

# Observation of the Early Structural Changes Leading to the Formation of Protein Superstructures

*Vito Foderà,<sup>†,‡,\*</sup> Valeria Vetri,<sup>§</sup> Thea S. Wind,<sup>†</sup> Wim Noppe,<sup>||</sup> Claus Cornett,<sup>⊥</sup> Athene M. Donald,<sup>‡</sup> Ludmilla A. Morozova-Roche<sup>#</sup> and Bente Vestergaard<sup>†</sup>*

<sup>†</sup>Department of Drug Design and Pharmacology, University of Copenhagen, Denmark.

<sup>‡</sup>Sector of Biological and Soft Systems, Department of Physics, Cavendish Laboratory, University of Cambridge, United Kingdom.

<sup>§</sup>Dipartimento di Fisica e Chimica, Università di Palermo, Italy.

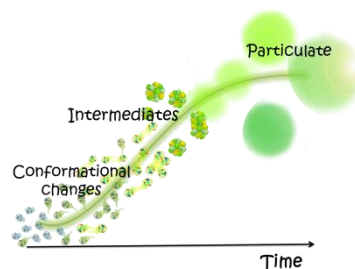
<sup>||</sup>IRF Life Sciences KuLeuven KULAK, Kortrijk, Belgium.

<sup>⊥</sup>Department of Pharmacy, University of Copenhagen, Denmark.

<sup>#</sup>Institute of Medical Biochemistry and Biophysics, Umeå University, Sweden

ABSTRACT. Formation of superstructures in protein aggregation processes has been indicated as a general pathway for several proteins, possibly playing a role in human pathologies. There is a severe lack of knowledge on the origin of such species in terms of both mechanisms of formation and structural features. We use Equine Lysozyme as a model protein and, combining spectroscopic techniques and microscopy with X-ray fibre diffraction and *ab initio* modelling of Small Angle X-ray Scattering data, we isolate the partially unfolded state from which one of these superstructures (i.e. particulate) originates. We reveal the low-resolution structure of the unfolded state and its mechanism of formation, highlighting the physico-chemical features and the possible pathway of formation of the particulate structure. Our findings provide a novel detailed knowledge of such a general and alternative aggregation pathway for protein, being this crucial for a basic and broader understanding of the aggregation phenomena.

## TABLE OF CONTENTS

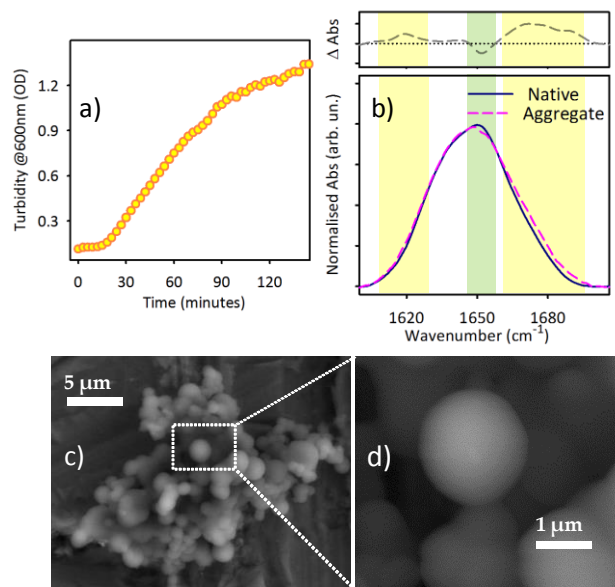


KEYWORDS: Protein Superstructures, Hydrophobicity, Conformational Changes, Unfolded State, Particulate.

Protein aggregation processes can lead to very complex structures distinct from amyloid-like fibrils, distinguished by the cross- $\beta$ -sheet signature.<sup>1</sup> Such complex species are generally denoted as superstructures (e.g. spherulites and particulates)<sup>1</sup> and widely observed both *in vitro* and in human tissues related to specific pathologies.<sup>2</sup> However, details of their nature, structure and mechanism of formation are unknown, precluding the development of strategies to inhibit or control the aggregation reaction. The formation of protein spherical particulates is a general pathway observed for a wide class of both model and medically-relevant proteins.<sup>1</sup> Here we elucidate the mechanism of formation of an unfolded state from which such a spherical structure originates. Using Equine Lysozyme (EL) as a model system and combining different techniques, we reveal the early conformational changes and single out the low-resolution structure of the partially unfolded state within the superstructure. Moreover, we demonstrate the hydrophobic nature and near-perfect spherical arrangement of such aggregates. All this suggests a novel non-conventional packing different from the classical fibril structure.

