Supporting Information for:

Local Chemistry of the Surfactant's Head Groups Determines Protein Stability in Reverse Micelles**

Michael Senske[a], Yao Xu[a,b], Alexander Bäumer[a], Sarah Schäfer[a], Hanna Wirtz[a], Janne Savolainen[a], Hermann Weingärtner[a], Martina Havenith[a,*]

[a]Department of Physical Chemistry II

Ruhr-Universität Bochum

44780 Bochum, Germany

e-mail: martina.havenith@rub.de

[b]Present address:

Department of Cell and Molecular Biology, Computational Biology and Bioinformatics

Uppsala Universitet

751 24 Uppsala, Sweden

** We acknowledge funding from the Cluster of Excellence RESOLV (EXC 1069) funded by the German Research Foundation (DFG). M.S. acknowledges financial support by the Fonds der Chemischen Industrie. Y.X. thanks W.A. Goddard III, S.-T. Lin, T.A. Pascal, H. Antila and

S. Abel for discussions on the 2PT method and the simulation setup of the reverse micelles, respectively.

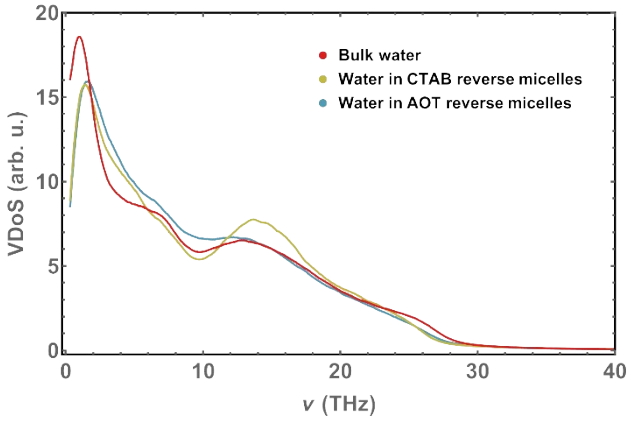

Figure SI1: VDoS of bulk water as well as AOT and CTAB reverse micelles.

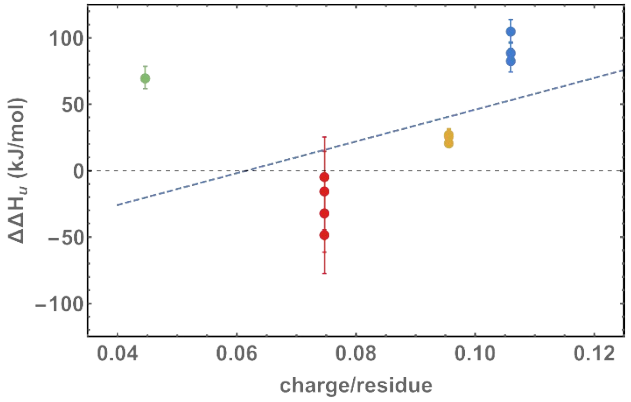

Figure SI2: $^{\Delta\Delta H_u}$ of all reverse micelle systems shown in Fig. 2. $^{\Delta\Delta H_u}$ was calculated at ^{T_m} of each protein in the respective buffer solution. The charge of each protein was calculated using the Adaptive Poisson-Boltzmann Solver [1] and PDB2PQR [2, 3] software and the pH values given in the Senske *et al.*[4] and Shastry *et al.*[5] The dashed line is a linear fit to the data to indicate the trend with increasing charge/residue. $^{\Delta\Delta H_u}$ was calculated as described below.

Table SI1: Densities of reverse micelle solutions, the used organic solvents and water. Densities were measured using a DMA 58 density meter (Anton Paar, Graz, Austria) at 20 °C and are given in g ml^{-1}. The maximum uncertainty given by the manufactures protocol is 0.0005 g ml^{-1}.

solutions	densities
AOT reverse micelle, $W_{7.5}$	0.79503
AOT reverse micelle, W_{10}	0.79649
AOT reverse micelle, $W_{12.5}$	0.79753
CTAB reverse micelle, $W_{7.5}$	0.79190
CTAB reverse micelle, W_{10}	0.79276
CTAB reverse micelle, $W_{12.5}$	0.79389
cyclohexane	0.77853
hexanol	0.81879
cyclohexane + hexanol (8 volume-%)	0.78050
water	0.99820

Table SI 2: Molecular composition of both simulated reverse micelles. R(H_2O) is the radius of the water shell in Å, N(hex) is the number of hexanol molecules, and m(surfact)/m(iso) is the total mass ratio of the surfactant over isooctane in the systems.

