

Supplementary Information

Optimized reverse micelle surfactant system for high-resolution NMR spectroscopy of encapsulated proteins and nucleic acids dissolved in low-viscosity fluids

Igor Dodevski,¹ Nathaniel V. Nucci,¹ Kathleen G. Valentine,¹ Gurnimrat K. Sidhu,¹ Evan S. O'Brien,¹ Arthur Pardi,² and A. Joshua Wand¹

¹Johnson Research Foundation and Department of Biochemistry & Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania USA

²Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado USA

Corresponding Author: A. J. Wand (wand@mail.med.upenn.edu)

Protein & tRNA expression and purification

Seven proteins were expressed on minimal media using established procedures: human ubiquitin (Ub),¹ horse cytochrome c (Cyt c),² the C55A mutant of flavodoxin from *C. anabaena* (Flv),³ malate synthase G from *E. coli* (MSG),⁴ N-terminal myristoylated human recoverin (Rec[+myr]),⁵ full-length maltose binding protein from *E. coli* (MBP),⁶ and human ald-keto reductase 1C2 (AKR)⁷ (courtesy of Professor Trevor Penning). The gene for arginine kinase from *L. polyphemus* (AK) was cloned into the pET11 plasmid. MBP, AK, and AKR were purified by DEAE-sepharose anion exchange chromatography and Superdex 200 size exclusion chromatography. Maltose was added to MBP following purification and refolding. NADP⁺ was added to AKR following purification. Native (fully modified) ¹⁵N-labeled *E. coli* tRNA^{Val} was prepared by overexpression in *E. coli* containing the pVALT7 plasmid,⁸ where a 100 mL culture was used to inoculate 1 L of minimal medium containing ¹⁵NH₄Cl as the only nitrogen source. The tRNA^{Val} was purified by denaturing polyacrylamide gel electrophoresis and anion exchange chromatography as described,⁹ and then was extensively dialyzed into 10 mM sodium phosphate (pH 6.8), 80 mM NaCl, 5 mM MgCl₂ and 0.1 mM EDTA.

Systematic screening of surfactants for optimal protein encapsulation

Mixtures of up to four different commercially available surfactants were screened featuring nonionic, ionic, or zwitterionic head groups and different types of linear, branched, or cyclic hydrophobic tails (Supplementary Table 1). In the first step, more than one thousand surfactant mixtures were screened for their ability to form stable reverse micelle solutions initially in hexane (which has minimal volatility) and then in pentane in the absence of encapsulated protein (hereafter termed “empty reverse micelles”). The screen focused on surfactant mixtures having higher nonionic than ionic content, since such mixtures would be less likely to show perturbing electrostatic interactions with an encapsulated protein. Hexanol, which is typically present as a stabilizing cosurfactant in RM preparations, was not included in this initial screen in order to emphasize the propensities of the individual surfactants to form stable reverse micelles. To carry out the initial screen a small defined volume of aqueous solution, containing bromophenol blue as a visual aide, was mixed with 400 μL of hexane containing 75 mM of a given surfactant mixture. Samples were initially prepared at a molar ratio of water to total surfactant of 10 (also known as water loading or W_0). For those initial conditions that turned clear blue upon RM formation without forming precipitate, W_0 was successively increased until a phase separation was observed. This initial screen revealed a variety of surfactant mixtures capable of forming stable empty reverse micelles at W_0 values of 10 to 30, which are typically used for protein RM NMR.

