Electronic Supplementary Material (ESI) for Soft Matter.

Supplementary information for

Structural heterogeneity of milk casein micelles: A SANS contrast variation study

Antoine Bouchoux,[ab*] Jorge Ventureira,[ab] Geneviève Gésan-Guiziou,[ab] Fabienne Garnier-Lambrouin,[ab] Peng Qu,[ab] Coralie Pasquier,[ab] Stéphane Pézennec,[ab] Ralf Schweins,[c] and Bernard Cabane[d]

[a]INRA, UMR 1253 Science et Technologie du Lait et de l'Œuf, F-35042 Rennes, France
[b]AGROCAMPUS OUEST, UMR1253 STLO, F-35042 Rennes, France
[c]Institut Laue-Langevin, DS/LSS group, F-38042 Grenoble CEDEX 9, France
[d]Laboratoire CBI, CNRS UMR 8231, Université Pierre et Marie Curie, Université René Diderot, ESPCI, 10 rue Vauquelin, 75231 Paris cedex 05, France

[*]Corresponding author. E-mail: Antoine.Bouchoux@insa-toulouse.fr. Present address: Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés / LISBP, UMR 5504/792 INRA-CNRS-INSA, 135 avenue de Rangueil, 31077 Toulouse cedex 04, France.

A - Centrifugation conditions for sample preparation

This section aims at giving the reader some justification for the different conditions of centrifugation (speed & time) that were used for preparing the casein micelle samples. Centrifugation was performed for several purposes; each one being briefly discussed in the following paragraphs. In all cases, the settling velocity v_s of the objects present in the dispersions is estimated through the Stokes law:

$$v_s = \frac{2}{9} R_p^2 \frac{\left(\rho_p - \rho_f\right)}{\eta} G \qquad \text{(S1)}$$

with R_p and ρ_p the radius and mass density of the particle, ρ_f and η the density and dynamic viscosity of the surrounding fluid (simply taken as 1 g.mL^{-1} and 0.001 Pa.s), and G the centrifugal acceleration.

- *Separating casein micelles from serum proteins:* In that case, centrifugations were performed at a *g*-force of 45.000 *g* for 1 h. For serum proteins (mainly α-lactalbumin, β-lactaglobulin) of ~2-3 nm in radius and ~1.35-1.4 g.mL^{-1} in density,[1,2] the maximum settling velocity is estimated to $v_{s,prot.}$ = 3.5 × 10^{-7} m.s^{-1}. For casein micelles of ~80 nm in radius, and ~1.08 g.mL^{-1} in density (as calculated from its average composition in caseins, CaP, and water, see next parts of this SI), the settling velocity is estimated to $v_{s,cm.}$ = 5.3 × 10^{-5} m.s^{-1}. So in 1 h, the average casein micelle has covered a settling distance of ~20 cm, while the serum proteins have covered only ~0.1 cm. Those values clearly indicate that the separation, as performed in centrifuge tubes of ~10 cm in size, is successful.

- *Separating casein micelles from possible casein aggregates:* Here centrifugations were performed at 30.000 *g* for 30 min to 1h (for samples from fresh milk), and at 15.000 *g* for 15 to 30 min (for samples from casein powder). For samples prepared from casein powder, the objective was to remove any possible micron-sized "aggregates" of caseins that could persist in the suspensions. With the conditions used, the maximum settling distance for casein micelles is estimated to ~3 cm, which is fully consistent with casein micelles that remain in the supernatant and micron-sized particles that are recovered in the pellet (settling distances > 100 cm). For samples prepared from fresh milk, our intention was not only to eliminate some potential aggregates but also to reduce the size polydispersity of the micellar population by removing the largest casein micelles. The centrifugation conditions were chosen according to works of other groups,[3,4] so that part of the micelles is eliminated in the pellet. Those conditions (30.000 *g*, 30 min - 1 h) correspond to a maximum settling distance of ~10 cm for an average casein micelle. In fact, only a small fraction of casein micelles was sedimented in that way (about 10% in mass), which led to a very limited decrease of the average size of the population (as measured through Dynamic Light Scattering (DLS), results not shown).

- *Separating precipitated caseins from residual fat:* Here centrifugation was performed at high speed (70.000 *g*) for 30 min with dispersions in which casein micelles have aggregated and precipitated at pH 4.6. As illustrated from the calculations above, these conditions are clearly sufficient to eliminate all the caseins in the pellet. The objects that could not be eliminated through such a centrifugation step have lower density than water and are probably residual fat globules, as discussed in the article.

B - SAXS experiments with FM samples at various D_2O contents

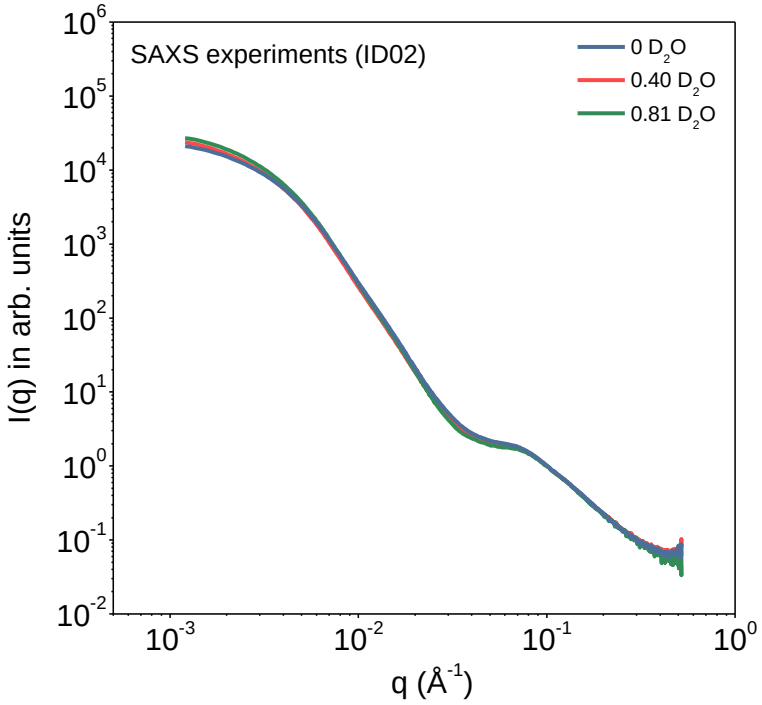

Fig. S1 The SAXS intensities of casein micelles from fresh milk (FM_s1) at casein concentration $C \approx 12.5$ g/L and at various volume fractions of heavy water. The intensities are slightly rescaled (normalization at $q = 0.1$ A^{-1}) to account for the small differences in concentration between the three samples. Experiments were performed at the ESRF ID02 beamline (Grenoble, France).

As opposed to SANS, SAXS is sensitive to variations in the electronic density of the objects. A SAXS curve is therefore not supposed to change with H-D substitution, unless this substitution does modify the internal structure and/or size of the object.

Fig. S1 shows Small Angle X-ray Scattering (SAXS) curves obtained from casein micelles of one of the experimental series presented in the article (fresh milk, FM_s1 series). Those spectra were obtained at three concentrations of D_2O, namely x_{D_2O} = 0, 0.41, and 0.81. From this figure, it appears that H-D exchange has virtually no effect on the SAXS profile of the casein micelle. It is then safe to consider in our model calculations (see the article and sections E and F of this SI) that the micellar size and internal structure are parameters that do not change with x_{D_2O}.