The aggregation propensity of EL has preliminarily been studied in the pH-range 4.5-11 at 57°C. At pH-values close to the isoelectric point ( $\sim$ 8.5) and higher, macroscopic floccules are readily formed. Figure 1a shows the turbidity signal detected at 600 nm during incubation at 57°C and pH 8.5. The kinetics profile presents a sigmoidal shape for the aggregate growth, with a lag time of  $\sim$ 13 minutes. Figure 1b shows the normalised FTIR spectra within the amide I' region for both the native and aggregate sample. Data indicate that upon aggregation the protein secondary structure is almost conserved. However, in the aggregated sample the presence of small shoulders at 1620  $\text{cm}^{-1}$ , 1670 $\text{cm}^{-1}$  and 1690  $\text{cm}^{-1}$  indicates a fraction of intermolecular  $\beta$ -sheets and  $\beta$ -turns.<sup>3</sup> Moreover, the growth of the signal above 1660  $\text{cm}^{-1}$  is compatible with the presence of not exchanged irregular helical structures.<sup>3</sup> The morphology is examined by using

**FIGURE 1**

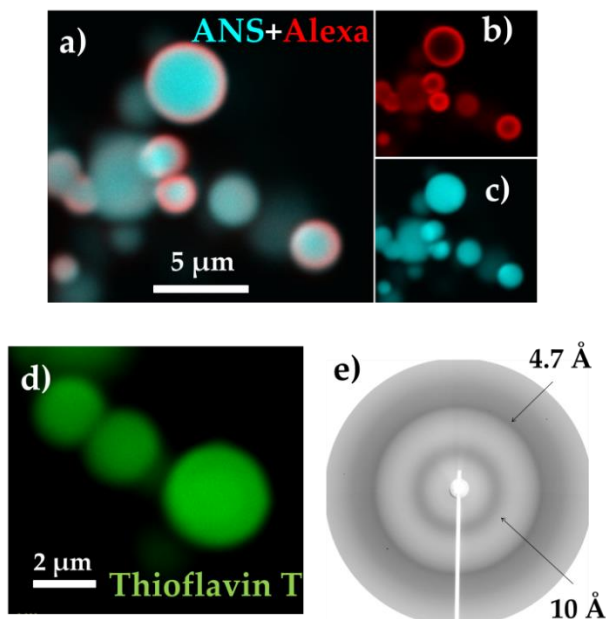


**Figure 1.** a) Turbidity detected at 600 nm during the aggregation of 5 mg/ml EL, in 20 mM sodium acetate, pH 8.5, 57°C. b) FTIR normalised spectra of the native (blue line) and aggregate (pink dashed line) EL state. Top panel: residual signal (grey dashed line) after subtraction of the native state signal ( $\Delta$ Abs). Decrease (green) and increase (yellow) of the signal. c) Morphology of the aggregate detected by ESEM. d) Magnified portion of c).

Environmental Scanning Electron Microscopy (ESEM, see Supporting Information, SI). Panels 1c-d show large clusters of spherical particles with varying diameters ( $\mu$ m range). Such particles are referred to as particulates.<sup>4</sup>

To probe the physico-chemical features of such aggregates further, we used confocal and two-photon excitation fluorescence microscopy (see SI for details). Experiments were performed in two colors and we used the hydrophilic dye Alexa 647 (red) to heuristically probe the

**FIGURE 2**



**Figure 2.** Two-photon excitation fluorescence microscopy and X-ray diffraction data of EL particulates. a) ANS (cyan) and Alexa (red) staining in two-color configuration b) Alexa and c) ANS, both in single color, showing that the interior of the aggregates is devoid of water. d) Thioflavin T staining. e) X-ray diffraction pattern of particulates.