Those surfactant mixtures revealed by the empty RM screen were subsequently examined for their ability to encapsulate protein. For this, the 18.8 kDa acidic ($pI \sim 4.5$) flavodoxin (Flv) from *C. anabaena*, which appears bright yellow ($\lambda_{\text{max}} = 462 \text{ nm}$) when the flavin mononucleotide (FMN) cofactor is bound was used. Ten μL of a 5 mM flavodoxin stock solution was directly injected into 500 μL solution of pentane containing 75 mM of a given surfactant mixture resulting in 100 μM flavodoxin samples at $W_0 = 15$. The procedure of injecting and mixing a defined volume of concentrated protein solution with a defined surfactant solution is termed the direct injection method (Fig. 1c). Many samples showed the characteristics of appropriately encapsulated flavodoxin, namely a transparent yellow pentane phase and little or no

precipitate. These apparently successful encapsulation conditions were then examined by ^{15}N -heteronuclear single quantum coherence (^{15}N -HSQC) spectroscopy (Supplementary Fig. 1). Comparison of the ^{15}N -HSQC of encapsulated flavodoxin in different surfactant mixtures to that in free aqueous solution was used to assess structural integrity.

To identify those surfactant mixtures that would provide the best performance in RM NMR we estimated the molecular correlation time τ_m of encapsulated flavodoxin using the ^{15}N -TROSY-based (TRACT) relaxation experiment.¹⁰ Although the ^{15}N -TRACT experiment systematically underestimates τ_m , it provides a fast, relative measure and such estimates are designated as $\tilde{\tau}_m$. Supplementary Table 2 provides a summary of surfactant mixtures tested and the corresponding $\tilde{\tau}_m$ values for encapsulated flavodoxin, which were found to be between 7.1 and 13.3 ns. This systematic surfactant screen identified a binary surfactant mixture featuring 1-decanoyl-*rac*-glycerol and lauryldimethylamine-N-oxide (10MAG/LDAO) as the most promising surfactant system for RM NMR.

10MAG/LDAO reverse micelle sample preparation

All surfactants were purchased at the highest commercially available grade of purity and they were used without further purification, except that LDAO was desiccated by lyophilization to remove residual water. d_{12} -pentane and d_6 -ethane were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Volatile solvents were handled with positive displacement pipettes.

Encapsulated protein was prepared by mixing the appropriate volume of concentrated protein solution with a defined amount of surfactant and alkane solvent to achieve the desired W_0 . For reverse micelle samples prepared in pentane, the surfactants 10MAG and LDAO, and the co-surfactant hexanol were dissolved in 600 μl pentane in a 1.8 ml glass vial and the appropriate volume of concentrated protein in aqueous buffer (see below) was added to the pentane/surfactant mixture (e.g., for a 600 μl sample comprising 75 mM 10MAG/LDAO at $W_0 = 20$, 16.2 μl of aqueous protein solution was injected). The sample was mixed by vigorous vortexing for up to 10 minutes. Typically at this point a phase transition from a cloudy

suspension to a clear solution would be visible. For certain samples (especially for the proteins in the 40 kDa range and above), encapsulation efficiency improved after overnight incubation at room temperature on a rotisserie. For NMR, encapsulated protein samples in pentane were transferred to glass NMR tubes sealed with a PTFE-lined screw cap. Teflon tape around the caps was used to prevent sample evaporation.

Reverse micelle samples in liquid ethane were prepared using specialized apparatus from Daedalus Innovations, LLC (Aston, PA). LDAO, 10MAG, and hexanol were dissolved in 300 μ L of d_{12} -pentane and transferred to the mixing chamber (1.65 ml nominal volume) of a Daedalus Innovations RM Synthesizer unit. A volume of concentrated protein solution yielding the desired final W_0 was added and the mixing chamber was sealed. Pressurized liquid ethane was pumped into the mixing chamber with an Isco 65D (Lincoln, NE) syringe pump while the sample was being stirred. The sample pressure was increased until a distinct phase transition from a cloudy suspension to a clear solution was observed. This generally occurred in the range of 275-450 bar though higher pressures were sometimes required for proteins of larger size and/or at higher protein concentrations. The sample pressure was adjusted to 14-20 bar above the observed encapsulation (transition) pressure. The sample was then transferred to a high-pressure 3.0 mm i.d. NMR cell rated to 1 kbar (Daedalus Innovations, LLC, Aston, PA) for further analysis by NMR. The final sample contained 80% (v/v) ethane and 20% (v/v) pentane.

