doi:10.1016/j.jmb.2006.11.008 J. Mol. Biol. (2007) 366, 258–274

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Kinetics of Different Processes in Human Insulin Amyloid Formation

Manno Mauro[1]*, Emanuela Fabiola Craparo[2], Alessandro Podestà[3] Donatella Bulone[1], Rita Carrotta[1], Vincenzo Martorana[1] Guido Tiana[4] and Pier Luigi San Biagio[1]

[1]*Institute of Biophysics at Palermo Italian National Research Council, via U. La Malfa 153, I-90146 Palermo Italy*

[2]*Dipartimento di Chimica e Tecnologie Farmaceutiche Università di Palermo via Archirafi 32 I-90123 Palermo, Italy*

[3]*Department of Physics and CIMAINA, University of Milano, via Celoria 16 I-20133 Milano, Italy*

[4]*Department of Physics University of Milano and INFN via Celoria 16, I-20133 Milano Italy*

**Corresponding author*

Human insulin has long been known to form amyloid fibrils under given conditions. The molecular basis of insulin aggregation is relevant for modeling the amyloidogenesis process, which is involved in many pathologies, as well as for improving delivery systems, used for diabetes treatments. Insulin aggregation displays a wide variety of morphologies, from small oligomeric filaments to huge floccules, and therefore different specific processes are likely to be intertwined in the overall aggregation. In the present work, we studied the aggregation kinetics of human insulin at low pH and different temperatures and concentrations. The structure and the morphogenesis of aggregates on a wide range of length scales (from monomeric proteins to elongated fibrils and larger aggregates networks) have been monitored by using different experimental techniques: time-lapse atomic force microscopy (AFM), quasi-elastic light-scattering (QLS), small and large angle static light-scattering, thioflavin T fluorescence, and optical microscopy. Our experiments, along with the analysis of scattered intensity distribution, show that fibrillar aggregates grow following a thermally activated heterogeneous coagulation mechanism, which includes both tip-to-tip elongation and lateral thickening. Also, the association of fibrils into bundles and larger clusters (up to tens of microns) occurs simultaneously and is responsible for an effective lag-time.

Keywords: amyloid fibrils; insulin; aggregation; atomic force microscopy; light-scattering

Introduction

A wide range of pathologies, including systemic amyloidoses,[1] type II diabetes,[2] and neurodegenerative diseases (e.g. Alzheimer's and Hungtington diseases, Spongiform Encephalopathies),[3] are characterized by the anomalous self-assembly and deposition of proteic material into insoluble (ordered) supramolecular structures, such as fibrils, tangles, or aggregates of fibrils.[4]

Formation of amyloid fibrils is often related to different processes, such as association/dissociation of protein oligomers, nucleation, elongation of "linear" aggregates (fibrils), and coagulation of fibrils into larger aggregates (floccules, dendrimers, entangled networks).[5,6] In order to understand and control the occurrence of a given (pathological) morphology, it is crucial to distinguish the different processes involved and their role in the overall aggregation. In the present work, we focused on a well-known protein, insulin, as a model to study the kinetics of amyloid fibrillation and aggregation (see Brange *et al.*, 1997 and references therein).[7]

The heat precipitation of insulin in acidic solutions has long been known,[8] and is associated with the formation of "flakes" or floccules. Since the pioneering work of D.F. Waugh,[9] this heat precipitation was found to involve insulin fibrils,[10–12] insoluble in acid but partially soluble in alkali,[13] and able to assemble in spherites or entangled networks,[11,14] as shown by the early electron microscopy images.[15] The aggregation pathway discovered by Waugh included three main steps[14,16]: (i) formation of oligomeric

Abbreviations used: AFM, atomic force microscopy; QLS, quasi-elastic light-scattering.

E-mail address of the corresponding author: mauro.manno@pa.ibf.cnr.it

0022-2836/$ - see front matter © 2006 Published by Elsevier Ltd.

active centers (nucleation); (ii) growth of fibrils; (iii) assembly of fibrils into floccules and precipitation. Interestingly, the fibril growth was a function of both protein concentration and surface area of the fibril population. This suggests a double mechanism, analogous to that formalized by Eaton and coworkers to explain Sickle Cell Hemoglobin polymerization: elongation from nuclei (homogeneous nucleation) and elongation and thickening by addition of monomers or oligomers to already formed fibrils (heterogenous nucleation).[17–19] The presence of such a secondary heterogeneous aggregation pathway in insulin amyloidosis, as implicitly hinted by Waugh, has found a recent validation in the case of bovine insulin,[20] and it is here also confirmed in the case of human insulin (see also Manno *et al.*, 2006).[21]

In the present work, the aggregation kinetics of human insulin is studied at low pH and different temperatures and concentrations by using different experimental techniques. The combined usage of small and large angle light-scattering and quasi-elastic light-scattering (QLS) allows us to monitor *in situ* the formation of structures from nanometers to millimeters. Further, in order to resolve more clearly the low molecular mass aggregates, other techniques have been used such as time-lapse atomic force microscopy (AFM) and thioflavin T fluorescence.

Thus, by addressing the overall morphogenesis on several length scales, we have clarified the general pathway of insulin aggregation from insulin oligomers to amyloid fibrils, fibril bundles, and large clusters of fibrils and bundles, until the eventual precipitation or gelation. We have found that different structures are formed simultaneously by diffusion and coalescence of (fibrillar) aggregates. The lag-time in the aggregation kinetics is due to the low experimental sensitivity, and mainly related to the formation of larger aggregates. The overall aggregation is controlled by the secondary heterogeneous pathway, and any nucleation step is limited to the very early stages. The temperature dependence of the aggregation rates suggests the existence of an activation barrier, which seems related to protein conformational changes, consistent with the observation made in insulin fibrils,[22,23] as well as in the early kinetic stages.[24,25]

Results

Kinetics of aggregate size monitored by elastic and QLS

Insulin solutions have been prepared at 5 °C at concentrations of 800, 210, and 90 μM, and quenched to temperatures of 50, 60, or 70 °C. Aggregation kinetics have been monitored by QLS and scattered intensity measurements at 90°. We have chosen the concentrations in the submillimolar range. At higher concentrations aggregation would be too fast for an extended kinetic study, while at lower concentrations the scattering signal would be extremely noisy.

From scattered intensity measurements, we determined the average molecular mass M_w. We obtained a mean mass $M_w = 10.5$, 10.0, and 8.5(±1.0) kDa, respectively for 800, 210, and 90 μM, which implies a mean aggregation number of 1.8, 1.7, and 1.5±0.2, since the molecular mass of insulin is 5806 Da. No significant differences have been observed at different temperatures.

The data obtained from static light-scattering measurements concern the weight average mass of aggregates, which is defined as $M_w = \sum_s M_s c_s / \sum_s c_s$, where M_s and c_s are, respectively, the molecular mass and the mass concentration of oligomers composed of s proteins. More information about the actual mass or size distribution of aggregates is achieved by QLS measurements. In Figure 1(a),

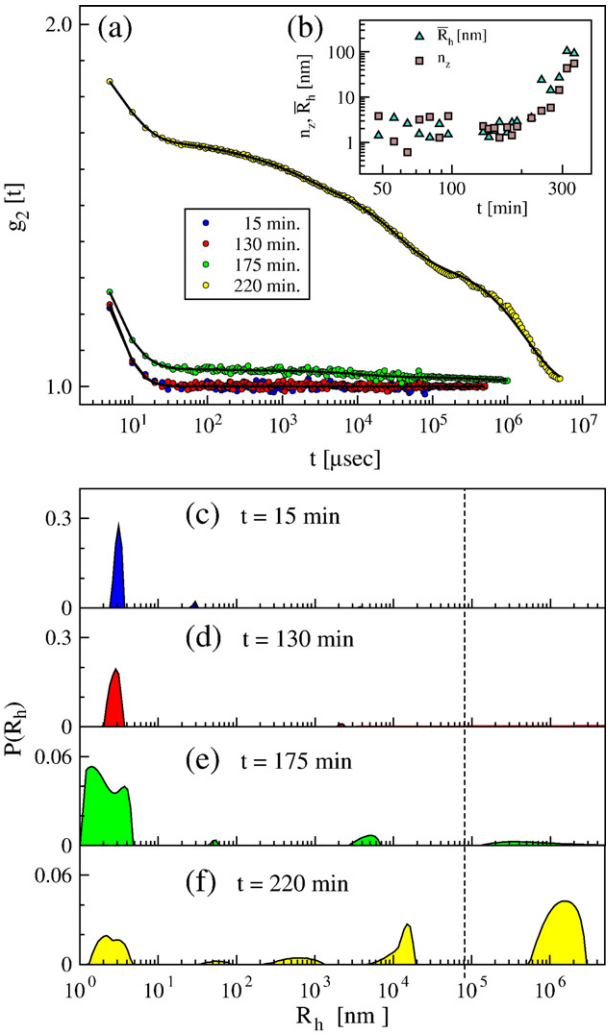

Figure 1. Aggregation kinetics of 800 μM insulin solution incubated at 60 °C: QLS experiments. (a) Intensity correlation functions at selected times as reported in the legend. (b) Kinetics of the average hydrodynamic radius $\bar{R}_h$ and of the z-averaged aggregation number n_z for the species in the range from 1 nm to 80 μm. (c)–(f) Distribution of hydrodynamic radii at selected times of kinetics. Analysis of QLS data was performed as reported in the text.

typical QLS measurements are reported for the case of $c = 800$ μM. In the early stages of aggregation, the distribution of hydrodynamic radii is limited to small oligomers, with size in the nanometer range (Figure 1(c) and (d)). After the lag-time, large-size aggregates (from hundreds of nanometers to millimeter) are observed. It is remarkable that large-size aggregates appear simultaneously at all length scale. Such changes occur within the experimental time resolution, which is about 10 min. However, the sensitivity of light-scattering measurements depends upon the square of aggregate mass. Therefore, it is conceivable that the early stages of aggregate growth are not visible at the studied concentrations. Experiments performed at other concentrations show the same qualitative behavior, apart from the duration of the lag-time.

In Figure 1(e), a broad peak is observable in the nanometer range. We may speculate whether this is due to smaller species, resulting from the degradation of insulin under the high temperatures and low pH conditions employed.[26] However, this could also be an artifact due to the fitting of correlation functions of Figure 1(a) with a regularization method. The experimental resolution is not sufficient to address this issue. Although this would be a key point in the study of the early events in insulin fibrillation, it is not crucial in the context of our work, which addresses the overall aggregation pathway without focusing on the nucleation process.

In Figure 1(b), we show the kinetics of the average hydrodynamic radius $\bar{R}_h$ and of the z-averaged aggregation number $n_z = M_w / M_1$, by taking into account aggregates up to tens of microns. The growth of both signals is approximately linearly correlated, even if the signal to noise ratio is quite low. This is a typical outcome in the case of amyloid fibrils or other elongated structures.[27–29]

Kinetics of aggregate size monitored by time-lapse AFM

In order to increase the experimental sensitivity and detect fibrillar and early stage aggregates, we used time-lapse AFM on 200 μM insulin solution incubated at 60 °C. This technique has been extensively used to observe the structure and growth of amyloid fibrils.[5,6,30] Deposition of a small aliquot of incubated solutions has been preferred to perform *in situ* AFM, in order to avoid the effects of surfaces, which are known to affect insulin aggregation.[31] Snapshots of the system from the beginning of the incubation (defined as time zero) up to 12 h are shown in Figure 2.

