Dynamics of pathological aggregates and fibers

Kevin Pounot

To cite this version:

Kevin Pounot. Dynamics of pathological aggregates and fibers. Neurobiology. Université Grenoble Alpes [2020-..], 2020. English. NNT : 2020GRALY014 . tel-03099901

HAL Id: tel-03099901

https://theses.hal.science/tel-03099901v1

Submitted on 6 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université
Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L’UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique pour les Sciences du Vivant

Arrêté ministériel : 25 mai 2016

Présentée par

Kevin POUNOT

Thèse dirigée par **Martin WEIK**
et codirigée par **Tilo SEYDEL**, Communauté Université Grenoble Alpes
et **Giorgio SCHIRO**, Chargé de recherche, CNRS

préparée au sein de l'**Institut Laue-Langevin**
et l'**Institut de Biologie Structurale**
dans **l'École Doctorale Physique**

Dynamique des agrégats et fibres pathologiques

Dynamics of pathological aggregates and fibers

Thèse soutenue publiquement le **10 juin 2020**
devant le jury composé de :

Madame Judith PETERS
Professeur des Universités, Université Grenoble-Alpes, Présidente
Monsieur Franz BRUCKERT
Professeur des Universités, Grenoble INP, Examinateur
Monsieur Michael DECRESSAC
Chargé de recherche, INSERM délégation Alpes, Examinateur
Madame Gabriele KAMINSKI SCHIERLE
Docteur-Chercheur, Université de Cambridge - Royaume-Uni Examinateur
Monsieur Andreas STADLER
Chercheur HDR, Centre de recherche Jülich - Allemagne, Rapporteur
Monsieur Fabio STERPONE
Directeur de recherche, CNRS délégation Paris-Centre, Rapporteur

Remerciements

En premier lieu, je tiens à remercier Martin Weik, qui m'a fait confiance et m'a donné l'opportunité de faire cette thèse sur ce sujet passionnant. Son aide pour l'écriture des articles m'a été très précieuse. De plus, il sait aussi bien régler les problèmes que mettre en avant le travail et les succès de chacun. Cela m'a permis d'apprendre beaucoup et de rester motivé tout au long de mon travail de thèse.

Je remercie Giorgio Schirò et Tilo Seydel qui ont été d'une grande aide. Ils ont su me donner des conseils et me guider, tout en me laissant réfléchir et parvenir au résultat par moi-même. Je pense que c'est important, pour une formation par la recherche, d'avoir cette liberté d'action, qui incite à prendre du recul sur son travail et ainsi à prendre de meilleures décisions. Pendant ces trois ans, ils ont toujours été très disponibles et réactifs pour discuter des problèmes rencontrés. N'étant pas de formation en physique, j'aurais difficilement pu obtenir ces résultats sans leurs précieux conseils et explications.

Je remercie aussi Vito Foderà, qui m'a accueilli dans son laboratoire à Copenhague. Vito est toujours à l'écoute et ses compétences dans le domaine de l'agrégation des protéines ont permis de grandement faire avancer le projet. Cela a été aussi pour moi l'occasion de découvrir cette belle capitale.

En outre, il me faut remercier aussi Douglas Tobias, qui m'a accueilli dans son laboratoire pour trois mois, avec le magnifique soleil de Californie. Doug est un excellent scientifique et j'ai beaucoup appris avec lui. Humainement aussi, il est difficile de ne pas s'entendre avec Doug, surtout après avoir goûté son excellent chili con (ou sin) carne. Il donne une importance particulière à l'entraide et à la cohésion dans son groupe. J'ai rencontré beaucoup de personnes qui m'ont fait apprécier la Californie, il me faut citer Ellen Burt, qui m'a accueilli chez elle pour ces trois mois, m'a fait découvrir de nombreux aspects de la vie en Californie et avec qui j'ai partagé de très bon repas. Il y avait aussi les thésards du groupe, notamment Saleh et Andrew qui m'ont emmené découvrir la côte pacifique en vélo. Enfin, Yann Fichou, ancien thésard de l'équipe qui m'a aidé sur le projet de la protéine tau, avec qui j'ai passé mon premier Thanksgiving à Santa Barbara et qui m'a initié à la fabrication du bon

pain au levain maison.
De retour en France, je tiens à remercier tout les membres du groupe DYNAMOP. Julie, Kyprianos, Elena et Daniel avec qui je partage la vie de thésard, Guillaume sans qui les pauses déjeuner ou thé ne serait pas les mêmes, Tadeo et sa bonne humeur, et notre super technicienne Ninon avec qui j'allais prendre le thé le matin et qui a été d'une efficacité redoutable pour la production de la protéine tau.
Toujours en production de protéine, Martine Moulin, au D-Lab, m'a beaucoup aidé et le projet sur l'α-synucléine ne serait pas allé bien loin sans elle.
Au-delà de la thèse, de nombreuses personnes m'ont permis d'arriver jusque là. Notamment pendant mes stages avec le groupe RMN de l'IBS et surtout Catherine Bougault, Cédric Laguri et Jean-Pierre Simorre. Et je dois citer aussi Robert Blanvillain, qui m'a permis de replonger dans l'extraordinaire monde des plantes avec un projet sur la biogenèse du chloroplaste.
Je remercie aussi ma famille et ma belle-famille, par le simple fait de passer du temps ensemble, pour les rires, pour les balades, pour leur générosité et les longues soirées à discuter et boire la goutte.
Enfin, et surtout, je remercie ma femme Frédérique Francillon, pour son soutien, sa fierté pour mon travail, sa patience, sa façon de me faire rire, ses bons gâteaux, sa tendresse. Elle a été, et continue à être, mon meilleur moteur pour ma thèse mais aussi pour tout le reste. Merci à toi, ce travail t'est dédié.

The Institut Laue-Langevin (ILL) and the Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) are greatfully acknowledged for funding the research for this thesis. We are grateful to the ILL and the Heinz-Maier Leibnitz Zentrum (MLZ) for the neutron beamtime, and to the European Synchrotron Radiation Facility (ESRF) for the X-ray beamtime.

Thesis abstract

Neurodegenerative diseases, such as Parkinson's and Alzheimer's, constitute a growing threat with ever increasing prevalence. These diseases are characterized by the presence of protein deposits in the patient's brain that are called amyloids. Several proteins were identified in these deposits as being the molecular hallmark of the disorder, among which we can cite α-synuclein for Parkinson's disease and tau for Alzheimers's disease. Protein amyloid aggregation is central to neurodegenerative diseases and hence constitutes a target of choice for diagnostic and therapeutic attempts. It is characterized by the formation of a structural cross-β pattern, which is a stack of β-sheets, usually forming long fibrils. Under specific conditions, larger aggregates can be obtained, such as micrometer-sized particles, including so-called particulates and spherulites. Several pieces of evidence suggest that the formation of such aggregates, and especially at early-stages, can be involved in protein toxicity. Yet, the reasons for the aggregation to occur are not well understood. In this work, we aimed at deciphering the fundamental principles underlying protein amyloid aggregation by studying the changes in protein and hydration water dynamics, the understanding of which might help in the development of water-dynamics based diagnostic methods.

We employed mainly incoherent neutron scattering (on SPHERES at the MLZ and IN16B at the ILL) and molecular dynamics simulations. The former provides ensemble averaged information on hydrogen motions in the system, and the latter provides a fully atomistic picture from which dynamical and structural aspects can be investigated.

Studying α-synuclein, we could show that protein backbone and side-chain motions - that is, internal dynamics - is barely affected by aggregation. However, hydration water motions are increased around amyloid fibrils. The increased dynamics originates from a fraction of water molecules being displaced from the protein hydrophobic core to the hydrophilic termini regions when fibrils are formed. Hence, it results in a higher water entropy in fibrils, where the central cross-β pattern appears highly efficient in protecting itself from interacting with the solvent.

For γS-crystallin, comparison of the internal protein dynamics of the wild-type protein with a G18V mutant revealed that the mutant is less dynamic, whatever its aggregation state. This observation, along with the comparison of protein dynamics with their relative hydropathy index, indicates that the internal dynamics depends strongly on the amino acid composition, but not on the aggregation state. In addition, other factors can affect protein dynamics, such as the presence of metal ions.

The measurements carried out on insulin, in the presence or absence of zinc show that the metal promotes protein hydration at pH 1.8, where it interacts loosely with the protein. The zinc affects also aggregate-aggregate interaction, probably by electrostatic screening as the formation of spherulites is facilitated in the absence of the metal.

Eventually, the possibility to unambiguously and simultaneously access internal dynamics and center-of-mass diffusion was demonstrated by carrying out so-called fixed-window scans on the IN16B instrument at the ILL. This novel technique applied to lysozyme showed that particulate formation occurs in a one-step process, and the internal dynamics remains constant all along. This pilot experiment opens up the possibility to study fibril formation of pathologically relevant proteins.

Taken together, the aforementioned results demonstrate that we can now study the amyloid aggregation process with great detail, and there is a great opportunity to extend this work within a biological context, in order to link the biophysical properties of protein amyloid aggregation with its effects and toxicity *in-vivo*.

Résumé de la thèse

Les maladies neurodégénératives, comme Parkinson ou Alzheimer, sont une menace croissante, avec une prévalence augmentant sans cesse. Ces maladies sont caractérisées par la présence de dépôts protéiques, appellés amyloïdes, dans le cerveau. Plusieurs protéines ont été identifiées dans ces dépôts comme étant des marqueurs de la maladie, dont l'α-synucléine pour Parkinson et tau pour Alzheimer. L'agrégation amyloïde est centrale dans les maladies neurodégénératives et constitue ainsi une cible privilégiée pour le diagnostic ou les essais cliniques.

L'agrégation amyloïde est caractérisée par la formation d'un motif cross-β, qui consiste en un empilement de brins β, pouvant former ainsi de longues fibres. Dans certaines conditions, des particules de taille micrométrique peuvent être obtenus, tels que les 'particulates' ou les 'sphérulites'. Plusieurs études montrent que la formation des agrégats, en particulier au stade précoce, est impliquée dans la toxicité. En revanche, les raisons de l'agrégation des protéines ne sont pas bien comprises. Dans ce travail, nous avons cherché à comprendre les principes fondamentaux impliqués dans l'agrégation amyloïde, en étudiant les changements de dynamique du système protéine-solvent, ce qui de plus, pourrait aider le développement de nouvelles méthodes de diagnostic.

Dans ce but, j'ai utilisé principalement la diffusion incohérente des neutrons et les simulations de dynamique moléculaire. La première fournit une dynamique moyennée sur l'ensemble des atomes d'hydrogène dans le système et la seconde fournit une vision atomique dans laquelle structure et dynamique peuvent être étudiées.

En étudiant l'α-synucléine, j'ai montré que les mouvements des chaines latérales et principales - dynamique interne - sont inchangés par l'agrégation. Cependant, les mouvements de l'eau sont accélérés autour des fibres, ce qui provient d'une fraction de l'eau étant déplacée du coeur hydrophobe vers les régions terminales hydrophiles lors de la formation des fibres. Ainsi, l'entropie de l'eau est augmentée dans les fibres, ou le motif cross-β central semble être très efficace pour se protéger lui-même de l'intéraction avec le solvent.

La comparaison de la γS-crystalline sauvage avec le mutant G18V montre que le

mutant est moins dynamique, quel que soit l'état d'agrégation. Cette observation, et la comparaison de la dynamique interne avec l'hydrophobicité des protéines, montre que la dynamique interne dépend fortement de la composition en acides aminés et non pas de l'état d'agrégation. En outre, les ions métalliques peuvent aussi influencer la dynamique interne.

Les mesures sur l'insuline, en présence ou absence de zinc montre que le métal aide à l'hydratation de la protéine, même à pH 1.8, où il intéragit faiblement avec la protéine. Le zinc affecte aussi les intéractions entre agrégats, probablement par écrantage électrostatique, étant donné que la formation de sphérulites est facilitée en son absence.

Enfin, la possibilité de suivre en simultané, et sans ambiguïté, la dynamique interne et la diffusion du centre de masse a été démontrée en utilisant des scans à fenêtre d'énergie fixe sur l'instrument IN16B à l'ILL. Cette nouvelle méthode, appliquée au lysozyme, montre que la formation des 'particulates' se déroule en une étape, avec la dynamique interne restant constante tout au long du processus. Cette expérience pilote ouvre la voie à des études de fibrillation de protéines ayant un intérêt médical.

Ensemble, ces résultats démontrent que l'on peut étudier le processus d'agrégation amyloïde avec beaucoup de détails, et il y a une grande opportunité d'étendre ce travail dans un contexte biologique afin de faire le lien entre les paramètres biophysiques de l'agrégation amyloïde, ses effets et sa toxicité *in-vivo*.

Contents

Introduction

1.1 The amyloid plague

The last decades showed an important increase in Alzheimer's, Parkinson's and other neurodegenerative diseases prevalence, such that Alzheimer was called the 21st century's plague [Dobson, 2015]. Numerous factors seem to be involved, including genetic polymorphisms, environment, and ageing. However, these risk factors remain poorly understood, and as they are very challenging to study, results are often inconclusive [Brown et al., 2005].

These diseases share a common feature, namely the presence of protein-rich deposits - called amyloid plaques - in the brain [Braak et al., 1989]. After the observation of these plaques in patient brains, several proteins were identified as being involved in their formation. Those include α-synuclein [Breydo et al., 2012], tau [Iqbal et al., 2010], or even insulin - outside of the brain [Swift et al., 2002]. These proteins, as well as many others, can aggregate, adopt a new conformation characterized by a cross-β pattern, and form fibrils [Chiti and Dobson, 2006] or super-structures [Vetri and Foderà, 2015] of micrometer size, including the so-called spherulites and particulates.

The aggregation process is thought to be the main cause of damage observed in Alzheimer or Parkinson patients. Even though visible effects of aggregation are mostly located in the brain, it is not clear whether these proteins start to aggregate there. For α-synuclein, it has been found that fibrillation can start in the gut, then fibrils migrate to the brain through the vagal nerve in a prion-like way [Chandra et al., 2017] [Liddle, 2018]. For tau, periodontitis was shown to be an important risk factor [Cerajewska et al., 2015], which might be associated with aggregation in the gums. Hence, the actual origin of amyloid aggregation is not trivial, as the process could potentially start in multiple organs, and have a physiological role, as suggested for tau [Dominy et al., 2019], and for α-synuclein [Liddle, 2018]. However, the consequences of aggregation on the cell integrity are becoming clearer.

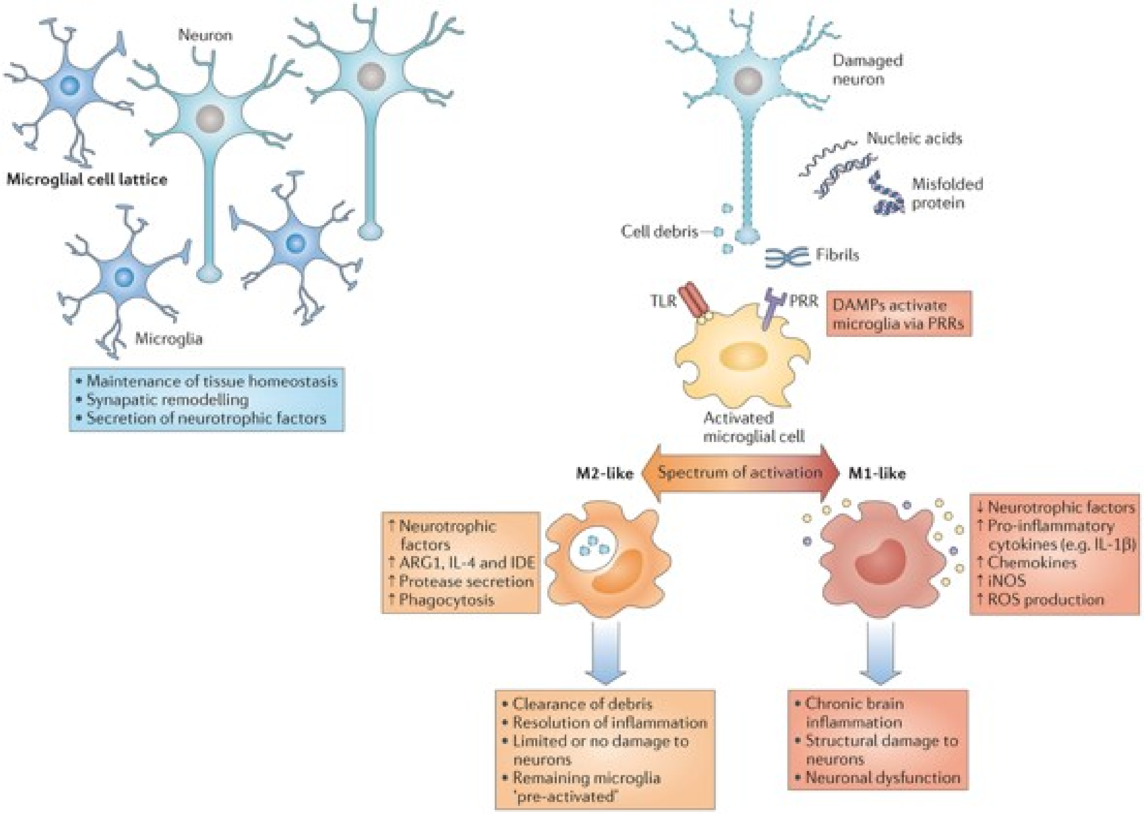


Figure 1.1: Glia activation after amyloid protein cell membrane disruption. Adapted from [Heneka et al., 2014].

A globally accepted model for toxicity could be drawn from several studies. Firstly, it has been shown for α-synuclein that cytotoxicity is higher for mutants presenting slower fibrillation kinetics, or mutants that do not form fibrils at all, but remain oligomeric [Lázaro et al., 2014]. A specific folding could even be observed with αB-crystallin, where three proteins form a cytotoxic cylinder [Laganowsky et al., 2012]. Therefore, oligomers, or elongating fibrils appear to be the actual toxic species. Secondly, important damages were observed on negatively charged lipid bilayers in the presence of fibrillating proteins, while no damage occurs when mature fibrils are used [Hellstrand et al., 2013] [Martel et al., 2017] [Jones et al., 2012], thereby highlighting membrane disruption as a direct cause for cytotoxicity.

Along with these two points, an additional effect should be mentioned here, that is, activation of microglia - especially for its immune system role - by damage associated molecular patterns (DAMP). When M1-like cells are activated, pro-inflammatory cytokines, and reactive oxygen species (ROS), are released, which cause deleterious and repeated inflammations (fig. 1.1) [Brück et al., 2016] [Heneka et al., 2014].

Most of the current treatment strategies for amyloid disorders are focusing on neuroprotection, such as the use of Mementine [Olivares et al., 2012] that helps maintaining a reasonable level of glutamate, the use of L-Dopa or similar agonists, a dopamine agonist that helps maintaining the neuronal activity and has neurotrophic effect [Oertel and Schulz, 2016] or the use of anti-inflammatory drugs [Moore et al., 2010]. Even though these treatments do improve patient life quality, they do not prevent the disease progression [Oertel and Schulz, 2016]. It appears more and more necessary to go further and find solutions to treat the causes and not only the consequences for these neurodegenerative disorders. More recently, the early steps of Parkinson disease were better understood as we shall see.