Specific Reverse Micelle Sample Details

The surfactant system was employed at a molar ratio of 10MAG/LDAO of 65%:35% and at a total surfactant concentration of 75-150 mM. The concentration of the co-surfactant hexanol depended on the total 10MAG/LDAO surfactant concentration and on the type of bulk alkane solvent used (pentane or ethane). Pentane samples were prepared with 10-20 mM hexanol at 75 mM surfactant and 20-30 mM hexanol at 150 mM surfactant. Ethane samples were typically prepared with 30-50 mM hexanol at 75 mM surfactant and 60-70 mM hexanol at 150 mM surfactant.

It is important to note that concentrated solutions of proteins used for direct injection do not need to be “good solutions” in the usual sense. Rather, they may be slurries. Unless otherwise

stated the concentrated aqueous protein or RNA solutions for preparing reverse micelle samples by the direct injection method were as follows. Ub—20 mM in 50 mM sodium acetate (pH 5.0), 50 mM NaCl; Cyt c—15 mM in 50 mM sodium acetate (pH 5.0), 50 mM NaCl, 10 mM ascorbic acid to maintain the reduced state of the heme; Flv—5.6 mM in 50 mM sodium phosphate (pH 5.8); AKR—2.1 mM in 25 mM MES (pH 5.8), 50 mM KCl, 1 mM EDTA, 1 mM TCEP, 3 mM NADP⁺; AK—2.2 mM in 25 mM MES (pH 6.0), 50 mM NaCl, 1 mM EDTA, 1 mM TCEP; MBP—7.7 mM in 50 mM sodium phosphate (pH 7.0), 50 mM NaCl, 7 mM maltose; MSG—2.8 mM in 25 mM MES (pH 6.4), 50 mM KCl, 1 mM EDTA, 1 mM TCEP; Rec(+myr)—4.6 mM in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM DTT, 1 mM EGTA, 30 mM CaCl₂; tRNA^{val}—4.6 mM in 10 mM sodium phosphate (pH 6.8), 80 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA.

Encapsulated proteins and tRNA were prepared at the following final sample conditions: Ub—270 μ M in 75 mM surfactant at $W_0 = 10$; Cyt c—200 μ M in 75 mM surfactant at $W_0 = 10$; Flv—150 μ M in 150 mM surfactant at $W_0 = 10$; AKR—100 μ M in 150 mM surfactant at $W_0 = 18$; AK—120 μ M in 150 mM surfactant at $W_0 = 20$; MBP—125 μ M in 75 mM surfactant at $W_0 = 12$; MSG—150 μ M in 150 mM surfactant at $W_0 = 20$; Rec(+myr)—100 μ M in 100 mM surfactant at $W_0 = 12$; tRNA^{val}—100 μ M in 150 mM surfactant at $W_0 = 9$.

Buffer conditions for proteins prepared in free aqueous solution were identical to those described above except that the pH for AK, AKR, MBP, and MSG was increased by 0.6-1.0 pH units above that of the encapsulation buffer. The pH of these aqueous samples was AKR—pH 6.4; AK—pH 7.0; MBP—pH 7.0, MSG—pH 7.0. This modification in the buffer pH was found to minimize pH-induced chemical shift differences observed between proteins in the encapsulated and non-encapsulated state. It should be noted that the pH can be set by pre-equilibration of aqueous solutions of surfactants followed by lyophilization.¹¹ All aqueous samples contained 0.02% (w/v) sodium azide.

Supplementary Table S1. List of surfactants screened for the formation of empty reverse micelles.