C - Guinier plots FM_s2, NPC_s1, NPC_s2, and NPC_s2_Residual

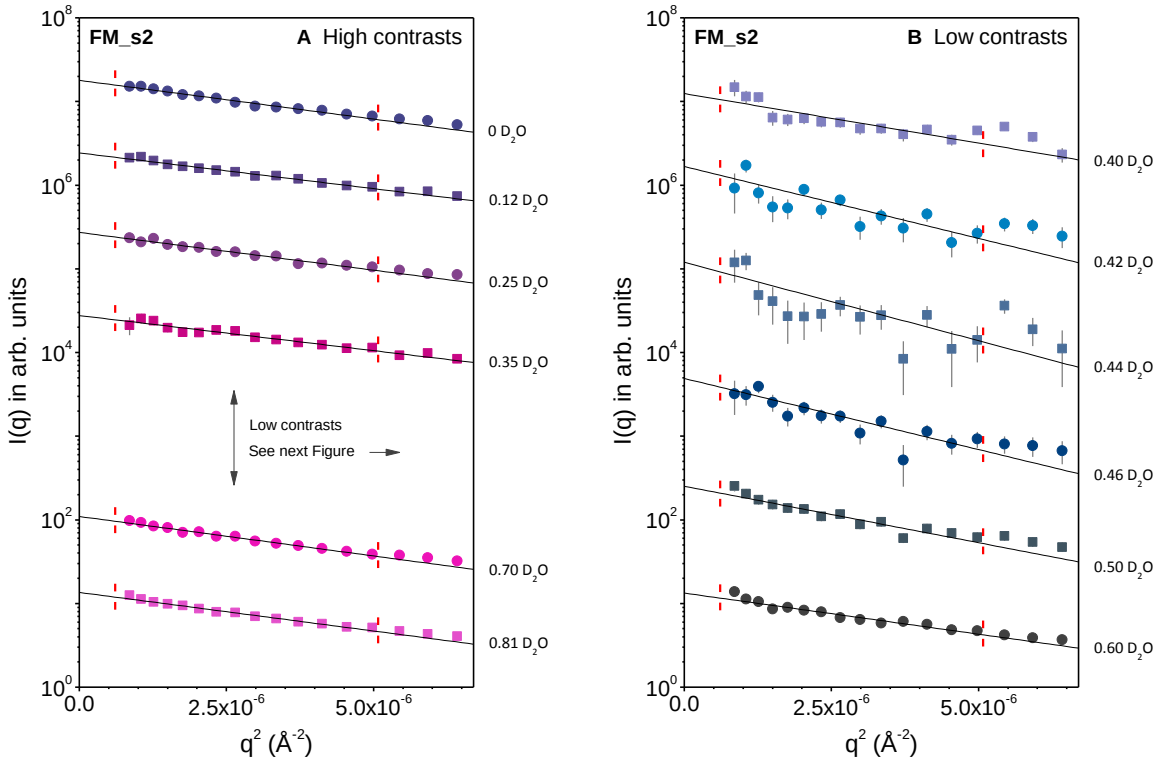

Fig. S2 The Guinier plots for the SANS intensities of sample set FM_s2.

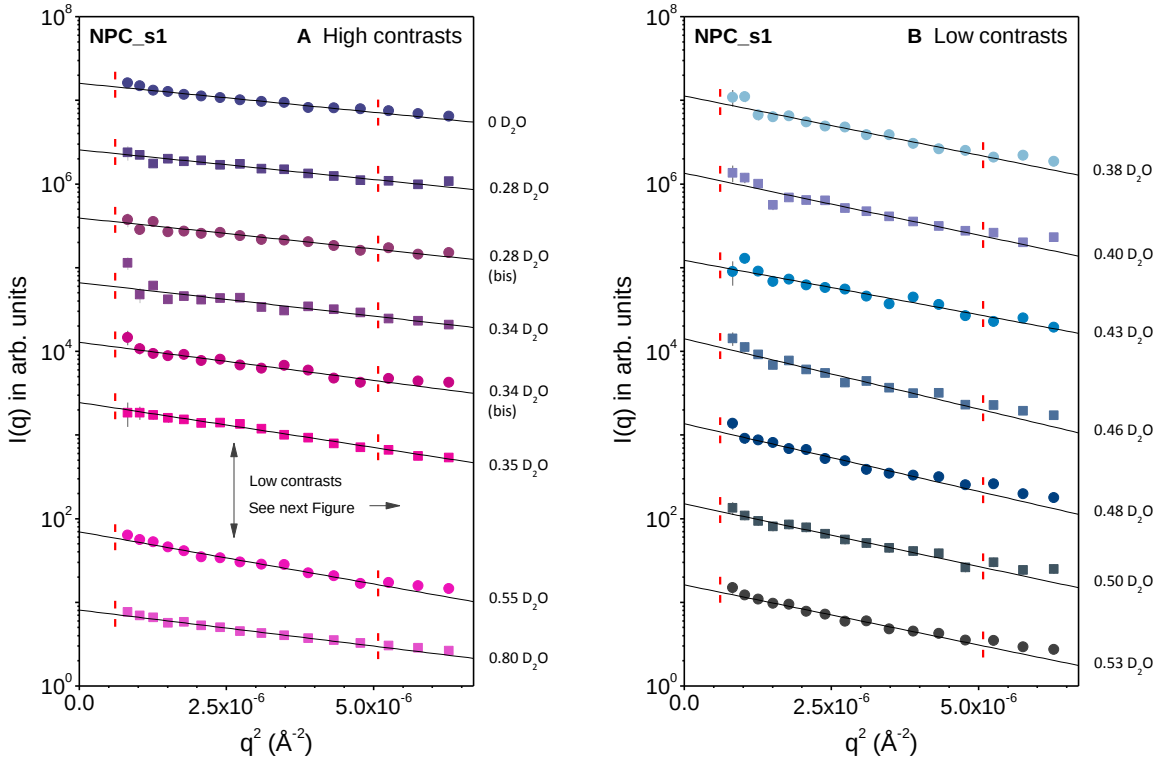

Fig. S3 The Guinier plots for the SANS intensities of sample set NPC_s1.

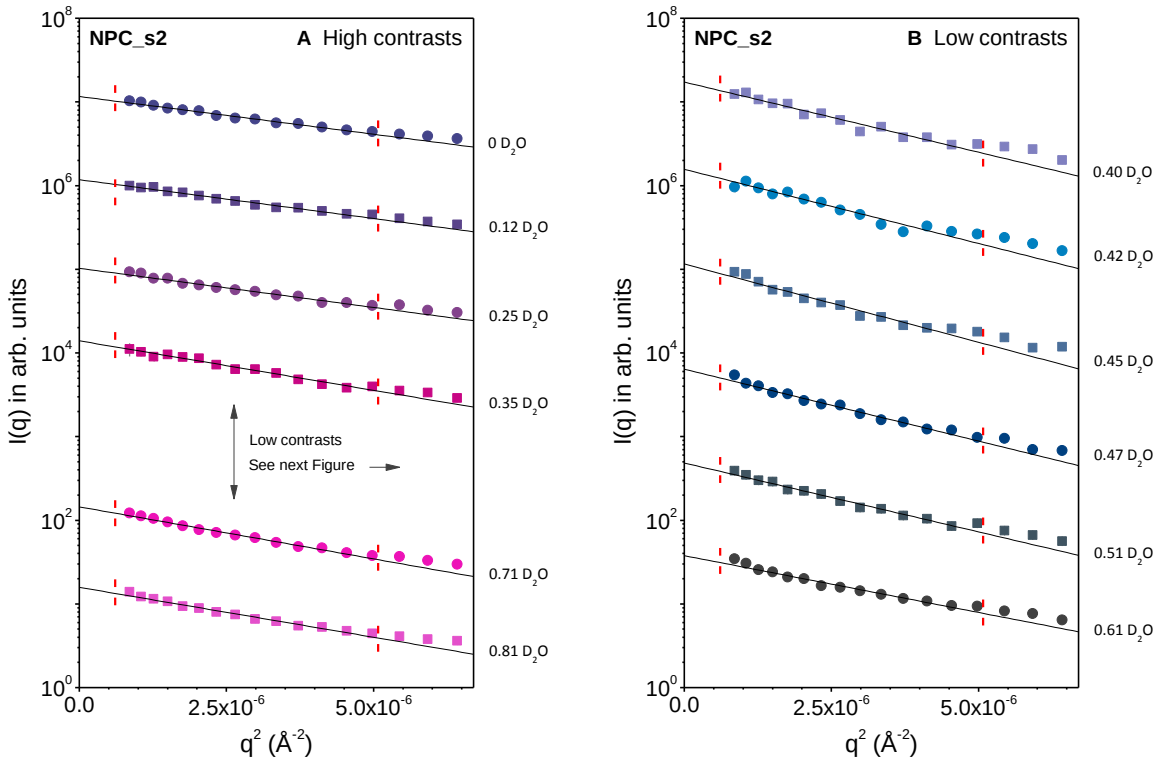

Fig. S4 The Guinier plots for the SANS intensities of sample set NPC_s2.

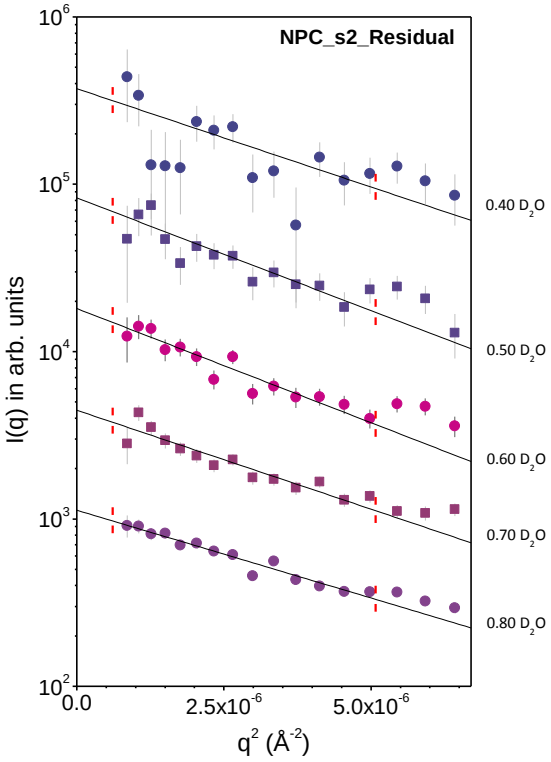

Fig. S5 The Guinier plots for the SANS intensities of sample set NPC_s2_Residual.

