Biophysical Journal Volume 90 June 2006 4585–4591

Kinetics of Insulin Aggregation: Disentanglement of Amyloid Fibrillation from Large-Size Cluster Formation

Mauro Manno, Emanuela Fabiola Craparo, Vincenzo Martorana, Donatella Bulone, and Pier Luigi San Biagio
Institute of Biophysics at Palermo, Italian National Research Council, Palermo, Italy

ABSTRACT Kinetics of human insulin aggregation has been studied at pH 1.6 and 60°C, when amyloid fibrils are formed. We developed a novel approach based on the analysis of scattered light intensity distribution, which allows distinguishing between small and large size aggregates. By this method, we observed an exponential growth of fibrillar aggregates implying a heterogeneous aggregation mechanism. Also, the apparent lag time observed, correlated with the major increase of thioflavin T fluorescence, has been assigned to the onset of large size cluster formation.

INTRODUCTION

Proteins or peptides can assemble into different morphologies, which have an important role in technological applications, or in organism physiology and pathology. A large class of proteins can form elongated insoluble aggregates, known as amyloid fibrils (1), which are involved in the clinical manifestations or in the etiology of many diseases (2). Insulin is known to form amyloid fibrils when heated at low pH or in low-concentration solutions (3–5). The knowledge of the mechanisms of insulin fibrillization is important to improve pharmaceutical formulations used in the treatment of diabetic patients (5), as well as to shed light into the basic molecular features of amyloid formation. However, the study of the aggregation mechanisms is often made difficult by the formation of a complex hierarchy of supramolecular assemblies (6), from single fibrils (7–9) to macroscopic floccules of fibrils (10).

In this work, we have studied the kinetics of insulin aggregation at 60°C and at pH 1.6, starting from a 210 μM solution of human insulin. In such conditions, insulin forms aggregates of different sizes (10,11). In Fig. 1, typical amyloid fibrils can be observed along with large size aggregates, formed after 9 h of incubation (9). To study the onset of the formation of small size fibrils, we have developed a novel approach based on the analysis of scattered light intensity distribution. By this method, the growth of small size fibrils has been distinguished from that of large size clusters.

The intensity signal related to amyloid fibrils has been found to increase exponentially after heating. Exponential growth occurs when the aggregation rate is proportional to the mass of aggregates, and it is a typical signature of a heterogeneous aggregation mechanism (12)—that is, an aggregation mechanism including both tip-to-tip elongation and lateral fiber growth. The same mechanism was already proposed in the early studies of insulin fibrillization (4,13).

In addition, our results show that the apparent lag time observed in the aggregation process is mainly due to large size aggregates. Thus, they give a warning against classical methods used for detecting fibril formation, such as thioflavin T binding.

METHODS

Sample preparation

Recombinant human insulin powder (purchased from Sigma-Aldrich, St. Louis, MO, and used without further purification) was directly dissolved at 5°C in buffer solution (50 mM KCl/HCl in SuperQ water, pH 1.6 at 60°C; Millipore, Bedford, MA). The solution was ultrafiltered 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 μm Millex-GV; Millipore) into a quartz cuvette. Insulin concentration was 210 μM, as measured by UV absorption at 276 nm (extinction coefficient: 1.0675 cm^{-1} $(mg/ml)^{-1}$). All chemicals were reagent grade.

Large-angle light scattering

After preparation, samples were incubated at 60°C and scattered light intensity at 90° was measured by using a model No. BI-9000 correlator equipped with an He-Ne laser (Brookhaven Instruments, Holtsville, NY), $\lambda_0 = 632.8$ nm. The scattered intensity I is given in terms of Rayleigh ratio $R = I(q)/I_s r^2/V_s$, 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) were obtained by normalization with respect to toluene, whose Rayleigh ratio at 632.8 nm was taken as 14.02×10^{-6} cm^{-1}.

Optical microscopy

At the end of scattering experiments, a few drops of solution were put on a glass for imaging with a phase-contrast Axioskop2 Plus microscope (Carl Zeiss, Oberkochen, Germany).

Submitted November 11, 2005, and accepted for publication March 6, 2006.

Address reprint requests to Pier Luigi San Biagio, Institute of Biophysics at Palermo, Italian National Research Council, via U. La Malfa 153, Palermo, I90146, Italy. Tel.: 39-091-680-9311; E-mail: pierluigi.sanbiagio@pa.ibf.cnr.it.

Fabiola Craparo's present address is Dipartimento di Chimica e Tecnologie Farmaceutiche, Università di Palermo, via Archirafi 32, Palermo, I90133, Italy.