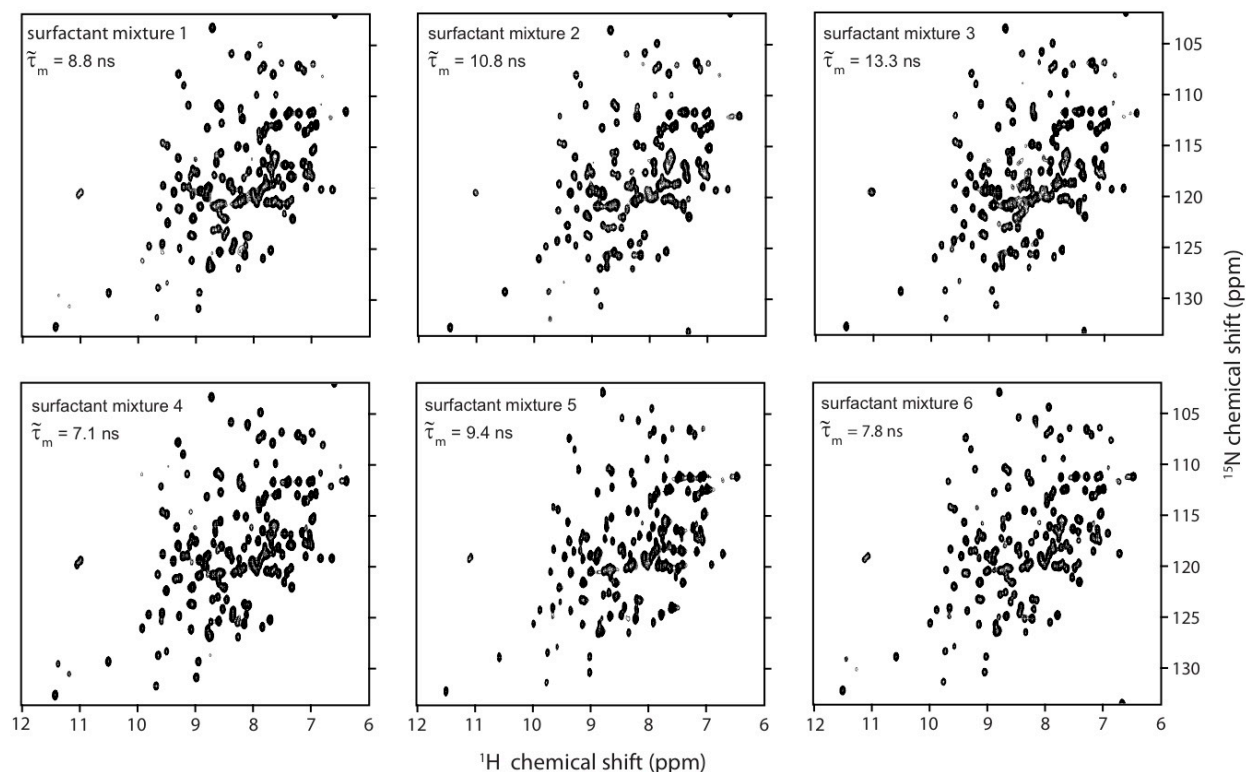
Surfactant	Vendor
<u>Nonionic</u>	
Sorbitane monooleate (Span 80)	Sigma-Aldrich
Triton X-45	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tergitol 15S-3	Sigma-Aldrich
Tergitol 15S-5	Sigma-Aldrich
Tergitol 15S-7	Sigma-Aldrich
Tergitol TMN-6	Sigma-Aldrich
1-(7Z-tetradecenoyl)- <i>rac</i> -glycerol (7.7MAG)	Avanti Lipids
1-(9Z-hexadecenoyl)- <i>rac</i> -glycerol (9.7 MAG)	Avanti Lipids
1-Dodecanoyl- <i>rac</i> -glycerol (12 MAG)	Avanti Lipids
1-Decanoyl- <i>rac</i> -glycerol (10MAG)	Sigma-Aldrich
<u>Zwitterionic</u>	
Lauryldimethylamine-N-oxide (LDAO)	Affymetrix
Dodecylphosphocholine (DPC)	Affymetrix
Tetradecylphosphocholine (Fos-14)	Affymetrix
Tripod amphiphile (Tripao)	Affymetrix
C-dodecafos	Affymetrix
Nopol-fos	Affymetrix
<u>Cationic</u>	
Cetyltrimethylammonium bromide (CTAB)	Sigma-Aldrich
Dodecyltrimethylammonium bromide (DTAB)	Sigma-Aldrich
Dodecyltrimethylammonium chloride (DTAC)	Sigma-Aldrich
Dihexadecyldimethylammonium bromide (DHAB)	Sigma-Aldrich
Didodecyldimethylammonium bromide (DDAB)	Sigma-Aldrich
<u>Anionic</u>	
bis(2-ethylhexyl)-sulfosuccinate (AOT)	Sigma-Aldrich
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich

