Supplementary Material

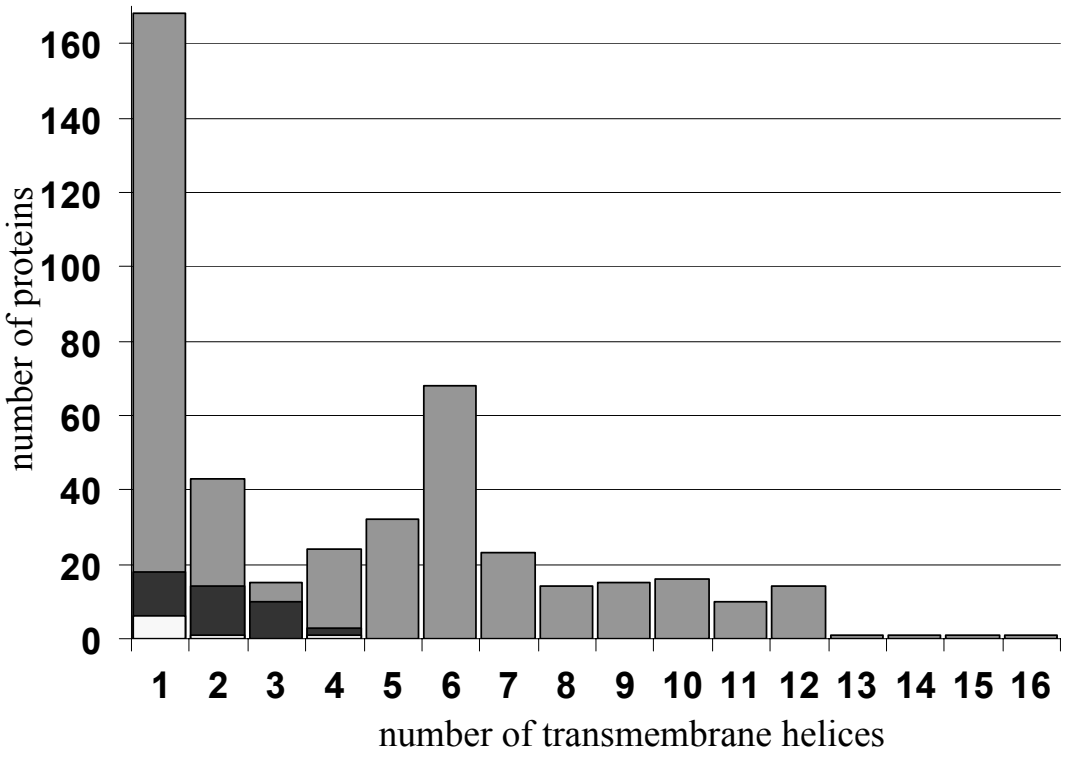
Expression, purification, and characterization of *Thermotoga maritima*

membrane proteins for structure determination

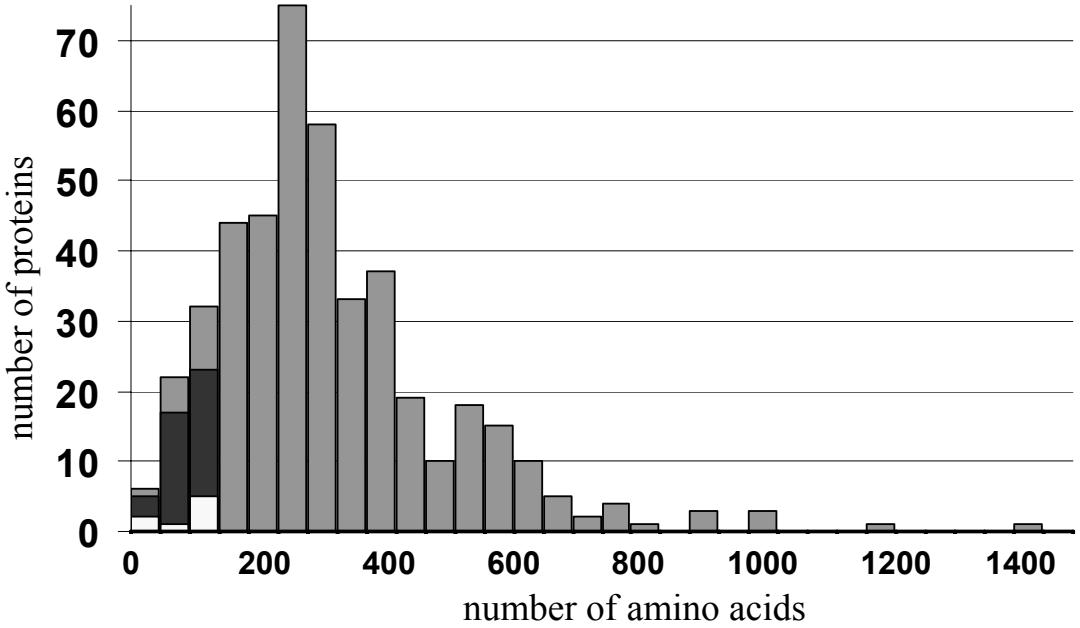
Linda Columbus, Jan Lipfert, Heath Klock, Ian Millett, Sebastian Doniach, and Scott A. Lesley

