did not need to be deuterated, and the ^{15}N, ^{1}H-TROSY spectrum is shown in Figure 6. The peak intensities are homogenous, and 110 of 140 expected cross-peaks were observed.

None of the 1D ^{1}H NMR spectra indicated promising detergent conditions for TM1514, predominantly because

Downloaded from www.proteinscience.org on April 6, 2006

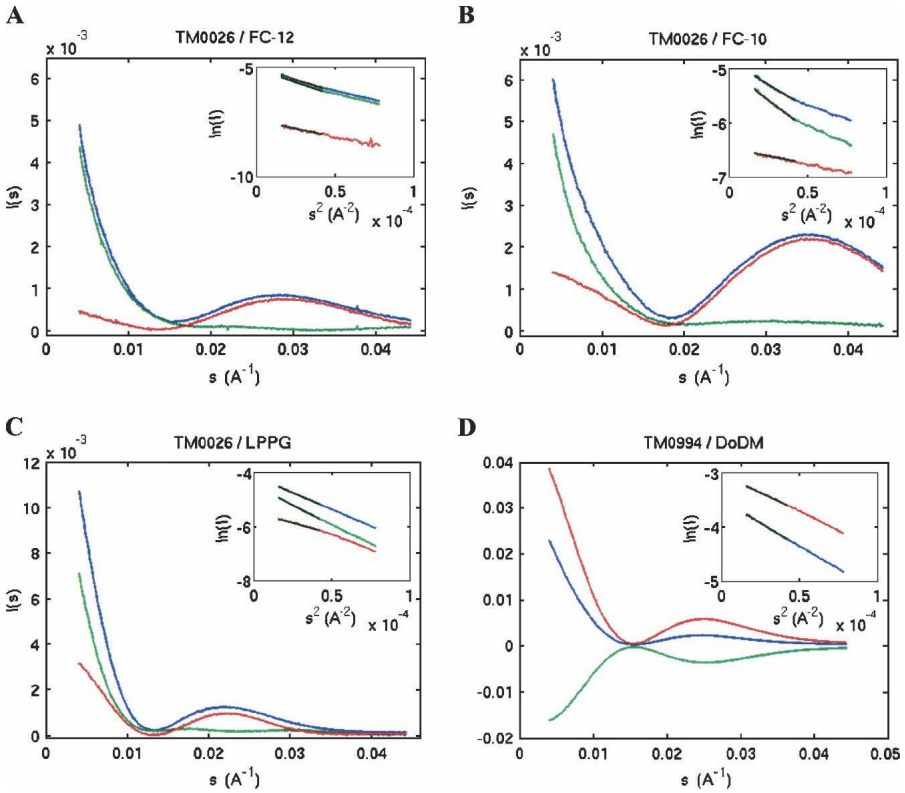

Figure 5. SAXS scattering profiles and Guinier analysis. Approximations to the PDC scattering profile I(complex − buffer) (blue) and I(complex − micelle) (green) and scattering profile of the detergent micelles in the absence of protein I(micelle) (red). (*Insets*) Guinier plots of ln(I) vs. s^2 for the low s region (same color code as main figures) and the linear fits used to obtain the forward scattering intensities and radii of gyration (black solid lines). Data shown for 0.33 mM TM0026 in 65 mM FC-12 (*A*), 0.47 mM TM0026 in 277 mM FC-10 (*B*), 0.26 mM TM0026 in 17 mM LPPG (*C*), and 0.3 mM TM0994 in 116 mM DoDM (*D*).

of line broadening. To confirm the interpretation of the 1D ^{1}H spectra was likely due to the molecular weight of the protein dimer (as determined by SAXS and chemical cross-linking), the protein was ^{15}N- or ^{15}N, ^{2}H-labeled and ^{15}N, ^{1}H-TROSY spectra were recorded in each detergent condition. Without deuteration only two or three sharp cross-peaks were observed; however, with ~75% deuteration of the protein, the best spectrum obtained is shown in Figure 6. Significant line broadening and 82 of 145 expected peaks were observed. These results are in agreement with the assessment based on the 1D ^{1}H NMR spectra that none of the detergent conditions were appropriate for NMR structure determination for this target most likely due to the molecular weight of the PDC. Therefore, TM1514 in DoDM seemed to be a candidate for X-ray crystallographic studies and initial crystal leads exhibit diffraction to better than 6–7 Å (data not shown).

3D ^{15}N-resolved NOESY spectroscopy

In addition to structure determination, NMR spectroscopy provides a means to investigate interactions between the

Table 4. *SAXS data obtained from Guinier fits shown in Figure 5*

					I(0) × 10^{-3}			R_g (Å)			
Protein	Detergent	c_{prot} (mM)	$c_{detergent}$ (mM)	c_{mic} (mM)[a]	Micelle	Comp − Buffer	Comp − Micelle	Micelle	Comp − Buffer	Comp − Micelle	Estimated (error)
TM0026	FC-10	0.47	277	6.9	1.4	8.0	7.0	21	40	45	43 (±5)
TM0026	FC-12	0.33	64	0.80	0.6	7.5	7.1	34	46	47	46 (±3)
TM0026	LPPG	0.26	17	0.17	4.5	16	12	35	44	49	47 (±5)
TM0994	DoDM	0.3	116	1.18	49	31	[b]	32	38	[b]	>38

[a] Micelle concentration computed as $C_{mic} = C_{det} / m_{mic}$ with m_{mic} from Table 1 or values obtained from SAXS data.
[b] Could not be computed because the free micelle scatters more strongly than the complex.