Supplementary Table S2. RM surfactant conditions and corresponding molecular reorientation times of encapsulated flavodoxin in pentane at 25 °C.

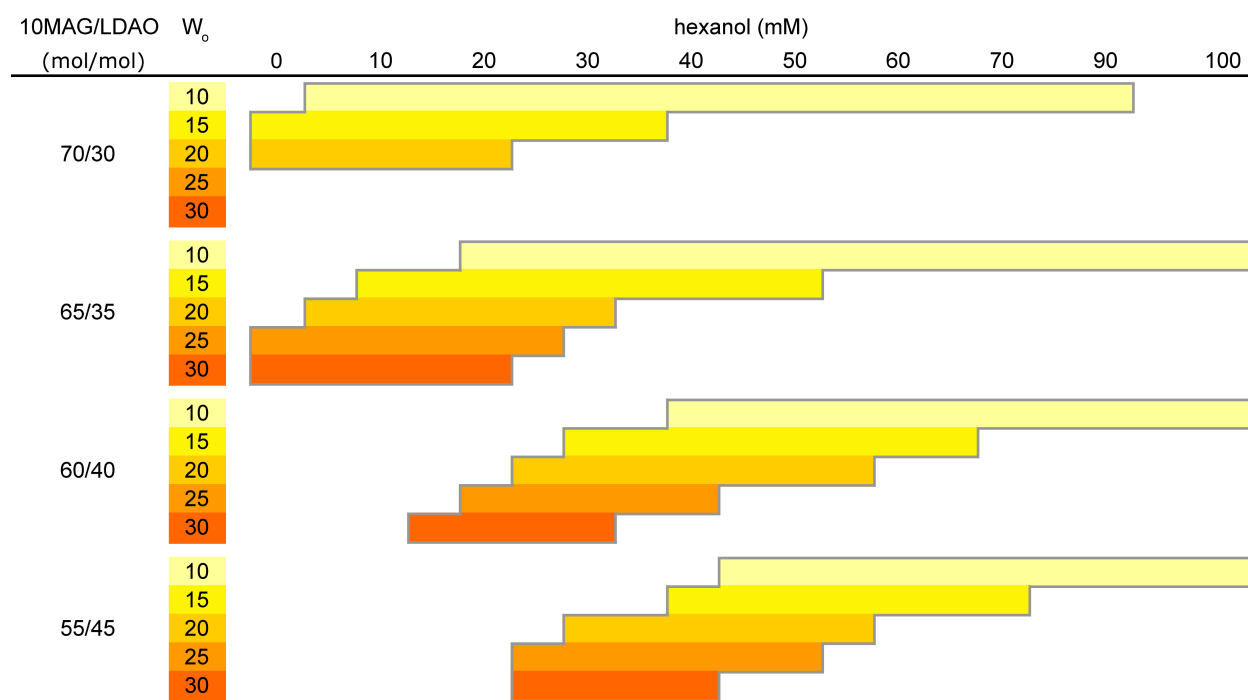
Surfactant mixture	$\tilde{\tau}_m^*$	Surfactant (mM) [†]									
		Tergitol 15S-5	Span80	Triton X-45	10MAG	77MAG	AOT	CTAB	LDAO	Tripao	DPC
1	8.8	21	32						23		
2	10.8	45	8				6	17			
3	13.3			53			6	17			
4	7.1				53				23		
5	9.4		21	32						23	
6	7.8					53					23

* Estimated molecular correlation time as measured by the ¹⁵N-TRACT experiment.