0006-3495/06/06/4585/07 $2.00

doi: 10.1529/biophysj.105.077636

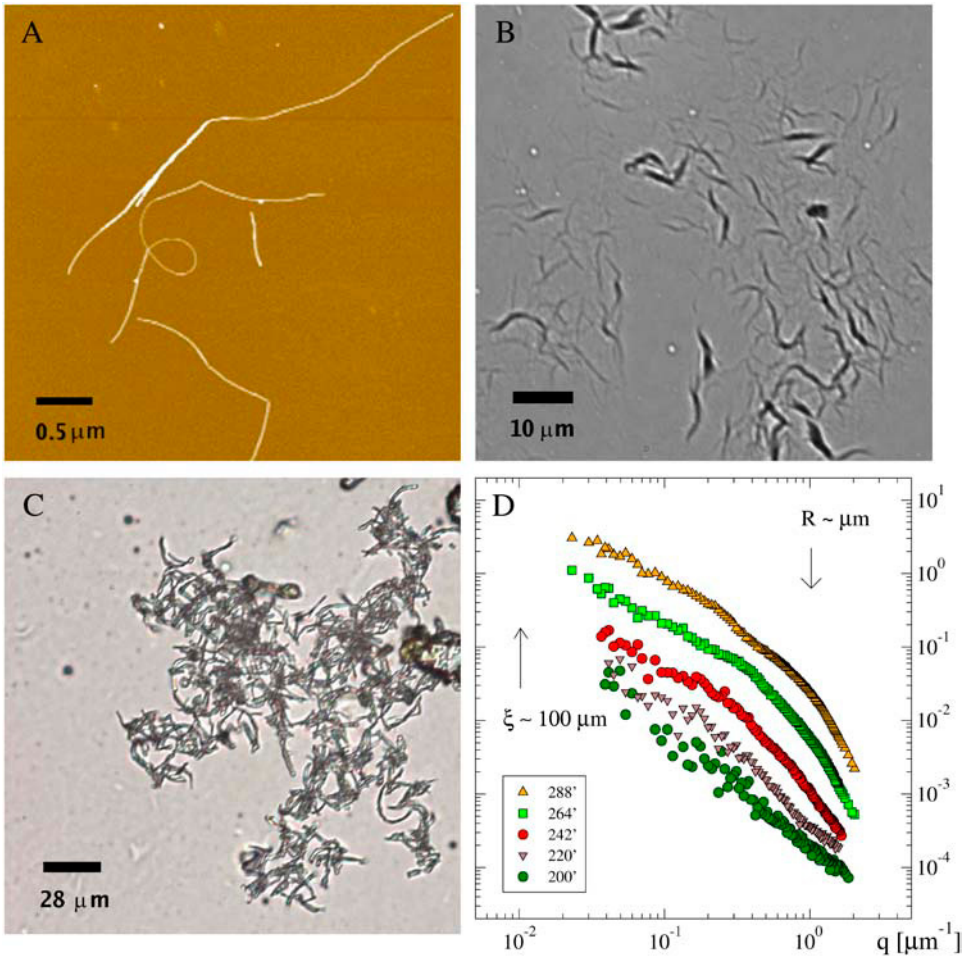

FIGURE 1 Insulin aggregates after 9 h of incubation at 60°C, pH 1.6 (insulin concentration 210 μM). (*A*) Atomic force microscopy image of amyloid fibrils (redrawn from (9)). (*B*,*C*) Phase-contrast microscopy image of large-size clusters. (*D*) Structure functions obtained by small-angle light scattering.

Thioflavin T fluorescence spectroscopy

Thioflavin T was added to the insulin solutions directly into the quartz cuvette (1-cm optical path) to reach the final concentration of 20 mM. Fluorescence emission spectra, from 470 to 560 nm, were recorded after quenching to 60°C with excitation at 450 nm and scan-speed of 100 nm/min by using a model No. FP-6500 spectrofluorometer (JASCO Europe, Cremella, Italy, Great Dunmow, Essex, UK). The signal was obtained from the thioflavin T fluorescence intensity at 482 nm, by subtraction of a background blank measurement. Fluorescence experiments were performed simultaneously with scattering experiments, by using the same sample preparation.

Small-angle light scattering

Small-angle light scattering measurements were performed using an experimental device built in our laboratory, and based on previous work (14,15). The device was equipped with a 30-mW helium-neon laser (Melles Griot, Carlsbad, CA) and a model No. TM765 charge-coupled device camera (PULNIX America, Sunnyvale, CA). Measurable intensities span a wide dynamical range (more than four orders of magnitude), thanks to a software integration of multiple exposure times (1/60 ÷ 1/10,000 s). After background subtraction (performed by homemade software), reliable measurements could be made over a range of scattering angles of 0.1° ÷ 11°, corresponding to scattering vector magnitudes of 0.02 ÷ 2 μm^{-1}, with a resolution of 0.01°. Small angle light scattering experiments were performed simultaneously with large angle light scattering experiments, by using the same sample preparation.

THEORY

Analysis of scattered intensity distributions

The 90° scattered intensity recorded during the aggregation kinetics yields a noisy signal due to the appearance of high molecular weight aggregates. The intensity contributions due respectively to low mass and high mass objects have been identified by the present approach, based on the analysis of scattered intensity distribution.

Instantaneous scattered intensity $I = |E(t)|^2$ is exponentially distributed with average $\langle I \rangle$, since the electric field $E(t)$ scattered by a large number of objects is a complex Gaussian variable (16). However, in scattering experiments the total measured intensity $I_{\delta T}$ is time-averaged over a small time-window δT, which is usually a fraction of second: $I_{\delta T} = \delta T^{-1} \int_{\delta T} |E(t)|^2 dt$.