A neurodegenerative process is usually associated with impairment of the nuclear pore complex, and retention of mRNA in the nucleus [Tamburrino and Decressac, 2016]. In addition, several genes involved in mitochondrial activity are down-regulated in dopaminergic neurons from substantia nigra that over-express α-synuclein, resulting in increased oxydative stress and loss of cells [Decressac et al., 2012]. Among these genes, nurr1 is of particular interest as it can help in rescuing damaged neurons from α-synuclein toxicity when over-expressed [Decressac et al., 2013b]. Physiologically, the accumulation of proteins within the cell is related to ageing in that the protein clearance mechanisms become less efficient as we are getting older. The reduced proteolysis and autophagy leads to protein accumulation in the cytoplasm. Adding potential shift from homeostasis and oxydative stress resulting from nurr1 down-regulation, the amyloid aggregation can be favored [Stefanis et al., 2019]. Using α-synuclein as an example, the relations between protein expression, clearance and aggregation are summarized in figure 1.2. In the case of α-synuclein, all the aforementioned results suggest that understanding what drives the protein amyloid aggregation and understanding the early steps of this process is of paramount interest to prevent disease progression.

The interest of this work is twofold, it provides valuable insights into the fundamental principles associated with protein amyloid aggregation, which can be studied as a function of various parameters, such as mutations, protein or ion concentration. Second, it might pave the way to new water dynamics based diagnostic methods [Snow et al., 2017], or new prophylactic strategies.

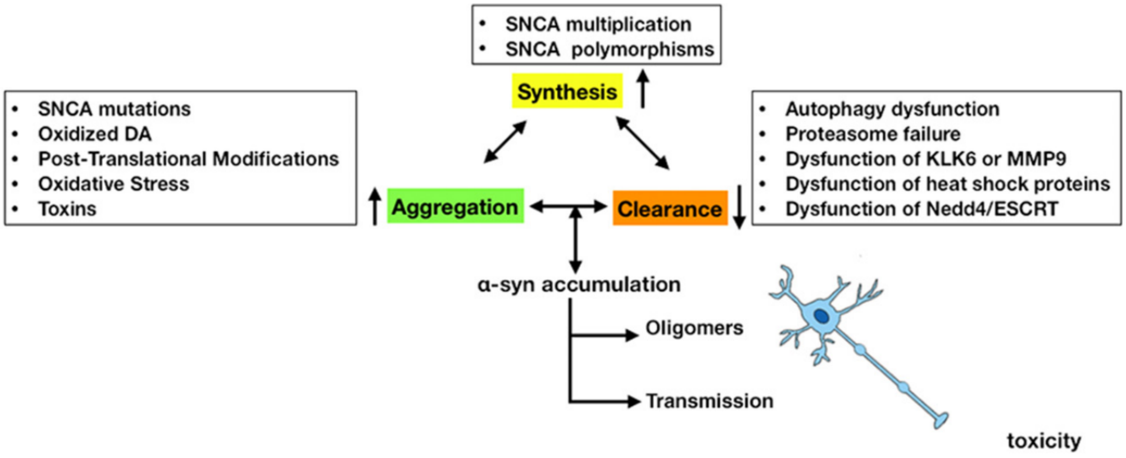


Figure 1.2: α-synuclein clearance and aging
The expression of the α-synuclein gene - SNCA - can be increased because of gene multiplication and polymorphism. In addition, protein clearance is impaired with aging as autophagy and proteolysis are reduced.
Adapted from [Stefanis et al., 2019]

1.2 Protein aggregation

1.2.1 Protein misfolding

The picture of a gene, being transcribed to an RNA, giving rise to a protein, whose well defined structure assures a specific function has been found to be too simplistic. After the observation of missing density in protein crystallography, clear signs of a disordered, yet functional, region were found using nuclear magnetic resonance (NMR) within histone H5 globular segment [Aviles et al., 1978]. From then on, a vast number of protein conformations were observed, ranging from the well-folded globular one to the intrinsically disordered proteins (IDP), among which α-synuclein and tau can be found. Specific features can be observed for different classes of proteins, that is, strong amino acid sequence determinants for whether a protein will be an IDP or well-folded (fig. 1.3). More specifically, the higher the hydrophobicity, and the lower the net charge, the more likely a protein will adopt a well-defined conformation [Uversky, 2011]. Moreover, IDPs are polyampholytes, with a homogeneous distribution of charged residues along the amino acid sequence [Das and Pappu, 2013]. They are almost never found in enzymes, or transporters, for which catalysis or channel formation across membranes require a well-defined structure. However, IDPs can adopt transient secondary structures upon interaction with specific targets - so-called folding upon binding [Wright and Dyson, 2009]-, and, in the case of α-synuclein, can also exhibit a globular shape *in-vivo*, where a small hydrophobic core is protected from solvent interaction by flanking disordered regions [Theillet et al., 2016]. The same globular, compact conformation was found *in-vitro* for α-synuclein [Brodie et al., 2019] and tau [Popov et al., 2019].

All proteins thus exhibit a set of conformational states, which relates to energetics and dynamics [Parak and Frauenfelder, 1993]. A protein energy landscape can be drawn, with several minima being separated by energy barriers. Thus, protein dynamics results in a wide range of motions, from femtosecond harmonic oscillations to milliseconds domain motions, and seconds to hours for protein-protein interactions and aggregation. Motions on these time-scales can be involved in biological function and are closely related to structure. This observation is particularly true for enzyme catalysis processes where a specific energy is needed for the reaction [Karplus and Kuriyan, 2005]. However, this dynamics-structure-function relationship is not always well understood, neither are the reasons for a specific folding or aggregation pathway for IDPs. In addition, further post-translational modifications can occur, such as chemical modification, pro-peptide cleavage, or complex formation, that will make the protein reach its actual active state.

Misfolding is not always trivial, and can be correctly defined only when the protein function is known. We will consider here a protein as misfolded, whenever a change in its structural or dynamical pattern, and possibly oligomerization state, results in a loss of function that is usually observed.

1.2.2 Molecular crowding and clustering

In addition to protein tertiary structure, proteins are always in interaction with each other, and can form quaternary, or oligomeric structures. The total protein concentration in cells lies between 100 and 300 mg/mL [Albe et al., 1990] on average. This high concentration implies a very crowded intracellular environment. In this environment, diffusion is greatly affected, as well as interactions, that can be either specific or non-specific. Molecular crowding can lead to clustering, driven by weak interactions, similar to a colloidal suspension. It has been suggested that an equilibrium results from the interplay between short-range patchy attractions, and long-range repulsions [Stradner et al., 2004]. The results obtained by small-angle scattering (SAS) with proteins are well reproduced using colloid-polymer mixtures (see fig. 1.4).

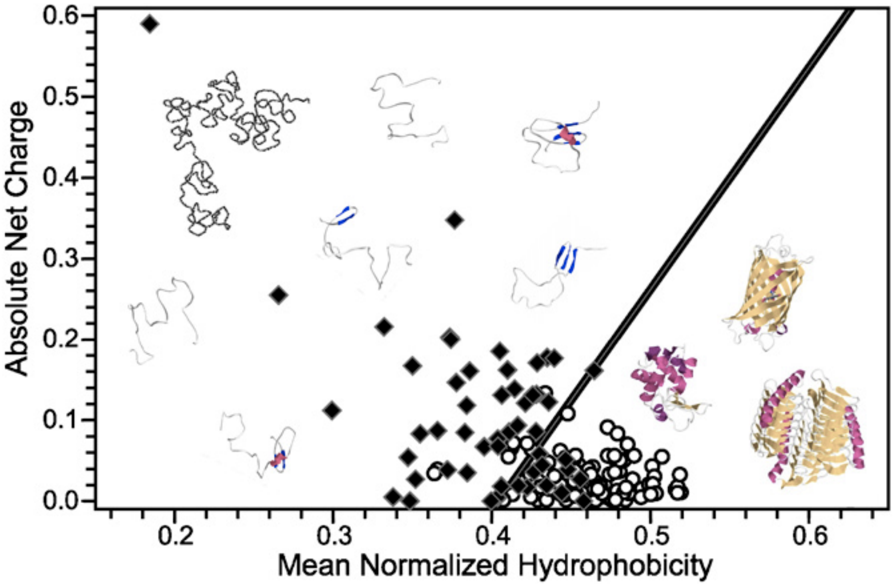


Figure 1.3: Sequence determinants for protein folding
Comparison of the mean net charge and mean hydrophobicity of 275 folded (open circles) and 91 natively unfolded (black diamonds) proteins. The line represents the border between extended IDPs and compact folded proteins.
Adapted from [Uversky, 2011].

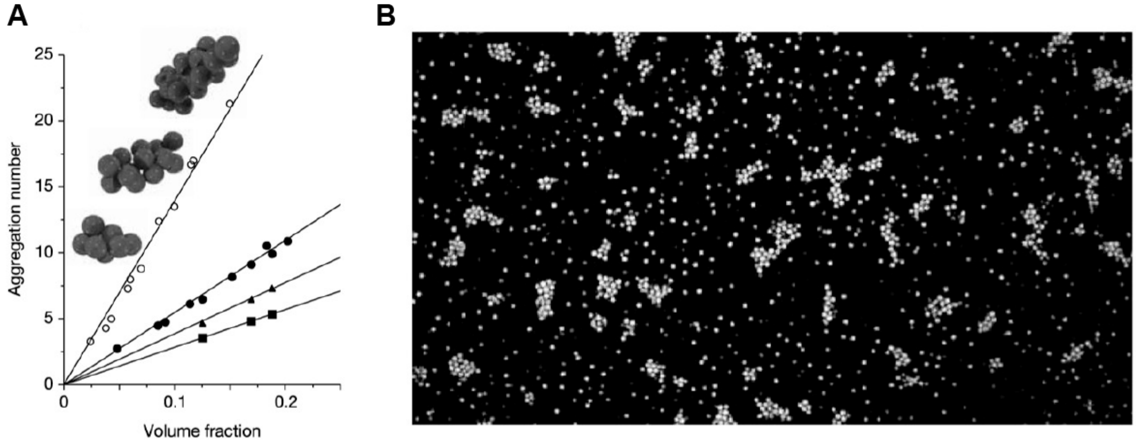


Figure 1.4: Cluster formation with proteins and colloid-polymer suspension
A, average aggregation number of clusters versus volume fraction in lysozyme samples at 5°C (filled circles), 15°C (triangles) and 25°C (squares), and in colloid-polymer mixtures (cp $\simeq$ 3 mg/ml, open circles). Typical cluster, reconstructed from confocal microscopy are shown in insets. **B,** confocal microscopy image from a colloid- polymer mixture.
Adapted from [Stradner et al., 2004]

The crowding effect was observed for several systems, and quasi-elastic neutron scattering has proven to be well suited for that [Grimaldo et al., 2014] [Braun et al., 2017]. The combination of experimental data with computer simulations has also shown, for crystallins, that anisotropic and patchy attractions are of major importance for the equilibrium distribution of cluster sizes [Bucciarelli et al., 2016]. However, no clear sign of phase separation was detected.

The absence of phase separation suggests that an additional phenomenon is involved in aggregation, as isotropic electrostatic potential and patchy weak attractions alone are not sufficient to explain the formation of stable amyloid fibrils. Indeed, DLVO theory (named after Boris Derjaguin, Lev Landau, Evert Verwey and Theodoor Overbeek), and in a general way, colloid theory appears as a good starting point to understand first steps of protein aggregation, as the effect of salt concentration can be well described [De Young et al., 1993] [Yoshimura et al., 2012]. However, these models consider only hard spheres with uniformly distributed charges, with the solvent being viewed as a continuum. More recently, the so-called ”patchy-particle” models gained attention as they proved to be accurate in describing the directional interaction, mediated by multivalent ions, between proteins in a crystal [Roosen-Runge et al., 2014].

1.2.3 Amyloid super-structures

Aggregation processes can follow two main pathways, one consists of ordered aggregates, among which I will present the so-called amyloid fibrils, and the other consists of amorphous aggregates that will be presented in the next section. Protein-rich deposits were identified in patients suffering from neurodegenerative disorders. These deposits exhibit a green birefringence after Congo red staining and appear as stacked fibrils under an electron microscope. These deposits were denoted amyloids [Glenner et al., 1988]. The common feature of amyloid proteins is the presence of a cross-beta pattern ([Eanes and Glenner, 1968] [Bonar et al., 1969] and fig. 1.5), which is a series of anti-parallel stacked beta strands, with possibly some additional three-dimensional features. This structural motif can be easily labelled using a fluorescent probe, Thioflavin T (ThT), which allows to follow in real-time the amyloid aggregation process. The aggregation process - as followed by ThT - usually presents a sigmoidal shape, characterized by an initial lag phase, followed by an exponential growth, and a final plateau phase. Finally, a fibrillar morphology is observed. Congo red birefringence, cross-beta pattern and fibrillar morphology defines the amyloid fibrils as described by Nilsson [Nilsson, 2004].

Several structures are available now, even using amyloid fibrils directly obtained from patients. Among them, we can mention the tau fibril core, obtained us-

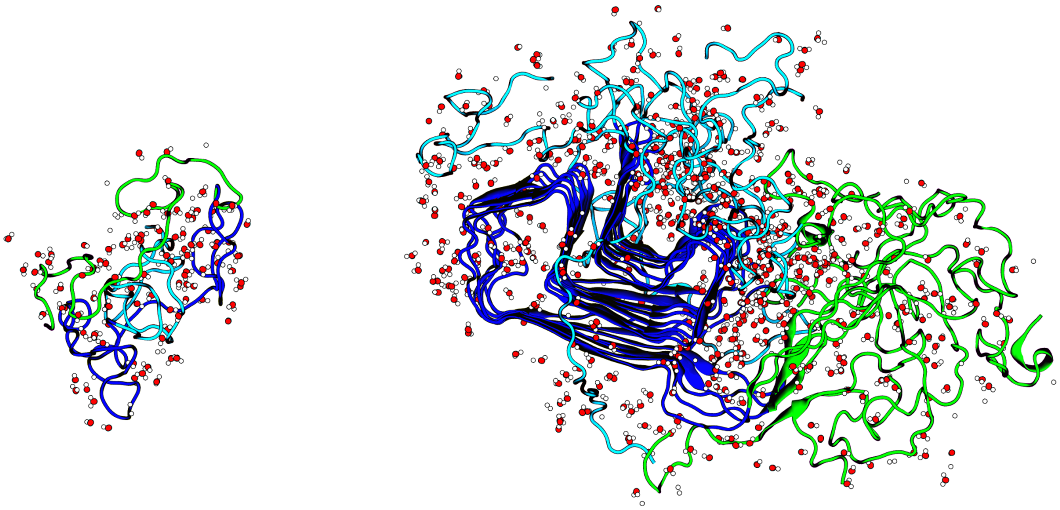


Figure 1.5: α-synuclein fibril formation
Left, native state, functional protein. **Right,** final fibrillar state. The N-ter region is depicted in green, the core region - residue 40 to 95 - in deep blue, and the C-ter region in cyan. Native ensemble of structures provided by Martin Blackledge, fibril structures from Tuttle et al. [Tuttle et al., 2016] and render made using VMD software [Humphrey et al., 1996].

ing cryo-electron microscopy (EM) [Fitzpatrick et al., 2017], where a specific three-dimensional folding can be observed, resembling to what has been called a greek-key topology. Here, approximately 16 % - 72 residues - of the protein's sequence is involved in the fibril core, which is surrounded by a disordered fuzzy coat. For α-synuclein, structures obtained using NMR, corresponding to the full-length protein [Tuttle et al., 2016], as well as a structure from EM, for the fibril core only [Li et al., 2018b][Li et al., 2018a] are available. Different topologies can be observed for the fibril core, with or without a twist along the fibril's main axis. Depending on the morphology, 40 to 50 residues are involved in the fibril core formation, which represents around 30 % of the sequence. No structural information is available yet concerning insulin or γ-crystallin fibrils.

It is interesting to note that the amyloid cross-β seems to be a structural pattern that can be adopted by various proteins [Dobson, 2003]. There is something universal in this very stable fold, that might make it very useful to understand the fundamental principles that govern protein folding and aggregation. In a study with the β_2-microglobulin, it is suggested that the nucleation-like pathway of amyloid aggregation is due to this cross-beta pattern formation, and not oligomerization [Yoshimura et al., 2012].

Surprisingly, a similar fold was observed, but with alpha-helices in place of beta sheets for *Staphylococcus aureus* PSMα3 protein [Tayeb-Fligelman et al., 2017]. This provides a clear example of functional amyloid structures, secreted by the bacteria as virulence factors.

It has been found that fibrils can associate and form even larger structures known as spherulites. These were first identified for bovine insulin [Krebs et al., 2004] [Krebs et al., 2005]. They consist in a radial arrangement of amyloid fibrils, and can easily reach 100 μm to mm in diameter, and are usually formed far from the protein's isoelectric point (pI) [Vetri and Foderà, 2015]. Spherulites of $A\beta_{42}$ were identified in Alzheimer's patients brain [Exley et al., 2009], making them also of prime medical interest. Another worth-to-mention super-structure - for which an amyloid nature is suspected but not yet demonstrated - is the particulate. They are spherical structures with a diameter ranging from hundreds of nanometers to around 2 μm. They can be formed at high temperature and at pH close to protein pI. For both of these super-structures, the hydrophobic interactions have been shown to play an important role, and most of the inner volume of the spherulites is devoid of water [Vetri and Foderà, 2015]. Several parameters have been shown to influence significantly the formation, the number, or the size of these assemblies, such as salt and protein concentrations or temperature [Smith et al., 2012].

Further work is still needed to understand super-aggregate formation, not only from a medical point of view. Indeed, these assemblies might be very relevant for bio-inspired materials, as a first study showed, where spherulites allow long-lasting, and controlled drug-delivery in the organism [Jiang et al., 2011].

1.2.4 Amorphous aggregates

For some aggregates, structural patterns might not be well resolved, present only as secondary structure or global, large scale structure, or there might not be any at all. In this case, aggregates are called amorphous. This kind of aggregates can be useful in biochemistry, especially with the formation of inclusion bodies, that are likely to play a biological role in bacteria as a protection mechanism against thermal stress [Villaverde and Carrió, 2003]. However, these structures can also be associated with serious issues, as it is the case for the γ-crystallin family. The crystallin aggregation process is responsible for eye-lens cataract formation, and several forms can be distinguished depending on the gene sequence and type of aggregate formed [Héon et al., 1999]. They come in three types, namely α, β, and γ, and genetic mutations associated with cataract were mostly located in the CRYG genes, coding for the γ-crystallins group. The monomeric protein exhibits a low frictional ratio, suggesting low propensity for interaction with solvent or other proteins [Vendra et al., 2016].
As for amyloids, the reasons for amorphous aggregates formation are not completely clear. Again, colloid theory provides some insights, highlighting the auxiliary role of high salt concentration [De Young et al., 1993]. Moreover, their aggregation kinetics can differ significantly from amyloids, that is, no lag phase appears and neither agitation nor seeding affect it, which indicates that growth does not follow a nucleation-dependent pathway [Yoshimura et al., 2012].

1.2.5 Relation to protein dynamics

Along with their structural properties, the dynamics of proteins is essential for biological function in the cell. Moreover, dynamics provides important information on the energy involved, and the thermodynamics of a particular change in the system, such as aggregation. Several time and length scales are involved, and parameters such as temperature, pH or salt concentration can influence the dynamics. Hartmann and coworkers measured - using X-ray crystallography and Mössbauer spectroscopy - mean-square displacements (MSD), of atoms in metmyoglobin [Hartmann et al., 1982]. They observed different kind of motions as a function of temperature, harmonic and anharmonic ones. Therefore, atoms in proteins experience harmonic potential, with weak restoring forces, but can also be involved in several conformational states, with transition rates k obeying the Arrhenius law: $k = Ae^{E_\mathrm{a}/RT}$, where A is a scalar, E_a the activation energy, R the ideal gas constant and T the temperature in Kelvin. Even at low temperature, Doster and coworkers found, using inelastic neutron scattering (INS), that harmonic, vibrational motions can be well described by a Debye-Waller factor, with MSD proportional to temperature, except for very low temperatures, where quantum tunnelling effects dominates [Doster et al., 1989]. In Doster's work, a transition is observed, around 200 K, which was described by considering an atom jump between two sites separated by a given distance d, and by an energy ΔG, and was called the protein dynamical transition (PDT).

