IL NUOVO CIMENTO **39 C** (2016) 304
DOI 10.1393/ncc/i2016-16304-0

Colloquia: SWGM 2015

Electrostatics promotes molecular crowding and selects the aggregation pathway in fibril-forming protein solutions

S. Raccosta([1]), M. Blanco([2]), C. J. Roberts([2]),
V. Martorana([1]) and M. Manno([1])([*])

([1]) *Institute of Biophysics, National Research Council of Italy (CNR) - 90126 Palermo, Italy*
([2]) *Department of Chemical and Biomolecular Engineering, University of Delaware Newark, Delaware 19716, USA*

received 19 July 2016

Summary. — The role of intermolecular interaction in fibril-forming protein solutions and its relation with molecular conformation are crucial aspects for the control and inhibition of amyloid structures. Here, we study the fibril formation and the protein-protein interactions for two proteins at acidic *p*H, lysozyme and α-chymotrypsinogen. By using light scattering experiments and the Kirkwood-Buff integral approach, we show how concentration fluctuations are damped even at moderate protein concentrations by the dominant long-ranged electrostatic repulsion, which determines an effective crowded environment. In denaturing conditions, electrostatic repulsion keeps the monomeric solution in a thermodynamically metastable state, which is escaped through kinetically populated conformational sub-states. This explains how electrostatics acts as a gatekeeper in selecting a specific aggregation pathway.

1. – Introduction

The conformation of protein molecules and the thermodynamic and environmental properties of protein solutions are both crucial aspects for determining protein stability and controlling self-assembly [1]. The structural details at molecular level typically define the morphologies of supramolecular assemblies. For instance, a well-established relation exists between the capability of a protein to form amyloid fibrils and its molecular conformation and sequence, or more specifically its amount of β-sheet structures [2]. On the other hand, the environmental conditions of protein solutions are important to determine the intermolecular interactions responsible for aggregation [3]. The physics of intermolecular interactions has been extensively studied for an important type of

([*]) Corresponding author. E-mail: mauro.manno@cnr.it

ordered aggregation, namely crystallization. The formation of crystals of native proteins is prompted by the onset of mild protein attractive interactions, along with conditions stabilizing the native conformation [4-7]. These studies were pioneered by the observation that proteins, which are prone to crystallisation, are encompassed in the so-called crystallisation slot. This slot corresponds to a window of slightly negative values of the second virial coefficient B_{22}, which is a parameter related to the solvent-mediated protein-protein potential of mean force U_{22} [8].

Several computational and theoretical methods have been developed to rationalize and predict fibrillation rates by taking into account hydrophobicity, electric charge or other properties, as well as the role of the conformation state of the protein [9, 10]. On the other hand, the intermolecular interaction potentials, which determine protein attraction and cause solution instability, have only been explored on limited cases [11-20].

In order to address this problem, we studied two well known proteins, Hen Egg White Lysozyme (HEWL) and bovine α-chymotrypsinogen A (aCgn) [21]. Both proteins are known to form fibrils at low *p*H, where they are positively charged, notwithstanding the intermolecular electrostatic repulsion, as observed for many different proteins [22-24]. Here, we present a new analysis and review our previous measurements on the protein-protein interaction for these two systems, by showing how electrostatic repulsions may play a key role in determining crowding conditions at moderately high concentrations, and in selecting a specific aggregation pathway [25, 26].

2. – Concentration fluctuations in metastable lysozyme solutions

HEWL is known to form amyloid fibrils at low *p*H and high temperature [27, 28]. At *p*H 2, HEWL is extremely charged, with $Z = +17$. Also, in order to highlight the effect of electrostatics, we started from low ionic strength where electrostatic interaction are not screened. In such a condition HEWL has a reduced thermal stability with respect to physiological *p*H. The mid point temperature for thermal unfolding, measured by DSC, is at about 71 °C at *p*H 7, while it reduces to 58 °C at *p*H 2 [25]. Also, the calorimetric peak is much broader at acidic *p*H, indicating a greater protein flexibility and/or the existence of multiple states during thermal unfolding [29].

