Electronic Supplementary Material (ESI) for RSC Advances.

Support Information

Lu et al

Dynamic Light Scattering

As larger particles with a larger hydrodynamic radius diffuse slower in bulk solution, the autocorrelation function can thus distinguish the particle sizes by calculating the diffuse time. However, this means it can not detect the precise shape of the measured molecules. In this work, the detection angle was 173° with respect to the incoming beam. Samples analyzed were contained in a 1 cm path length quartz cell, and the data were analyzed using Malvern Instruments Dispersion Technology Software. The polymer refractive index was taken to be 1.45 with an absorbance of 0.001. The viscosity and refractive index of water were taken as 0.8872 cPa and 1.330, respectively.

Small Angle Neutron Scattering

SANS measures the differential scattering cross section which contains information about particle size, shape and interactions between particles. The scattering intensity from a sample solution can be expressed as

$$I(q) = NV^2\Delta\rho^2 P(q)S(q) + B \qquad (9)$$

where q, N, V and $\Delta\rho$ are the momentum transfer, the number concentration of scattering particles, the volume of one scattering particle and the contrast of scattering length density between the scattering particles and the bulk solution, respectively. P(q) is the form factor that represents the interference of neutrons scattered from different parts of the same object while

S(q) indicates the interference of neutrons scattered from different objects. B is the background signal. As S(q) is determined by interference effect between scattering particles, it is dependent on the concentrations of scattered particles in solution. In this work, the concentration of keratin solution is so small that S(q) can be regarded as unity. Thus the scattered intensity I(q) is only determined by proportion to the variable P(q). Here, an ellipsoidal model was used as a fitting model for keratin solution. The output of the 2D scattering intensity function for oriented ellipsoids is given by [2]

$$P(q,\alpha) = \frac{scale}{V} f^2(q) + bkg \tag{10}$$

$$f\ (q) = \frac{3(\Delta\rho)V(\sin\left[qr\left(R_a,R_b,\alpha\right)\right] - qrcos[qr\left(R_a,R_b,\alpha\right)]}{qr\left(R_a,R_b,\alpha\right)^3} \tag{11}$$

$$r\left(R_a,R_b,\alpha\right) = [{R_b}^2 sin^2\alpha + {R_a}^2 cos^2\alpha]^{1/2} \tag{12}$$

where α is the angle between the axis of the ellipsoid and the beam; V is the volume of the ellipsoid; R_a and R_b are the lengths along and perpendicular to the rotation axis of the ellipsoid respectively; Δρ is again the difference of scattering length density between the scattered particle and bulk solution.

The fitting process uses the iterated method until an acceptable fit is produced. By comparing the calculated profile from the ellipsoidal model with the measured scattering profile, the geometrical shape and size of the protein molecule in bulk solution can be obtained. To make SANS data comparable to the NR work, the measurements are conducted at pH 6 with NaCl concentration fixed at 5 mMs.

The normal ellipsoidal model did not contain the volume fraction of the scattered particles in solvent. Thus we first used a Hayter-MSA structure factor that was introduced to complement

the fitting parameters needed in this study. In general, the mean spherical approximation (MSA)[3] method can describe the structure of a dispersion of charged colloidal particles interacting through a screened Coulombic potential. It was found from MSA fittings that the scattered particle in solvent was very weakly charged and could be treated as no charge under low salt concentrations. Thus the MSA method is no longer appropriate in high salt systems. In the zero charge limits, the Percus-Yevick hard sphere solution can then be recovered[4]. It assumes the inter-particle potential is given by

$$U(r) = \begin{cases} \infty, r < 2R \\ 0, r > 2R \end{cases} \quad (13)$$

The Percus-Yevick method provides very accurate approximation for particle volume fractions of up to ~0.45 [5].The hard sphere structure method combined with ellipsoidal model calculates the inter-particle structure factor for mono-disperse spherical particles that interact through hard sphere interactions. In this case, the form factor S(Q) can no longer be treated as unity but be approximated as [6]

$$S(q) = [1 + \frac{24\eta_{HS}G(2R_{HS}q)}{2R_{HS}q}]^{-1} \quad (14)$$

The two variables in the S(q) equation are the hard sphere volume fraction η_{HS} and the corresponding hard sphere radius R_{HS}. In data analysis, the fitting parameters contained the background, the two radii of the ellipsoid, the scale factor, the volume fraction and the scattering length densities of the scattered particles and solvents. The scale factor and scattering length density of solvent were fixed at 1 and 6.35×10^{-6}, respectively. Thus the only variables left in the fitting process were the ellipsoidal radii a and b, the scattering length density and the volume fraction of the scatters.