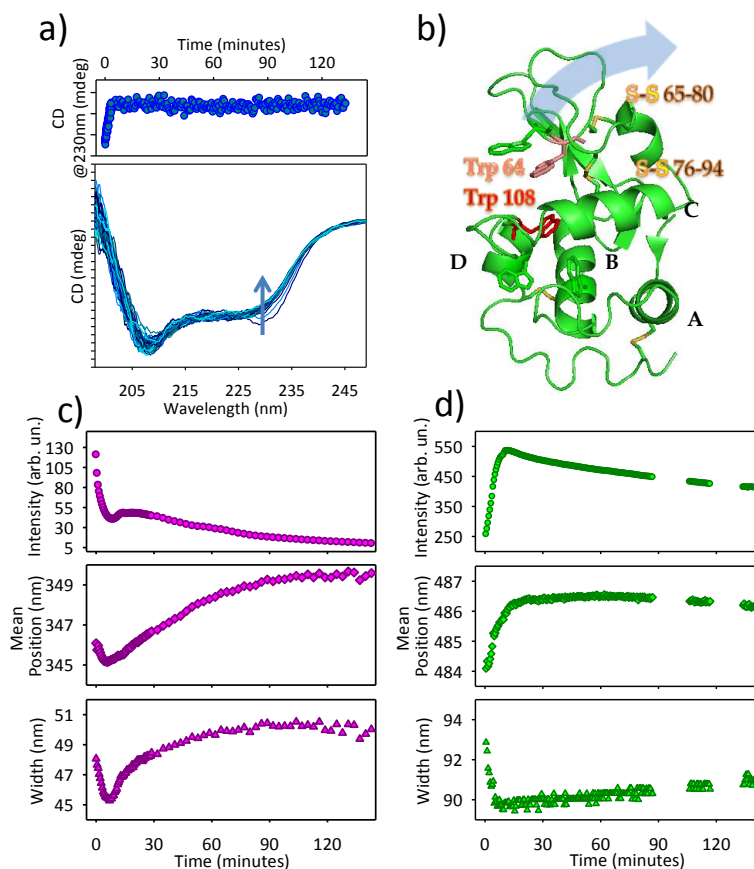
accessibility of the solvent within the aggregates, and 8-Anilino-1-naphthalene-sulfonate (ANS, cyan) to probe hydrophobic regions. Figure 2a-c shows that the ANS dye is uniformly distributed within the particulate, while Alexa does not have access to the internal part of the aggregate even during experiments including longer observation times. This is a fingerprint of physico-chemical properties of the superstructures. In fact, under the hypothesis that no preferential interaction exists between Alexa and the observed aggregates, it is possible to infer

that solvent diffusion is the driving force for Alexa uptake. As a consequence, data in Figure 2 a-c suggest a highly dense structural packing, possibly excluding water from the internal regions. Particulates show a positive Thioflavin T (ThT) staining (Fig. 2d) and their X-ray fibre diffraction (XFD) pattern presents two signals at 4.7 Å and 10 Å (Fig. 2e), i.e. the typical distances observed for amyloid fibrils. However, while in the case of the elongated fibrils the reflections show a very preferential alignment, in the case of the EL aggregates in Figure 2e, isotropic rings are obtained. As for  $\beta$ -lactoglobulin particulates,<sup>4</sup> this indicates that no preferential alignment is present in the sample.

In order to investigate the mechanism of formation of such a superstructure, we further analyzed the early structural changes, indicated by our initial FTIR analysis. Figure 3a shows the temporal evolution of the far-UV Circular Dichroism (CD) spectrum during the kinetics. Overall, the CD spectra do not change dramatically, and only a disappearance with time of the shoulder at 230 nm is observed, corresponding to the lag time of Figure 1a (Fig. 3a, top). This suggests that the secondary structure is mostly conserved upon heating at 57°C, further corroborating the FTIR analysis and in line with previous studies on EL structure persistence.<sup>5</sup> Specifically in the early stages of the aggregation process studied here, EL structural rearrangements lead to an only partially unfolded state, which is similar to the molten globule observed for this protein.<sup>5</sup> This is confirmed by the changes observed in the NMR spectra during the kinetics (see Figure S1 in the SI) that, according to previously reported peaks' assignments and data, stem from a transition between native and molten globule state.<sup>6</sup>

The native crystal structure reveals that EL is formed by the  $\alpha$ -helical domain composed by four major helices (A, B, C and D in Fig. 3b) and the  $\beta$ -domain containing a three-stranded anti-parallel  $\beta$ -sheet.<sup>7</sup> During EL denaturation, helices A, B and D remain unvaried while the least