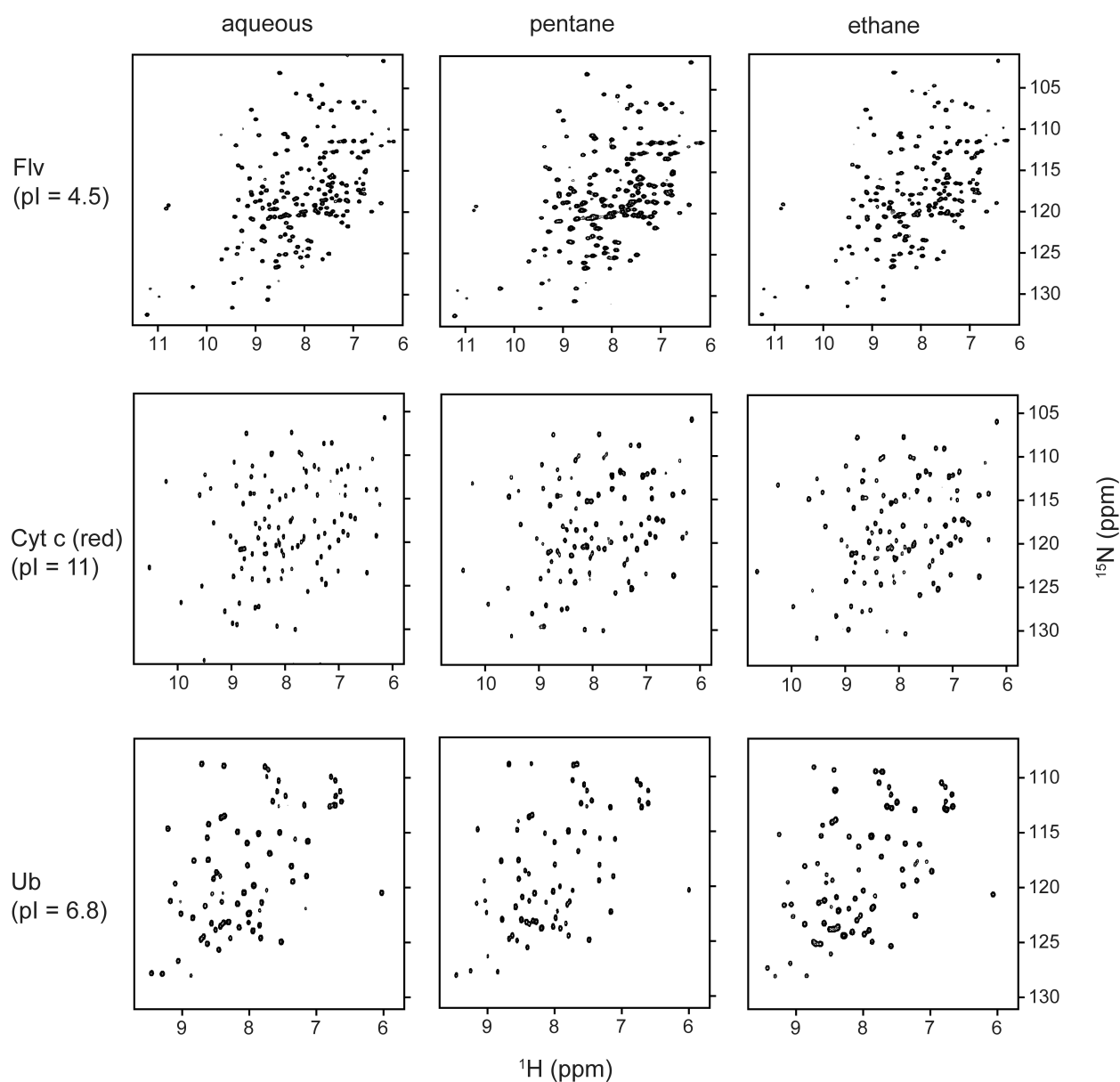
† 75 mM total surfactant concentration with 70% (mol/mol) nonionic surfactant in all mixtures.