Table SI1 Comparison of the scattering length density (SLD) values as calculated from three different keratins and BSA. This illustrates that despite different sequences their SLDs in a given solvent (H_2O, D_2O or their mixtures) are virtually the same. [7]

	SLD in D_2O ($\times 10^{-6}$ Å^{-2})	SLD in H_2O ($\times 10^{-6}$ Å^{-2})	SLD in NRW ($\times 10^{-6}$ Å^{-2})	Molecular Weight ($\times 10^3$g/mol)	Molecular Volume ($\times 10^3$Å^3)
Human Hair Keratin (a3)	3.46	1.94	2.07	46.2	54.9
Mouse Hair Keratin (a1)	3.42	1.94	2.06	47.9	56.8
Sheep Wool Keratin (47.6)	3.40	1.91	2.03	47.3	56.6
Bovine Serum Albumin	3.45	1.97	2.09	69.3	83.4

The sequence of the sheep wool keratin (47.6 kD) taken from the work by Wilson et al (Wilson, B.W.; Edwards, K.J.; Sleigh, M.J.; Byrne, CR.; Ward, K.A., *Gene* 1988, *73*, 21-31) and the helical forming segments were marked in red (bold) following the work by Yu et al (Yu, J.; Yu, D.-W.; Checkla, D. M.; Freedberg, I. M.; Bertolino, A. P. *Journal of investigative dermatology* 1993, *101*, 56S)

MSPAPCLPALSPASSCSSAPCVPSSCCGTT 29

LPGACAIPASVGSCATPCGGSPAGAG**LGTM** 59

GPLAAALASTLGLVAGLGAGAAGLGAAILG 89

ASGGGGPLVC**PATGSTPATIGGLGGLILCG** 119

LSGAAALVVGIAAALLASAAPATLTGTGVS 149

LAGLVGAALAGLAAILAGLTLCLSALGAAV 179

GSLLGGLICLLGAHGGGVATLASGLGAALA 209

VGVAAAPTVAL**AHVLAGTAAGTGALVGT**AA 239

AAVGGT**TIAGTGGLALGVVSSSGGLGSCGA** 269

GIIGLAATVAALGVGLGAGHALAASLGATL 299

TGTGAATSCGLAGVGSLIVSVGSGLAGIAS 329

ALGAGAGGTGVLLAVAAALGCGIATTAGLL 359

ASGACLLPCAPCATTATCGLPIGPCISA 387

PCVSATACGPCATPVH 403

Table SI2 Two layer model fits to the reflectivity profiles of keratin measured with 5 mM and 500 mM NaCl under NRW.

[NaCl] (mM)	layer	Thickness (Å) ±1	SLD ($\times10^{-6}$ Å^{-2}) ±0.02	Volume fraction ±0.02	Layer mass (mg/m^2) ±0.03	Total mass (mg/m^2) ±0.03
500	1	19	1.05	0.52	1.3	1.65
	2	25	0.2	0.1	0.35	
5	1	25	1.1	0.55	1.9	2.29
	2	26	0.22	0.11	0.39	

Table SI3 Ellipsoidal model fits to the scattering profiles of keratin measured with 5 mM NaCl in D_2O.

Ellipsoid model	Keratin 1 mg/ml	Keratin 0.5 mg/ml	Keratin 0.25 mg/ml
SLD Ell ($\times10^{-6}$ Å^{-2}) ±0.02	4.48	4.48	4.48
SLD Solv ($\times10^{-6}$ Å^{-2}) ±0.02	6.35	6.35	6.35
Radius a (Å) ±2	60	60	59
Radius b (Å) ±8	138	139	138

Table SI4 Ellipsoidal model fits to the scattering profiles of keratin measured at 1 mg/ml, with NaCl changed from 5 mM to 0.5 M in D_2O, pD 6.3.