When the coherence time of the scattered field τ is much larger than δT, the measured intensity is equivalent to the instantaneous intensity, and therefore exponentially distributed. On the other hand, when the field coherence time τ is short compared to the interval δT, we can partition the integration time in $N = \delta T/\tau$ subintervals of length τ. The intensity related to each subinterval is commonly approxi-

mated to a constant and considered not correlated with that of other subintervals (16,17); thus, time-averaged intensity $I_{\delta T}$ can be taken as the sum of N statistically independent contributions with the same exponential distribution function. The resulting distribution function $P(I_{\delta T})$ is the γ-distribution function (16), with average $\langle I\rangle$ and cumulants κ_n: $\kappa_n = (n-1)!\langle I\rangle^n/N^{n-1}$. For $N = 1$, $P(I_{\delta T})$ is equal to the exponential distribution function as for the instantaneous intensity, while for $N \to \infty$, a δ-distribution function is obtained. For N sufficiently large, only the first two cumulants—that is, average and variance—are appreciably non-null, and a Gaussian distribution is obtained, as prescribed by the central limit theorem (18).

In a solution of freely diffusing aggregates, the coherence time τ is essentially given by the diffusional properties of scatterers of size R_h (hydrodynamic radius) and diffusion coefficient D, given by the Stokes-Einstein relation $\tau = (Dq^2)^{-1} = (k_BTq^2)^{-1}6\pi\eta R$, where η is the medium viscosity, T is the temperature, k_B is the Boltzmann constant, and q is the scattering vector (19). More properly, the equivalence between diffusional coherence time and aggregate size is unequivocal when dealing with compact (spherelike), not interacting Brownian objects. In such cases, the hydrodynamic radius and the radius of gyration have close values, and the Stokes-Einstein relation holds. However, in a solution of supramolecular aggregates, as in the present work, the separation of large and small size is inevitably blurred by other factors such as aggregate shape and interactions. In the experimental condition of the present work ($T = 60°C$, $q = 18.7\ \mu m^{-1}$, $\eta = 0.466 \times 10^{-3}$ Pa s), we have $\tau = \alpha R$ with $\alpha = 5.5 \times 10^{-3}$ s/μm, and $\delta T = 0.9$ s. Thus, for an object of size 150 μm one obtains $\delta T/\tau \approx 1$ and, hence, a scattered intensity distribution that is exponential. On the other hand, a smaller object yields a γ-distribution, which can be univocally approximated as a Gaussian distribution when $\delta T/\tau \gg 1$, that is, when $R_h \ll 100\ \mu$m. As a rule of thumb, we can say that when $R_h \leq 3\ \mu$m, one obtains $N = \delta T/\tau > 50$, and the γ-distribution can be univocally approximated with a Gaussian distribution. 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, whereas the appearance of non-Gaussian noise is related to the formation of large size, slowly diffusing aggregates, such as clusters of fibrils or floccules.

To distinguish between the contributions of large and small aggregates (or more correctly slow and fast aggregates), we assume that the scattered field E is the sum of two terms:

1. A fast fluctuating field E_F due to small size objects with an average intensity $\langle I_F\rangle$, and a coherence time $\tau_F < \delta T$
$\langle I_F\rangle$.
2. A slow decaying field E_S due to large size, slowly diffusing objects, with an average intensity $\langle I_S\rangle$, and a coherence time $\tau_S > \delta T$. The total time-integrated intensity is given by the sum of these two terms: $I_{\delta T} = \int_{\delta T} |E_F + E_S|^2 dt$.

We have recently shown (20) that the cumulants of the distribution function of $I_{\delta T}$ can be written, to leading order $1/N$, as

$$\begin{aligned}\kappa_1 &= \langle I_S\rangle + \langle I_F\rangle\\ \kappa_2 &= \langle I_S\rangle^2 + \frac{4}{N}\langle I_S\rangle\langle I_F\rangle + \frac{1}{N}\langle I_F\rangle^2 + O\left(\frac{1}{N^2}\right)\\ \kappa_{n\geq 3} &= (n-1)!\left\{\langle I_S\rangle^n + \frac{2n}{N}\langle I_S\rangle^{n-1}\langle I_F\rangle + O\left(\frac{1}{N^2}\right)\right\},\end{aligned} \quad (1)$$

where $O(1/N^2)$ are the terms of order $1/N^2$ or higher. In general, the meaning of this approximation is to consider only second-order temporal coherence, or more precisely, the first-order in $1/N = \tau_F/\delta T$ and the zero-order in $\delta T/\tau_S$.

