PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Intermediates caught in the act: Tracing insulin amyloid fibril formation in time by combined optical spectroscopy, light scattering, mass spectrometry and microscopy

A. Gladytz,[1] E. Lugovoy,[1,2] A. Charvat,[1,2] T. Häupl,[2] K. R. Siefermann,[1] and B. Abel[1,2]*

1) Leibniz-Institute of Surface Modification (IOM), Permoserstr. 15, 04318 Leipzig, Germany

2) Wilhelm-Ostwald-Institut für Physikalische und Theoretische Chemie, Universität Leipzig, Linné-Strasse 2, 04103 Leipzig, Germany

ABSTRACT

Interest in the topic of amyloid formation by peptides and proteins has increased dramatically in recent years, transforming it from a puzzling phenomenon associated with a small number of diseases into a major subject of study in disciplines ranging from material science to biology and medicine. The tendency of numerous (also non-pathogenic) proteins such as insulin to self-assemble into amyloid-like fibrils is well known. While fibrils are usually easily detected, the observation of transient intermediates is a big challenge in general. They are the key and the 'holy grail' for a molecular understanding of mechanisms in this context. Here we show that intermediates, i.e. oligomers, can be detected and their hydrodynamic radius R_H as well as their overall conformation and structure can be monitored and the aggregation dynamics as well as structure formation can be detected in time with a suitable combination of experimental techniques. We have observed transient intermediates that resemble large oligomers held together in solution by non-covalent forces. The oligomers appear to convert into building blocks for mature fibrils with largely beta-sheet conformation resembling key players in a mechanism, which is termed 'nucleated conformation conversion' in the literature. Structural transformations of oligomers in time towards dominant beta-sheet conformations have been observed for the first time. The structures can even be observed in liquid phase AFM experiments. With this approach we have successfully shed new light into the aggregation and fibrilization process of insulin being possibly a model system for other amyloid systems.

*) for correspondence: bernd.abel@iom-leipzig.de

INTRODUCTION

Neurodegenerative diseases such as Alzheimer's, Parkinson's and the transmissible spongiform encephalopathies (TSEs) are characterized by abnormal protein deposits in the form of highly symmetric amyloid fibrils.[1,2] Amyloid fibrils are polypeptide wire-like assemblies in which the polypeptide backbone is arranged in a specific and characteristic cross-β-sheet quaternary structure. However, the molecular mechanism(s) of fibril formation are still not fully understood and are currently the subject of intense research. In particular, it is not clear if various different mechanisms exist nor what the intermediates are. There is evidence that not the full-grown fibrils are the dangerous species but smaller sub-fibrillar oligomers that likely cause the disease via a toxic gain-of-function phenotype.[3]

In the past couple of years there have been reports about amyloid formation of non-pathogenic proteins not assigned to any disease.[3] This led to the conclusion that amyloid formation is no specific property of only some few proteins and polypeptides, namely those that have been described to assemble into clinical amyloid deposits. Instead, is has been found that many different systems are readily able to form amyloid fibrils in a process which is accelerated by autocatalysis. [3]

The physiological and therapeutic importance of insulin, and the fact that insulin aggregates in vitro under destabilizing conditions (heat, low pH), have made the hormone an excellent model for protein aggregation studies.[4-9] As early as 1940 Langmuir and Waugh[10,11] observed that, in acidic solutions, insulin can be modified to form fibrils, and a few years later Waugh proposed that a nearly simultaneous interaction of three or four insulin monomers is necessary to forms a nucleus.[12] Beyond being a model system for amyloid systems,[13] insulin aggregation is known to cause injection amyloidosis[14] in vivo and therefore received some attention.

Under physiological conditions, insulin is a hexamer,[15] being able to bind two or four Zn^{2+} ions. Decreasing pH and protein concentration shift the equilibrium toward smaller oligomers. The insulin monomer consists of two chains: a 21-residue-long A-chain and a 30-residue-long B-chain. The chains are connected by two disulfide bridges: A7–B7 and A20–B19, whereas a third bridge binds cysteine residues 6 and 11 of the A-chain. The insulin aggregation involves only non-covalent interactions, and the disulfide bridges are not broken during the aggregation process. As a consequence, the conserved

A6–A11 SS-bond is expected to cause a significant topological constraint and bending of the A-chain. The insulin aggregation is promoted under conditions favoring partly destabilized monomers and dimers, such as low pH, high temperature, and to some extent contact with hydrophobic media.[16] Several results suggest the importance of particular amino acid residues for insulin aggregation; for example, the bovine insulin, differing from the human protein at three amino acid positions (A8 Ala^{Bov}/Thr^{Hum}, A10 Val^{Bov}/ Ile^{Hum}, and B30 Ala^{Bov}/Thr^{Hu}), aggregates more easily in vitro than the latter.[17] Studies on insulin derivatives established that the two B-chain's termini might play distinct roles in the aggregation. Elsewhere it was shown that insulin fibrils' morphologies and the corresponding infrared spectra vary depending on conditions of their growth.[18] Insulin fibrils display characteristic features common to all amyloids: x-ray diffraction patterns; affinity to Congo Red and Thioflavin T; birefringence; biological dysfunction and high stability against temperature; pressure; low pH; and proteases. The infrared spectra of the amyloid suggest a parallel arrangement of the β-strands, and this appears to be in agreement with a three dimensional model proposed for insulin amyloid fibrils.[7] Considerable debate exists as to the mechanism of fibril assembly in general and for insulin in particular as well as intermediate species and their structures.[8]

Currently, there are three 'classical' popular mechanistic models[3] describing amyloid fibril formation. The first is referred to as template assembly (TA), wherein a preassembled nucleus binds a soluble state peptide in a random coil conformation in a rapid pre-equilibrium step. This is followed by a rate-determining structure change to add the peptide to the growing end of the fibril or filament, presumably as part of the β-sheet-rich quaternary structure. The second model is the monomer directed conversion (MDC) model, which implies that a monomeric peptide can adopt a conformation called the A state that is analogous to the conformation adopted in the fibril. In the rate-determining step this structured monomer (A) then binds and converts an S state monomer, resulting in an A state dimer. The dimer then dissociates and the structured monomers rapidly add to the end of the growing fibril. A third model, the often invoked nucleated polymerization (NP) mechanism, is characterized by the rate limiting formation of a nucleus resulting from an equilibrium between monomers that are and are not assembly competent. Once a nucleus is established, assembly occurs by the addition of assembly competent monomers to the growing end of the fibril (i.e., the nucleus).

Limitations of the three putative mechanisms of amyloid fibril assembly for explaining Sup35 prion formation have led to the introduction of an additional model by Serio *et al.*[19] – featuring 'nucleated conformational conversion' of loosely bound oligomers, which then (after conformational conversion into β-sheets) serve as building blocks for fibril assembly. This model we will simply refer to as the 'oligomer conversion model' (OCM) in the following. Eisenberg has pointed out that one of the main mechanisms and driving forces for amyloid fibril formations is the steric zipper motif, which appears to be the model with the most predictive power and largest molecular insight at present.[20]

Early work connected amyloid fibrils with dementia,[21] but recent studies have shifted attention to small oligomeric species as possible toxic agents responsible for the disease.[22] There is no consensus for the definition of these entities, especially when it relates to their size, but in general it is assumed that small to medium sized and loosely bound oligomers/aggregates are soluble, which then, via a nucleation cascade, assemble into insoluble amyloid fibrils. Detailed structural analysis of these oligomeric species is still missing, mainly because of the experimental challenges associated with identification, quantification, isolation and characterization of individual components in a dynamic multi-component mixture. The literature on this topic is relatively sparse in general. For insulin the investigation of oligomers with mass spectrometry,[23] electrospray differential mobility analysis,[24] nuclear magnetic resonance,[25] dynamic light scattering (DLS),[26], [27] small-angle neutron scattering (SANS)[28] and x-ray scattering (SAXS) are described.[29] Also microscopy approaches, including cryogenic transmission electron microscopy (TEM) [30] and atomic force microscopy (AFM) are reported.[17,31,32] Remarkable is also a series of high-resolution separation techniques for fractionating and recovering oligomer components, including size exclusion chromatography, electrophoresis, capillary electrophoresis and membrane filtration.[33, 42]

The goal of the present paper is to characterize possible oligomers and intermediates in the insulin aggregation and fibril formation process – which may be relevant for other amyloid systems as well.

EXPERIMENTAL SECTION

Aggregation assays: Bovine insulin was purchased from Sigma–Aldrich (Product No. I5500) and was used without further purification. Hydrochloric and sulfuric acid was purchased from Merck. Solvents

were prepared by diluting the necessary amount of sulfuric acid (unless otherwise noted sulfuric acid was employed in the experiments) with double distilled water to give the required pH and were then filtered through a syringe filter (0.45 mm pore size). Insulin (5.8 mg) was dissolved in the solvent of the required pH value to give the final solution concentration. The pH values of the solvent and the final solution were verified to be within 0.1 of the desired value with a pH meter (Hanna Instruments). Aliquots (1 mL) were transferred into polypropylene reaction vessels (1.5 mL) and shaken at 200 rpm and 65ºC in a block heater. Samples (50 µL) were taken from the solution and stored at 25°C for further analysis, e.g., mass spectrometric analysis.