Downloaded from www.proteinscience.org on April 6, 2006

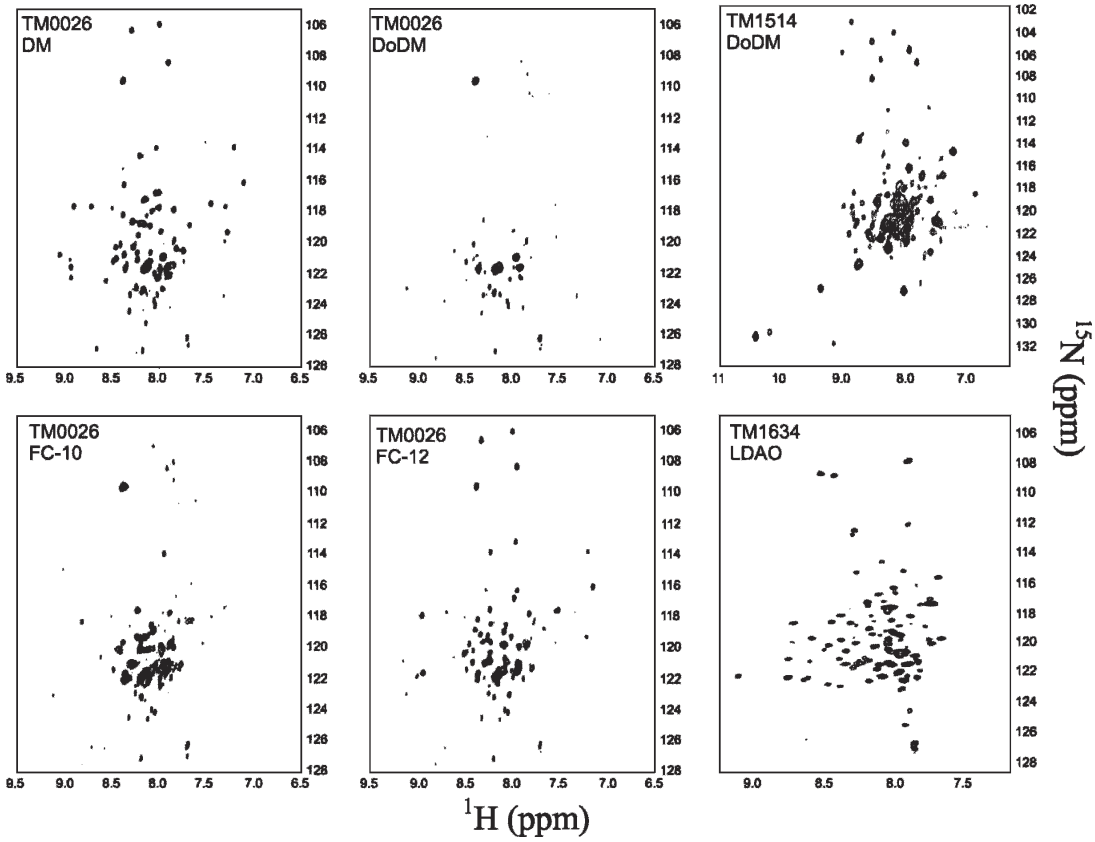

Figure 6. Two-dimensional ^{15}N, ^{1}H-TROSY spectra of TM0026 in DM, DoDM, FC-10, and FC-12, TM1514 in DoDM, and TM1634 in LDAO. For TM0026, the two indole cross-peaks are not shown in order to better view the backbone resonances.

detergent micelle and the protein (Fernandez et al. 2002). Nuclear Overhauser effects (NOEs) between the membrane protein and the detergent molecules provide a detailed description of protein–detergent interactions. In particular, the NOEs between the hydrophobic alkyl chain and the backbone amide proton indicate the protein is folded within the interior of the micelle. Selected strips from a ^{15}N-resolved NOESY spectrum of TM0026 in DM are shown in Figure 7. The 1D ^{1}H NMR spectrum of the detergent is aligned on the right with the resonances from the alkyl chain protons assigned to the structure of DM. NOEs between 25 protein amide protons and detergent protons in the hydrophobic tail are observed. Without a backbone assignment, the identity of the amino acids that are in the micelle interior is not known. However, the knowledge that a significant portion of the protein (~35%) is located in the interior of the micelle is indicative that the protein is properly encapsulated in the micelle rather than solubilized in an alternative conformation, such as on the surface of the micelle. The extent of deuteration is ~75%, and spin diffusion is not likely to be a major factor even at a mixing time of 200 msec.