Since in the case of a sum of two statistically independent processes, the sum of their cumulants is equal to the cumulants of the sum (18), we derived from Eq. 1 that the integrated intensity $I_{\delta T}$ can be interpreted as the sum of two statistically independent contributions (20): an exponentially distributed slow contribution with average

$$\langle I_E\rangle = \langle I_S\rangle + \frac{2}{N}\langle I_F\rangle, \quad (2)$$

and a Gaussian-distributed fast contribution with average and variance

$$\begin{aligned}\langle I_G\rangle &= \langle I_F\rangle - \frac{2}{N}\langle I_F\rangle\\ \sigma^2 &= \frac{\langle I_G\rangle^2}{N}.\end{aligned} \quad (3)$$

The expression for the intensity distribution function is thus the convolution of a Gaussian with an exponential function,

$$P(I_{\delta T}) = \langle I_E\rangle^{-1}\exp\{-[I_{\delta T} - \langle I_G\rangle]/\langle I_E\rangle\} \quad \Theta(I_{\delta T}), \quad (4)$$

where

$$\begin{aligned}\Theta(I_{\delta T}) = \Bigg[&\mathrm{erf}\left(\frac{I_{\delta T}}{\sigma\sqrt{2}} - \frac{\langle I_G\rangle\langle I_E\rangle + \sigma^2}{\langle I_E\rangle\sigma\sqrt{2}}\right)\\ &+ \mathrm{erf}\left(\frac{\langle I_G\rangle\langle I_E\rangle + \sigma^2}{\langle I_E\rangle\sigma\sqrt{2}}\right)\Bigg]\frac{1}{2}e^{\frac{\sigma^2}{2\langle I_E\rangle^2}},\end{aligned} \quad (5)$$

with $erf(x)$ being the error function $\mathrm{erf}(x) = 2/\sqrt{\pi}\int_0^x e^{-t^2}dt$.

We remark that the tails of the Gaussian distribution must be short enough to avoid negative values for the intensity. This is actually implicit in our approximation, since $\tau_F/\delta T$ is assumed to be small.

It is worth noting that to apply this method of analysis, the experimental duration should be sufficiently long to allow an efficient sampling of slow diffusing objects. In the present work we are not concerned with the limiting case in which the size of aggregates becomes huge, so that they span the

entire sample, and gelation occurs. In this case, the probability distribution function of scattered intensity is not correctly restored by a conventional experimental implementation, due to sample nonergodicity. Nonergodic systems such as gels or glasses are characterized by a long coherence time (21). In such cases, when structural relaxations are longer than the experiment duration, a correct ensemble average of scattered intensity should be obtained by spatial averaging over different positions of the arrested sample (20,21).

RESULTS

Kinetics of insulin aggregation

We have studied the aggregation kinetics of recombinant human insulin heated in acidic solution (pH 1.6 at 60°C), by light scattering and thioflavin T fluorescence. In such conditions insulin is known to assemble into a variety of supramolecular aggregates, such as amyloid fibrils (3,7,9) or floccules (4,10). At the end of kinetics, aggregates of different sizes of up to hundreds of microns can be observed, as shown in the examples of Fig. 1, *A–C*.

Here we report the detailed analysis of a single kinetics experiment. However, the same experiments have been performed a few times to check reproducibility. To obtain reliable and reproducible results, it is very important that the sample is completely dissolved, and that no spurious aggregates are left in the solution, since they may act as seeds for a templated aggregation (22). The absence of seeds has been carefully verified at the beginning of each experiment.

The scattered intensity signal (namely the Rayleigh ratio $R(q)$) is related to the weight-average molecular mass M_w of aggregates through the relation $R = 4\pi^2\tilde{n}^2(d\tilde{n}/dc)^2 \lambda_0^{-4} N_A^{-1} c M_w P(q)$, 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 (19). At the onset of kinetics, we obtained $M_w = 10 \pm 1$ kDa, by taking $(d\tilde{n}/dc) = 0.18$ cm^3 g^{-1}, (23) and $P(q) = 1$ (since the initial size of solutes is much smaller than q^{-1}).

This corresponds to a mean aggregation number of 1.7 ± 0.2, by taking the molecular mass of insulin as 5806 Da. After ~180 min, the growth of intensity yields a noisy signal. We have calculated the distribution of scattered intensity over a small time interval of ~5 min (Fig. 2). The chosen time-interval should be large enough to collect a statistically meaningful amount of data; at the same time, it should be sufficiently small to avoid dramatic changes in the average quantities. In the studied kinetics, a time interval of 5 min was the best compromise to study the onset of the signal growth. The distributions are mainly Gaussian in the early stages of kinetics, and they develop a marked skewedness as aggregation proceeds.

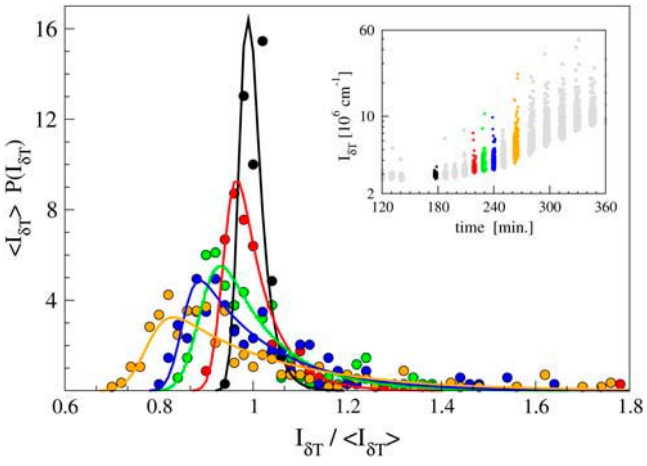

FIGURE 2 Scattered intensity distributions at different times of kinetics. Data are fit to Eq. 4 (*solid lines*). The inset shows the temporal evolution of 90° scattered intensity. Time intervals and distributions are related by matching colors.