The formation of amyloid fibrils occurs upon incubation at high temperatures for several days. Figure 1 shows typical AFM images of HEWL fibrils formed after 1 week of incubation at 65 °C of 18.4 gL^{-1} HEWL solution (HEWL was purchased from Sigma Chemical Co.). After 1:10000 dilution, a 20 μL drop was deposited in a mica substrate and dried. Images were collected by a Multimode Nanoscope V atomic force microscope (Veeco Instruments, Santa Barbara, CA, USA) operating in air tapping mode and

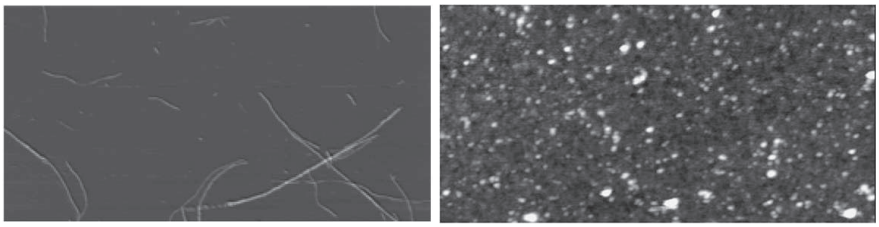

Fig. 1. – AFM image of 18.4 gL^{-1} lysozyme solution incubated at 65 °C for 7 days. Left panel: no NaCl, width = 5 μm. Right panel: 20 mM NaCl, width = 2 μm.

equipped with rigid cantilevers (Nanosensor PPP-NCHR-50, resonance 330 kHz). Note that the amount of fibrillar structure is considerably reduced if the solution is incubated at higher temperature (*e.g.* 70 °C) or higher salt concentration (*e.g.* 20 mM NaCl). In such cases, one observes only amorphous globular aggregates (fig. 1).

Due to the long latency before HEWL fibrillation, we were able to study the interaction occurring at both low and high temperatures when the solution is still monomeric. In a previous work [25], we performed experiments at different temperatures up to 70 °C. Figure 2(a) shows the excess Rayleigh ratio R_{90}, measured by scattered intensity at 90° scattering angle and normalized by the protein apparent molecular weight M_w and the constant $K = (2\pi\tilde{n}[\mathrm{d}\tilde{n}/\mathrm{d}c_2]\lambda^{-2})^2 N_A^{-1}$, which is determined by the solution refractive index $\tilde{n}$, the refractive index increment $[\mathrm{d}\tilde{n}/\mathrm{d}c_2]$, the laser wavelength λ, and the Avogadro number N_A.

In the classical approach, the excess Rayleigh ratio R_{90} is directly proportional to the protein mass concentration c_2 and to an equilibrium thermodynamic parameter, the osmotic isothermal compressibility κ_T (note that for small protein molecules the Rayleigh ratio does not depend upon the scattering angle) [30]

$$\frac{R_{90}}{K} = M_w c_2 \frac{\kappa_T}{\kappa_0}, \tag{1}$$

where M_w is the weight-averaged protein mass and $\kappa_0 = (\rho_2 k_B T)^{-1}$ is the ideal gas compressibility, which depends upon the protein number concentration ρ_2, the temperature T and the Boltzmann factor k_B. The inverse of the compressibility can be expanded in powers of concentration (the so-called virial expansion), and in the limit of infinite dilution the expansion may be truncated to the first two terms