Ellipsoid model	[NaCl] = 5mM	[NaCl] = 20mM	[NaCl] = 40mM	[NaCl] = 100mM	[NaCl] = 200mM	[NaCl] = 500mM
SLD Ellip ($\times10^{-6}$ Å^{-2}) ±0.02	4.48	4.5	4.53	4.58	4.69	5.28
Radius a (Å) ±2	60	59	59	55	50	37
Radius b (/Å) ±8	138	168	201	260	280	350
Volume fraction	0.001	0.001	0.001	0.001	0.001	0.001
Particle volume ($\times10^{-6}$/Å^3) ±0.02	1.94	2.37	2.55	3.54	2.93	2.01
Volume fraction of keratin	0.58	0.59	0.6	0.61	0.61	0.39
No. of keratin molecules in one aggregate	**21**	**25**	**27**	**38**	**31**	**14**

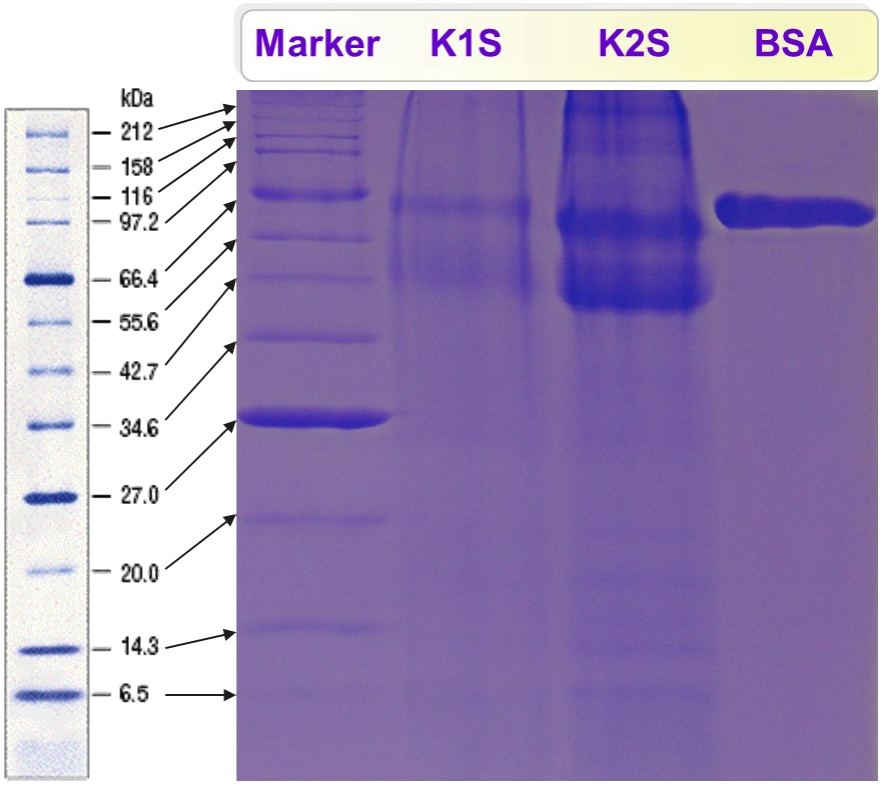

Figure SI1 SDS PAGE of keratin extracted from the method utilized in this work (K1S) and that extracted by a slightly modified method without SDS (K2S).