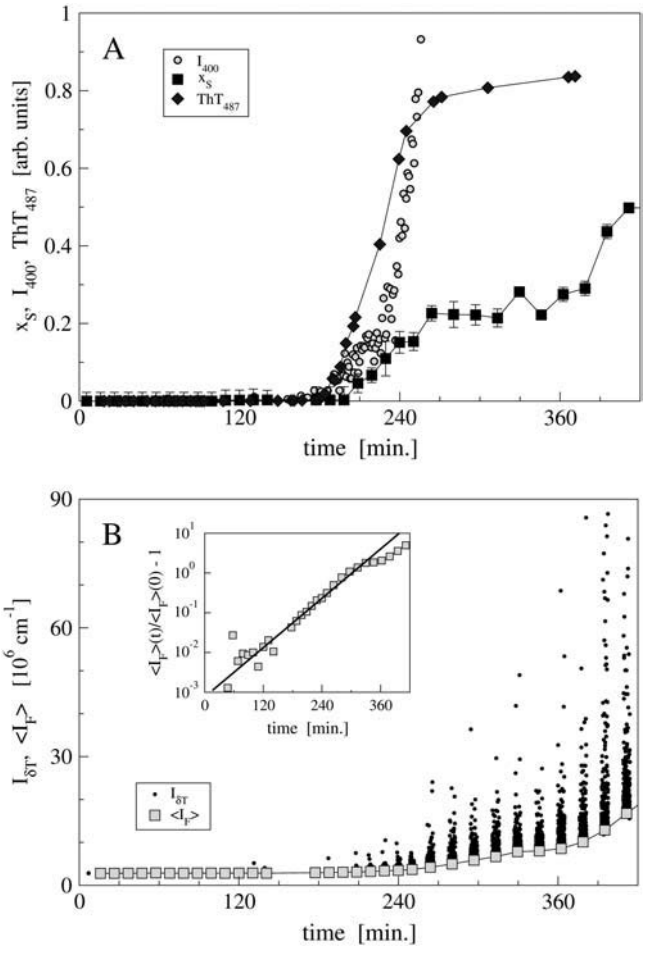

FIGURE 3 Kinetics of insulin aggregation at 60°C, pH 1.6. (*A*) Thioflavin T fluorescence at 487 nm (*diamonds*); small-angle scattered light intensity at 400 cm^{-1} (*circles*); slow fraction of 90° scattered intensity (*squares*). (*B*) 90° scattered intensity (*dots*); fast contribution to 90° scattered intensity (*squares*). (*Inset*) Growth of fast contribution to the intensity with respect to the initial signal. Data are fit to an exponential function (*solid line*).

The onset and the growth of non-Gaussian scattered intensity are accompanied by the growth of a structure function revealed by small angle light scattering (Fig. 1 *D*). The structure functions $S(q)$ exhibit a quasi-linear slope in the low q-range (approximately from 0.02 to 0.2 μm^{-1}), which corresponds to lengths of the orders of tens of microns. The slope changes at scattering vectors q of the order of microns and it reaches a high value in the submicron range. The latter features of the structure functions evidence the existence of compact objects with a characteristic length in the micron range. Also, the low q-range, where no distinct characteristic lengths are observed, evidences the existence of structures with a size larger than tens of microns, that is, on the order of hundreds of microns. Therefore, large-size aggregates, such as those directly observed by microscopic imaging at the end of the kinetics (Fig. 1 *C*), are present in solution at least since the onset of non-Gaussian large angle light scattering. Note, however, that the images collected at the end of kinetics are not exhaustive of the possible aggregate morphologies. The presence of large-size aggregates at the earlier stages cannot be stated by the small-angle light scattering technique due to the low experimental sensitivity.

The kinetics of aggregation has been also monitored by thioflavin T fluorescence. Thioflavin T is a classical marker for amyloid fibrils (24), since it fluoresces upon binding to fibrils. It has been used for both bovine (25) and human (11) insulin fibrillation. Our results show that the major increase of thioflavin fluorescence occurs concurrently with the onset of non-Gaussian scattered intensity and the appearance of small-angle structure functions, as shown in Fig. 3 *A*.

To filter out the contribution of large aggregates, we have applied the method described above. We have assumed that the scattered field is the sum of two terms: a fast fluctuating, Gaussian contribution (mainly due to small-size objects), and a slow decaying, exponentially distributed contribution (mainly due to large-size slowly diffusing objects). By fitting intensity distribution to Eqs. 2–5, we have obtained the slow fraction of intensity $x_S = \langle I_S \rangle / \langle I_{\delta T} \rangle$, that is, the fraction of intensity contributed by large objects, and the fast Rayleigh ratio $\langle I_F \rangle$ due to the formation of small aggregates, which are known to be amyloid fibrils (5,7,9,11,26). These data are displayed in Fig. 3 along with the fluorescence of thioflavin T and the small-angle light scattering at small q.