Using high-resolution neutron backscattering (cf. section 2.1), a first transition typically becomes visible around 110 K (fig. 1.6), and corresponds to the onset of methyl groups rotation [Roh et al., 2005] [Wood et al., 2010] [Schiró et al., 2010] [Hong et al., 2013]. Subsequently, a second transition occurs, corresponding to the already mentioned PDT, associated with side-chains motions and backbone fluctuations. The temperature dependence of the MSD after this second transition is highly dependent on amino acid composition, as shown in a neutron scattering study [Schirò et al., 2011].

Interestingly, the MSD presents a relatively small contribution from the backbone compared to the contribution coming from amino acid composition, as it was shown by Schirò and co-workers [Schirò et al., 2011]. The backbone and amino acid sequence might mostly influence function-related structure and slow motions.

We can consider that fast dynamics - less than nanosecond time-scale - provides the most significant contribution to conformational stability - except for some proteins that perform electron transport for instance. Besides, slow, collective, dynamics is mostly related to function [Bahar et al., 1998]. Fast dynamics contributions can be understood in terms of free energy, using

$$\mathrm{F} = \mathrm{H} - \mathrm{TS} \tag{1.1}$$

with F, the free energy, H the enthalpy, T the temperature and S the entropy. In the canonical ensemble, the entropy is given by:

$$\mathrm{S} = -k_B \sum_\mathrm{i} p_\mathrm{i} \log(p_\mathrm{i}) \tag{1.2}$$

with k_B the Boltzmann constant and p_i given by Boltzmann distribution for the energy E_i of state i

$$p_\mathrm{i} = \frac{\exp^{-E_\mathrm{i}/k_\mathrm{B}T}}{\sum_\mathrm{j} \exp^{-E_\mathrm{j}/k_\mathrm{B}T}} \tag{1.3}$$

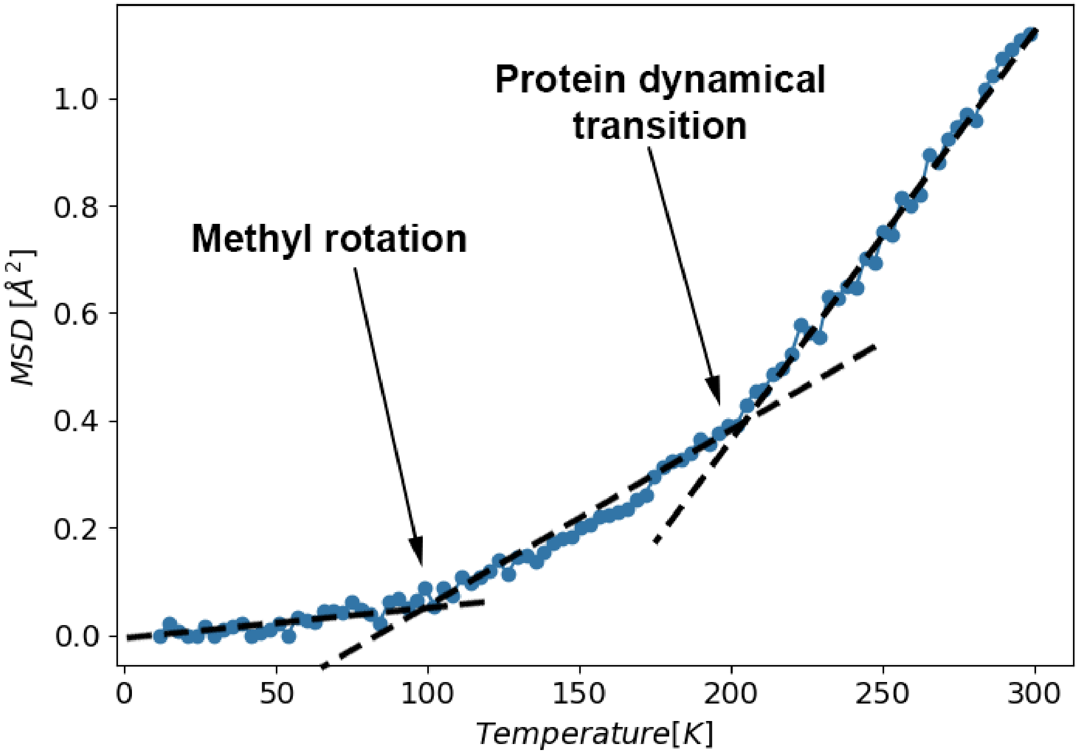


Figure 1.6: Dynamical transition resulting from anharmonic motions activation in tau protein, seen in the apparent mean-squared displacement (MSD).
Typical experimental result obtained with elastic fixed windows scans (EFWS) obtained on a neutron backscattering spectrometer are shown in blue - here, tau protein. Dashed lines were added for better visualisation of the transitions.

we can see that for slow motions, with $E_i \ll k_\mathrm{B}T$, p_i will be higher than for $E_\mathrm{i} \gg k_\mathrm{B}T$. As $-\log(p_\mathrm{i})$ increases very fast with decreasing argument compared to p_i, the contribution to entropy, and therefore to free energy will be much higher for fast motions than for slow motions. This observation shows that entropy resulting from fast dynamics will contribute significantly to conformational stability.

Turning our attention back to amyloid fibrils, it appears that they are highly stable structures in their mature form. We shall then expect some changes in fast dynamics. Some interesting effects were observed with concanavillin A on mean-squared displacements, depending on whether the protein was in native, amorphous aggregated or fibrillated state [Schiró et al., 2012]. Indeed, dynamics is enhanced when amyloid fibrils are formed, as compared to the native form. However, amorphous aggregates show similar behaviour - however with slightly lower MSD -, while this form is considered highly stable as well.

Several other studies showed that dynamics barely differs, or even not at all, between monomers and fibrils, we can cite work on lysozyme [Fujiwara et al., 2013], on α-synuclein [Fujiwara et al., 2016], or tau [Fichou et al., 2015b]. This might be surprising at first, but we did not mention an essential component of the system so far, hydration water. We will review its role in the next section.

1.3 Water and protein aggregation

Water is essential for life, and the original matrix in which it appeared. It possess numerous peculiar properties - its amphoteric nature, its electric dipole moment, its geometry, its phases to name a few - that render it one of the most challenging, and yet extremely interesting systems to study, despite its apparent simplicity. Moreover, it furnishes electrons and protons to what might be the most important process for life, photosynthesis, and it is directly involved in numerous biological processes. Most of water aspects in biology were thoroughly reviewed by Philip Ball [Ball, 2008], we will summarize some of them here.

1.3.1 Water as a solvent

Water presents some differences with simple liquids, for which molecules are described as hard-spheres, interacting via direct contacts and isotropic potential. Their capacity to form hydrogen bonds actually makes interactions highly oriented, and the inter-atomic distances distribution shows an additional peak, not present in simple liquids, typical for this type of interaction. Geometrically, liquid water is highly dynamic, and several defects can be found in the tetrahedral geometry, that affect positively the mobility of the molecule [Sciortino et al., 1991] [Sciortino et al., 1992]. In contrast, below the freezing point, the tetrahedral configuration is more stable, leading to a lower density and larger volume.

When hydrophilic solutes are added, especially ions, the high dielectric constant of water makes it very efficient at reorienting so that the ions are correctly solvated. This effect tends to disrupt the water structure only up to the first hydration shell [Omta, 2003]. Conversely, the effect of hydrophobic solutes is not completely explored. Kauzmann proposed a model, in which water molecules are arranged such that they form a cage around the solute in a highly ordered manner [Kauzmann, 1959]. Aggregation of hydrophobic particles would then be a way to reduce order in the water structure, and thus, increase entropy. However, this model was strongly challenged in studies, in which the perturbation of the water bonding network was shown to be weak and probably mostly located in the second hydration layer than the first [Dixit et al., 2002] [Finney et al., 2003]. More recently, simulations showed an important rotational anisotropy for water molecules within 10 Å of the surface of proteins ([Shin and Willard, 2018a] [Shin and Willard, 2018b] and fig. 1.7).

Considering the aforementioned results, the entropic increase upon protein folding and aggregation remains plausible, but its origin is a bit more subtle and complicated than postulated by Kauzmann. The situation is not clearer when considering water on large hydrophobic - or hydrophilic - surfaces, although rotational dynamics is still very likely to play a role [Ball, 2008].

1.3.2 The protein hydration shell

A simplistic model of protein folding contains contributions from the internal hydrogen bonding network, as proposed by Pauling in the late 1930ies [Mirsky and Pauling, 1936], and contributions from hydrophobic interactions in line with Langmuir's suggestions [Langmuir, 1939]. For globular proteins, this usually results in a hydrophobic pocket protected from the solvent by hydrophilic residues that are in direct contact with water. The surrounding water differs from bulk, with several works suggesting a higher density, not only in the first layer but up to 10 Å from the protein surface [Levitt and Sharon, 1988][Badger, 1993] [Perticaroli et al., 2017] [Svergun et al., 1998]. Moreover, relaxation times for rotational dynamics can be much slower in the hydration shell as compared to bulk water [Likić and Prendergast, 2001]. In addition, a molecular dynamics simulation study showed that the diffusion coefficient perpendicular to surface is more affected than the parallel one, and this effect depends mostly on surface geometry rather than specific amino acids [Makarov et al., 1998]. More

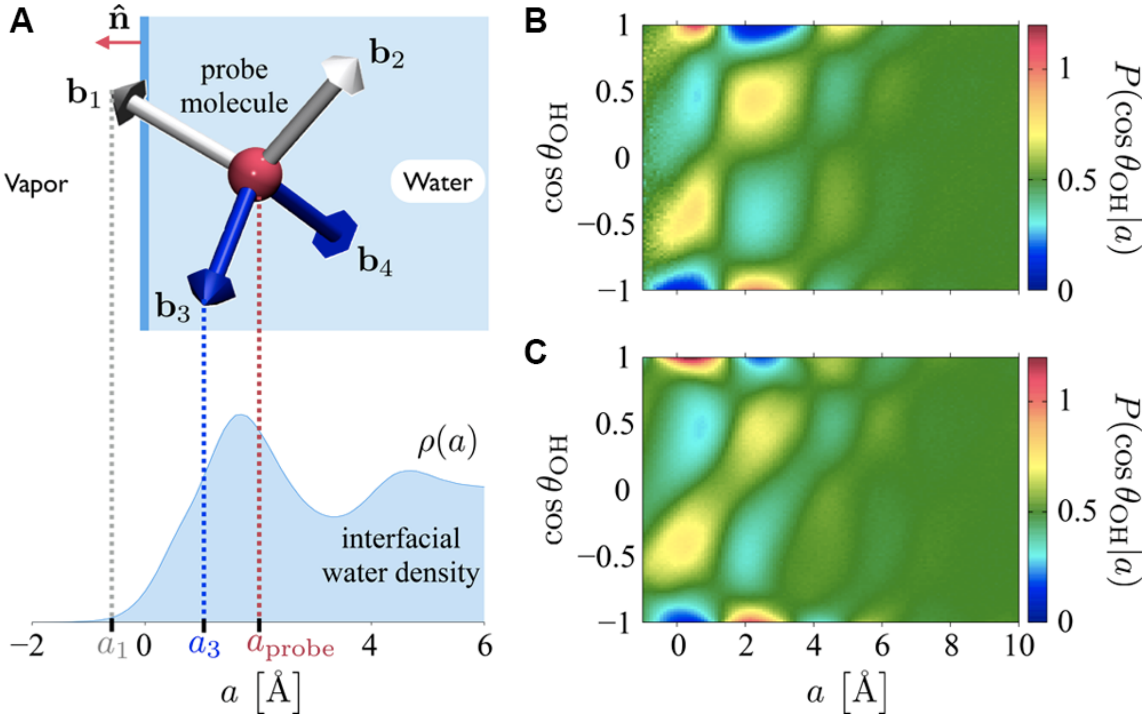


Figure 1.7: Water in hydration shell shows a high rotational anisotropy
A, Schematic depiction of the mean-field model showing a probe molecule with tetrahedrally coordinated bond vectors (white for donor, blue for acceptor) within the liquid (blue shaded region) at a distance a, probed from the position of the instantaneous interface (solid blue line). A plot of the interfacial density profile, ρ(a), obtained from the atomistic simulation with TIP5P water, is shown with dotted lines indicating the termination points of bond vectors b1 and b3. Panels **B** and **C** contain plots of the orientational distribution function, P(cos θ_{OH}|a), as indicated by shading, computed from atomistic simulation and from the rigid tetrahedral model, respectively.
Reproduced from [Shin and Willard, 2018b].

recently, the hydration shell was thoroughly investigated using molecular dynamics simulations. The authors showed an important broadening in the rotational correlation time for water in the hydration shell as compared to bulk - with slower times being more populated, associated with a more oriented dipole moment - consistent with higher anisotropy [Mukherjee et al., 2017]. Other parameters show an interesting behaviour, that is, the effective dielectric constant increases and the local specific heat decreases as we move away from the protein surface.

The protein hydration shell can be very stable. Indeed, γS-crystallin has evolved such that it can remain folded and monomeric over an entire lifetime. A recent study, making use of electron paramagnetic resonance, showed that the hydration layer is barely affected upon protein aggregation for wild-type γS-crystallin, while the aggregation-prone G18V mutant presents a substantially different hydration shell [Huang et al., 2016]. Hence, hydration water can influence protein structure and stability, but also the protein can affect the hydration water, as the composition in amino acids of proteins influences the hydration layer dynamics.

Interestingly, it has been shown that NMR and incoherent neutron scattering can be well suited to study water within living cells (or intracellular water) [M. Stadler et al., 2016] [Jasnin et al., 2010] [Martinez et al., 2016]. By studying the water dynamics within red blood cells, it appears that a small fraction - 10% - of water is slowed down as compared to the bulk water, which constitutes the other fraction [M. Stadler et al., 2016]. The slowed down fraction was attributed to the water located at the cell surface. Moreover, by comparing the intracellular water dynamics in *Escherichia coli* and the extreme halophile *Haloarcula marismortui*, it was observed that a significant fraction of water molecules exhibits a lower dynamics in the halophile [Jasnin et al., 2010]. The decreased dynamics was shown to be independent on the NaCl or KCl concentration. Hence, it can be attributed to biological components of the cell, such as proteins or glycans. The evolution of macromolecules towards their current folding and organization within the cell is likely to have been partly driven by an increase of water entropy, where the influence of macromolecules on water would lead to ever more favorable entropy as they evolve.

In conclusion, water can influence protein folding and dynamics as well as protein can influence water in its hydration shell, thereby showing that an important coupling exists. The coupling between protein and water has been well documented these last decades, in particular from the dynamical point of view. We will now discuss some of the results obtained so far.

1.3.3 Coupling to protein dynamics and aggregation

As it might shine through the previous section, studying water itself is quite challenging, and studying how it couples to protein motions might be even worse. Nonetheless, interesting results have been obtained thus far, mainly thanks to neutron scattering and computer simulations. It became clear that the protein dynamical transition was linked to water. Indeed, hydration has a significant effect on the amplitude of this transition, by allowing atoms to undergo anharmonic motions [Ferrand et al., 1993] [Steinbach and Brooks, 1993] [Tarek and Tobias, 2000] [Tarek and Tobias, 2008]. In addition, using a dual heat bath with simulations showed that activation of solvent translational motions drives the dynamical transition for bacteriorhodopsin [Tarek and Tobias, 2002] [Tournier et al., 2003]. More recently, the hydrogen bond (HB) network of water around lysozyme was studied using simulations [Rahaman et al., 2017]. It was found that the configurational entropy increases with temperature, and is reflected by the number of states the HB network can access.

New experimental evidences arose by having access to hydration water self-dynamics with neutron scattering. Indeed, it was observed that proteins can be functional even in the powder state, with at least 0.4 g of water per gram of dry protein [Rupley and Careri, 1991]. This allows to study the dynamics of the hydration shell only using neutron scattering, removing all contributions from bulk water, which would otherwise completely dominate the signal. Consistent with simulation data, the PDT was observed experimentally with maltose binding protein [Wood et al., 2008]. However, a different behaviour was found for a membrane protein, for which lipid dynamics might equally influence protein motions [Wood et al., 2007] [Gallat et al., 2012b]. The coupling between protein and its environment is summarized in figure 1.8.

The first transition at 110 K was shown to be independent on hydration [Hong et al., 2013]. The origin of the PDT at 230 K was further explored using neutron scattering applied to samples of poly-peptide chains containing either only glycine, alanine or various residues. This approach provided direct evidence that side-chain motions activation is coupled with the onset of water translational motions [Schiró et al., 2010]. Moreover, this effect is not only present at the protein surface, where amino acids are in direct contact with water, but is also observed in the protein core [Wood et al., 2013]. In addition, a transition was also observed with the cytochrome P450 and the green fluorescent protein [Liu et al., 2017], which indicates that the coupling with water motions might be of major importance but is not the only effect at play for anharmonic motions in proteins.

More recently, the protein aggregation process was studied using the protein tau. It has been unexpected to find that amyloid fibril formation does not correlate with

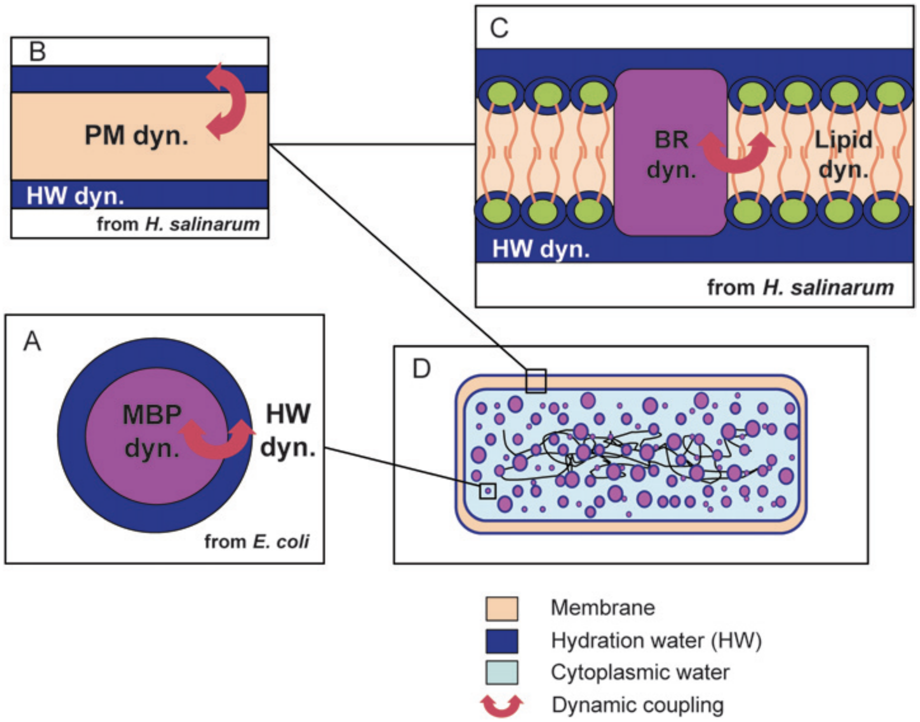


Figure 1.8: Dynamical coupling between water and proteins
A, Hydration around a soluble protein and coupling with protein dynamics (PM - Purple membrane). **B,** Hydration around a membrane not coupled to membrane protein dynamics (BR - Bacteriorhodopsin). **C,** Dynamic coupling between protein and lipid in the membrane (MBP - Maltose binding protein). **D,** Channels of free water flowing around the hydrated macromolecules in the cell.
Reproduced from [Frölich et al., 2009]

any change in protein internal dynamics, especially when we consider the extreme stability of amyloid fibrils. The origin of aggregate stability is to be found in hydration water. Indeed, hydration translational motions are enhanced in tau aggregates [Schirò et al., 2015], which was interpreted as an increased entropy of water [Fichou et al., 2015b]. How the cross-β structure influences hydration water is not completely clear. It appears that the fibril core - that is, the central hydrophobic region of the protein - alone is not sufficient to explain the hydration water mobility increase [Fichou et al., 2015a]. Moreover, it was found, using *in-vivo* NMR, that α-synuclein presents a very globular shape in cells, with the hydrophobic core being protected from the solvent by flanking disordered regions [Theillet et al., 2016]. The compact folding of α-synuclein and tau was also demonstrated with using experimental distance constraints with simulations [Brodie et al., 2019] [Popov et al., 2019]. The dramatic structural change upon aggregation highlights the interplay between water entropy increase and protein structure that can influence hydrophobic interaction. More thorough studies are still needed to better understand the physical origin(s) of protein aggregation and neutron scattering in combination with molecular dynamics simulation provides important insights in that regard.