	Wo	N(surfact)	N(H_2O)	R(H_2O)/Å	N(hex)	N(iso)	m(surfact)/m(iso)
AOT	10	98	980	19	-	3500	10.9%
CTAB	10	135	1350	20	184	2604	16.5%

Data Analysis of Protein Stability Data

We analyzed protein stability data in reverse micelles of the following systems: N-terminal SH3 domain of the protein drk encapsulated in CTAB reverse micelles ($W_0 = 15, 20, 25$, pH 7.6 and $W_0 = 20$, pH 4.2) and 10MAG/LDAO reverse micelles ($W_0 = 15, 20, 25$, pH 7.0)[4] as well as ribonuclease (RNase) T1 in AOT reverse micelles ($W_0 = 4.94,\ 6.17, 7.40, 12.0$, pH 7.0).[5] Marques *et al*. observed that the pH inside AOT reverse micelles is 5-5.5 regardless of the pH of the injected buffer solution.[6] Therefore, the charge/residue in Fig. 3 was calculated using a pH of 5.25. In Fig. SI2, the charge/residue was calculated using the pH value given in Shastry *et al*. (pH 7.0).[5]

Calculation of $\Delta\Delta G_{u,max}$ *and* ΔT_s

For data of SH3, $\Delta\Delta G_{u,max} = \Delta G_{u,max,reverse\ micelle} - \Delta G_{u,max,buffer}$ and $\Delta T_s = T_{s,reverse\ micelle} - T_{s,buffer}$ which are shown in Fig. 2, could be calculated via $\Delta G_u(T_s) = \Delta G_{u,max}$ and T_s obtained from the global fit of the stability curves (Table SI3). Errors were calculated via Gaussian error propagation using the standard errors of $\Delta G_u(T_s)$ and T_s obtained from the weighted fit with respect to the uncertainty of $\Delta G_u(T)$ given in Senske *et al*.[4] SH3 cosolute data shown in Fig. 2 were not part of the global fit and were taken from Senske *et al*.[4]

For RNase T1, $\Delta\Delta G_{u,max}$ and ΔT_s were determined via T_m, $\Delta H_u(T_m)$, and ΔC_p. At T_s, $\Delta S_u = 0$ and T_s is given by Eq. SI1. Knowing T_s, $\Delta G_u(T_s)$ can be calculated via Eqs. 8 and 9. T_m, $\Delta H_u(T_m)$, and ΔC_p ($\Delta C_p = 1.2 \pm 0.2$ kcal mol^{-1} K^{-1}) were taken as given by Shastry *et al*.[5] A constant ΔC_p was used to calculate $\Delta\Delta G_{u,max}$ and ΔT_s. The difference of calculating T_s and $\Delta G_u(T)$ using ΔC_p or $\Delta C_p(T)$ are small.[7, 8] Furthermore, T_s of RNase T1 is at subzero temperatures at which $\Delta C_p(T)$ is difficult to be determined experimentally. Gaussian error propagation was used to calculate the errors of $\Delta\Delta G_{u,max}$ and ΔT_s using the error estimates of T_m and $\Delta H_u(T_m)$. The error of ΔC_p can be neglected in this case as we are using

the assumption $\Delta C_{p,reverse\ micelle} = \Delta C_{p,buffer}$ and are not interested in the absolute values of $\Delta\Delta G_{u,max}$ and ΔT_s but in the difference between the reverse micelle and the respective buffer solutions.

$$T_s = T_m e^{\left[-\Delta H_u(T_m)/\Delta C_p T_m\right]} \quad \text{(SI1)}$$

Calculation of $\Delta\Delta H_u$ *and* $\Delta\Delta S_u$

In Figs. 3, SI2 and SI3, the excess enthalpy ($\Delta\Delta H_u$) and entropy ($\Delta\Delta S_u$) are given. The excess parameter ($\Delta\Delta X_u$, $X = H$ or S) describes the difference between the protein stability data in the reverse micelle and the buffer solution: $\Delta\Delta X_u(T) = \Delta X_{u,reverse\ micelle}(T) - \Delta X_{u,buffer}(T)$. $\Delta X_u(T)$ is given by Eqs. 9 and 10. Data given in Figs. 3, SI2, and SI3 were calculated at $T_{m,buffer}$. Since ΔC_p is not largely affected by CTAB and 10MAG/LDAO reverse micelles,[4] we used the assumption $\Delta C_{p,reverse\ micelle}(T) = \Delta C_{p,buffer}(T)$ for all systems. Therefore, $\Delta\Delta H_u$ is temperature independent because $\partial\Delta\Delta H_u/\partial T = \Delta\Delta C_p = 0$ and the temperature dependence of $T\Delta\Delta S_u$ is given by $\partial T\Delta\Delta S_u/\partial T = \Delta C_p \ln\left(T_{s,buffer}/T_{s,reverse\ micelle}\right)$ for a temperature independent ΔC_p.