D - The neutron scattering length density (SLD) of the "average" casein micelle

α_{s1}-, α_{s2}-, β-, κ-caseins

The following table gives the neutron SLDs estimated for the four major casein proteins that compose the casein micelle: α_{s1}-, α_{s2}-, β-, κ-caseins. The chemical formulas are determined taking into account the protonation state of each amino acid residue at pH 6.6. The density of the caseins in H_2O is taken as $d_{cas} = 1/v^*_{cas}$ =1/0.736 mL.g^{-1}.[5] The mass densities of the caseins in D_2O are estimated from their value in H_2O and by considering their density as proportional to their molecular weight.

Table S1 The scattering length densities of α_{s1}-, α_{s2}-, β-, κ-caseins in light and heavy water.

Protein		**Chemical Formula**	**Mass Density (g.mL^{-1})**	**SLD (10^{10} cm^{-2})**	**Mass fraction in the casein micelle**[6]
α_{s1}-casein *CASA1_BOVIN, P02662*[*]	in H_2O	$C_{1035}H_{1575}N_{265}O_{317}S_5$	1.359	1.8975	0.45
	in D_2O	$C_{1035}H_{1237}D_{338}N_{265}O_{317}S_5$	1.379	3.1513	
α_{s2}-casein *CASA2_BOVIN, P02663*[*]	in H_2O	$C_{1083}H_{1709}N_{287}O_{338}S_6$	1.359	1.8393	0.12
	in D_2O	$C_{1083}H_{1297}D_{412}N_{287}O_{338}S_6$	1.382	3.2805	
β-casein *CASB_BOVIN, P02666*[*]	in H_2O	$C_{1080}H_{1684}N_{268}O_{310}S_6$	1.359	1.8072	0.33
	in D_2O	$C_{1080}H_{1370}D_{314}N_{268}O_{310}S_6$	1.377	2.9415	
κ-casein *CASK_BOVIN, P02668*[*]	in H_2O	$C_{849}H_{1321}N_{223}O_{262}S_4$	1.359	1.8645	0.10
	in D_2O	$C_{849}H_{1029}D_{292}N_{223}O_{262}S_4$	1.380	3.1757	

[*]Protein identifier in the UniProt database (http://www.uniprot.org)

The mass fractions of each type of casein in the casein micelle are estimated using the average concentrations of caseins in skim milk.[6] From those values, we can calculate the SLD of the "average" casein material that constitutes the micelle:

$$\rho_{cas}\left[10^{10} cm^{-2}\right] = 1.8573 + 1.2416\, x_{D_2O} \quad \text{(S2)}$$

which gives a contrast match point at x_{D_2O} = 0.426 in a mixture of pure D_2O/H_2O.

CaP nanoclusters

There is still no clear consensus about the exact composition of the CaP nanoclusters, with propositions ranging from the "simple" brushite, to hydroxyapatite or the more sophisticated amorphous $Ca(HPO_4)_{0.7}(PO_4)_{0.2,}xH_2O$.[7–9] Here we adopt the first proposition (brushite), with a chemical formula $CaHPO_4.2H_2O$, and a density of 2.31 g/mL.[7] Assuming that all the brushite protons are exchangeable, the SLD of the CaP nanoclusters is:

$$\rho_{CaP}\left[10^{10} cm^{-2}\right] = 2.0980 + 4.2076\, x_{D_2O} \quad \text{(S3)}$$

which gives a contrast match point at x_{D_2O} = 0.981 in a mixture of pure D_2O/H_2O.

Casein micelle

The SLD of a casein micelle of composition ϕ_{cas}, ϕ_{CaP} and ϕ_{solv} is simply given by:

$$\rho_{cm} = \phi_{cas}\rho_{cas} + \phi_{CaP}\rho_{CaP} + \phi_{solv}\rho_{solv} \quad \text{(S4)}$$

with ϕ_{cas} and ϕ_{CaP} the volume fractions occupied by the casein chains and the CaP nanoclusters in the micelle's interior, respectively, and $\phi_{solv} = 1-(\phi_{cas} + \phi_{CaP})$ the volume fraction of water inside the edifice.

The SLD of this water is:

$$\rho_{solv}\left[10^{10} cm^{-2}\right] = -0.5583 + 6.9161 x_{D_2O} \quad \text{(S5)}$$

as calculated for a mixture of pure D_2O/H_2O at 25°C.

From eqns (S2-S5) and condition $\phi_{solv} = 1-(\phi_{cas} + \phi_{CaP})$, we obtain the SLD of any casein micelle of composition ϕ_{cas} and ϕ_{CaP}:

$$\rho_{cm}\left[10^{10} cm^{-2}\right] = \left[-0.5583 + 2.4156\phi_{cas} + 2.6563\phi_{CaP}\right] + \left[6.9161 - 5.6745\phi_{cas} - 2.7085\phi_{CaP}\right] x_{D_2O} \quad \text{(S6)}$$

The contrast match point of such a micelle is the D_2O volume fraction at which $\rho_{cm} - \rho_{solv} = 0$, which gives:

$$x_{D_2O} = \frac{2.4156 + 2.6563\dfrac{\phi_{CaP}}{\phi_{cas}}}{5.6745 + 2.7085\dfrac{\phi_{CaP}}{\phi_{cas}}} \quad \text{(S7)}$$

For an "average" casein micelle, the casein volume fraction ϕ_{cas} can be estimated to $\phi_{cas} = v^*_{cas} / v^*_{mc} \approx$ 0.167, with v^*_{mc} = 4.4 mL/g the specific volume of the micelle,[10] and v^*_{cas} =0.736 mL/g the specific volume of a casein molecule.[5] The CaP volume fraction can in turn be estimated from the CaP density d_{CaP}, the casein density d_{cas} and volume fraction ϕ_{cas}, and the average mass of micellar calcium per gram of caseins in a casein micelle, *i.e.*, $m_{CaP/cas}$ = 0.77 mM.g^{-1}.[9] This gives:

$$\phi_{CaP} = \frac{\phi_{cas} d_{cas} m_{CaP/cas} MW_{CaHPO_4.2H_2O}}{d_{CaP}} \approx 0.013 \quad \text{(S8)}$$

For this "average" casein micelle, eqn (S6) then turns into:

$$\rho_{cm}\left[10^{10} cm^{-2}\right] = -0.1204 + 5.9332 x_{D_2O} \quad \text{(S9)}$$

which gives a contrast match point at x_{D_2O} = 0.446 in a mixture of pure D_2O/H_2O.

E - The variation of I_0 with contrast *vs.* the distribution in composition of the casein micelle

Our SANS results indicate that the zero angle scattering of casein micelle dispersions does not vanish when contrast is varied, but rather passes through a minimum value that is not zero (Figs. 2 and 4(A) of the article). This can have two origins: (1) The composition of the casein micelles varies from one micelle to another (as observed by Stuhrmann for ferritin[11]), (2) The dispersions not only contain casein micelles, but also other objects of distinct SLD. Both cases result in a distribution of SLD within the sample, which makes it impossible to find a contrast at which all the objects become "invisible" to neutrons. Here we examine the first proposition at the light of what we know of the compositional characteristics of the casein micelle.

Casein micelles are commonly defined as particles that all contain about the same quantity of caseins (~17% in volume, part D of this SI), minerals (~1%), and water (~82%).[7,12] The relative proportion of each casein type (α_{s1}-, α_{s2}-, β-, κ-caseins) in the total volume of casein possibly changes from a micelle to another, in particular as a function of the size of the micelle.[13] But the potential effect of such a disparity on the variation of I_0 in SANS is very limited as the SLDs (or match points) of the different caseins are very close to each other. So we do not consider this variation in our calculations.

In fact, what has potentially a strong effect on the match point distribution of the casein micelles is the variation of the ratio minerals/caseins among the micelle population (eqn (S7)). However, very little is known about that variation, and this specific question has never been addressed as such in the literature (at least to our knowledge). At most, some indirect information can be found in two studies of D. Dalgleish and co-workers, in which "native" populations of casein micelles were fractionated into 6-8 subpopulations of different sizes through successive centrifugations.[13,14] In both works, the concentration ratio [CaP/caseins] is found to differ a bit from a fraction to another; the large micelles containing more CaP than the small ones (probably because the relative proportion of κ-casein molecules is higher in small casein micelles than in large ones,[13] and that κ-casein has less affinity towards minerals than the other caseins). The reported range of variation is 0.65-0.96 mM of CaP per gram of caseins (*i.e.*, ϕ_{CaP} = 0.011-0.016).[13,14] Importantly, the variation of the ratio [CaP/casein] between casein micelles belonging to a same subpopulation is never considered or discussed in those two articles. This is probably because this variation is assumed to be small compared to the "inter-fraction" variation discussed by the authors.