As an example, along with the favorable entropic contribution of water upon protein aggregation, neutron scattering and simulations have shown that hydrodynamic interactions can strongly influence protein diffusion and aggregation in solution [Chiricotto et al., 2016] [Grimaldo et al., 2019]. Indeed, the combination of neutron incoherent scattering and simulations proved that the ps-ns self-diffusion of macromolecules is significantly affected if their radius differs from the average radius R_{eff} of all molecules in solution. But the self-diffusion is similar to that of a monodisperse solution if the radius is close to R_{eff} [Grimaldo et al., 19ed]. Moreover, using the OPEP force field with lattice Boltzmann molecular dynamics, an extensive study of β-amyloid peptide (16-22) oligomerization and aggregation was conducted [Chiricotto et al., 2016]. The study showed that hydrodynamic effects promote the fusion of oligomers and exchanges between them. In addition, an elongated and twisted structure could be observed. These two studies reveal that hydrodynamic interactions can strongly influence the protein self-diffusion and large protein aggregate formation.

1.4 Metal ions and aggregation

1.4.1 Metals and biology

An increasing number of studies show that metal ions can play various essential roles in biology, among which there is a structural role, by stabilizing a specific fold or oligomer. Notably, bivalent cations seem to be prevalent. The most studied system so far is likely to be calmodulin and its binding to Ca^{2+}. Indeed, calmodulin is central to various cellular processes, and calcium can act on it by regulating its subcellular location, its structure, and its interaction with other proteins [Chin and Means, 2000]. It has been found as well that pH can also be a key factor for metal induced refolding of superoxide dismutase [Fee and Phillips, 1975]. It appears that metal ions can affect proteins at pH well below their pKa, where proper refolding of superoxide dismutase is more efficient when occuring at low pH, in presence of copper and zinc. Another example is hemoglobin, for which an iron ion is coordinated by an exogenous molecule within the protein, the heme [Perutz et al., 1960]. The heme allows the metalloproteins to transport various molecules such as oxygen or carbon dioxide. Numerous other examples could be cited, including cytochrome, iron-sulphur centres, zinc fingers. Many of them were thoroughly reviewed already [Holm et al., 1996]. Notwithstanding the beneficial effects of metal ions in biology, they have also their counterpart, as they can have deleterious effects on the organism. Iron, for instance, can bridge a molecule to oxygen, and promote the formation of super-oxide that can damage various essential components in the cell [Welch et al., 2002]. Thus, the level of iron has to be finely regulated, so that induced oxidative stress remains low. In addition, metals might affect protein aggregation.

1.4.2 Metals and protein aggregation

Metal ions have been proposed as risk factors for Alzheimer's disease, with more or less confidence depending on the species involved [Frederickson, 1989] [Tomljenovic, 2011]. While *in vivo* studies are challenging, due to the difficulty to obtain reliable statistics on patients, the effect on aggregation is very clear *in vitro*. Indeed, the fibrillation kinetics of β-amyloid peptide, involved in Alzheimer's, is increased by a factor of 100 to 1000 when aluminum, iron or zinc are present in solution [Mantyh et al., 1993]. Interestingly, it seems that β-amyloid peptides can play a protective role against metal induced oxidative stress [Zou et al., 2002]. On the other hand, the very same metals can make the β-amyloid peptide lose its protective function, and eventually become neurotoxic. Thus, there is a subtle balance here, that is likely to be metal homeostasis dependent.

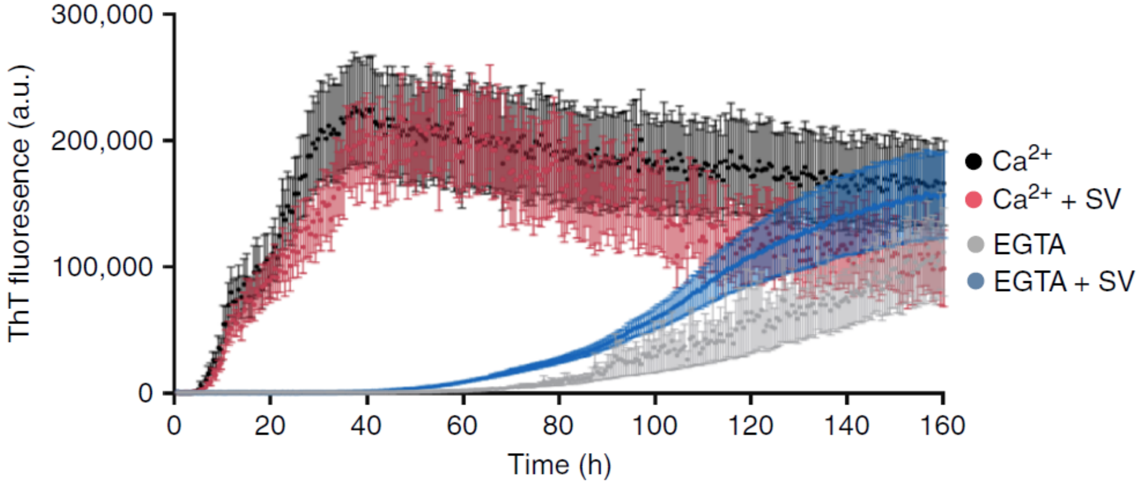


Figure 1.9: Calcium significantly accelerates α-synuclein fibrillation
α-synuclein aggregation measured by ThT fluorescence using 100 µM monomeric alpha-synuclein under shaking conditions. The presence of 2.5mM calcium increased the aggregation kinetics of α-synuclein compared to 1mM EGTA, both in the presence or absence of synaptic vesicles (black and red). For the EGTA-containing groups, there was a trend towards faster aggregation of α-synuclein in the presence of synaptic vesicles (blue vs. gray). Three biological repeats, n= 18 for all conditions, where n represents single wells. The values represent mean ± s.e.m. Adapted from [Lautenschläger et al., 2018]

Similarly, a substantial effect of copper was observed on α-synuclein fibrillation. The cation was found to promote pore-forming aggregated species formation, which are highly toxic for cells *in vitro* [Zhang et al., 2015]. Also, calcium bound to α-synuclein accelerates the amyloid fibrillation, promotes synaptic vesicle lipid binding, and a deviation from physiological metal concentration induces synaptic vesicles clustering (fig. 1.9) [Lautenschläger et al., 2018].

The effects of metal ions on aggregation is of major importance, yet are still barely studied, especially when we consider fast dynamics and hydration water. We still lack information about how they bind for some proteins like α-synuclein or insulin, and what the effects are on protein clustering, on dynamics and stability. It is likely that neutron scattering studies, in combination with X-ray absorption spectroscopy and other spectroscopic methods might provide valuable insights regarding these questions. A first step toward the resolution of the effects of multi-valent ions in protein solution was made using neutron scattering in a study that shows a significantly affected protein diffusion coefficient depending on the presence of metal [Grimaldo et al., 2015a].

1.5 Concluding remarks on protein aggregation

In summary, amyloid protein aggregation is a process involving a great diversity of factors and physical phenomena. Considering the amyloid fibrillation, a lag phase is usually observed at the beginning of the fibrillation kinetics and might be strongly dependent on protein clustering, and more generally on electrostatic and hydrodynamic effects that can promote or prevent the formation of pre-fibrillar oligomers. In addition, conformational changes can be a limiting step that can explain the initial lag phase, as the gain in stability - or the decrease of free energy - obtained by forming the cross-β structure, will influence the kinetics of fibrillation.

If the kinetics of the amyloid aggregation has been widely studied, the origin of protein aggregation remains elusive, especially regarding the internal dynamics of the protein and its hydration water shell. The results can differ from one protein to another, such as concanavalin and tau. Indeed, differences in fast internal dynamics were measured in concanavalin [Schiró et al., 2012] upon aggregation, suggesting a favorable energetic change coming from the protein internal motions, while no difference appeared at all for tau [Fichou et al., 2015b].

Regarding the hydration shell, the entropy gain seems to be of major importance as observed in the case of the protein tau. However, we still do not know how different parts of the protein are involved, in particular, how the proportions of fuzzy coat and fibril core can influence the dynamics of hydration water upon amyloid fibrillation.

Along with the protein properties, metal ions can significantly influence aggregation kinetics [Lautenschläger et al., 2018], but little is known about their effect on protein and water dynamics.

We still need to determine how both intrinsic - protein sequence - and extrinsic - metal ions, salt concentration - factors act on the aggregation and the ensemble-averaged dynamics.

1.6 Scope and objectives of the thesis

In this work, we have focused on the changes that occur in protein and hydration water dynamics during the amyloid aggregation of various proteins, α-synuclein, insulin, γS-crystallin, lysozyme and tau. We made use of incoherent neutron backscattering spectroscopy, molecular dynamics simulations, X-ray powder diffraction, dynamic light scattering and X-ray absorption fine structure measurements. Our experimental results concern mainly the ensemble-averaged dynamics, and the local, atomic-scale information is brought by the simulations. We used hydrated protein powders in either the monomeric or fibrillar state. The powder state strongly facilitates the analysis and the interpretation of data at the cost of losing the information about the center-of-mass self-diffusion that is suppressed in powders. In addition, the powder state is the only way hydration water can be studied with neutron scattering as the signal would be completely dominated by the bulk water in the liquid state. However, we have started to study protein dynamics in liquid state as well, to follow the aggregation in a time-resolved manner using the so-called elastic/inelastic fixed-window scans (E/IFWS).

Among the proteins studied, we focused on the hydration water dynamics of α-synuclein, which is involved in Parkinson's disease and for which we aimed at performing an enlightning comparison with the protein tau previously studied in the DYNAMOP group at the IBS [Fichou et al., 2015a] [Fichou et al., 2015b]. Moreover, we took advantage of the published full-length structure of α-synuclein [Tuttle et al., 2016] to explore the effect of the fibrillation process on hydration water dynamics in atomistic details using simulations.

We also studied γS-crystallin, which is involved in the eye-lens cataract and which can form not only amyloid fibrils, but also other types of aggregates that can be amorphous or contain a cross-β structure. The γS-crystallin also presents the advantage of being easily produced in large quantities, i.e. hundreds of milligrams. This choice of protein allows us to investigate different aggregation pathways and their effect on the protein and hydration water dynamics. In addition, the mutant G18V, associated with early-onset cataract, is studied to determine how the dynamics is affected by the point mutation. The main difficulty that arises with γS-crystallin is that the different types of aggregates can form concurrently. These aggregates can then be challenging to isolate to study the different types of aggregates separately.

In addition to α-synuclein and γS-crystallin, we studied insulin, which presents a medical interest as well but that is also studied for its potential use as a biomaterial. In particular, we aimed at better understanding how the zinc cations can affect the insulin internal dynamics and its aggregation kinetics. Moreover, insulin is commercialy available, which greatly facilitates the sample preparation of the hydrogenated samples. However, similarly to γS-crystallin, insulin can form different types of aggregates in a concurrent way.

Furthermore, as the transient oligomeric species that occur on the amyloid aggregation pathway are of prime interest, we explored the possibility to measure incoherent

neutron scattering data in a time-resolved manner on liquid protein samples that are undergoing aggregation. To this purpose, we use the model system, lysozyme, for which aggregation is well studied and which is highly reproducible. This first time-resolved experiment was to be followed by another, for which we used a medically more relevant system, the protein tau. Our aim was to capture both the center-of-mass self-diffusion and protein-internal dynamics of any transient oligomeric species that appears during the course of amyloid aggregation.
More generally, we want to obtain a detailed information on the changes in protein and hydration water dynamics that occur during amyloid aggregation, and on the effect of various intrinsic factors - protein sequence, aggregation pathway - and extrinsic factors, such as metal ions. In addition, this work might provide a general framework to study amyloid aggregation in the light of the protein and water dynamics, and thus of the energy changes, and, also pave the way to new diagnostic methods of amyloid diseases based on the dynamics of hydration water [Walton et al., 2011].

1.7 My contribution to the various projects

We mentioned the different projects and the associated objectives in the previous section, these are (1), the dynamics of the hydration water of α-synuclein monomers and amyloid fibers, (2), the dynamics of γS-crystallin monomers and aggregates, (3), the effect of zinc on the dynamics of insulin monomers and aggregates, (4), the time-resolved study of the formation of lysozyme particulates and (5), the time-resolved study of the heparin-triggered fibrillation of tau.
For (1), I produced and purified the protein with Martine Moulin at the D-Lab, I optimized the fibrillation conditions in 1.5 mL eppendorf tubes and prepared the samples for neutron scattering experiment and biophysical characterizations. At the MLZ neutron reactor in Garching, Germany, I collected the ENS and the QENS data at 300 K in May 2018 and Giorgio Schirò collected the QENS data at 260 K in February 2020. I analyzed the neutron scattering data under the guidance of Giorgio Schirò. To this purpose, I wrote a Python application programming interface (API) (nPDyn - github.com/kpounot/nPDyn) that can handle various types of data from neutron backscattering spectroscopy. Concerning the MD simulations, I spent three months at the laboratory of Douglas Tobias at UC Irvine, USA, where I built the powder model for monomers and fibers based on several discussions we had with Douglas Tobias and Martin Weik and I then performed the MD simulations following an established procedure used previously for tau [Fichou et al., 2015a]. To analyze the MD simulations data, I have written a Python/C++ API (NAMDAnalyzer - github.com/kpounot/NAMDAnalyzer), for which the procedure used and the

results were validated by Douglas Tobias. The X-ray powder diffraction measurement was performed at the ESRF synchrotron by Giorgio Schirò and myself and I analyzed the data using the pyFAI library [Kieffer and Wright, 2013]. The article was written by Martin Weik and myself, with input from all other authors (section 3.1).

For (2), I have hydrated the protein powders for the neutron scattering experiments at the MLZ and I performed the neutron experiment in September 2017 with Giorgio Schirò and the one from September 2018 on my own. I analyzed the data using the nPDyn API and wrote the chapter in this manuscript (section 3.2).

For (3), I determined the conditions to obtain pure spherulite or pure fiber samples under the supervision of Vito Foderà during a stay of two weeks in his laboratory at the University of Copenhagen, Denmark. I prepared the samples for the neutron experiment in September 2017 and the neutron scattering data were collected by Giorgio Schirò and myself. I have suggested to investigate the role of the zinc atom and I prepared new samples for this purpose. The neutron sacttering data on the new samples were collected by Giorgio Schirò on June 2018 at the MLZ and I analyzed the data. The XAFS measurements were performed by Alessandro Longo, Giorgio Schirò and myself. I built the structural model that Alessandro Longo used for XAFS data analysis. I have performed the circular dichroïsm, the fluorescence microscopy and the DLS experiments. The article was written by Giorgio Schirò and myself, with input from all other authors (section 3.3).

For (4), the idea of performing a time-resolved E/IFWS measurement came up during a discussion between Tilo Seydel and myself. The experimental conditions (protein and salt concentrations, pH, temperature) were determined by Hussein Chaaban and myself. I performed the E/IFWS experiment in June 2018 at the ILL with Tilo Seydel and I analyzed the data under the supervision of Tilo Seydel. I performed the DLS experiment. I also built a model of lysozyme in solution and performed the MD simulations to obtain the structural model I used to compute the theoretical center-of-mass diffusion coefficient at 90°C. The theoretical center-of-mass diffusion coefficient at 7°C was obtained by myself using a structure of native lysozyme from the PDB (code 3ijv). The article was written by Tilo Seydel, Martin Weik and myself, with input from all other authors (section 3.4).

For (5), I assisted Ninon Zala in the production and the purification of the protein. I optimized the experimental conditions (protein concentration, buffer, pH, temperature) for fibrillation with Yann Fichou. I performed the time-resolved E/IFWS experiment end of January 2020 at the ILL with Tilo Seydel and I analyzed the data. I suggested a role for the aluminum in our neutron experiments and I tested my hypothesis in a first ThT kinetics experiment. The corresponding chapter in this manuscript was written by myself (section 3.5).

Methods

2.1 Quasi-elastic neutron scattering

Neutrons are fermions that can be produced using either nuclear fission or spallation. The former makes use of fuel elements consisting of heavy fissile atoms. Several sources are operating on this principle, among them we can cite the Institut Laue-Langevin in France, or the FRM-II in Germany. Spallation consists in hitting a target, e.g. mercury, with accelerated protons. The impact allows to disintegrate the nucleus, which produces several particles, including neutrons. Some sources operating on this principle are the Spallation neutron source (SNS) in the USA, ISIS in the UK, and in a near future the European Spallation Source in Sweden.
Neutrons possess a magnetic moment and are electrically neutral, which allow them to penetrate deeply into materials, and probe various properties such as dynamics and structures of atoms or spins. Moreover, neutrons can be produced at various energies, depending on the moderator used (table 2.1). The use of cold neutrons ($\simeq$ 2 meV energy) allows to match the energies encountered in proteins, providing access to both spatial and temporal information. Cold neutrons also have the advantage of preserving chemical bonds due to their low energy.

Moderator	Energy [meV]	Temperature [K]	Wavelength λ [10^{-10}m]
cold	0.1-10	1-120	30-3
thermal	5-100	60-1000	4-1
hot	100-500	1000-6000	1-0.4

Table 2.1: Energy, temperature and wavelength ranges from three types of sources. Adapted from [Squires, 1996]

.

2.1.1 Scattering cross-section

Following the book by Squires [Squires, 1996] or the one by Bée [Bee, 1988], we consider a beam of neutrons passing through a sample, which is basically a collection of atoms, from which the neutron wave can be scattered in a given direction with a given energy exchange between the neutron and the atom. The general principle of neutron scattering is to count the number of neutrons being scattered into a solid angle Ω with energy E'. This gives the partial differential cross-section,

$$\frac{d^2\sigma}{d\Omega dE'}\ \Phi\ d\Omega dE' = \begin{array}{l}\text{number of neutrons scattered per second}\\ \text{into solid angle d}\Omega,\ \text{and with}\\ \text{energy in the interval } [E',\ E'+dE'].\end{array} \tag{2.4}$$

where Φ is the flux of incident neutrons, as the number passing through a flat area, perpendicular to the beam propagation vector.
Integrating over energy, we can obtain the total scattering cross-section, which is the total number of neutrons scattered per second, divided by incident flux. Assuming isotropic scattering for angle ϕ, this quantity is therefore given by

$$\sigma_{\text{tot}}\ \Phi = \int_{\text{all space}} d\Omega \int_0^\infty dE'\ \frac{d^2\sigma}{d\Omega dE'} = \int_0^\pi d\theta\ 2\pi\sin(\theta)\ \frac{d\sigma}{d\Omega} \tag{2.5}$$

All atoms in the sample will contribute to this quantity, but we should expect them to scatter neutrons with various efficiencies. That is, for a given nucleus, if we consider an incoming wave propagating along z given by

$$\Psi_{\text{in}} = e^{ikz} \tag{2.6}$$

then, the scattered wave will be spherical with a scaling factor b such that

$$\Psi_{\text{out}} = -\frac{b}{r}e^{ikr} \tag{2.7}$$

The factor b is known as the scattering length and is to be determined experimentally. It varies erratically from one nucleus to another. If we consider elastic scattering, the number of neutrons passing through unit area per unit time is

$$v\ dS|\Psi_{\text{out}}|^2 = v\ dS\ \frac{b^2}{r^2} = v\ b^2\ d\Omega \tag{2.8}$$

Using (2.4), with $\Phi = v|\Psi_{\text{in}}|^2$, we have

$$\frac{d\sigma}{d\Omega} = \frac{vb^2 d\Omega}{v d\Omega} = b^2 \tag{2.9}$$

and the total cross-section $\sigma_{\text{tot}} = 4\pi b^2$. Thus, the scattering length is directly related to elastic scattering probability for a given nucleus.