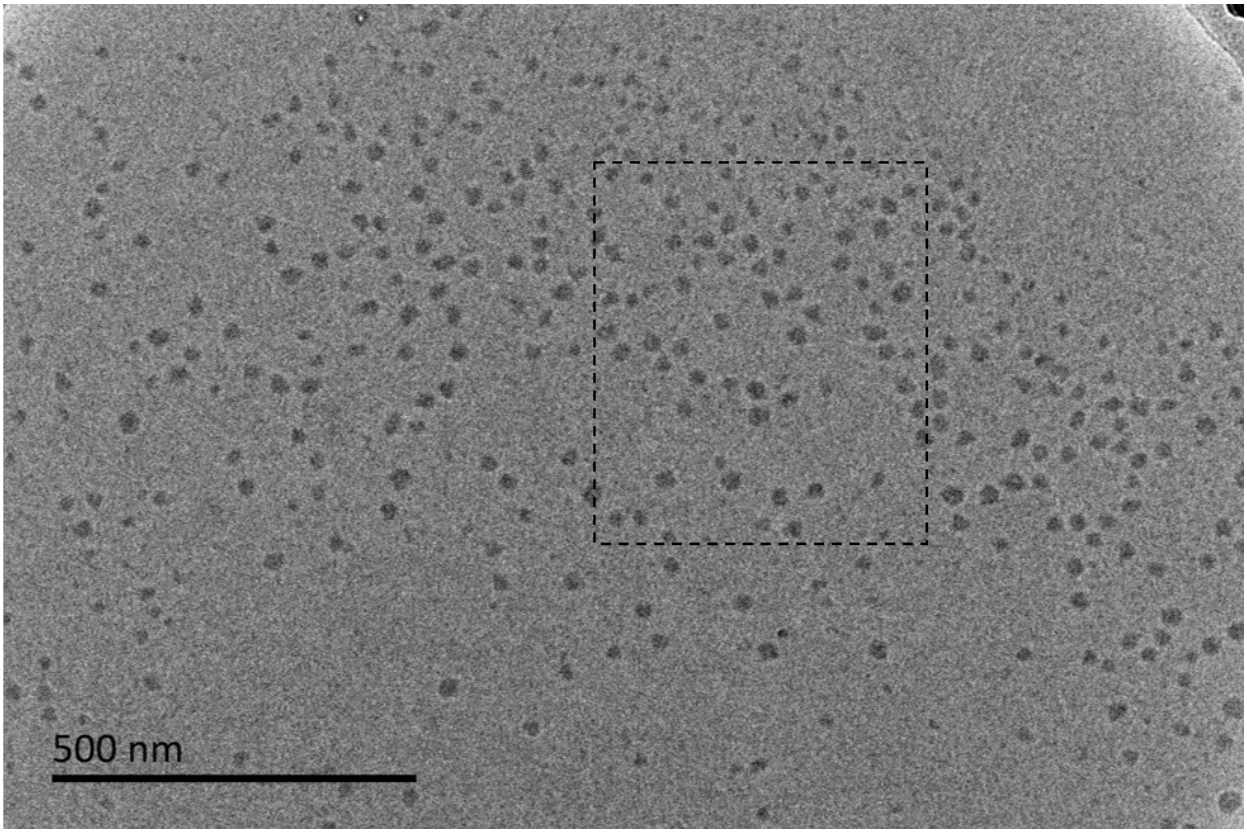

Figure SI2 Cryo-EM image of 1 mg/ml keratin in 5 mM NaCl solution. The dashed line represented the area that was shown in Figure 4.