Liquid beam laser desorption mass spectrometry (MS): The MS-based method for measuring the bio-kinetics of insulin in the present work was time-resolved liquid beam desorption mass spectrometry. For technical details we refer to earlier publications.[34-36] The samples (50µl) were introduced into the liquid water jet in vacuum [35-38] by using an HPLC injection valve (*Rheodyne MX9925*). For the laser desorption of protonated insulin from the liquid jet the idler wave of a $LiNbO_3$ optical parametric oscillator (OPO, GWU) with a wavelength of 2800 nm was employed. The OPO was pumped by the fundamental of a Nd:YAG laser (Spectra Physics, Quanta Ray INDI). Each injection resulted in 80 usable single shot spectra at 20 Hz repetition rate. The median average of the monomer peak intensities was calculated for further data analysis. The obtained traces usually were averaged in order to obtain a satisfactory signal-to-noise ratio.

Atomic force microscopy: The solution containing the aggregation products at predefined times (see results and discussion part) was diluted by a factor of ten, 5–10 µL of which were either spin coated to dryness on a freshly cleaved mica surface or a droplet was investigated directly on mica. A Digital Instruments (VEECO) MultiMode scanning probe microscope IIIa was used. Samples were imaged in air (dry samples or droplet samples) at a 1 Hz scan rate by using silicon tips.

X-ray microscopy: X-ray microscopy has been performed at the 3rd generation synchrotron source at BESSY in Berlin (U41 - TXM).[39] The full-field x-ray microscope (λ=2.4nm) features an ultimate resolution in the range of 30-40nm. The instrument was installed at undulator U41 and was developed specifically for samples in a gas or condensed phase (aqueous samples) environment. In this setup two vacuum windows supported by pinholes separated the condenser and the x-ray objective chambers.

The distance of these windows was typically 300µm to avoid a too strong absorption of the soft x-rays by air.

Dynamic light scattering (DLS): For the dynamic light scattering (DLS) measurements a Viscotek Model 802DLS™ equipped with a single mode fiber optic detection modul and dual attenuation technology was employed. The instrument uses a 544 channel multi-tau correlator and a Single Photon Counting Module/detector, as well as a 50mW fiber coupled diode laser operating at 830nm. The 802DLS™ was operated with the OmniSIZE™ software. A small volume of the samples of the kinetic experiments was measured in a 12µl cuvette, which is introduced into the sample compartment and placed in a temperature controlled cell holder. A complete DLS size distribution analysis could be accomplished in 5 minutes. The size range that could be measured was between 0.5nm and 10µm hydrodynamic radius.

Circular dichroism spectroscopy (CD): The aggregation studies have been performed with a *JASCO J-810A spectral photometer*, equipped with a *JASCO PTC423S temperature controlling unit.* The sample compartment was continuously flushed with dry nitrogen gas (3 - 5 L/min). The temperature was measured directly in the thermostat block close to the cuvette. Characteristic conditions for all measurements were 5 averaged scans, at 1.00 nm *bandwidth*, 1.00 s *response,* and a scan speed of 100 nm/min. In all experiments quartz glass cuvettes were employed. (l = 1.00 cm).

RESULTS AND DISCUSSION

In order to measure the concentration of free insulin as a function of time in aqueous solution under denaturation conditions we have employed liquid beam mass spectrometry for time resolved investigations. An advantage of the present technique over previous approaches using mass spectrometry may be the linear response of the liquid beam desorption MS technique, [40] that is the linear relationship between the concentration of a substance in solution and its mass spectrometry detector signal in the gas phase. This feature facilitates kinetic measurements and - in principle - the determination of rate constants. It is well known that insulin aggregation can be initiated by heat and a low pH. Therefore, aqueous solutions of insulin at pH 2 in the concentration range roughly between 10^{-5}M and $5{\cdot}10^{-4}$M were weakly stirred/shaken at a temperature of 65°C. Samples of the solution were

injected in the liquid beam source and analyzed in the TOF mass spectrometer as a function of time. A typical experiment ($c=1x10^{-4}M$) is displayed in Fig. 1. As can be easily seen, the parent insulin mass peak at t=0 decreases within the course of the experiment until it vanishes for longer times. The typical sigmoidal trace is normalized (C/Co) and the time for the relative intensity of C/Co = 0.5 is referred to as the "lag-time" in the following (Fig. 1). This lag time is characteristic for the kinetics of the system, i.e., the aggregation process of insulin, which is viewed here through the loss of the monomer.

The kinetic traces for the protein aggregation have been measured for different concentrations (range between 10^{-5} - $5{\cdot}10^{-4}$ mol/l). The concentration dependence of the lag-time displayed in Fig. 2 under our experimental conditions is only moderate (within a factor of 2 in the concentration rage). A somewhat similar approach using ESI-MS was reported recently by Robinson and co-workers[23]. The agreement between the two data sets is fair. Although, the features of the kinetic traces are similar the concentration dependence (in particular for lowest concentrations) is not as pronounced as reported in Ref. [23], which we attribute to the slightly different experimental conditions (stirring vs. non stirring of the sample). We will come back to the data of Nettleton et al. below. In order to show that the monomer decay is actually correlated with fibril formation we have imaged the fibrils after about 2h with AFM. As can be clearly seem in Fig. 3 the monomer decay in the experiments ultimately leads to characteristic fibrils – which are expected for the present conditions.

In the course of our work we have also investigated the aggregation of insulin in time with x-ray microscopy (λ=2.4nm). Aggregation has been initiated in a weakly stirred/shaken external reaction tube ($1x10^{-4}$M, pH=2, HCl), and aggregation was measured by taking samples and measuring them off-line at different times. We had hoped to identify some intermediate structures as well as the formation of fibrils. The results are displayed in Fig. 4. During the lag phase of the aggregation no characteristic structures can be observed, which may be due to the resolution or the contrast at the wavelength employed. At intermediate times broad structures are visible and at longer times - significantly beyond the lag time τ - fine but extended brushes of fibrils are clearly visible (Fig. 4). Unfortunately, the exposure time was limited to several tens of seconds, due to radiation damage of

the fibrils. These data again show that the observed monomer decay in Fig. 1 is directly correlated with the formation of fibrils.

It is proposed that smaller and larger oligomers may be intermediates in the aggregation and fibrillation process of amyloid systems – also for insulin. In principle, intermediates should be detectable with soft ionization mass spectrometry – such as the laser induced liquid beam desorption mass spectrometry employed here. As displayed in Fig. 5 we could detect multimers of insulin in the initial phase of the experiment at t=0 in small abundance ($t<\tau$), which however, simply vanish for longer times together with the monomers. Dimers and multimers up to the tetramer and possibly the pentamer are visible in the mass spectra ($Co=10^{-4}M$). It is the question, however, if these complexes are real intermediates or oligomers, which we aim to detect in such an experiment. It is known that soft mass spectrometry is maybe not soft enough to detect larger weakly bound species (insulin complexes or aggregates) – the detected ones could be simply unspecific complexes formed in the MS process or just fragments of larger intermediate oligomers. In any case our experiments point towards a very low abundance of such complexes – which may also be expected for transient species from a kinetic point of view. However, this may be a first indication that oligomer structures may play a role in the aggregation process of insulin – an issue that is highly debated and which needs more attention from an experimental point of view through a strategy that employs a series of complementary techniques.

In a subsequent study we used light scattering for the detection of intermediates in the course of the aggregation and fibril formation process. We anticipated that the chance of observing intermediates (oligomers) would be significant because larger aggregates are easier to detect than monomers due to the higher scattering cross sections and the largest fibrils precipitate from the solution. For this purpose we employed dynamic light scattering (DLS). Dynamic light scattering (DLS) is a laser scattering technique capable of unbiased analysis of size distributions of diffusing particles in the nanometer to micron size range. When coherent light rays are scattered from Brownian particles, interference causes temporal fluctuations in the detected intensity. The timescale of these temporal fluctuations correlate with diffusion time constants from which particle size is inferred. The ability to resolve multimodal size distributions and make absolute size measurements makes DLS a powerful

technique for systems with heterogeneous species. DLS has been used to study fibril formation in the past (e.g., Ref. [27]).

The fibrillation process has been monitored (off-line) as a function of time such as displayed in Fig. 6 ($C=1\cdot10^{-4}$ M, pH=2, 65ºC, weakly stirred). For early times the monomer is clearly visible. After about 45min the distribution of hydrodynamic radii changes. The monomer is hardly visible anymore. A new peak at 2.34 arises which may be due to the denatured insulin or a dimer. At times beyond 45min the distribution of hydrodynamic radii changes significantly and becomes broader and inhomogeneous. From 60-80min on fibrillar structures or real fibrils may be visible in the range above 1μm. The present data leave little doubt that transient oligomers are present during the aggregation process and play a role in the fibrilization of insulin. It appears also clear that these oligomers are on-pathway species on the way to the fibrils.

