

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

March 07, 2015

# Influence of poly(ethylene glycol) on the micelle formation of alkyl maltosides used in membrane protein crystallization<sup>†</sup>

Frank Müh,<sup>‡\*</sup> Dörte DiFiore,<sup>§</sup> and Athina Zouni<sup>&</sup>

<sup>‡</sup>*Institut für Theoretische Physik, Johannes Kepler Universität Linz, Altenberger Strasse 69, A-4040 Linz, Austria,* <sup>§</sup>*Max-Volmer-Laboratorium für Biophysikalische Chemie, Technische Universität Berlin, Strasse des 17. Juni 135, D-10623 Berlin, Germany,* <sup>&</sup>*Institut für Biologie, Humboldt Universität zu Berlin, Leonor-Michaelis-Haus, Philippstrasse 13, D-10095 Berlin, Germany.*

\*To whom correspondence should be addressed. E-mail: frank.mueh@jku.at.

---

<sup>†</sup> Electronic supplementary information (ESI) available: Data concerning the analysis of titration curves (Tables 2–4), parameters used in molecular thermodynamic modeling (Tables 5, 6) as well as plots of transfer free energies (Fig. 9) and molecular surface areas (Fig. 10) of alkanes over the number of carbon atoms, density (Fig. 11) and molarity (Fig. 12) of buffer solutions and the molarity correction term (Fig. 13) as a function of PEG concentration.

## ABSTRACT

With the aim to better understand the phase behavior of alkyl maltosides (n-alkyl- $\beta$ -D-maltosides,  $C_nG_2$ ) under the conditions of membrane protein crystallization, we studied the influence of poly(ethylene glycol) (PEG) 2000, a commonly used precipitating agent, on the critical micelle concentration (CMC) of the alkyl maltosides by systematic variation of the number  $n$  of carbon atoms in the alkyl chain ( $n = 10, 11, 12$ ) and the concentration of PEG2000 ( $\chi$ ) in a buffer suitable for the crystallization of cyanobacterial photosystem II. CMC measurements were based on established fluorescence techniques using pyrene and 8-anilino-naphthalene-1-sulfonate (ANS). We found an increase of the CMC with increasing PEG concentration according to  $\ln(\text{CMC}/\text{CMC}_0) = k_p\chi$ , where  $\text{CMC}_0$  is the CMC in the absence of PEG and  $k_p$  a constant that we termed “polymer constant”. In parallel, we measured the influence of PEG2000 on the surface tension of the detergent-free buffer solutions. At PEG concentrations  $\chi > 1$  % w/v, the surface pressure  $\pi_s(\chi) = \gamma(0) - \gamma(\chi)$  was found to depend linearly on the PEG concentration according to  $\pi_s(\chi) = \kappa\chi + \pi_s(0)$ , where  $\gamma(0)$  is the surface tension in the absence of PEG. Based on a molecular thermodynamic modeling, CMC shifts and surface pressure due to PEG are related, and it is shown that  $k_p = \kappa c(n) + \eta$ , where  $c(n)$  is a detergent-specific constant depending *inter alia* on the alkyl chain length  $n$  and  $\eta$  is a correction for molarity. Thus, knowledge of the surface pressure in the absence of detergent allows for a prediction of the CMC shift. The PEG effect on the CMC is discussed concerning its molecular origin and its implications for membrane protein solubilization and crystallization.

## 1 Introduction

Detergents (or surfactants)<sup>1, 2</sup> are used in diverse applications ranging from detergency, paint formulations, food, and pharmacy to the biochemistry of membrane proteins.<sup>3</sup> In the latter field, nonionic amphiphiles, which possess a polar but uncharged headgroup, are of particular usefulness. They are mild and not only allow for isolation of membrane proteins in a functional state, but have proven essential in membrane protein crystallization, which is still a prerequisite for attaining structural information about larger enzymes.<sup>4</sup> Sugar surfactants<sup>5</sup> are made of a hydrophilic sugar headgroup and a hydrophobic *n*-alkyl chain. Among the detergents of widespread use in membrane protein research are those sugar surfactants, where the two parts are connected via an ether bond.<sup>6</sup> They are referred to as  $C_nG_s$  with *n* being the number of carbon atoms in the alkyl tail and *s* the number of sugar (usually hexose) units in the head. An advantage of these detergents in many applications is the relative temperature-insensitivity of their physico-chemical properties such as the critical micelle concentration (CMC), i.e., the detergent concentration, at which the detergent monomers start to form globular aggregates known as micelles. The significance of the CMC is that micelles are required to solubilize hydrophobic compounds in the aqueous phase, e.g., dyes<sup>7</sup> or membrane proteins.<sup>8</sup>

The *n*-alkyl- $\beta$ -D-maltosides ( $C_nG_2$ ), which are in the focus of the present study, have been used in many cases as solubilizing agents to achieve membrane protein crystallization. Representative examples of a successful application of  $C_{12}G_2$  are photosystem I (PSI)<sup>9</sup> and photosystem II (PSII)<sup>10-12</sup> as well as the cytochrome (*cyt*) *b<sub>6</sub>f* complex<sup>13</sup> of oxygenic photosynthesis, a bacterial multidrug efflux transporter,<sup>14</sup> a *cyt c* quinol dehydrogenase,<sup>15</sup> human leukotriene C<sub>4</sub> synthase,<sup>16, 17</sup> and the rotor of the V-type Na<sup>+</sup>-ATPase.<sup>18</sup> The homologous detergent  $C_{11}G_2$  was used in the case of the yeast *cyt bc<sub>1</sub>* complex,<sup>19, 20</sup>  $C_{10}G_2$  for a glutamate transporter,<sup>21</sup> and  $C_8G_2$  for a chloride channel.<sup>22</sup> In addition, mixtures of alkyl maltosides have been applied in the crystallization of bacterial *cyt c* oxidase,<sup>23</sup> fumarate reductase,<sup>24</sup> and mitochondrial respiratory complex II.<sup>25</sup>

Protein crystallization requires the addition of a precipitating agent to change the protein-protein interactions in a supersaturated solution such as to promote crystal assembly in favor of amorphous aggregation. In essentially all of the above cited work, the soluble polymer poly(ethylene glycol) (PEG) has been used for this purpose. Despite the extensive use of PEG in connection with alkyl maltosides, information about the effects of this cosolute on the aggregation behavior of  $C_nG_2$  is scarce. Among the few studies that we were able to find in the literature, one work reports the influence of PEG20K (average molecular weight of approximately 20,000 Da) on the CMC of  $C_{12}G_2$ , but only for one concentration of PEG.<sup>26</sup> In another paper, the CMC of  $C_8G_2$  at three different concentrations of PEG400 was determined in the context of membrane protein crystallization.<sup>27</sup> Both studies showed that PEG increases the CMC of  $C_nG_2$ . These results are in

agreement with studies of n-alkyl- $\beta$ -D-glucosides ( $C_nG_1$ ) reporting an increase of the CMC due to PEG3350 for  $n = 8$  and  $9$  (in the presence of  $0.5$  M ammonium sulfate)<sup>28</sup> and an increase of the CMC of mixtures of  $C_9G_1$  with n-octyl-2-hydroxyethylsulfoxide due to PEG2000.<sup>29</sup>

Yet, the CMC shift induced by the precipitating agent is usually not the central object of these investigations. Rather, the focus is on the effective interactions between detergent micelles and protein-detergent complexes (PDCs) induced by PEG. The interest in these interactions originates from the seminal work of George and Wilson<sup>30</sup> who used the second osmotic virial coefficient ( $B_{22}$ ) as a measure of intermolecular potential to define crystallization conditions for soluble proteins. They found the  $B_{22}$  values under crystallization conditions to fall into a fairly narrow range below zero, which they referred to as “crystallization slot” and which indicates small attractive interactions between protein molecules. Later, the existence of a crystallization slot for membrane proteins was demonstrated, and it was concluded that the detergent portion of the PDC is important in determining the  $B_{22}$  values.<sup>29, 31</sup> This finding implies that the detergent belt surrounding the hydrophobic part of the protein surface in the PDC remains fully intact in the course of crystallization and that detergent-detergent interactions (i.e., between detergent belts) are weakly attractive. Then, a simple picture emerges, according to which the PDC is a fixed entity behaving like a soluble protein and its aggregation behavior should be similar to and can be inferred from that of free (i.e., protein-free) micelles. However, there are several objections that can be raised against this simple picture: (i) Protein-protein interactions contribute to crystal contacts. Detergent belts in crystals usually show a high degree of disorder, and their role in stabilizing a crystal is difficult to assess. (ii) There are examples for membrane protein crystallization under conditions, where the detergent-detergent interactions are essentially absent<sup>32</sup> or even repulsive.<sup>27</sup> (iii) Many crystals show a packing that does not allow for the accommodation of a fully developed detergent belt between the protein molecules. Prototypical examples are the type-I crystals<sup>33</sup> of plant light-harvesting complex II<sup>34</sup> and the hexagonal crystal form of bacterial porins,<sup>35</sup> in which the protein seems to be completely stripped of bound detergent. In other cases, the detergent belt is partially removed or squeezed at certain contact points as in the trigonal crystal form of bacterial porins<sup>35</sup> or the monomeric cyanobacterial PSII.<sup>11</sup>

Very recently, it was demonstrated for the first time that detergent-rich type-II crystals can be transformed into detergent-depleted type-I crystals.<sup>36</sup> Normally, however, it is believed that type-I crystals are formed directly. The latter process requires crystallizing membrane proteins under conditions that promote the degradation of the detergent belt. Then, the question arises of how this can be accomplished without an amorphous precipitation of the protein. A prerequisite for an understanding of these complicated processes is a sound knowledge about the influence of the precipitating agent on the aggregation behavior of the detergent, as this will likely also affect the

formation of detergent belts and the location of crystallizations slots. To this end, we investigated CMC shifts of  $C_nG_2$  ( $n = 10, 11, 12$ ) caused by PEG2000 by using two different fluorescence probes. One is the well-established  $I_1/I_3$ -fluorescence ratio<sup>7, 26, 37, 38</sup> of the polycyclic aromatic hydrocarbon pyrene and the other the fluorescence enhancement of 8-anilino-naphthalene-1-sulfonate.<sup>39-41</sup> The observed CMC shifts are analysed on the basis of two different strategies of molecular thermodynamic modeling and related to the change of the surface tension of detergent-free buffer induced by PEG. It will be shown that knowledge of the surface pressure of PEG solutions can be used to predict CMC shifts. On the basis of these results, the role of the CMC and its relation to the critical solubilization concentration (CSC) of membrane proteins<sup>8</sup> under crystallization conditions will be discussed.

## 2 Materials and Methods

Detergents were purchased from Glycon (Luckenwalde, Germany), all other chemicals from Sigma-Aldrich, and used without further purification. All experiments were performed with buffered aqueous solutions containing 100 mM piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES), adjusted to pH 7.0 with NaOH, and 5 mM  $CaCl_2$  corresponding to the crystallization conditions of PSII.<sup>10</sup> Pyrene was suspended in buffer as in earlier work<sup>8</sup> to give an excimer-free stock solution of approximately 0.5  $\mu$ M. Fluorescence spectra were taken with a Horiba Jobin Yvon FluoroMax-2 spectrometer. The determination of the CMC was carried out by exploiting the  $I_1/I_3$ -fluorescence ratio<sup>7</sup> of pyrene as in earlier work<sup>8</sup> as well as by using the fluorescence enhancement of 8-anilino-naphthalene-1-sulfonate (ANS).<sup>41</sup> For pyrene, fluorescence spectra were recorded between 360 and 400 nm with the excitation wavelength set to 320 nm and an integration time of 0.3 s. In the case of ANS, the fluorescence was recorded in the range between 460 and 530 nm with an integration time of 0.5 s, and excitation was at 370 nm.

The surface tension of buffer solutions with various concentrations of PEG2000 was measured by drop shape analysis using a DataPhysics OCA 15 instrument with the pendant drop method and Young-Laplace fitting (drop size 10 – 25  $\mu$ L, needle diameter 1.83 mm). For density correction, the density of the same buffer solutions was determined by using a Mettler Toledo DA-100 M digital density meter. All measurements (and theoretical modeling, see Section 3.2) were performed at room temperature ( $T = 298$  K).