In agreement with QLS experiments, AFM images display oligomers during the lag-time observed in scattering experiments,[32] but not fibril-like structures (Figure 2(a) and (b)). The lag-time shown by scattering experiments on a 210 μM solution quenched at 60 °C was of 190 min. In the present

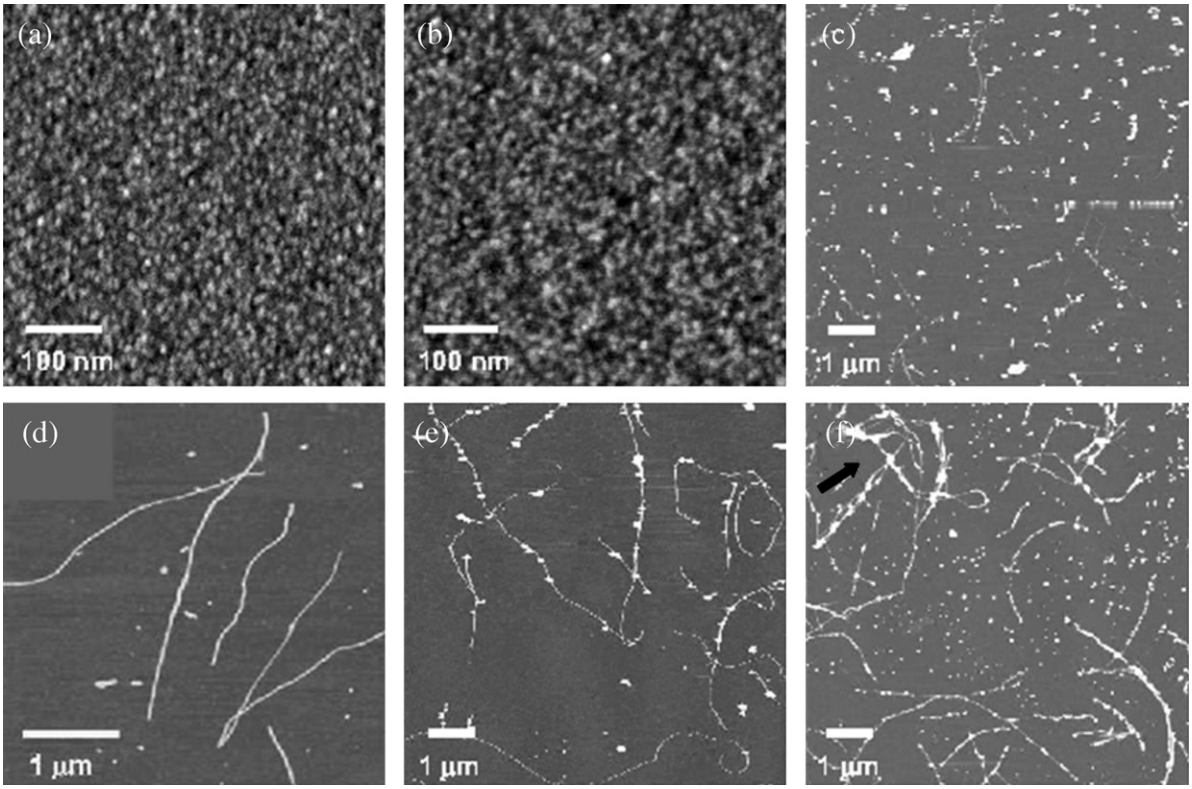

Figure 2. Snapshots of aggregation kinetics of 200 μM insulin solution at 60 °C monitored by time-lapse AFM. Times elapsed after incubation: (a) 1 min, (b) 250 min, (c) 305 min, (d) 350 min, (e) 365 min, (f) 740 min. The vertical color scale is (a) and (b) 5 nm, (c) 20 nm, (d) 30 nm, (e) 10 nm, and (f) 20 nm. (a) and (b) are redrawn from Podestà *et al.* (2006).[32] The arrow points to fibril bundles.

case, the lag-time observed by AFM experiments was between 250 and 280 min. However, it is worth noting that this value is affected by the experimental sensitivity and by the sample treatment performed in each technique. Indeed, as it will be shown and clarified in the following discussion, other techniques may experience a different lag-time. Also, note that, in order to obtain a detailed comparison of kinetic data, the same sample preparation should be used.

The early stages of insulin aggregation have been studied in a previous work.[32] AFM images during the lag-time show a distribution of ellipsoidal oligomeric aggregates (Figure 2(a) and (b)), consistent with analogous findings in other amyloidogenic systems or other proteins,[33–38] as well as in prefibrillar solutions of bovine insulin.[39]

After the lag-time, mature amyloid fibrils are observed (Figure 2(c)–(f)), in agreement with previous studies on bovine insulin by AFM,[6,39–46] or electron microscopy,[15,22,47–51] and also with electron microscopy studies on human insulin.[23,24,49,52,53] Observed fiber diameters are typically few tens of nanometers, as expected from the lateral association or growth of many thinner protofilaments.[39,50] The length of fibrils can reach several microns (Figure 3). Some fibrils are decorated with ellipsoidal clusters, but it is not clear if this represents a step in the aggregation pathway or it is just an artifact due to the sample deposition on the mica surface.

AFM experiments have been also performed on an 800 μM insulin solution incubated at 60 °C, by using the same procedure described for the lower concentration. We did not observe differences in the main structural features of amyloid fibrils at the two concentrations after several hours of incubation, as shown in Figure 3. However, concentration-dependent differences at shorter times cannot be excluded.

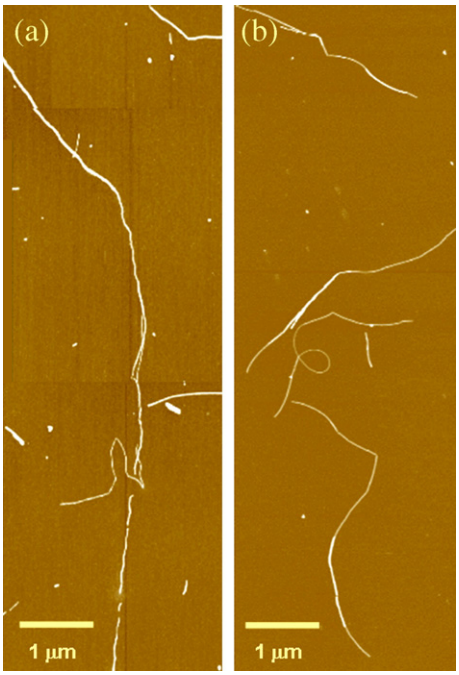

Figure 3. AFM images of insulin solution incubated at 60 °C for 9 h. Insulin concentrations are 800 μM (a), and 200 μM (b). The vertical color scale is 30 nm. (b) is redrawn from Podestà *et al.* (2006).[32]

Morphological study of amyloid fibrils by double quenching experiments

In order to gain a more detailed description of amyloid fibril morphology, we used a double quenching technique: insulin solutions at 200 μM concentration were incubated at 60° for 280 min and then diluted and incubated at 5° for more than 5 h.

This procedure has been derived from the pioneering work of D.F. Waugh,[11,14] who established that, at low temperatures and concentrations, fibril nucleation is inhibited, while elongation continues on preformed protofibrils. Therefore, a double quench would allow the growth of a few long fibrils. Such a method is also challenging to investigate the actual role of the two processes of nucleation and elongation. In the present case, however, we set the focus on single fibrils to improve the capability of morphological studies.

The AFM images performed after such double quenching show fibers with a length of several microns and a relatively small diameter (Figure 4). Even if a more extended statistics of deposited samples would be required to obtain general morphological properties, we may obtain a useful information from these images. Fibrils can form loops with a diameter of the order of half a micron and in general they have a persistence length of the order of 100 nm. The diameters are about 10 to 20 nm consistent with the six-protofilaments observed by Jimenez *et al.* (2002).[50] Also, fibrils have the structure of a right-handed helix with a pitch of the order of 10 nm (Figure 4(b) and (c)).

Kinetics of large-size aggregates monitored by multiple angle light-scattering

Kinetics of large-size clusters have been monitored by small and large angle light-scattering. Insulin solutions with a concentration of 800, 210, and 90 μM have been prepared at 5 °C and quenched to temperatures of 50, 60, or 70 °C. The combined usage of small angle and large angle light-scattering allowed us to observe the growth of structures on many length scales, from hundreds of nanometers to hundreds of microns.

As already pointed out by QLS experiments, no large aggregates are observable at the onset of the kinetics. The structure functions after the lag-time are shown for one typical experiment in Figure 5(a) in terms of the Rayleigh ratio $I_R(q)$. The absolute value of intensity grows in time, while the main structural features of aggregates do not exhibit dramatic changes. Therefore, we are mainly observing a growth in the number concentration of structured aggregates. The details regarding the

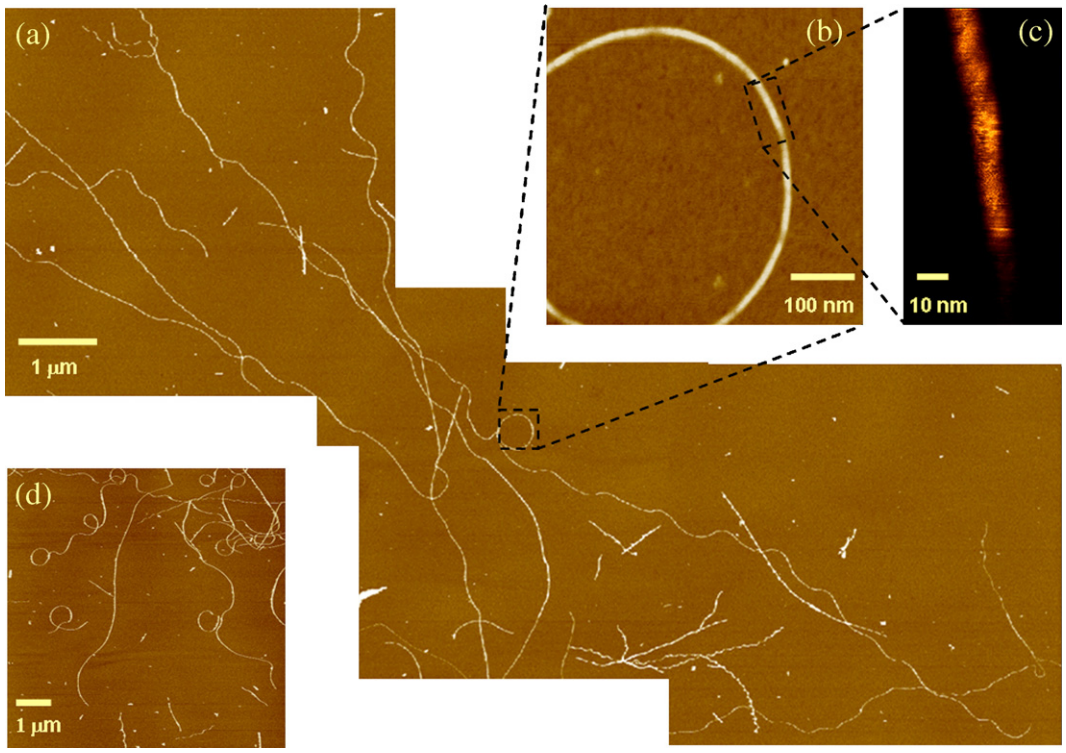

Figure 4. (a) AFM image of a 200 μM insulin solution incubated at 60 °C for 280 min and at 0 °C for 180 min. (b) Zoom of (a). (c) Zoom of (b). (d) AFM image of a 200 μM insulin solution incubated at 60 °C for 320 min and at 0 °C for 330 min. The vertical color scale is (a) 30 nm, (b) 8 nm, (c) 8 nm, and (d) 10 nm. The contrast in (c) has been enhanced to highlight the helical structure.

formation of such structures are not resolved by the present experiments due to instrumental sensitivity.