2.1.2 Incoherent scattering function

The system under study can exchange kinetic energy with the incoming neutron. This can be expressed by the transition rate for a given state $|k, \lambda\rangle$ to turn into the final state $|k', \lambda'\rangle$, with k being the neutron state and λ the nucleus state. We use the so-called *Fermi's golden rule*, which states that

$$\sum_{\mathbf{k}' \text{ in } \Omega} W_{\mathbf{k},\lambda \to \mathbf{k}',\lambda'} = \frac{2\pi}{\hbar} \rho_{\mathbf{k}'} | \left\langle \mathbf{k}'\lambda' \right| V \left|\mathbf{k}\lambda\right\rangle |^2 \tag{2.10}$$

It can be shown that this expression yields the following for the differential cross-section [Squires, 1996],

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\lambda \to \lambda'} = \frac{k'}{k} \left(\frac{m}{2\pi\hbar^2}\right)^2 | \left\langle \mathbf{k}'\lambda' \right| V \left|\mathbf{k}\lambda\right\rangle |^2 \delta(E_\lambda - E_{\lambda'} + E - E') \tag{2.11}$$

Here, the nucleus is considered as a point in space, therefore, the potential V takes the the form of the Fermi pseudopotential

$$V(\mathbf{r}) = \frac{2\pi\hbar^2}{m} b\delta(\mathbf{r}) \tag{2.12}$$

with r being the distance between the neutron and the nucleus.
Using this potential, representing the delta on energy by an integral on a complex exponential, and taking the averaged scattering length for the system, it can be shown that the double differential cross-section for coherent scattering is given by [Squires, 1996]

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\text{coh}} = \frac{\sigma_{\text{coh}}}{4\pi} \frac{k'}{k} \frac{1}{2\pi\hbar} \sum_{\text{j}\neq\text{j}'} \int_{-\infty}^{\infty} \langle e^{i\boldsymbol{q}(r_{\text{j}}(t) - r_{\text{j}'}(0))} \rangle e^{i\omega t} dt \tag{2.13}$$

and is related to collective dynamics and mostly used with neutron crystallography for example. While incoherent scattering can be described by

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\text{inc}} = S(\boldsymbol{q},\ \omega) = \frac{\sigma_{\text{inc}}}{4\pi} \frac{k'}{k} \frac{1}{2\pi\hbar} \sum_{\text{i}} \int_{-\infty}^{\infty} \langle e^{i\boldsymbol{q}(r_{\text{i}}(t) - r_{\text{i}}(0))} \rangle e^{i\omega t} dt \tag{2.14}$$

and is related to self dynamics, with $\sigma_{\text{coh}} = 4\pi\overline{b}^2$, $\sigma_{\text{inc}} = 4\pi(\overline{b^2} - \overline{b}^2)$, **q** = **k'** - **k**, and r_{i} denotes the position of nucleus i.
Therefore, incoherent scattering will be the one interesting for us, as self dynamics is the one we probe using quasi-elastic neutron scattering.
Some values for scattering cross-sections are given in table 2.2. We can readily see that hydrogen has a significantly higher cross-section than other nuclei. This feature can be used to probe specifically some parts of the system using specific deuteration. Indeed, using perdeuterated protein and hydrating the powder with H_2O, all contributions but the one from hydration water will be considered negligible, such that only the dynamics of water is modelled, and conversely.

Nucleus	Z	σ_{coh} $[10^{-28}m\check{s}]$	σ_{inc} $[10^{-28}m\check{s}]$
^1H	1	1.8	80.2
^2H	2	5.6	2.0
C	6	5.6	0.0
N	7	11.01	0.5
O	8	4.2	0.0
Al	13	1.5	0.0

Table 2.2: Scattering cross-section for experimentally relevant nuclei
Adapted from [Squires, 1996].

.

2.1.3 Instrumentation

Backscattering spectroscopy has come a long way to reach its actual performance, in terms of flux, resolution and possibilities of measurement schemes. Originally proposed by Maier-Leibnitz, and developed at Munich research center FRM, the technique makes use of Bragg's law to obtain an efficient monochromatisation of the incident beam and of the scattered neutrons. This can be achieved with good success using certain types of crystals for backscattering. For example, Si(111) hexagonal single crystals plane, mounted spherically around the sample, allow to achieve a resolution with full width at half maximum (FWHM) of 0.85 μeV [Frick, 2002].
Several improvements were made over the years, starting with neutron guides. Indeed, neutrons have a very low total reflection critical angle, which make them difficult to guide and focus without losses. The development of super-mirrors, which consist of multi-layered deposit of material, that evanescent-wave reflection can be used, was a major improvement in neutron optics [Hayter and Mook, 1989]. From

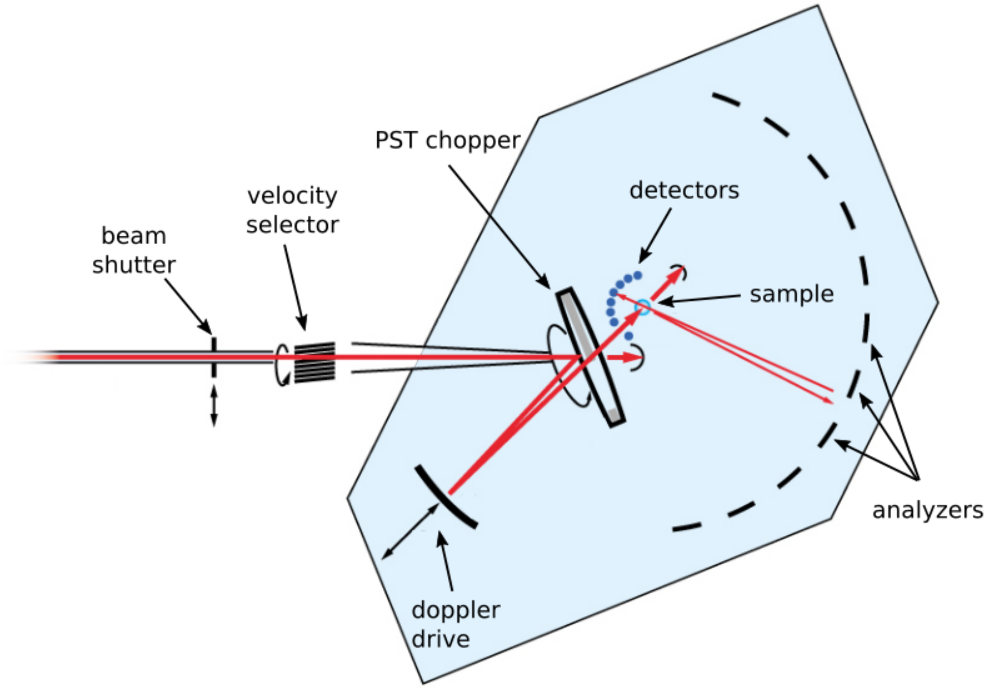


Figure 2.1: SPHERES backscattering spectrometer
Typical layout representative of a reactor-based backscattering spectrometer
Adapted from https://www.mlz-garching.de/spheres

then, focusing guides and incident flux were greatly improved.
A typical backscattering instrument layout is depicted in figure 2.1. As it can be seen, a converging guide, on which some choppers and/or a so-called velocity selector (not shown) can be found for wavelength selection and background lowering, leads neutrons towards a phase space transformation (PST) chopper. Both SPHERES at MLZ, Münich and IN16B at ILL, Grenoble use also a velocity selector between the instrument shutter and converging guide (see [Frick et al., 2010]).
The PST chopper was designed to address the problem that the beam scattered toward the monochromator presents a phase space volume that does not fully match the accepted volume for backscattering [Schelten and Alefeld, 1984]. Therefore, the high resolution of this kind of instrument is obtained at the cost of the flux, as a significant part of incoming neutrons is simply not backscattered toward the sample. Schelten and Alefeld proposed this solution, which consists in rotating graphite mosaic crystals. Due to the transformation from the laboratory reference frame to the

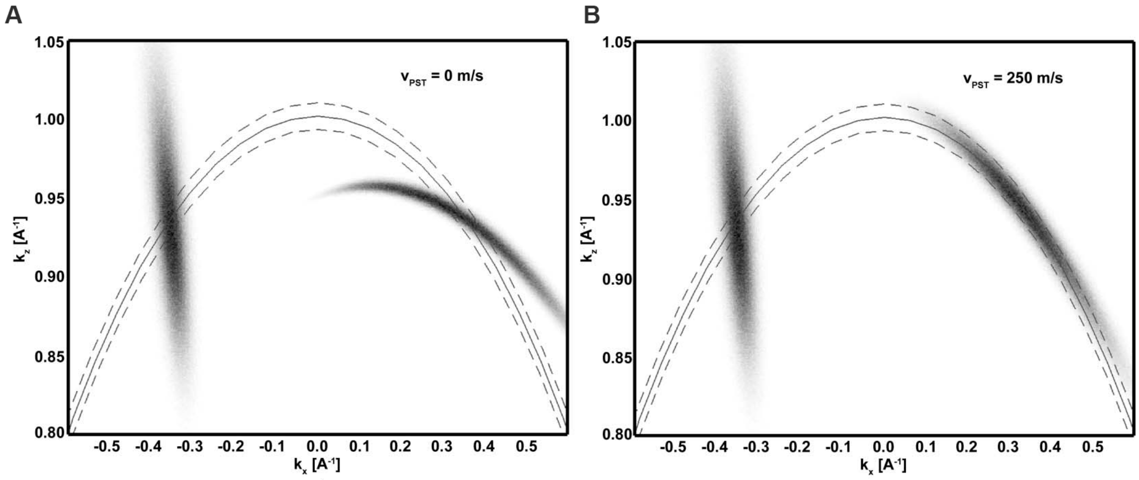


Figure 2.2: Simulated incident and reflected intensity in momentum space
Plotted is the two-dimensional projection of a phase-space element prior (left clouds) and subsequent (right clouds) to the reflection from a mosaic crystal at rest (**A**) and at $v_{\mathrm{PST}} = 250$ m/s (**B**), respectively. The point density is proportional to the neutron density. The solid line denotes the elastic line at the monochromator, and the dashed lines mark the limits of the acceptance range of the moving monochromator.
Adapted from [Hennig et al., 2011]

moving frame, it results in the fact that higher energy neutrons will be scattered at lower energy, and conversely. Thus, the momentum dispersion on one axis is made much lower, the phase space volume will be transformed as in figure 2.2, such that more neutrons will match the backscattering conditions. This component allows to increase the incident flux by a factor comprised between 4 and 5 [Hennig et al., 2011].

The monochromator is a crystal mounted on a Doppler drive. This permits to move it back and forth using different velocity profiles, which, by the Doppler effect, induces a small shift in energy of backscattered neutrons. On most instruments, only a sinusoidal profile can be used, but recently, the IN16B spectrometer gained the possibility to use custom profiles, including constant velocity, which allows to perform fixed window scans (FWS) to measure specific energy transfer between the neutron and the system [Frick et al., 2012]. After passing through the sample, neutrons with a defined energy are backscattered by the analysers towards detectors. In the elastic case, the monochromator is at rest, and only neutrons that are elastically scattered by the system are analysed and detected. While when the monochromator is moving, only neutrons that have exchanged the right amount of energy with the system, so that they match the analysis conditions, are detected.

2.1.4 Modelling the elastic scattering signal

If we consider only the elastic scattering signal, useful information can already be extracted. Using equation 2.14, the correlation function in the integrand can be expanded, considering the actual measured signal by convolution with the resolution function $S_{\text{exp}}(q,\ \omega) = S(q,\ \omega) \otimes R(q,\ \omega)$, to obtain [Roosen-Runge and Seydel, 2015]

$$\begin{aligned} S(q,\ \omega)_{\text{exp}} = \hat{R}(\omega) & \\ & - \frac{q^2}{2\sqrt{2\pi}} \int dt e^{-i\omega t} \langle \Delta r^2 \rangle (t) R(t) \\ & + \frac{q^4}{24\sqrt{2\pi}} \int dt e^{-i\omega t} \langle \Delta r^4 \rangle (t) R(t) \\ & + \mathcal{O}(r^6) \end{aligned} \tag{2.15}$$

In line with the Debye-Waller factor $e^{-\langle u_{\text{vib}}^2 \rangle q^2/3}$, the mean-squared displacement can be obtained by

$$\begin{aligned} \langle MSD \rangle_\omega &= - \lim_{\text{q}\to 0} \frac{3}{q^2} \log \left[S_{\text{exp}}(q,\omega)/\hat{R}(\omega) \right] \\ &= \frac{3}{2\sqrt{2\pi}} \int dt e^{i\omega t} \langle \Delta r^2 \rangle (t) R(t)/\hat{R}(\omega) \end{aligned} \tag{2.16}$$

This is the most general mean-squared displacement. In the case where we are interested only in the elastic signal, $\omega = 0$, and the equation becomes

$$\langle MSD \rangle_0 = \frac{3}{2\sqrt{2\pi}} \frac{1}{\hat{R}(0)} \int dt \langle \Delta r^2 \rangle (t) R(t) \tag{2.17}$$

We see that the MSD provides a model-free indicator of the type of dynamics that strongly depends on the time window given by the resolution function. This indicator can be used with both elastic and inelastic fixed-window scans.
Other methods were proposed [Yi et al., 2012] [Kneller and Hinsen, 2009], the simplest of which makes use of a simple Gaussian, being valid only on a limited q-range.

Using cumulant expansion on the auto-correlation function, it is possible to derive either a simple Gaussian expression for the MSD, or a bit more complicated one that includes also higher terms in the expansion [Yi et al., 2012], thereby giving the following,

$$S(q,\ \omega = 0) \simeq e^{\frac{1}{6}q^2\langle\Delta r^2\rangle}\left(1 + \frac{q^4}{72}\sigma^2\right) \tag{2.18}$$

where σ^2 is the variance of $\langle\Delta r^2\rangle$ over all atoms. This model was successfully applied to α-lactalbumin to highlight the heterogeneity of motions in hydrated powder samples [Zeller et al., 2018]. Also, a gamma distribution can be used to take into account the heterogeneity of motions in the sample [Peters and Kneller, 2013] [Zeller et al., 2020]. In practice, the data to be fitted can be noisy and exhibit few features. Therefore the fitting process can quickly become unstable when the model gets more complicated. The Gaussian model presents the advantage to give a simple, and quick way to compare dynamics of different samples, but it is limited to momentum transfer $\mathrm{q}^2\langle\mathrm{r}^2\rangle < 1$, where $\langle r^2\rangle$ is the mean square displacement.

2.1.5 Modelling quasi-elastic spectra

Resolution function

Let us consider again the measured signal $S_{\mathrm{exp}}(q,\ \omega) = S(q,\ \omega) \otimes R(q,\ \omega)$. The resolution function $R(q, \omega)$ gives the minimum energy exchange that can be probed by the instrument, as any line shape will be broadened after convolution with it. This function takes several parameters into account and its expression will strongly depend on the instrument set-up. On SPHERES at the MLZ, or on IN16B at the ILL, the resolution function is mostly Gaussian. And common models consist of a sum of two Gaussians, or a pseudo-Voigt profile (Gaussian + Lorentzian) which reads

$$R(\boldsymbol{q}, \omega) = e^{-q^2\langle u^2_{\mathrm{vib}}\rangle}\left[d\frac{1}{\pi}\frac{\gamma(\boldsymbol{q})}{\gamma^2(\boldsymbol{q}) + \omega^2} + (1-d)\frac{1}{\Gamma(\boldsymbol{q})\sqrt{2\pi}}e^{-\omega^2/2\Gamma^2(\boldsymbol{q})}\right] \tag{2.19}$$

where the leading term is the Debye-Waller factor, γ is the Lorentzian linewidth, Γ is the Gaussian linewidth, and d is a scalar in the interval [0, 1].
In addition, detector efficiency, as well as the Debye-Waller factor, can be included in a leading scaling factor C, which is fitted for each detector independently. Thus, the expression in brackets in the above equation being normalized, the scaling factor provides the total incoherent scattering signal. This quantity is particularly useful for data normalization. Eventually, the final model also includes a q-dependent

background term B($\mathbf{q}$), which takes into account fast motions that are not probed by the energy transfer range of the instrument, so that the final model reads

$$R(\boldsymbol{q},\omega) = S(\boldsymbol{q}) \left[d\frac{1}{\pi}\frac{\gamma(\boldsymbol{q})}{\gamma^2(\boldsymbol{q})+\omega^2} + (1-d)\frac{1}{\Gamma(\boldsymbol{q})\sqrt{2\pi}}e^{-\omega^2/2\Gamma^2(\boldsymbol{q})} + B(\boldsymbol{q}) \right] \tag{2.20}$$

Even though the background term is not normalized here, its magnitude is usually at least three orders of magnitude lower than 1, such that it does not contribute much to the integral, and S remains a reliable normalization factor. The fitting procedure can be performed using vanadium data, which has the advantage to scatter almost exclusively incoherently and shows a pure elastic peak modulated by the Debye-Waller factor.

Brownian diffusion

As it can be seen from equation 2.14, we are basically looking at correlations of atom positions at different times. Thus, we need to understand how atoms move to correctly model the experimental data. Protein dynamics involves motions on broad time, and length, scales. Using hydrated protein powders simplifies analyses, as we do not have to deal with global diffusion.

For powders, dynamics can be modelled using a Debye-Waller factor for fast vibrational motions for which energy exchanged lies beyond the energy transfer range probed by the instrument, which multiplies other energy dependent factors.