**FIGURE 3**



**Figure 3.** Structural changes observed for 5 mg/ml EL in 20 mM sodium acetate, pH 8.5 incubated at 57°C. a) CD spectra during the process. The arrow indicates the direction of the signal development at 230 nm over time. Inset: CD signal at 230 nm as a function of aggregation time. b) EL crystal structure, (PDB 2EQL<sup>7</sup>). The blue arrow indicates the unfolding of the  $\beta$ -domain upon heating. c) Trp and d) ANS emission during the aggregation process: intensity (top), wavelength (centre) of the peak position and width (bottom).

protected C-helix has been suggested to be loosened,<sup>5</sup> forming a polyproline II helix<sup>8</sup>. This leads to changes in both the orientation of the S-S bonds (65-80 and 76-94, sticks in Fig. 3b) and/or hydrophobicity of the tryptophan (Trp) environment in the proximity of the C-helix. In

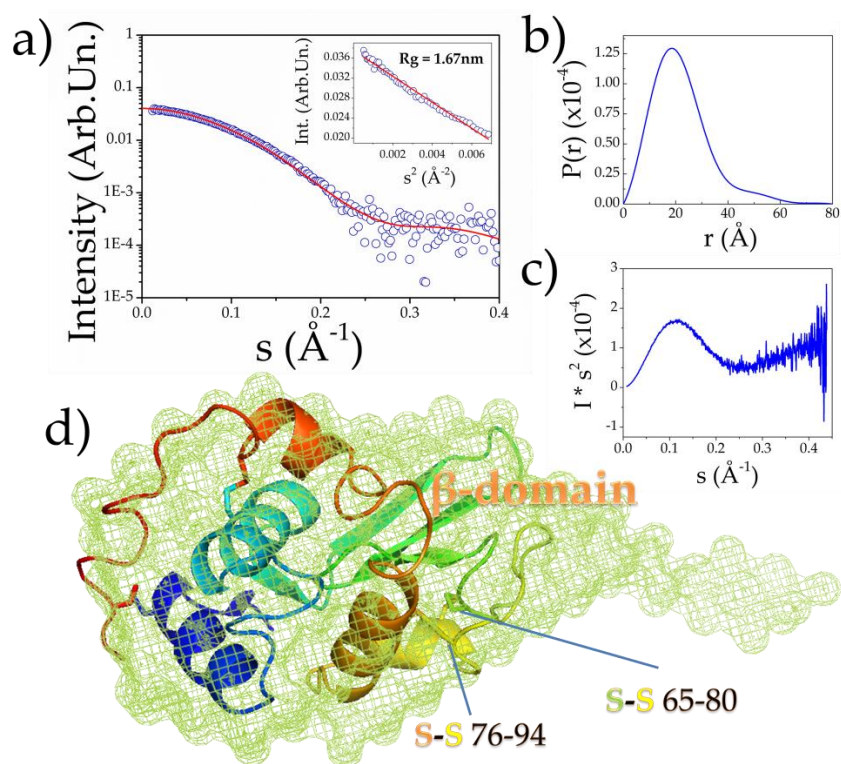
particular, Trp108 (red in Fig. 3b) has been suggested to give rise to a CD signal in the 230 nm region<sup>5</sup> which is the signal observed to change in this study. To corroborate these observations, in Figure 3c we show the momenta analysis of the sample intrinsic fluorescence (see details in SI) during the aggregation process. EL contains five Trps.<sup>6</sup> However, in the light of the above considerations and data, during thermal denaturation, changes in the intrinsic fluorescence signal would reflect changes in mobility/polarity of Trp close to the less protected C-helix, i.e. Trp 64 and 108 (red and salmon in Fig. 3b). Interestingly, in hen egg lysozyme (HEWL), these two Trps are the main emitters.<sup>9</sup> During the lag phase a reduction of fluorescence intensity and a blue shift of Trp emission band are observed (Fig. 3c). This could be due to conformational changes with the loss of tertiary contacts and variations in the distance between Trp108 and Trp64, also affecting the overall stability and/or position of the  $\beta$ -domain adjacent to the C-helix. Importantly, a non-radiative energy transfer between correspondent Trps in HEWL has been proposed.<sup>9, 10</sup> Similarly, an increase in the Trp64-Trp108 distance in EL may determine a reduction of transfer efficiency. This would lead to a reduction of Trp64 signal and an apparent blue shift of the overall intrinsic emission, as it is exactly observed. This explanation also foresees a decrease of the emission bandwidth, as indeed observed (Fig. 3c). After the lag-phase a slight increase and then a decrease in the intensity are detected together with a red shift and a broadening of the band. According to previous studies,<sup>11</sup> this can be interpreted as further rearrangement of interacting molecules leading to the exposure of internal residues. Notably, the temporal changes in intensity of the intrinsic fluorescence are similar to the kinetics of both turbidity and CD signal at 230 nm. A two-phase kinetics with the same temporal profile is also observed for ANS fluorescence emission (Fig. 3d). ANS binding to proteins is due to hydrophobic interactions, normally resulting in an increase of fluorescence emission and a blue