Discussion

Expression and localization

To date, three published studies reported high-throughput expression screening of bacterial membrane proteins in *E. coli* (Dobrovetsky et al. 2005; Eshaghi et al. 2005; Korepanova et al. 2005). The reported successes in expression vary from 30% to 71%; however, several technical aspects need to be considered when comparing the numbers from each of these studies: (1) the number of constructs (i.e., C-terminal and N-terminal His-tags, different promoters, or different fusion proteins) and the number of *E. coli* cell lines attempted for each protein; (2) the protein detection method used to assay expression (Coomassie or Western blot); and (3) the localization of the expressing protein (i.e., the insoluble or the membrane fraction). In this study, 22% of the proteins overexpressed as determined with Coomassie staining. Eshaghi et al. (2005) and Dobrovetsky et al. (2005) reported 71% and 30%, respectively, of protein targets expressed in *E. coli* as assayed with Western blot analysis. The high expression success of Eshaghi et al. (2005) is likely

Downloaded from www.proteinscience.org on April 6, 2006

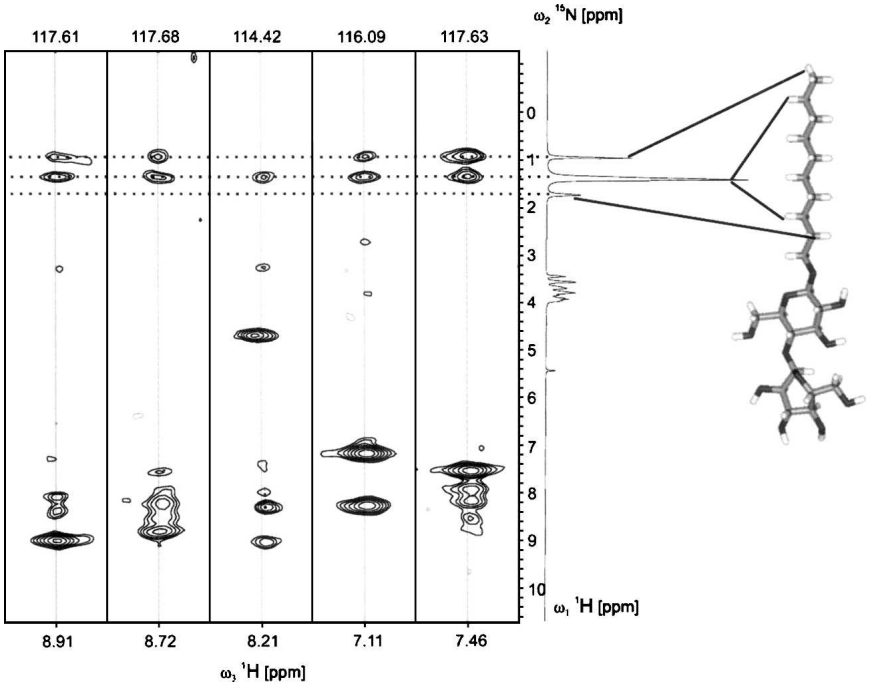

Figure 7. Selection of $\omega_1(^1H)/\omega_3(^1H)$ strips from a 3D ^{15}N-resolved [^1H,^1H]-NOESY spectrum. The strips were taken at the ^{15}N chemical shifts of the residues indicated at the *top* and are centered about the respective amide proton chemical shifts. The 1D ^1H NMR spectrum of DM, measured with the same sample and the same experimental conditions, is aligned to the *right* of the NOESY strips, and the chemical structure of DM is shown. The detergent resonances of interest are assigned to the detergent protons using lines.

attributed to the use of five different fusion tags for each protein and three different cell lines. In addition, these studies utilized Western blot analysis for evaluation of expression, which also detects protein levels below the Coomassie detection method used in this study. For comparison, Korepanova et al. (2005) compared the staining methods and found 25% of the proteins they investigated overexpressed, as assayed with Coomassie, and 50% expressed, if detected by Western blot analysis.

Only one of the three published studies separated the insoluble fraction from the membrane fraction. Korepanova et al. (2005) reported that almost all of the protein targets were detected in the insoluble fraction, 9% of which also overexpressed to the membrane. Our results are similar in that all of the proteins expressed to the membrane were also detected in the insoluble fraction; however, 18% overexpressed to the membrane. These data suggest the importance of isolating the insoluble fractions (presumably unfolded proteins) before solubilizing the membrane fraction (properly folded proteins), if a refolding protocol is not implemented. Establishing membrane localization in the case of the abundant hypothetical putative membrane proteins (~55% of *T. maritima* membrane proteins, including the eight proteins in this study, are of unknown function) is important because the association to the membrane is predicted based on hydrophobicity, charge bias, and helical length, and needs to be confirmed with experimental evidence.

Solubility

High-resolution membrane protein structures have been determined in a select few detergents. These include OG, DM, DoDM, and LDAO for crystallography (Wiener 2004) and DHPC (Fernandez et al. 2004), FC-12 (Hwang et al. 2002; Oxenoid and Chou 2005), and OG (Hwang et al. 2002) for solution NMR spectroscopy. One study implicated LPPG (Krueger-Koplin et al. 2004) as a potential detergent for solution NMR spectroscopy structure determination, although a structure has not been reported in this detergent. In this work, glucosides, maltosides, DHPC, CHAPS, LDAO, phosphocholines, and LPPG were surveyed for solubility of the eight overexpressing membrane proteins. In terms of solubility, LPPG and DoDM were the most successful; in fact, all eight proteins are soluble in LPPG (Table 2). A general trend in the chain length and the ability to solubilize was observed; the detergents with the longer alkyl chains solubilized more proteins than those with shorter chains. The hydrophobic effect is stronger for the longer chain detergents, thereby providing a closer mimic to the lipid bilayer, which is likely to increase the solubility and stability of

Downloaded from www.proteinscience.org on April 6, 2006

the PDC. Several studies found similar results and have directly measured increases in stability and activity of membrane proteins in detergents with longer (>10 carbons) alkyl chains (Knudsen and Hubbell 1978; Vinogradova et al. 1998). However, detergents that are too long may alter helix–helix interactions due to the protein adapting to the size and shape of the micelle (Therien and Deber 2002). In addition to chain length, the size of the head groups must also play a role in solubility influenced by the drastic results between DG and DM.