### 3 Results

#### 3.1 Determination of CMC

We determined the CMC of the alkyl maltosides as a function of PEG2000 concentration,  $\chi$ , by using two different fluorescence probes. ANS exhibits a marked increase of its fluorescence intensity with a peak maximum around 500 nm,  $I_{\text{ANS}}$ , in the presence of micelles.<sup>40</sup> This property can be used to determine the CMC as illustrated for  $\text{C}_{12}\text{G}_2$  in Fig. 1A. We observe  $I_{\text{ANS}}$  to remain essentially constant below a certain threshold value,  $C_{\text{ANS}}$ , of the total detergent concentration and to increase above that value.  $C_{\text{ANS}}$  is determined graphically by linear extrapolation of the constant part and the initial rise of the titration curve as shown in Fig. 1A. In Fig. 2A are shown various titration curves for  $\text{C}_{11}\text{G}_2$  at different concentrations of PEG2000. It can be seen that the rise of the curve above  $C_{\text{ANS}}$  becomes flatter with increasing PEG concentration, and  $C_{\text{ANS}}$  increases (ESI,† Tables 2–4). This behavior is observed for all  $\text{C}_n\text{G}_2$  (Fig. 3). In the following, we shall identify  $C_{\text{ANS}}$  with the CMC. For later analysis, we show in Fig. 4 the dependence of  $\ln(C_{\text{ANS}}/C_0)$  on  $\chi$ , where  $C_0$  is the CMC of the respective detergent in the used buffer in the absence of PEG. Within the error limits,  $\ln(C_{\text{ANS}}/C_0)$  depends linearly on the PEG concentration according to

$$\ln\left(\frac{C_{\text{ANS}}}{C_0}\right) = k_p \chi \quad (1)$$

where by analogy with the salt constant, we introduced the polymer constant  $k_p$ . We find  $k_p = 0.040 \pm 0.005$  ( $\% \text{ w/v}$ )<sup>-1</sup> for  $n = 10$ ,  $k_p = 0.045 \pm 0.002$  ( $\% \text{ w/v}$ )<sup>-1</sup> for  $n = 11$ , and  $k_p = 0.046 \pm 0.005$  ( $\% \text{ w/v}$ )<sup>-1</sup> for  $n = 12$ .

The richly structured fluorescence spectrum of pyrene is known to be sensitive to the polarity of the environment of the pyrene molecule. In particular, the intensity ratio of the first vibronic band at 370–372 nm to the third band at 381–383 nm ( $I_1/I_3$ -ratio) is decreased, when the pyrene molecule is transferred from the highly polar aqueous phase to the less polar interior of a micelle. This effect allows to monitor micelle formation via the  $I_1/I_3$ -ratio.<sup>7</sup> A typical titration curve for  $\text{C}_{12}\text{G}_2$  is shown in Fig. 1B. At low detergent concentration in the absence of micelles,  $I_1/I_3 \approx 1.7$ , whereas after micelle formation at high detergent concentration,  $I_1/I_3 \approx 1.2$ . In between, there is an interval of the detergent concentration, in which the  $I_1/I_3$ -ratio decreases gradually. There are different prescriptions in the literature for the extraction of the CMC from such a titration curve. In principle, there are three characteristic points in the curve: (i) the inflection point (denoted  $x_0$  in Fig. 1B), (ii) the interception of the tangent to the inflection point with the horizontal line representing the final  $I_1/I_3$ -ratio ( $x_2$ , green arrow in Fig. 1B), and (iii) the interception of the tangent to the inflection point with the horizontal line representing the initial  $I_1/I_3$ -ratio ( $x_1$ ). The characteristic detergent concentrations related to these points can be determined graphically as indicated by the

green lines in Fig. 1B. However, Aguiar et al.<sup>38</sup> suggested a more precise method that is based on an approximation of the titration curve by a Boltzmann-type sigmoid

$$y = \frac{y_1 - y_2}{1 + e^{(x-x_0)/\Delta x}} + y_2 \quad (2)$$

where the variables  $x$  and  $y$  are the total detergent concentration and the  $I_1/I_3$ -ratio, respectively,  $y_1$  and  $y_2$  are the upper and lower limits of the sigmoid,  $x_0$  is the inflection point, and  $\Delta x$  is the slope parameter. The meaning of the latter parameter can be illustrated by calculating the first derivative at  $x = x_0$ :

$$\left(\frac{dy}{dx}\right)_{x=x_0} = \frac{y_2 - y_1}{4\Delta x} \quad (3)$$

which is the slope of the tangent to the inflection point. This slope is negative for  $y_2 < y_1$  as in the case of  $y = I_1/I_3$ , and the decrease of  $y$  is steeper the smaller  $\Delta x$ . The tangent itself is given by the equation

$$y_t = \frac{y_1 + y_2}{2} + \frac{y_2 - y_1}{4\Delta x}(x - x_0) \quad (4)$$

Equating  $y_t$  with  $y_1$  and  $y_2$  allows for a straightforward determination of the interceptions of the tangent with the lines representing the initial and final  $I_1/I_3$ -ratio, respectively.<sup>38</sup>

$$x_{1,2} = x_0 \mp 2\Delta x \quad (5)$$

There is no consensus as to whether  $x_0$ ,  $x_1$  or  $x_2$  should be identified with the CMC. Comparisons with other methods for the determination of the CMC yielded a variety of assignments. Here, we compare with the ANS data and find

$$C_{\text{ANS}} = x_1 = x_0 - 2\Delta x \quad (6)$$

(Fig. 3 and ESI,† Tables 2–4). This result is in agreement with Aguiar et al.<sup>38</sup> who suggested  $\text{CMC} = x_1$  for nonionic detergents. We note that we used  $\text{CMC} = x_2$  (green arrow on Fig. 1B) in our earlier work.<sup>8</sup> A possible meaning of  $x_2$  in the context of membrane proteins is discussed below. In accordance with Aguiar et al.,<sup>38</sup> we find  $x_0/\Delta x \leq 10$  for the nonionic detergents (ESI,† Tables 2–4).

The effect of PEG2000 is to increase both  $x_0$  and  $\Delta x$  (ESI,† Tables 2–4) so that the titration curves are shifted to higher detergent concentrations and widened (Fig. 2B). The changes are such that  $x_1$  is increased in accordance with  $C_{\text{ANS}}$  (Fig. 3). Thus, the increase of the CMC due to PEG2000 as inferred from the ANS data is confirmed by the pyrene data.

### 3.2 Thermodynamic Modeling

3.2.1. *Modeling of  $C_0$* . The CMC is related to the free energy change per detergent molecule  $\Delta\mu_{\text{mic}}^0$  for the transfer of the detergent monomer into the micelle by

$$g_{\text{mic}} = \frac{\Delta\mu_{\text{mic}}^0}{k_{\text{B}}T} = \ln X_{\text{CMC}}, \quad (7)$$

where  $\text{CMC} = X_{\text{CMC}} \cdot c_{\text{tot}}(0)$ ,  $X_{\text{CMC}}$  is the mole fraction of detergent at the CMC,  $k_{\text{B}}$  is Boltzmann's constant,  $T$  the absolute temperature, and  $c_{\text{tot}}(0)$  the total molarity of the solution defined further below in Section 3.2.2. The application of eq. (7) implies the assumption that  $X_{\text{CMC}} \approx X_1$  with  $X_1$  being the mole fraction of detergent monomers at the CMC. Modeling of the CMC is possible by using the traditional molecular thermodynamic (TMT) approach<sup>42</sup> or the more recently proposed so-called computer simulation-molecular thermodynamic (CS-MT) approach.<sup>43</sup>

3.2.1.1. *TMT approach to  $C_0$* . In this model,  $g_{\text{mic}}$  is decomposed into different contributions:

$$g_{\text{mic}} = g_{\text{tr}} + g_{\text{int}} + g_{\text{pack}} + g_{\text{st}}. \quad (8)$$

Here, the transfer term  $g_{\text{tr}}$  determines the free energy change of transferring the alkyl chain of the detergent molecule from water into a liquid hydrocarbon phase representing the hydrophobic core of the micelle, the interfacial term  $g_{\text{int}}$  represents the free energy change associated with the formation of an interface between the hydrophobic alkyl tail or micelle core and the surrounding water, the packing term  $g_{\text{pack}}$  accounts for free energy changes due to packing constraints of the alkyl chains in the micelle core forcing them to adopt different conformations than in a pure hydrocarbon liquid, and the steric term  $g_{\text{st}}$  describes headgroup interactions in the micelles. The exponential dependence of the CMC on the alkyl chain length can be traced back to  $g_{\text{tr}}$ , which is a linear function of  $n$ , the number of carbon atoms in the alkyl chain:

$$g_{\text{tr}} = g_{\text{tr}}(\text{CH}_3) + g_{\text{tr}}(\text{CH}_2) (n - 1) \quad (9)$$

with  $g_{\text{tr}}(\text{CH}_3)$  and  $g_{\text{tr}}(\text{CH}_2)$  being the contributions of the methyl and methylene groups, respectively, to the transfer term. At this point, we note that an error has occurred in the theoretical estimation of the CMCs in ref. 8, where a factor of  $10^3$  got lost for unknown reasons. In fact, the values of  $g_{\text{mic}}$  deduced from the experimental CMCs based on eq. 7 and the values of  $g_{\text{tr}}$  calculated from eq. 9 differ by  $\sim 7$  (depending slightly on  $n$ ). Thus, if the drastic approximation  $g_{\text{mic}} \approx g_{\text{tr}}$  is made as in ref. 8, the CMC values are underestimated by three orders of magnitude, which demonstrates the importance of the other terms in eq. 8. Recently, Stephenson et al.<sup>43</sup> performed molecular dynamics (MD) simulations of alkane droplets in water and inferred values of  $g_{\text{tr}}$  for octane, dodecane and hexadecane. From these data, we can deduce  $g_{\text{tr}}(\text{CH}_3) = -3.72 \pm 0.07$  and  $g_{\text{tr}}(\text{CH}_2) = -1.45 \pm 0.02$  (ESI,† Fig. 9) in agreement with earlier estimates.<sup>42</sup> The values of  $g_{\text{tr}}$  calculated from these values on the basis of eq. 9 are denoted as  $g_{\text{tr}}(\text{TMT})$  in ESI,† Table 5.

The interfacial term is modeled as<sup>42</sup>

$$g_{\text{int}} = \frac{\sigma_{\text{hw}}}{k_{\text{B}}T} (A - A_0) \quad (10)$$

where  $\sigma_{\text{hw}}$  is the interfacial tension between the hydrophobic micellar core and the aqueous environment,  $A$  is the surface area of the micellar core per detergent molecule (see below), and  $A_0$  is the area per molecule of the core shielded from the aqueous phase by the sugar head group of the detergent. Following Nagarajan and Ruckenstein,<sup>42</sup> we choose  $\sigma_{\text{hw}}$  as the macroscopic interfacial tension between an aliphatic hydrocarbon (h) of the same molecular weight as the detergent's alkyl tail and water (w). This interfacial tension can be calculated in terms of the surface tensions  $\gamma_{\text{h}}$  and  $\gamma_{\text{w}}$ , respectively, of the hydrocarbon and water (i.e., liquid-air interfacial tensions) according to<sup>42, 44, 45</sup>

$$\sigma_{\text{hw}} = \gamma_{\text{h}} + \gamma_{\text{w}} - 2\Phi\sqrt{\gamma_{\text{h}}\gamma_{\text{w}}} \quad (11)$$

with  $\Phi \approx 0.55$  for  $n = 10 - 12$ ,

$$\gamma_{\text{h}} = 35.0 - 325 M^{-2/3} - 0.098 (T - 298) \quad (12)$$

and

$$\gamma_{\text{w}} = 72.0 - 0.16 (T - 298) \quad (13)$$

where  $T$  is the absolute temperature and  $M$  the molecular weight of the hydrocarbon in atomic mass units. From these eqs., we obtain  $\sigma_{\text{hw}} = (50 \pm 1)$  mN/m at  $T = 298$  K.

In contrast to other applications of the TMT approach, it is not our intention here to predict aggregation numbers and micellar shape. Rather, we use constraints from experimental information to model the various free energy contributions to  $C_0$ . Lipfert et al.<sup>46</sup> analysed the aggregation number  $m$  and micellar shape of  $\text{C}_{10}\text{G}_2$  and  $\text{C}_{12}\text{G}_2$  by means of small-angle X-ray scattering (SAXS). In particular, the forward scattering intensity was shown to be a rather direct measure of the aggregation number that is independent of form factor models. We use their values for  $\text{C}_{10}\text{G}_2$  and  $\text{C}_{12}\text{G}_2$  (ESI,† Table 5) together with an interpolated value for  $\text{C}_{11}\text{G}_2$ . The values are in good agreement with data from independent sources<sup>47, 48</sup> (see also ref. 8). The SAXS data suggest a model, in which the micellar core is an oblate spheroid with minor radius  $a$  and major radius  $b$ , and the shell of detergent head groups has a thickness of  $d = (6.15 \pm 0.15)$  Å (Fig. 5). The experimental values<sup>46</sup> of  $a$  and  $b$  for  $\text{C}_{10}\text{G}_2$  and  $\text{C}_{12}\text{G}_2$  are listed in ESI,† Table 5 together with the eccentricity (or ellipticity)

$$\varepsilon = \sqrt{1 - \frac{a^2}{b^2}} \quad (14)$$

Note that  $a$ ,  $b$ , and  $\varepsilon$  depend on  $n$ . Note also that  $a < l_{\text{c}} < b$ , where  $l_{\text{c}}$  is the maximum extension of an alkyl chain of length  $n$  given approximately as<sup>42</sup>  $l_{\text{c}} = (1.50 + 1.265 n)$  Å, but  $l_{\text{c}}$  is closer to  $a$  than to

$b$  (ESI,† Table 5). The molecular volume  $V_c$  of the alkyl chain can be calculated from the group contributions according to<sup>42</sup>

$$V_c = V(\text{CH}_3) + V(\text{CH}_2) (n - 1) \quad (15)$$

with  $V(\text{CH}_3) = 54.6 \text{ \AA}^3$  and  $V(\text{CH}_2) = 26.9 \text{ \AA}^3$  at  $T = 298 \text{ K}$ . It is found that  $m V_c = V_s$  within the error limits (ESI,† Table 5), where

$$V_s = \frac{4}{3} \pi a b^2 \quad (16)$$

is the volume of the spheroidal core. Thus, we can calculate  $a$  and  $b$  for  $n = 11$  by setting  $V_s = m V_c$  and assuming a value of  $\varepsilon = 0.86$  midway between the eccentricity values for  $\text{C}_{10}\text{G}_2$  and  $\text{C}_{12}\text{G}_2$ , resulting in  $a = (13.0 \pm 0.2) \text{ \AA}$  and  $b = (25.5 \pm 0.4) \text{ \AA}$ . From these data, we can determine the surface area of the micellar core as

$$A_c = 2\pi b^2 + \pi \frac{a^2}{\varepsilon} \ln\left(\frac{1+\varepsilon}{1-\varepsilon}\right) \quad (17)$$

With  $A = A_c/m$  and  $A_0 = 21 \text{ \AA}^2$  (see ref. 42) we have all we need to calculate  $g_{\text{int}}$  (see ESI,† Table 5).