Protein Science (2006) **15**(5)

A



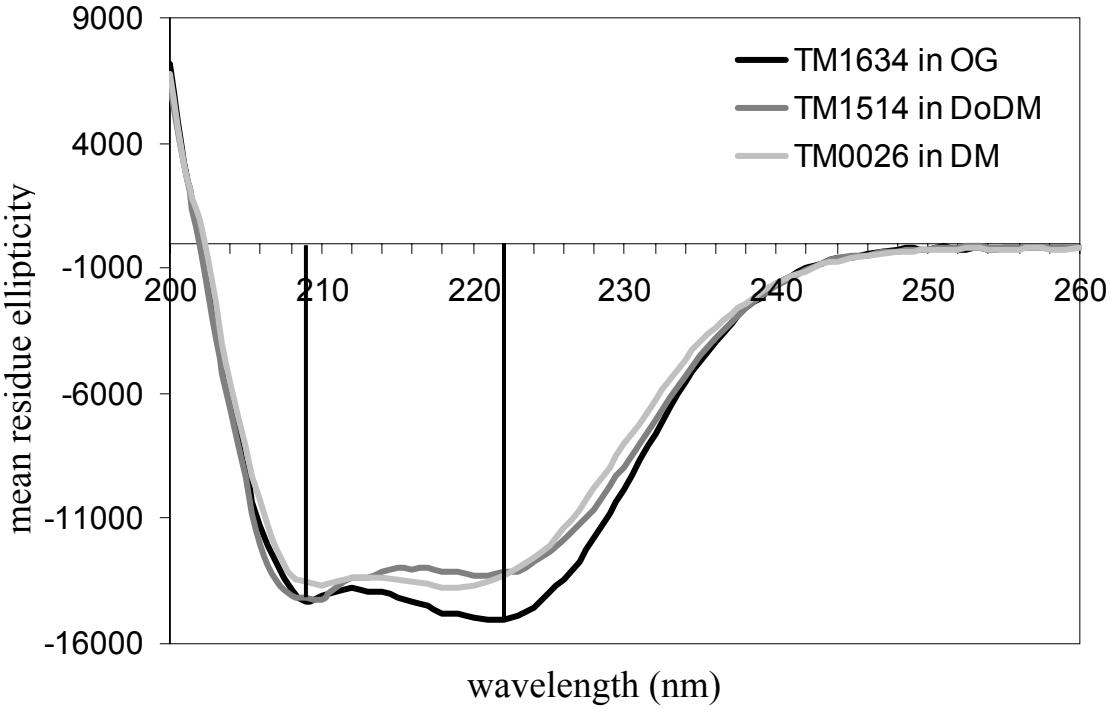
B



Supplementary Figure 1. Distribution of the number of transmembrane helices and the overall length of the membrane proteins in *T. maritima*. (A) The number of proteins is plotted versus the number of transmembrane helices for the total membrane proteome (light gray bars), the forty-five NMR-target membrane proteins (dark gray bars), and the eight proteins that over-expressed and localized to the membrane (white bars). (B) The number of proteins is plotted versus the number of amino acids in the protein for the total membrane proteome (light gray bars), the forty-five target membrane proteins (dark gray bars), and the eight proteins that over-expressed and localized to the membrane (white bars).