Overall fold, size, and oligomerization state of PDCs

The CD (data shown in Supplemental Material) and 1D ^{1}H NMR spectroscopy indicated that all the proteins have helical structures in the micelles in which they are soluble. Although complicated by the detergent proton resonances, the 1D ^{1}H spectrum of PDCs was useful in evaluating the PDCs for NMR structure determination. In the case of TM0026 and TM1634, the PDCs that yielded high-quality ^{15}N, ^{1}H-TROSY spectra were in fact the conditions that had the high-quality 1D ^{1}H NMR spectra. Although the oligomerization states are estimates, in most cases the SAXS and the cross-linking data agree. Both techniques suggest that TM1634, TM1514, and TM0026 are dimeric and TM0361, TM0746, and TM0859 are monomeric. The SAXS data suggest that the oligomers are destabilized by LPPG or stabilized by the other detergents.

The radius of gyration measured with SAXS is similar to the hydrodynamic radius of gyration and is a global measure of the size and shape of the molecular complex. The difference is that the R_g from SAXS measurements is electron density contrast weighted and the hydrodynamic R_g is mass weighted; however, both measurements are correlated to the diffusion properties of the molecular complex. It is interesting to compare the R_g values between different PDCs to attempt to understand the dominating factors that dictate the size of the PDC and, hence, the stoichiometry of detergent to protein. For instance, TM0026 has a relatively constant R_g value between the five different detergents with a range of 40–47 Å. Conversely, TM1634 has a much larger range of R_g values between 38 and 65 Å. Since the protein is a dimer in all detergents (except LPPG), these differences must be due to the detergent. The data suggest that properties of both the detergent and the protein dictate the size and shape of the PDC. The R_g values do not correlate with molecular weight or the number of transmembrane helices (i.e., the hydrophobic surface area). However, a slight bias observed throughout the data is that the detergents that form larger micelles also form larger PDCs. A more extensive data set including larger proteins may provide correlations between the detergent and the protein properties that mediate the formation of the PDC. In the future, to investigate the ''matching'' of detergents to proteins, SAXS techniques and 3D NMR ^{13}C and ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectroscopy could be combined to determine the protein oligomeric state, the number of detergent molecules, and the specific interactions between the proteins and the detergent in the PDC. Currently, a method is under development to obtain a more accurate estimate of the PDC scattering profile over the entire measured angle range from a singular value decomposition (SVD) (Doniach 2001) of scattering data collected at different protein and detergent stoichiometries. An accurate estimate of the full scattering profile allows the possibility to construct and test molecular models of the PDC based on SAXS alone, or in combination with NMR data similar to that presented in Figure 7. However, as these SVD studies require several measurements and modeling for each protein–detergent combination, they are less well suited for the high-throughput screening approach presented in this study, which only requires a single measurement of the PDC.

Line broadening is a significant impediment to NMR structure determination, and distinguishing between molecular weight (e.g., the overall correlation time) and conformational exchange (microsecond to millisecond protein fluctuations between two or more conformations) contributions to the line width is important, especially for PDCs for which the membrane proteins are solubilized in unnatural conditions to mimic the membrane environment and have unknown oligomeric states. Different micellar environments may modulate protein dynamics differently. For example, if the helix–helix packing were destabilized in one micelle versus another, the dynamics of the helices would be affected differently in each micelle. Hence, the different dynamics could result in line broadening from conformational exchange in one micelle but not be observed in another. However, line broadening could also be due to a high molecular weight of the PDC either from the detergent micelle, protein oligomerization, or both. The line widths in the 1D ^{1}H and 2D ^{15}N, ^{1}H-TROSY NMR spectra of TM0026 in FC-10 and DoDM were significantly broader than those observed in DM and FC-12. The SAXS data and cross-linking indicate that TM0026 is a monomer in all detergent conditions, implying that the spectral differences are not due to the oligomerization of the protein. Therefore, the line broadening observed for the FC-10 and DoDM PDCs is either due to differences in the micelle size or from conformational exchange. The SAXS data provide an estimate of the size and shape of the PDC and, in the case of TM0026, the R_g is similar for all detergent conditions at $\approx$46 Å. Therefore, the likely explanation is that conformational exchange occurs in FC-10 and DoDM that is not present in FC-12 and DM. The explanation for this result is not the alkyl chain length, because they are identical for each pair FC-10/DM and

Downloaded from www.proteinscience.org on April 6, 2006

FC-12/DoDM. However, the aggregation number differs between the detergents. The two detergent conditions that lack significant line broadening have aggregation numbers of ~70 detergent molecules; however, DoDM has ~135 and FC-10 has ~40 detergent molecules (data not shown). It is possible that the packing density of the detergents is important due to steric constraints of the hydrophobic surface.