$$\frac{\kappa_0}{\kappa_T} = 1 + 2B_{22}c_2. \tag{2}$$

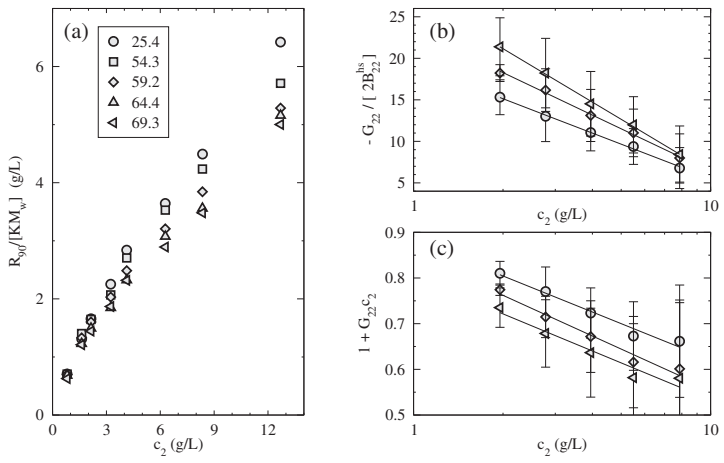

Fig. 2. – Light scattering quantities as a function of lysozyme concentration c_2 for different temperatures. (a) Normalized Rayleigh ratio. (b) Normalized G_{22} values. (c) Concentration fluctuations calculated from fitted values of G_{22}. Straight lines are guides to the eye.

The parameter B_{22} is the second virial coefficient; it is related to the solvent-mediated protein-protein potential of mean force U_{22}, upon integration over the distance r of the protein centres of mass [31]

$$B_{22} = -2\pi \frac{N_A}{M_0} \int_0^\infty \left(\mathrm{e}^{-U_{22}/k_B T} - 1 \right) r^2 \mathrm{d}r, \tag{3}$$

where M_0 is the protein molecular mass.

We have used expressions (1) and (2) to fit light scattering data and determine the weight-averaged mass and the second virial coefficient. The weight-averaged mass is consistent with the actual molecular mass of lysozyme, 14.3 kDa. B_{22} is highly positive implying the prevalence of repulsion over attraction, and constant up to 50–55 °C, where it exhibits a slight increase. The value found for B_{22} exceeds by more than one order of magnitude the equivalent hard-sphere value B_{22}^{hs} related to the effective interaction volume V_I ($B_{22}^{\mathrm{hs}} = 4V_I(N_A/M_0)$). We found that for ellipsoidal molecules, a good estimation of the interaction volume may be taken by assuming a sphere with a radius equal to the hydrodynamic radius [15], which in the present case is 1.65 nm. In Raccosta *et al.* [25], we were able to match the experimental B_{22} value by adding the expected electrostatic repulsion, which can be modelled by using Debye-Hückel potential and the effect of ion condensation [32].

Given the intrinsic assumptions of eq. (2), B_{22} is only related to protein-protein interactions at dilute conditions, but it becomes less informative at high protein concentrations. In order to extend the information contained in our data to higher concentrations, one needs to re-examine the actual dependence of scattering intensity upon thermodynamic quantities, as in Blanco *et al.* [33]. If one considers the effect of fluctuations of all species in solution except protein molecules, the excess Rayleigh ratio is related to the fluctuations in protein number N_2 via the relation

$$\frac{R_{90}}{K} = M_w c_2 \frac{\langle N_2^2 \rangle - \langle N_2 \rangle^2}{\langle N_2 \rangle}. \tag{4}$$

In the grand-canonical ensemble concentration fluctuations may be conveniently expressed in terms of Kirkwood-Buff integrals G_{22} [34]

$$\frac{\langle N_2^2 \rangle - \langle N_2 \rangle^2}{\langle N_2 \rangle} = 1 + G_{22} c_{22}, \tag{5}$$

where G_{22} is related to the orientation-averaged protein-protein pair correlation function g_{22} integrated over the distance r between the protein centres of mass

$$G_{22} = 4\pi \frac{N_A}{M_0} \int_0^\infty (g_{22} - 1)\, r^2 \mathrm{d}r. \tag{6}$$

Therefore, $G_{22}c_2$ represents the excess concentration around a protein relative to an ideal solution with bulk concentration c_2. In the limit of infinite dilution, or more precisely in the limit of both low concentration and weak interaction ($G_{22}c_2 \ll 1$), the pair correlation function coincides with the Boltzmann factor of the potential of mean force, and G_{22} is equal to $-2B_{22}$.