In a first and reasonable approach, we consider that the variation range given by Dalgleish *et al.* is a good estimate of the composition distribution of the casein micelle in "native" tank milk. For the convenience of the calculation, we can treat this composition distribution as a variation of CaP volume fraction ϕ_{CaP} among a population of casein micelles that have the same composition in caseins (ϕ_{cas} = 0.167, see part D in this SI). We finally assume that the variation of ϕ_{CaP} follows a log-normal distribution as the one given in Fig. S6(A), with *i* classes of average composition $\phi_{CaP,i}$ and relative

proportion (in number) N_i. In this case, we choose a distribution that is centered on the variation range given by Dalgleish et al.,[13,14] with about 80% of the micelles belonging to this range. For the sake of the demonstration, we also assume that ~20% of the micelles have CaP compositions that slightly go beyond this range, with limits at $\phi_{CaP} \approx 0.008$ and 0.021. The distribution is also chosen so that it yields an average concentration $\phi_{CaP} \approx 0.013$ for the whole collection of casein micelles.[9]

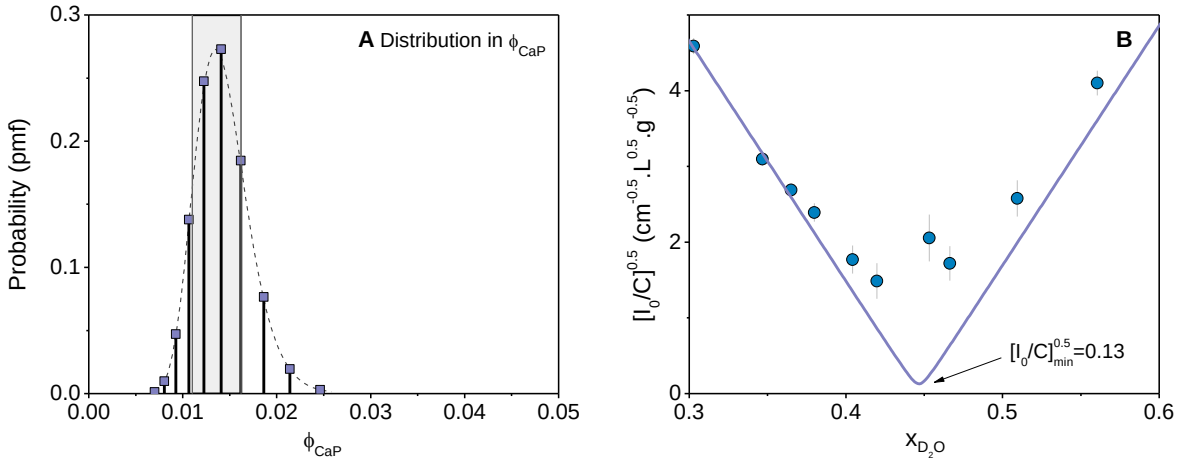

Fig. S6 Effect of composition distribution on the variation of I_0 with contrast. (A) The micelles have CaP compositions that follow a log-normal distribution covering the variation range estimated from works of other groups (shaded area). (B) The resulting variation of zero angle scattering intensity (line) as compared to the experimental points of the FM_s1 series (symbols).

The zero angle scattering intensity of the whole population of micelles is then simply given by:

$$I_0 = A \times \sum_i N_i \left(\rho_{cm,i} - \rho_s\right)^2 \qquad \text{(S10)}$$

where $\rho_{cm,i}$ is the SLD of the micelles of class i (eqn (S6)) and A is a constant that depends on the average size of the micelles and their total concentration. The variation of I_0 with contrast is simply constructed by calculating eqn (S10) at various values of x_{D_2O} and by adjusting A so that the intensities calculated at high contrasts match those measured experimentally.

Fig. S6(B) gives the result of this calculation. As expected, the minimum in zero angle scattering intensity is not zero. However, the obtained value is much lower than the minimum value measured experimentally (only the results of the FM_s1 experimental series are represented here), with $\sqrt{I_0/C}$ = 0.13 $cm^{-0.5}.L^{0.5}.g^{-0.5}$ from our calculations against 1.49 $cm^{-0.5}.L^{0.5}.g^{-0.5}$ in the experiment. So clearly, the variation of composition that is "naturally" expected for the casein micelle is not able to explain our I_0 results at low contrast. This is even more obvious when looking at the experimental results obtained with casein micelles from casein powder (Fig. 4(A) of the article), where the difference between the expected value for $I_{0,min}$ and the measured one is even more important: 0.13 $cm^{-0.5}.L^{0.5}.g^{-0.5}$ against 3.50 or 5.18 $cm^{-0.5}.L^{0.5}.g^{-0.5}$.

In a second approach, we proceed in the opposite direction and vary the composition distribution so that we can reproduce the experimental values of I_0 obtained at low contrast. The average CaP concentration in the whole population is maintained to $\phi_{CaP} \approx 0.013$. Also the distribution is still chosen to be of a log-normal type.

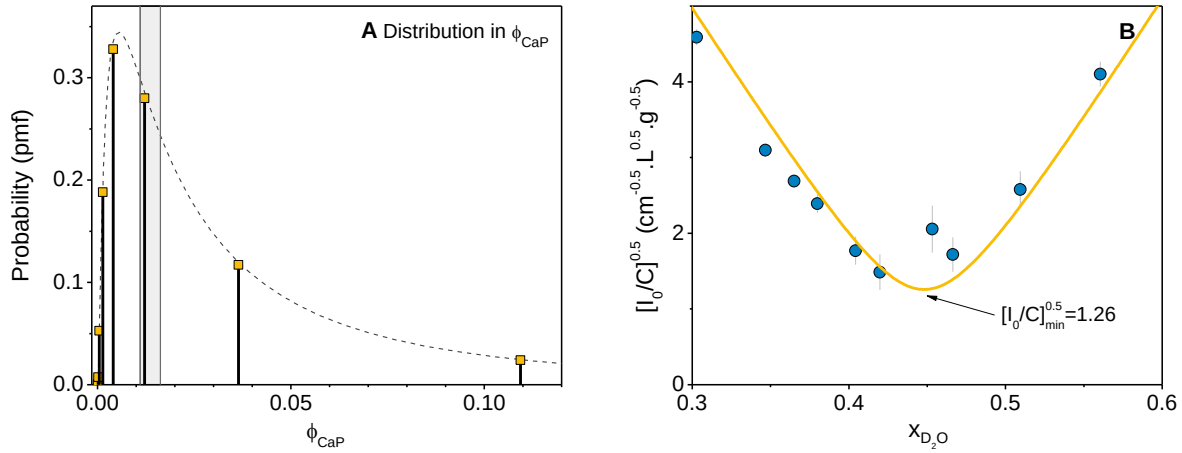

Fig. S7 Same as Fig. S6, but this time the distribution in ϕ_{CaP} among the population of micelles has been varied so that the calculated variation of I_0 best matches the experimental one.

Fig. S7(B) gives the best fit that is possible to obtain following that procedure. The modeled variation of I_0 is slightly shifted towards higher D_2O contents as compared to the experiments. But the fit is overall very satisfactory, with values of $I_{0,min}$ that clearly match the experimental ones. The resulting distribution in CaP composition is given in Fig. S7(A). Here, only 25-30% of the casein micelles have a CaP composition that lies in the range of the values reported by Dalgleish *et al*. The other micelles have compositions that are distributed over an extremely wide range of ϕ_{CaP} values, with ~15% of the micelles expected to contain 4 to 10 times more CaP nanoclusters than the "average" micelle, and ~20% of them showing a CaP content that is 10 times lower than the commonly admitted value of $\phi_{CaP} \approx 0.013$. Here it seems clear to us that such a distribution cannot be realistic. The first simple reason is that the obtained compositional distribution is colossal as compared to the one suggested by Dalgleish and co-workers. A second reason is that such a distribution in CaP content would be visible in transmission microscopy images of casein micelles; which is clearly not the case.[15–17] Also it is not certain that such a huge variation is CaP composition, and consequently in mass density, between individual casein micelles would allow to separate them successfully by centrifugation according to their size (as performed by Dalgleish *et al.* for instance[13,14]). Finally, we can add a last argument based on the Laplace pressure inside the CaP nanoclusters, which, according to their size (~5 nm in diameter[18]), is probably very large and in the order of ~100 atm. Such a pressure implies that the solubility of CaP in the dispersion is significant. This in turn suggests that the CaP/casein ratio easily and rapidly reaches its equilibrium value everywhere in the dispersion, and consequently has about the same value from a micelle to another.