In the case of Brownian diffusion, the self correlation function obeys the diffusion equation [Fick, 1855],

$$\frac{\partial}{\partial t}G_{\mathrm{s}}(\boldsymbol{r},t) = D\nabla^2 G_{\mathrm{s}}(\boldsymbol{r},t) \tag{2.21}$$

Following Einstein's random walk model, this is equivalent to

$$\tau\frac{\partial}{\partial t}G_{\mathrm{s}}(\boldsymbol{r},t) = \frac{\langle l^2\rangle}{6}\nabla^2 G_{\mathrm{s}}(\boldsymbol{r},t) \tag{2.22}$$

with τ being the time between two collisions, and l the distance a particle travels during this time. Thus $\langle l^2\rangle/6\tau$ can be identified as the diffusion coefficient D. Requiring the self correlation to be normalized, and delta distributed at t=0, a solution to the above is

$$G_{\mathrm{s}}(\boldsymbol{r},t) = \frac{1}{(4\pi Dt)^{3/2}}e^{-\boldsymbol{r}^2/4Dt} \tag{2.23}$$

The Fourier transform of this equation in both space and time then yields a Lorentzian of width $D\boldsymbol{q}^2$. Thus, taking also into account slow motions that are within the resolution function, the final model reads [Stoeckli et al., 1986]

$$S(\boldsymbol{q},\omega) = e^{-\boldsymbol{q}^2\langle u_{\mathrm{vib}}^2\rangle} R(\boldsymbol{q},\omega) \otimes \left[A_0\delta(\omega) + \sum_{\mathrm{i}} \frac{A_{\mathrm{i}}}{\pi} \frac{D_{\mathrm{i}}\boldsymbol{q}^2}{(D_{\mathrm{i}}\boldsymbol{q}^2)^2 + \omega^2} \right] + B(\boldsymbol{q}) \tag{2.24}$$

where D_{i} is the i^{th} atom diffusion coefficient, ω is the energy offset, A_0 the elastic incoherent structure factor (EISF), A_{i} a contribution factor for the quasi-elastic signal from atom i, and B(**q**) a background term. In principle, the contribution of each atom in the system is implicitly contained in the above formula, but in practice, experimental limitations are such that only one or two terms should be used. This can be proven using Bayesian theory, by taking the posterior probability for a model to correctly fit the data as a function of the number of terms in the sum, we have [Sivia et al., 1992],

$$P(N|d) \propto \frac{N!}{(\omega_{\mathrm{max}} A_{\mathrm{max}})^N} \frac{(4\pi)^N e^{-\chi^2_{\mathrm{min}}/2}}{\sqrt{Det(\nabla\nabla\chi^2)}} \tag{2.25}$$

where $p(N|d)$ is the probability that N Lorentzians can fit the data d, ω_{max} and A_{max} are maximum energy offset and intensity, respectively, and χ^2 is the cost function used for fitting. The result of this analysis is shown in figure 2.3. It can be readily seen that beyond two Lorentzians, one cannot tell which model best fits the data, a result that strongly suggests over-fitting. It can be shown that Γ in 2.24 is linear with respect to the **q** vector, with $\Gamma = D\boldsymbol{q}^2$ [Springer, 1977]. Thus, neutron scattering offers a powerful way to probe the global diffusion coefficient for atoms that are following the aforementioned model. It actually corresponds to Fickian diffusion type. However, atoms can deviate significantly from free diffusion, because of crowding or confinement. A more involved analysis for translational motions inside an impermeable sphere of radius a was performed by Volino and Dianoux [Volino and Dianoux, 1980]. They showed that the actual signal strongly deviates from the $D\boldsymbol{q}^2$ law when the (q.a) product becomes greater than one.

There is still ongoing work to go beyond these models [Kneller and Calandrini, 2007] [Kneller and Chevrot, 2012] [Kneller, 2018]. In the present work, we have made use mostly of the so-called jump-diffusion model [Singwi and Sjölander, 1960], when analysing data from liquid samples.

Jump diffusion model

The jump diffusion model is based on the assumption that particles undergo oscillatory motion for a given time t, followed by a diffusive motion for a time t',

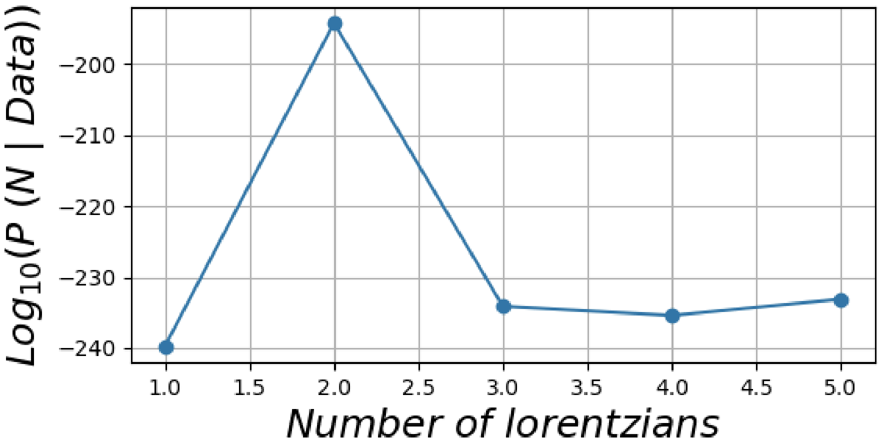


Figure 2.3: Example of bayesian analysis giving a model with two Lorentzians as the most reliable
The aforementioned bayesian posterior probability formula was used with different model to obtain the probability for each that they fit our data, taking into account the deviation from the data, and model complexity.
This plot is obtained from data acquired in the lab.

then an oscillatory motion again, and so on and so forth [Singwi and Sjölander, 1960]. The derivation goes as follows, taking the following quantities,

- $g(\boldsymbol{r}, t)$ is the probability to find a particle at position **r** at time t when it is performing oscillatory motions, starting from origin at t=0.
- $p(t)$ is the probability that a particle remains in the same oscillatory state after time t.
- $h(\boldsymbol{r}, t)$ is similar to $g(\boldsymbol{r}, t)$, but for diffusive motions.
- $q(t)$ is the probability that a particle, starting from diffusive motions, stays in the same state during time t.

The probability to find a particle at position **r** after time t, provided that it was at the origin at t=0 can be found by integrating step by step, which gives,

$$\begin{aligned}
F_0(\boldsymbol{r},t) &= g(\boldsymbol{r},t)p(t) \\
F_1(\boldsymbol{r},t) &= -\int_0^t dt_1 \int d\boldsymbol{r_1}\ q(t-t_1)h(\boldsymbol{r-r_1}, t-t_1)p'(t_1)g(\boldsymbol{r_1},t_1) \\
F_{\mathrm{2N}}(\boldsymbol{r},t) &= -(1^{2N}) \int_0^t dt_{\mathrm{2N}} \int_0^{t_{\mathrm{2N}}} dt_{\mathrm{2N-1}} \int_0^{t_{\mathrm{2N-1}}} dt_{\mathrm{2N-2}} \cdots \int_0^{t_2} dt_1 \int d\boldsymbol{r}_{\mathrm{2N}} d\boldsymbol{r}_{\mathrm{2N-1}} \cdots d\boldsymbol{r}_1 \\
&\quad \times p(t-t_{\mathrm{2N}})g(\boldsymbol{r-r_{2N}}, t-t_{\mathrm{2N}})q'(t_{\mathrm{2N}}-t_{\mathrm{2N-1}})h(\boldsymbol{r_{2N}}-\boldsymbol{r_{2N-1}}, t_{\mathrm{2N}}-t2N-1) \\
&\qquad \cdots p'(t_1)g(\boldsymbol{r_1},t_1)
\end{aligned} \tag{2.26}$$

Taking the Fourier transform and using the substitution $t - t_{\mathrm{2N}} = \tau_{\mathrm{2N+1}}, \cdots, t_2 - t_1 = \tau_2,\ t_1 = \tau_1$, this can be reduced to a simple form,

$$\begin{aligned}
\int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} F_{\mathrm{2N}}(\boldsymbol{r},t) &= AC^N D^N \\
\int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} F_{\mathrm{2N+1}}(\boldsymbol{r},t) &= BC^{N+1} D^N
\end{aligned} \tag{2.27}$$

with

$$\begin{aligned}
A &= \int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} p(t)g(\boldsymbol{r},t) \\
B &= \int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} q(t)h(\boldsymbol{r},t) \\
C &= \int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} p'(t)g(\boldsymbol{r},t) \\
D &= \int_0^\infty dt \int d\boldsymbol{r} e^{i(\boldsymbol{q}.\boldsymbol{r}-\omega t)} q'(t)h(\boldsymbol{r},t)
\end{aligned} \tag{2.28}$$

Therefore, the full scattering cross-section, using $\frac{\tau_0}{\tau_0+\tau_1}$ for the fraction of particles starting at time t=0 by oscillatory motion and $\frac{\tau_1}{\tau_0+\tau_1}$ for particles starting with diffusive motion, is given by [Singwi and Sjölander, 1960]:

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{\sigma_{\mathrm{inc}}}{2\pi}\frac{k}{k_0}\left\{\frac{\tau_0}{\tau_0+\tau_1}\left(\frac{A+BC}{1-CD}\right)+\frac{\tau_1}{\tau_0+\tau_1}\left(\frac{B+AD}{1-CD}\right)\right\} \tag{2.29}$$

To describe the positional probability of particles $g(\boldsymbol{r},t)$ and $h(\boldsymbol{r},t)$, Singwi and Sjölander used a Gaussian, with time dependent width, t. Concerning probabilities for particles to stay in the same state, an inverse exponential law was used. In the case $\tau_1 \gg \tau_0$, that is when diffusive motions prevail, the simple Lorentzian form is retrieved, with a width following the Dq^2 law.
For $\tau_1 \ll \tau 0$, the line shape still appears Lorentzian, but the width is now given by

$$\gamma(\boldsymbol{q}) = \frac{\boldsymbol{q}^2 D}{1+\boldsymbol{q}^2 D\tau_0} \tag{2.30}$$

where D can be related to the apparent diffusion coefficient. This model was used to analyse data from liquid protein solution samples, with both full QENS spectra and inelastic fixed window scans (IFWS). We took into account global diffusion, following Fick's law and described by the parameter $\Gamma = \boldsymbol{q}^2 D$, which was convoluted with the expression for the internal protein dynamics, containing the jump diffusion behaviour for oscillatory prevailing motions. Thus, the final model reads,

$$S(\boldsymbol{q},\omega) = R(\boldsymbol{q},\omega)\otimes\left\{S\left[a_0\mathcal{L}_\Gamma(\boldsymbol{q},\omega)+(1-a_0)\mathcal{L}_{\Gamma+\gamma}(\boldsymbol{q},\omega)\right]+S_{\mathrm{D_20}}\mathcal{L}_{\mathrm{D_2O}}\right\} \tag{2.31}$$

where a_0 is the EISF, S and $S_{\mathrm{D_2O}}$ are scalars, giving contributions from protein and D_2O solvent respectively, $\mathcal{L}_\Gamma$ is the Lorentzian describing global diffusion, $\mathcal{L}_{\Gamma+\gamma}$ describes internal dynamics, $\mathcal{L}_{\mathrm{D_2O}}$ the solvent line shape, and $R(\boldsymbol{q},\omega)$ is the resolution function. This model has been used successfully to determine diffusion coefficients in crowded environments [Grimaldo et al., 2014] [Grimaldo et al., 2015b][Beck et al., 2018].

Rotational motions

When water motions are observed, it is common that not only translational motions are modelled, but also rotational ones. It is worth to mention that both types of motions cannot always be distinguished. Rotational and translational motions can be separated only when they occur in sufficiently different length, and time scale. The analysis of rotational motions can be performed for hydration water. The process is simplified under the assumption that translational and rotational motions are uncorrelated, which implies that the intermediate scattering function can be factorized as [Springer, 1977]

$$I_{\mathrm{s}}(\boldsymbol{q},t) = I_{\mathrm{rot}}(\boldsymbol{q},t) I_{\mathrm{trans}}(\boldsymbol{q},t) \tag{2.32}$$

The measured signal is then a convolution between the scattering law for rotations and the Lorentzian shaped scattering law for translational diffusion, plus a Lorentzian term accounting for internal dynamics.
In the case of hydrated powders, the model is simplified, as we do not have to deal with global diffusion, such that the model consists of a convolution between resolution function and the following

$$S(\boldsymbol{q},\omega) = a_{\mathrm{rot}}(\boldsymbol{q}) S_{\mathrm{rot}}(\boldsymbol{q},\omega) \otimes a_{\mathrm{trans}}(\boldsymbol{q}) \mathcal{L}_{\gamma}(\boldsymbol{q},\omega) \tag{2.33}$$

where the first term on the right has to be determined, and a gives the relative contribution of both types of motions to the signal.
Scattering cross-sections for rotations were thoroughly studied for diatomic molecules by Sears [Sears, 1966a] [Sears, 1966b] [Sears, 1967]. The resulting model for water can be understood as free rotors, which exhibit angular momentum l. In the assumption of isotropic scattering, only the radial component affects the scattering and we are left with an expansion of Lorentzians weighed by spherical Bessel function,

$$S_{\mathrm{rot}}(\boldsymbol{q},\omega) = j_0^2(qb)\delta(\omega) + \sum_{\mathrm{l}=1}^{N} (2l+1) j_{\mathrm{l}}^2(qb) \frac{1}{\pi} \frac{l(l+1)\Gamma_{\mathrm{rot}}}{\omega^2 + [l(l+1)\Gamma_{\mathrm{rot}}]^2} \tag{2.34}$$

where $j_{\mathrm{l}}(qb)$ are Bessel functions, and b is the distance from center of mass for rotation around it, which is set to the O-H distance for water. Moreover, it was found that rotational and translational contributions can be deconvolved, such that the convolution in equation 2.33 becomes a sum. Thus, the model used to study hydration water around proteins reads [Schirò et al., 2015]

$$S(\boldsymbol{q},\omega) = \; e^{-q^2\langle u^2\rangle} R(\boldsymbol{q},\omega) \otimes [S_{\mathrm{trans}}(\boldsymbol{q},\omega) + S_{\mathrm{rot}}(\boldsymbol{q},\omega)] + b(\boldsymbol{q}) \tag{2.35}$$

where $b(\boldsymbol{q})$ is a q-dependent background. This model has shown good agreement with the quasi-elastic signal analysis and simulations in terms of conclusions that can be drawn for water behaviour around amyloid proteins [Schirò et al., 2015] [Fichou et al., 2015b]. All of the mentioned models so far can be fitted to experimental data, with good robustness regarding starting parameters.

2.1.6 Data analysis

In order to perform data treatment and analysis, I have developed an application programming interface (API), called nPDyn (see github.com/kpounot/nPDyn), and written in Python/C. It consists of several classes, which can handle data from either QENS, temperature ramp EFWS, or IFWS. Thus far, .inx and hdf5 files can be read. Concerning the latter, it is assumed that the so-called 'unmirroring' of the Doppler-signal was performed with Mantid [Arnold et al., 2014] already. Subsequently, resolution function, D_2O signal, and empty cell signal can be fitted using either an already existing model from the API or a user-defined model. All of these can be used for corrections. Functions are available for quick corrections, such as empty cell subtraction or Paalman-Ping corrections. The latter makes use of existing C code by Joachim Wuttke, libabsco [Wuttke, 2012], for which a Python wrapper was created so that it can be used directly within the API.

A model can be assigned to each loaded dataset and fitted. The models already present in the API all use a basin-hopping minimization procedure [Wales and Doye, 1997]. The idea of basin-hopping - as used in the Scipy basinhopping method - is sketched in fig. 2.4. It basically consists in performing several iterations of the following sequence. First, randomly chose a set of initial parameters. Second, use of a minimization procedure to find the local minimum of the cost function, which can be defined as a χ^2 distance, for example. Third, compare this minimum with the previously obtained one using the Metropolis criterion. If the test is validated, the new minimum is defined as being the global minimum. The algorithm can rather quickly find best parameters, and provides a powerful way to optimize a model while being quite independent of initial parameters.

Finally, nPDyn offers also several functions to plot the fitted data, and can be used with the NAMDAnalyzer package (github.com/kpounot/NAMDAnalyzer) to compare neutron experimental data with results from molecular dynamics simulations.

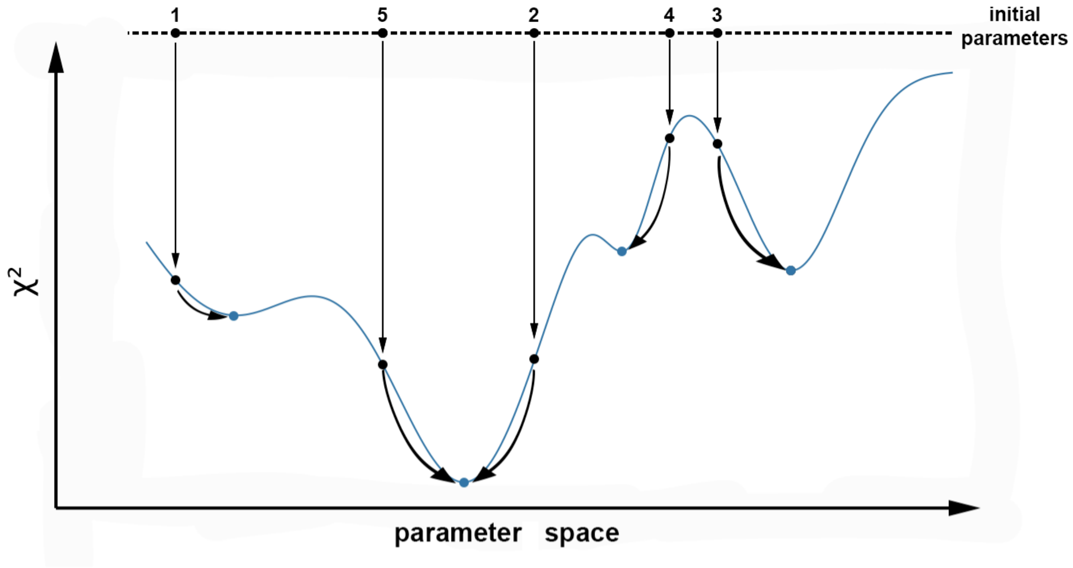


Figure 2.4: Illustration of the principle of basin-hopping minimization
The χ^2 distance - (model - experimental)^2 - is plotted as a blue line as a function of model parameters. The numbers refer to the iteration sequence, where a set of initial parameters is randomly chosen, then a minimization procedure is applied so that the local minimum is reached. The next iteration starts with a new set of initial parameters. The new minimum is compared with the previous one using the Metropolis criterion, and kept as the global minimum in case the test is validated. Figure rendered in Python and modified with Adobe Photoshop.

2.2 Molecular dynamics simulations

The principles governing the dynamics and electronic structure of atoms are known since the beginning of the 20^{th} century. However, for atoms bigger than helium, and a fortiori molecules, the problem becomes too complex to be solved exactly. With the advent of computers, it has been possible to calculate, under some simplifying assumptions, the electronic structure of various small molecules. This was done using quantum mechanical principles within the framework of the Hartree-Fock matrix and Roothan-Hall equation [Leach, 2001]. More recently, density functional theory has gained attention as a good alternative way to compute electronic structures. However, these quantum mechanical methods are too demanding in terms of computational power when we want to deal with macromolecules.

Setting aside electrons, molecular mechanics models were developed, considering only nuclei positions, and making use of classical physics to compute energy and motions.

2.2.1 Equation of motion

The basis of molecular dynamics simulations is to use Newton's second law,

$$\frac{d^2}{dt^2}\boldsymbol{r}_{\mathrm{i}} = \frac{F_{\mathrm{i}}}{m_{\mathrm{i}}} \tag{2.36}$$

where $\boldsymbol{r}_{\mathrm{i}}$ is the i^{th} atom position, F_{i} the force acting on it and m_{i} its mass. Using a proper potential function, the force can be easily obtained by derivation with respect to coordinates. Thus, the strategy is to use small time steps to integrate the second law. Given an atom at a certain position at time t, we can, using the computed force, determine the acceleration to be applied on, and obtain its position and velocity at time $t + \delta t$. The most widely used algorithm for this purpose is probably the Verlet algorithm [Verlet, 1967], and subsequent improvements. The algorithm makes use of the following expansions,

$$\begin{aligned} \boldsymbol{r}(t+\delta t) &= \boldsymbol{r}(t) + \delta t\dot{\boldsymbol{r}}(t) + \frac{\delta t^2}{2}\ddot{\boldsymbol{r}}(t) + \mathcal{O}(\delta t^3) \\ \boldsymbol{r}(t-\delta t) &= \boldsymbol{r}(t) - \delta t\dot{\boldsymbol{r}}(t) + \frac{\delta t^2}{2}\ddot{\boldsymbol{r}}(t) + \mathcal{O}(\delta t^3) \end{aligned} \tag{2.37}$$

Adding these therefore leads to an expression giving the new position in terms of known quantities,

$$\boldsymbol{r}(t+\delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t-\delta t) + \delta t^2\ddot{\boldsymbol{r}}(t) \tag{2.38}$$

Then, velocities can be obtained from using

$$\dot{\boldsymbol{r}}(t) = \boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t-\delta t)/\delta t \tag{2.39}$$

Various similar algorithms do exist, some of them might be more accurate. An errorless algorithm would perfectly give energy conservation throughout the simulation. However, because of the truncated expansion, the computation does give rise to an erroneous result. A solution to this problem is to compare the predicted acceleration from the previous step using Taylor expansion with the computed one at the current step. The $\Delta\ddot{\boldsymbol{r}}$ can then be used to apply a simple linear correction for current values of position, velocity and acceleration.