For the analysis of the PEG effect in the framework of the TMT approach (see below), it is advantageous to treat  $g_{\text{tr}}$  on an equal footing with  $g_{\text{int}}$  by expressing it phenomenologically in terms of the interfacial tension  $\sigma_{\text{hw}}$  as

$$g_{\text{tr}} = -\frac{\sigma_{\text{hw}}}{k_{\text{B}}T} S = -\frac{\sigma_{\text{hw}}}{k_{\text{B}}T} [S(\text{CH}_3) + S(\text{CH}_2) (n - 1)] \quad (18)$$

where  $S$  is the molecular surface (solvent-excluded surface (SES) or Connolly surface<sup>49</sup>) of the hydrocarbon and  $S(\text{CH}_3)$  and  $S(\text{CH}_2)$  are the group contributions. Tuñón et al.<sup>50</sup> found a strictly linear correlation between the molecular surface and the transfer free energies of alkanes for  $n = 1 - 10$  with a slope of  $(69.2 \pm 1.1) \text{ cal}/(\text{mol \AA}^2)$ . This can be translated into  $(48 \pm 1) \text{ mN/m}$  in remarkable agreement with the value for  $\sigma_{\text{hw}}$  calculated above for  $n = 10 - 12$ . To obtain the group contributions, we can plot the surface area values over  $n - 2$  (ESI,† Fig. 10) resulting in  $S(\text{CH}_3) = 35.02 \text{ \AA}^2$  and  $S(\text{CH}_2) = 19.85 \text{ \AA}^2$ . However, the values of  $g_{\text{tr}}$  calculated in this way are consistently larger than  $g_{\text{tr}}(\text{TMT})$ . The reason is that the transfer free energies used in ref. 50 are somewhat larger than those in ref. 42 and 43. Since the smaller values have proven to be suitable for the modeling of micelle formation, we introduce here a correction factor  $\alpha$  to match the surface-based model of  $g_{\text{tr}}$  to the traditional model. We set  $\alpha = 0.71$  corresponding to the ratio of transfer free energies of octane, which is the only alkane that occurs in both, ref. 43 and 50. The resulting values of  $g_{\text{tr}}$  are labeled  $g_{\text{tr}}(\text{SES})$  in ESI,† Table 5. Thus, we can understand  $g_{\text{tr}}$  and  $g_{\text{int}}$  jointly as a term representing the contribution to  $g_{\text{mic}}$  due to a change of the effective molecular surface of hydrocarbons exposed to water:

$$g_{\text{tr}} + g_{\text{int}} = \frac{\sigma_{\text{hw}}}{k_{\text{B}}T} (A - A_0 - \alpha S) \quad (19)$$

The packing term  $g_{\text{pack}}$  is used to model the energetic and entropic consequences of a different conformational distribution of the alkyl chains in the micellar core compared to a liquid hydrocarbon phase. This is necessary, because the spatial constraints affect the conformational distribution and one end of the alkyl chain is required to stay at the surface of the micellar core due to its connection to the maltose head group. Nagarajan and Ruckenstein<sup>42</sup> proposed an empirical formula based on lattice models

$$g_{\text{pack}} = \frac{3\pi^2 R_s^2}{80NL^2} \quad (20)$$

Here,  $L = 4.6 \text{ \AA}$  is the lattice constant,  $N = (n + 1)/3.6$ , and  $R_s$  is the radius of the micellar core assumed to be a sphere. For simplicity, we assume this formula to be approximately valid also for an oblate spheroid with  $R_s = (ab^2)^{1/3}$  representing the radius of a sphere that has the same volume as the spheroid.

Finally,  $g_{\text{st}}$  is modeled by assuming a hard core repulsion interaction between headgroups at the micellar surface, resulting in<sup>42</sup>

$$g_{\text{st}} = -\ln\left(1 - \frac{A_p}{A}\right) \quad (21)$$

where  $A_p = 40 \text{ \AA}^2$  is the effective cross-sectional area of the maltose headgroup.

The  $C_0$  values calculated on the basis of  $g_{\text{tr}}(\text{SES})$  are closer to experiment than those calculated from  $g_{\text{tr}}(\text{TMT})$  (ESI,† Table 5). Thus, the decomposition of  $g_{\text{mic}}$  involving  $g_{\text{tr}}(\text{SES})$  is a reasonable basis for the further analysis of the PEG effect (see below).

*3.2.1.2. CS-MT approach to  $C_0$ .* Stephenson et al.<sup>43, 51, 52</sup> studied micelle formation by applying MD simulations. Based on their analysis, which they referred to as computer simulation-molecular thermodynamic (CS-MT) modeling approach, they proposed a different decomposition of  $g_{\text{mic}}$ :

$$g_{\text{mic}} = g_{\text{dehydr}} + g_{\text{hydr}} + g_{\text{pack}} + g_{\text{st}}. \quad (22)$$

in which the traditional terms  $g_{\text{tr}}$  and  $g_{\text{int}}$  are replaced with  $g_{\text{dehydr}}$  and  $g_{\text{hydr}}$ . The dehydration term  $g_{\text{dehydr}}$  accounts for the change in free energy associated with the dehydration of detergent monomers that accompanies the self-assembly into micelles. In the approach by Stephenson et al.,<sup>43</sup>  $g_{\text{dehydr}}$  is further decomposed according to

$$g_{\text{dehydr}} = \sum_{j=1}^{n_{\text{hydr}}} (1 - f_j) g_{\text{tr}}^{(j)} \quad (23)$$

where  $n_{\text{hydr}}$  is the total number of hydrophobic groups in the detergent molecule,  $(1 - f_j)$  is the fractional dehydration associated with group  $j$  upon micelle formation, and  $g_{\text{tr}}^{(j)}$  is the free energy

for the transfer of group  $j$  from the aqueous solution into the micelle. Hydrophilic groups are assumed to have a negligible effect on the dehydration free energy. Nonetheless, the contributions to  $g_{\text{dehydr}}$  do not solely originate from the alkyl tail, depending on the chemical nature of the headgroup. Finally, the hydration free energy term  $g_{\text{hydr}}$  in eq. 22 accounts for free energy changes experienced by water molecules that remain in a hydrating contact with the detergent upon micelle formation.

Stephenson et al.<sup>51</sup> analysed  $g_{\text{dehydr}}$  and  $g_{\text{hydr}}$  for  $C_{12}G_2$  on the basis of MD simulations. We use their values for  $C_{12}G_2$  as well. As there are no such data for the other two detergents and MD simulations are beyond the scope of the present work, we assume for simplicity that  $g_{\text{hydr}}$  is the same for all  $n$  between 10 and 12. Values of  $g_{\text{dehydr}}$  for  $C_{11}G_2$  and  $C_{10}G_2$  are approximated by subtracting from the value for  $C_{12}G_2$  once and twice, respectively, the term  $(1 - f) g_{\text{tr}}(\text{CH}_2)$ , where  $g_{\text{tr}}(\text{CH}_2)$  is the transfer free energy of a methylene group (see Subsection 3.2.1.1.) and  $f = 0.19$  is the average fractional hydration of a group in the alkyl tail as determined for  $C_{12}G_2$  from the MD simulations.<sup>51</sup> The values obtained in this way are denoted  $g_{\text{dehydr}}(\text{CS-MT})$  in ESI,† Table 6.

Similar to the treatment of  $g_{\text{tr}}$  above, it will be advantageous for the analysis of the PEG effect to express  $g_{\text{dehydr}}$  in terms of the interfacial tension  $\sigma_{\text{hw}}$ . This is possible by rewriting eq. 23 as

$$g_{\text{dehydr}} = -\frac{\alpha \sigma_{\text{hw}}}{k_{\text{B}} T} \sum_{j=1}^{n_{\text{hydr}}} h_j \quad (24)$$

where  $h_j = (1 - f_j) S_j$ ,  $\alpha$  is the correction factor and  $S_j$  is the molecular surface of group  $j$ . The sum is evaluated by averaging over similar groups in the detergent molecule:

$$\sum_{j=1}^{n_{\text{hydr}}} h_j = h(\text{CH}_3) + (n - 1)h_{\text{c}}(\text{CH}_2) + 2h_{\text{malt}}(\text{CH}_2) + 10h_{\text{malt}}(\text{CH}) \quad (25)$$

Here,  $h(\text{CH}_3)$  represents the final  $\text{CH}_3$  group of the alkyl tail,  $h_{\text{c}}(\text{CH}_2)$  the methylene groups in the alkyl tail,  $h_{\text{malt}}(\text{CH}_2)$  the methylene groups in the maltose headgroup and  $h_{\text{malt}}(\text{CH})$  the CH units of the sugar moieties in the maltose headgroup. The average  $f$  values for the different group types are derived from the data of Stephenson et al.,<sup>51</sup> resulting in  $h(\text{CH}_3) = 0.81 S(\text{CH}_3)$ ,  $h_{\text{c}}(\text{CH}_2) = 0.81 S(\text{CH}_2)$ ,  $h_{\text{malt}}(\text{CH}_2) = 0.17 S(\text{CH}_2)$ , and  $h_{\text{malt}}(\text{CH}) = 0.30 S(\text{CH})$  with  $S(\text{CH}_3)$  and  $S(\text{CH}_2)$  as above (see Section 3.2.1.1.) and  $S(\text{CH})$  estimated to  $5 \text{ \AA}^2$  by extrapolating the difference between  $S(\text{CH}_3)$  and  $S(\text{CH}_2)$ . The values obtained in this way are denoted  $g_{\text{dehydr}}(\text{CS-SES})$  in ESI,† Table 6.

The  $C_0$  values calculated on the basis of  $g_{\text{dehydr}}(\text{CS-SES})$  are clearly closer to experiment than those calculated from  $g_{\text{dehydr}}(\text{CS-MT})$  (ESI,† Table 6), but the quality of the data depends on the estimate of  $S(\text{CH})$  as well as the way of calculating  $g_{\text{pack}}$  and  $g_{\text{st}}$ . Note that we use different values of  $g_{\text{pack}}$  and  $g_{\text{st}}$  than Stephenson et al.,<sup>51</sup> which explains the difference between our value of  $C_0(\text{CS-MT})$  for  $C_{12}G_2$  and their value of  $(0.14 \pm 0.01) \text{ mM}$ .

**3.2.2. Modeling of  $k_{\text{p}}$ .** The CMC shift as a function of PEG concentration  $\chi$  can be related to the shift in free energy change by

$$\ln\left(\frac{C_{\text{ANS}}}{C_0}\right) = \ln\left(\frac{X_{\text{CMC}}(\chi)}{X_{\text{CMC}}(0)}\right) + \ln\left(\frac{c_{\text{tot}}(\chi)}{c_{\text{tot}}(0)}\right) \quad (26)$$

where  $c_{\text{tot}}(\chi)$  and  $c_{\text{tot}}(0)$  are the total molarities of the buffer in the presence and absence, respectively, of PEG. Neglecting the molarities of other components of the solution besides the buffer (PIPES), we have

$$c_{\text{tot}}(\chi) = \frac{1}{M_{\text{wat}}}(\rho - c_{\text{PIP}}M_{\text{PIP}}) + \left(\frac{1}{M_{\text{PEG}}} - \frac{1}{M_{\text{wat}}}\right)\chi + c_{\text{PIP}} \quad (27)$$

Here,  $\rho$  is the mass density of the solution (see ESI,† Fig. 11),  $c_{\text{PIP}}$  the molar concentration of the buffer and  $M_{\text{wat}}$ ,  $M_{\text{PIP}}$ , and  $M_{\text{PEG}}$  are the molar masses of water, PIPES, and PEG2000 respectively.