As a conclusion, we find that (i) a realistic variation in composition among the population of casein micelles is not able to describe the minimum in zero angle scattering that is measured experimentally; (ii) the distribution of composition that is necessary to describe such I_0 values is totally improbable. So clearly, the I_0 behavior at low casein contrasts can only be explained by the presence of additional objects in the dispersion (Fig. 5 of the article).

F - I_0 & R_g modeling

In our article, three different scenarios are considered for modeling the variation of I_0 and R_g with contrast. In the first one (*homogeneous model*), the casein micelles are taken as polydisperse spheres of the same uniform composition. In the two others (*core-shell 1* and *2*), the casein micelles are taken as polydisperse core-shell particles, with a shell that contains either more or less CaP than the core. In all of these three scenarios, the casein micelles are supposed to be accompanied with a certain number of objects made of lipids or phospholipids (vesicles or residual fat droplets/globules). In the following, we give some details about the assumptions made and the equations used for performing our calculations. In Figs. S8-S11, we give a direct comparison between the experimental results and the models. Tables S2-S5 list the parameters injected in the models in each case.

Homogeneous model

. The casein micelles are supposed to be polydisperse spheres with a log-normal distribution; as it is commonly assumed and measured.[10] The probability density function $f(r)$ of the distribution is characterized by a polydispersity index σ and a mass-averaged hydrodynamic radius $\overline{R}$:

$$f(r) = \frac{1}{r\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln\left(r / \overline{R}\right)}{\sigma} \right)^2 \right] \quad \text{(S11)}$$

. This distribution is divided into i classes (typically 20) of average radius r_i and mass fraction m_i.

. The micelles are considered as fully uniform in density. Then in each class, the radius of gyration is simply:

$$R_{g,\mathrm{cm}_i} = \sqrt{\frac{3}{5}} r_i \quad \text{(S12)}$$

and the intensity scattered at zero angle is:

$$I_{0,\mathrm{cm}_i} = N_i V_{cm_i}^2 \left(\rho_{cm_i} - \rho_s \right)^2 \quad \text{(S13)}$$

where V_{cm_i} is the volume of the micelles in class i, and N_i is the number density of micelles, as calculated from m_i and the known mass of casein per micelle. Also each micelle has the same average composition, *i.e.*, ϕ_{cas} = 0.167, ϕ_{CaP} = 0.013, and ϕ_{water} = 0.820 (see part D of this SI), and thus have the same SLD $\rho_{cm_i} = \rho_{cm}$, as given in eqn. S9.

. The micelles are surrounded by fat droplets or vesicles of external radius R_{fat} or R_{ves}. The concentration ratio of (phospho-)lipids over caseins in the sample is $x_{fat/cas}$. Here we do not take into account the size distribution of those objects as (1) we do not know it exactly and we are not interested in determining it, (2) it has only a small influence on the fitted value of $x_{fat/cas}$, (3) our goal is only to have an estimation of this ratio $x_{fat/cas}$.

. The contrast match point of those objects is 0.10 (see Fig. 5 of the article). This gives a direct estimation of the SLD of the material they are made of, *i.e.*, $\rho_{lip}\left[10^{10} cm^{-2}\right] = 0.1333$ (assuming that this material cannot exchange protons with deuterium molecules).

. For the fat droplets, we have:

$$R_{g,fat} = \sqrt{\frac{3}{5}} R_{fat} \tag{S14}$$

and

$$I_{0,\mathrm{fat}} = N_{fat} V_{fat}^2 \left(\rho_{lip} - \rho_s\right)^2 \tag{S15}$$

where N_{fat} and V_{fat} are calculated from R_{fat} and $x_{fat/cas}$; and assuming that the density of the fat droplets is close to 1 g/mL.

. As for the vesicles, we assume that they have a bilayer membrane of thickness t_{ves} = 60 Å,[19] and internal radius $R_{ves,core} = R_{ves} - t_{ves}$. We then have:

$$R_{g,ves} = \sqrt{\frac{3}{5}\left(\frac{R_{ves}^5 - R_{ves,core}^5}{R_{ves}^3 - R_{ves,core}^3}\right)} \tag{S16}$$

and

$$I_{0,ves} = N_{ves}\left(\frac{4}{3}\pi\left(R_{ves}^3 - R_{ves,core}^3\right)\right)^2 \left(\rho_{lip} - \rho_s\right)^2 \tag{S17}$$

where N_{ves} is calculated from $x_{fat/cas}$.

. The average intensity scattered by the samples at zero angle is finally obtained by summing the intensities of each class of casein micelles, plus the intensity of the fat droplets or vesicles:

$$I_0 = \sum_i I_{0,cm_i} + I_{0,fat} \quad \text{or} \quad I_0 = \sum_i I_{0,cm_i} + I_{0,ves} \tag{S18}$$

. While the average radius of gyration that is measured from the sample is given by:

$$R_g = \sqrt{\frac{\sum_i I_{0,cm_i} R_{g,cm_i}^2 + I_{0,fat} R_{g,fat}^2}{I_0}} \quad \text{or} \quad R_g = \sqrt{\frac{\sum_i I_{0,cm_i} R_{g,cm_i}^2 + I_{0,\mathrm{ves}} R_{g,\mathrm{ves}}^2}{I_0}} \tag{S19}$$

. The full variation of R_g and I_0 with contrast is built from those two last equations.

. In our case, we are interested in knowing if the model can describe our experimental data in a reasonable manner. For that purpose, we vary the parameters $\overline{R}$, σ, $x_{fat/cas}$ and R_{fat} (or R_{ves}) of the model to fit the experimental data. The fitting procedure is stopped when the relative average deviation between the model and experimental values is minimum.

Core-Shell models

. The casein micelles are polydisperse core-shell particles with external radii that follow a log-normal distribution of parameters $\overline{R}$ and σ (eqn. S11).

. In all cases, we consider that the casein and CaP concentrations in the shell and in the core do not vary from one micelle to another. The thickness t of the core is also taken as constant within the population of micelles.

. In the "Core-Shell 1" model, the shell corresponds to the κ-casein hairy layer, which is usually reported as being 100 Å thick.[12] This shell is exempt of CaP nanoclusters, *i.e.*, $\phi_{CaP,shell}$ = 0.

. In the "Core-Shell 2" model, inspired from the work of Shukla et al.,[20] the shell is about 115 Å thick and its concentration in CaP nanoclusters is taken as 5 times the CaP concentration in the core:

$$\phi_{CaP,shell} = 5\phi_{CaP,\mathrm{core}} \tag{S20}$$

. In both models, the casein concentration in the core and shell are taken as identical to $\phi_{cas,shell}$ = $\phi_{cas,core}$ = 0.167. The water concentration in both the core and shell is given by $\phi_{solv} = 1-(\phi_{cas}+\phi_{CaP})$.

. The collection of micelles is divided into *i* classes (typically 20) of average radius r_i and mass fraction m_i.

. In each class of size, the radius of gyration of the micelles is given by:

$$R_{g,\mathrm{cm}_i} = \sqrt{\frac{\overline{R}^5(\rho_{shell}-\rho_{solv})-(\overline{R}-t)^5(\rho_{shell}-\rho_{core})}{\overline{R}^3(\rho_{shell}-\rho_{solv})-(\overline{R}-t)^3(\rho_{shell}-\rho_{core})}} \qquad \text{(S21)}$$

while the zero angle scattering is:

$$I_{0,\mathrm{cm}_i} = N_i\left[\frac{4}{3}\pi\left[\overline{R}^3(\rho_{shell}-\rho_{solv})-(\overline{R}-t)^3(\rho_{shell}-\rho_{core})\right]\right]^2 \qquad \text{(S22)}$$

In those expressions, ρ_{shell} and ρ_{core} are calculated from the volume fractions of casein, CaP, and water in each compartment (see eqn. S4). N_i is the number density of micelles, as calculated from m_i and the known mass of casein per micelle.

. In the same manner as in the "homogeneous" model, the micelles are supposed to be surrounded by fat droplets or phospholipid vesicles with radii of gyration and zero angle scattering intensities given by eqs. S14-S17.

. The average R_g and I_0 for the whole sample are given by eqs. S18 and S19. Their variation with contrast is obtained by performing the calculations at various D_2O contents.

. The procedure then consists in varying the parameters $\overline{R}$, σ, $x_{fat/cas}$ and R_{fat} (or R_{ves}) of the model to fit the experimental data of contrast variation. In all cases, the CaP volume fraction in the core of the micelles is also adjusted so that the average CaP concentration in the micelles (core+shell) is ϕ_{CaP} = 0.013. The fitting procedure is stopped when the relative average deviation between the model and the experimental values is minimum.