shift for the band peak.<sup>12</sup> Here, in contrast, we observe a modest, but notable red shift in concert with the fluorescence increase during the lag phase. Simultaneously, the bandwidth is reduced, indicating that the accessible hydrophobic binding sites to the dye become progressively more homogeneous. The observed trend for EL is thus non-conventional. The intensity increase may be due to the enhanced accessibility of new hydrophobic regions of the protein as a consequence of the unfolding process and/or to the formation of new hydrophobic clusters. Regarding the peak position, besides the obvious temperature effect during heating, the red shift can be ascribed to the formation/accessibility of different ANS binding sites with progressively less favorable protein-dye interaction.<sup>11</sup> In the second phase of the process, aggregate reorganisation is observed and leads to a partial exposure of the dye.

Data in Figure 1, 2 and 3 together with previous structural studies on EL<sup>5, 6</sup> suggest the following scenario: during the lag time native EL rapidly converts into a more “open” molten globule conformation in which a large part of the  $\alpha$ -domain retains a native-like structure, while the C-helix adapts a more flexible conformation causing the destabilization of the neighboring  $\beta$ -domain. All this brings us to the following question: can we single out the structure of such an intermediate species, validate it against the spectroscopic data and suggest a possible pathway for particulate formation? To do this, we perform SAXS measurements coupled with *ab initio* modelling.<sup>13</sup> The high viscosity of the sample and the time scale of the EL aggregation process did not allow the use of the same real time SAXS approach previously reported.<sup>14</sup> However, the EL ‘molten globule state’ is partially native, and has been reported to be stable.<sup>6</sup> We hence pursued a mechanical breaking of the particulates by soft sonication aiming at dispersing into solution their constituent elements. Imaging analysis indeed showed that, after sonication, few particulates with a decreased diameter are left (see Figure S2). After the sonication,

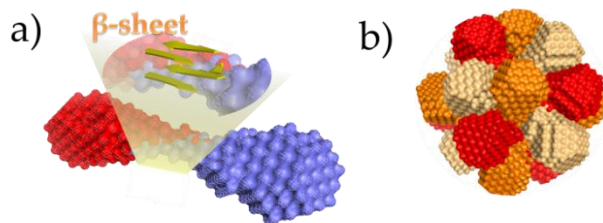
**FIGURE 4**



**Figure 4.** a) SAXS data for the aggregated EL sample after sonication and centrifugation. Red line is the calculated scattering curve of the average low-resolution *ab initio* structure (see panel d). Inset: Guinier range and  $R_g$  estimation. b)  $P(r)$  function and c) Kratky plot from data in panel a). d) *ab initio* model with the native EL crystal structure (PDB-2EQL<sup>7</sup>) superimposed.

we centrifuged the sample to remove the final particulates, and the SAXS experiment was performed (see SI for details). Importantly, the final sample does not present high molecular weight species that can generate artefacts in the SAXS measurement (see Figure S3 and S4 and text in SI). Figure 4a shows the radially averaged SAXS data. Data analysis has been performed using PRIMUS<sup>15</sup> and from the Guinier approximation, an  $R_g = 1.67 \pm 0.06 \text{ nm}$  is estimated (inset). This value is higher than the one calculated for the native EL monomer ( $R_g = 1.42 \text{ nm}$ ).

## SCHEME 1



**Scheme 1.** Model for particulate formation. a)  $\beta$ -sheet interaction between unfolded EL molecules and b) reorganization of 12 unfolded molecules within spherical oligomeric species (nm range) from which particulates may originate.