CD Spectroscopy. The secondary structure content of the soluble membrane protein samples was approximated by measuring the molar elipticity between 190 and 260 nm. The spectra were recorded with an Aviv (model 202) circular dichroism spectropolarimeter at 25°C. Protein concentrations were approximately 15 μM in 20 mM phosphate buffer at pH 7.0, 150 mM NaCl, and various concentrations of detergents.

CD analysis programs quantify secondary structure of membrane proteins poorly (Wallace et al. 2003) most likely because of the limited number of membrane protein references (Sreerama and Woody 2004). Regardless, the overall evaluation of α-helicity can be qualitatively analyzed based on the negative intensities observed at 222 and 209 nm, corresponding to the n→π* and the parallel component of the π→π* transition, respectively. For membrane proteins, the transmembrane α-helices are characterized by more negative intensities at 208 and 222 nm compared to soluble α-helices; in addition, Park *et al.* report that the 208 nm intensity is slightly less negative than that at 222 nm (Park et al. 1992; Fasman 1993). The CD spectra recorded for the PDCs all had characteristic α-helical intensities; however, the ratio of the intensity (I) at 208 nm to 222 nm varied. Three CD spectra corresponding to TM1634 in OG (black trace), TM1514 in DoDM (gray trace), and TM0026 in DM (light gray trace) are shown in Figure 2. Each spectrum represents one of the three ratios observed: $I_{n\rightarrow\pi^*}/I_{\pi\rightarrow\pi^*} > 1$, $I_{n\rightarrow\pi^*}/I_{\pi\rightarrow\pi^*} \approx 1$, and $I_{n\rightarrow\pi^*}/I_{\pi\rightarrow\pi^*} < 1$. In addition, the wavelength of the n→π* transition ranged from 219 nm to 224 nm.



Supplementary Figure 2. Representative CD spectra of PDCs. TM1634 in OG (black), TM1514 in DoDM (gray), and TM0026 in DM (light gray) are shown.

SAXS: theoretical. We measure the scattering intensity as a function of momentum transfer s, with $s = 2\sin(\theta)/\lambda$, where 2θ is the total scattering angle and λ, the photon wavelength, is 1 Å for all measurements. In the absence of interparticle interference, the scattering at low scattering angles can be expanded in the Guinier approximation (Guinier 1939; Svergun and Koch 2003)

$$\ln(I(s)) \approx \ln(I(0)) - \frac{(2\pi s R_g)^2}{3} \tag{1}$$

and from a linear fit of $\log(I)$ as a function of s^2 the forward scattering intensity $I(0)$ and radius of gyration R_g is obtained. The forward scattering intensity can be expressed in terms of molecular weight M, concentration c and average electron density contrast $\Delta\rho$ (Glatter and Kratky 1982)

$$I(0) = K\, c\, (\Delta\rho\, M)^2 \tag{2}$$

The electron density contrast $\Delta\rho$ is given by the average electron density of the scatterer minus the electron density of the displaced volume of solvent. The proportionality constant K depends on experimental parameters that are held constant throughout the measurements and can be determined by comparison with a molecular weight standard. As a direct consequence of equation (2), oligomerization increases the forward scattering intensity. For example, dimerization leads to a two-fold reduction in concentration and a two-fold increase in molecular weight, which results in a two-fold increase in the forward scattering intensity.

In the absence of interparticle interference, the scattering signal from a mixture of two different molecular species is

$$I(s) = c_1 I_1(s) + c_2 I_2(s) \tag{3}$$

in which I_1 and I_2 are the scattering contributions from the two different species and c_1 and c_2 are their respective concentrations. The apparent radius of gyration of the mixture is given by

$$R_g^2 = \frac{c_1 I_1(0) R_{g,1}^2 + c_2 I_2(0) R_{g,2}^2}{c_1 I_1(0) + c_2 I_2(0)} \tag{4}$$

where $R_{g,1}$ and $R_{g,2}$ are the radii of gyration of the two species.