As in Blanco *et al.* [33], we adopt a novel approach to compute the value of G_{22} as a function of protein concentration without any *a priori* assumption for the functional form of the potential of mean force. One may consider a series of small "windows" of c_2, and locally fit G_{22} by using expressions (4) and (5) for a given window, assuming that G_{22} is effectively constant in that range. The validity of such a fitting procedure was demonstrated by a Taylor expansion for eq. (5) around each c_2 [33]. Figure 2(b) displays for selected temperatures the value of G_{22} (with respect to the hard-sphere second virial coefficient $B_{22}^{\rm hs}$) obtained by such a model-free approach. Note that G_{22} is normalized by the hard-sphere value at infinite dilution ($-2B_{22}^{\rm hs}$). Under such scale, positive values (larger than unity) indicate net repulsive protein-protein interactions. Also, the protein concentration fluctuations are reported in fig. 2(c) in terms of $1 + G_{22}c_2$. Thus, the main observation is that fluctuations are reduced by increasing concentration due to electrostatic repulsion, and this repulsion is enhanced at high temperatures. Also in fig. 2(b), (c) we note that a relative increment of concentration (the axis is here on a lin-log scale) equally affects fluctuations. We may argue that such a concentration-dependent regime is due to an effective *crowding* due to electrostatic repulsion.

3. – Screened electrostatic repulsion in α-chymotrypsinogen solutions

In order to validate the actual role of electrostatic repulsion at high concentration, we focus on another protein able to form fibril at acidic *p*H, aCgn [35-37]. In a previous work [26], we performed experiments at *p*H 3.5, 25 °C and at different concentrations of NaCl, up to 100 mM, to progressively screen the effect of electric charge. Analogously to lysozyme experiments, we measured the Rayleigh ratio at different protein concentrations (fig. 3(a)). We used the same approach to compute the Kirkwood-Buff integral G_{22}, shown in fig. 3(b) with respect to the hard-sphere second virial coefficient (calculated by assuming a hydrodynamic radius of 2 nm). Figure 3(c) displays the concentration fluctuations in terms of G_{22}. As fig. 3(a) illustrates, the decrease of the

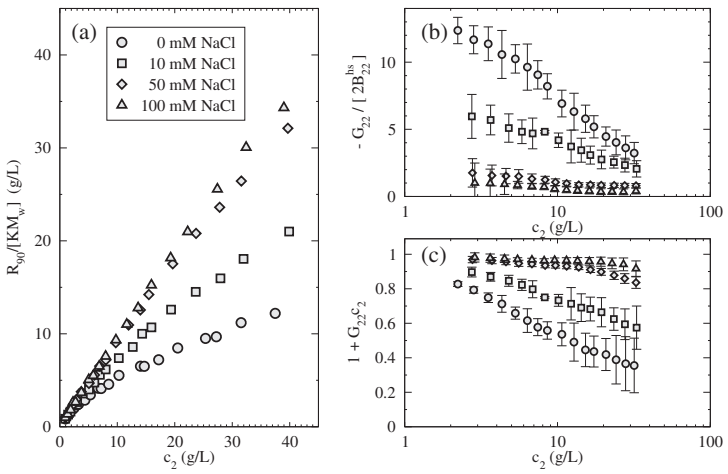

Fig. 3. – Light scattering quantities as a function of α-chymotrypsinogen concentration c_2 for different temperatures. (a) Normalized Rayleigh ratio. (b) Normalized G_{22} values. (c) Concentration fluctuations calculated from fitted values of G_{22}.

curvature of R_{90} vs protein concentration by increasing salt concentration clearly indicates the screening of electrostatic interactions. Also, we observe again the damping of concentration fluctuations as protein concentration is increased. Without salt, the same concentration-dependent crowding-like behaviour found for lysozyme can be observed. It is interesting that by addition of salt, such regime is shifted towards higher concentrations, confirming that electrostatics is the actual origin of such an effect.