Stuhrmann Representations

. In the next Figures (Figs. S8-S11), subfigures (A4), (B4) and (C4) are representations similar to the one proposed by Stuhrmann in his 1974's seminal paper.[11] Such representations consist in plotting the squared radius of gyration of the casein micelle population $R^2_{g,CM}$ (*i.e.*, without the contribution of fat droplets) as a function of the reciprocal of the contrast of the micelles, $1/\sqrt{\rho_{CM}}$. This contrast is directly proportional to $1/\sqrt{I_{0,CM}/C}$, where $I_{0,CM}$ is the contribution of the casein micelles to the total intensity.

. The modeled values of $R^2_{g,CM}$ and $I_{0,CM}$ are simply given by:

$$I_{0,CM} = \sum_i I_{0,cm_i} \qquad \text{(S23)} \qquad \text{and} \qquad R^2_{g,CM} = \frac{\sum_i I_{0,cm_i} R^2_{g,cm_i}}{I_{0,CM}} \qquad \text{(S24)}$$

. The experimental values are calculated knowing the contribution of the fat droplets to the measured values of I_0 and R_g:

$$I_{0,CM} = I_0 - I_{0,fat} \qquad \text{(S25)} \qquad \text{and} \qquad R^2_{g,CM} = \frac{I_0 R_g^2 - I_{0,fat} R^2_{g,fat}}{I_{0,CM}} \qquad \text{(S26)}$$

. Note that there is a possible source of errors if equation S26 is applied to systems in which the apparent radii of gyration are very different. Indeed, if the *q*-range is set, one of the radii may get out of this range, because the curvature of the corresponding intensity becomes too strong (it goes to lower *q* values) or too weak (it goes to higher *q* values). In this case the Guinier plot reflects the curvature caused by one component of the dispersions only.

. There is another possible error if one of the curvatures becomes negative and the corresponding radius is still counted as positive in using equation S26. However, this is easily checked, since only the micelle apparent radius can become negative, and only when the SLD of the solvent becomes quite close to that of casein (i.e., in a very narrow range of D_2O content, see next figures).

. These additional sources of error may explain why, in the Stuhrmann representations, the points near zero contrast are away from the model predictions.

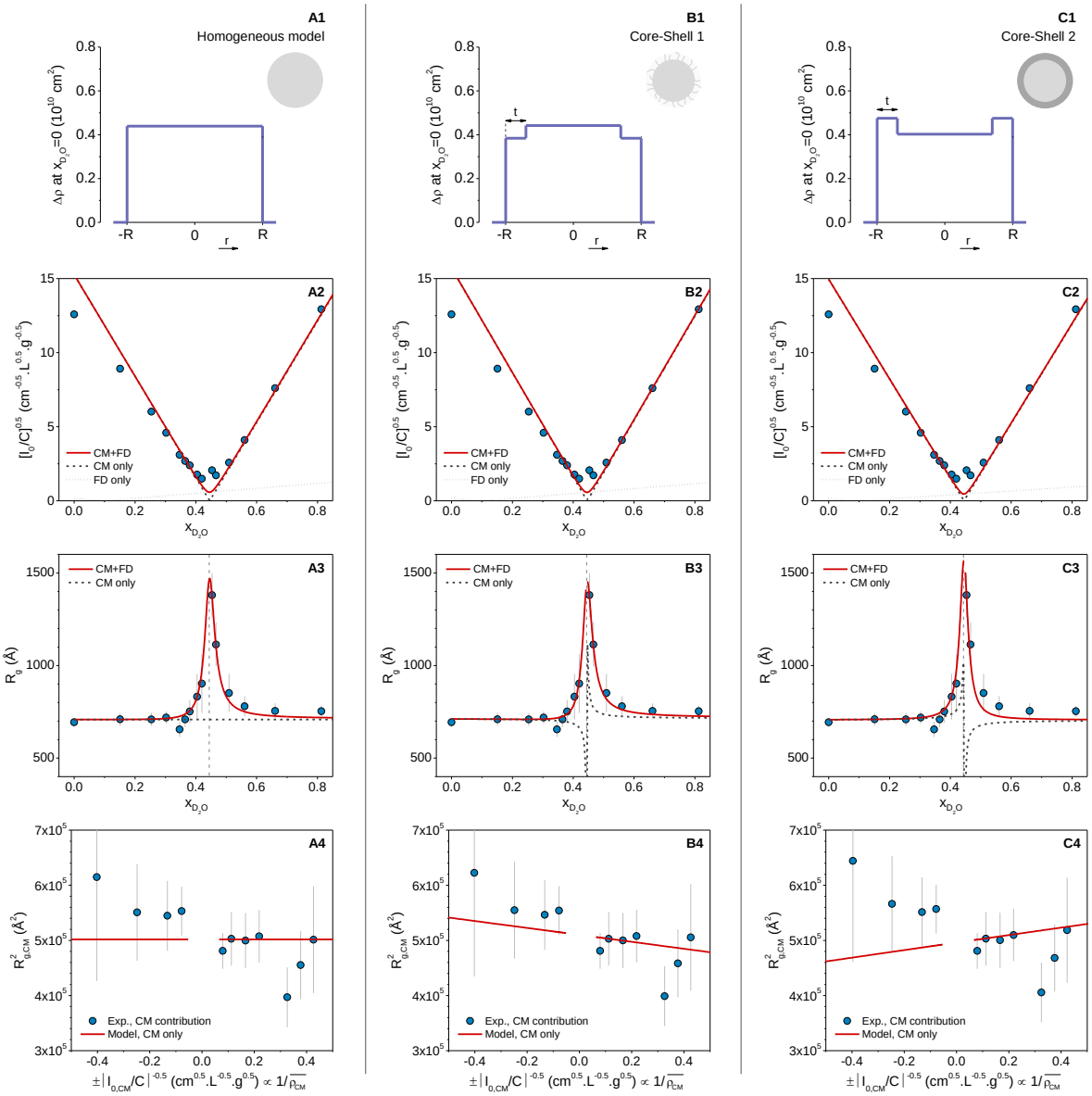

Fig. S8 Modeling the contrast variation of I_0 and R_g: casein micelles from fresh milk, FM_S1. CM = casein micelles, FD = fat droplets. Identical fits are obtained by considering phospholipid vesicles instead of fat droplets. Figures (A4), (B4) and (C4) use the representation of H.B. Stuhrmann and give the model and experimental variations of the apparent radius R_g of the casein micelle population as a function of the reciprocal of its average contrast. In these plots, all the experimental points for which the contribution of the fat droplets to the total scattered intensity is less than 8% are reported.

Table S2 The parameters used for the modeling: FM_s1 series.

CM from fresh milk / FM_S1			A - Homogeneous	B - Core-Shell 1	C - Core-Shell 2
Casein Micelle	$\overline{R}$ [Å]		843	861	832
	σ		0.14	0.13	0.15
	Core	ϕ_{cas_core}	0.167	0.167	0.167
		ϕ_{CaP_core}	0.013	0.019	0.005
	Shell	t [Å]	-	100	115
		ϕ_{cas_shell}	-	0.167	0.167
		ϕ_{CaP_shell}	-	-	0.027
w/ Fat Droplets		R_{fat} [Å]	1897	1878	2031
		$x_{fat/cas}$ [%]	2.1×10^{-3}	2×10^{-3}	1×10^{-3}
w/ Phospolipid Vesicles		R_{ves} [Å]	1499	1484	1613
		$x_{fat/cas}$ [%]	3.6×10^{-2}	3.6×10^{-2}	1.8×10^{-2}