Choosing detergents for solution NMR studies of membrane proteins

The extensive screen presented here indicates that there is not one detergent that is well suited for NMR studies of membrane proteins; solubility, dynamics, the hydrophobic surface area of the protein, and other physical properties differ for each PDC, and the proper combination still needs to be empirically determined. For instance, in this detergent screen, the longer chain lipids solubilize well but do not necessarily correlate with quality samples for NMR structure determination, as seen for LPPG. This is likely due to the rather large micelle and the resulting large complex as indicated by the measured R_g. Another interesting observation is the lack of aromatic proton dispersion (compared with other conditions) in LPPG PDCs, as shown in Figure 3, A and B, with a horizontal bar. The lack of dispersion indicates the aromatic side chains are in similar environments, which is not usually observed in NMR spectra of folded proteins.

Conclusion

This study emphasizes the importance of screening a variety of detergents and characterizing protein–detergent complexes with a battery of techniques for successful structure determination. We have established SAXS and NMR spectroscopy as valuable screening tools in evaluating PDCs for successful structural investigations. A method of estimating the scattering profile of the PDC in the low angle region from a single measurement of the protein is presented in order to obtain a reasonably accurate measurement of the radius of gyration and protein oligomerization state (via the forward scattering intensity) for most protein–detergent combinations. NMR methods traditionally used to evaluate the fold of soluble proteins were shown to be applicable to membrane proteins in detergent micelles. Cumulatively, the data acquired have provided solution conditions for TM0026, TM1514, and TM1634, in which the proteins are folded and the PDC is stable. Based on the quality of the 2D ^{15}N, ^{1}H-TROSY NMR spectra, TM0026 in DM and TM1634 in LDAO are promising candidates for NMR structure determination. In addition, the physical characterization and the preliminary crystals of TM1514 in DoDM suggest the PDC is suitable for crystallographic studies.

Materials and methods

Cloning and expression

Genes representing 97% of the *T. maritima* genome were cloned into expression vectors as an effort of The Joint Center for Structural Genomics (JCSG) (Lesley et al. 2002). These vectors encode a purification tag (MGSDKIHHHHHH) at the N terminus of the full-length protein. Protein expression was performed with TB media (24 g/L yeast extract and 12 g/L tryptone) containing 1% glycerol (v/v), 50 mM MOPS (pH 7.6), and 100 μg/mL ampicillin in the 65 mL/96 cultures GNFermentor (DiDonato et al. 2004). Expression was induced by the addition of 0.20% arabinose for 3 h. For uniform ^{15}N-labeling, Celtone media (Spectra Stable Isotopes) supplemented with 50 mM MOPS (pH 7.6), 50 mM NaCl, 1% glycerol, and 100 μg/mL ampicillin was used. For deuterated, ^{15}N-labeled proteins, published protocols using conventional shakers and minimal media in D_2O supplemented with $^{15}NH_4Cl$ were used. The GNFermentor could not be used because evaporation was observed for the lengthy times required for growth and expression (~2 d). Fresh transformations were required for efficient and reproducible expression. Expression was assayed on an SDS-PAGE gel (4–20%, Invitrogen) with an addition of 8 M urea to the standard loading buffer.

Localization

Bacteria were lysed by sonication in lysis buffer (50 mM Tris at pH 8.0, 100 mM NaCl, 1 Complete protease inhibitor pellet [Roche]), and cell debris was pelleted by low-speed centrifugation at 15,000*g* for 30 min. The insoluble pellet was washed with lysate buffer containing 0.4% n-decyl-β-D-maltoside (DM, Anatrace, Inc) and centrifuged again at 15,000*g* for 30 min. The pellet was additionally washed with lysate buffer to remove excess detergent and the remaining pellet was resuspended in lysate buffer. The soluble fraction from the first low-speed spin containing the membrane was centrifuged at 100,000*g* for 1 h. The membrane pellet was resuspended in lysis buffer and centrifuged again at 100,000*g* for 1 h. The remaining pellet was resuspended in lysate buffer with 0.4% DM and incubated at room temperature for 3 h. The insoluble, supernatant, and membrane fractions were analyzed by SDS-PAGE gel and the localization of the protein was determined.

Protein purification

Bacteria were lysed by sonication in lysis buffer (50 mM Tris at pH 8.0, 100 mM NaCl, 1 Complete protease inhibitor pellet [Roche]), and cell debris was removed by low-speed centrifugation at 15,000*g* for 30 min. DM (0.4%) was added directly to the supernatant containing the cell membranes and incubated for 3 h. The solubilized protein was purified using standard cobalt-chelating chromotography (GE Healthcare). The wash buffer (25 mM sodium phosphate at pH 7.8, 100 mM NaCl, 25 mM imidazole) and elution buffer (wash buffer with 600 mM imidazole) each contained one of 11 detergents at a concentration above the reported critical micelle concentration (CMC) and with the micelle concentration approximately twofold greater than the protein concentration. DHPC and LPPG were purchased from Avanti Polar Lipids, and the remaining detergents were purchased from Anatrace. The elution was concentrated to 600 μL, and the imidazole was removed by dialysis (20 mM

Downloaded from www.proteinscience.org on April 6, 2006

sodium phosphate at pH 7.0, 150 mM NaCl). For detergents with high CMCs (>0.5%), the protein solution was dialyzed against 2 × 4L for 30 min for each 4 L dialysis. For detergents with low CMCs, the protein solution was dialyzed against 3 × 4 L for 1 h for each 4 L dialysis. Table 1 lists the CMC, the micelle size, and the ionic property of the detergents used.