In order to literally shed more light into the structures of Fig. 6 during the kinetics we have employed time-resolved circular dichroism (CD) spectroscopy to obtain structural information as a function of time (experimental conditions as for the DLS measurements). It is known that insulin has to convert from an overall rich helical to dominant beta-sheet structure during the fibrillation - which has to be present before insulin can assemble into fibrils. This transition should be detectable in a CD-spectrum. In order to study this we employed the same aggregation protocol as outlined before and measured CD spectra during the aggregation process. The results are shown in Fig. 7. During the initial phase of the experiment two characteristic overlapping bands at 208nm and 222nm are observed, which indicate significant and typical contributions from α-helical secondary structures of insulin – as it is expected. The spectrum does not change significantly below the lag time τ. After 50 minutes the band structure and its intensity changes towards a single band with a maximum at 215-216nm. This spectral signature is characteristic for dominant a β-sheet structure in insulin at this stage of the aggregation process. If we compare the data with the scattering data in Fig. 6 we learn that between 45min and 60min a structural transition is visible – which is color coded in Fig. 6 (green). This may be considered to be direct proof for the fact that at 45min the structures belong to larger oligomers but which still do not have a dominant or overall beta-sheet structure. In the picture of the nucleated conformation conversion or the oligomer conversion models the oligomers at 45min are not converted yet while the

oligomers at or beyond 60 min appear to be converted (green scattering spectrum in Fig. 6). This is a very valuable and exciting observation and insight, not only into the appearance of oligomers but also into their transient structure and conformation. To the best of our knowledge this is the first time, at least in the case of bovine insulin (or similar), where the presence of oligomer distributions in combination with conformational and dynamical information has been measured at the same time, i.e., a transition between overall conformations of oligomers.

In an attempt to actually capture the "intermediate structures" with nanoscale resolution - for which we have obtained strong evidence from mass spectrometry, DLS, CD-spectra and light scattering experiments, we employed AFM in the liquid water phase (droplet) with *in situ* atomic force microscopy. Characteristic diffuse and "cloudy" structures at $t \approx \tau$ (i.e., t≈50-60min) are visible with AFM in water solution, which are not full grown fibrils and not monomers either, which appear to assemble into fibrils and which look like building blocks of mature fibrils. They are displayed in Fig. 8a. According to the CD-spectra the cloudy structures in Fig. 8a correspond to beta-sheet rich 'on-pathway' oligomers in the fibrillation process of insulin. Shown is a 450nm×450nm section of the liquid sample on a mica plate (off-line measurement). The diffuse structures are on the order of 50nm in diameter, see the circle in Fig. 8a. In the lower part of Fig. 8a diffuse spherical structures or nuclei appear to be arranged in a row, which are able to form fibrils. Again, we interpret these structures as intermediates in the fibrilization that are structurally (at least in part) converted and which are building blocks to assemble proto-fibrils and in turn mature fibrils with extended over-structures. At reaction times beyond the lag-time τ (i.e., $t > 80$min) short fibrils with a characteristic screw turn (see also the inset of Fig. 8b) are detected in the liquid aqueous phase (Fig. 8b).

We interpret the data to be evidence for the presence of a nucleated conformation conversion (NCC) or oligomer conversion mechanism. Yet, insulin appears to be another example besides the Sup35 system that displays an NCC aggregation mechanism[19] in liquid water at pH=2 and 65ºC. The driving force for an aggregation of a water soluble protein is destabilization and hydrophobic forces. The insulin molecules can stabilize themselves under native state destabilizing conditions by exposing hydrophobic sites of the molecule not to water but towards other hydrophobic sites of other molecules and thus excluding water in a way that the process is also entropy driven. These aggregates probably

convert and display a more or less rapid nucleated (oligomer) conformation conversion (towards beta-sheet). The picture that emerges from the present studies is that of an 'oligomer conversion model' (OCM) depicted in Fig. 9a). Loosely bound clusters (B) of insulin monomers (A) form and structurally convert towards beta-sheet rich clusters/aggregates (C') at some point in the course of 'reaction' towards fibrils These structurally converted aggregates (C') with largely beta-sheet internal structure can be direct building blocks of larger protofibrillar structures (displayed in Fig. 9b). They may also combine with and convert (C'->C'') other not yet fully converted aggregates/clusters (B) and assemble in a way like one of the first models proposed by M. Eigen[41] for single molecules (monomer driven conversion model) – here adapted for molecular aggregates. The converted aggregates (C') may in turn even add and convert monomers (E-F) – which resembles the template assembly models described above. This holds as well for the larger structures in Fig. 9b. Although, the present experimental results favor a fast nucleation step and a conversion of conformation to be the main mechanism of aggregation we cannot rule out stochastic (step-by-step) nucleation events[43], which may be present and not detectable due to sensitivity reasons but we believe that they are not contributing much to the main mechanism.

Finally, the question may arise why many of the fibrils (for many pathogenic and model systems) are similar in size – at least in thickness. Taken the hypothesis that fibrils are assembled from converted spherical aggregates as building blocks/units towards protofibrils and mature fibrils with over structures it is the question why these aggregates/clusters do not grow much beyond 100nm or even above 1µm and in turn form much thicker primary fibrils? All fibrils found for a number of amyloidic systems are usually below 100nm – other even significantly below that value (10nm) – even if they assemble in overstructures. What could possibly limit the size of the aggregates/clusters in the aggregation process in the first place? In order to explain this we invoke a simple nucleation model such as depicted in Fig. 10. Initially, larger aggregates/clusters are formed from smaller ones through aggregation of monomers. For smaller aggregates the probability of dissociation (into a monomer and an (n-1)-aggregate) is larger than that of monomer addition – due to the surface energy. A nucleated conformation conversion is governed by a more or less concerted isomerization of the molecular aggregate. Since a concerted conformation change is quite unlikely in general, this process can hardly

compete with dissociation of monomers at smaller size of the aggregates – which is fast in general. As the aggregates grow the surface energy decreases and the molecular/aggregate degrees of freedom increase, the dissociation of aggregates into monomers and fragments becomes less likely (a simple argument from reaction kinetics). In any case we are talking here about loosely bound aggregates with a near native protein structure. When the dissociation step slows down the isomerization step is more competitive (in terms of reaction rates k_{diss} and k_{iso}) and possibly taking over at a certain size range of the aggregate. No matter how fast it is, this step is certainly irreversible as opposed to the dissociation step. We believe that this process and ultimately thermodynamics and kinetics limit the growth and the size of the aggregates and the building blocks of the protofibrils and ultimately the mature fibrils. At a certain size they just isomerize and are out of the business. This may explain why a number of proteins with different amino acid residues and structure assemble fibrils of similar thickness. The length of the fibrils depends crucially on the aggregation protocol and in particular if the solution is stirred.

The experimental findings appear to document that insulin aggregation and fibrilization is actually governed by intermediates, i.e. oligomers. The size of the oligomers appears to be much beyond that of dimers or smaller multimers. In table 1 we have collected and compared techniques in the literature for the identification and characterization of oligomers and intermediates in insulin fibrilization. In the following we briefly want to compare our results with findings in the literature on the insulin system with respect to amyloid fibril formation and its mechanisms. As stated above intermediates are intrinsically difficult to detect, most likely due to their low concentration and stability. They are just not stable intermediates that can be captured easily.

In 2000, for the first time, Nettleton et al. [23] took advantage of the soft electrospray method in combination with MS to provide a quantitative characterization of bovine insulin oligomers assembly in vitro under both native conditions and conditions promoting fibril growth. Not surprisingly, at pH=4 and in the presence of Zn^{2+} ions, insulin was predominantly a hexamer. At pH 2, and at about 65°C oligomers up to the 12-mer were observed at highest concentrations without the presence of Zn^{2+}. Under acidic conditions and elevated temperature (favorable aggregation conditions) the formation of oligomers could be followed in real time. Qualitatively, the experiments by Nettleton et al [23] are close to the results obtained here. The aggregation protocol is slightly different, which may

explain the different concentration dependence. The fact that they observe larger aggregates than detected in this study may reflect the softness and differences of the laser desorption/ionization and electrospray ionization (LDI and ESI). However, we believe that *both* techniques most likely do *not* capture the true oligomer size (such as measured here with AFM and DLS) because in the MS experiments the non-covalently and weakly bound complexes may still fragment in the source, such as it is known for other weakly bound non-covalently bound complexes. On the other hand non-specific complexes are easily formed in the source, which may have nothing to do with oligomers in solution. It is interesting to note here that a different technique called electrospray differential mobility analysis also did only monitor oligomers up to the trimer,[24] possibly due to the same reason and problems outlined above. Unfortunately, in NMR studies no oligomers could be detected.[25] Concomitant with the MS work, researchers have tried to use DLS to investigate first the size distribution of native insulin assemblies, and later the development of larger aggregates under fibrillation conditions.[26,27] The latter distribution was quite broad and indicated the presence of a heterogeneous population of oligomers, which was difficult to capture. Moreover, the experiments suggested a relatively low-abundance population of smaller oligomers present in addition to larger oligomers. The DLS experiments in the literature (see table 1) appear to support our present findings – at least qualitatively, although, some caution is necessary because the outcome of an aggregation experiment depends crucially upon the reaction conditions, i.e. the aggregation protocol.