In particular, the structure functions $I_R(q)$ exhibit a power law dependence in the range from 10^{-2} to 10^{-1} μm^{-1}: $I_R(q) \approx q^{-d}$, where d is about 1.5 to 1.6. At q of the order of μm^{-1}, we observe a change in the slope of $I_R(q)$, which reaches the value of 4 up to scattering vectors of the order 0.1 μm^{-1}. The exponent d is related to the packing of molecular aggregates. A low d value indicates a loose structure, while higher values (up to 4) indicate more compact objects. These features suggest the

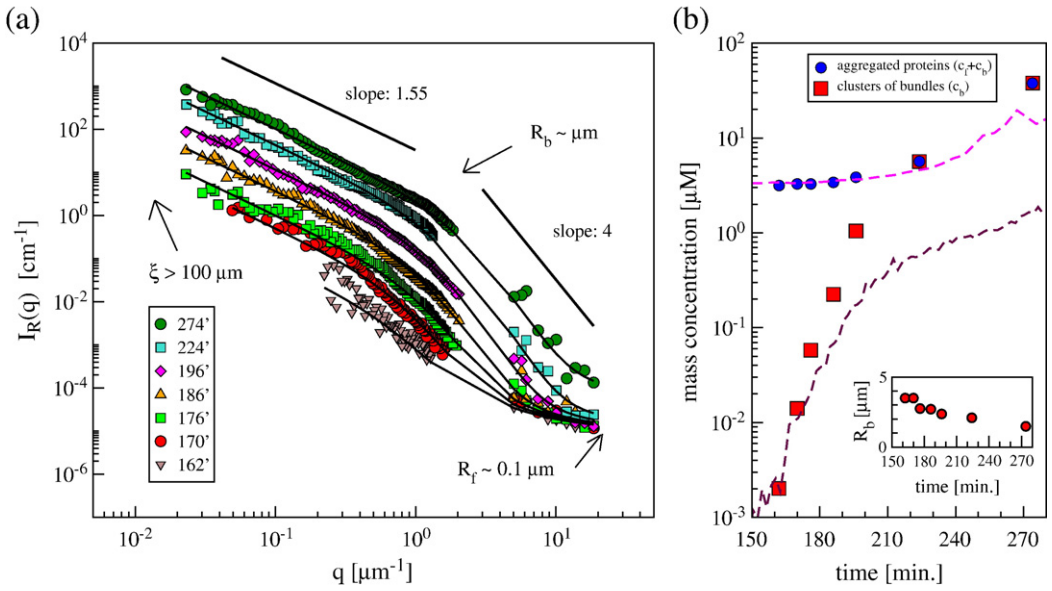

Figure 5. (a) Scattering structure function (Rayleigh ratio) of a 800 μM insulin solution incubated at 60 °C at selected times, as reported in the legend. (b) Kinetics of the mass concentration of protein aggregates (c_f+c_b) and fractal of fiber bundles (c_b). Dashed lines are the scattered light intensity, respectively (from top to bottom) at 18.7 and 0.2 μm^{-1} (in arbitrary units). Inset: Kinetics of the radius of gyration of fiber bundles.

existence of huge fractal clusters with a low fractal dimension d and a correlation length ξ of the order of hundreds of microns or higher. The word fractal refers to structures which have, on a limited length scale, no characteristic length and therefore a power law dependence of the structure function. The building blocks of such clusters can be either fiber elements with a characteristic length of about 100 nm or compact bundles of fibers with a characteristic size of microns. This scenario is confirmed by optical microscopy images recorded at the end of the aggregation, as those shown in Figure 6, and, on a smaller length scale, by the AFM image of Figure 2(f). Other images of similar fibril bundles are reported in the literature.[6,49]

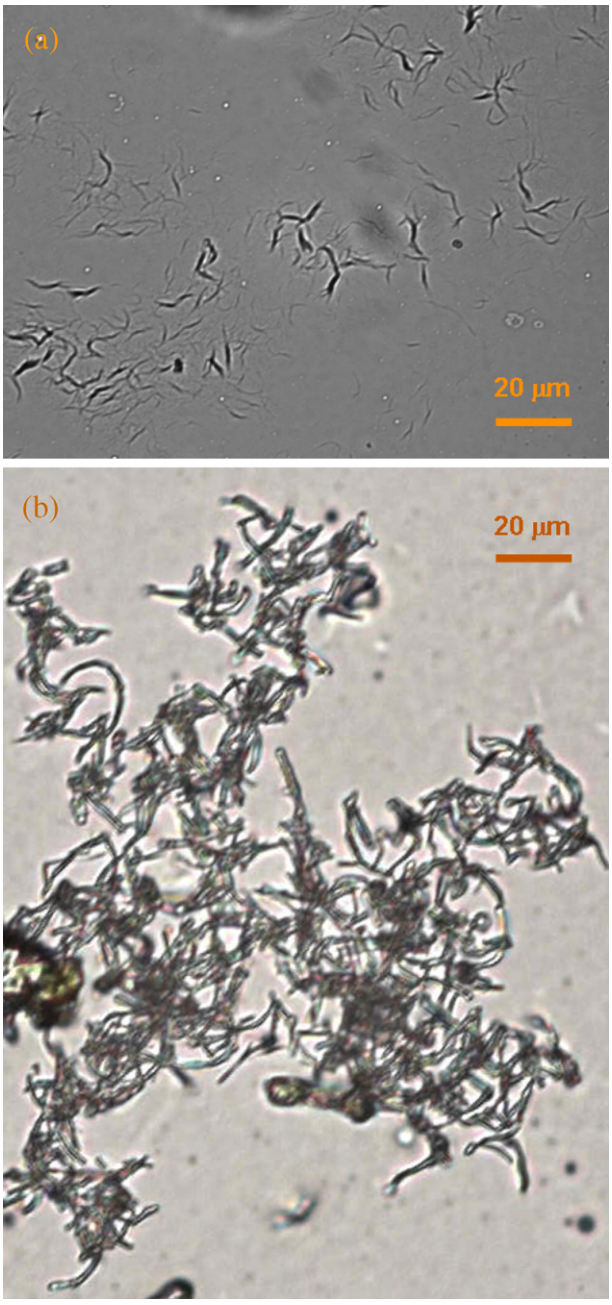

Figure 6. Phase contrast microscopy images of 210 μM insulin solution incubated at 60 °C for 9 h. Redrawn from Manno *et al.* (2006).[21]

We have built a model for the structure functions based on the observations made with AFM and optical microscopy as well as on the basic features of the structure functions themselves. We take into account protein monomers of mass M_p and concentration c_p, fiber elements of mass M_f and concentration c_f, and fiber bundles of mass M_b and concentration c_b. The data of Figure 5 have been fit to the following expression, which includes these main contributions:

$$I_R(q) = K_R[c_p M_p + c_f M_f S_f(q) P_f(q) + c_b M_b S_b(q) P_b(q)] \quad (1)$$

where K_R is the experimental factor defined in the Materials and Methods section, and S_f (P_f) and S_b (P_b) are the structure (form) factor of respectively the fiber elements and the fiber bundles.

The structure and form factor cannot be uniquely derived from the data. However, we have modeled them and simplified by making a few reasonable assumptions, based on the microscopy and scattering experiments, as well as on literature results. The details of the model and the expressions for S_f (P_f) and S_b (P_b) are reported in the Appendix A. The results fit coherently in the general framework obtained from all our experiments.

The concentration of aggregated proteins ($c_f + c_b$) starts growing from an initial value of about 3 μM (Figure 5(b)). The concentration of proteins in the clusters of bundles (c_b) grows very rapidly from a nanomolar concentration to tens of micromolar, so that it becomes comparable with the initial protein concentration, which is 800 μM in the case of Figure 5.

The radius of gyration of fibril bundles is between 1.5 and 3.5 μm, with only a slight decrease in time (inset of Figure 5(b)). This performance reveals the approximation inherent in the model. Indeed, as to equation (1), we have considered two classes of large aggregates: clusters of fibril elements and clusters of bundles of such elements. This simpleminded partition is chosen to simplify the fitting procedure. However, it can be expected that bundles of fibers keep forming by assembling of single fibril elements or by thickening of already clustered fibrils. Therefore, sizes and section diameters of fibril bundles are likely to span the entire range from hundreds of nanometers to microns. While the contribution related to clusters of bundles becomes prevalent compared to that related to fibril clusters, it may partially embody the other one.

Fibrillation kinetics monitored by analysis of scattered intensity distributions

The scattered intensity signal is related to the weight average molecular mass M_w of aggregates, as already recalled in previous sections.[54] At the onset of the aggregation, this value was obtained by the average scattered intensity, which is reported in terms of the Rayleigh ratio. However, the growth of

intensity yields a noisy signal, mainly due to the appearance of large-size or interacting aggregates.

We now turn to studying the growth of fibrillar aggregates, avoiding the noise of larger species. The information related to the actual molecular mass of such aggregates can be extracted from this noisy signal by applying a method, recently developed by our group.[21] Such approach, which is based on the analysis of scattered intensity distribution, will be briefly reviewed in Appendix B.

We have calculated the distribution of scattered intensity over a small time interval of about 5 min (Figure 7) in the same experiments performed with QLS. The distributions are mainly Gaussian in the early stages of the kinetics, but they develop a marked skewness as aggregation proceeds.

We have pointed out in a previous work[21] that the existence of a non-Gaussian (exponential) contribution in the distribution of integrated intensity is related to the presence of objects with a "slow" dynamic relaxation. These may be large-size diffusing objects (bigger than tens of microns), or interacting aggregates (such as entangled long fibers). We can partition the scattered intensity $I_{\delta T}$ into two contribution, I_F and I_S, related respectively to the Gaussian ("Fast") and to the non-Gaussian ("Slow") part of the intensity distribution. The analytical expression for the scattered intensity distribution is reported and described in Appendix B.

By fitting intensity distribution, we have obtained the "slow" fraction of intensity $x_S = \langle I_S \rangle / \langle I_{\delta T} \rangle$, which is the fraction of intensity contributed by large objects, and the "fast" Rayleigh ratio $\langle I_F \rangle$ due to the formation of smaller aggregates, which are known to be amyloid fibrils. The partition performed by these techniques does not provide a sharp separation of objects of small and large size. For example, in this approximation, the objects with a size in the intermediate range (between few microns and few hundreds of microns) can give both a fast and a slow contribution to scattered intensity. In other words, the method does not provide a cut-off length, but it allows us to filter out the contribution of larger objects with a high efficiency for aggregate sizes of hundred of microns or higher, and with a decreasing efficiency for aggregate sizes of tens of microns. An example of such approach is given in Figure 7, for one experiment ($T = 60$ °C, $c = 800$ μM). The results for the experiments $T = 60$ °C and $c = 200$ μM have been already given in Manno *et al.* (2006).[21] The kinetics of x_S and $\langle I_F \rangle$ for the three experiments at 60 °C are shown in Figure 8(a)–(c).