Also, a new method to approximately separate the scattering contribution of the PDC from that of the "empty" micelle (N.B. this population will be referred to as micelle in the remaining text) is presented, in particular for the low s region used to fit the radius of gyration and forward scattering intensity. In general, the scattering contributions from monomeric (i.e. not micellar) detergent molecules is neglected, which is justified because i) the concentration of monomeric detergent is expected to be approximately equal to the CMC, which is much less than the micelle concentration in these experiments and ii) the scattering signal from monomeric detergent molecules is much smaller than that of the micelles, due to the large difference in molecular weight. In the following, we distinguish two different concentration regimes: (1) where the PDC concentration is greater than the micelle concentration and (2) where the micelle concentration is in excess over the PDC concentration.

Method I: PDC concentration in excess; $I_{complex}$-I_{buffer}. In the case that most detergent molecules in the solution are bound in the PDC, we can obtain an estimate of its scattering profile by subtracting the buffer (20 mM sodium phosphate and 150 mM NaCl) profile from the full scattering profile obtained for the protein-detergent solution. The buffer subtraction is always necessary to remove parasitic scattering and other artifacts. The intensity profile has contributions

$$I(s) = c_{complex} I_{complex}(s) + c'_{micelle} I_{micelle}(s) = c_{complex} I_{complex}(s)(1+\varepsilon) \quad (5)$$

where $c'_{micelle}$ denotes the concentration of micelles remaining after the detergent molecules in the PDC are accounted for and

$$\varepsilon = \frac{c'_{micelle} I_{micelle}}{c_{complex} I_{complex}}. \quad (6)$$

In general, the micelle scattering can be as large as or larger than the signal from the protein detergent complex. In the Guinier region, i.e. for small *s*, we expect the complex to typically scatter more strongly than the detergent micelle, as it is usually larger and more electron dense than the micelle. Hence in the case where $c'_{micelle}$ is small relative to $c_{complex}$ and/or when $\frac{I_{micelle}}{I_{complex}} << 1$, we can neglect the second term and to first order simply use the buffer-subtracted profile to estimate the forward scattering and radius of gyration of the protein-detergent complex from the Guinier fit. The radius of gyration in this case is given by

$$R_g^2 = R_{g,complex}^2 \frac{1 + \varepsilon \frac{R_{g,micelle}^2}{R_{g,complex}^2}}{1+\varepsilon} \quad (7)$$

Typically $R_{g,micelle}$ is smaller than $R_{g,complex}$ and the fitted apparent radius of gyration is smaller than the true $R_{g,complex}$.

Method II: Micelle in excess; $I_{complex}$-$I_{micelle}$. As an alternative approach in the limit that micelles are present at much higher concentration than the PDC, we can subtract the scattering profile of the micelle solution measured in the absence of protein. In the absence of interference, this yields

$$I(s) = c_{complex} I_{complex}(s) + c'_{micelle} I_{micelle}(s) - c_{micelle} I_{micelle}(s) = c_{complex} I_{complex}(s)(1+\varepsilon') \quad (8)$$

where

$$\varepsilon' = \frac{(c'_{micelle} - c_{micelle}) I_{micelle}(s)}{c_{complex} I_{complex}(s)} \tag{9}$$

and $c_{micelle}$ is the concentration of micelles in the absence of protein, which is given by $c_{micelle} \approx c_{\mathrm{det}ergent} / m_{mic}$ where m_{mic} is the micelle aggregation number, and $c'_{micelle}$, the micelle concentration in the presence of protein. In the case that there is detergent in excess and only a small fraction of the detergent molecules bind to the protein, we can expect ε' to be small and the micelle-subtracted profile will provide a reasonable approximation for calculating the forward scattering and R_g in the Guinier region. In this limit, the fitted apparent radius of gyration is

$$R_g^2 = R_{g,complex}^2 \frac{1 + \varepsilon' \dfrac{R_{g,micelle}^2}{R_{g,complex}^2}}{1 + \varepsilon'} \tag{10}$$

As $\varepsilon' < 0$ the fitted radius of gyration exceeds the true $R_{g,complex}$, if the micelle is smaller than the PDC.