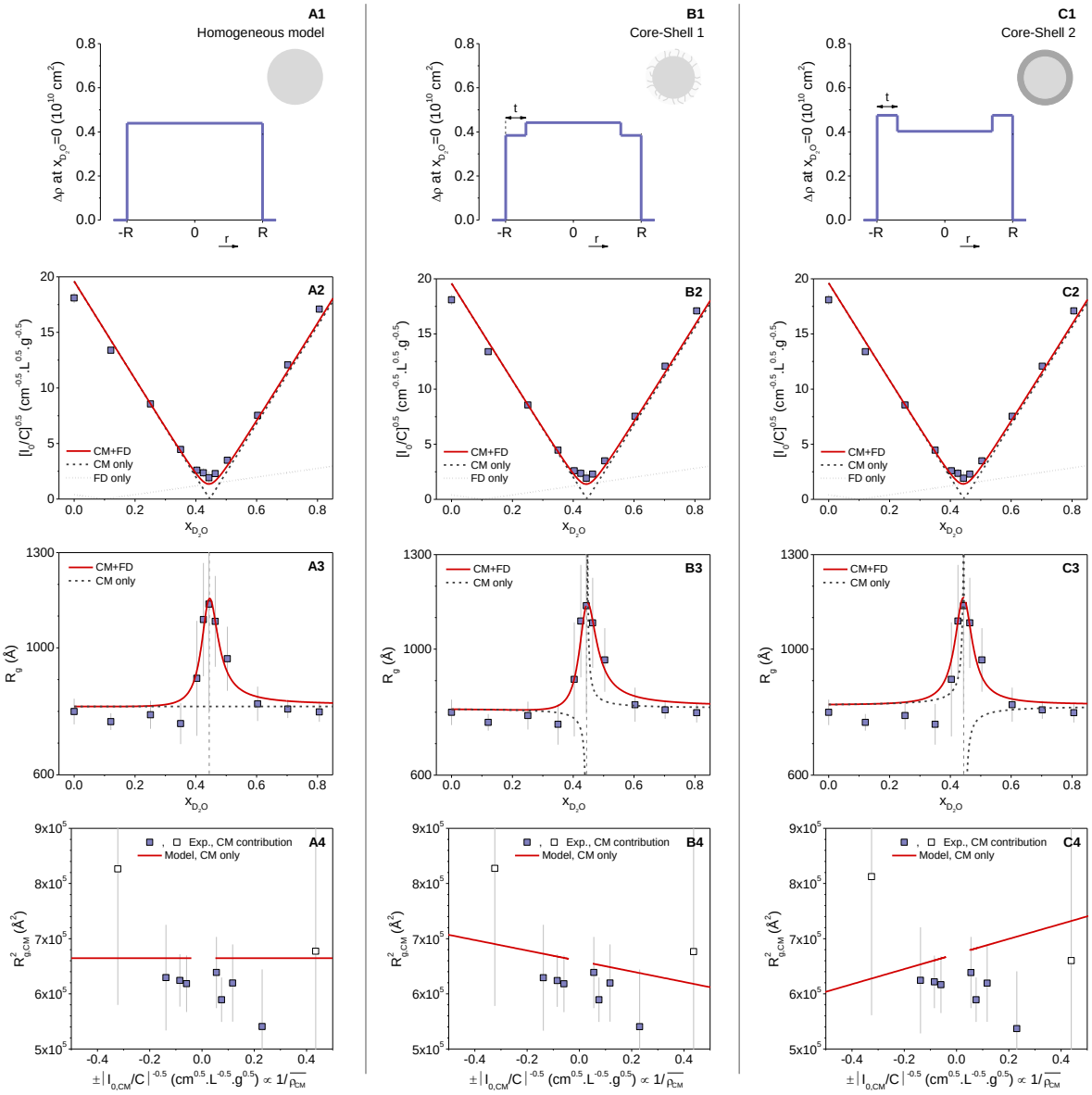

Fig. S9 Modeling the contrast variation of I_0 and R_g: casein micelles from fresh milk, FM_S2. CM = casein micelles, FD = fat droplets. Identical fits are obtained by considering phospholipid vesicles instead of fat droplets. Figures (A4), (B4) and (C4) use the representation of H.B. Stuhrmann and give the model and experimental variations of the apparent radius R_g of the casein micelle population as a function of the reciprocal of its average contrast. In these last plots, the empty symbols are the experimental points for which the contribution of the fat droplets to the total scattered intensity is more than 8%.

Table S3 The parameters used for the modeling: FM_s2 series.

CM from fresh milk / FM_S2			A - Homogeneous	B - Core-Shell 1	C - Core-Shell 2
Casein Micelle	$\overline{R}$ [Å]		1011	1010	1013
	σ		0.1	0.1	0.1
	Core	ϕ_{cas_core}	0.167	0.167	0.167
		ϕ_{CaP_core}	0.013	0.018	0.006
	Shell	t [Å]	-	100	115
		ϕ_{cas_shell}	-	0.167	0.167
		ϕ_{CaP_shell}	-	-	0.029
w/ Fat Droplets		R_{fat} [Å]	1492	1479	1500
		$x_{fat/cas}$ [%]	2.3×10^{-2}	2.5×10^{-2}	2.5×10^{-2}
w/ Phospolipid Vesicles		R_{ves} [Å]	1195	1175	1190
		$x_{fat/cas}$ [%]	0.27	0.34	0.35

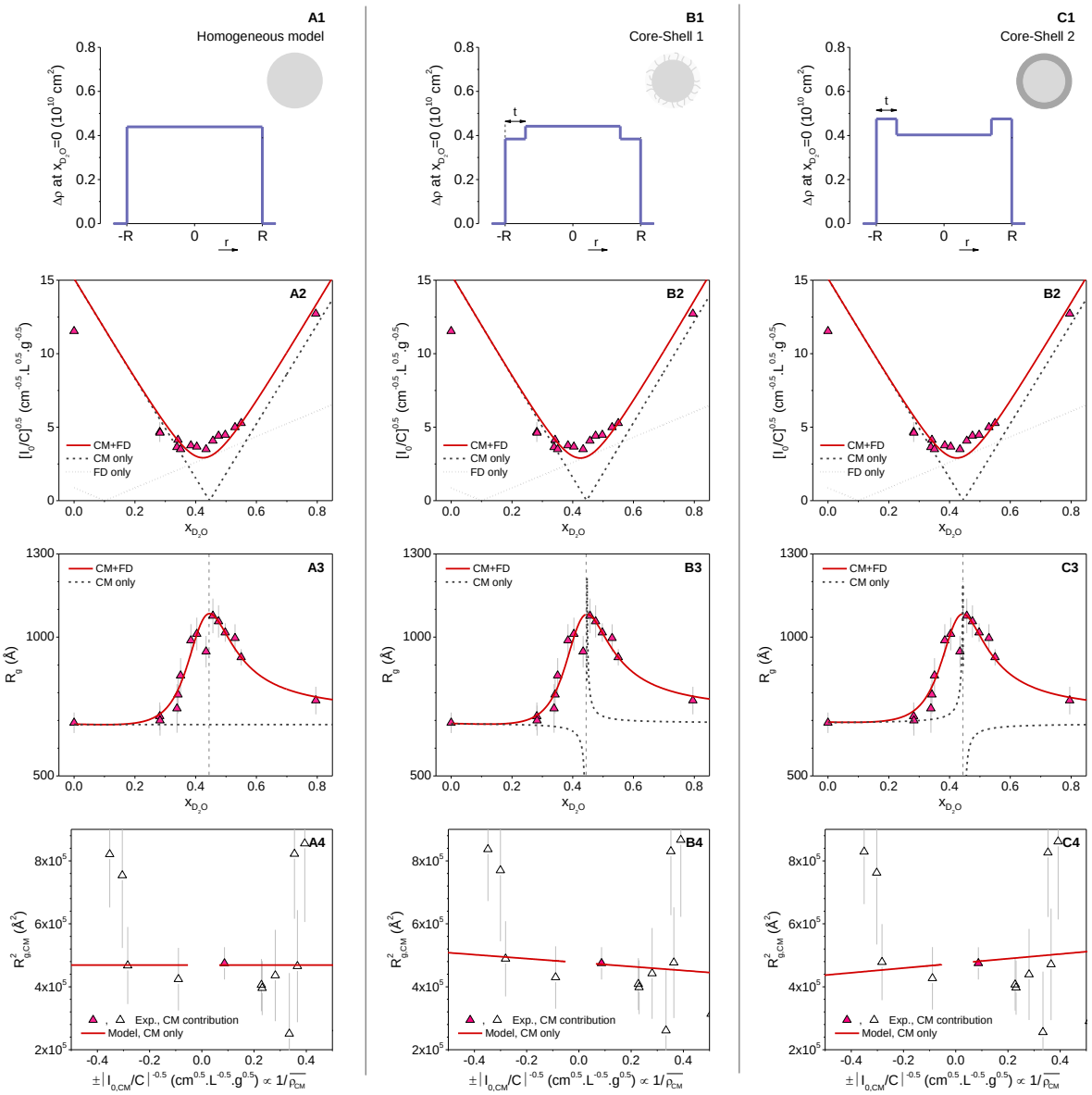

Fig. S10 Modeling the contrast variation of I_0 and R_g: casein micelles from casein powder, NPC_s1. CM = casein micelles, FD = fat droplets. Identical fits are obtained by considering phospholipid vesicles instead of fat droplets. Figures (A4), (B4) and (C4) use the representation of H.B. Stuhrmann and give the model and experimental variations of the apparent radius R_g of the casein micelle population as a function of the reciprocal of its average contrast. In these last plots, the empty symbols are the experimental points for which the contribution of the fat droplets to the total scattered intensity is more than 8%.

Table S4 The parameters used for the modeling: NPC_s1 series.