For  $\chi = 0$ , we have

$$c_{\text{tot}}(0) = \frac{\rho}{M_{\text{wat}}} + \left(1 - \frac{M_{\text{PIP}}}{M_{\text{wat}}}\right)c_{\text{PIP}} \quad (28)$$

that is,  $c_{\text{tot}}(0) = 54.68 \text{ mol L}^{-1}$  for 100 mM PIPES.  $c_{\text{tot}}(\chi)$  is a linear function of  $\chi$  (ESI,† Fig. 12).

For  $\chi \leq 25 \text{ \% w/v}$ , the logarithmic correction term in eq. (26) is approximately proportional to  $\chi$  with proportionality constant  $\eta = -0.01 \text{ (\% w/v)}^{-1}$  (ESI,† Fig. 13), so that

$$\ln\left(\frac{C_{\text{ANS}}}{C_0}\right) = \frac{\Delta\mu_{\text{mic}}^0(\chi) - \Delta\mu_{\text{mic}}^0(0)}{k_{\text{B}}T} + \ln\left(\frac{c_{\text{tot}}(\chi)}{c_{\text{tot}}(0)}\right) = g_{\text{mic}}(\chi) - g_{\text{mic}}(0) + \eta\chi \quad (29)$$

where  $g_{\text{mic}}(\chi) = \ln X_{\text{CMC}}(\chi)$  and  $g_{\text{mic}}(0) = \ln X_{\text{CMC}}(0)$  are the free energy changes in the presence and absence, respectively, of PEG. In the following, we will evaluate the PEG effect based on the decomposition of  $g_{\text{mic}}$  according to the TMT and the CS-MT approach. In both cases, we assume that the geometric characteristics of the micelles are not affected by PEG, so that  $g_{\text{pack}}$  and  $g_{\text{st}}$  remain unchanged.

*3.2.2.1. TMT approach to  $k_{\text{P}}$ .* In this model, the CMC change is due a shift of  $g_{\text{tr}} + g_{\text{int}}$ , which according to eq. (19) can be expressed as

$$\ln\left(\frac{C_{\text{ANS}}}{C_0}\right) = \frac{\Delta\sigma_{\text{hw}}(\chi)}{k_{\text{B}}T}(A - A_0 - \alpha S) + \eta\chi \quad (30)$$

where  $\Delta\sigma_{\text{hw}} = \sigma_{\text{hw}}(\chi) - \sigma_{\text{hw}}(0)$  is the change in interfacial tension due to PEG and  $\sigma_{\text{hw}}(0)$  is given by eq. (11). As shown in Section 3.3, the surface tension  $\gamma_{\text{w}}$  depends linearly on  $\chi$  for PEG concentrations  $\chi > 1 \text{ \% w/v}$ :

$$\gamma_{\text{w}}(\chi) = \gamma_{\text{w}}(0) - \pi_{\text{s}}(0) - \kappa\chi \quad (31)$$

where  $\pi_{\text{s}}(0)$  and  $\kappa$  are constants characterizing the linear part of the function (see Section 3.3). We assume the same linear dependence to be valid for  $\sigma_{\text{hw}}(\chi)$ , i.e., we make the ansatz

$$\sigma_{\text{hw}}(\chi) = \sigma_{\text{hw}}(0) - \kappa\chi \quad (32)$$

with the same constant  $\kappa$  as in eq. (31). This assumption is discussed in Section 4.1. The analysis of experimental data based on eq. (31) yields  $\kappa = 0.179 \pm 0.009 \text{ mN m}^{-1} (\% \text{ w/v})^{-1}$ . Thus,

$$\Delta\sigma_{\text{hw}}(\chi) = -\kappa\chi \quad (33)$$

Inserting into eq. (30) and comparing with eq. (1) results in a formula for the polymer constant:

$$k_{\text{p}} = \frac{A - A_0 - \alpha S}{k_{\text{B}}T} \kappa + \eta = \frac{g_{\text{tr}}(0) + g_{\text{int}}(0)}{\sigma_{\text{hw}}(0)} \kappa + \eta \quad (34)$$

where  $g_{\text{tr}}(0)$  and  $g_{\text{int}}(0)$  are the values for  $g_{\text{tr}}(\text{SES})$  and  $g_{\text{int}}$ , respectively, for  $\chi = 0$  as listed in ESI,† Table 5 (The sum  $g_{\text{tr}}(0) + g_{\text{int}}(0)$  is listed in Table 1). The values of  $k_{\text{p}}$  calculated in this way ( $k_{\text{p}}(\text{TMT})$  in Table 1) are in remarkable agreement with the experimental values  $k_{\text{p}}(\text{exp.})$  and show the same tendency to slightly increase with increasing  $n$ .

3.2.2.2. *CS-MT approach to  $k_{\text{p}}$* . Here, we make the assumption that the CMC change is exclusively due to a shift of  $g_{\text{dehydr}}$ . Accordingly, we write

$$\ln\left(\frac{C_{\text{ANS}}}{C_0}\right) = -\frac{\alpha \Delta\sigma_{\text{hw}}}{k_{\text{B}}T} \sum_{j=1}^{n_{\text{hydr}}} h_j + \eta\chi \quad (35)$$

With eq. (33), this results in

$$k_{\text{p}} = -\frac{\alpha}{k_{\text{B}}T} \sum_{j=1}^{n_{\text{hydr}}} h_j \kappa + \eta = \frac{g_{\text{dehydr}}(0)}{\sigma_{\text{hw}}(0)} \kappa + \eta \quad (36)$$

where  $g_{\text{dehydr}}(0)$  corresponds to  $g_{\text{dehydr}}(\text{CS-SES})$  for  $\chi = 0$  as listed in ESI,† Table 6 and Table 1. The calculated values of  $k_{\text{p}}$  ( $k_{\text{p}}(\text{CS-MT})$  in Table 1) are slightly higher than those of  $k_{\text{p}}(\text{TMT})$  and the experimental values  $k_{\text{p}}(\text{exp.})$ .

### 3.3 Effect of PEG2000 on surface tension

To obtain a model for the dependence of  $\sigma_{\text{hw}}$  on  $\chi$ , we measured the surface tension of the used buffer as a function of  $\chi$ . The result is shown in Fig. 6. It can be seen that there is a very steep decrease of  $\gamma_{\text{w}}$  for small values of  $\chi$  followed by a flatter linear decrease at higher PEG2000 concentrations. Thus, the surface pressure  $\pi_{\text{s}}(\chi) = \gamma_{\text{w}}(0) - \gamma_{\text{w}}(\chi)$  is characterized by a straight line with slope  $\kappa$  and intercept  $\pi_{\text{s}}(0)$  for  $\chi > 1 \text{ \% w/v}$ :

$$\pi_{\text{s}}(\chi) = \gamma_{\text{w}}(0) - \gamma_{\text{w}}(\chi) = \kappa\chi + \pi_{\text{s}}(0) \quad (\chi > 1 \text{ \% w/v}) \quad (37)$$

Here, the constant  $\pi_{\text{s}}(0)$  accounts for the strong decrease of  $\gamma_{\text{w}}(\chi)$  in the range  $\chi < 1 \text{ \% w/v}$ . (The blue line shown in Fig. 6 is  $\gamma_{\text{w}}(0) - \pi_{\text{s}}(\chi)$ , i.e., eq. (31).) As discussed below, we use the linear dependence characterized by the slope  $\kappa$  as model for the dependence of the interfacial tension  $\sigma_{\text{hw}}$  on the PEG concentration  $\chi$ .

## 4 Discussion

### 4.1 Modeling of the polymer constant and its relation to the surface tension

In order to understand the effect of PEG2000 on the CMC of alkyl maltosides, we performed a molecular thermodynamic modeling following two different strategies, the traditional MT<sup>42</sup> and the more recent CS-MT.<sup>43, 51, 52</sup> In both approaches, the free energy for micelle formation,  $g_{\text{mic}}$  (in units of  $k_{\text{B}}T$ ), which is the relevant thermodynamic quantity that determines the CMC, is decomposed into various contributions. Only part of these contributions depend on the hydrophobic effect. Upon micelle formation, detergent molecules are dehydrated in the sense that fewer hydrophobic groups of one detergent molecule are exposed to the aqueous phase in the micelle compared to the detergent monomer. Hence, the hydrophobic effect is central to the understanding of the CMC. Those parts of  $g_{\text{mic}}$  that represent this dehydration are  $g_{\text{tr}} + g_{\text{int}}$  in TMT (eq. (19)) and  $g_{\text{dehydr}}$  in CS-MT (eq. (23)).

Our modeling of the polymer constant  $k_{\text{p}}$ , characterizing the PEG effect on the CMC, is based on the assumption that PEG influences the interfacial tension  $\sigma_{\text{hw}}$  related to the molecular surface of hydrophobic groups in the detergent molecule in the same way as it influences the surface tension  $\gamma_{\text{w}}$  of the solution at PEG concentrations  $\chi > 1$  % w/v. This assumption results in a fairly good agreement between measured and calculated polymer constants (Table 1). It allows relating  $\kappa$  and  $k_{\text{p}}$  in a simple way via the terms of  $g_{\text{mic}}$  that depend on  $\sigma_{\text{hw}}$ . The details of this relation depend on the type of molecular thermodynamic modeling employed, but in both approaches, TMT and CS-MT, the essential result is

$$k_{\text{p}} = \kappa c(n) + \eta \quad (38)$$

where  $c(n)$  is a detergent-specific constant that depends *inter alia* on the alkyl chain length  $n$  (see eq. (34) for TMT and eq. (36) for CS-MT) and  $\eta$  is a correction for molarity valid for PEG concentrations  $\chi \leq 25$  % w/v. A practical consequence of eq. (38) is that knowledge of the surface pressure of buffers with varying PEG concentrations (eq. (37)) entails the possibility to predict CMC shifts in these buffers via eq. (1). In the following sections 4.2 and 4.3, we shall discuss, why the CMC shifts might be of importance in the context of membrane protein research.

Nonetheless, a correlation between  $k_{\text{p}}$  and  $\kappa$  according to eq. (38) is not self-evident. To understand why, we shall have a closer look on the underlying thermodynamics. Assuming that only water (component 1) and PEG (component 2) enter the buffer-air interface with mole numbers  $n_1^{\text{s}}$  and  $n_2^{\text{s}}$  in the surface and chemical potentials  $\mu_1$  and  $\mu_2$ , the Gibbs-Duhem equation for this surface at constant temperature reads

$$-A_{\text{w}}d\gamma_{\text{w}} = n_1^{\text{s}}d\mu_1 + n_2^{\text{s}}d\mu_2 \quad (39)$$

By dividing eq. (39) through by the surface area  $A_{\text{w}}$ , we obtain

$$-d\gamma_w = \Gamma_1 d\mu_1 + \Gamma_2 d\mu_2 \quad (40)$$

where  $\Gamma_i = n_i^s / A_w$  are the surface concentrations. For the bulk phase, the Gibbs-Duhem equation at constant temperature reads

$$VdP = n_1^b d\mu_1 + n_2^b d\mu_2 \quad (41)$$

with  $n_i^b$  being the mole number of component  $i$  in the bulk liquid. At constant pressure, the relationship

$$d\mu_1 = -\frac{n_2^b}{n_1^b} d\mu_2 \quad (42)$$

results, which upon insertion into eq. (40) yields

$$-d\gamma_w = \left( \Gamma_2 - \Gamma_1 \frac{n_2^b}{n_1^b} \right) d\mu_2 = \left( \Gamma_2 - \Gamma_1 \frac{c_2}{c_1} \right) d\mu_2 \quad (43)$$

where  $c_1$  and  $c_2$  are the molar concentrations of water and PEG, respectively, in the bulk. Note that  $c_2 = \beta\chi$ , where  $\beta = 10/2000$  (mol L<sup>-1</sup>)/(% w/v) for PEG2000. Introducing the partition coefficients  $K_i = \Gamma_i/c_i$  and assuming  $d\mu_2 \approx RTd\ln c_2$ , we obtain

$$-\frac{d\gamma_w}{d\chi} = RT\beta (K_2 - K_1) = \kappa \quad (44)$$

Hence,  $\kappa$  is a constant, if  $K_2 - K_1$  is a constant (at constant  $T$ ). This is at least the case, if  $K_1$  and  $K_2$  are independently constants. In other words, the surface tension becomes an approximately linear function of  $\chi$  at higher PEG concentrations, if the ratio of surface to bulk concentrations becomes approximately fixed for both components. Note that  $K_2 > K_1$  in order for  $\kappa$  to be positive. This implies a significant excess of PEG in the surface. For  $K_2 = 0$  (no PEG in the surface), the surface tension would increase with increasing PEG concentration and then could be traced back to the decreasing water activity in the bulk solution. The decrease of the surface tension with increasing PEG concentration indicates that PEG is located in the surface.