Detergent concentrations of the final sample were measured with 1D ^{1}H NMR spectroscopy by comparing the spectra of the sample with spectra of solutions with known detergent concentrations. The measured detergent concentrations for the PDC solutions were used to prepare the appropriate solution for the SAXS measurements in which the scattering of the micelle solutions are subtracted for the protein–detergent solution. Protein concentrations were measured from the absorbance at 280 nm in 6 M guanidine hydrochloride and with the BCA protein assay (Pierce).

Detergent screen

After dialysis, the protein solutions were incubated at room temperature for 5 d (selection of this time period was dictated by the typical time required for recording triple resonance experiments). The solutions were spun at 18,000*g* for 30 min, and both the pellet and supernatant were examined with SDS-PAGE gel. Samples were never heated before loading on SDS-PAGE gels. If the protein was in the supernatant, then 5% D_2O was added and 1D ^{1}H NMR spectra were recorded.

Chemical cross-linking

Soluble PDC complexes were reacted with the hydrophobic cross-linker DSG and/or with a hydrophilic cross-linker EDC. Protocols from Pierce were followed using the following reaction conditions: 20 mM phosphate buffer (pH 7), 150 mM NaCl, 0.6 mM DSG, and ~10 μM protein for ~15 min or 50 mM HEPES (pH 8), 150 mM NaCl, 2.5 mM EDC, and ~10 μM protein for ~90 min. Each reaction was evaluated with SDS-PAGE denaturing gels, and the oligomeric state was estimated based on the migration of the reacted protein compared with unreacted protein. Reaction times and protein concentrations were optimized for each PDC.

SAXS

All data were recorded on beam line BESSERC CAT 12-ID at the Advanced Photon Source, employing a 2 m camera and a CCD detector at a photon energy of 12 keV. The measurements were performed at 25 ± 1°C using a custom-made thermostated cell with a sample volume of 16 μL. For each data point, a total of five measurements of 0.1 sec integration time were recorded. Data were image-corrected and circularly averaged; the five profiles for each condition were averaged to improve signal quality. We tested for possible radiation damage by comparing subsequent exposures of the same sample, and no change was detected. A theoretical treatment of the SAXS methods is provided in the Supplemental Material.

NMR spectroscopy

All NMR data were recorded at 313 K on a 600 MHz Bruker Avance spectrometer. The 1D ^{1}H NMR spectra were recorded with the Watergate W5 pulse sequence (Liu et al. 1998), 2048 accumulated transients, and a time domain data size of 4096 complex points. The large number of scans needed was in part due to the strong intensity of the detergent, which required the receiver gain to be low (~32). The 2D ^{15}N, ^{1}H-TROSY spectra were recorded with 128 transients per increment, $t_{1max}(^{15}N)$ = 42 msec, $t_{2max}(^{1}H)$ = 285 msec, and a time domain data size of $64(t_1) \times 2048(t_2)$ complex points. 3D ^{15}N-resolved [^{1}H, ^{1}H]-NOESY were recorded with a time domain data size of $64(t_1) \times 16(t_2) \times 1024(t_3)$ complex points, $t_{1max}(^{1}H)$ = 8.9 msec, $t_{2max}(^{15}N)$ = 10.5 msec, $t_{3max}(^{1}H)$ = 143 msec, and τ_m = 200 msec.

Acknowledgments

We thank Dr. Kurt Wüthrich for support, Dr. Peter Wright for the use of the CD spectrometer, Drs. Gerard Kroon and Bernhard Geierstanger for helpful discussion and technical NMR assistance, Dr. Andreas Kreusch for crystallographic work, and Michael Geralt for technical laboratory support. We thank the DOE for beam time on the BESSRC-CAT beam line 12-ID at the APS, Argonne National Lab. We also thank Sönke Seifert and Vincent Chu for help with the data collection at the APS. Support for this research was provided by the NIH Protein Structure Initiative, grants P50 GM62411 and U54 GM074898 (S.L.), NSF award no. PHY-0140140 (S.D.) and NIH grant 1F32GM068286 (L.C.).

References

Bendedouch, D., Chen, S.H., and Koehler, W.C. 1983. Determination of interparticle structure factors in ionic micellar solutions by small-angle neutron scattering. *J. Chem. Phys.* **87:** 2621–2628.

Bezzobotnov, V.Y., Borbely, S., Cser, L., Farago, B., Gladkih, I.A., Ostanevich, Y.M., and Vass, S. 1988. Temperature and concentration dependence of properties of sodium dodecyl sulfate micelles determined from small-angle neutron scattering experiments. *J. Phys. Chem.* **92:** 5738–5743.