Furthermore, various motions, with different time scales are involved in biological molecules, many of which do not require the usual 1 or 2 femtosecond time steps, but could be very well approximated using larger time steps. To address this point, a commonly used algorithm is known as r-RESPA [Tuckerman et al., 1992], which consists in a reversible application of time translation operators on the system in the form

$$\Gamma(t) = e^{i(L_1+L_2+L_3+L_4)t}\Gamma(0) \tag{2.40}$$

where $\Gamma(t)$ is the system state at time t, and L_i denotes the Liouville operator associated with a given term i in the potential energy function,

$$iL = \sum_{\mathrm{i}=1}^{3N}\left[\frac{\partial x_\mathrm{i}}{\partial t}\frac{\partial}{\partial x_\mathrm{i}} + F_\mathrm{i}(x)\frac{\partial}{\partial p_\mathrm{i}}\right] \tag{2.41}$$

with x_i and p_i denoting position and conjugate momentum of atom i, respectively, and F_i the force acting on it. Each type of force comes with a given time step, being a multiple of the smallest one. The standard Verlet algorithm is retrieved if all time steps are equal. This method substantially improves simulation performances, and shows a good stability, even for long runs.

In all the aforementioned algorithms, we see that we always require to compute the force. This is done by means of a defined force field, which I will review now.

2.2.2 Force field

When we consider atoms in macromolecules, we can look at several parameters, such as bond lengths, angles, torsions, to name a few. These parameters usually come with an equilibrium value and atoms are oscillating around it. Various force fields are available now, but they usually all contain at least the following terms,

$$\begin{aligned} V(\boldsymbol{r}) = &\sum_{\text{bonds}}\frac{k_\mathrm{i}}{2}(l_\mathrm{i}-l_{\mathrm{i},0})^2 + \sum_{\text{angles}}\frac{k_\mathrm{i}}{2}(\theta_\mathrm{i}-\theta_{\mathrm{i},0})^2 + \sum_{\text{torsions}}\frac{V_\mathrm{n}}{2}(1+\cos(n\omega-\gamma)) \\ &+ \sum_{\mathrm{i,j=i+1}}^{N}\left(4\epsilon_{\mathrm{ij}}\left[(\sigma_{\mathrm{ij}}/r_{\mathrm{ij}})^{12}-(\sigma_{\mathrm{ij}}/r_{\mathrm{ij}})^{6}\right] + \frac{q_\mathrm{i}q_\mathrm{j}}{4\pi\epsilon_0 r_{\mathrm{ij}}}\right)\end{aligned} \tag{2.42}$$

The first term corresponds to the bond length, with an equilibrium length $l_{\mathrm{i},0}$ for an i atom pair. The second term is similar, but for bond angles - such as the H-O-H angle in a water molecule for instance. The third term deals with rotations of a group around a bond axis, and the last one contains Coulomb and Lennard-Jones potentials for electrostatic and van der Waals interactions, respectively. Cross-terms

	SPC	SPC/E	TIP3P	BF	TIP4P	ST2
r(OH) [Å]	1.0	1.0	0.9572	0.96	0.9572	1.0
HOH angle [deg]	109.47	109.47	104.52	105.7	104.52	109.47
$A.10^{-3}$ [kcal Å^{12}/mol]	629.4	629.4	582.0	560.4	600.0	238.7
C [kcal Å^{6}/mol]	625.5	625.5	595.0	837.0	610.0	268.9
q(0)	-0.82	-0.8472	-0.834	0.0	0.0	0.0
q(H)	0.41	0.4238	0.417	0.49	0.52	0.2375
q(M)	0.0	0.0	0.0	-0.98	-1.04	-0.2375
r(OM) [Å]	0.0	0.0	0.0	0.15	0.15	0.8

Table 2.3: Geometry and potential parameters for several water models
Adapted from [Jorgensen et al., 1983]

can be used as well. For instance, stretching of two bonds with a definite angle between the two bond axes can be controlled using an expression of the form

$$V(l_1, l_2, \theta) = \frac{k_{l_1,l_2,\theta}}{2}\left[(l_1 - l_{1,0})(l_2 - l_{2,0})\right](\theta - \theta_0) \tag{2.43}$$

Concerning non-bonded interactions, the multipole expansion offers an accurate way to compute electrostatic interactions, including dipole and quadrupole moments. However, this method is usually limited to small molecules, and is not suited for intra-molecular interactions. Thus, electrostatic interactions are usually computed using point-charge models. The idea is to place partial charges on the molecule so that its properties can be correctly reproduced. In the case of the N_2 molecule, placing a charge -q at each nucleus center and a +2q charge at the center of mass permits to retrieve the neutrality, absence of dipole moment and presence of quadrupole moment of the molecule. For proteins, partial charges might be calculated and optimized for each amino acid separately, either using quantum mechanical calculations or other methods, such as electronegativity [Gasteiger and Marsili, 1980]. In addition, various models were proposed to compute effects such as anisotropy, polarisation, van der Waals forces, and others. Several of them are reviewed in Leach's book [Leach, 2001].

Nonetheless, water will be discussed, as it is of paramount interest for biology and it possesses very peculiar properties that make it challenging to model accurately. Several models were proposed, showing different accuracy depending on the conditions being investigated. The most commonly used models are reviewed in table 2.3. The most simple models, SPC, SPC/E and TIP3P show partial charges located at the center of each nucleus, and slightly differ by their geometry parameters. Four-site models, BF and TIP4P, present a negative charge shifted toward the hydrogen along the bisector of the HOH angle. Five-site models are also available, such as ST2, or more recently, TIP5P. Some properties are rather well described by all these models, such as density or enthalpy of vaporisation. Some others are model dependent such as phase transition, especially ice formation, which is better described using the TIP4P model [Matsumoto et al., 2002].

It is important to keep in mind that all force fields are always empirical, a general form is tried and parameters are optimized by comparison with experiment. Therefore, some might be more appropriate depending on the system being studied or the conditions used, and over-fitting might be a problem to take care of. Once the equation of motion and force field are set, it can be necessary to minimize the initial structure to avoid atom contacts and large energy terms due to bad geometry. Subsequently, simulations can be run.

2.2.3 Running the simulation

Minimization

Several algorithms exist to minimize the energy. A simple way to locate a minimum is the steepest-descent, which consists in using derivatives of energy with respect to parameters using the following expression,

$$\boldsymbol{v}_{\mathrm{k}} = -\boldsymbol{g}_{\mathrm{k}} + \gamma_{\mathrm{k}} \boldsymbol{v}_{\mathrm{k}-1} \tag{2.44}$$

with $\boldsymbol{g}_{\mathrm{k}}$ being the gradient in direction k, $\boldsymbol{v}_{\mathrm{k}-1}$ the displacement from the previous iteration in the direction k-1, perpendicular to k. The weighing factor γ is given by

$$\gamma_{\mathrm{k}} = \frac{\boldsymbol{g}_{\mathrm{k}} \boldsymbol{g}_{\mathrm{k}}}{\boldsymbol{g}_{\mathrm{k}-1} \boldsymbol{g}_{\mathrm{k}-1}} \tag{2.45}$$

Other methods make use of the second derivative. The best known example is the Newton-Raphson algorithm, where the derivative of the potential function is used along with its expansion

$$V'(x) = V'(x_{\mathrm{k}}) + (x - x_{\mathrm{k}}) V''(x_{\mathrm{k}}) \tag{2.46}$$

such that the minimum can be found at x* using

$$x^* = x_{\mathrm{k}} - \frac{V'(x_{\mathrm{k}})}{V''(x_{\mathrm{k}})} \tag{2.47}$$

Notwithstanding its accuracy, this method is computationally demanding because of the Hessian matrix computation. Thus, quasi-Newton methods are often used to reduce algorithm complexity. These methods are based on an approximation of the Hessian matrix using only current, next position and gradients. The most famous ones are known as Davidon-Fletcher-Powell (DFP), Broyden-Fletcher-Goldfarb-Shanno (BFGS), and Murtaugh-Sargent(MS) algorithms [Press et al., 1992].
Once energy is minimized, iterations can be started. From then on, there are several possibilities regarding the thermodynamic ensemble that is required.

Thermodynamic ensembles

Depending on the kind of study we want to carry out, we might desire to control some parameters to stick to a certain thermodynamic ensemble. A commonly used one is the isothermal-isobaric (NPT) ensemble, here both pressure and temperature are controlled. In the canonical ensemble (NVT), temperature and cell volume are maintained constant.
Regarding constant temperature simulations, we note that a system temperature is related to the average kinetic energy by the following equation:

$$\langle K\rangle_{\mathrm{NVT}} = \frac{3}{2}Nk_{\mathrm{B}}T \tag{2.48}$$

where k_{B} is the Boltzmann constant and N the number of particles. One way to control temperature is thus to rescale velocities using a factor $\lambda = \sqrt{T_{\mathrm{wanted}}/T_{\mathrm{current}}}$. Another possibility to control temperature is to couple the system to an external heat bath with fixed temperature - Berendsen's thermostat [Berendsen et al., 1984] -, where the scaling parameter is then given by

$$\lambda^2 = 1 + \frac{\delta t}{\tau}\left(\frac{T_{\mathrm{bath}}}{T_{\mathrm{sys}}} - 1\right) \tag{2.49}$$

where τ is a coupling parameter, the magnitude of which has to be determined such that the temperature is properly controlled without affecting too much velocities throughout the simulation.
However, the aforementioned temperature control methods do not generate a rigorous canonical ensemble, and other algorithms were proposed to control temperature more efficiently. A commonly used one is known as *extended system*, initially developed by Nosé [Nosé, 1984] and improved by Hoover [Hoover, 1985]. The basic idea is to add a new degree of freedom s to the Hamiltonian, which reads

$$\mathcal{H} = \mathcal{H}^0 + \frac{p_{\mathrm{s}}^2}{2Q} + gkT\ln(s) \tag{2.50}$$

where p_s denotes the momentum, Q is an artificial mass and g the number of degrees of freedom in the extended system. Introducing these terms, and switching to NVT ensemble, usually implies for the system to become non-Hamiltonian. That is, it does not satisfy the following relations anymore,

$$\begin{aligned} \dot{\boldsymbol{p}} &= \frac{\partial \mathcal{H}}{\partial \boldsymbol{x}} \\ \dot{\boldsymbol{x}} &= -\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}} = \boldsymbol{F}(\boldsymbol{r}) \end{aligned} \tag{2.51}$$

In consequence, the total energy is not necessarily conserved, but the Liouville equation is still satisfied, although it appears in a different form [Tuckerman and Martyna, 2000]. More recently, velocity rescaling methods regained some interest, and an algorithm was proposed by Bussi and coworkers [Bussi et al., 2007]. Therein, velocity rescaling was done using a factor $\alpha = \sqrt{\frac{K_\mathrm{t}}{K}}$ where K_t is kinetic energy drawn from canonical distribution

$$\bar{P}(K_\mathrm{t})dK_\mathrm{t} = K_\mathrm{t}^{N_\mathrm{f}/2-1} e^{\beta K_\mathrm{t}} dK_\mathrm{t} \tag{2.52}$$

where N_f is the number of degrees of freedom in the system and $\beta = 1/k_\mathrm{B}T$. Using stochastic dynamics methods, the change dK in kinetic energy can be obtained, which gives a form similar to Berendsen's thermostat plus an additional stochastic term, weighed by an arbitrary factor.

Along with temperature, pressure can also be kept constant such that the probed ensemble is the NPT. Very similarly to constant temperature, pressure can be controlled using an extended system by adding appropriate terms to the system Hamiltonian. This has led to the Nosé-Hoover-Langevin piston algorithm which was used to run simulations in this work [Martyna et al., 1994][Feller et al., 1995].

Periodic boundary conditions

If we consider a protein in a box filled with water molecules, two problems arise. First, water molecules can escape the box, and eventually end up as free particles in vacuum. Second, it might be difficult to deal with non-bonded forces. The problem of particles leaving the box is easily solved by making their image at opposite sides enter the box. For non-bonded forces, the issue comes from the slow convergence of summation to compute pairwise interactions for all atoms.

A typical solution is to use the Ewald sum employing the following identity

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r} \tag{2.53}$$

In addition, all charges are surrounded by a Gaussian distribution of neutralizing charges in real space, plus another distribution - of opposite charges from the first one - in reciprocal space. The complementary error function is usually used for f(r), and some correction terms are needed, such that the full expression is given by

$$V = \frac{1}{2}\sum_{\mathrm{i,j}=1}^{N} \left\{ \begin{array}{l} \sum_{|\mathrm{n}|=0}^{\infty} \frac{q_{\mathrm{i}} q_{\mathrm{j}}}{4\pi\epsilon_0} \frac{erfc(\alpha|\boldsymbol{r}_{\mathrm{ij}}+\boldsymbol{n}|)}{|\boldsymbol{r}_{\mathrm{ij}}+\boldsymbol{n}|} \\ \\ + \sum_{\mathrm{k}\neq 0} \frac{q_{\mathrm{i}} q_{\mathrm{j}}}{4\pi^2 L^3 \epsilon_0} \frac{4\pi^2}{k^2} e^{-k^2/(4\alpha^2)} \cos(\boldsymbol{k}.\boldsymbol{r}_{\mathrm{ij}}) \\ \\ - \frac{\alpha}{\sqrt{\pi}} \sum_{\mathrm{k}=1}^{N} \frac{q_{\mathrm{k}}^2}{4\pi\epsilon_0} + \frac{2\pi}{3L^3} \left| \sum_{\mathrm{k}=1}^{N} \frac{q_{\mathrm{k}}}{4\pi\epsilon_0} r_{\mathrm{k}} \right|^2 \end{array} \right. \tag{2.54}$$

Improvements were proposed, and other methods do exist to efficiently compute non-bonded interactions. Some of them are reviewed by Leach [Leach, 2001].
Being now all set, the simulation can be run. In this work, NAMD [Phillips et al., 2005] was used to this purpose, with the CHARMM36 force field [Huang et al., 2013], and TIP4P water model, which was shown to give results in good agreement with experimental data regarding dynamics, structure factor and phase changes [**?**] [Sanz et al., 2013].

2.2.4 Complementing experimental data

In order to run simulations that are comparable with experiments, we need first to correctly set up the system. This is usually a challenging part, especially for amyloid systems, for which initial structures are not always known. Indeed, α-synuclein is intrinsically disordered, furthermore, several fibril morphologies can be formed, or even co-exist [Li et al., 2018b] [Li et al., 2018a] [Skamris et al., 2019].
In this work, several conformers were extracted from NMR refined ensembles for monomers. In the case of fibrils, we used two different models, the first assays were carried out using an NMR structure of full-length, single-stranded fibrils [Tuttle et al., 2016], for which the cross-beta core was fitted on a double-stranded structure obtained with electron microscopy [Li et al., 2018a]. This first model did not show good agreement with the experiment. Therefore, we then used single-stranded fibrils. Both systems were hydrated to the same level as for neutron scattering measurements (0.4 g H_2O / g protein).
Once simulations are finished, the results can be easily compared with neutron experiment using the autocorrelation function to compute simulated QENS spectra. In addition, the MSD can be directly obtained using the following

$$MSD = \frac{1}{N}\sum_{\mathrm{i}} \langle (\boldsymbol{r}_{\mathrm{i}}(t+\Delta t) - \boldsymbol{r}_{\mathrm{i}}(t))^2 \rangle \tag{2.55}$$

where the sum ranges over atoms, N is the total number of atoms being considered, and angular brackets denote the average over time origins. Moreover, the MSD can also be extracted from the simulated QENS spectra - using the elastic peak dependence on the momentum-transfer q - where instrumental resolution function is taken into account. The MSD allows to perform an experimental validation of the results before starting a more thorough analysis.

As discussed before, neutron scattering data provide thermally averaged information. However, local structure and dynamics might also be of great help in deciphering the aggregation process. This is where molecular dynamics simulations comes to play, by providing full trajectories of each individual atom. Various analysis algorithms are available now, either in the VMD software [Humphrey et al., 1996], or in other packages such as MDAnalysis [Gowers et al., 2016] [Michaud-Agrawal et al., 2011], and its related modules for water dynamics analysis [Araya-Secchi et al., 2014].

2.3 Complementary biophysical characterizations

2.3.1 X-ray powder diffraction

Having a protein powder at hand, it is necessary to characterize it prior to dynamics measurements. More precisely, we want to make sure that we got the right aggregated - or native - species. To this purpose we can take advantage of very well defined distances in the cross-beta pattern of amyloid fibrils [Eanes and Glenner, 1968]. Even for randomly oriented fibrils in the powder, these distances will contribute to form diffraction rings at specific angles that are characteristic of cross-beta patterns as it can be seen in figure 2.5. The exact angle and intensity may vary depending on the protein considered or aggregate morphology but two rings are usually visible. The first appears at 4.5 Å and corresponds to the distance between individual β-strands along the fibril axis. The second ring, at 10 Å, corresponds to distances arising from the three dimensional fold of the β-sheet stack.

2.3.2 Proton-induced X-ray emission

Originally proposed by Sven Johansson in 1970 [Johansson et al., 1970], the micro beam proton induced X-ray emission (microPIXE) makes use of the interaction of MeV protons with the electron cloud. The induced perturbation causes electron excitation or ionization. The subsequent relaxation then results in the emission of photons of energy typically in the range of 1 to 30 keV. The ionization cross-section for K and L shells is highest when the proton velocity matches that of the electron in its orbit, usually around 10 % of the speed of light. With protons, only a small fraction of the initial energy is lost at each interaction. Thus, proton motion is barely affected along its way through the material and almost no bremsstrahlung is emitted. Hence, the background signal is very low compared to other methods, and small amounts of elements can be detected.