We can perform a similar analysis for the hydrophobic surface of a molecule like a detergent monomer in the PEG solution, resulting in

$$-\frac{d\sigma_{hw}}{d\chi} = RT\beta (K_2^\sigma - K_1^\sigma) \quad (45)$$

Here,  $K_1^\sigma$  simply represents the partitioning of water between the bulk phase and the hydration shell of the molecule. However,  $K_2^\sigma$  then represents the partitioning of PEG between the bulk phase and the hydration shell. Thus,  $K_2^\sigma$  actually is an association constant that describes the interaction of detergent monomers with PEG molecules in a way that water molecules of the hydration shell are displaced. The detergent-PEG association results in a “pre-dehydration” of the hydrophobic

molecular surface of the detergent monomer, so that the free energy gain of forming micelles is reduced and the CMC increased. This effect is modeled indirectly via the terms of  $g_{\text{mic}}$  that depend on the interfacial tension. The actual problem of this approach is not eq. (45), but the use of the same  $\kappa$  for both, surface and interfacial tension, for this implies

$$K_2 - K_1 = K_2^\sigma - K_1^\sigma \quad (46)$$

This equation states that the partitioning of both components is the same for the macroscopic buffer-air interface and the microscopic molecular interface. It is not clear, why this should be true. Note that if  $K_2 = K_2^\sigma = 0$  (no component 2 in any surface), the use of the same  $\kappa$  would be easier to justify, since it could be explained with the decreasing water activity at increasing concentration of component 2 in the bulk. This is probably true for component 2 being an electrolyte. However, in the case of PEG, we have to explain why  $K_2 \approx K_2^\sigma \neq 0$ .

One possible explanation for the observed dependence of  $\gamma_w$  on the PEG concentration (Fig. 6) is that at  $\chi > 1$  % w/v the surface is saturated with PEG in the sense that any further PEG molecule entering the surface is actually adsorbed to another PEG molecule. Then, the driving force for going into the surface, and hence  $K_2$ , is determined by interactions between PEG molecules that likely occur between ethylene units. This interaction is probably similar to that between an ethylene unit of PEG and the alkyl tail of a detergent monomer, so that the driving force for going into the surface is similar to that for detergent-PEG association and  $K_2 \approx K_2^\sigma$ . At present, this explanation is hypothetical, and further work needs to be done to evaluate it. Other issues to be addressed in future work are the following:

*4.1.1. Analysis of titration curves.* The determination of the CMC is based on a phenomenological assignment of characteristic points in the function  $y(x)$  to the CMC, where  $x$  is the total detergent concentration and  $y$  is a suitable observable such as the fluorescence intensity or a fluorescence ratio. The relationship between these characteristic points and the explicit concentrations of particular molecular species in the sample such as detergent monomers, micelles or detergent aggregates of various sizes is usually unknown. It is therefore not surprising that methods employing different observables often yield different values for the CMC (see, e.g., the discussion by Al-Soufi et al.<sup>53</sup>). Consequently, any attempt to link experimental CMCs to the free energy of micellization as in eq. (7) bears the risk that the measured quantity is not exactly what is described by  $g_{\text{mic}}$ . To obtain a better quantitative analysis of CMC data, it will be necessary in the future to properly define an observable-independent CMC based on a thermodynamic analysis of self-assembly and then, in a second step, to link this CMC to specific observables by explicit modeling of the experimental conditions.

4.1.2. *Molecular thermodynamic modeling.* Even if the relationship between experimental CMCs and the free energy of micellization is clarified, there remains the formidable task to link  $g_{\text{mic}}$  to molecular properties. The traditional MT approach<sup>42</sup> is a basis for a decomposition of  $g_{\text{mic}}$  into contributions that depend on the hydrophobic effect and those that do not and for a description of the PEG effect in terms of the former. However, in TMT, concepts from macroscopic physics are applied that are not necessarily suitable at a molecular scale. The more recent CS-MT approach<sup>43, 51, 52</sup> offers a promising alternative as it helps to link  $g_{\text{mic}}$  more directly to molecular properties based on MD simulations. Nonetheless, the CS-MT approach is not yet fully developed, and more simulations will have to be done to improve its accuracy. For example, a necessary next step is to perform explicit MD simulations of  $C_{10}G_2$  and  $C_{11}G_2$  in comparison to  $C_{12}G_2$ .

4.1.3. *Understanding the surface tension of PEG solutions.* Adsorption of PEG into the air-water interface is normally studied at very low PEG concentrations (see, e.g., Gilányi et al.<sup>54</sup>), and information about the surface composition of highly concentrated PEG solutions is lacking. The latter would help to develop a clear theoretical model of the correlation between  $k_p$  and  $\kappa$ .

## 4.2 Implications for membrane protein solubilization

In ref. 8, we studied the resolubilization of a detergent-depleted membrane protein in aqueous solution. The investigated cyanobacterial PSI has the peculiar property of changing its fluorescence spectrum upon formation of the detergent belt due to a response of protein-bound pigments to the detergent-protein interaction by a yet unknown mechanism. This effect was quantitatively analyzed for its dependence on the detergent and protein concentrations in conjunction with turbidimetry and pyrene fluorescence. In this way, it became possible to correlate changes of the  $I_1/I_3$ -ratio of pyrene with the resolubilization of the membrane protein. The titration curves in the presence of PSI are similar to those in its absence, but shifted systematically to higher detergent concentrations with increasing protein concentration. The point  $x_2$  in the titration curves correlates with the detergent concentration, where the turbidity approaches its lowest value indicating the resolubilization of the membrane protein. Hence, it was assigned to the critical solubilization concentration (CSC, see Fig. 7). On the basis of the present data, we interpret  $x_1$  as the detergent concentration, where micelles start to form, that is, the CMC. As argued in ref. 8, the free energy of the formation of detergent belts should be very similar to  $g_{\text{mic}}$ . Therefore, we interpret the  $I_1/I_3$ -titration curves in the presence of membrane protein in the following way: At  $x_1$  (the CMC), micelles and belts start to form. The latter implies that PDCs start to form. In the region between  $x_1$  and  $x_2$ , the concentrations of PDCs and micelles increase. At  $x_2$  (the CSC), the formation of PDCs reaches its limit, i.e., the limit of saturation of the hydrophobic (originally membrane-spanning) protein surfaces with detergent. Hence,  $x_2$  is the critical detergent concentration for membrane protein solubilization. Above  $x_2$ , the concentration of PDCs remains essentially constant,

but the concentration of (protein-free) micelles may further increase. According to this interpretation, there is not enough aggregated detergent in the concentration range between  $x_1$  and  $x_2$  to fully saturate the hydrophobic protein surfaces. Therefore, this range might be suitable for the formation of type-I crystals as discussed in Section 4.3.

An important result of ref. 8 is that the CSC shows a logarithmic dependence on the protein concentration  $c_{\text{prot}}$  similar to the dependence of the CMC on the PEG concentration observed here:

$$\ln \left( \frac{\text{CSC}}{\text{CMC}_0} \right) = k_{\text{prot}} c_{\text{prot}} \quad (47)$$

In this formula,  $\text{CMC}_0$  is the CMC in the absence of protein and  $k_{\text{prot}}$  a constant (the ‘‘protein constant’’, termed  $\tilde{n}_0$  in ref. 8). Note that the CSC is associated with  $x_2$ , but  $x_1$  depends on  $c_{\text{prot}}$  as well, albeit weaker. The behavior is qualitatively the same as observed for PEG (Fig. 3). The weak dependence of  $k_{\text{prot}}$  on the alkyl chain length  $n$  of the detergent  $\text{C}_n\text{G}_2$  could not be resolved, but likely exists. The important point is that such an influence of the membrane protein on the CSC and CMC cannot be explained with the binding of detergent molecules to the hydrophobic (originally membrane-spanning) protein surfaces, as this would require the formation of too large detergent belts for smaller  $n$  (in particular,  $n = 10$ , see ref. 8). Instead, the logarithmic dependence on  $c_{\text{prot}}$  is a sign of a change of the free energy of detergent aggregation (into belts or micelles) by the protein in a way similar to the change of  $g_{\text{mic}}$  by PEG. Therefore, we suggest that proteins influence the free energy of micellization by binding detergent monomers to their surface in addition to the formation of the detergent belt. Then, also a water-soluble protein should increase the CMC of a nonionic detergent. We are presently testing this hypothesis with further experiments.

### 4.3 Implications for membrane protein crystallization

In Fig. 8A is shown a qualitative phase diagram of a membrane protein solution containing detergent and PEG with the CMC and CSC as a function of PEG concentration indicated according to our present knowledge. Note that CMC and CSC depend exponentially on the PEG concentration  $\chi$  with  $\text{CMC} = \text{CMC}_0 \exp[k_{\text{P}}\chi]$  and  $\text{CSC} = \text{CSC}_0 \exp[k'\chi]$ , where  $\text{CSC}_0$  is the CSC in the absence of PEG and depends on the protein concentration, and  $k' > k_{\text{P}}$ . The dependence of the CSC on  $\chi$  is still hypothetical, but preliminary experiments on PSI performed as in ref. 8 suggest that PEG increases the CSC (F. Müh, D. DiFiore, A. Zouni, unpublished data). Type-II crystals contain a fully developed detergent belt and, therefore, can be expected to form in a region of the phase diagram, where the membrane protein is fully solubilized, i.e., above the CSC curve (Fig. 8B, left). Since at a given protein concentration, the interactions between PDCs suitable for crystal formation depend on the PEG concentration, there should be a crystallization slot according to George and Wilson,<sup>30</sup> i.e., a certain range of  $\chi$ , where  $B_{22}$  is slightly negative and crystals form preferentially. For type-II

crystals we call this range “slot II”. Very likely, the concentration of protein-free micelles is increased with increasing detergent concentration above the CSC. If free micelles perturb crystallization, the optimal region for the formation of type-II crystals should lie in slot II right above the CSC curve as indicated by the green diamond in Fig. 8A (see also Fig. 8B, left). In a batch method, the detergent and PEG concentrations have to be adjusted to these optimal values. In a vapor diffusion method (hanging or sitting drop), the concentrations of solutes are gradually increased and the starting conditions have to be adjusted so that the system develops as indicated by the green arrow in Fig. 8A to reach the optimal region in the phase diagram.

Type-I crystals are depleted of detergent and thus can form directly only in a region of the phase diagram, where the detergent belts are destabilized. This is likely the region between the CSC and the CMC curves in Fig. 8A, i.e., below the CSC curve (see also Fig. 8B, right). There is also likely a certain range of PEG concentrations optimal for crystal formation, which we shall term “slot I”. The optimal region is indicated by a blue diamond in Fig. 8A. Since the membrane protein tends to aggregate below the CSC, a batch method might be problematic and a vapor diffusion method could be advantageous, in which the system starts above the CSC with stable detergent belts and is then driven slowly below the CSC to destabilize the detergent belts (blue arrow in Fig. 8A).

According to these considerations, the CSC curve plays a crucial role in both, type-I and type-II crystallization. The precipitating agent (here: PEG) plays two roles: It not only influences the protein-protein interaction, but also the stability of the detergent belt. It might, therefore, be possible to optimize crystallization conditions by tuning the detergent concentration taking the effect of the precipitating agent into account. However, there are two problems in practical applications. (i) The CSC curve needs to be known at protein concentrations suitable for crystallization. It is possible that the fluorescence assay suggested here may not be applicable under these conditions (e.g., if protein-bound chromophores obscure the pyrene fluorescence) and more involved techniques (e.g., dynamical light scattering) have to be used to determine the solubility boundary of the membrane protein. (ii) Membrane protein samples contain an usually unknown amount of detergent originating from the preparation and concentrating procedures, which perturbs the determination of phase diagrams, and methods for detergent quantification are required (e.g., thin layer chromatography). For the moment, the only advice that can be given is to consider the detergent as a critical component in the crystallization setup, whose concentration should be carefully adjusted in screening experiments.

Recently, it was demonstrated that type-I crystals can be obtained *indirectly*, i.e., by transformation of initially formed type-II crystals<sup>36</sup> (red arrow in Fig. 8B). So far, the observation of such a crystal transformation is unique, and it is unclear at present, whether it can be considered a general new route to type-I crystals. The transformation requires an extraction of detergent from the

crystal, which is supported by PEG. More precisely, type-II crystals of cyanobacterial PSII were incubated successively in detergent-free solutions with stepwise increasing concentrations of PEG 5000 monomethyl ether (MME).<sup>36</sup> The transformed crystals showed a significant improvement of the diffraction quality, which is likely due to a tightening of protein contacts in the crystal between originally membrane-spanning surfaces (Fig. 8B). This result demonstrates the potential advantage of type-I over type-II crystals for crystallographic structure analysis. To further optimize the indirect formation of type-I crystals, it is of interest to know the mechanism of the PEG-induced detergent extraction. So far, the transformation does not work with  $C_{12}G_2$ , but only with crystals containing a different type of nonionic detergent, octaethyleneglycolmonododecylether ( $C_{12}E_8$ ). The headgroup of this detergent is itself a PEG-like molecule. It was proposed that PEG stabilizes the detergent monomers in the solution surrounding the crystal as represented by the increase of the CMC and, in this way, promotes extraction of the detergent from the crystal. Indeed, it was found that the polymer constant of PEG 5000 MME is larger for  $C_{12}E_8$  than for  $C_{12}G_2$ .<sup>36</sup> We are presently further analyzing this effect and its relation to the surface tension of the transformation buffer.