In small-angle neutron scattering (SANS) experiments[28] researchers found for similar experimental conditions as in our study that dimers were dominant at the start of the aggregation and they then converted to hexamers and ultimately to fibrils. Small angle x-ray scattering experiments[29] under similar aggregation conditions also found only smaller oligomers such as for electrophoresis[33] experiments. While in AFM experiments[31] researchers found evidence for a broad oligomer distribution for pH=2 and 65°C conditions, cryogenic transmission electron microscopy ‚observed' oligomers up to 15-20nm in diameter.[30] Heldt and co-workers[42] used membrane filtration, with a 100kDa molecular mass cut-off membrane, and successfully isolated an oligomeric structure (about 50nm measured by AFM) with interesting properties. This would be a nice confirmation of the results in this study. Unfortunately, the study did not reveal whether these oligomers pre-exist or were created

during fibril dissolution or assembled as intermediates in the aggregation process. So, taken together all literature data we came to the conclusion that the overall picture is still somewhat diffuse – with our present results it just begins to clear up somewhat. Part of the problems with different results on the same model system is the fact that experiments are difficult to control because they very much depend upon the experimental conditions and artifacts are easily produced if one does not have full control on the experimental parameters. Impurities also play a big role for the results obtained in an experiment. While there is in general agreement on the product, i.e. the structure of the fibrils, the intermediates of amyloid system aggregation and fibril formation still represent the ‚holy grail' in the understanding of amyloid system's kinetics and mechanisms. The separation, fractionation, isolation, and characterization of insulin oligomers from a mixture of different conformations and sizes are still very critical challenges. These oligomers are unstable and species far from equilibrium.

Summary and Conclusions

In summary, we highlighted different techniques applied to measure insulin aggregation and fibrilization and intermediate oligomers under fibril-forming conditions. We have successfully shed new light into the aggregation and fibrilization process of insulin being a model system for other amyloid systems that have a relatively stable native structure. We have observed intermediates that resemble large oligomers held together in solution by non-covalent forces. The oligomers (consisting of loosely bound monomers) appear to convert in these aggregates after some time resembling an oligomer conversion mechanism, which is termed 'nucleated conformation conversion' in the literature. The transient structures can be even observed in AFM experiments. The structures are likely on pathway structures because their appearance correlates with the decay of monomers and the formation of fibrils, which are also identified in the present experiments. Future experiments and theory may provide more evidence that this scenario is also present and relevant for other systems beyond insulin.

Finally, it became very clear that only a combined analytical approach can decipher the molecular mechanisms of amyloid aggregation and fibril formation. There is hardly any technique that can resolve the problem alone. A combination of methods should also be considered to leverage the

advantages of each technique and to validate the results of one technology, such as done in this study. Identification, isolation and purification of transient oligomer species have been very difficult tasks in the past due to their very low abundance and strong inherent reactivity.[28] The present work demonstrates that at least identification and kinetic tracing of oligomeric intermediates may be possible with a combined spectroscopic and imaging experimental approach – providing important input for theoretical understanding and modeling.

Acknowledgments

Discussions of several aspects of the present work with colleagues within the Sonderforschungsbereich SFB-TR102 (Halle-Leipzig, B1) are gratefully acknowledged. We thank Drs. A. Bögehold and J. Thieme for assistance and help with the MS- and dynamic light scattering experiments and the X-ray imaging experiments at BESSY. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) within the Transregio SFB-TR102 (Project B1).

References

(1) Dobson, C. M. Protein folding and misfolding. *Nature* **2003**, *426*, 884.

(2) Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **1982**, *216*, 136-144.

(3) Kelly, J. W. Mechanisms of amyloidogenesis. *Nat. Struct. Biol.* **2000**, *7*, 824-826.

(4) Ahmad, A.; Millett, I. S.; Doniach, S.; Uversky, V. N.; Fink, A. L. Partially folded intermediates in insulin fibrillation. *Biochemistry* **2003**, *42*, 11404-11416.

(5) Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V. N.; Fink, A. L. Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **2001**, *40*, 8397-8409.

(6) Whittingham, J. L.; Scott, D. J.; Chance, K.; Wilson, A.; Finch, J.; Brange, J.; Dodson, G. G. Insulin at pH 2: structural alalysis of the conditions promoting insulin fibre formation. *J. Mol. Biol.* **2002**, *318*, 479-490.

(7) Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R. The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci.*

2002, *99*, 9196-9201.

(8) Brange, J.; Andersen, L.; Laursen, E.; Mayn, G.; Rasmussen, E. Toward understanding insulin fibrillation. *J. Pharmac. Sci.* **1997**, *86*, 517.

(9) Hua, Q.-X.; Weiss, M. A. Mechanism of insulin fibrillation. *J. Biol. Chem.* **2004**, *279*, 21449-21460.

(10) Langmuir, I.; Waugh, D. F. Pressure-soluble and pressure-displaceable components of monolayers of native and denatured proteins. *J. Am. Chem. Soc.* **1940**, *62*, 2771-2793.

(11) Waugh, D. F. A fibrous modification of insilin. I. The heat precipitate of insulin. *J. Am. chem. Soc.* **1946**, *68*, 247-250.

(12) Waugh, D. F. A mechanism for the formation of fibrils from protein molecules. *J. Cell Physiol.* **1957**, *49*, 145-164.

(13) Groenning, M.; Frokjaer, S.; Vestergaard, B. Formation mechsnism of insulin fibrils and structural aspacts of the insulin fibrillation process. *Curr. Protein Peptide Sci.* **2009**, *10*, 509-528.

(14) Dische, F. E.; Wernstedt, C.; Westermark, G. T.; Westermark, P.; Pepys, M. B.; Rennie, J. A. Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia* **1988**, *31*, 158-161.

(15) Chang, X.; Jorgensen, A. M.; Bardrum, P.; Led, J. J. Solution structures of the R6 human insulin hexamer. *Biochemistry* **1997**, *36*, 9409-9422.

(16) Sluzky, V.; Tamada, J. A.; Llibanov, A. M.; Langer, R. Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. *Proc. Natl. Acad. Sci.* **1991**, *88*, 9377-9381.

(17) Jansen, R.; Dzwolak, W.; Winter, R. Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy. *Biophys. J.* **2005**, *88*, 1344-1353.

(18) Ohnishi, S.; Takano, K. Amyloid fibrils from the viewpoint of protein folding. *Cell. Mol. Life Sci.* **2004**, *61*, 511-524.

(19) Serio, T. R.; Cashikar, A. G.; Kowal, A. S.; Sawicki, G. J.; Moslehi, J. J.; Serpell, L. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **2000**, *289*, 1317-1321.

(20) Eisenberg, D.; Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Sambashivan, S.; Ivanova, M. I.; Madsen, A.; Riekel, C. The structural biology of protein aggregation diseases: Fundamental questions and some answers. *Acc. Chem. Res.* **2006**, *39*, 568-575.

(21) Alzheimer, A. Über eine eigenartige erkrankung der Hirnrinde. *Allgemine zeitschrift für Psychiatrie und physisch-gerichtlich Medizin* **1907**, *64*, 146-148.

(22) Ferreira, S. T.; Vieira, M. N.; De Felice, F. G. Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases. *IUBMB Life* **2007**, *59*, 332-345.

(23) Nettleton, E. J.; Tito, P.; Sunde, M.; Bouchard, M.; Dobson, C. M.; Robinson, C. V. Characterization of the oligomeric states of insulin in self-assembly and amyloid formation by mass spectrometry. *Biophys. J.* **2000**, *79*, 1053-1065.

(24) Pease, L. F.; Sorci, M.; Guha, S.; Tsai, D. H.; Zachariah, M. R.; Tarlov, M. J. Probing the nucleus model for oligomer formation during insulin amyloid fibrillogenesis. *Biophys. J.* **2010**, *99*, 3979-3985.

(25) Hua, Q. X.; Weiss, M. A. Mechanisms of insulin fibrillation: the structure of insilin under amyloidogenic conditions resembles a protein-folding intermediate. *J. Biol. Chem.* **2004**, *279*, 21449-21460.

(26) Bumagina, Z.; Gurvits, B.; Artemova, N.; Muranov, K.; Kurganov, B. Paradoxical acceleration of dithiothreitol-induced aggregation of insulin in the presence of a chaperone. *Int. J. Mol. Sci.* **2010**, *11*, 4556-4579.

(27) Manno, M.; Craparo, E. F.; Podesta, A.; Bulone, D.; Carrotta, R.; Martorana, V. Kinetics of different processes in human insulin amylid formation. *J. Mol. Biol.* **2007**, *366*, 258-274.

(28) Nayak, A.; Sorci, M.; Krueger, S.; Belfort, G. A universal pathway for amyloid nucleus and precursor formation for insulin. *Proteins* **2009**, *74*, 556-565.

(29) Vestergaard, B.; Groenning, M.; Roessle, M.; Kastrup, J. S.; van de Weert, M.; Flink, J. M. A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *PLoS* **2007**, *5*, e134.

(30) Sorci, M.; Grassucci, R. A.; Hahn, I.; Frank, J.; Belfort, G. Time-dependent insulin oligomer reaction pathway prior to fibril formation:cooling and seeding. *Proteins* **2009**, *77*, 62-73.