The growth of the slow fraction x_S marks the formation of large size aggregates (Fig. 2), such as those observed in Fig. 1, *B* and *C*, as also confirmed by the simultaneous increase of small-angle light scattering. It is remarkable that the apparent lag time of 180 min (Fig. 3), measured by the growth of thioflavin T fluorescence and the onset of non-Gaussian noise of scattered intensity, are both related to the formation of large clusters, and not only to the fibrillogenesis process. Thioflavin T is a dye commonly used to indicate fibril formation (25). However, its affinity to fibrils and its selectivity should be carefully checked for each specific system (27), or perhaps accompanied by an experimental technique sensitive to masses or size of aggregates, such as the light scattering technique used in the present case. Even though the increase of thioflavin T fluorescence marks an increase in the amount of fibrillar structures, the present results warn against the effects that large size aggregates may have on the correct labeling of fibrils.

The mechanism of insulin fibrillogenesis can be elicited by the kinetics of the fast Rayleigh ratio $\langle I_F \rangle$. In this quantity, the noisy effect of slow and large objects have been filtered out (Fig. 3 *B*), and we can extend the analysis over a wider time window, across the lag time. The fibrillation kinetics is of course cluttered at the very early stages, due to the low experimental sensitivity, and at the late stages, when it is overwhelmed by the effect of large-size clusters, or eventually by entanglement or sedimentation. For example, a partial precipitation of larger aggregates may be responsible for the anomalous decrease of the growth rate of the scattered intensity after 300 min (Fig. 3 *B*). In the inset of Fig. 3 *B*, the fast scattered intensity is shown with respect to its initial value, $\langle I_F \rangle(t) - \langle I_F \rangle(0)$, to highlight the intensity contribution due to the supramolecular aggregates, which emerges out of the initial scattering due to a monomeric-oligomeric solution ($\langle I_F \rangle(0)$). We show that starting from a solution of insulin oligomers, containing mainly monomers and dimers, fibril formation (related to $\langle I_F \rangle$) proceeds with an exponential growth and a rate of $4 \times 10^{-4}\ s^{-1}$. This is a clear symptom of a heterogeneous aggregation, i.e., having a rate proportional to the mass (12). These results confirm the first model proposed for insulin fibrillation, which includes both tip elongation and lateral thickening (4,13), as well as more recent findings on bovine insulin (28).

The slowly diffusing scatterers, which have been revealed by small-angle light scattering, contribute to the slow fraction. Indeed, as explained above, the diffusional coherence time of such aggregates exceeds the chosen value for δT. However, the exact size and coherence time of the fast aggregates is hindered by many factors. For instance, the hydrodynamic radius is strongly affected by the shape of aggregates. In particular, the translational diffusion coefficient of elongated fibril-like structures is related to an effective hydrodynamic radius that is much smaller than the radius of gyration (29,30). Also, large-size interacting aggregates may undergo structural relaxations not related solely to Brownian diffusion, so that $\tau > (Dq^2)^{-1}$. Therefore, the parameter δT does not provide a clearcut partition between large and small size objects. In this respect, our method faces the same drawbacks of dynamic light scattering, when it is used to measure a distribution of molecular sizes. Indeed, the actual reason for choosing such a high value of δT is not that the corresponding hydrodynamic radius represents a significant size in the system. The aim is to filter out the noise due to slow (large) objects as much as possible, without throwing out the baby with the bath water, that is without washing out any physically meaningful signal.

Thus, although δT (and namely its ratio to the coherence time) is a key parameter, it cannot be finely tuned to make a

sharp separation of aggregate sizes. It is worth noting that the partition of the scattered electric field in two contributions is univocally defined only when there are two types of objects, having two different timescales τ_F and τ_S, so that $\tau_F \ll \tau_S$. In such cases, a trivial choice for δT is $\tau_F \ll \delta T \ll \tau_S$. Thus, the distribution of scattered intensity can be approximated to the zero-order in both $\tau_F/\delta T$ and $\delta T/\tau_S$, and Eq. 5 reduces to a step function (20): $\Theta(I_{\delta T})$ is zero for $I_{\delta T} < \langle I_E \rangle$ and one otherwise. This approach was early developed by Pusey and van Megen for strong chemical gels (21). The rationale for our more extended approach is that the two timescales are not completely separated. This allows us to apply the method to aggregation kinetics and to systems with a complex hierarchy of supramolecular structures. Moreover, we remind the reader that objects in the range between a few microns and hundreds of microns ($\sim 3 \div 150$ microns, as we noted in the previous section) can give a contribution to both the small and the fast parts of the intensity distribution. Their distribution is neither purely Gaussian nor purely exponential. In other words, the statistical filter which is set up by the present approach has an high efficiency for sizes bigger than hundreds of microns, which are filtered out, and for sizes smaller than few microns, which are preserved, although it is not sharp in the intermediate range.