## 5 Conclusions and outlook

We have analyzed the influence of PEG2000 on the micelle formation of alkyl maltosides ( $C_nG_2$  with  $n = 10, 11, 12$ ). The approximately linear dependence of  $\ln(\text{CMC}/\text{CMC}_0)$  on the PEG concentration  $\chi$  is traced back to a linear shift of the free energy of micellization that correlates with the change of the surface tension of detergent-free aqueous solutions at  $\chi > 1$  % w/v. The effect is interpreted as due to binding of detergent monomers to PEG, which is postulated to be similar to the binding of PEG to a PEG-saturated buffer-air interface at bulk concentrations of PEG between 1 and 25 % w/v. A similar mechanism is postulated for the influence of proteins on micelle formation. Membrane proteins influence the formation of micelles and the detergent belt surrounding their originally membrane-spanning surface by additional interaction of their water-exposed surfaces with detergent monomers. Consequently, they change the free energy of micelle and belt formation, which explains the approximately linear dependence of  $\ln(\text{CSC}/\text{CMC}_0)$  on the protein concentration  $c_{\text{prot}}$ . When used as precipitating agent in membrane protein crystallization experiments, PEG determines the CSC curve in the phase diagram, which is hypothesized to be crucial for a quantitative understanding of the crystallization process and the finding of optimal crystallization conditions. In particular, the existence of different crystallization slots for type-I and type-II membrane protein crystals is postulated, where the former is at detergent concentrations below the CSC and the latter above the CSC. To further test the roles of CMC and CSC, future experiments will aim at a direct determination of the CMC and CSC curves for  $C_nG_2$  and  $C_{12}E_8$  in the presence

of the membrane proteins PSI and PSII and the effects of PEG variants with different molecular weights.

### **Acknowledgments**

We thank Dr. E. Schlodder for allowing us to use his fluospectrometer. The experiments of this work were performed over several years and benefited from the financial support by the Deutsche Forschungsgemeinschaft through diverse Collaborative Research Centers (SFB 429, 498 and 1078, project A5) as well as the Cluster of Excellence “Unifying Concepts in Catalysis” coordinated by the Technische Universität Berlin. Further support comes from HFSP (RGP 0063/2013) and the Bundesministerium für Bildung und Forschung (BMBF), Germany (Biotechnologie 2020+, project 031A154B).

## References

1. K. Holmberg, B. Jönsson, B. Kronberg and B. Lindman, *Surfactants and Polymers in Aqueous Solution*, Wiley, Chichester, 2003.
2. R. J. Hunter, *Foundations of Colloid Science*, Oxford University Press, Oxford, New York, 2001.
3. R. M. Garavito and S. Ferguson-Miller, *J. Biol. Chem.*, 2001, **276**, 32403-32406.
4. G. G. Privé, *Methods*, 2007, **41**, 388-397.
5. C. Stubenrauch, *Curr. Opin. Colloid Interface Sci.*, 2001, **6**, 160-170.
6. P. Rosevear, T. VanAken, J. Baxter and S. Ferguson-Miller, *Biochemistry*, 1980, **19**, 4108-4115.
7. K. Kalyanasundaram and J. K. Thomas, *J. Am. Chem. Soc.*, 1977, **99**, 2039-2044.
8. F. Müh and A. Zouni, *Biochim. Biophys. Acta*, 2008, **1778**, 2298-2307.
9. P. Jordan, P. Fromme, H. T. Witt, O. Klukas, W. Saenger and N. Krauss, *Nature*, 2001, **411**, 909-917.
10. B. Loll, J. Kern, W. Saenger, A. Zouni and J. Biesiadka, *Nature*, 2005, **438**, 1040-1044.
11. M. Broser, A. Gabdulkhakov, J. Kern, A. Guskov, F. Müh, W. Saenger and A. Zouni, *J. Biol. Chem.*, 2010, **285**, 26255-26262.
12. Y. Umena, K. Kawakami, J. R. Shen and N. Kamiya, *Nature*, 2011, **473**, 55-60.
13. D. Stroebel, Y. Choquet, J. L. Popot and D. Picot, *Nature*, 2003, **426**, 413-418.
14. S. Murakami, R. Nakashima, E. Yamashita and A. Yamaguchi, *Nature*, 2002, **419**, 587-593.
15. M. L. Rodrigues, T. F. Oliveira, I. A. C. Pereira and M. Archer, *EMBO J.*, 2006, **25**, 5951-5960.
16. H. Ago, Y. Kanaoka, D. Irikura, B. K. Lam, T. Shimamura, K. F. Austen and M. Miyano, *Nature*, 2007, **448**, 609-612.
17. D. M. Molina, A. Wetterholm, A. Kohl, A. A. McCarthy, D. Niegowski, E. Ohlson, T. Hammarberg, S. Eshaghi, J. Haeggstrom and P. R. Nordlund, *Nature*, 2007, **448**, 613-616.
18. T. Murata, I. Yamato, Y. Kakinuma, A. G. W. Leslie and J. E. Walker, *Science*, 2005, **308**, 654-659.
19. C. Hunte, J. Koepke, C. Lange, T. Rossmann and H. Michel, *Structure*, 2000, **8**, 669-684.
20. H. Palsdottir, C. G. Lojero, B. L. Trumpower and C. Hunte, *J. Biol. Chem.*, 2003, **278**, 31303-31311.
21. D. Yernool, O. Boudker, Y. Jin and E. Gouaux, *Nature*, 2004, **431**, 811-818.
22. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait and R. MacKinnon, *Nature*, 2002, **415**, 287-294.
23. L. Qin, C. Hiser, A. Mulichak, R. M. Garavito and S. Ferguson-Miller, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 16117-16122.
24. C. R. D. Lancaster, A. Kroger, M. Auer and H. Michel, *Nature*, 1999, **402**, 377-385.
25. F. Sun, X. Huo, Y. Zhai, A. Wang, J. Xu, D. Su, M. Bartlam and Z. Rao, *Cell*, 2005, **121**, 1043-1057.
26. M. Aoudia and R. Zana, *J. Colloid Interface Sci.*, 1998, **206**, 158-167.
27. J. Blouwolff and S. Fraden, *J. Crystal Growth*, 2007, **303**, 546-553.
28. M. G. Santonicola, M. A. Yocum, A. M. Lenhoff and E. W. Kaler, *Langmuir*, 2007, **23**, 5358-5366.
29. C. Hitscherich, J. Kaplan, M. Allaman, J. Wiencek and P. J. Loll, *Protein Sci.*, 2000, **9**, 1559-1566.
30. A. George and W. W. Wilson, *Acta Crystallogr. D*, 1994, **50**, 361-365.
31. P. J. Loll, M. Allaman and J. Wiencek, *J. Crystal Growth*, 2001, **232**, 432-438.
32. P. Thiagarajan and D. M. Tiede, *J. Phys. Chem.*, 1994, **98**, 10343-10351.
33. H. Michel, *Trends Biochem. Sci.*, 1983, **8**, 56-59.

34. R. Standfuss, A. C. T. van Scheltinga, M. Lamborghini and W. Kühlbrandt, *EMBO J.*, 2005, **24**, 919-928.
35. G. Rummel and J. P. Rosenbusch, in *Methods and Results in Crystallization of Membrane Proteins*, ed. S. Iwata, International University Line, La Jolla, 2003.
36. J. Hellmich, M. Bommer, A. Burkhardt, M. Ibrahim, J. Kern, A. Meents, F. Müh, H. Dobbek and A. Zouni, *Structure*, 2014, **22**, 1607-1615.
37. K. P. Ananthapadmanabhan, E. D. Goddard, N. J. Turro and P. L. Kuo, *Langmuir*, 1985, **1**, 352-355.
38. J. Aguiar, P. Carpena, J. A. Molina-Bolívar and C. C. Ruiz, *J. Colloid Interface Sci.*, 2003, **258**, 116-122.
39. K. S. Birdi, H. N. Singh and S. U. Dalsager, *J. Phys. Chem.*, 1979, **83**, 2733-2737.
40. E. De Vendittis, G. Palumbo, G. Parlato and V. Bocchini, *Anal. Biochem.*, 1981, **115**, 278-286.
41. E. B. Abuin, E. A. Lissi, A. Aspée, F. D. Gonzalez and J. M. Varas, *J. Colloid Interface Sci.*, 1997, **186**, 332-338.
42. R. Nagarajan and E. Ruckenstein, *Langmuir*, 1991, **7**, 2934-2969.
43. B. C. Stephenson, A. Goldsipe, K. J. Beers and D. Blankschtein, *J. Phys. Chem. B*, 2007, **111**, 1025-1044.
44. L. A. Girifalco and R. J. Good, *J. Phys. Chem.*, 1957, **61**, 904-909.
45. R. J. Good and E. Elbing, *Ind. Eng. Chem.*, 1970, **62**, 54-&.
46. J. Lipfert, L. Columbus, V. B. Chu, S. A. Lesley and S. Doniach, *J. Phys. Chem. B*, 2007, **111**, 12427-12438.
47. L. Bamber, M. Harding, P. J. G. Butler and E. R. S. Kunji, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 16224-16229.
48. C. Dupuy, X. Auvray and C. Petipas, *Langmuir*, 1997, **13**, 3965-3967.
49. M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548-558.
50. I. Tunón, E. Silla and J. L. Pascual-Ahuir, *Prot. Eng.*, 1992, **5**, 715-716.
51. B. C. Stephenson, A. Goldsipe, K. J. Beers and D. Blankschtein, *J. Phys. Chem. B*, 2007, **111**, 1045-1062.
52. B. C. Stephenson, K. J. Beers and D. Blankschtein, *J. Phys. Chem. B*, 2007, **111**, 1063-1075.
53. W. Al-Soufi, L. Pineiro and M. Novo, *J. Colloid Interface Sci.*, 2012, **370**, 102-110.
54. T. Gilányi, I. Varga, M. Gilányi and R. Mészáros, *J. Colloid Interface Sci.*, 2006, **301**, 428-435.

## Tables

Table 1 Parameters used to relate the effect of PEG2000 on the CMC of  $C_nG_2$ , represented by the polymer constant  $k_p$ , to the effect on the surface tension of detergent-free buffer solutions.

$n$	10	11	12
$g_{tr}(0) + g_{int}(0)$	$-14.0 \pm 0.6$	$-16.3 \pm 0.6$	$-18.2 \pm 0.7$
$g_{dehydr}(0)$	$-16.8 \pm 0.4$	$-18.2 \pm 0.4$	$-19.6 \pm 0.4$
$k_p$ (TMT) (% w/v) $^{-1}$	$0.040 \pm 0.006$	$0.048 \pm 0.006$	$0.055 \pm 0.007$
$k_p$ (CS-TMT) (% w/v) $^{-1}$	$0.050 \pm 0.006$	$0.055 \pm 0.006$	$0.060 \pm 0.007$
$k_p$ (exp.) (% w/v) $^{-1}$	$0.040 \pm 0.005$	$0.045 \pm 0.002$	$0.046 \pm 0.005$

## Figure captions

**Figure 1.** Representative titration curves of the ANS fluorescence intensity  $I_{\text{ANS}}$  (A) and the pyrene  $I_1/I_3$ -fluorescence ratio (B) versus total detergent concentration for  $C_{12}G_2$  in 100 mM PIPES (pH 7.0), 5 mM  $\text{CaCl}_2$ . The straight lines illustrate the extrapolation procedures to determine the CMC discussed in the text (Section 3.1). The sigmoidal curve in B is a fit of the experimental data points to eq. (2) to determine the parameter  $x_1$  that is identified with the CMC of the detergent and equals the value  $C_{\text{ANS}}$  determined with ANS as indicated in A. The parameter  $x_0$  is the inflection point of the sigmoid, and  $x_2$  is of interest for membrane protein solubilization as discussed in Section 4.2 (see also Fig. 7).

**Figure 2.** Representative titration curves of the ANS fluorescence intensity  $I_{\text{ANS}}$  (A) and the pyrene  $I_1/I_3$ -fluorescence ratio (B) versus total detergent concentration for  $C_{11}G_2$  in 100 mM PIPES (pH 7.0), 5 mM  $\text{CaCl}_2$  for different concentrations of PEG2000. The straight lines illustrate the extrapolation procedures to determine the CMC discussed in the text (Section 3.1) and the effect of PEG on the shape of the titration curves.

**Figure 3.** Dependence of parameters characterizing the titration curves of the ANS fluorescence intensity ( $C_{\text{ANS}}$ ) and the pyrene  $I_1/I_3$ -fluorescence ratio ( $x_0$ , inflection point, and  $x_1 = x_0 - 2\Delta x$ ) on the PEG2000 concentration  $\chi$  for three different alkyl maltosides  $C_nG_2$  in 100 mM PIPES (pH 7.0), 5 mM  $\text{CaCl}_2$ . The straight lines serve as guides to the eye. Numerical values are listed in ESI,<sup>†</sup> Tables 2 – 4. The good match of  $C_{\text{ANS}}$  and  $x_0 - 2\Delta x$  support the assignment of the latter to the CMC.