Bowie, J.U. 2001. Stabilizing membrane proteins. *Curr. Opin. Struct Biol.* **11:** 397–402.

Bu, Z. and Engelman, D.M. 1999. A method for determining transmembrane helix association and orientation in detergent micelles using small angle X-ray scattering. *Biophys. J.* **77:** 1064–1073.

Burns, R.A., Roberts, M.F., Dluhy, R., and Mendelsohn, R. 1982. Monomer-to-micelle transition of dihexanoylphosphatidylcholine: ^{13}C NMR and Raman studies. *J. Am. Chem. Soc.* **104:** 430–438.

Chattopadhyay, A. and London, E. 1984. Fluorimetric determination of critical micelle concentration avoiding interference from detergent charge. *Anal. Biochem.* **139:** 408–412.

Chen, Y., Song, J., Sui, S.F., and Wang, D.N. 2003. DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in *Escherichia coli*. *Protein Expr. Purif.* **32:** 221–231.

Chou, J.J., Baber, J.L., and Bax, A. 2004. Characterization of phospholipid mixed micelles by translation diffusion. *J. Biomol. NMR* **29:** 299–308.

DeGrip, W.J. and Bovee-Geurts, P.H.M. 1979. Synthesis and properties of alkylglucoside with mild detergent action-improved synthesis and purification of β-1-octyl-glucose, β-1-nonyl-glucose and β-1-decyl-glucose-synthesis of β-1-undecylglucose and β-1-dodecymaltose. *Chem. Phys. Lipids* **23:** 321–335.

DiDonato, M., Deacon, A.M., Klock, H.E., McMullan, D., and Lesley, S.A. 2004. A scaleable and integrated crystallization pipeline applied to mining the *Thermotoga maritima* proteome. *J. Struct. Funct. Genomics* **5:** 133–146.

Dobrovetsky, E., Lu, M.L., Andorn-Broza, R., Khutoreskaya, G., Bray, J.E., Savchenko, A., Arrowsmith, C.H., Edwards, A.M., and Koth, C.M. 2005. High-throughput production of prokaryotic membrane proteins. *J. Struct. Funct. Genomics* **6:** 33–50.

Doniach, S. 2001. Changes in biomolecular conformation seen by small angle X-ray scattering. *Chem. Rev.* **101:** 1763–1778.

Dupuy, C., Auvray, X., and Petipas, C. 1997. Anomeric effects on the structure of micelles of alkyl maltosides in water. *Langmuir* **13:** 3965–3967.

Eshaghi, S., Hedren, M., Nasser, M.I., Hammarberg, T., Thornell, A., and Nordlund, P. 2005. An efficient strategy for high-throughput expression

Downloaded from www.proteinscience.org on April 6, 2006

screening of recombinant integral membrane proteins. *Protein Sci.* **14:** 676–683.

Fernandez, C., Hilty, C., Wider, G., and Wüthrich, K. 2002. Lipid-protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. *Proc. Natl. Acad. Sci.* **99:** 13533–13537.

Fernandez, C., Hilty, C., Wider, G., Guntert, P., and Wüthrich, K. 2004. NMR structure of the integral membrane protein OmpX. *J. Mol. Biol.* **336:** 1211–1221.

Fernando, S.A., Selvarani, P., Das, S., Kumar, C.K., Mondal, S., Ramakumar, S., and Sekar, K. 2004. THGS: A web-based database of transmembrane helices in genome sequences. *Nucleic Acids Res.* **32:** D125–D128.

Folkers, G.E., van Buuren, B.N., and Kaptein, R. 2004. Expression screening, protein purification and NMR analysis of human protein domains for structural genomics. *J. Struct. Funct. Genomics* **5:** 119–131.

Hayter, J.B. and Zem, T. 1982. Concentration-dependent structure of sodium octanoate micelles. *Chem. Phys. Lett.* **93:** 91–94.

He, L., Garamus, V.M., Funari, S.S., Malfois, M., Willumeit, R., and Niemeyer, B. 2002. Comparison of small-angle scattering methods for the structural analysis of octyl-β-maltopyranoside micelles. *J. Phys. Chem. B* **106:** 7596–7604.

Helenius, A., McCaslin, D.R., Fries, E., and Tanford, C. 1979. Properties of detergents. *Methods Enzymol.* **56:** 734–749.

Herrmann, K.W. 1962. Non-ionic-cationic micelar properties of dimethydodecyamine oxide. *J. Phys. Chem.* **66:** 295–300.

———. 1966. Micellar properties of some zwiterionic surfactants. *J. Colloid Interface Sci.* **22:** 352.

Hjelmeland, L.M., Nebert, D.W., and Osborne, J.C. 1983. Sulfobetaine derivatives of bile-acids: Nondenaturing surfactants for membrane biochemistry. *Anal. Biochem.* **130:** 72–82.

Hwang, P.M., Choy, W.Y., Lo, E.I., Chen, L., Forman-Kay, J.D., Raetz, C.R., Prive, G.G., Bishop, R.E., and Kay, L.E. 2002. Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci.* **99:** 13560–13565.

Kiefer, H. 2003. In vitro folding of α-helical membrane proteins. *Biochim. Biophys. Acta* **1610:** 57–62.