Another typical source of a non-Gaussian signal in scattering experiments may be the fluctuations of particle numbers (31). However, to appreciate this effect the number of objects in the scattering volume must be of the order of a few units. Therefore, number fluctuations can give a contribution comparable to that of the initial oligomeric solution only if they come from large-size objects. Moreover, the correlation time τ_N related to number fluctuations of diffusive objects is very large, since it is roughly proportional to the time necessary to diffuse through the scattering volume, whose typical size is slightly less than a few millimeters. Therefore, in our experimental conditions the ratio between the correlation time τ_N and the particle size R is on the order of 10 s/nm, so that it would be difficult to observe such fluctuations in 5 min. This correlation time would be much shorter in the case of a nondiffusive flux. The latter could be due to either thermal convection or sedimentation. We can reasonably exclude any thermal convection in our sample, since the sample holder in large-angle scattering experiments is sunk into the index-matching liquid, which is efficiently thermostated by recirculating water. This allows a homogeneous heating of the sample and the absence of thermal gradients. As for sedimentation, only very large size objects would be involved in such a process, if any. Thus, the non-Gaussian signal would add to the slow part of scattering. It would be interesting to extend our theory to explicitly include this effect as well (even if an experimental implementation might be unreliable). However, in our experiments sedimentation is likely to be a second-order effect. Indeed, protein aggregates have a density quite close to that of sample buffer and they are also likely not to be fully packed. The present experiments, e.g., small-angle light scattering, suggests that large-size aggregates may be a loose net of entangled linear (fibrillar) objects. Further, sedimentation is easily detected by light scattering, through the loss of scattering signal. As noted above, it is likely that at time 300 min, the anomalous change in the slope of scattering growth may be due to aggregate precipitation.

DISCUSSION

In the present work, we have studied the aggregation of human insulin, which brings forth the formation of amyloid fibrils and large-size clusters (Fig. 1) (7,9,10). We developed a method based on the analysis of scattered intensity distribution, which allows us to filter out the scattering contributions due to large-size objects.

Another powerful experimental technique suited to measure sizes of diffusing aggregates is dynamic light scattering, which has been successfully applied to amyloid formation (30,32). However, this technique requires relatively long integration times and high signal intensities, so that it is not appropriate for fast kinetics. In this work, for example, the scattered intensity signal is very low at the onset of kinetics, and it grows rapidly after the lag time. The statistical filter that we have developed and applied to protein aggregation kinetics has of course a poorer resolution with respect to dynamic light scattering technique. However, it can be easily implemented on simple static light scattering measurement, and it is applicable to fast kinetics. Further, the technique developed is essentially based on the filtration of non-Gaussian noise. In this regard, it is quite general, and could be applied to many other processes as well as different kinds of molecular aggregation.

By applying this method, we singled out the formation of insulin fibrils, and evidenced the formation of large-size clusters (Fig. 2). Our results show an apparent lag time in the aggregation kinetics, as monitored by different techniques (Fig. 3 *A*):

1. Thioflavin T fluorescence, which is specifically suited to reveal amyloid structures.
2. Small-angle light scattering, which is particularly applicable to large-size structures.
3. Non-Gaussian scattered intensity distribution, especially for the fraction of slow scattering, which has been identified in our approach and is directly related to slow (and hence somehow large) objects.

During the lag time, the presence of fibrils or larger aggregates is not directly observed, due to the low experimental sensitivity of such techniques. Also, the same experimental conditions have been explored in a previous work by time-lapse atomic force microscopy (9), without observing any fibrillar aggregate during the lag time.

The aggregation mechanism has been elicited by the filtered scattered light intensity (Fig. 3 *B*). The intensity signal corresponds initially to a solution of small insulin oligomers (mainly monomers and dimers), and then starts growing

exponentially. The statistical approach has allowed us to observe such growth even before the apparent lag time of ~180 min, and to assign a correct value to the aggregation rate. The exponential growth of the signal is obtained when the rate of the signal increase is proportional to the signal itself (12). The scattering signal is related to the mass average distribution, and more precisely to the second moment of the mass aggregate distribution. Therefore, these findings confirm the first model for insulin fibrillation, which was proposed by Waugh (3,4,13), and recently supported by thioflavin T fluorescence experiments (28): Insulin fibrillation proceeds not only by tip-to-tip addition of monomers to fibrils, but also by lateral assembly of monomers or larger protofibrils into fibril surface (heterogeneous nucleation/aggregation).

In summary, we have developed a novel approach to treat the scattered intensity distribution during an aggregation kinetics. By applying this technique to the formation of insulin aggregates in amyloid-forming conditions, we observed a heterogeneous aggregation of fibrils (Fig. 3), as proposed in the early model of insulin fibrillation (4). The implications of these findings on elucidating the mechanisms of insulin fibrillization will be exploited in a forthcoming article. Also, we have been able to ascribe the apparent lag time observed in light scattering and thioflavin T fluorescence kinetics to large cluster formation. This stresses that, in the studies of protein aggregation, it is very important to use experimental techniques, like scattering or microscopies, which are sensitive to masses and sizes of aggregates.

We thank A. Podestà, G. Tiana, and P. Milani for collaboration and access to unpublished data. This work was partially supported by the Italian Ministero della Salute through the project ''Animal Neuropathies: Molecular and Functional Analysis of Prion Protein in Sicilian Bovine Species''.