4. – Kinetic pathways out of metastability for lysozyme solutions

In the case of Lysozyme at acidic *p*H, we have shown that electrostatic repulsion keeps the solution of monomeric proteins in a thermodynamic state, which is stable at low temperatures, and metastable at high temperatures. Indeed, at high temperatures the fibrillar state is definitively more stable than monomeric solution, since proteins eventually aggregates. Given that protein-protein interactions are predominately repulsive at these conditions (as shown in sect. 2), there is an outstanding question to be answered: what is the driving force for aggregation? We addressed this question by going beyond the framework of colloidal theory and considering the changes in protein conformation caused by an increase of temperature.

In our previous work [25], we used far-UV Circular Dichroism (CD) to observe the changes in the secondary structure caused by protein incubation at high temperature. The CD spectra have been analysed by using CONTINLL software [38], which allows to determine the fraction of alpha helical content f_α in the overall protein secondary structure. Upon incubation at 60 °C, lysozyme undergoes a partial and progressive unfolding, as monitored by the changes in alpha helical fraction with respect to secondary structure at $t = 0$: $1 - f_\alpha(t)/f_\alpha(0)$ (fig. 4). Also, we performed intrinsic fluorescence experiments to probe the tertiary structure and in particular the environment of tryptophan residues [39]. Upon incubation at 60 °C, one observes a red shift of photoluminescence emission, as described by the variations of the first momentum M_1 of the emission spectrum: $M_1(t)/M_1(0) - 1$ (fig. 4). Lysozyme has 6 tryptophans, but it is well established that the main contribution to photoluminescence comes from tryptophan 108 which is buried in the hydrophobic cleft and tryptophan 62 which is solvent exposed [40]. Therefore, the observed red shift indicates the exposure of tryptophan 108 to the solvent and the loosening protein of tertiary structure.

Figure 4 also shows the increase of the weight-averaged mass of lysozyme aggregates upon incubation at 60 °C. This was measured by static light scattering, as described in Raccosta *et al.* [25]: $R_{90}(t)/R_{90}(0) - 1$. It is important to note that the conformational changes revealed by the changes of the spectroscopic signals continue progressively along the incubation and largely occur before the onset of fibrillation, revealed by light scattering measurements. The conformational change could be ascribed to the average changes of one conformationally unstable molecular species. However, we may credit the work by the Dobson group [29], that highlighted the existence of isoenergetic conformational substates in lysozyme solution at *p*H 2 and 60 °C. Hence, we may argue that the escape from the monomeric state is due to a kinetic balance among almost isoenergetic states: a charge-stabilized state and a fibrillation-prone conformation. The rationale for solution metastability is given by the strong electrostatic repulsion, which provides high energies of activation for aggregation. The system of non-native proteins is driven out of metastability through specific conformational substates, which are kinetically populated and experience lower activation energy for fibril formation, reasonably due to a more conspicuous amount of β-sheet structures (as revealed by CD kinetics experiments).

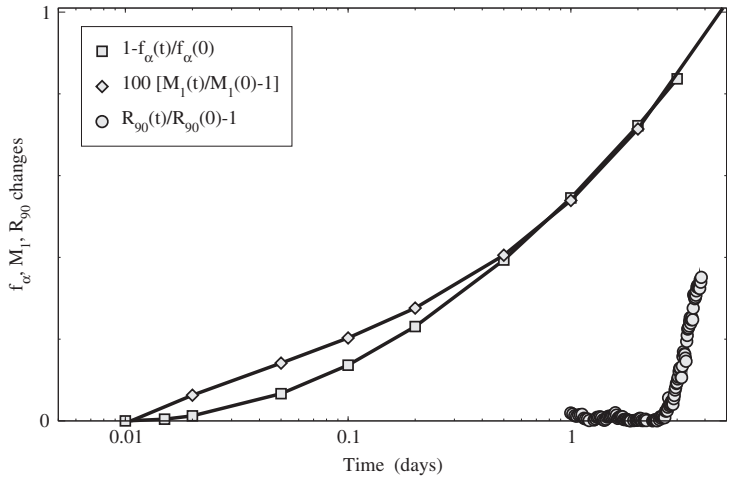

Fig. 4. – Variation of kinetic parameters of lysozyme solutions upon incubation at 60 °C: $f(\alpha)$, percent fraction of alpha helix structure ($c_2 = 0.21\,\mathrm{g\,L^{-1}}$); M_1, first momentum of tryptophan emission band ($c_2 = 0.18\,\mathrm{g\,L^{-1}}$); R_{90}/K, normalised Rayleigh ration ($c_2 = 18.4\,\mathrm{g\,L^{-1}}$).