T_m and $\Delta H_u(T_m)$, or T_s and $\Delta H_u(T_s)$, as well as ΔC_p of SH3 were obtained by a global fit of the stability curves given in Senske *et al*. [4] (Table SI3). T_m, ΔH_u, and ΔC_p of RNase T1 were used as given by Shastry *et al*.[5] Errors of $\Delta\Delta X_u$ were calculated using Gaussian error propagation of the experimental uncertainties of T_m and ΔH_u which are given in the reference.[5] Errors of T_m and ΔH_u of SH3 were obtained by a global weighted fit of the stability curves given in Senske *et al*.[4] (Table SI3, Figs. SI3-SI5). The error of ΔC_p can be neglected in this case as we are using the assumption $\Delta C_{p,reverse\ micelle} = \Delta C_{p,buffer}$ and are not interested in the absolute values of ΔH_u and ΔS_u but the difference between reverse micelle and buffer solution.

Global fit of stability curves of SH3

Senske *et al*. showed that ΔC_p of SH3 encapsulated in CTAB and 10MAG/LDAO revere micelles and of SH3 in the unconfined buffer solution are nearly identical within the experimental uncertainty. Here, we used all SH3 data (reverse micelle and buffer solutions) given in Senske *et al*. to globally fit ΔC_p in order decrease the uncertainty of estimating ΔC_p.[4] We determined T_m, $\Delta H_u(T_m)$, T_s, $\Delta H_u(T_s)$ and ΔC_p of the different systems by a global weighted (with respect to the error bars) fit using a single ΔC_p value to describe the data of all systems (Table SI3). The fitted values of T_m, $\Delta H_u(T_m)$, T_s and $\Delta H_u(T_s)$ do not differ much compared to the published data and ΔC_p could be obtained with a higher accuracy (Table SI3).[4] All data with the corresponding fits are shown in Figs. SI3-SI5.

Table SI3: Fitting parameters of a global fit of Eq. 8 to all SH3 data (reverse micelle and buffer solutions) given in Senske *et al*. $\Delta H_u(T_m)$ and $\Delta H_u(T_s)$ are given in kcal/mol to ease the comparison with the published data.[4] The global ΔC_p was estimated to be 0.89 ± 0.04 kcal mol^{-1} K^{-1}. T_m and $\Delta H_u(T_m)$ of SH3 for CTAB, $W_0 = 20$, pH 7.6 and CTAB, $W_0 = 25$, pH 7.6 could not be determined because the corresponding stability curves do not cross the x-axis. The errors correspond to the standard error of each parameter obtained by the weighted fit.

	T_m/K	$\Delta H_u(T_m)$	T_s/K	$\Delta H_u(T_s)$
buffer, pH 7.6	306.5 ± 0.6	18.5 ± 0.9	286 ± 1	0.62 ± 0.05
CTAB, $W_0 = 15$, pH 7.6	272 ± 2	13 ± 3	258 ± 3	0.3 ± 0.2
CTAB, $W_0 = 20$, pH 7.6	NA	NA	261 ± 2	-0.2 ± 0.1
CTAB, $W_0 = 25$, pH 7.6	NA	NA	263 ± 3	-0.4 ± 0.2

buffer, pH 4.2	324.7 ± 0.9	37 ± 2	285 ± 3	2.3 ± 0.3
CTAB, $W_0 = 20$, pH 4.2	288.5 ± 0.3	22.1 ± 0.9	265 ± 2	0.93 ± 0.08
Buffer, pH 7.0	307.9 ± 0.3	18.4 ± 0.6	287.9 ± 0.4	0.60 ± 0.02
10MAG/LDAO, $W_0 = 15$, pH 7.0	302.4 ± 0.4	20 ± 1	281 ± 2	0.71 ± 0.08
10MAG/LDAO, $W_0 = 20$, pH 7.0	301.8 ± 0.5	19.5 ± 0.8	281 ± 1	0.69 ± 0.05
10MAG/LDAO, $W_0 = 25$, pH 7.0	300.9 ± 0.5	17.2 ± 0.5	282.1 ± 0.7	0.54 ± 0.02