The  $P(r)$  function reveals a  $D_{MAX}$  of  $\sim 8$  nm and is in accordance with a mainly globular shape, with a smaller proportion in a slightly elongated conformation. The Kratky plot is indicative of a mainly folded particle (see  $s < 0.2 \text{ \AA}^{-1}$  range) albeit with an increase at higher scattering angles, suggesting some flexibility (Fig. 4b-c). After this preliminary analysis, *ab initio* modelling was performed. Using the program DAMMIF,<sup>16</sup> an average of 20 different models (see Fig. S5 in SI) was obtained and is presented as a mesh surface in Figure 4d together with the native EL monomer crystal structure. The low-resolution structure confirms the presence of a globular part and an extended part. This structure can be linked to the spectroscopic data. From Figure 3, we have argued that loosening of the C-helix<sup>8</sup> and possibly of the S-S bridges protecting the  $\beta$ -domain (labels in Fig. 4d) takes place, which would allow a part of the molecule to be more exposed (Fig. 3). This information suggests that the extended part visualized in our *ab initio* model likely reflects a rearrangement of the  $\beta$ -domain, protruding away from the native-like hydrophobic core, which is maintained. This is in line with the highly conserved CD and FTIR signals (Fig. 1 and 3).

Based on our data we can postulate a pathway for particulate formation. Figure 2a shows that the aggregates are highly hydrophobic and their internal part is barely accessible to the solvent. Moreover, they are perfectly spherical with a variable final diameter. This can only be formed if the particulate building block has no preferential orientation and interactional directionality, hence reasonably suggesting the building block being also near-spherical. We suggest that the exposed  $\beta$ -strands in the repositioned  $\beta$ -domain interact (Scheme 1a) and further rearrange. This allows for the formation of smaller spherical aggregates with a spherically symmetric  $\beta$ -sheet core formed from the  $\beta$ -strands of individual molecules (Scheme 1b). From this specie, particulates may originate. This is in line with  $\beta$ -structures formed by the interaction of a number of protomers proposed for other proteins, e.g.  $A\beta(1-42)$ <sup>17</sup> and would explain the isotropy of the XFD signal (Fig. 2). If the EL  $\beta$ -strands constitute the core of the spherical intermediate oligomers, the exterior would be constituted by the more stable  $\alpha$ -helical domains. The latter will allow the interaction with other oligomers and the formation of the non-polar interior of particulates (Fig. 2). In agreement with this idea, a number of hydrophobic residues are present in the exterior of the  $\alpha$ -domains.<sup>5</sup> Specifically, Val 112, 2 and 99 (D-helix/loop interface) together with Leu 120, 124 and 129 (A-helix) are candidate for the formation of such hydrophobic interactions.

In summary we singled out the partially unfolded state from which protein particulate originates, providing unprecedented structural insights into the aggregation process leading to superstructures. Furthermore we postulate an explanation for the packing of such intermediate structures into particulates. The detailed knowledge on the unfolding process that can be obtained by our novel approach offers the opportunity to design targeted strategies to control the

aggregation reaction already in the early stages. This is of key significance, especially in the light of recent reports indicating that superstructures can have a role in human pathologies.<sup>2</sup>

## EXPERIMENTAL METHODS

Equine Lysozyme (EL) was isolated from equine milk. All the measurements were performed in aqueous solution with 20 mM sodium acetate. pH was adjusted with HCl and NaOH to the desired pH value. Please refer to the SI for full details on sample preparation, experimental techniques and data analysis.

## ASSOCIATED CONTENT

**Supporting Information.** Supporting data and details on experiments and data analysis are available in the supplementary information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\* Vito Foderà: [vito.fodera@sund.ku.dk](mailto:vito.fodera@sund.ku.dk).

## ACKNOWLEDGMENT

VF acknowledges support from the FP7 Marie-Curie Actions Intra European Fellowship (IEF) 2012-2014, Project Nr. 299385 “FibCat”, at the University of Copenhagen, Denmark. LAM-R thanks VR-M and Insamlingstiftelsen. TSW and BV acknowledge funding from ‘Safir’-Sapere Aude DFF-Starting Grant, Danish Council for Independent Research-Medical Sciences. TSW thanks the IDA Travel and Pureskov legatet for funding. Authors thank A.E. Langkilde and M. Groenning (University of Copenhagen) for useful discussions. The Lundbeck Foundation (grant R77-A6742) and the Danish Research Council for Independent Research (grant 10-085264) are acknowledged for the NMR equipment.