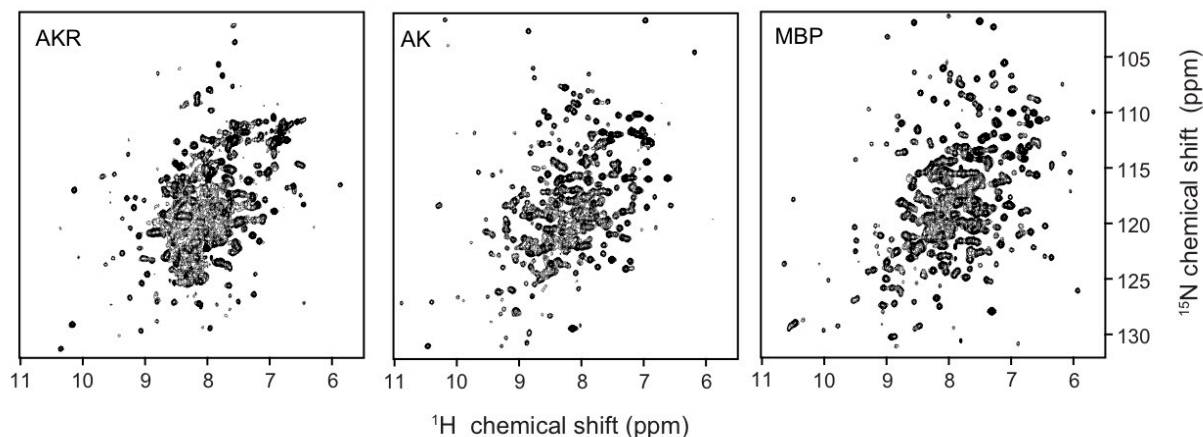
Supplementary Figure S1. ^{15}N -HSQC screening of encapsulated Flv in six different surfactant mixtures dissolved in pentane. The surfactant compositions for each mixture are given in Supplementary Table 2. The molecular reorientation times $\tilde{\tau}_m$ were estimated by the ^{15}N -TRACT experiment and are indicated for each surfactant mixture. Although Flv is encapsulated with high structural fidelity in all six mixtures, not all mixtures are characterized by fast $\tilde{\tau}_m$ values. Surfactant mixture 4 excelled in both fast $\tilde{\tau}_m$ and spectral matching with the aqueous Flv sample and it was further developed to become the 10MAG/LDAO surfactant system described in this paper. The samples were 100 μM Flv and spectra were recorded at 25 $^{\circ}\text{C}$ at 500 MHz (^1H) with 48 complex points in the indirect ^{15}N dimension and 16 or 32 transients per FID. All spectra were processed identically.



Supplementary Figure S2. Systematic examination of the effects of 10MAG/LDAO molar surfactant ratio, water loading (W_0), and hexanol concentration on the formation of reverse micelles. These parameters were tested in an empty reverse micelle screen (i.e., the encapsulation solution contained buffer but no protein) in 75 mM total surfactant concentration using pentane as the bulk solvent. Hexanol concentration was tested from 0 to 100 mM and W_0 values were tested from 10 to 30. The colored areas, reflecting the associated W_0 values, indicate experimental conditions in which stable empty reverse micelles are formed, i.e. in which a clear reverse micelle solution is formed without any visual signs of precipitate. The largest region of such conditions was found at a surfactant ratio of 10MAG/LDAO = 65/35 (mol/mol). This surfactant ratio was subsequently used for protein encapsulation in 10MAG/LDAO for reverse micelle protein NMR. Stable reverse micelles can also be obtained in the absence of the co-surfactant hexanol within a narrow range of surfactant ratios and water loadings.