REFERENCES

1. Stefani, M., and C. M. Dobson. 2003. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* 81:678–699.
2. Kelly, J. W. 1998. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* 8:101–106.
3. Waugh, D. F. 1946. A fibrous modification of insulin. I. The heat precipitate of insulin. *J. Am. Chem. Soc.* 68:247–250.
4. Waugh, D. F., D. F. Wilhemson, S. L. Commerford, and M. L. Sackler. 1953. Studies of the nucleation and growth reactions of selected types of insulin fibrils. *J. Am. Chem. Soc.* 75:2592–2600.
5. Brange, J., L. Andersen, E. D. Laursen, G. Meyn, and E. Rasmussen. 1997. Towards understanding insulin fibrillation. *J. Pharm. Sci.* 86: 517–525.
6. Khurana, R., C. Ionescu-Zanetti, M. Pope, J. Li, L. Nielson, M. Ramirez-Alvarado, L. Regan, A. L. Fink, and S. A. Carter. 2003. A general model for amyloid fibril assembly based on morphological studies using atomic force microscopy. *Biophys. J.* 85:1135–1144.
7. Jimenez, J. L., E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil. 2000. The protofilament structure of insulin amyloid fibrils. *Protein Sci.* 9:1960–1967.
8. Jansen, R., W. Dzwolak, and R. Winter. 2004. Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy. *Biophys. J.* 88:1344–1353.
9. Podestà, A., G. Tiana, P. Milani, and M. Manno. 2006. Early events in insulin fibrillization studied by time-lapse atomic force microscopy. *Biophys. J.* 90:589–597.
10. Krebs, M. R. H., C. E. MacPhee, A. F. Miller, I. E. Dunlop, C. M. Dobson, and A. M. Donald. 2004. The formation of spherulites by amyloid fibrils of bovine insulin. *Proc. Natl. Acad. Sci. USA.* 101: 14420–14424.
11. Hua, Q.-X., and M. A. Weiss. 2004. Mechanism of insulin fibrillation: the structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate. *J. Biol. Chem.* 279:21449–21460.
12. Eaton, W. A., and J. Hofrichter. 1990. Sickle cell hemoglobin polymerization. *Adv. Protein Chem.* 40:63–279.
13. Waugh, D. F. 1957. A mechanism for the formation of fibrils from protein molecules. *J. Cell. Comp. Physiol.* 49:145–164.
14. Ferri, F. 1997. Use of a charge-coupled device camera for low-angle elastic light scattering. *Rev. Sci. Instrum.* 68:2265–2274.
15. Bulone, D., D. Giacomazza, V. Martorana, J. Newman, and P. L. San Biagio. 2004. Ordering of agarose near the macroscopic gelation point. *Phys. Rev. E.* 69:0411401-1–0411401-9.
16. Goodman, J. W. 2000. Statistical Optics. Wiley, New York.
17. Rice, S. O. 1945. Mathematical analysis of random noise. *Bell Sys. Tech. J.* 23–24:1–162.
18. Van Kampen, N. G. 1992. Stochastic Processes in Physics and Chemistry. North-Holland, Amsterdam.
19. Berne, B. J., and R. Pecora. 1976. Dynamic Light Scattering. Wiley Interscience, New York.
20. Manno, M., D. Bulone, V. Martorana, and P. L. San Biagio. 2004. Ergodic to non-ergodic transition monitored by scattered light intensity statistics. *Physica A.* 341:40–54.
21. Pusey, P. N., and W. van Megen. 1989. Dynamic light scattering by non-ergodic media. *Physica A.* 157:705–741.
22. Harper, J. D., and P. T. J. Lansbury. 1997. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* 66:385–407.
23. Doty, P., and G. E. Myers. 1953. II. Low molecular weight proteins. thermodynamics of the association of insulin molecules. *Discuss. Faraday Soc.* 13:51–58.
24. Levine, H. 1999. Quantification of β-sheet amyloid structures with thioflavin T. *Methods Enzymol.* 309:256–274.
25. Nielsen, L., R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V. N. Uversky, and A. L. Fink. 2001. Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry.* 40:6036–6046.
26. Grudzielanek, S., R. Jansen, and R. Winter. 2005. Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. *J. Mol. Biol.* 351:879–894.
27. Naiki, H., and F. Geiyo. 1999. Kinetic analysis of amyloid fibril formation. *Methods Enzymol.* 309:305–319.
28. Librizzi, F., and C. Rischel. 2005. The kinetic behavior of insulin fibrillation is determined by heterogeneous nucleation pathways. *Protein Sci.* 14:3129–3134.
29. Russo, P. S. 1993. Dynamic light scattering from rigid and nearly rigid rods. *In* Dynamic Light Scattering: The Method and Some Applications. W. Brown, editor. Clarendon Press, Oxford. 512–533.
30. Carrotta, R., M. Manno, D. Bulone, V. Martorana, and P. L. San Biagio. 2005. Protofibril formation of amyloid β-protein at low pH via a non-cooperative elongation mechanism. *J. Biol. Chem.* 280:30001–30008.
31. Pusey, P. N. 1977. Photon Correlation Spectroscopy and Velocimetry. H. Z. Cummings and E.R. Pike, editors. Plenum Press, New York and London.
32. Lomakin, A., D. S. Chung, G. B. Benedek, D. A. Kirschner, and D. B. Teplow. 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.