**Figure 4.** Dependence of  $\ln(C_{\text{ANS}}/C_0)$  on the PEG2000 concentration  $\chi$ , where  $C_0$  is  $C_{\text{ANS}}$  (i.e., the CMC) at  $\chi = 0$ , for three different alkyl maltosides  $C_nG_2$  in 100 mM PIPES (pH 7.0), 5 mM  $\text{CaCl}_2$ . The straight lines are fits to eq. (1) to determine the polymer constant  $k_P$  ( $k_P(\text{exp.})$  in Table 1). The error bars at each point indicate the interval  $\pm 0.1$  corresponding to the error of  $\ln(C_{\text{ANS}}/C_0)$  due to the error of  $\pm 5\%$  of  $C_{\text{ANS}}$  estimated from the graphical analysis of the titration curves.

**Figure 5.** Micelle model used for the molecular thermodynamic modeling of the free energy of micelle formation,  $g_{\text{mic}}$  (see Section 3.2). The ellipsoidal shape (oblate spheroid) is motivated by SAXS experiments<sup>46</sup> and the parameters  $a$  and  $b$  characterizing the size of the micellar core as well as  $d$  determining the thickness of the layer of headgroups are inferred from these experiments.

**Figure 6.** Dependence of the surface tension  $\gamma_w$  of the buffer-air interface (100 mM PIPES, pH 7.0, 5 mM  $\text{CaCl}_2$ ) on the concentration of PEG2000  $\chi$ . The blue line is a fit of data points for  $\chi > 1\%$  w/v to eq. (37).

**Figure 7.** Schematic titration curve (blue) of the  $I_1/I_3$ -ratio of pyrene as a function of detergent concentration in the presence of a membrane protein as determined in ref. 8. The green lines illustrate the graphical determination of the two characteristic points  $x_1$  and  $x_2$ , which are related to a Boltzmann-type sigmoid as described in Section 3.1.  $x_1$  is assigned in this work to the critical micelle concentration (CMC) of the detergent based on a comparison with ANS fluorescence data.  $x_2$  is assigned in ref. 8 to the critical solubilization concentration (CSC) of the detergent/membrane protein combination based on a comparison with turbidity data. The detergent concentration range between  $x_1$  and  $x_2$  is hypothesized to be suitable for the formation of type-I membrane protein crystals, since the detergent belts surrounding the originally membrane-spanning protein surfaces can be expected to be incomplete.

**Figure 8.** (A) Schematic phase diagram of a membrane protein solution containing detergent and PEG with the CMC and CSC as a function of PEG concentration indicated according to our present knowledge. The green and blue diamond indicate the optimal region for the formation of type-II and type-I membrane protein crystals, respectively. With each region is associated a hypothetical crystallization slot,<sup>30</sup> in which the second virial coefficient  $B_{22}$  of protein-protein interactions is slightly negative. The arrows indicate the development of the system during a vapor diffusion crystallization experiment. Note that for type-I crystals, it might be necessary to cross the CSC curve during this development. (B) Illustration of the crucial role of the stability of detergent aggregates (belts in PDCs and free micelles) in membrane-protein crystallization. Both, belts and micelles, are stable above the CSC curve, but the concentration of the latter is likely minimal close to the CSC curve. In this region of the phase diagram, type-II crystals form (green arrow), which contain fully developed detergent belts (left). Direct formation of type-I crystals (blue arrow), which are depleted of detergent and allow for tight protein contacts, require a destabilization of detergent belts (right). This is likely the case below the CSC curve. An alternative route to type-I crystals, discovered recently,<sup>36</sup> is indirect via type-II crystals and requires detergent extraction from these crystals (red arrow). Agents like PEG that stabilize detergent monomers in aqueous solution (and increase the CMC) likely promote this extraction.