Supplementary Figure S3. ^{15}N -HSQC spectra of Ub, Flv, and reduced Cyt c in aqueous solution and encapsulated in 10MAG/LDAO reverse micelles in pentane and ethane. The possibility of encapsulating these three proteins, which are characterized by distinctly different isoelectric points ranging from 4.5 to 11, demonstrates efficient protein encapsulation in 10MAG/LDAO independent of protein net surface charge. The pH of the encapsulation buffers was such that the net surface charge was negative for Flv and positive for Cyt c and Ub.



Supplementary Figure S4. ^{15}N -HSQC spectra of AKR, AK and MBP encapsulated in 10MAG/LDAO reverse micelles and dissolved in pentane. All samples were 150 mM 10MAG/LDAO at a molar percent ratio of 65%:35%, 30 mM hexanol, $W_0 = 20$ and protein concentrations were 100 μM AKR, 110 μM AK, and 150 μM MBP. The spectra were recorded with identical parameters (100 complex ^{15}N points and 32 scans per FID) and processed identically. All spectra were recorded at 750 MHz (^1H) at 25 $^\circ\text{C}$. The resonances are relatively broad due to the “volume penalty” imposed by encapsulation and the viscosity of pentane. This is overcome by encapsulation in ethane solutions (see Figure 3).

- (1) Wand, A. J.; Urbauer, J. L.; McEvoy, R. P.; Bieber, R. J. *Biochemistry* **1996**, *35*, 6116-6125.
- (2) Rumbley, J. N.; Hoang, L.; Englander, S. W. *Biochemistry* **2002**, *41*, 13894-13901.
- (3) Liu, W.; Flynn, P. F.; Fuentes, E. J.; Kranz, J. K.; McCormick, M.; Wand, A. J. *Biochemistry* **2001**, *40*, 14744-14753.
- (4) Howard, B. R.; Endrizzi, J. A.; Remington, S. J. *Biochemistry* **2000**, *39*, 3156-3168.
- (5) Valentine, K. G.; Peterson, R. W.; Saad, J. S.; Summers, M. F.; Xu, X.; Ames, J. B.; Wand, A. J. *Structure* **2010**, *18*, 9-16.
- (6) Nucci, N. V.; Marques, B. S.; Bédard, S.; Dogan, J.; Gledhill, J. M., Jr.; Moorman, V. R.; Peterson, R. W.; Valentine, K. G.; Wand, A. L.; Wand, A. J. *J. Biomol. NMR* **2011**, *50*, 421-430.
- (7) Jin, Y.; Stayrook, S. E.; Albert, R. H.; Palackal, N. T.; Penning, T. M.; Lewis, M. *Biochemistry* **2001**, *40*, 10161-10168.
- (8) Chu, W. C.; Horowitz, J. *Nucleic Acids Res.* **1989**, *17*, 7241-7252.
- (9) Vermeulen, A.; McCallum, S. A.; Pardi, A. *Biochemistry* **2005**, *44*, 6024-6033.
- (10) Lee, D.; Hilty, C.; Wider, G.; Wüthrich, K. *J. Magn. Reson.* **2006**, *178*, 72-76.
- (11) Marques, B. S.; Nucci, N. V.; Dodevski, I.; Wang, W. C.; Athanasoula, E. A.; Jorge, C.; Wand, A. J. *J. Phys. Chem. B* **2013**, DOI: jp4103349.