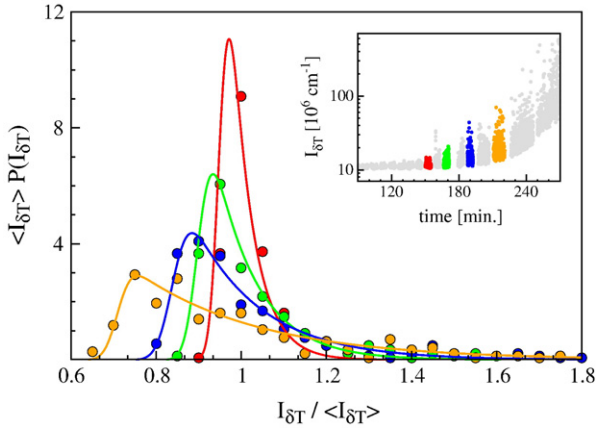

Figure 7. Scattered intensity distributions of 800 μM insulin solution incubated at 60 °C, at selected times of kinetics. Data are fit to equation (1) (solid lines). The inset shows the temporal evolution of 90° scattered intensity. Time intervals and distributions are related by matching colors.

Intensity distribution as a survey of seed-free conditions

A main problem in insulin aggregation experiments has been the reproducibility of experimental results (aggregation rates, lag-times, etc.). This is actually a severe limitation to quantitative studies on amyloid formation in many systems.[5] Experimental reproducibility is affected by the initial sample preparation. In particular, the incomplete dissolution of protein powder, the accidental mechanical stress given by inaccurate filtration, or the long-term storage of protein solution before incubation often causes the presence of protein aggregates in the solution. Such aggregates act as seeds for a templated aggregation competing with ordinary self-assembly, and change the initial monomer concentration.

The analysis of scattered intensity distribution has emerged as a powerful tool to test for the presence of seeds. Indeed, the observation of non-Gaussian noise at the onset of kinetics (just after incubation) means that the solution is not seed-free, and the aggregation rates and pathways derived from such kinetics are not reliable. In Materials and Methods, we describe the all preparation procedure apt to avoid the presence of seeds. Apart from this, such analysis has been performed for all sample preparations, and used as a rejection criterion to reach an acceptable level of reproducibility.

Fibrillation kinetics monitored by thioflavin T fluorescence

Aggregation kinetics have been monitored by fluorescence spectroscopy, upon addition of the dye thioflavin T to the solution. This is a classical probe for amyloid fibrils, since it fluoresces when bound to fibrils.[55,56]

The molecular details of thioflavin T binding are not completely understood, and its ability to reveal fibril formation should be tested on each specific system. Also, the addition of thioflavin T to the protein solution could affect the aggregation. In our experimental conditions, addition of thioflavin T caused a delay on insulin aggregation only at the lower temperature (50 °C), or at lower concentrations ($c < 100$ μM). At higher temperatures, no significant effect was observed (cf. Nielsen *et al.*, 2001).[40]

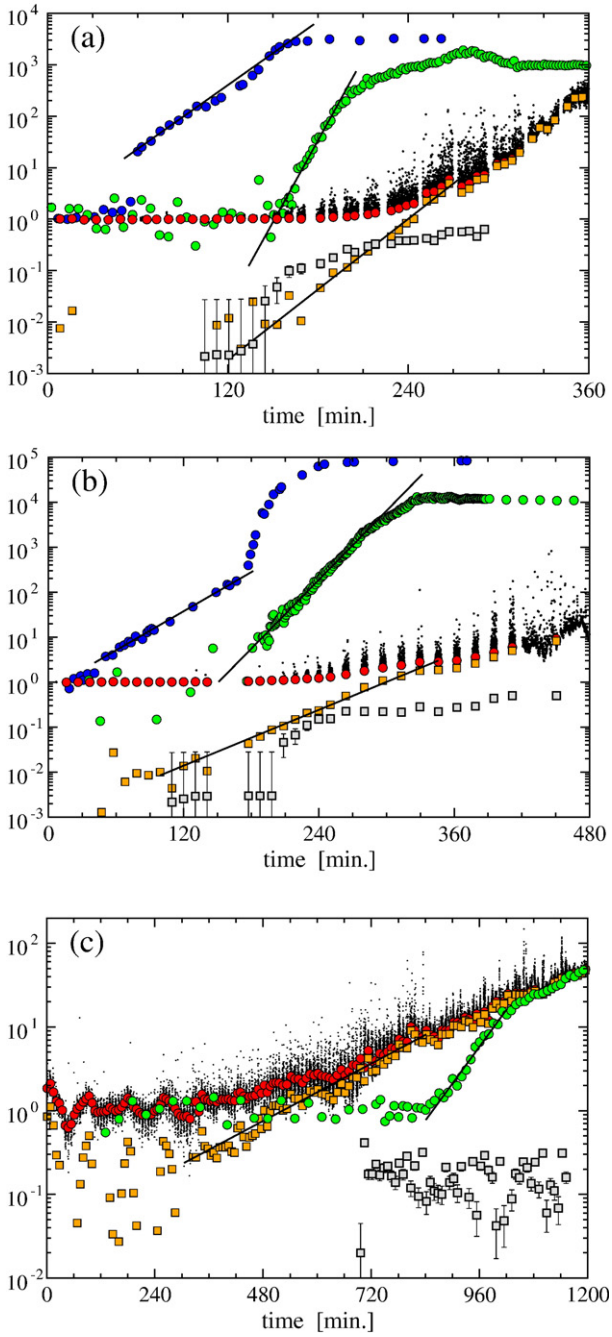

Figure 8. Kinetics of insulin aggregation at 60 °C. Blue circles: Thioflavin T fluorescence at 487 nm. Green circles: Small angle light scattered intensity at 2000 cm^{-1}. Black dots: 90° scattered intensity, $I_{\delta T}$. Red circles: Average fast (Gaussian distributed) intensity, $\langle I_F \rangle$. Orange squares: Average fast intensity, with respect to the initial value: $\langle I_F \rangle$ $(t) - \langle I_F \rangle$ (0). Grey squares: Slow fraction of 90° scattered intensity. Solid lines: Fits to exponential functions. All the intensity signals have been normalized with respect to the initial value. Insulin concentrations are 800 μM (a), 210 μM (b), and 90 μM (c).

As shown in Manno *et al.* (2006),[21] the large increase in thioflavin T fluorescence is correlated with the growth of large-size aggregates and not only with the growth of amyloid filaments and fibrils (Figure 8). If plotted on linear scale, the kinetics of thioflavin T fluorescence would exhibit a lag-time, which is coincident with the onset of non-Gaussian noise in scattered light intensity. Indeed, the thioflavin T fluorescence is not straightforwardly related to the mass concentration of protein assembled in amyloid fibrils. Such a close relation has been carefully tested in some important cases, such as the Alzheimer's amyloid β-protein.[57] However, the present results stress that such a relation should be verified for each specific system and condition. In general, the thioflavin T fluorescence is due to the binding of the dye to the repeated β-sheet structure of fibrillar elongated aggregates. Thus, it is closely related to the available surface of fibrils. In the case of thin filaments, the relation between available surface and mass concentration is more direct. In other cases, as in the present one, fibrils may be thick and assembled in bundles. Therefore, even though the growth of thioflavin T fluorescence does mark the growth of fibrils, one needs to be cautious about the correct labeling of fibril mass concentration.

By plotting the signals on a semilogarithmic scale, we note that thioflavin T fluorescence grows from almost the very beginning of the kinetics, and it exhibits no lag-time (Figure 8, blue circles). In other words, the lag-time is due to the sensitivity of the different experimental techniques, as clearly elicited by classical nucleation theories.[18,58,59] Moreover, the first part of signal growth is exponential and it exhibits approximately the same rate of the first part of the scattered intensity signal. Thus, we can consistently take the two signals as both related to the same aggregation mechanism.

Aggregation rates

In Figure 8, we show a summary of the different techniques and approaches used to monitor insulin aggregation at 60 °C. The 90° scattered intensity (black dots) is shown along with the average "fast" intensity $\langle I_F \rangle$ (red circles). The latter has been also displayed by subtracting the initial value (orange squares) in order to highlight its exponential growth and its resemblance to the thioflavin T signal at the early stages (blue circles). Such data have been fit by the following expression: $\langle I_F \rangle$ $(t) - \langle I_F \rangle$ $(0) = A\exp(\kappa t)$.

The aggregation rates κ obtained from the fast intensity signal for each temperature and concentration are shown in Figure 9. The rates are mainly proportional to the mass concentration, as expected from a process which is due to diffusion and coalescence of clusters of different sizes. Moreover, on a semilogarithmic scale, such rates are linearly correlated with the inverse temperature T^{-1}, with a slope of 12.5±1 (Figure 9). This suggests that there is one main aggregation mechanism with an activation barrier of 25(±2) kcal/mol.

In more detail, we can use aggregation theory[60] and assume that the aggregation rate is given by $\kappa = \kappa_0 \exp(\Delta G / k_B T)$, where ΔG is the free energy barrier for coagulation of two given clusters of

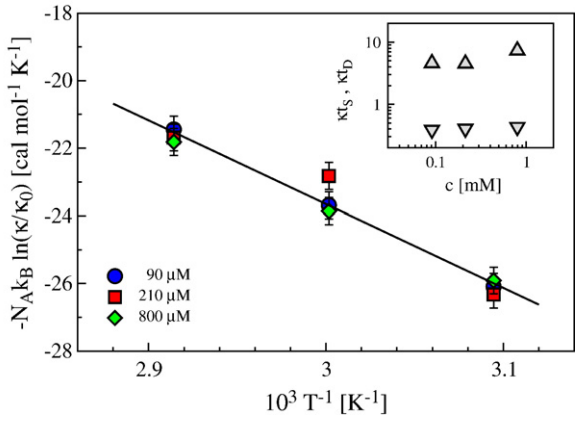

Figure 9. Arrhenius plot of fibril aggregation rates κ. Solid line is a linear fit to data. Inset: inverse exponential rates t_S (down triangles) and delay times t_D (up triangles), of slow aggregates for solutions incubated at 60 °C. Both times are multiplied by the rate κ.

proteins or fibrils, and the prefactor κ_0 is related to the number of objects which enter the "sphere of influence" of another object per unit time.[29,61] From classical coagulation theories,[60] we can assume that the prefactor is proportional to number concentration ρ and to DR, where $D = k_B T / (6\pi\eta R)$ is the diffusion coefficient of an object whose sphere of influence is R, at temperature T and with medium viscosity η. Thus, we put $\kappa_0 = 8\pi DR\rho = 1.33 k_B T \eta^{-1} c N_A M_1^{-1}$, where k_B is the Boltzmann factor, N_A is the Avogadro number, c is the mass concentration, and M_1 is the mass of insulin. With such a reasonable assumption, we can estimate the order of magnitude of both the enthalpy and entropy costs associated with coagulation, by fitting data in Figure 9 to the expression:

$$-N_A k_B \ln(\kappa/\kappa_0) = \Delta S - \Delta H T^{-1} \qquad (2)$$

The latter expression implicitly assumes that an Arrhenius law holds in the considered range, and therefore a single activation barrier is involved. This approximation is actually committed by the narrow temperature range. We obtained $\Delta H = 25(\pm 2)$ kcal mol^{-1}, and $\Delta S = 50(\pm 2)$ cal mol^{-1} K^{-1}, with a corresponding average free energy of $\Delta G = 8(\pm 3)$ kcal mol^{-1}. The error associated with ΔH is mainly due to the experimental error on the rates κ, which is on the order of 10%. On the other hand, the error in the entropic term ΔS and consequently in the free energy is affected also by the intrinsic indetermination in the prefactor κ_0. If we assume an error in κ_0 of 100%, the error in ΔS is on the order of $N_A k_B$. Also, it should be recalled that the values of the rates are inferred from a limited number of temperatures and concentrations.