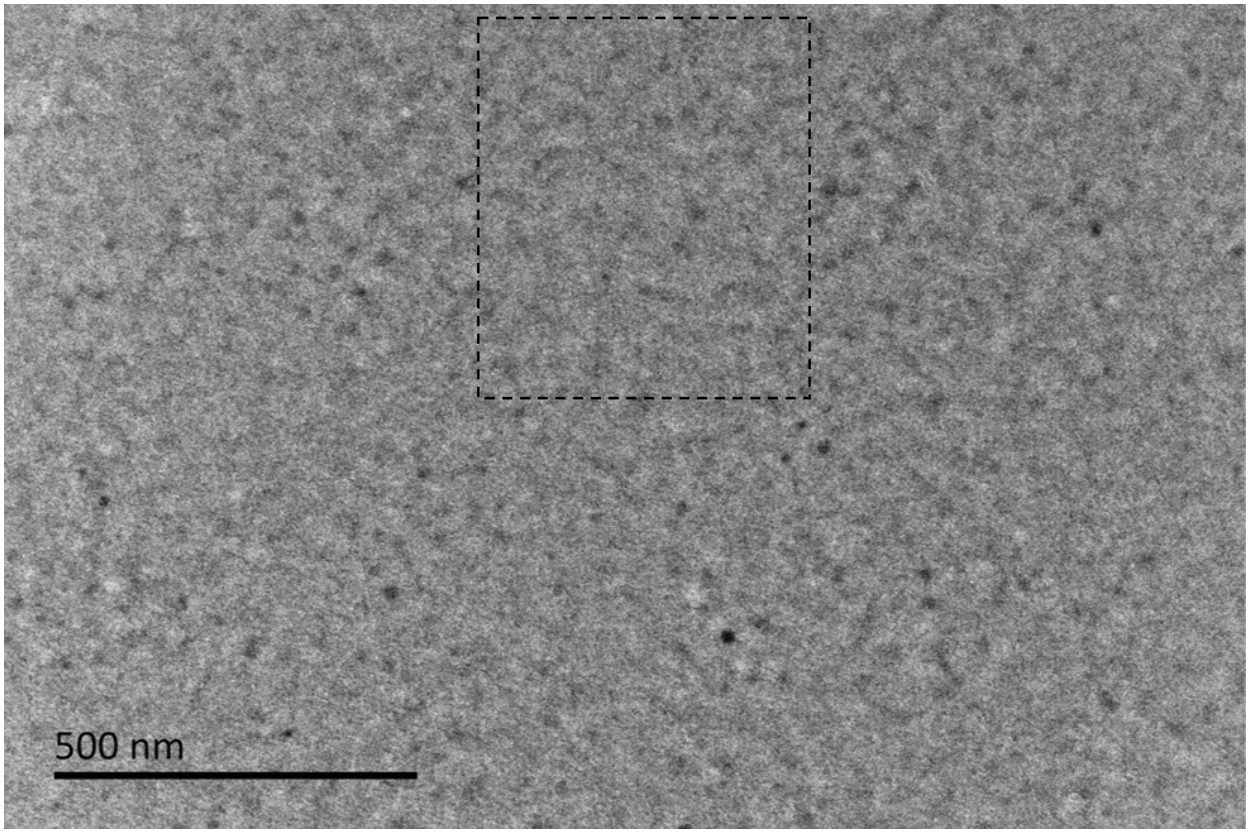

Figure SI3 Cryo-EM image of 1 mg/ml keratin in 500 mM NaCl solution. The dashed line represented the area that was shown in Figure 7.

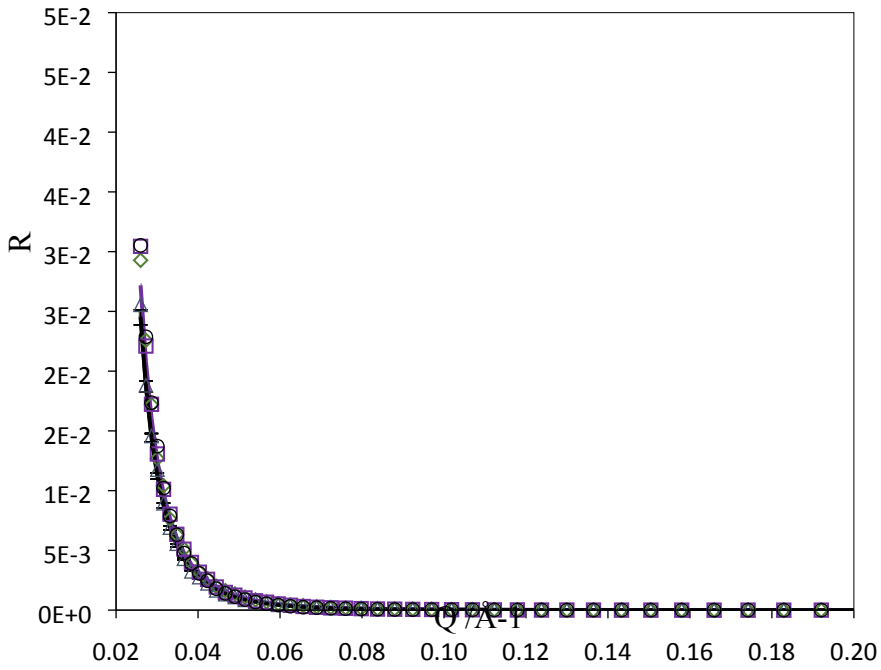

Figure SI4 NR reflectivity profiles measured from keratin adsorption onto the D_2O surface at concentrations of 3×10^{-3}(◊), 1×10^{-2} (□), 3×10^{-2} (Δ), 0.3 (o) mg/ml. Solutions were prepared in 5 mM NaCl at pH 6.7±0.1. Solid lines indicate the best fits to the data measured.