(31) Podestra, A.; Tiana, G.; Milani, P.; Manno, M. Early events in insulin fibrilisation studied by time-lapse atomic force microscopy. *Biophys. J.* **2006**, *90*, 589-597.

(32) Haas, J.; Vohringer-Martinez, E.; Bogehold, A.; Matthes, D.; Hensen, U.; Pelah, A.; Abel, B.; Grubmuller, H. Primary steps of pH-dependent insulin aggregation kinetics are governed by conformational flexibility. *Chembiochem : a European journal of chemical biology* **2009**, *10*, 1816-1822.

(33) Veltova, A.; Tatarek-Nossol, M.; Andreetto, E.; Kapurniotu, A. Exploiting cross-amyloid interactions to inhibit protein aggregation but not function: Nanomolar affinity inhibition of insulin aggregation by an IAPP mimic. *Angew. Chem int. Ed.* **2008**, *47*, 7114-7118.

(34) Abel, B.; Charvat, A.; Diederichsen, U.; Faubel, M.; Grimann, B.; Niemeyer, J.; Zeeck, A. Applications, features, and mechanistic aspects of liquid water beam desorption mass spectrometry. *Int. J. Mass Spectrom.* **2005**, *243*, 177-188.

(35) Charvat, A.; Lugovoj, E.; Faubel, M.; Abel, B. Analytical laser induced liquid beam desorption mass spectrometry of protonated amio acids and their non-covalently bound aggregates. *Eur. Phys. J. D* **2002**, *20*, 573.

(36) Charvat, A.; Lugovoj, E.; Faubel, M.; Abel, B. New design for a time-of-flight mass spectrometer, RSI. *Rev. Sci. Instr.* **2004**, *75*, 1209-1218.

(37) Faubel, M.: Photoelectron spectroscopy at liquid surfaces. In *Adv. Series in Physical Chemistry*; Ng, C. Y., Ed.; World Scientific: New York, 2000; Vol. 10A.

(38) Faubel, M.; Kisters, T. Non-equilibrium molecular evaporation of carboxylic acid dimers. *Nature* **1989**, *339*, 527-529.

(39) Guttmann, P.; Schmahl, G.; Niemann, B.; Rudoph, D.; Schneider, G.; Bahrdt, J.: The X-Ray Microscopy project at BESSY II. In *X-Ray Microscopy and Spectromicroscopy*; J. Thieme, G. S., D. Rudoph, E. Umbach, Ed.; Springer: Berlin-Heidelberg, 1998.

(40) Charvat, A.; Abel, B. How to make big molecules fly out of liquid water: applications, features and physics of laser assisted liquid phase dispersion mass spectrometry. *Physical chemistry chemical physics : PCCP* **2007**, *9*, 3335-3360.

(41) Eigen, M. Prionics or the kinetic basis of prion diseases. *Biophys. Chem.* **1996**, *63*, A1-A18.

(42) Heldt, C. L.; Kurouski, D.; Sorci, M.; Grafeld, E.; Lednev, I. K.; Belfort, G. Isolating toxic insulin amyloid reactive species that lack beta-sheetsand have a wide pH-Stability *Biophys. J.* **2011**, *100*, 2792-2800.

(43) Foderà, V.; Librizzi, F.; Groenning M.; van de Weert, M.; Leone, M., Secondary Nucleation and Accessible Surface in Insulin Amyloid Fibril Formation, J. Phys. Chem. B **2008**, 112, 3853–3858

Figures and Table with Captions

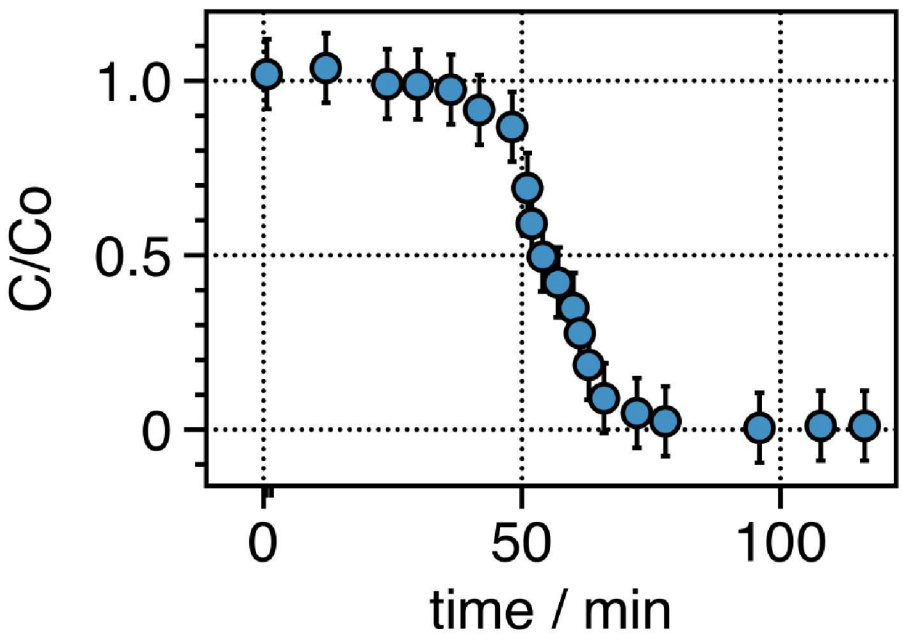


Fig. 1 Mass signal of the protonated insulin monomer ($1 \cdot 10^{-4}$M, T=65˚C, pH=2, weakly stirred) as a function of time during fibril formation under destabilizing conditions (pH=2 and 65˚C). Error bars mainly reflect the uncertainty from one experiment to the other. The mass spectrum can be inspected in Fig. 5. The typical sigmodial trace is normalized (C/C_0) and the time for the relative intensity of C/C_0=0.5 is referred to as the "lag time τ" in the following. A linear response of gas phase signal as a function of solution concentration has been found. For more details see the text.

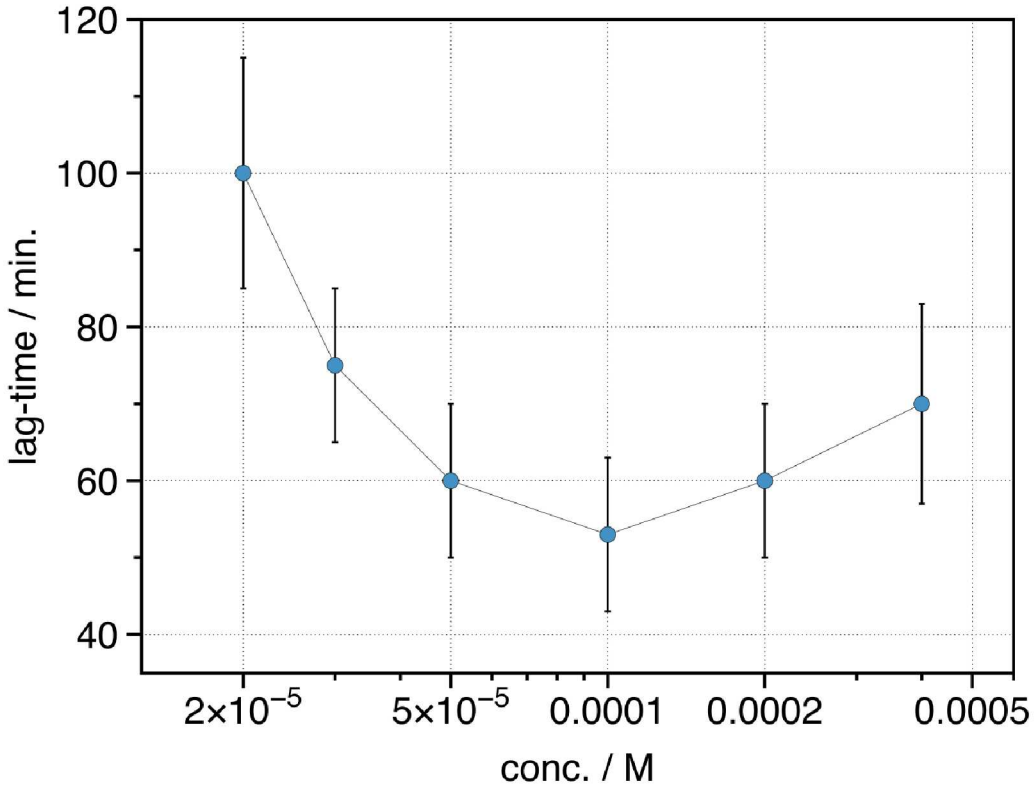


Fig. 2 Lag time τ (measured with liquid beam laser desorption MS) as a function of insulin initial concentration ($2x10^{-5}M < c < 5x10^{-4}M$). The concentration dependence is only small to moderate. The solid line is meant to guide the readers eyes. Error bars reflect the uncertainties in the average of 4 experiments making up one data point here.