Knudsen, P. and Hubbell, W.L. 1978. Stability of rhodopsin in detergent solutions. *Membr. Biochem.* **1:** 297–322.

Korepanova, A., Gao, F.P., Hua, Y., Qin, H., Nakamoto, R.K., and Cross, T.A. 2005. Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli*. *Protein Sci.* **14:** 148–158.

Krueger-Koplin, R.D., Sorgen, P.L., Krueger-Koplin, S.T., Rivera-Torres, I.O., Cahill, S.M., Hicks, D.B., Grinius, L., Krulwich, T.A., and Girvin, M.E. 2004. An evaluation of detergents for NMR structural studies of membrane proteins. *J. Biomol. NMR* **28:** 43–57.

le Maire, M., Champeil, P., and Moller, J.V. 2000. Interaction of membrane proteins with solubilizing detergents. *Biochim. Biophys. Acta* **1508:** 86–111.

Lesley, S.A., Kuhn, P., Godzik, A., Deacon, A.M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H.E., McMullan, D., Shin, T., et al. 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proc. Natl. Acad. Sci.* **99:** 11664–11669.

Lin, T.-L., Chen, S.-H., Gabriel, N.E., and Roberts, M.F. 1986. Use of small-angle neutron scattering to determine the structure and interaction of dihexanoylphosphatidylcholine micelles. *J. Am. Chem. Soc.* **108:** 3499–3507.

Liu, M., Mao, X., He, C., Huang, H., Nicholson, J.K., and Lindon, J.C. 1998. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *J. Magn. Reson.* **132:** 125–129.

Loll, P.J., Allaman, M., and Wiencek, J. 2001. Assessing the role of detergent-detergent interactions in membrane protein crystallization. *J. Cryst. Growth* **232:** 432–438.

Lorber, B., Bishop, J.B., and Delucas, L.J. 1990. Purification of octyl β-glucopyranoside and reestimation of its micellar size. *Biochim. Biophys. Acta* **1023:** 254–265.

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.

Nilsson, F., Södermann, O., and Hansson, P. 1998. Physical-chemical properties of C_9G_1 and $C_{10}G_1$ β-alkylglucosides. Phase diagrams and aggregate size/structure. *Langmuir* **14:** 4050–4058.

Oxenoid, K. and Chou, J.J. 2005. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl. Acad. Sci.* **102:** 10870–10875.

Peti, W., Etezady-Esfarjani, T., Herrmann, T., Klock, H.E., Lesley, S.A., and Wüthrich, K. 2004. NMR for structural proteomics of *Thermotoga maritima:* Screening and structure determination. *J. Struct. Funct. Genomics* **5:** 205–215.

Rehm, T., Huber, R., and Holak, T.A. 2002. Application of NMR in structural proteomics: Screening for proteins amenable to structural analysis. *Structure* **10:** 1613–1618.

Riek, R., Pervushin, K., and Wuthrich, K. 2000. TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci.* **25:** 462–468.

Stafford, R.E. 1989. Interfacial properties and critical micelle concentration of lysophospholipids. *Biochemistry* **28:** 5113–5120.

Svergun, D.I. and Koch, M.J. 2003. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* **66:** 1735–1782.

Tausk, C.L., van Esch, J., Karmiggelt, J., Voordouw, G., and Overbeek, J.T.G. 1974. Physical chemical studies of short-chain lecithin homologues. II: Micellar weights of dihexanoyl- and dihetanoyllecithin. *Biophys. Chem.* **1:** 184–203.

Therien, A.G. and Deber, C.M. 2002. Interhelical packing in detergent micelles. Folding of a cystic fibrosis transmembrane conductance regulator construct. *J. Biol. Chem.* **277:** 6067–6072.

Thiyagarajan, P. and Tiede, D.M. 1994. Detergent micelle structure and micelle-micelle interactions determined by small-angle neutron scattering under solution conditions used for membrane protein crystallization. *J. Phys. Chem.* **98:** 10343–10351.

Tian, C., Karra, M.D., Ellis, C.D., Jacob, J., Oxenoid, K., Sonnichsen, F., and Sanders, C.R. 2005. Membrane protein preparation for TROSY NMR screening. *Methods Enzymol.* **394:** 321–334.

Vinogradova, O., Sonnichsen, F., and Sanders, C.R. 1998. On choosing a detergent for solution NMR studies of membrane proteins. *J. Biomol. NMR* **11:** 381–386.

Weiss, H.M. and Grisshammer, R. 2002. Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur. J. Biochem.* **269:** 82–92.

Wiener, M.C. 2004. A pedestrian guide to membrane protein crystallization. *Methods* **34:** 364–372.

Wishart, D.S. and Sykes, B.D. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239:** 363–392.

Womack, M.D., Kendall, D.A., and MacDonald, R.C. 1983. Detergent effects on enzyme-activity and solubilization of lipid bilayer-membranes. *Biochim. Biophys. Acta* **733:** 210–215.

Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons, Inc., New York.

Zhang, R.T., Marone, P.A., Thiyagarajan, P., and Tiede, D.M. 1999. Structure and molecular fluctuations of n-alkyl-β-D-glucopyranoside micelles determined by X-ray and neutron scattering. *Langmuir* **15:** 7510–7519.