The fibril-prone substates, which are in equilibrium with the other partially denatured conformational substates, are progressively sequestered by the fibrillation process. The propensity for fibril formation is reduced by increasing the ionic strength, which implies a reduction of Debye screening length and hence a reduction of the activation energy for amorphous aggregation.

5. – Conclusions

We have studied the intermolecular interactions of lysozyme and α-chymotrypsinogen which are capable of forming amyloid fibrils at acidic pH, when they are largely charged. The strong electrostatic repulsion at low ionic strength determines the stability of monomeric solutions at low temperatures as well as their meta-stability at high temperatures. Protein concentration fluctuations are measured by light scattering and analyzed in terms of the classical second virial coefficient approach and a new model-free approach based on Kirkwood-Buff integrals. The latter method allows to point out a concentration-dependent behaviour of concentration fluctuations that are damped by increasing protein concentration and by reducing salt concentration. Such effective crowding is observed even at low concentration due to electrostatic repulsion. Finally, we explained how the electrostatic repulsion causes a kinetic instability and is able to act as a gatekeeper in selecting the appropriate pathway of aggregation. Our results highlight the importance of considering both protein conformation and solution thermodynamics when addressing the important issue of protein stability against aggregation or fibrillation, and the importance of leading the physics of protein stability beyond the classical framework of colloidal theory.

* * *

We acknowledge the valuable collaboration and discussions with R. Noto, L. Randazzo, A. Emanuele, and T. Perevozchikova.