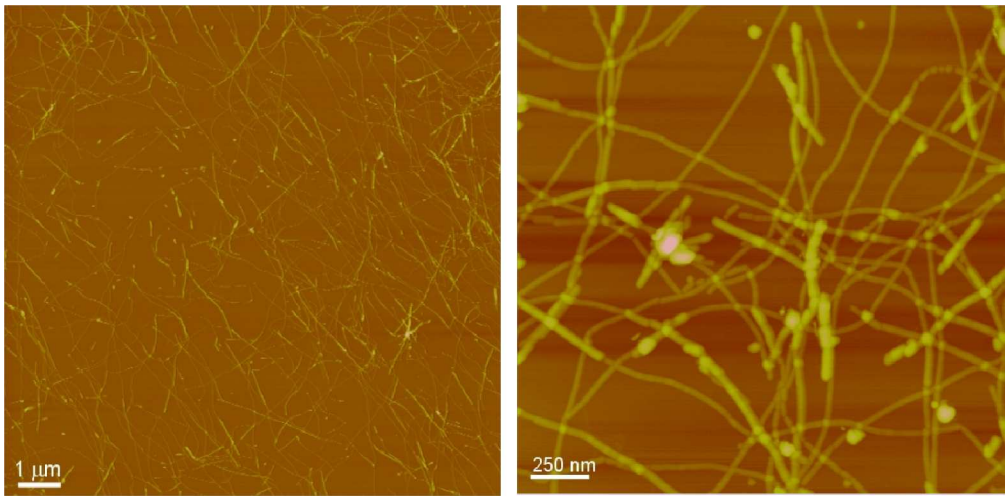


Fig. 3 Insulin fibrils (dry samples) in high resolution after the aggregation process (t >1.5h, $1 \cdot 10^{-4}$M, T=65˚C, pH=2, weakly stirred). The scale bar on the left corresponds to 1µm and the scale bar on the right corresponds to 250nm, respectively.

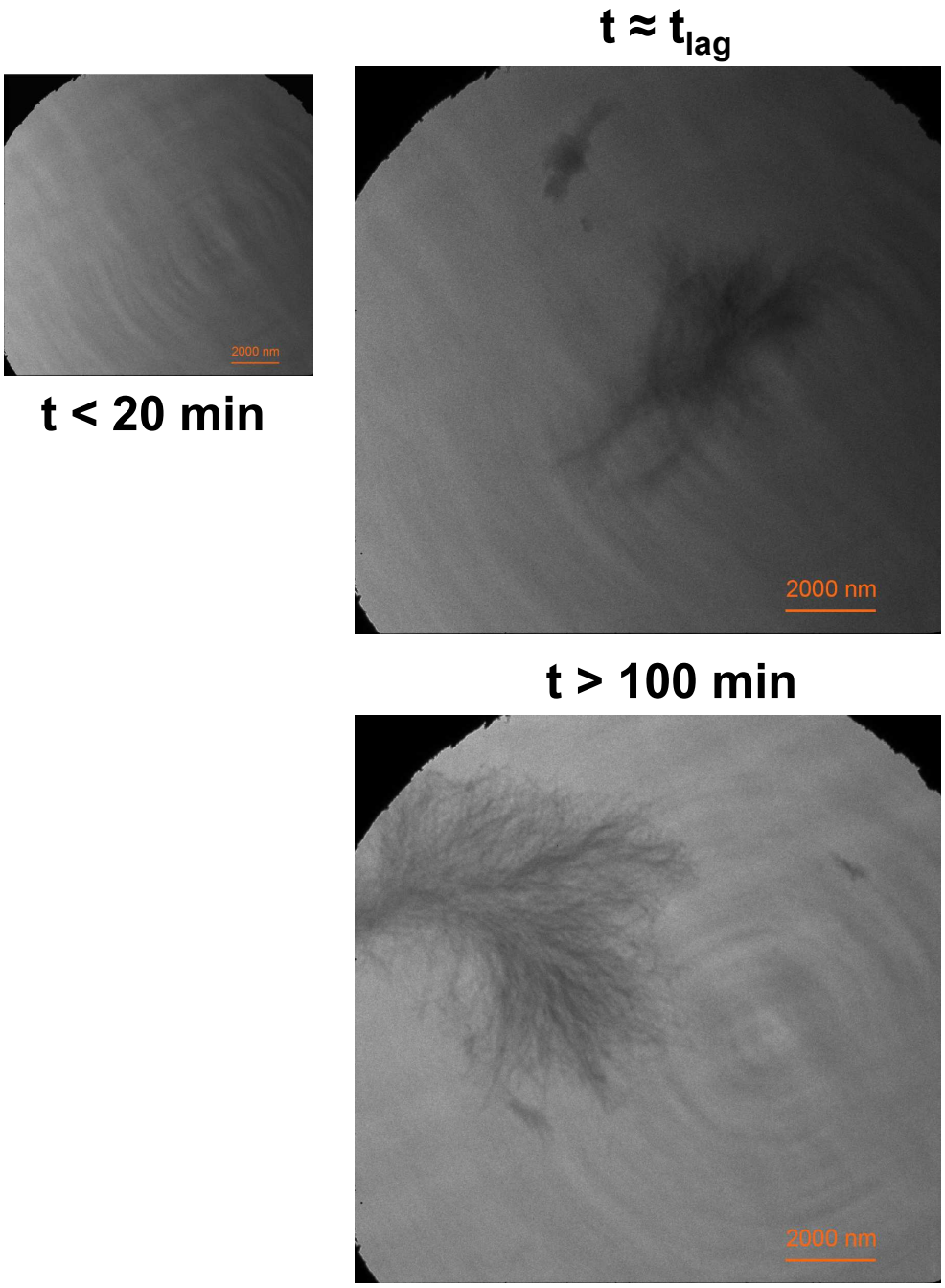


Fig. 4 Series of images as obtained from the TXM x-ray microscope (BESSY, Berlin) during the insulin aggregation process (T=65˚C, pH=2, HCl, $1 \cdot 10^{-4}$M). Scalebar is 2000 nm in all figures.

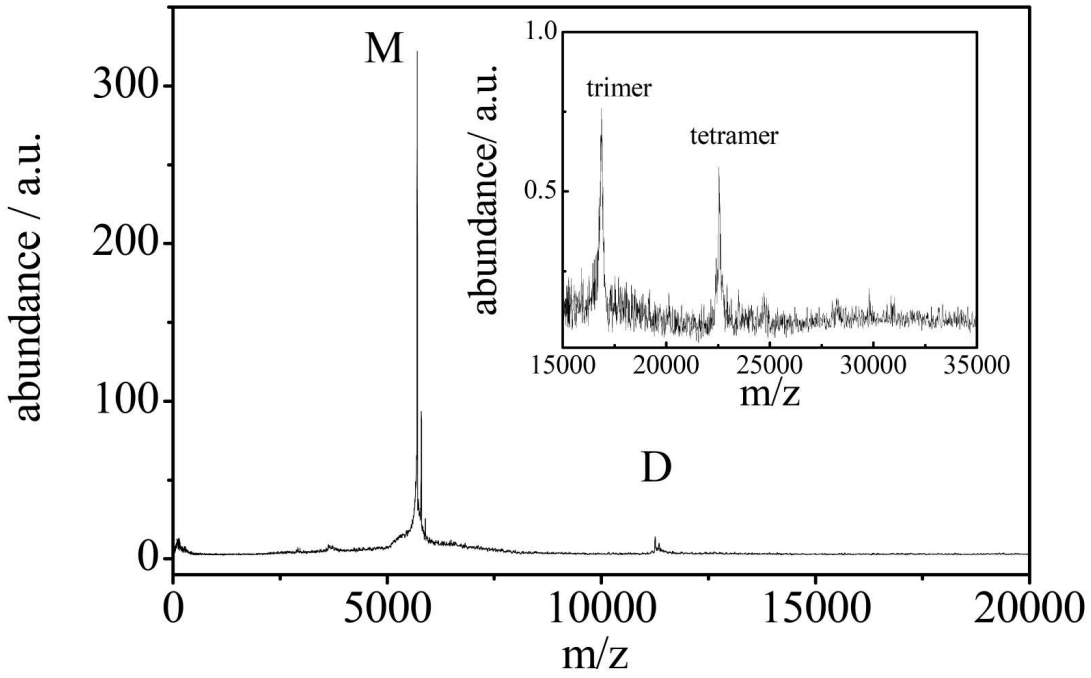


Fig. 5 Mass spectrum of insulin in aqueous solution (at 10^{-4}M, T=65˚C, pH=2,) measured at early times (t=0) of the aggregation process (t << lag time τ) with liquid beam laser desorption mass spectrometry. The protonated monomer and the dimer can be easily observed (singly charged). At higher amplification, in the inset, also protonated and singly charged trimmers, tetramers and possibly even pentamers are detectable in small abundance.