In Figure 8, the lag-time is marked by the emergence of the slow fraction of 90° scattered intensity (gray squares) as well as by the onset of growth of the small angle light scattered intensity at 2000 cm^{-1} (green circles). The latter signal can be related to the concentration of large-size clusters of bundles: the growths of these two signals are akin as shown in Figure 5(b). The inverse rates of the exponential growth displayed by the small angle scattered intensity (t_s) and the lag-times (t_d), measured as the onset of non-Gaussian noise, are driven by the rates of 90° scattered intensity κ, as highlighted by the inset of Figure 9. Indeed, these two characteristic times are not related to different aggregation mechanisms or processes: the first one (the inverse rate t_s) is associated with the concentration of large aggregates, whose growth is taken into account in the 90° scattered intensity, and the second one (the lag-time t_d) is only due to a lack of experimental sensitivity.

Therefore, the significant parameter for insulin aggregation is the rate κ of the exponential growth measured by large angle light-scattering, and filtered by our statistical analysis.

Discussion

In the present work, we have studied the kinetics of aggregation of human insulin in acidic solution (pH 1.6 at 60 °C), by incubation at high temperatures (between 50 and 70 °C). In such conditions, insulin is known to form amyloid fibrils.[10,16,62] However, the overall aggregation process leads to morphological heterogeneity, involving amyloid fibrils and filaments, clusters of fibrils, floccules and spherites, as well as the formation of a thixotropic transparent gel.[14]

Notwithstanding the general scheme established by Waugh, who showed that there were three main aggregation steps (nucleation, elongation, flocculation),[14] the formation of the different structures can be concurrent and intertwined in the course of the aggregation kinetics. Aware of this warning, we have studied insulin aggregation using different experimental techniques.

In particular, the combined use of small and large angle light-scattering and QLS allows us to monitor *in situ* the formation of structures from nanometers to millimeters. Moreover, since scattering is more sensitive to high molecular masses, we have performed time-lapse AFM and thioflavin T fluorescence experiments to increase the resolution of small size aggregates.

We have highlighted the different stages of kinetics and the morphogenesis of the different structures involved in insulin aggregation.

The onset of kinetics

At the onset of the kinetics, upon incubation at high temperatures, we have found that the insulin molecules are mainly in monomeric and dimeric form, in agreement with previous studies.[63] In particular, the weight-averaged molecular mass measured by our light-scattering experiments corresponds to 1.8, 1.7, and 1.5±0.2 insulin molecules,

respectively, at 800, 210, and 90 μM. However, these values are related to the second moment of the actual distribution of insulin oligomers. Indeed, both QLS and AFM experiments at the onset of the kinetics point out the existence of a wider oligomer distribution, including up to several tens of insulin molecules (Figures 1 and 2).[32] This is in agreement with previous experiments performed by electrospray mass-spectrometry on bovine insulin in analogous conditions, which revealed the presence of oligomers of up to 12 units.[64]

In a previous paper,[32] we have analyzed the oligomer distribution observed by AFM experiments (Figure 2(a)). Such a distribution was shown to exhibit an exponential tail and was reasonably fit to an exponential times a power law, thus suggesting a coagulation-evaporation mechanism operating at the very early stages.[32,65]

Across the lag-time

The most evident outcome of the present experimental results is the existence of a lag-time in the aggregation kinetics. This means that the molecular species and aggregates that are observable in solution are dramatically different before and after a given time. Of course this is not a sharp phase transition and the so-called lag-time (or delay time) is actually a relatively small time window. The changes seem abrupt only due to the low experimental sensitivity or low time resolution. For example, in both QLS and AFM time-resolved experiments (Figures 1 and 2), one can observe fibrillar or large-size aggregates only after a given time (the lag-time), while the solution is latent in the early stages.

A clear hallmark of large molecular aggregates is the presence of non-Gaussian noise in the scattered light intensity. We have used an approach based on the analysis of scattered light intensity distribution to filter out the non-Gaussian noise and obtain the contribution ($\langle I_F \rangle$) due to aggregates with a fast dynamic, which are mainly protein oligomers or small fibrils.

With this filtering procedure, we have been able to observe the growth of the scattering signal immediately before and after the appearance of large aggregates, that is across the lag-time. More specifically, the lag-time is related to the increase in the mass and concentration of fibril bundles and clusters, as revealed by multiple angle light-scattering and QLS (Figures 1, 5, and 8), and it is also related to the increase in the mass concentration of elongated fibrils as revealed by the thioflavin T fluorescence and time-lapse AFM (Figures 2 and 8). All these structures are encompassed by the parameter x_S, obtained from the analysis of scattered light intensity, which measures the fraction of non-Gaussian (exponential) contribution to scattering. The onset of non-zero x_S can be unequivocally used as a measure of the lag-time. As seen in Figure 8, the increase of small angle scattering (green circles) and the major increase of thioflavin T fluorescence (blue circles) are both closely related with this value.

One can observe an exponential growth of the scattering signal with respect to the initial scattering (orange squares in Figure 8), in accordance with a preliminary report in our previous work.[21] An exponential increase occurs when the rate of growth of a signal is proportional to the signal itself. In the present case, the scattering signal is related to the weight averaged mass of protein aggregates. Therefore, this means that the aggregation process is catalyzed by the amount of aggregate mass. Such a mechanism has been extensively described by Ferrone, Hofrichter, and Eaton,[17–19] and applied to the explanation of Sickle Cell Hemoglobin polymerization. The simple addition of proteins to a fibril nucleus gives rise to a power law dependence of the mass as a function of time.[58,59]. In addition to this "homogeneous" nucleation and elongation, Ferrone *et al.* (1985)[17] includes a secondary "heterogeneous" pathway, which allows the addition of other nuclei to the surface of fibrils and consequent growth along the surface. In this way, the fibrils are not only elongated by tip-to-tip addition of monomers, but are also laterally thickened. This "heterogeneous" aggregation pathway is prompted by the preformed fibrils, and therefore it is actually responsible for the exponential growth.

Such an aggregation mechanism has been highlighted by D.F. Waugh for the formation of insulin amyloid fibrils.[14,16] His works on insulin fibrillation are excellent and inspiring, yet they date back to the fifties, and therefore a confirmation of his proposals with more recent techniques could be useful. Such a scheme has been also recently confirmed by Librizzi and Rischel on bovine insulin in acetic acid.[20]

In Figure 8, the scattered intensity at a low q (green circles) is also shown. This signal is useful to visualize the increase of the concentration of largest objects, which are the clusters of fibril bundles. The growth of this signal is approximately exponential as a function of time. Its rate is well correlated with the rate of the fast scattered intensity. In the inset of Figure 9, the inverse of such rate t_s, as well as the lag-time measured by the onset of the non-Gaussian noise t_d, times the rate of fibrillar aggregation κ is shown to be quite concentration-independent. This fact points out directly that the formation of larger aggregates and smaller fibrils is due to the same mechanism, and gives a rationale for the simultaneous appearance of aggregates on all the length scales.

We have studied insulin aggregation under different conditions of initial monomer concentration and temperatures. The rate of the exponential growth of the scattering intensity has been found to be proportional to the concentration, thus suggesting a process involving diffusion and coalescence of clusters of different sizes. Also, even though the explored temperature range is quite limited, we have attempted an estimate of the energetic barriers involved in the aggregation process by making an Arrhenius plot of the aggregation rates. These

calculations provide an enthalpy of $\Delta H \approx 25$ kcal mol^{-1}, and a corresponding average free energy of $\Delta G \approx 8$ kcal mol^{-1}. Analogous values have been found for the fibrillogenesis of the Aβ(1-40) peptide involved in Alzheimer's disease.[29,66] These values of activation energy are reminiscent of those expected for protein conformational changes[67] (see e.g. Kusumoto *et al.*, 1998 and references therein).[66] Indeed, the formation of insulin fibrils has been found to involve intermediate protein conformations with a partial unfolding of the α-helical secondary structures.[22–25]

The early stages

In the present work, we have focused on the formation of aggregates of different sizes and we are not deeply addressing the early stages of aggregation. However, an interesting suggestion can be derived from the observation of the thioflavin T fluorescence experiments. While on a linear plot, the major increase of thioflavin T fluorescence is correlated with the lag-time,[21] a logarithmic plot, as in Figure 8 (blue circles), shows a low, yet monotonic, increase of the signal. Such growth is exponential and has a rate which is essentially equivalent to that of the fast scattered intensity.

This equivalence encourages us to ascribe the growth of the intensity signal to the growth of the mass concentration of amyloid fibrils, even if it really measures the second moment of the mass distribution of all kinds of molecular aggregates. Also, the growth of the scattering signal has been extended to times preceding the lag-time, thanks to the statistical approach. Yet, the extrapolation of the exponential growth to the onset of kinetics is not guaranteed. The thioflavin T signal suggests that the exponential growth initiates in the early stages and that a nucleation step, if any, is confined to the first minutes of kinetics.

In a previous work, we analyzed the early stages of insulin aggregation by AFM.[32] We found that a distribution of slightly elongated (prefibrillar) insulin oligomers was reached in a few minutes through an evaporation-condensation kinetic mechanism. Also, we have shown that aggregates have a mean eccentricity of 0.75. This means that the ratio between the major and the minor axis of the ellipsoidal aggregates is 1.5. On average, aggregate size and eccentricity are correlated so that larger oligomers exhibit a larger eccentricity. This observation suggests that such oligomers are prefibrillar aggregates. The present results support those findings by assigning to such aggregates the role of prefibrillar structures. In other word, the amyloidogenic aggregates are formed in the very early stages, and the subsequent aggregation kinetics are controlled by the secondary heterogeneous mechanism discussed above. Another observation of fibril growth in the early kinetics stage has been recently reported in a valuable paper by Winter and coworkers, by time-resolved experiments of fluorescence anisotropy.[44]

Such results as well as the present findings coherently suggest a non-cooperative fibrillogenesis process. However, these experiments appear too preliminary for an exhaustive explanation of the early aggregation stages, which need further investigation.

In this framework, it could be interesting to clarify the role of the conformational changes observed in amyloid-forming insulin. More specifically, it would be interesting to clarify whether they are related to the formation of a nucleus, or to the formation of an active intermediate, as the present results would suggest.

Aggregate morphology at later stages and the overall kinetic scheme

At the concentrations under study (in the submillimolar range), the aggregation process leads to the formation of a variety of structures spreading over many length scales. After several hours, aggregates become so large and sample-spanning that gelation occurs. The sample is then stuck in a stationary (non-equilibrium) state, with almost no flux. At this extreme stage, it is possible to observe floccules embedded in the gel matrix even by eye (data not shown).

At later kinetic stages (preceding the eventual structural arrest in the gel phase), a number of different structures can be observed, as highlighted by the scheme of Figure 10.

(i) Amyloid fibrils with a diameter in the range of 10 to 20 nm and with lengths of a few microns (Figures 2–4), analogous to early observations by electron microscopy,[10,15] and to more recent studies on bovine,[6,22,39–51] and human[23,24,49,52,53] insulin. In particular, the fibril diameter is consistent with the six-protofilaments observed by Jimenez *et al.* (2002).[50] By performing double quenching experiments (Figure 4), we have observed that elongated fibrils can form loops with a diameter of the order of half a micron and with a persistence length of about 100 nm. The resolution of AFM images was sufficiently high to show that fibrils have the structure of a right-handed helix with a pitch of the order of 10 nm.

(ii) Bundles of fibrils with a typical size of a few microns (Figure 6(a)), analogous to other reports in the literature.[6,49] As suggested by the kinetics of multiple angle light-scattering (Figure 5), such aggregates are rather compact objects. We may speculate that they are formed by the collapse of different fibrils by alignments of fibrils along their main axes, or they may be due to lateral growth of fibrils.