In the case that the micelle scatters very weakly as compared to the protein-detergent complex, i.e. $I_{micelle} / I_{complex} << 1$, both approximations will converge to the same result. Two types of errors affect the measured radius of gyration $R_{g,complex}$ and forward scattering intensity $I(0)_{complex}$: (1) experimental errors are always associated with obtaining these parameters from a buffer subtracted scattering profile, such as detector noise, beam drift, uncertainties in the Guinier fit, and errors from parasitic scattering due to imperfect background subtraction. These errors were minimized by averaging over multiple sample and background profiles and quantified by performing multiple Guinier fits for slightly different fitting ranges for each measurement; (2) the approximation made when correcting for the micelle scattering signal

contributes additional errors and can be estimated by comparing the two approximations. As $\varepsilon > 0$, the I(complex-buffer) treatment yields –in the absence of other experimental errors- an upper bound on the forward scattering intensity, whereas the I(complex-micelle) approximation provides a lower bound. Similarly, the true $R_{g,complex}$ can be bracketed by comparing the two estimates. In addition to the radius of gyration and forward scattering intensity, the SAXS measurement provides information about monodispersity and possible aggregation of the sample. Non-linearity in the Guinier representation (the plot of $\log(I)$ as a function of s^2 for small s) is a sign of polydispersity, i.e. a broad particle size distribution (Svergun and Koch 2003). Particle aggregation leads to a very strong scattering signal in the low s region, which corresponds to large spatial dimensions.

SAXS: Estimating oligomeric states. For a complex of N protein molecules of known (monomer) molecular weight M_{prot} and m_{PDC} detergent molecules of known (monomer) molecular weight $M_{\det}$, the total forward scattering intensity is given by

$$I_{PDC}^{theory}(0) = K\, c_{PDC}(N\, M_{prot}\, \Delta\rho_{prot} + m_{PDC}\, M_{\det}\, \Delta\rho_{\det})^2 \qquad (11)$$

Where K is a proportionality constant (the same as in Equation 3) that is determined from comparison with a molecular weight standard and $\Delta\rho_{prot}$ and $\Delta\rho_{\det}$ are the excess electron density for protein and detergent, respectively. The electron density contrast $\Delta\rho_{\det ergent} = \rho_{\det ergent} - \rho_{water}$ is determined from the known electron density of water, ρ_{water} = 0.334 e/Å^3 and the detergent electron densities in Table 1. The detergent electron density values were obtained by summing the number of electrons from the chemical composition and dividing by the molecular volume given in Table 1. The volumes were calculated from the published specific densities (le Maire et al. 2000) and using the Tanford formula for alkyl chain volumes to

adjust for different chain lengths (Tanford 1980). For LPPG, the molecular volume was computed by summing the partial chemical group volumes (Reynolds and McCaslin 1985).

The theoretical forward scattering intensity value needs to be compared to the experimentally obtained estimates. However, the associated errors are quite large because the approximation used to obtain the PDC scattering profile gives a less accurate estimate for the I(0) than for the R_g. Also, Equation 1 depends on the protein concentration, $c_{PDC} = c_{prot} / N$, which is difficult to measure accurately. A unique determination is impossible because N and m_{PDC} can vary independently. Nonetheless, solutions for different N values can be evaluated with the restriction that m_{PDC} is positive and is on the order of the micelle aggregation number m_{mic}.

References

Fasman, G.D. 1993. Distinguishing transmembrane helices from peripheral helices by circular dichrosim. *Biotechnol. Appl. Biochem.* **18 (Pt 2):** 111-138.

Glatter, O. and Kratky, O. 1982. *Small-Angle X-ray Scattering*. Academic Press, London, UK.

Guinier, A. 1939. La diffraction des rayons X aux tres petits angles: Application a l'etude de phenomenes ultramicroscopiques. *Ann. de Phys.* **12:** 161-237.

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

Park, K., Perczel, A., and Fasman, G.D. 1992. Differentiation between transmembrane helices and peripheral helices by the deconvolution of circular dichroism spectra of membrane proteins. *Protein Sci.* **1:** 1032-1049.

Reynolds, J.A. and McCaslin, D.R. 1985. Determination of protein molecular-weight in complexes in detergent without knowledge of binding. *Methods Enzymol.* **117:** 41-53.

Sreerama, N. and Woody, R.W. 2004. On the analysis of membrane protein circular dichroism spectra. *Protein Sci.* **13:** 100-112.

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

Tanford, C. 1980. *The hydrophobic effect: formation of micelles and biological membranes*, 2nd ed. Wiley, New York.

Wallace, B.A., Lees, J.G., Orry, A.J., Lobley, A., and Janes, R.W. 2003. Analyses of circular dichroism spectra of membrane proteins. *Protein Sci.* **12:** 875-884.