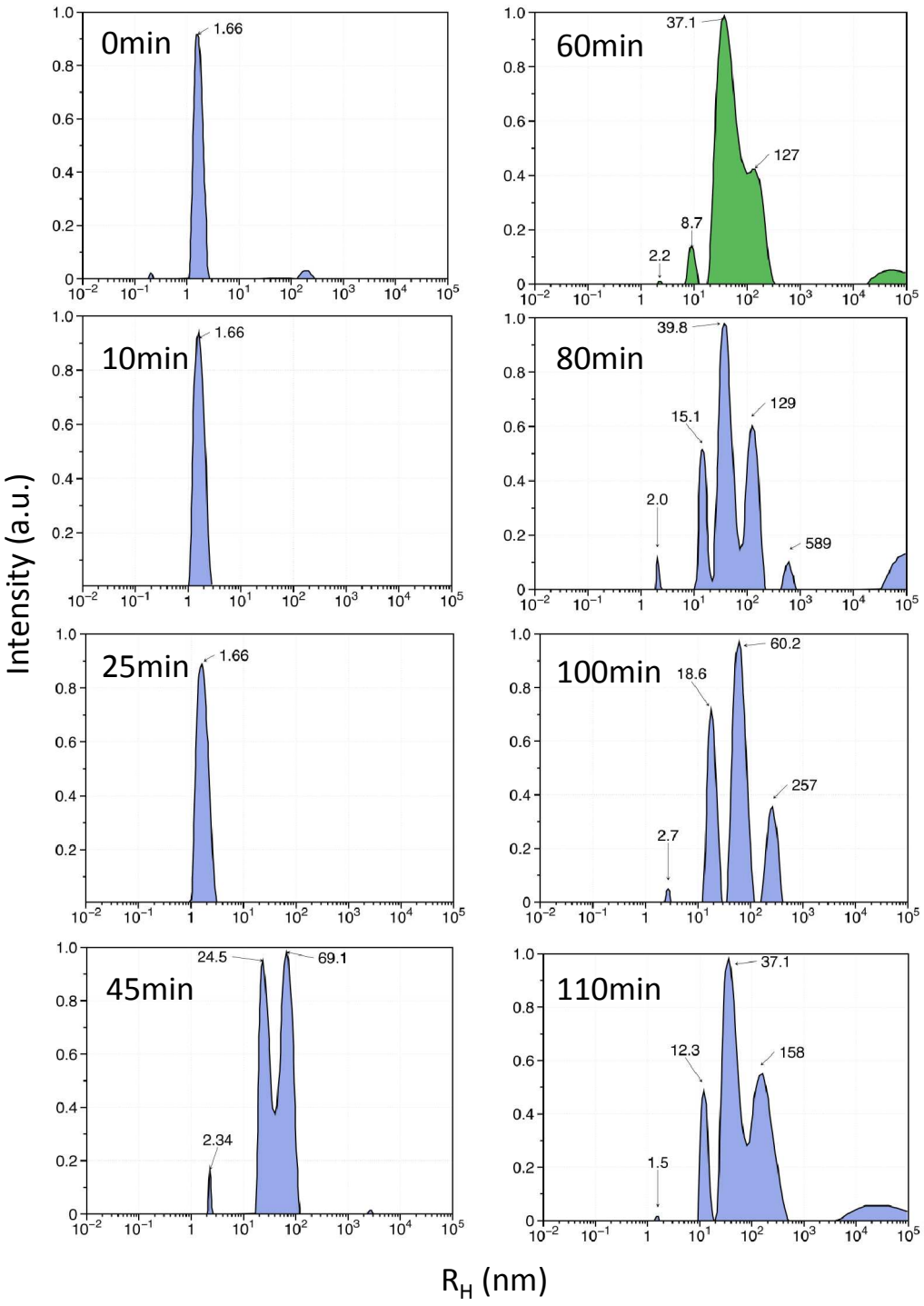


Fig. 6 Hydrodynamic radius R_H (distribution of radii for an inhomogeneous distribution of scatterers) measured during the insulin fibrillation process ($1 \cdot 10^{-4}$M, pH=2 and 65˚C, H_2SO_4) as obtained from the dynamic light scattering experiments (DLS) as a function of time. For more details and color coding see the text.

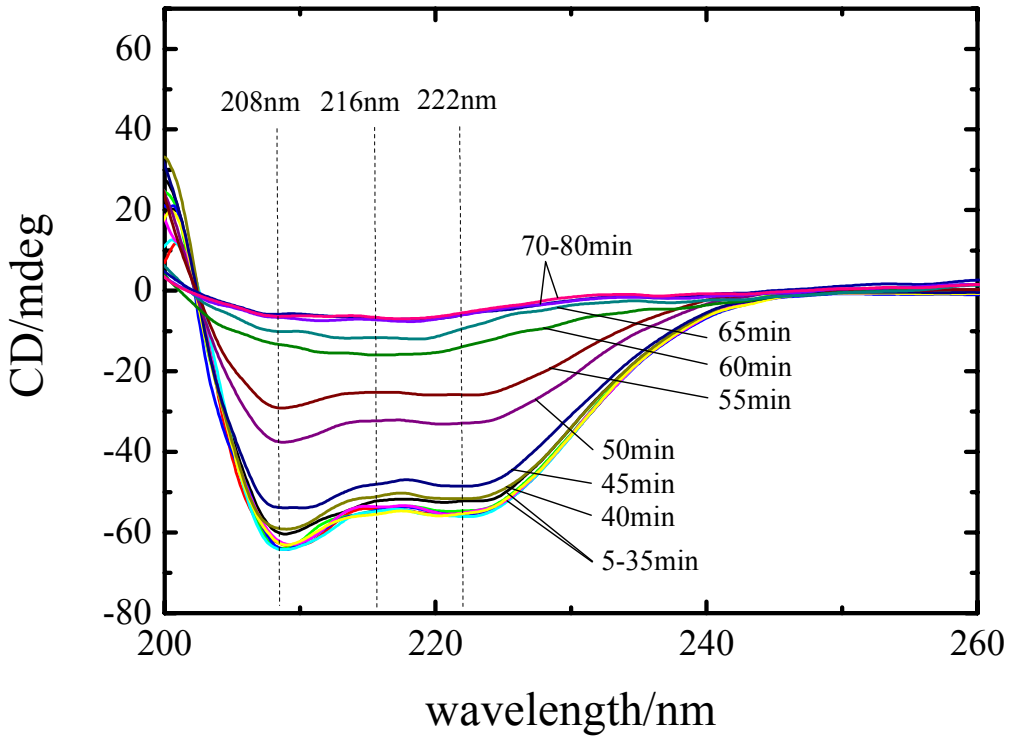


Fig. 7 CD spectra recorded during the insulin aggregation process (10^{-4}M, pH=2 and 65˚C, H_2SO_4). For more details see the text.

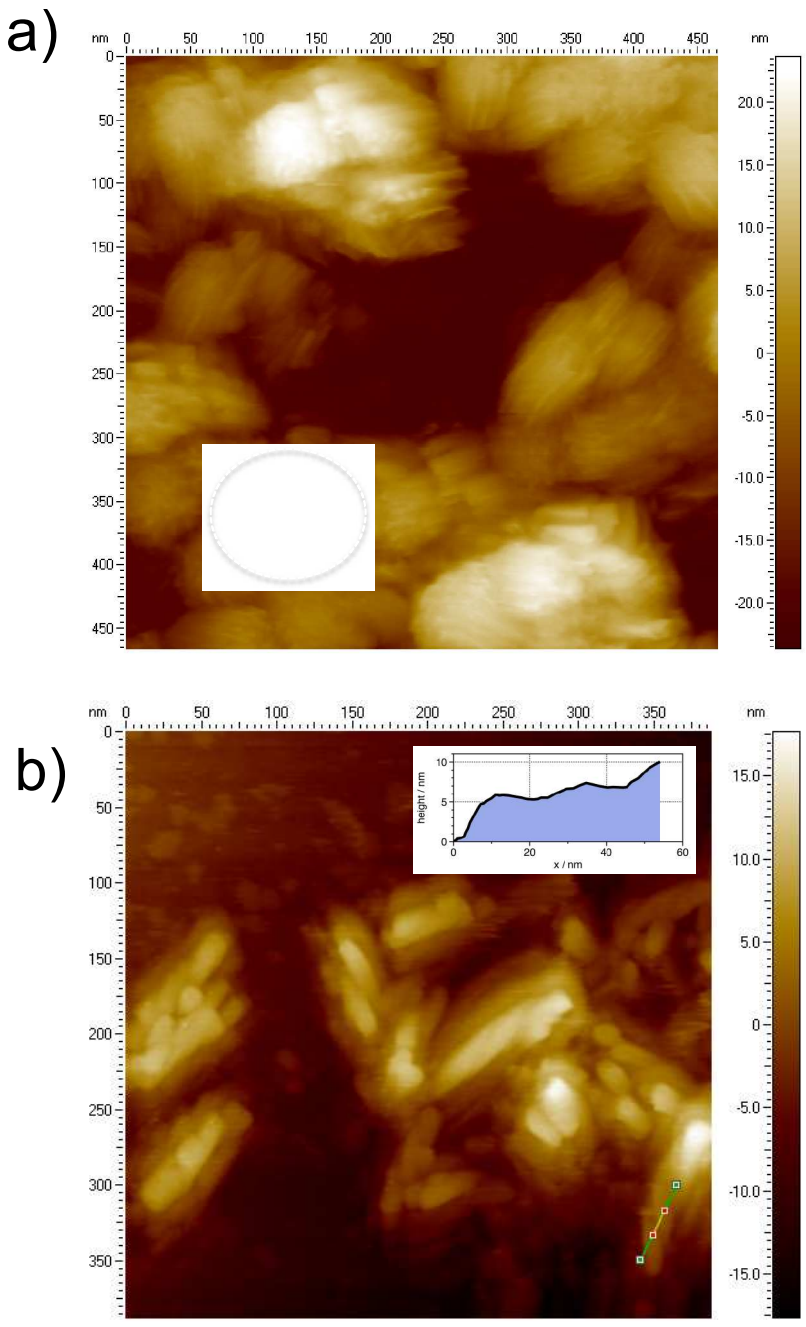


Fig. 8 a) Structures observed during the aggregation process with AFM at t≈τ (50-60min) of the insulin aggregation (pH=2 and 65˚C, H_2SO_4). Note, the structures are not artefacts and measured for samples taken and measured offline with the AFM (tapping mode) in the water phase (small droplet on mica). The circle marks structures on the order of 50nm in diameter – an order of magnitude also measured in DLS-experiments. **b)** Short fibrils observed after t>80min in liquid phase. The lateral resolution of the images is somewhat reduced and the structures appear somewhat fuzzy in the aqueous phase as opposed to Fig. 3 (dry samples). In the inset a height plot of a scan in the right lower corner of Fig. 8b is shown. It displays the characteristic height and the screw-turn of early fibrils in aqueous solution.