CM from NPC powder / NPC_S1			A - Homogeneous	B - Core-Shell 1	C - Core-Shell 2
Casein Micelle	$\overline{R}$ [Å]		849	860	851
	σ		0.1	0.1	0.1
	Core	ϕ_{cas_core}	0.167	0.167	0.167
		ϕ_{CaP_core}	0.013	0.019	0.005
	Shell	t [Å]	-	100	115
		ϕ_{cas_shell}	-	0.167	0.167
		ϕ_{CaP_shell}	-	-	0.027
w/ Fat Droplets		R_{fat} [Å]	1400	1397	1399
		$x_{fat/cas}$ [%]	0.14	0.14	0.14
w/ Phospolipid Vesicles		R_{ves} [Å]	1114	1113	1113
		$x_{fat/cas}$ [%]	1.8	1.7	1.8

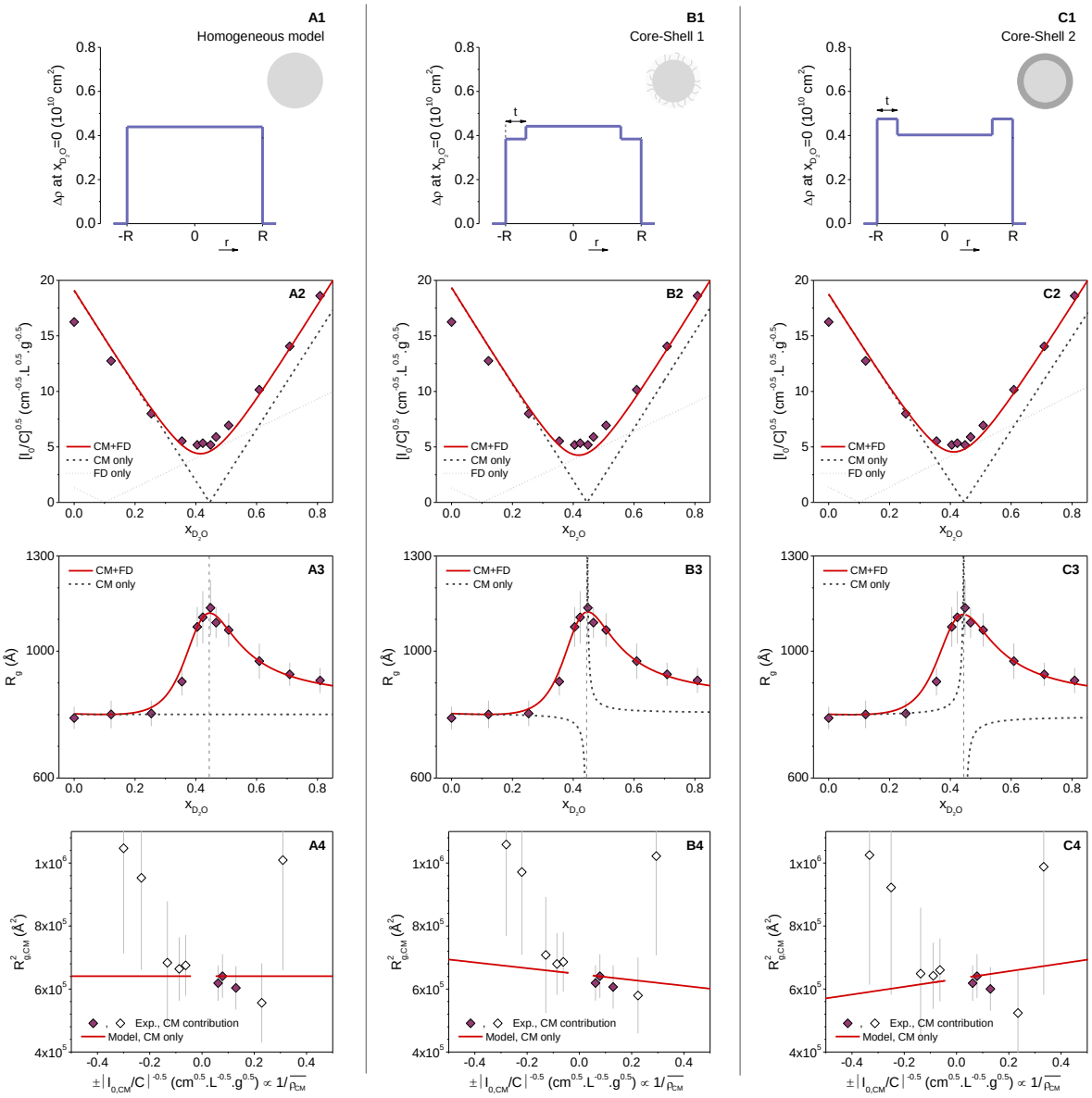

Fig. S11 Modeling the contrast variation of I_0 and R_g: casein micelles from casein powder, NPC_s2. CM = casein micelles, FD = fat droplets. Identical fits are obtained by considering phospholipid vesicles instead of fat droplets. Figures (A4), (B4) and (C4) use the representation of H.B. Stuhrmann and give the model and experimental variations of the apparent radius R_g of the casein micelle population as a function of the reciprocal of its average contrast. In these last plots, the empty symbols are the experimental points for which the contribution of the fat droplets to the total scattered intensity is more than 8%.

Table S5 The parameters used for the modeling: NPC_s2 series.

CM from NPC powder / NPC_S2			A - Homogeneous	B - Core-Shell 1	C - Core-Shell 2
Casein Micelle	$\overline{R}$ [Å]		993	1001	982
	σ		0.1	0.1	0.1
	Core	ϕ_{cas_core}	0.167	0.167	0.167
		ϕ_{CaP_core}	0.013	0.018	0.006
	Shell	t [Å]	-	100	115
		ϕ_{cas_shell}	-	0.167	0.167
		ϕ_{CaP_shell}	-	-	0.029
w/ Fat Droplets		R_{fat} [Å]	1446	1451	1441
		$x_{fat/cas}$ [%]	0.29	0.27	0.32
w/ Phospolipid Vesicles		R_{ves} [Å]	1149	1153	1145
		$x_{fat/cas}$ [%]	3.9	3.6	4.3

References

1. H. Fischer, I. Polikarpov, and A. F. Craievich, *Protein Sci.*, 2004, **13**, 2825–2828.
2. K. Gast, D. Zirwer, M. Muller-Frohne, and G. Damaschun, *Protein Sci.*, 1998, **7**, 2004–2011.
3. E. Robson, D. S. Horne, and D. G. Dalgleish, *J.Dairy Res.*, 1985, **52**, 391–400.
4. S. Marchin, J. L. Putaux, F. Pignon, and J. Léonil, *J.Chem.Phys.*, 2007, **126**, 045101–045110.
5. H. M. Farrell, H. Pessen, E. M. Brown, and T. F. Kumosinski, *J.Dairy Sci.*, 1990, **73**, 3592–3601.
6. H. M. Farrell Jr., R. Jimenez-Flores, G. T. Bleck, E. M. Brown, J. E. Butler, L. K. Creamer, C. L. Hicks, C. M. Hollar, K. F. Ng-Kwai-Hang, and H. E. Swaisgood, *J.Dairy Sci.*, 2004, **87**, 1641–1674.
7. C. G. de Kruif, T. Huppertz, V. S. Urban, and A. V. Petukhov, *Adv.Colloid Interface Sci.*, 2012, **171-172**, 36–52.
8. C. Holt, *Eur.Biophys.J.*, 2004, **33**, 421–434.
9. F. Gaucheron, *Reprod. Nutr. Dev.*, 2005, **45**, 473–483.
10. C. G. De Kruif, *J.Dairy Sci.*, 1998, **81**, 3019–3028.
11. H. Stuhrmann, *J.Appl.Crystallogr.*, 1974, **7**, 173–178.
12. D. G. Dalgleish, *Soft Matter*, 2011, **7**, 2265–2272.
13. D. G. Dalgleish, D. S. Horne, and A. J. R. Law, *Biochim. Biophys. Acta BBA - Gen. Subj.*, 1989, **991**, 383–387.
14. D. G. Dalgleish, J. Brinkhuis, and T. A. J. Payens, *Eur. J. Biochem.*, 1981, **119**, 257–261.
15. D. J. McMahon and B. S. Oommen, *J.Dairy Sci.*, 2008, **91**, 1709–1721.
16. D. J. McMahon and W. R. McManus, *J.Dairy Sci.*, 1998, **81**, 2985–2993.
17. R. Trejo, T. Dokland, J. Jurat-Fuentes, and F. Harte, *J.Dairy Sci.*, 2011, **94**, 5770–5775.
18. C. Holt, P. A. Timmins, N. Errington, and J. Leaver, *Eur.J.Biochem.*, 1998, **252**, 73–78.
19. M. Malmstem, B. Bergenstähl, L. Nyberg, and G. Odham, *J.Am.Oil Chem.Soc.*, 1994, **71**, 1021–1026.
20. A. Shukla, T. Narayanan, and D. Zanchi, *Soft Matter*, 2009, **5**, 2884–2888.