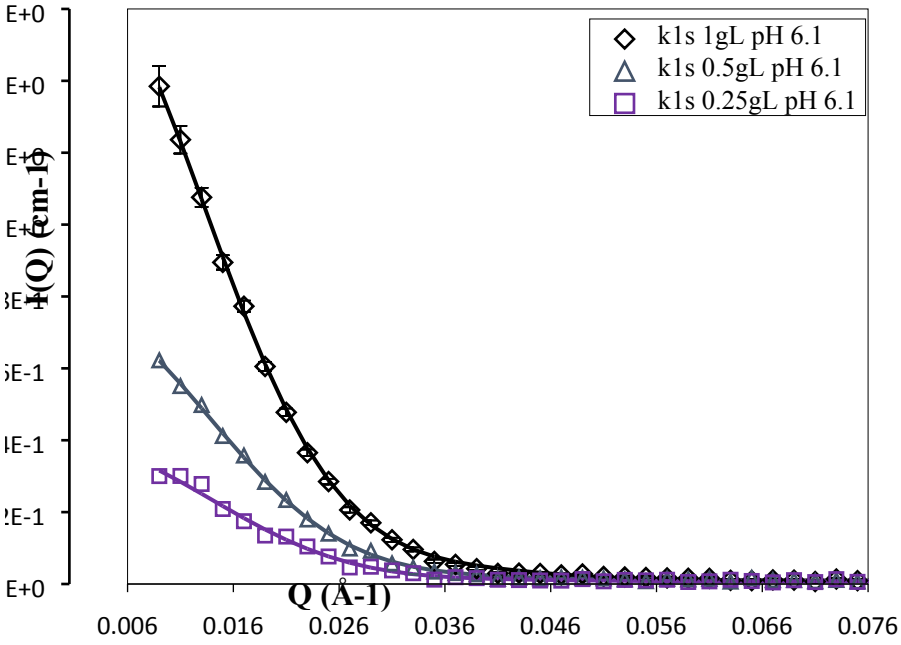

Figure SI5 The effect of keratin concentration on its aggregation studied at 0.25 (□), 0.5 (Δ) and 1 (◊) mg/ml by SANS. All solutions were prepared in buffers with 5 mM NaCl at pH 6.1. All the scattering curves were able to be fitted with the same size and shape of aggregates, indicating that they remained the nanostructure over this concentration range.

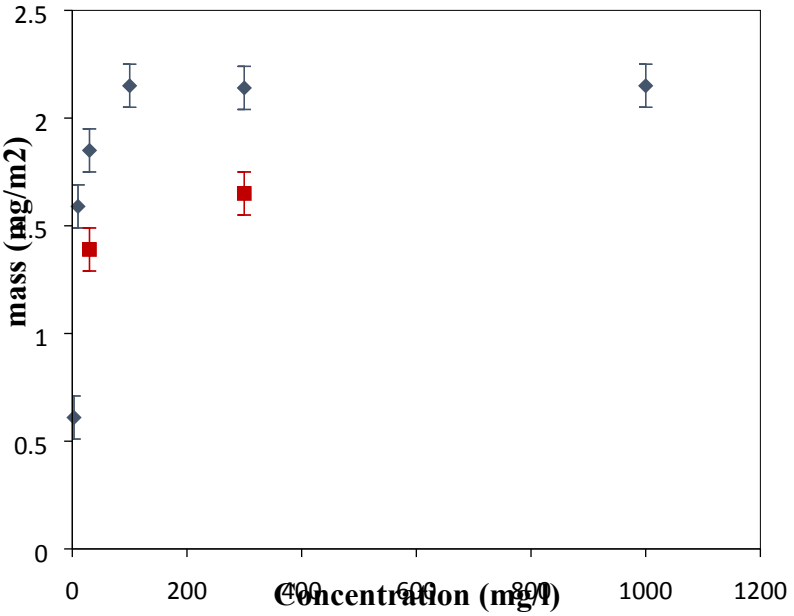

Figure SI6 Keratin surface adsorbed amounts obtained from NR plotted against solution concentration, with all data measured after 60 minutes adsorption. The blue diamond dots are measured at keratin solution with 5 mM NaCl at concentrations from 0.003 to 1 mg/ml while the red square dots are keratin solution with 0.5 M NaCl at concentrations of 0.03 and 0.3 mg/ml.

References

(1) Vertesse, B. G.; Magazu, S.; Mangione, A.; Migliardo, F.; Brandt, A. *Macromolecular bioscience* **2003**, *3*, 477.
(2) Feigin, L.; Svergun, D.; Plenum Press, New York.
(3) Hayter, J. B.; Penfold, J. *Molecular Physics* **1981**, *42*, 109.
(4) Percus, J. K.; Yevick, G. J. *Physical Review* **1958**, *110*, 1.
(5) Frenkel, D.; Vos, R.; De Kruif, C.; Vrij, A. *The Journal of chemical physics* **1986**, *84*, 4625.
(6) Stieger, M.; Pedersen, J. S.; Lindner, P.; Richtering, W. *Langmuir* **2004**, *20*, 7283.
(7) (a) Yu, J.; Yu, D.-W.; Checkla, D. M.; Freedberg, I. M.; Bertolino, A. P. *Journal of investigative dermatology* **1993**, *101*, 56S(b) Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.-i. *Biochemical and biophysical research communications* **1990**, *173*, 639.