(iii) Clusters of fibrils and bundles (Figure 6(b)). The presence of clusters of fibrillar objects is supported by the kinetics of multiple angle

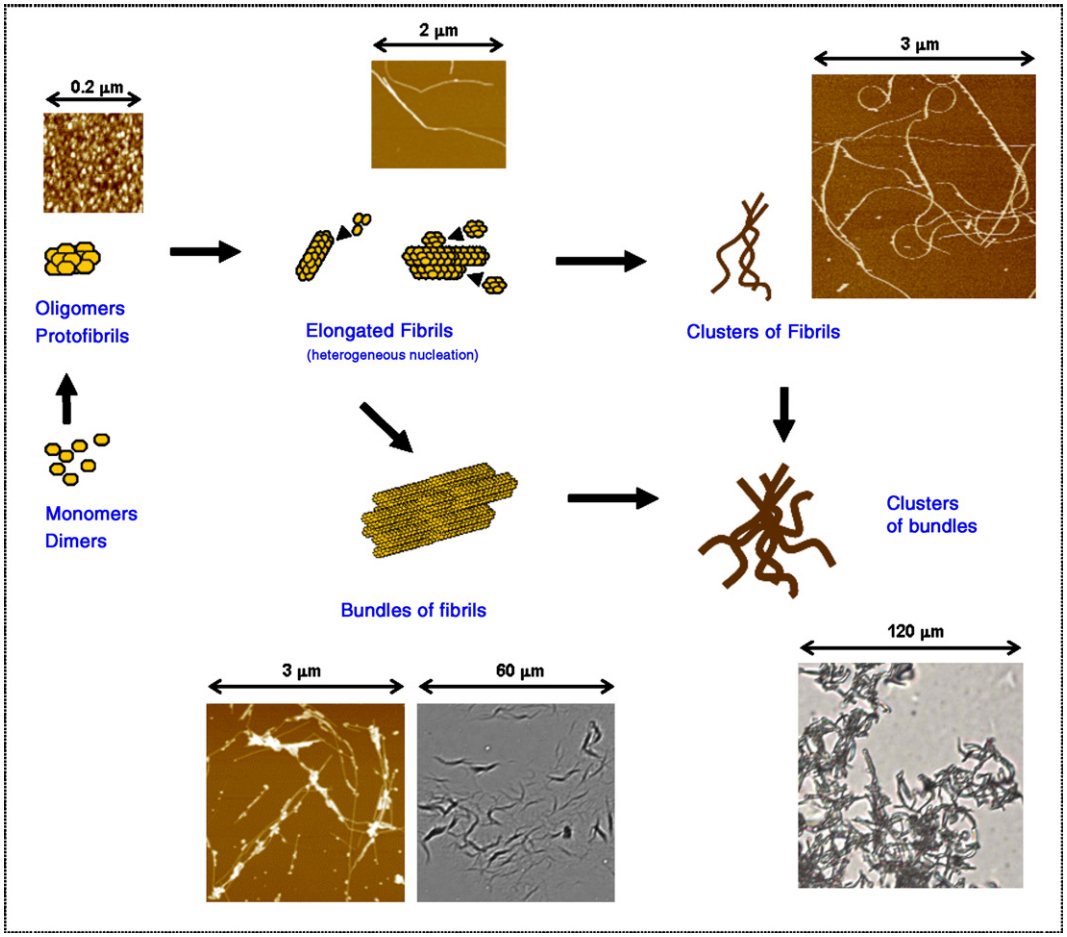

Figure 10. Scheme of insulin aggregation: from amyloid fibrils to clusters of fibrils and bundles. The AFM and optical microscopy images are taken from Figures 2, 4 and 6.

light-scattering (Figure 5). Such aggregates are likely to have different morphologies and to span different length scales. Indeed, we may argue that they originate either from the coalescence of fibril bundles to form loose tangles, or by the entanglement of more elongated fibrils, which eventually leads to gelation. In both cases, we deal with amorphous fractal-like clusters, and we are not aware of the existence of spherulites. These are peculiar structures made of fibrils radially assembled around a center, which exhibit birefringence.[11] The formation of spherulites occurs typically when a large number of short fibrils are present in solution, and thus it is enhanced by an increase in temperature, concentration, and ionic strength.[11]

In summary, the present experiments have addressed the formation of insulin amyloids and aggregates. The general pattern of aggregate formation is synthesized in the diagram of Figure 10. The aggregates have been monitored on different length scales by using different experimental techniques. We have found that different structures are formed simultaneously likely by diffusion and coalescence of (fibrillar) aggregates. The lag-time in the aggregation kinetics is due to the low experimental sensitivity, and any nucleation step probably arises in earlier stages.

Materials and Methods

Sample preparation

Recombinant human insulin powder (purchased from Sigma Chemical Co. and used without further purification) was directly dissolved at 5 °C in buffer solution (50 mM KCl/HCl in Millipore SuperQ water, pH 1.6 at 60 °C). The solution was ultra-filtered in a centrifuge (5 °C, 15 min, 5000 rpm) using a Millipore ultrafiltration unit with a cutoff of 100 kDa, and then filtered (0.22 m Millex-GV) into quartz cuvettes or glass vials for incubation at the desired temperature. Ultrafiltration was initially used to guarantee a seed-free condition at the onset of kinetics. We verified that insulin dissolution was not significantly improved by this step, which was then skipped in AFM experiments. Insulin concentration was measured by UV absorption at 276 nm by using an extinction coefficient of 1.0675 cm^{-1} $(mg/ml)^{-1}$. The final concentrations were consistent with those calculated by

weighting insulin powder, thus confirming that essentially no material was lost by filtering. All chemicals were analytical grade.

Atomic force microscopy

After given time intervals, 10 μL of incubated protein solution was diluted into 1 ml buffer solution, quenched to 0 °C to rapidly inhibit further aggregation, and used for AFM experiments. A few microliters of the insulin solution were dropped onto a freshly cleaved mica substrate (quality ruby muscovite). After a few minutes, the sample was washed dropwise with Millipore SuperQ water, and then dried with a gentle stream of dry nitrogen. The drying is expected to remove most of the surface water, while keeping the protein aggregates substantially hydrated locally. Images of the protein aggregates were recorded with a Multimode Nanoscope IIIa Atomic Force Microscope (Veeco Instruments, Santa Barbara, CA, USA), operating in Tapping Mode inside a sealed box where a dry nitrogen atmosphere was maintained. We used rigid cantilevers with resonance frequencies of about 300 kHz, and equipped with single crystal silicon tips with a nominal radius of curvature of 5 to 10 nm. Typical scan size was 500×500 nm^2 (512×512 points), and scan rate was 1 to 2 Hz.

Quasi-elastic light-scattering

Time-resolved QLS experiments were performed at different temperatures and concentrations, immediately after sample preparation. Samples were placed in a thermostated cell compartment of a light-scattering goniometer (Brookhaven Instruments BI200-SM), equipped either with a 100 mW Ar laser tuned at $\lambda_0 = 514.5$ nm or with a He-Ne laser ($\lambda_0 = 632.8$ nm). The temperature was controlled within 0.05 °C with a thermostated recirculating bath. Scattered light intensity at 90° and its time autocorrelation function $g_2(t)$ were measured simultaneously by using a Brookhaven BI-9000 correlator. Correlation functions $g_2(t)$ were analyzed by using a constrained regularization method[68] in order to determine the distribution of the apparent diffusion coefficients D: $g_2(t) = 1 + |\int P(D)\exp(-Dq^2t)\mathrm{d}t|^2$, where q is the scattering vector. The z-average hydrodynamic radius R_h was calculated by averaging over diffusion coefficient distributions and assuming the Stokes-Einstein relation: $k_B T D_z^{-1} = 6\pi\eta R_h$, where k_B is the Boltzmann constant, T is the temperature, and η is the solution viscosity.[54]

Large angle light-scattering

The experimental set up described for QLS experiments was also used for large angle static light-scattering measurement. Multiple angle experiments were performed in order to determine the scattered intensity $I(q)$ at different scattering vector $q = 4\pi\tilde{n}\lambda_0^{-1}\sin(\theta/2)$, where θ is the scattering angle and $\tilde{n}$ is the medium refractive index.[54] The scattered intensity $I(q)$ is given in terms of Rayleigh ratio $I_R(q) = I(q)/I_s r^2/V_s$, where I_s is the intensity of the laser source, V_s is the scattering volume, and r is the distance of the detector from the sample. Absolute values for scattered intensity (Rayleigh ratio I_R) were obtained by normalization with respect to Toluene, whose Rayleigh ratio was taken as 32×10^{-6} cm^{-1} and 14×10^{-6} cm^{-1} respectively at 514.5 nm and 632.8 nm.

Measurements of the scattered intensity provide the Rayleigh ratio $I_R(q)$, which is related to the weight averaged molecular mass M_w of aggregates by the relation: $I_R(q) = K_R c M_w P(q)$, with the instrumental factor $K_R = 4\pi^2\tilde{n}^2(\mathrm{d}\tilde{n}/\mathrm{d}c)^2\lambda_0^{-4}N_A^{-1}$, where c is the mass concentration, $P(q)$ is the z-averaged form factor, $\tilde{n}$ is the medium refractive index, λ_0 is the incident wavelength, and N_A is the Avogadro's number.[54] We calculated the average molecular mass M_w at the onset of kinetics by taking $(\mathrm{d}\tilde{n}/\mathrm{d}c) = 0.18$ cm^3 g^{-1},[63] and $P(q) = 1$. The form factor is related to the average shape and size of scatterers. However, it is equal to 1 when the size of solutes is much smaller than q^{-1} (see Berne and Pecora, 1976).[54]

Small angle light-scattering

Small angle light-scattering measurements were performed using an experimental device built in our laboratory, and equipped with a 30 mW Melles Griot helium-neon laser and a charge coupled device Pulnix TM765 camera. The apparatus was constructed following previous works.[69,70] Measurable intensities span a wide dynamical range (from 1 to 33105, in arbitrary units), thanks to a software integration of multiple exposure times (1/60 ÷ 1/10000 s). After background subtraction (performed by home-made software), reliable measurements could be made over a range of scattering angles of 0.1° to 11°, corresponding to scattering vector magnitudes of 0.02 to 2 μm^{-1}, with a maximum resolution of 0.01°. Small angle light-scattering experiments were performed simultaneously with large angle light-scattering experiments, by using the same sample preparation.

Optical microscopy

At the end of the scattering experiments, a few drops of solution were put on a glass slide for imaging with a phase-contrast Zeiss Axioskop2 plus microscope.

Thioflavin T fluorescence spectroscopy

A stock solution of thioflavin T was prepared at a concentration of 1 mM in SuperQ Millipore water and stored at 4 °C protected from light until usage. Aliquots of thioflavin T stock solution were added to the insulin solutions directly into the quartz cuvette (1 cm optical path). The final concentration of thioflavin T was 20 μM. Fluorescence emission spectra (from 470 to 560 nm) were recorded after quenching to the incubation temperature by using a Jasco FP-6500 spectrofluorometer with excitation at 450 nm and scan speed at 100 nm/min. For each sample, the signal was obtained as thioflavin T intensity at 482 nm after subtraction of a blank measurement. Also, fluorescence experiments have been performed simultaneously with scattering experiments, by using the same sample preparation.

Acknowledgements

We thank C. Rishel and F. Librizzi for collaboration and access to unpublished data, and D. Giacomazza and R. Noto for several discussions

and J. Newman for collaboration and critical reading of the manuscript. We gratefully acknowledge the help of L. Finzi and P. Milani. This work was partially supported by the Italian Ministero della Salute through the project Neuropatie animali-analisi molecolari e funzionali della proteina prionica in razze bovine siciliane.