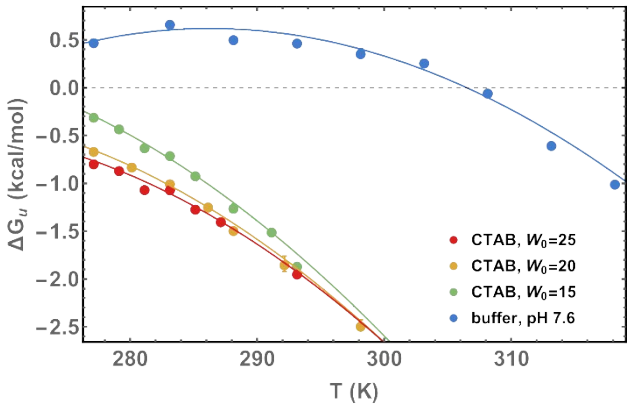

Figure SI3: Experimental data of SH3 in buffer solution and encapsulated in CTAB reverse micelles at pH 7.6 published by Senske *et al.*[4] Solid lines represent fitted stability curves obtained by a global fit of Eq. 8 of the main text to all datasets shown in Figs. SI3-SI5 using the constraint that ΔC_p is equal in all curves.

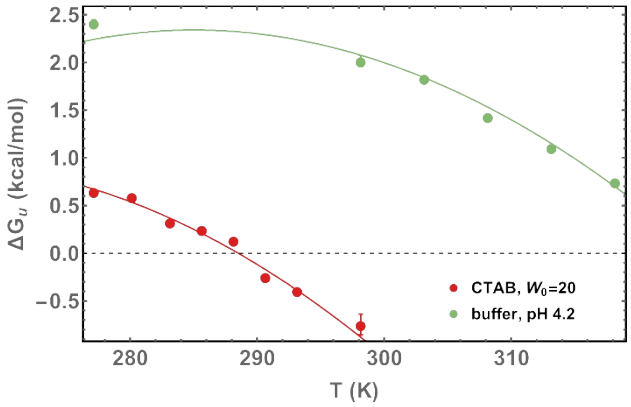

Figure SI4: Experimental data of SH3 in buffer solution and encapsulated in CTAB reverse micelles at pH 4.2 published by Senske *et al*.[4] Solid lines represent fitted stability curves obtained by a global fit of Eq. 8 of the main text to all datasets shown in Figs. SI3-SI5 using the constraint that $^{\Delta C_p}$ is equal in all curves.

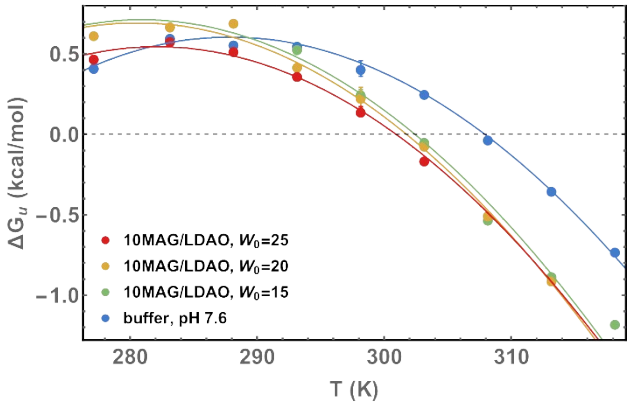

Figure SI5: Experimental data of SH3 in buffer solution and encapsulated in 10MAG/LDAO reverse micelles at pH 7.0 published by Senske *et al*.[4] Solid lines represent fitted stability curves obtained by a global fit of Eq. 8 of the main text to all datasets shown in Figs. SI3-SI5 using the constraint that $^{\Delta C_p}$ is equal in all curves.

References

1. N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 10037-10041.
2. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, **32**, W665-W667.
3. T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, *Nucleic Acids Res.*, 2007, **35**, W522-W525.
4. M. Senske, A. E. Smith and G. J. Pielak, *Angew. Chem. Int. Edit.*, 2016, **55**, 3586-3589.
5. M. C. R. Shastry and M. R. Eftink, *Biochemistry*, 1996, **35**, 4094-4101.
6. B. S. Marques, N. V. Nucci, I. Dodevski, K. W. C. Wang, E. A. Athanasoula, C. Jorge and A. J. Wand, *J. Phys. Chem. B*, 2014, **118**, 2020-2031.
7. P. L. Privalov, *Crit. Rev. Biochem. Mol. Biol.*, 1990, **25**, 281-305.
8. M. Senske, D. Constantinescu-Aruxandei, M. Havenith, C. Herrmann, H. Weingärtner and S. Ebbinghaus, *Phys. Chem. Chem. Phys.*, 2016, **18**, 29698-29708.