Figure 1

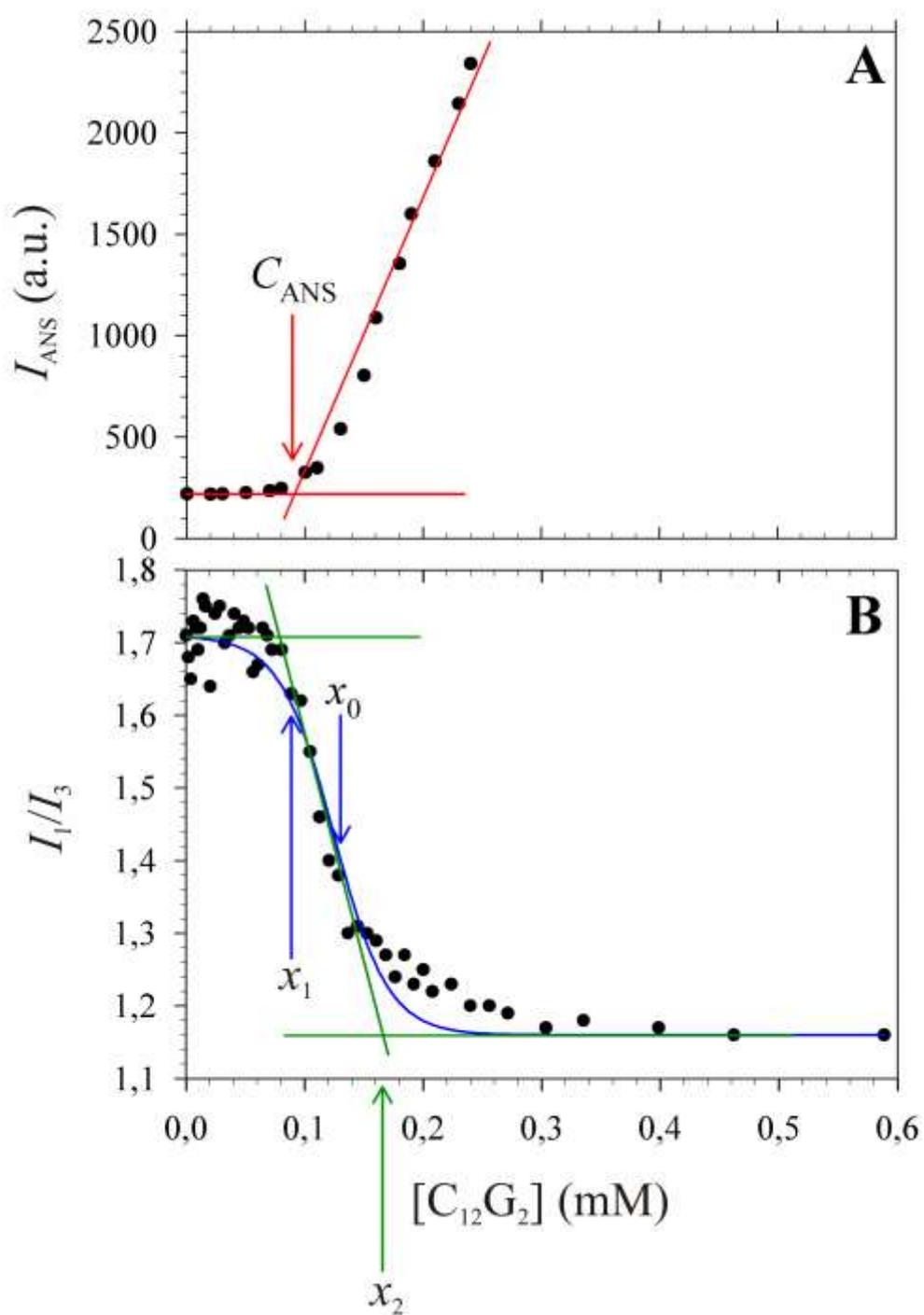


Figure 2

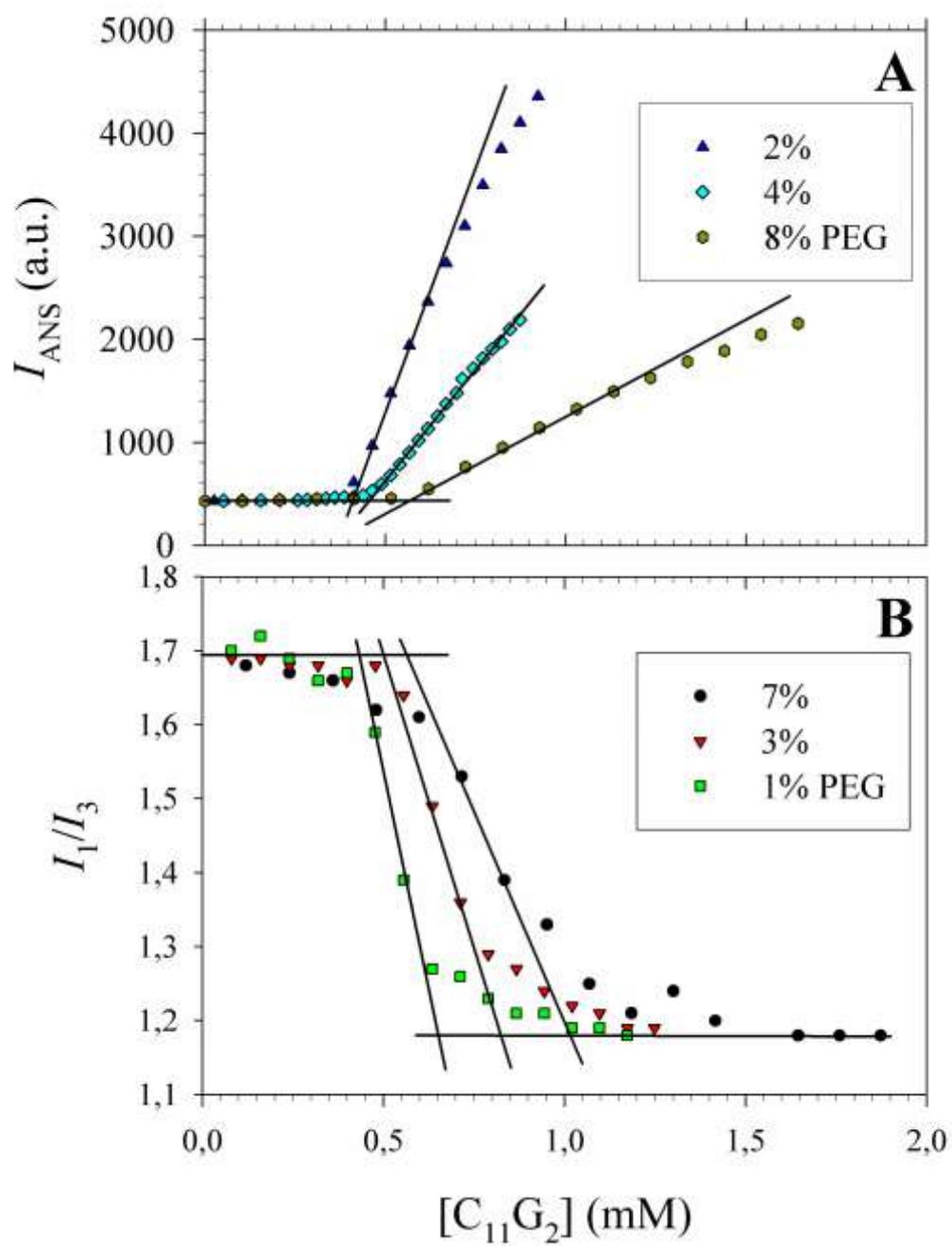


Figure 3

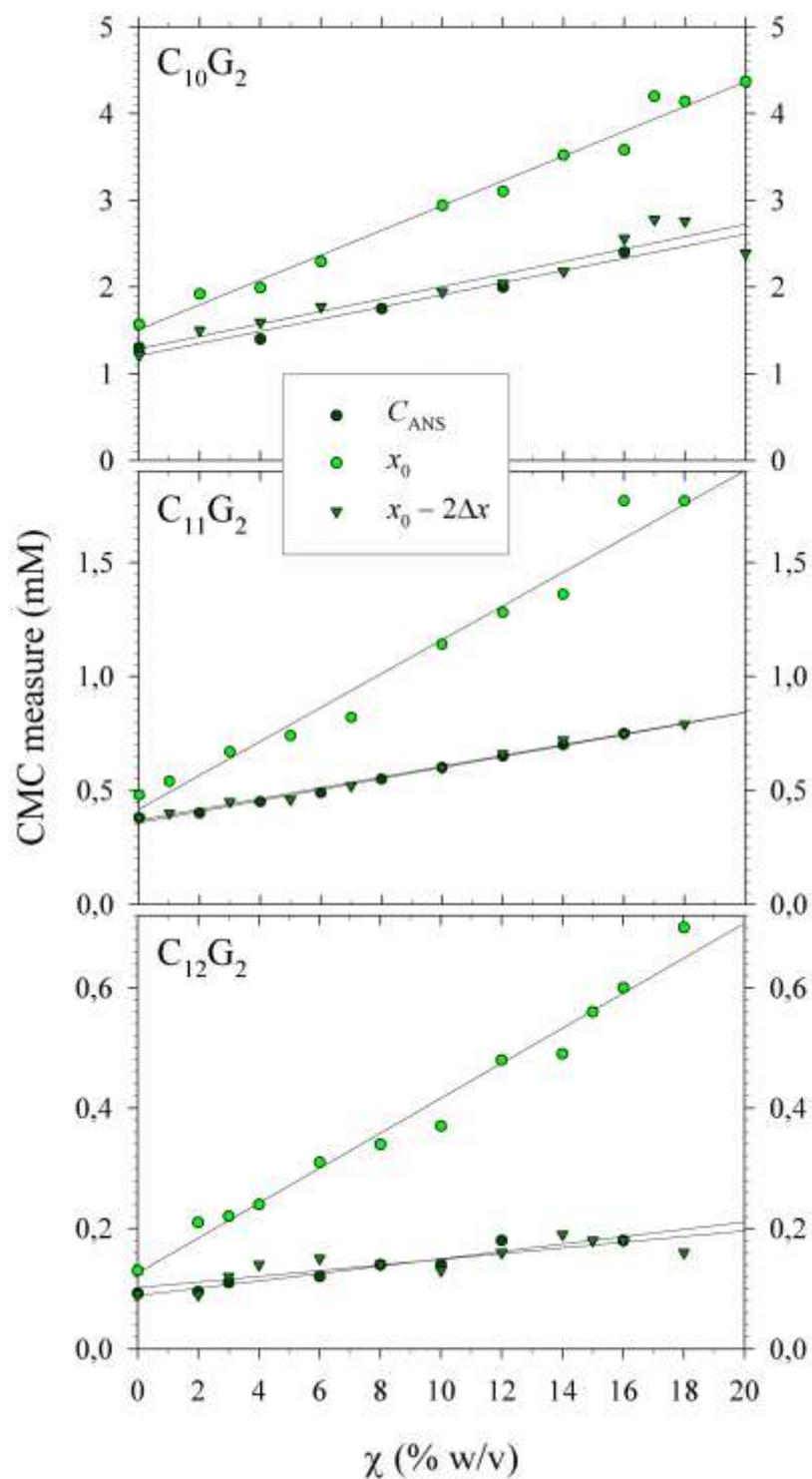


Figure 4

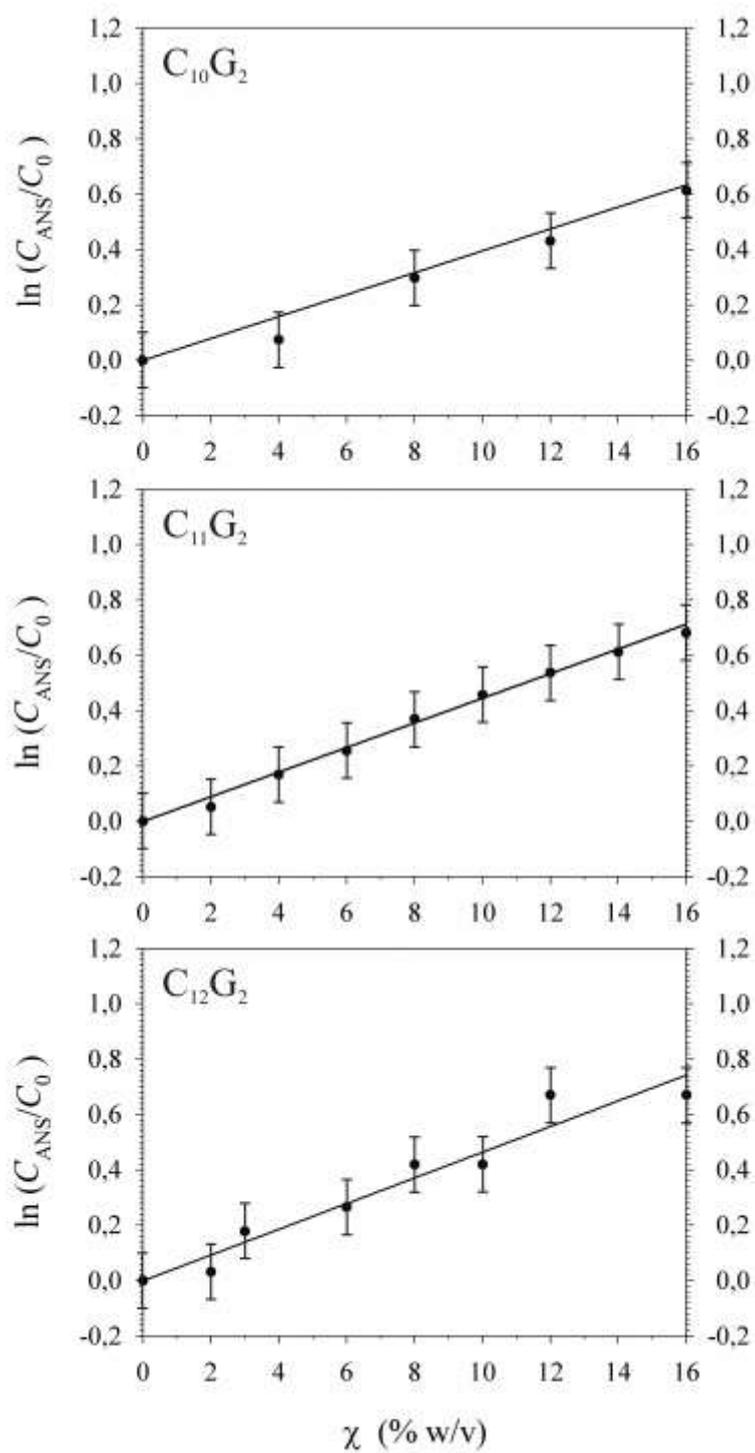


Figure 5

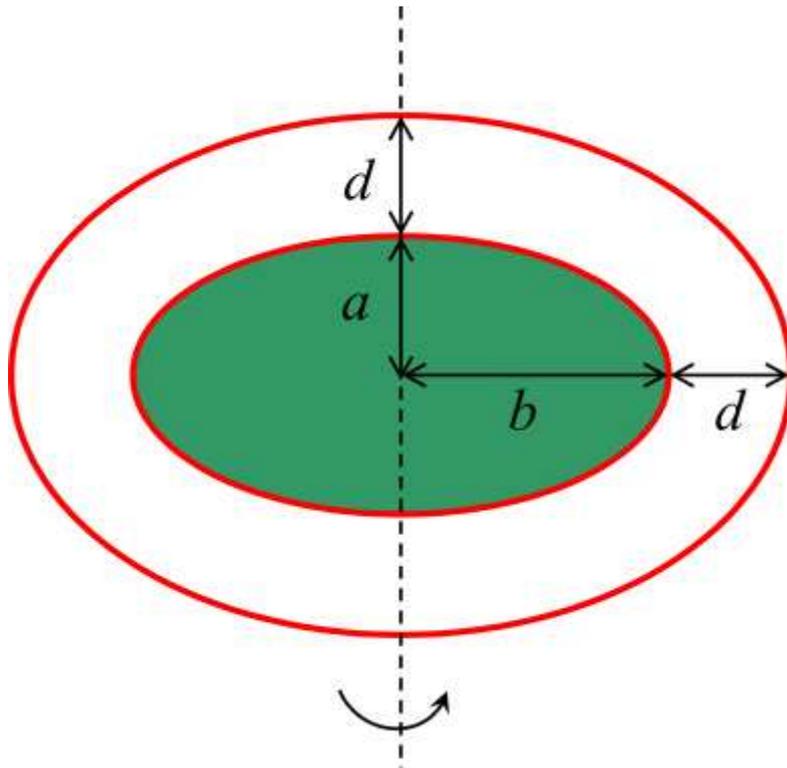


Figure 6

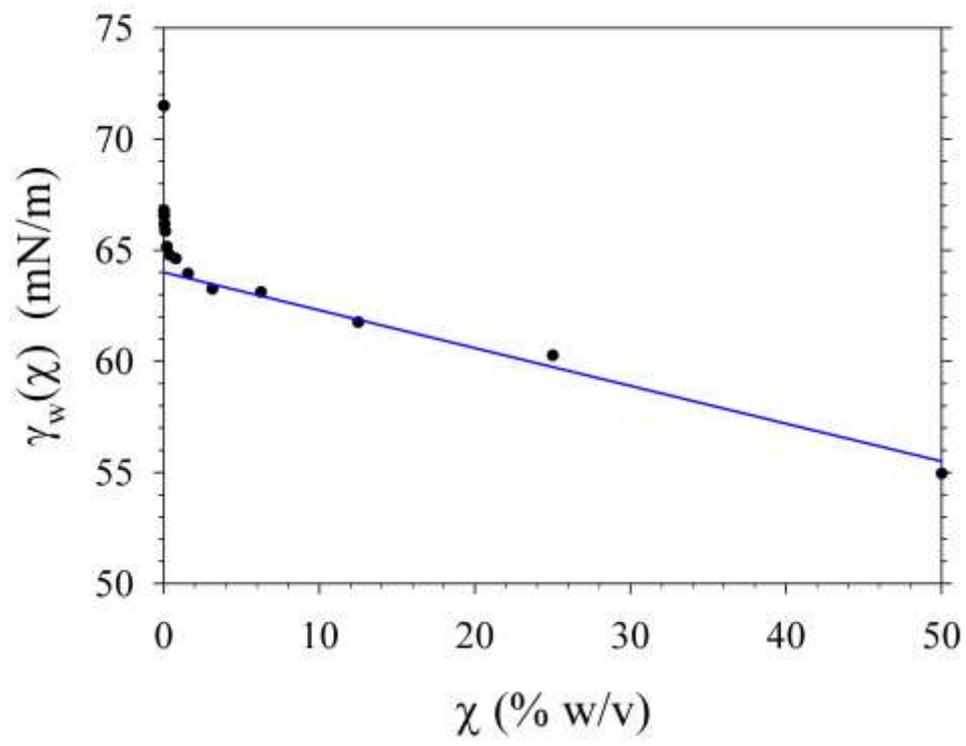


Figure 7

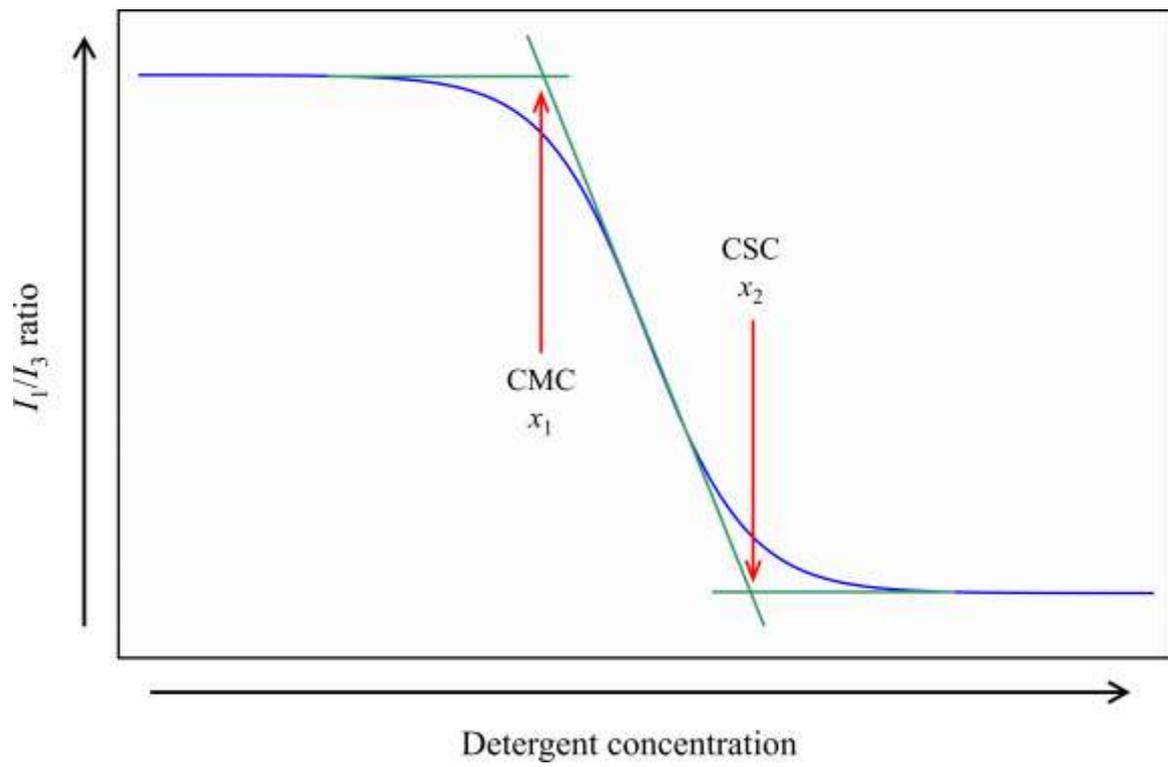


Figure 8