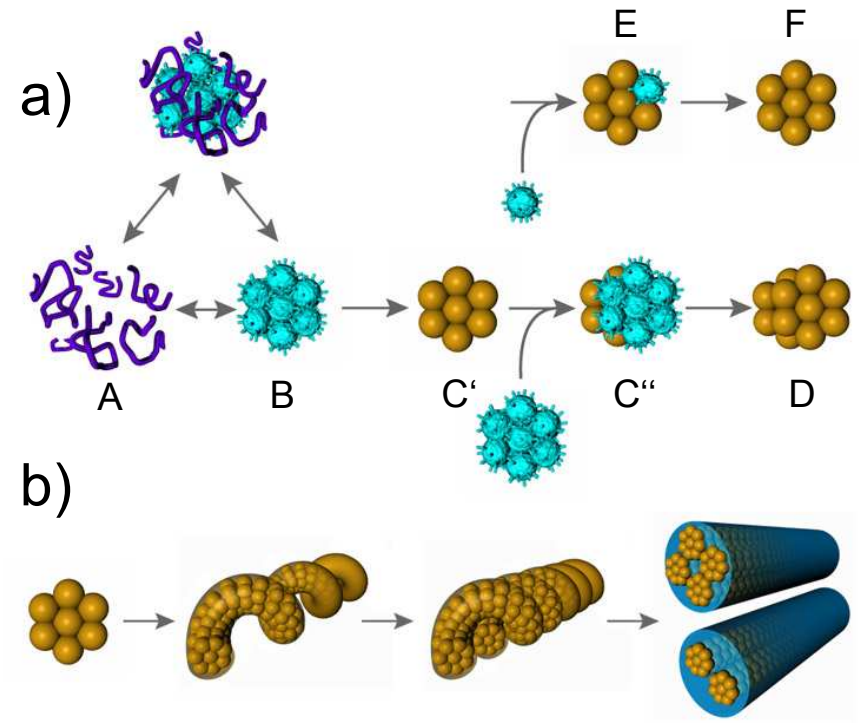


Fig. 9 Schematic picture of the nucleated conformation conversion aggregation process. a) initial steps and a) formation of larger structures from initial building blocks. For more details (labels A-F) see the text.

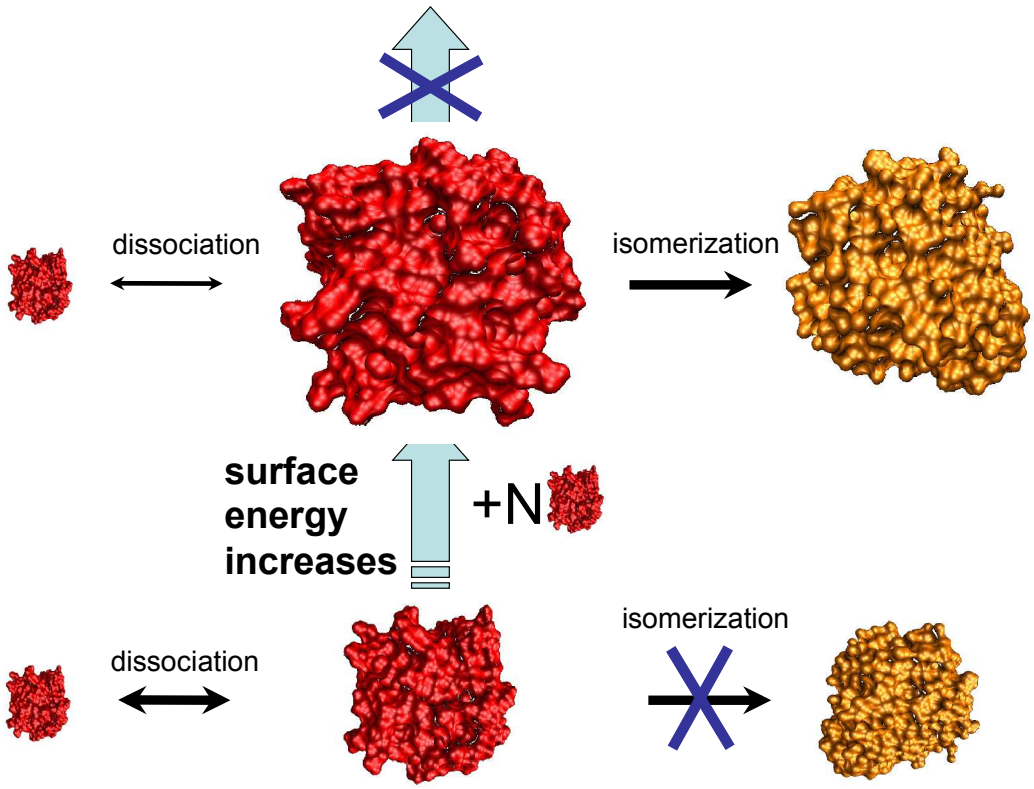
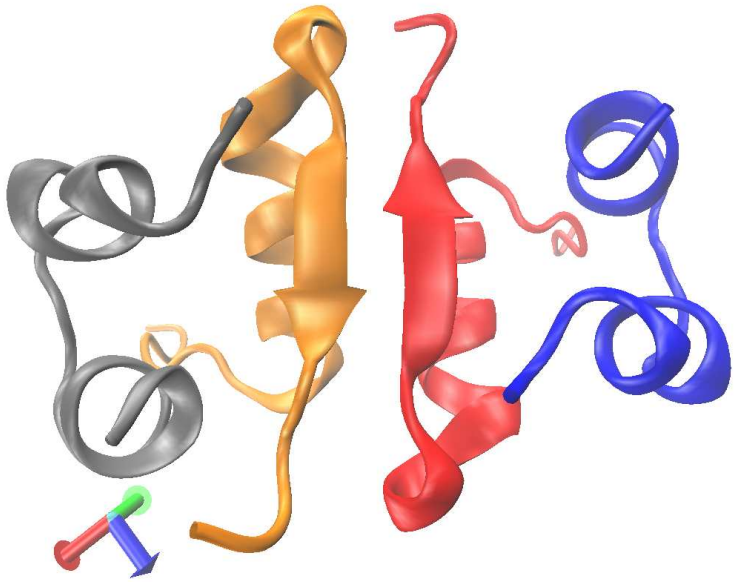


Fig. 10 Kinetic picture **(„cartoon“)** of insulin aggregation and nucleation. Note, the aggregate size (loose structures not yet converted) is limited and determined by the surface energy. While dissociation is more likely at smaller sizes – in agreement with nucleation theory – the contribution of isomerization is larger at larger aggregate sizes, limiting the upper size of the aggregates. This may limit the thickness for primary fibrils. They may nevertheless display larger thicknesses though formation of intertwined overstructures. For more details see the text.



TOC (figure for table of contents): Insulin under acidic conditions. PDB-Databank structure visualized with VMD.

method	**references**	**exp. conditions**	**detect. aggregates**
mass spectrometry	[23]	HCl, pH=2, 70°C	up to 7-mer
ESI differential mobility analysis	[24]	HAc, pH=2.1, 65°C	1-mer,2-mer, 3-mer
dynamic light scattering (DLS)	[26], [27]	HCl, pH=1.6-2, 60°C	Oligomers, R_g=1.3nm, 94nm
small angle neutron scattering	[28]	HCl, pH=1.6, 65°C	6-mer
AFM	[31]	HCl, pH=1.6, 60°C	oligomer distribution
cryogenic TEM	[30]	HCl, pH=1.6, 65°C	oligomers up to 15-20nm in diameter
electrophoresis	[33]	HCl, pH=2, 60°C	small-MW oligomers visible, high-MW oligomers not visible
membrane filtration	[42]	HCl, pH=1.6, 65°C	isolated ≈50nm toxic oligomer

Table 1: Analytical methods and attempts to trace insulin oligomers.