REFERENCES

[1] Roberts C. J., *Biotechnol. Bioeng.*, **98** (2007) 927.
[2] Chiti F. and Dobson C. M., *Annu. Rev. Biochem.*, **75** (2006) 333.
[3] Weiss IV W. F., Young T. M. and Roberts C. J., *J. Pharm. Sci.*, **98** (2009) 1246.
[4] Rosenbaum D., Zamora P. and Zukoski C., *Phys. Rev. Lett.*, **76** (1996) 150.
[5] Foffi G., McCullagh G. D., Lawlor A., Zaccarelli E., Dawson K. A., Sciortino F., Tartaglia P., Pini D. and Stell G., *Phys. Rev. E*, **65** (2002) 1.
[6] Tessier P. M., Johnson H. R., Pazhianur R., Berger B. W., Prentice J. L., Bahnson B. J., Sandler S. I. and Lenhoff A. M., *Proteins: Struct. Funct. Gen.*, **50** (2003) 303.
[7] Manno M., Xiao C., Bulone D., Martorana V. and San Biagio P. L., *Phys. Rev. E*, **68** (2003) 011904.
[8] George A. and Wilson W. W., *Acta Crystallogr. D*, **50** (1994) 361.
[9] Trovato A., Chiti F., Maritan A. and Seno F., *PLoS Comput. Biol.*, **2** (2006) 1608.
[10] Tartaglia G. G., Pawar A. P., Campioni S., Dobson C. M., Chiti F. and Vendruscolo M., *J. Mol. Biol.*, **380** (2008) 425.
[11] Bendedouch D. and Chen S.-H., *J. Phys. Chem.*, **87** (1983) 1473.
[12] Liu W., Cellmer T., Keerl D., Prausnitz J. M. and Blanch H. W., *Biotechnol. Bioeng.*, **90** (2005) 482.
[13] Grudzielanek S., Jansen R. and Winter R., *J. Mol. Biol.*, **351** (2005) 879.
[14] Javid N., Vogtt K., Krywka C., Tolan M. and Winter R., *Phys. Rev. Lett.*, **99** (2007) 2.
[15] Carrotta R., Manno M., Giordano F. M., Longo A., Portale G., Martorana V. and Biagio P. L. S., *Phys. Chem. Chem. Phys.*, **11** (2009) 4007.
[16] Routledge K. E., Tartaglia G. G., Platt G. W., Vendruscolo M. and Radford S. E., *J. Mol. Biol.*, **389** (2009) 776.
[17] Mezzenga R., Jung J. M. and Adamcik J., *Langmuir*, **26** (2010) 10401.
[18] Hill S. E., Miti T., Richmond T. and Muschol M., *PLoS ONE*, **6** (2011) e18171.
[19] Prausnitz J., *Biophys. J.*, **108** (2015) 453.
[20] Sarangapani P. S., Hudson S. D., Jones R. L., Douglas J. F. and Pathak J. A., *Biophys. J.*, **108** (2015) 724.
[21] Velev O. D., Kaler E. W. and Lenhoff a. M., *Biophys. J.*, **75** (1998) 2682.
[22] Carrotta R., Manno M., Bulone D., Martorana V. and San Biagio P. L., *J. Biol. Chem.*, **280** (2005) 30001.
[23] Manno M., Craparo E. F., Podesta A., Bulone D., Carrotta R., Martorana V., Tiana G., San Biagio P. L., Mauro M. and Podestà A., *J. Mol. Biol.*, **366** (2007) 258.
[24] Andersen C. B., Yagi H., Manno M., Martorana V., Ban T., Christiansen G., Otzen D. E., Goto Y. and Rischel C., *Biophys. J.*, **96** (2009) 1529.
[25] Raccosta S., Martorana V. and Manno M., *J. Phys. Chem. B*, **116** (2012) 12078.
[26] Blanco M. A., Perevozchikova T., Martorana V., Manno M. and Roberts C. J., *J. Phys. Chem. B*, **118** (2014) 5817.
[27] Arnaudov L. N. and de Vries R., *Biophys. J.*, **88** (2005) 515.
[28] Hill S. E., Robinson J., Matthews G. and Muschol M., *Biophys. J.*, **96** (2009) 3781, http://dx.doi.org/10.1016/j.bpj.2009.01.044.
[29] Dhulesia A., Cremades N., Kumita J. R., Hsu S. T. D., Mossuto M. F., Dumoulin M., Nietlispach D., Akke M., Salvatella X. and Dobson C. M., *J. Am. Chem. Soc.*, **132** (2010) 15580.
[30] Schmitz K. S., *An Introduction to Dynamic Light Scattering by Macromolecules* (Academic Press, Inc., San Diego) 1990.
[31] Hansen J. P. and McDonald I. R., *Theory of Simple Liquids* (Academic Press, London, New York, San Diego) 1986.
[32] Manning G. S., *J. Phys. Chem. B*, **111** (2007) 8554.
[33] Blanco M. A., Sahin E., Li Y. and Roberts C. J., *J. Chem. Phys.*, **134** (2011) 1.

[34] Kirkwood J. G. and Buff F. P., *J. Chem. Phys.*, **19** (1951) 774.
[35] Andrews J. M., Weiss IV W. F. and Roberts C. J., *Biochemistry*, **47** (2008) 2397.
[36] Weiss IV W. F., Hodgdon T. K., Kaler E. W., Lenhoff A. M. and Roberts C. J., *Biophys. J.*, **93** (2007) 4392.
[37] Weiss IV W. F., Zhang A., Ivanova M. I., Sahin E., Jordan J. L., Fernandez E. J. and Roberts C. J., *Biophys. Chem.*, **185** (2014) 79, http://dx.doi.org/10.1016/j.bpc.2013.11.005.
[38] Provencher S. W., *Comput. Phys. Commun.*, **27** (1982) 229.
[39] Lakowicz J. R., *Principles of Fluorescence Spectroscopy* (Springer) 2006.
[40] D'Amico M., Raccosta S., Cannas M., Martorana V. and Manno M., *J. Phys. Chem. B*, **115** (2011) 4078.