Appendix A

In this section, we give a more detailed explanation of the terms and analytic expression included in equation (1).

The first contribution in equation (1) concerns free proteins with mass M_p and concentration c_p. It is important to include this term, since the asymptotic value of I_R at high q is quite close to KcM_p (which is about 6.3×10^{-6} cm^{-1} in the case considered in Figure 5), and therefore the greater portion of the protein molecules is still free in solution at the end of the lag-time.

The second term in equation (1) refers to fibril elements with mass M_f, concentration c_f, structure factor $S_f(q)$, and form factor $P_f(q)$. Such fibril elements have been modeled by taking into account the AFM images (Figure 4), where one can observe fibrils with an average diameter of about 15 nm and an average persistence length of the order of 100 nm. Such a diameter is consistent with the six-protofilament fibril observed by Jimenez *et al.* (2002).[50] Thus, assuming a van der Waals diameter of 1.5 nm for an insulin molecule,[71,72] we consider such fiber elements as quasi-rigid (or semi-flexible) rods made up of about 200 molecules ($M_f=200M_p$), and with a characteristic size $R_f=100$ nm. Since our data are in the range $qR_f<1$, we can approximate the form factor $P_f(q)=1$.[54] A given fraction γ_f of such fibril elements is taken to be isolated and the remaining part to be assembled in larger structures, which can be either entangled fibrils or longer, branched clusters of fibril elements. The overall structure factor is:

$$S_f(q) = \gamma_f + (1-\gamma_f) n_f F(q;\, \xi,\, d,\, R_f) \qquad \text{(A1)}$$

where $F(q; \xi, d, R_f)$ is a function which indicates the structure factor of a fractal aggregate with a correlation length ξ and a fractal dimension d, made of n_f objects of size R_f (the fibril elements).[73] This expression is often used, since it introduces the minimum of information, that is the two characteristic lengths and the exponent of the power law dependence. Therefore, by using this expression, we are not pushing any artifact in the model. The low value of the fractal dimension ($d=1.55$), taken from the initial slope of scattering data (Figure 5), is typical of clusters of linear filaments and it is close to the value for self avoiding random polymers,[74] which has been already observed in amyloid aggregates.[29] The value of the correlation length cannot be correctly identified from our data, whose range is $q\xi>1$. In the limit of high $q\xi$, one obtains:[73]

$$n_f F(q;\, \xi\to\infty,\, d,\, R_f) = 1 + \frac{\Gamma(d+1)}{d-1}\left(qR_f\right)^{-d} \qquad \text{(A2)}$$

where $\Gamma(x)$ is the gamma function.

The last term in equation (1) refers to bundles of fiber elements, with mass $M_b=NM_f$, concentration c_b, structure factor $S_b(q)$, and form factor $P_b(q)$. Such compact elongated bundles, which have been observed by optical microscopy (Figure 6), are assembled in large fractal aggregates with a correlation length ξ and a fractal dimension $d=1.55$, and they are made of n_b bundles of radius of gyration R_b. The structure functions can be modeled by the same expression used for fibril clusters equation (4), in the limit of high $q\xi$:

$$S_b(q) = 1 + \frac{\Gamma(d+1)}{d-1}\left(qR_b\right)^{-d} \qquad \text{(A3)}$$

In this case, we are assuming that all the bundles are assembled into fractal clusters. Such bundles are compact objects, as elicited by the slope of structure functions for $q>\mu\text{m}^{-1}$. Thus, we can model them as pseudo-fractals with a dimension of $d=3$, radius of gyration $R_b=\xi_b\sqrt{6}$, made of N fibril element of characteristic size R_f.[73]

$$P_b(q) = 1 + 6\left(\frac{\xi_b}{R_f}\right)^3 \times \left[1+(q\xi_b)^2\right]^{-1} \frac{\sin(2\arctan(q\xi_b))}{2q\xi_b} \qquad \text{(A4)}$$

Appendix B

In this section, we will briefly recall the main conceptual outline which brings to the expression of the distribution of scattering intensity, which is used for data analysis (Figure 7).

A large number of scatterers give a scattered electric field which is a complex Gaussian variable and thus an instantaneous intensity $I(t)$ which is exponentially distributed.[75] In scattering experiments, the total measured intensity $I_{\delta T}$ is time averaged over a small time windows δT: $I_{\delta T}=\delta T^{-1}\int_{\delta T} I(t)\mathrm{d}t$. In our experimental condition, we put $\delta T=0.9$ s. When the coherence time of the scattered field is short compared to the interval δT, $I_{\delta T}$ can be taken as the sum of several statistically independent contributions with the same exponential distribution function. From the central limit theorem, this results in a Gaussian distribution. On the other hand, when the field coherence time is much larger than δT, the measured intensity is equivalent to the instantaneous intensity, and it is exponentially distributed.[21,75,76]

In a solution of freely diffusing aggregates, the coherence time τ can be derived from the diffusional properties of a scatterer of size R_h (hydrodynamic

radius), and diffusion coefficient D: $\tau = (Dq^2)^{-1} = (k_B T q^2)^{-1}\, 6\pi\eta R_h$, where η is the medium viscosity, T is the temperature, k_B is the Boltzmann constant, and q is the scattering vector.[54]

In the present work, the temperatures are between 50 and 70 °C, and the scattering vector is either 18.7 μm^{-1} (He-Ne laser) or 22.8 μm^{-1} (Argon laser). Thus, when $R_h > 200$ μm, the coherence time is larger than 1 sec, so that $\delta T / \tau \leq 1$, and therefore the scattered intensity distribution is exponential. On the other hand, when $R_h \leq 4$ μm, we have $\delta T / \tau > 50$, and the scattered intensity distribution is Gaussian. A scatterer with a size in the intermediate range (i.e. tens of microns) gives a scattered intensity with a gamma distribution.[75] Therefore, objects with a size smaller than a few microns, or more generally much smaller than hundreds of microns, give a Gaussian contribution to scattered intensity, while the appearance of non-Gaussian noise is related to the formation of large-size slowly diffusing aggregates, such as huge clusters of fibrils or floccules.

In order to filter out the contribution of large aggregates, we assume that the scattered field is the sum of two terms: (i) a "Fast" fluctuating, Gaussian contribution (due to small-size objects), with average $\langle I_F \rangle$ and variance $\langle I_F \rangle^2\, \tau_F / \delta T$ (where τ_F is the coherence time of this fast field); and (ii) a "Slow" exponentially distributed contribution (due to large-size slowly diffusing objects), with average $\langle I_S \rangle$. An analogous approach has been originally developed for non-ergodic systems.[76,77]

We have recently shown that up to the second order temporal coherence the distribution of the total measured intensity is equivalent to the sum of two statistically independent terms[76]: an exponentially distributed contribution with average $\langle I_E \rangle = \langle I_S \rangle + 2\tau_F\, (\delta T)^{-1}\, \langle I_F \rangle$ and a Gaussian distributed contribution with average $\langle I_G \rangle = \langle I_F \rangle - 2\tau_F\, (\delta T)^{-1}\, \langle I_F \rangle$ and variance: $\sigma^2 = \langle I_F \rangle^2\, \tau_F / \delta T$. Therefore, the intensity distribution function is obtained by the convolution of a Gaussian with an exponential function:[21,76]

$$P(I_{\delta T}) = \langle I_E \rangle^{-1} \exp\{-[I - \langle I_G \rangle]/\langle I_E \rangle\} \times \left[\operatorname{erf}\left(\frac{I}{\sigma\sqrt{2}} - \frac{\langle I_G \rangle \langle I_E \rangle + \sigma^2}{\langle I_E \rangle \sigma \sqrt{2}} \right) + \operatorname{erf}\left(\frac{\langle I_G \rangle \langle I_E \rangle + \sigma^2}{\langle I_E \rangle \sigma \sqrt{2}} \right) \right] \frac{1}{2} e^{\frac{\sigma^2}{2\langle I_E \rangle^2}} C \quad \text{(B1)}$$

$\operatorname{erf}(x)$ being the error function $\operatorname{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$, and $C \approx 1$ is a normalization constant: $C = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\langle I_G \rangle}{\sigma\sqrt{2}} \right) \right]$

References

1. Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struck. Biol.* **8**, 101–106.
2. Cooper, G. J., Willis, A. C., Turner, R. C., Sim, R. B. & Reid, K. B. (1987). Purification and characterization of a peptide from amyloid-rich pancreas of type 2 diabetic patients. *Proc. Natl Acad. Sci. USA*, **84**, 8628–8632.
3. Prusiner, S. B. (1998). Prions. *Proc. Natl Acad. Sci. USA*, **95**, 13363–13383.
4. Rochet, J. C. & Lansbury, P. T. J. (2000). Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struck. Biol.* **10**, 60–68.
5. Harper, J. D. & Lansbury, P. T. J. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Ann. Rev. Biochem.* **66**, 385–407.
6. Khurana, R., Ionescu-Zanetti, C., Pope, M., Li, J., Nielson, L., Ramirez-Alvarado, M. *et al.* (2003). A general model for amyloid fibril assembly based on morphological studies using atomic force microscopy. *Biophys. J.* **85**, 1135–1144.
7. Brange, J., Andersen, L., Laursen, E. D., Meyn, G. & Rasmussen, E. (1997). Towards understanding insulin fibrillation. *J. Pharm. Sci.* **86**, 517–525.
8. du Vigneaud, V., Sifferd, R. H. & Sealock, R. R. (1933). The heat precipitation of insulin. *J. Biol. Chem.* **102**, 521–533.
9. Langmuir, I. & Waugh, D. F. (1940). Pressure-soluble and pressure-displaceble components of monolayers of native and denatured proteins. *J. Am. Chem. Soc.* **62**, 2771–2793.
10. Waugh, D. F. (1944). The linkage of corpusular protein molecules. i. fibrous modification of insulin. *J. Am. Chem. Soc.* **66**, 663.
11. Waugh, D. F. (1946). A fibrous modification of insulin i. the heat precipitate of insulin. *J. Am. Chem. Soc.* **68**, 247–250.
12. Waugh, D. F., Thompson, R. E. & Weimer, R. J. (1950). Assay of insulin in vitro by fibril elongation and precipitation. *J. Biol. Chem.* **185**, 85–95.
13. Waugh, D. F. (1948). Regeneration of insulin from insulin fibrils by the action of alkali. *J. Am. Chem. Soc.* **70**, 1850–1857.
14. Waugh, D. F., Wilhemson, D. F., Commerford, S. L. & Sackler, M. L. (1953). Studies of the nucleation and growth reactions of selected types of insulin fibrils. *J. Am. Chem. Soc.* **75**, 2592–2600.
15. Farrant, J. L. & Mercer, E. H. (1952). Electron microscopical observation of fibrous insulin. *Biochim. Biophys. Acta*, **8**, 355–359.
16. Waugh, D. F. (1957). A mechanism for the formation of fibrils from protein molecules. *J. Cell. Physiol.* **49**, 145–164.
17. Ferrone, F. A., Hofrichter, J. & Eaton, W. A. (1985). Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol.* **183**, 611–631.
18. Eaton, W. A. & Hofrichter, J. (1990). Sickle cell hemoglobin polymerization. *Advan. Protein. Chem.* **40**, 63–279.
19. Ferrone, F. (1999). Analysis of protein aggregation kinetics. *Methods Enzymol.* **309**, 256–274.
20. Librizzi, F. & Rischel, C. (2005). The kinetic behavior of insulin fibrillation is determined by heterogeneous nucleation pathways. *Protein Sci.* **14**, 3129–3134.
21. Manno, M., Craparo, E. F., Martorana, V., Bulone, D. & San Biagio, P. L. (2006). Kinetics of insulin aggregation: disentanglement of amyloid fibrillation from large-size cluster formation. *Biophys. J.* **90**, 4585–4591.
22. Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M. & Robinson, C. V. (2000). Formation of