## REFERENCES

- (1) Krebs, M.R.; Domike, K.R.; Donald, A.M. Protein Aggregation: More Than Just Fibrils. *Biochem. Soc. Trans.* **2009**, *37*, 682-686.
- (2) Exley, C.; House, E.; Collingwood, J.F.; Davidson, M.R.; Cannon, D.; Donald, A.M. Spherulites of Amyloid-Beta42 In Vitro and in Alzheimer's Disease. *J. Alzheimer's Dis.* **2010**, *20*, 1159-1165
- (3) Chalmers, J.M.; Griffiths P. *Handbook of Vibrational Spectroscopy* Vol. 5, WILEY-VCH, Weinheim, 2011.
- (4) Krebs, M.R.; Devlin, G.L.; Donald, A.M. Amyloid Fibril-Like Structure Underlies the Aggregate Structure Across the pH Range for  $\beta$ -Lactoglobulin. *Biophys. J.* **2009**, *96*, 5013-5019.
- (5) Morozova-Roche, L.A.; Jones, J.A.; Noppe, W.; Dobson C.M. Independent Nucleation and Heterogeneous Assembly of Structure During Folding of Equine Lysozyme. *J. Mol. Biol.* **1999**, *289*, 1055-1073.
- (6) Morozova-Roche, L.A.; Arico-Muendel, C.C. ; Haynie, D.T.; Emelyanenko, V.I.; Van Dael, H.; Dobson C.M. Structural Characterisation and Comparison of the Native and A-states of Equine Lysozyme. *J. Mol. Biol.* **1997**, *268*, 903-921.
- (7) Tsuge, H.; Ago, H.; Noma, M.; Nitta, K.; Sugai, S.; Miyano M. Crystallographic Studies of a Calcium Binding Lysozyme From Equine Milk at 2.5 Å resolution. *J. Biochem.* **1992**, *111*, 141-143.

- (8) Blanch, E.W.; Morozova-Roche, L.A.; Cochran, D.A.E.; Doig, A.J.; Hecht, L.; Barron, L.D. Is Polyproline II Helix the Killer Conformation? A Raman Optical Activity Study of the Amyloidogenic Prefibrillar Intermediate of Human Lysozyme. *J. Mol. Biol.* **2000**, *301*, 553-563.
- (9) Imoto, T.; Forster, L.S.; Rupley, J.A.; Tanaka, F. Fluorescence of Lysozyme: Emissions from Tryptophan Residues 62 and 108 and Energy Migration. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 1151-1155.
- (10) Formoso, C.; Forster, L.S. Tryptophan Fluorescence Lifetimes in Lysozyme. *J. Biol. Chem.* **1975**, *250*, 3738-3745.
- (11) Vetri, V.; Leone, M.; Morozova-Roche, L.A.; Vestergaard, B.; Foderà, V. Unlocked Concanavalin A Forms Amyloid-like Fibrils from Coagulation of Long-Lived "Crinkled" Intermediates. *PLoS One*, **2013**, *8*, e68912
- (12) Stryer, L. The Interaction of a Naphthalene Dye with Apomyoglobin and Apohemoglobin. A Fluorescent Probe of Non-Polar Binding Sites. *J. Mol. Biol.* **1965**, *13*, 482-495.
- (13) Mertens, H.D.T.; Svergun D.I. Structural Characterization of Proteins and Complexes Using Small-Angle X-ray Solution Scattering. *J. Struct. Biol.* **2010**, *172*, 128-141.
- (14) Vestergaard, B.; Groenning, M.; Roessle, M.; Kastrup, J.S.; van de Weert, M.; Flink, J.M.; Frokjaer, S.; Gajhede, M.; Svergun, D.I. A Helical Structural Nucleus Is the Primary Elongating Unit of Insulin Amyloid Fibrils. *PLoS Biol.*, **2007**, *5*, e134.



(15) Konarev, P.V. ; Volkov, V.V.; Sokolova, A.V.; Koch M.H.J.; Svergun, D.I. PRIMUS: a Windows PC-Based System for Small-Angle Scattering Data Analysis. *J. Appl. Cryst.* **2003**, *36*, 1277-1282.

(16) Franke, D.; Svergun, D.I. DAMMIF, a Program for Rapid ab-initio Shape Determination in Small-Angle Scattering. *J. Appl. Cryst.* **2009**, *42*, 342-346.

(17) Stroud, J.C.; Liu, C.; Teng, P.K.; Eisenberg D. Toxic Fibrillar Oligomers of Amyloid- $\beta$  Have Cross- $\beta$  Structure. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7717-7722.