insulin amyloid fibrils followed by ftir simultaneously with cd and electron microscopy. *Protein Sci.* **9**, 1960–1967.
23. Whittingham, J. L., Scott, D. J., Chance, K., Wilson, A., Finch, J., Brange, J. & Dodson, G. G. (2002). Insulin at pH 2: structural analysis of the conditions promoting insulin fibre formation. *J. Mol. Biol.* **318**, 479–490.
24. Hua, Q.-X. & Weiss, M. A. (2004). Mechanism of insulin fibrillation: the structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate. *J. Biol. Chem.* **279**, 21449–21460.
25. Ahmad, A., Uversky, V. N., Hong, D. & Fink, A. L. (2005). Early events in the fibrillation of monomerics insulin. *J. Biol. Chem.* **280**, 42669–42675.
26. Nilsson, M. R. & Dobson, C. M. (2003). Chemical modification of insulin in amyloid fibrils. *Protein Sci.* **12**, 2637–2641.
27. Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A. & Teplow, D. B. (1996). On the nucleation and growth of amyloid β-protein fibrils: detection of nuclei and quantitation of rate constants. *Proc. Natl Acad. Sci. USA*, **93**, 1125–1129.
28. Lomakin, A., Benedek, G. B. & Teplow, D. B. (1999). Kinetic analysis of amyloid fibril formation. *Methods Enzymol.* **309**, 429–459.
29. Carrotta, R., Manno, M., Bulone, D., Martorana, V. & San Biagio, P. L. (2005). Protofibril formation of amyloid β-protein at low pH via a non-cooperative elongation mechanism. *J. Biol. Chem.* **280**, 30001–30008.
30. Goldsbury, C., Aebi, U. & Frey, P. (2001). Visualizing the growth of Alzheimer's aβ amyloid-like fibrils. *Trends Mol. Med.* **7**, 582.
31. Sluzky, V., Tamada, J. A., Klibanov, A. M. & Langer, R. (1991). Kinetics of insulin aggregation in aqueous solution upon agitation in the presence of hydrophobic surfaces. *Proc. Natl Acad. Sci. USA*, **88**, 9377–9381.
32. Podestà, A., Tiana, G., Milani, P. & Manno, M. (2006). Early events in insulin fibrillization studied by time-lapse atomic force microscopy. *Biophys. J.* **90**, 589–597.
33. Walsh, D. M., Lomakin, A., Benedek, G. B., Condron, M. M. & Teplow, D. B. (1997). Amyloid β-protein fibrillogenesis: detection of a protofibrillar intermediate. *J. Biol. Chem.* **272**, 22364–22372.
34. Huang, T. H. J., Yang, D. S., Plaskos, N. P., Go, S., Yip, C. M., Fraser, P. E. & Chakrabartty, A. (2000). Structural studies of soluble oligomers of the Alzheimer's β-amyloid peptide. *J. Mol. Biol.* **297**, 73–87.
35. Westlind-Danielsson, A. & Arnerup, G. (2001). Spontaneous in vitro formation of supramolecular β-amyloid structures, "βamy balls," by β-amyloid 1–40 peptide. *Biochemistry*, **40**, 14736–14743.
36. Hoshi, M., Sato, M., Matsumoto, S., Noguchi, A., Yasutake, K., Yoshida, N. & Sato, K. (2003). Spherical aggregates of β-amyloid (amylospheroid) show high neurotoxicity and activate tau protein kinase i/glycogen synthase kinase-3β. *Proc. Natl Acad. Sci. USA*, **100**, 6370–6375.
37. Relini, A., Torrassa, S., Rolandi, R., Gliozzi, A., Rosano, C., Canale, C. *et al.* (2004). Monitoring the process of hypf fibrillization and liposome permeabilization by protofibrils. *J. Mol. Biol.* **338**, 943–957.
38. Xu, S., Bevis, B. & Arnsdorf, M. F. (2001). The assembly of amyloido-genic yeast sup35 as assessed by scanning (atomic) force microscopy: an analogy to linear colloidal aggregation? *Biophys. J.* **81**, 446–454.
39. Jansen, R., Dzwolak, W. & Winter, R. (2004). Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy. *Biophys. J.* **88**, 1344–1353.
40. Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S. *et al.* (2001). Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry*, **40**, 6036–6046.
41. Sharp, J. S., Forrest, J. A. & Jones, R. A. L. (2002). Surface denaturation and amyloid fibril formation of insulin at model lipid-water interfaces. *Biochemistry*, **41**, 15810–15819.
42. Arora, A., Ha, C. & Park, C. B. (2004). Insulin amyloid fibrillation at above 100 c: new insights into protein folding under extreme temperatures. *Biochemistry*, **42**, 11347–11355.
43. Jansen, R., Grudzielanek, S., Dzwolak, W. & Winter, R. (2004). High pressure promotes circularly shaped insulin amyloid. *J. Mol. Biol.* **338**, 203–206.
44. Grudzielanek, S., Jansen, R. & Winter, R. (2005). Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. *J. Mol. Biol.* **351**, 879–894.
45. Dzwolak, W., Ravindra, R. & Winter, R. (2004). Hydration and structure—the two sides of the insulin aggregation process. *Phys. Chem. Chem. Phys.* **6**, 1938–1943.
46. Dzwolak, W., Jansen, R., Smirnovas, V., Loksztejn, A., Porowski, S. & Winter, R. (2005). Template-controlled conformational patterns of insulin fibrillar self-assembly reflect history of solvation of the amyloid nuclei. *Phys. Chem. Chem. Phys.* **7**, 1349–1351.
47. Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N. & Fink, A. L. (2001). Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry*, **40**, 8397–8409.
48. Nielsen, L., Frokjaer, S., Carpenter, J. F. & Brange, J. (2001). Studies of the structure of insulin fibrils by fourier transform infrared (ftir) spectroscopy and electron microscopy. *J. Pharm. Sci.* **90**, 29–37.
49. Garriques, L. N., Frokjaer, S., Carpenter, J. F. & Brange, J. (2002). The effect of mutations on the structure of insulin fibrils studied by fourier transform infrared (ftir) spectroscopy and electron microscopy. *J. Pharm. Sci.* **91**, 2473–2480.
50. Jimenez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M. & Saibil, H. R. (2000). The protofilament structure of insulin amyloid fibrils. *Protein Sci.* **9**, 1960–1967.
51. Krebs, M. R. H., MacPhee, C. E., Miller, A. F., Dunlop, I. E., Dobson, C. M. & Donald, A. M. (2004). The formation of spherulites by amyloid fibrils of bovine insulin. *Proc. Natl Acad. Sci. USA*, **101**, 14420–14424.
52. Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N. & Fink, A. L. (2003). Partially folded intermediates in insulin fibrillation. *Biochemistry*, **42**, 11404–11416.
53. Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N. & Fink, A. L. (2004). Stimulation of insulin fibrillation by urea-induced intermediates. *J. Biol. Chem.* **279**, 14999–15013.
54. Berne, B. J. & Pecora, R. (1976). *Dynamic Light Scattering*, Wiley, Interscience, New York, NY.
55. Levine, H. (1999). Quantification of b-sheet amyloid structures with thioflavin t. *Methods Enzymol.* **309**, 256–274.
56. Krebs, M. R. H., Bromley, E. H. C. & Donald, A. M. (2005). The binding of thioflavin-t to amyloid fibrils: localisation and implications. *J. Struct. Biol.* **149**, 30–37.
57. Naiki, H. & Geiyo, F. (1999). Kinetic analysis of amyloid fibril formation. *Methods Enzymol.* **309**, 305–319.

58. Oosawa, F. & Asakura, S. (1975). *Thermodynamics of the Polymerization of Proteins,* Academic Press, London.
59. Flyvbjerg, H., Jobs, E. & Leibler, S. (1996). Kinetics of self-assembling microtubules: an "inverse problem" in biochemistry. *Proc. Natl Acad. Sci. USA*, **93**, 5975–5979.
60. Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15**, 1–89.
61. Lomakin, A., Teplow, D. B., Kirschner, D. A. & Benedek, G. B. (1997). Kinetic theory of fibrillogenesis of amyloid β-protein. *Proc. Natl Acad. Sci. USA*, **94**, 7942–7947.
62. Brange, J. (1987). *Galenics of Insulin: The Physico-Chemical and Pharmaceutical Aspects of Insulin and Insulin Preparations,* Springer-Verlag, Berlin.
63. Doty, P. & Myers, G. E. (1953). Ii. low molecular weight proteins. Thermodynamics of the association of insulin molecules. *Discuss. Faraday Soc.* **13**, 51–58.
64. Nettleton, E. J., Tito, P., Sunde, M., Bouchard, M., Dobson, C. M. & Robinson, C. V. (2000). Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectroscopy. *Biophys. J.* **79**, 1053–1065.
65. Krapivsky, P. L. & Redner, S. (1996). Transition aggregation kinetics in dry and damp environments. *Phys. Rev. E*, **54**, 3553–3561.
66. Kusumoto, Y., Lomakin, A., Teplow, D. B. & Benedek, G. B. (1998). Temperature dependence of amyloid β-protein fibrillization. *Proc. Natl Acad. Sci. USA*, **95**, 12277–12282.
67. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2: 1. Evidence for a two-state transition. *Biochemistry*, **30**, 100435–104281.
68. Stepanek, P. (1993). Dynamic light scattering: the method and some application. In *Chapter Data analysis in Dynamic Light Scattering* (Brown, W., ed), pp. 177–240, Clarendon Press, Oxford.
69. Ferri, F. (1997). Use of a charge coupled device camera for low-angle elastic light scattering. *Rev. Sci. Instrum.* **68**, 2265–2274.
70. Bulone, D., Giacomazza, D., Martorana, V., Newman, J. & San Biagio, P. L. (2004). Ordering of agarose near the macroscopic gelation point. *Phys. Rev. E*, **69:0411401**, 1–9.
71. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G. *et al.* (1988). The structure of 2zn pig insulin crystals at 1.5 a resolution. *Philos. Trans. R. Soc. Lond., B*, **319**, 369–456.
72. Kadima, W., Ogendal, L., Bauer, R., Kaarsholm, N., Brodersen, K., Hansen, J. F. & Porting, P. (1993). The influence of ionic strength and pH on the aggregation properties of zinc-free insulin studied by static and dynamic laser light scattering. *Biopolymers*, **33**, 1643–1657.
73. Nicolai, T., Durand, D. & Gimel, J.-C. (1996). Light scattering: principles and development. In *Chapter Scattering Properties and Modeling of Aggregating and Gelling Systems* (Brown, W., ed), pp. 201–231, Clarendon Press, Oxford.
74. de Gennes, P.-G. (1979). *Scaling Concepts in Polymer Physics,* Cornell University Press, Ithaca and London.
75. Goodman, J. W. (2000). *Statistical Optics,* Wiley, New York.
76. Manno, M., Bulone, D., Martorana, V. & San Biagio, P. L. (2004). Ergodic to non-ergodic transition monitored by scattered light intensity statistics. *Physica, A*, **341**, 40–54.
77. Pusey, P. N. & van Megen, W. (1989). Dynamic light scattering by non-ergodic media. *Physica, A*, **157**, 705–741.

Edited by J. Weissman

(Received 29 April 2006; received in revised form 4 August 2006; accepted 2 November 2006)
Available online 9 November 2006