Another process can be detected, i.e. Rutherford backscattering (RBS) that allows to determine atom masses and thicknesses of the samples. The number of X-ray photons detected for a thin layer of sample at depth x is given by [Garman and Grime, 2005]

$$dN_{\mathrm{z}} = KQc_{\mathrm{z}}\sigma_{\mathrm{I}}(Z,E)\omega(Z,E_{\mathrm{Z}})\Omega\epsilon(E_{\mathrm{Z}})e^{-\mu(\boldsymbol{M},E_{\mathrm{Z}}\frac{x}{cos\theta})}dx \tag{2.56}$$

where K is a constant, Q the total proton charge in the exposure, m_{Z} the total mass of element exposed to the beam, σ_{I} the ionization cross-section, ω the fluorescence yield, Ω the solid angle of detection, ϵ the detector efficiency, **M** the sample matrix composition that gives elements depth distribution in the sample, and θ the angle

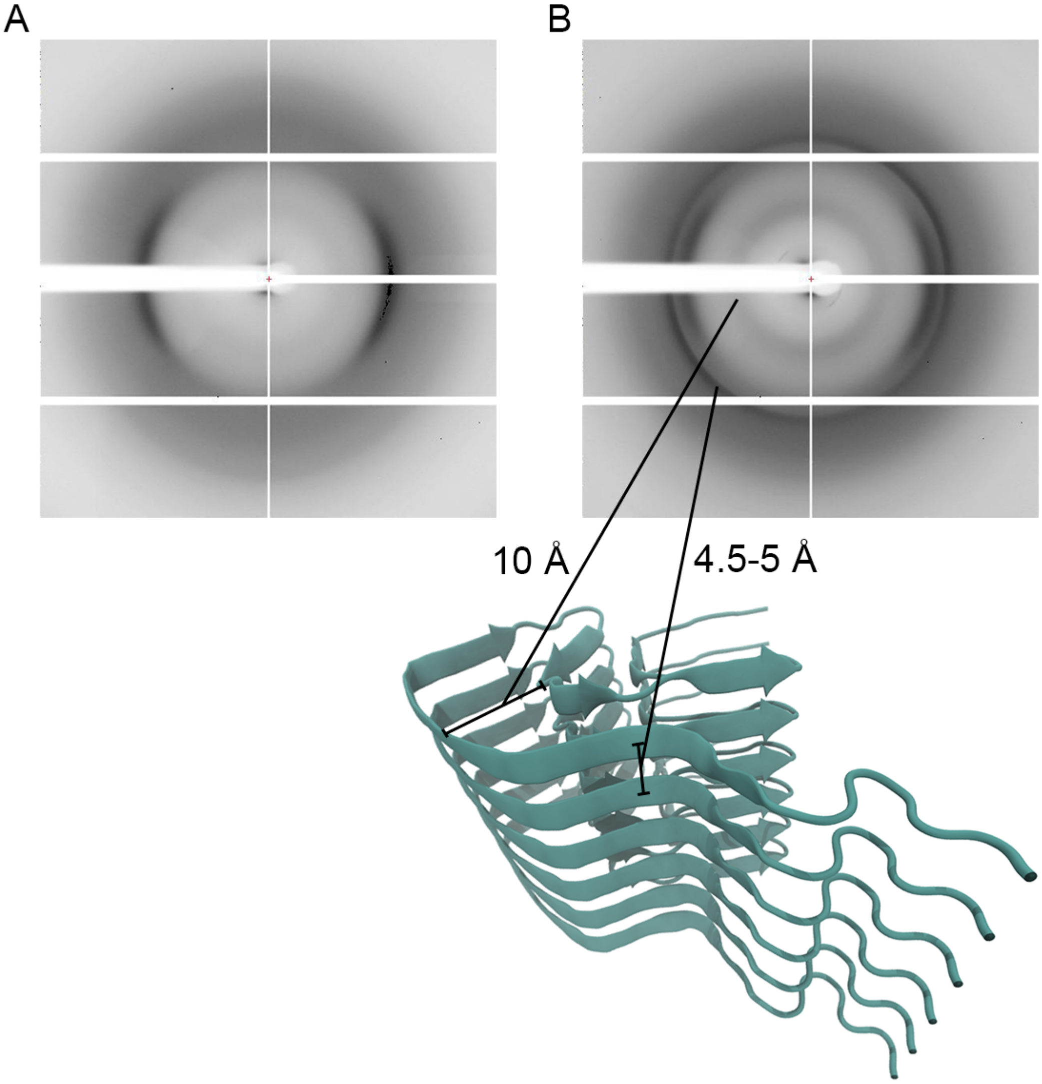


Figure 2.5: Amloid fibrils typical X-ray diffraction pattern
Typical X-rays diffraction pattern - recorded here on ID30A beamline at the ESRF - obtained for **A**, a natively disordered protein powder - α-synuclein here. And **B**, its amyloid fibrils. The two equatorial circular arcs originate from the Mylar capillary containing the powder.

between sample normal and detector direction. By integrating the above equation, the concentration of each element can - in theory - be obtained.

In practice, the determination of **M** and Q is challenging, especially for protein samples, where element densities can vary significantly at microscopic scale. Therefore, the actual analysis is usually done by taking the ratio of the unknown element concentration with that of known sulphur concentration in the protein - given by the number of cysteines in the amino acid sequence. To determine **M**, the Rutherford backscattering spectrum can be used. A typical result is shown in figure 2.6. The method provides a good accuracy with a low detection limit. Moreover, it has the non negligible advantage to require only tiny amounts of sample (typically a hundred of micrograms).

2.3.3 X-ray absorption spectroscopy

X-ray absorption fine structure (XAFS) refers to the oscillatory structure of the absorption coefficient at energies above an X-ray absorption edge. The oscillatory sctructure results from the process of photon absorption by a specific atom and subsequent emission of an excited electron. The latter will be scattered by neighbouring atoms, undergo interference effects with incoming and scattered waves and gives rise to a spectrum specific of the local structure around the absorber [Rehr and Albers, 2000]. The development of the method is relatively new, many advances occurred during the 90s, both theoretically and experimentally. The method can provide very accurate measurement of distances between atoms, and is particularly useful in biology when metal ions are involved.

XAFS basically consists in measurement of the absorption coefficient μ by comparing the decrease of X-ray beam intensity after it passes through the sample

$$I(E) = I_0(E)e^{-\mu(E)x} \tag{2.57}$$

where I_0 is the incident beam intensity, μ the absorption coefficient, E the energy of incident beam and x the distance travelled by the beam through the sample. The XAFS spectrum χ is computed using the following - and possibly other corrections terms [Rehr and Albers, 2000]

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\Delta\mu_0} \tag{2.58}$$

where μ_0 is a background absorption and $\Delta\mu_0$ a normalization factor. The abscissa is usually converted to wave vector k units for later Fourier transform of the function $\chi(k)$, which then provides a radial density distribution of scatterers around the

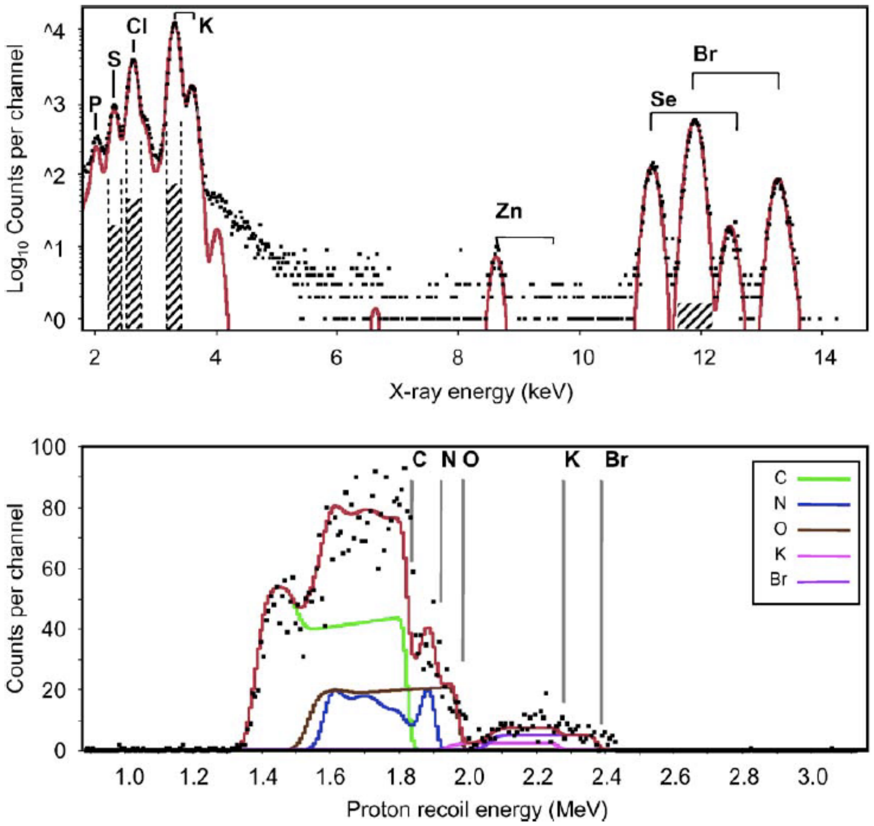


Figure 2.6: microPIXE and RBS spectra of protein sample

PIXE (top) and RBS spectra (bottom) from a 2 mm diameter point on a thin liquid protein sample from which the S/Se ratio was sought, and which had been buffer exchanged from KCl to KBr to avoid overlap of the sulphur and chlorine peaks. The spectra were recorded simultaneously in 5 min using a beam of 2.5 MeV protons at a current of 100 pA. In each spectrum, the recorded data are shown as black points and the theoretical fit generated by the spectrum processing software is shown as a solid red line. In the PIXE spectrum, the identity of the major peaks is marked: note the log y scale of this spectrum. The RBS spectrum is modelled as a 6.15 mm thick layer of material with chemical formula $C_5N_{1.15}O_{2.28}K_{0.21}Br_{0.13}$ mounted on a 1.85 mm polymer film with composition $C_5O_{1.8}$. Coloured lines indicate the partial spectra corresponding to each element in the sample structure and the surface energy of each element is indicated by vertical lines.
Adapted from [Garman and Grime, 2005].

absorbing atom. Data analysis is not straightforward, and it is actually not possible with protein samples to directly extract structural information from XAFS. Instead, a guessed structural model is constructed, then used to compute a theoretical spectrum that can be compared with the experimental one. For this purpose, a standard formula was proposed by Sayers and coworkers in 1971 [Sayers, 1971], and subsequently improved to yield the following general form

$$\chi(k) = \sum_{\mathrm{R}} S_0^2 N_{\mathrm{R}} \frac{|f(k)|}{kR^2} \sin(2kR + 2\delta_{\mathrm{c}} + \Phi) e^{-2R/\lambda(k)} e^{-\sigma^2 k^2} \tag{2.59}$$

where S_0^2 is an overall amplitude factor coming from many-body effects, N_{R} the coordination number, R the inter-atomic distance, f(k) a backscattering amplitude, δ_{c} and Φ denote phase shifts, $\lambda(k)$ the mean-free path of electrons, and σ a temperature dependent fluctuation term. Our data were processed using GNXAS suite [Filipponi and Cicco, 2000], that allow to account for single and multiple scattering to fit the experimental data.

2.3.4 Dynamic light scattering

Dynamic light scattering (DLS) is now a routinely used method in most biochemistry labs. This is due to its relatively low cost, for a good throughput and useful information that is provided regarding either quality control or sample characterisation. The basic principle is to send a laser on the sample, and monitor the intensity fluctuations in time of the outgoing signal at a given angle. Hence, we will have to make use of the autocorrelation function here. The DLS signal is usually described by means of two autocorrelation functions, the second order one describes the actual intensity measured, which depends mostly on particle motions, and is given by [Stetefeld et al., 2016]

$$g_2(\tau) = \frac{\langle I(t) I(t+\tau) \rangle}{\langle I(t) \rangle^2} \tag{2.60}$$

However, this equation includes various effects - including instrumental ones. It can be related to the scattered electric field correlation function - also known as first order correlation function - which directly depends on coherent scattering from particles moving relative to each other,

$$g_1(\tau) = \frac{\langle E(t) E(t+\tau) \rangle}{\langle E(t) \rangle^2} \tag{2.61}$$

using the following Siegert relation

$$g_2(\tau) = B + \beta|g_1(\tau)|^2 \tag{2.62}$$

where B is a baseline term, usually close to unity, and β is a coherence factor, that depends on instrumental parameters, and also on some of the studied system parameters. The first order correlation function is usually represented by a decaying exponential for diffusive motions, with a parameter Γ, related to the diffusion coefficient by $\Gamma = Dq^2$, with q^2 being the momentum transfer.

For monodisperse systems, a single exponential can be used. However, polydisperse systems are better represented using g_1 as an intensity weighed integral over the decay rate distribution

$$g_1(\tau) = \int_0^\infty G(\Gamma)e^{-\Gamma_\tau}d\Gamma \tag{2.63}$$

The most widely used approach for data analysis is the cumulant method, introduced by Koppel [Koppel, 1972]. The cumulants are related to the moments of a distribution. It takes the form of a series-expanded exponential. Indeed, let us consider the following moment generating function,

$$\mathcal{M}(-\tau, \Gamma) = \langle e^{-\Gamma\tau} \rangle = 1 + \sum_{\mathrm{m}=1} \frac{(-\tau)^m \langle \Gamma^m \rangle}{m!} \tag{2.64}$$

where angular brackets denote averages. We can readily see that the m^{th} cumulant $\langle \Gamma^m \rangle$ can be obtained by

$$k_{\mathrm{m}} = \frac{d^m}{d(-\tau)^m} \ln(\mathcal{M}(-\tau, \Gamma)) \bigg|_{\tau=0} = \frac{\langle \Gamma^m \rangle}{m!} \tag{2.65}$$

and the cumulant generating function is simply the natural logarithm of the moment generating function $\mathcal{K}(\tau) = \ln\Big(\langle e^{-\Gamma\tau)} \rangle\Big)$.

Using the Siegert relation, we can obtain the following for the measured signal $g_2(\tau)$

$$\ln(g_2(\tau) - B) = \ln\beta + 2\left(-\bar{\Gamma}\tau + \frac{k_2}{2!}\tau^2 + \frac{k_3}{3!}\tau^3 + \dots\right) \tag{2.66}$$

Thus, the different moments of the Gaussian distribution can be fitted on the measured signal. Importantly, the polydispersity index is given by $k_2/\bar{\Gamma}^2$.

Nowadays, data acquisition and analysis is well implemented and almost fully automated. Various parameters can be quickly obtained on a protein solution, such as collective translational diffusion coefficients, and the polydispersity index. Moreover, DLS presents the great advantage of requiring only tiny amounts of sample, typically 10 μL of a solution at 5-10 mg/mL are sufficient to obtain accurate measurements. Particles of size lower or comparable to the wavelength of the incoming beam are required to obtain reliable results, where the relation between the momentum transfer q and the particle radius R, qR « 1, is satisfied. The standard range for measurable particle size goes from 1 nm to 1 μm.

Results

3.1 α-synuclein and hydration water dynamics

Parkinson's disease is a neurodegenerative disorder which is usually characterized by shaking and rigidity. These symptoms are usually associated with additional problems such as depression, anxiety, dementia, diabetes, or cancer [Ristow, 2004] [Lim and Lang, 2010] [Landau et al., 2016]. Similarly to other amyloid disorders, amyloid deposits are found in patient brains that contain an important fraction of α- synuclein protein [Spillantini et al., 1997]. The protein is 140 amino acids long, with three major regions. A N-terminal region usually defined up to residue 60 involved in membrane interaction, a center NAC region - residue 60 to 95 - that is aggregation prone and a C-terminal region exhibiting phosphorylation and calcium-binding sites [Lautenschläger et al., 2018].
It is unclear how this protein can become toxic and trigger Parkinson. Various scenarios were proposed based on experimental data. Interestingly, the mature fibril seems to be a safe species, while early stage oligomers, and the fibrillation process itself might give rise to the toxicity. Indeed, membrane disruption is likely to be one of the toxicity mechanisms of α-synuclein. It has been shown that mature fibrils of IAPP peptides or tau are safe for membranes, while damages occur when the fibrillation process is taking place in presence of negatively charged membranes [Martel et al., 2017] [Jones et al., 2012]. Moreover, various species of α-synuclein can interact with lipids as shown in works by Hellstrand et al. and Skamris et al. [Hellstrand et al., 2013] [Skamris et al., 2019]. Additional cellular-leveled studies provide support to the hypothesis of toxicity arising from early-stage species, see works by Winner and Lázaro for examples [Winner et al., 2011] [Lázaro et al., 2014]. Eventually, α-synuclein can affect gene expression and homeostasis long before any cellular damage is seen [Decressac et al., 2013b]. Obviously, the disease is age-related, and how aging and α-synuclein fibrillation are linked is not clear, even though decreased efficiency of protein degradation might play an important role

[Stefanis et al., 2019].
Along with the understanding of what happens inside the cells, it is necessary to understand what makes the protein fibrillate in the first place. And subsequently understand how various factors as mutations, or environment can affect this process. In line with previous work in the group, we used a combined approach involving neutron backscattering and molecular dynamics simulations to get insights in the biophysical origin of fibrillation. It has been shown for tau that protein dynamics remains unchanged after fibrillation formation, while hydration water dynamics was significantly increased [Fichou et al., 2015b]. However, the lack of full- length tau fibril structure made further analyis based on simulations unfeasible. With the α-synuclein, both native form conformation ensembles and full-length fibril structure - produced in similar conditions as ours - are available [Tuttle et al., 2016]. We therefore embark on a comprehensive study on this protein and its hydration water dynamics to better understand the physical determinants of its amyloid fibrillation.

Initially, the protein has been produced at the D-Lab - ILL, Grenoble - and purified at Bente Vestergaard's lab in Copenhagen. Protein purification and samples production posed major challenge, mainly due to the requirement of an experimented workforce. The main issue was the quantity required for a neutron experiment, that is, around 100 mg per sample measured. To address this issue, I performed the protein expression and purification with the help of Martine Moulin at the D-Lab, who had already optimized the protocol for high amounts. Subsequently, the fibrillation protocol had be to be revised completely as it was designed only for fibrillation monitoring in 100 μL with low concentration by either Thioflavin T or SAXS measurements. It took some time to be able to produce α-synuclein fibrils in 1.5 mL eppendorf tubes with reasonable yield - around 50 %.
Once samples were produced, measurement on SPHERES at the MLZ, Münich and data analysis went smoothly using a well defined methodology. However, the next difficulty came from simulations. Indeed, if native form is relatively well defined, it is not the case for fibrils, which can exhibit various forms in the same tube [Pansieri et al., 2018]. Electron microscope images suggest a relatively good homogeneity in our conditions but we could not determine whether it is single or double-stranded fibrils. Thus, the two configurations were built and used in MD simulations. To achieve this, either the Tuttle's single-stranded model [Tuttle et al., 2016] was directly used, or two copies of this model were fitted on the core domain of double-stranded model published recently [Li et al., 2018b]. After having performed simulations as described in the following paper, we could show that only the single-stranded model can reproduce our experimental data. This form was then used for

all analysis.
To briefly summarize the results from the following article, we found that dynamics is increased for both native and fibril forms of α-synuclein. However in a larger extent for hydration water. The elastic incoherent structure factor shows that this arises partly from a larger fraction of water molecules going from an apparent immobile to a mobile state, for both rotational and translational motions. Analysis of MD simulations showed a good agreement in computed MSD and QENS with experimental data. Moreover, analysis of hydrogen bonds lifetime showed a general increase in monomers for both rotational and translational motions, and the number of hydrogen bonds between water and monomers is higher than for fibrils for the residues above the residue 40. Interestingly, we could show that hydration water is denser around termini regions than around protein core in fibrils, while density is rather homogeneous in monomers. It seems that the observed dynamical, and entropy, increase is strongly associated with the fact that water interacts more with the central hydrophobic region in the highly compact monomeric state, while the extended fibrillar structure allows the water to gain in mobility. This is supported by the fact that the core is segregated during fibrillation - as shown by averaged distances -, hence resulting in a less compact structure, but with a central part being self- sufficient and efficient in protecting itself from interacting with the solvent.
This study constitutes the material of the following paper, which will be submitted to the Journal of American Chemistry Society (JACS).

Article 1

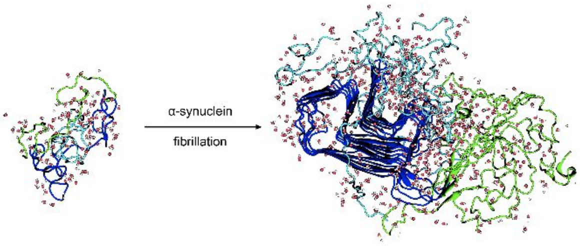
Water restructuring upon α-synuclein fibril formation induces an increase of dynamics and entropy

Water restructuring upon α-synuclein fibril formation induces an increase of dynamics and entropy

Kevin Pounot[1,2], Giorgio Schiro[1], Martine Moulin[3], Carlotta Marasini[4], Daria Noferini[5], Michaela Zamponi[5], Michael Haertlein[3], Trevor Forsyth[3], Martin Blackledge[1], Bente Vestergaard[4], Tilo Seydel[2], Annette E. Langkilde[4], Douglas J. Tobias[6], Martin Weik[1]

[1]Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France
[2]Institut Max von Laue – Paul Langevin (ILL), CS 20156, F-38042 Grenoble, France
[3]Life Sciences Group, Institut Laue–Langevin, 6 Rue Jules Horowitz, 38042 Grenoble, France
[4]Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark
[5]Jülich Centre for Neutron Science, outstation at Heinz Maier-Leibnitz Zentrum, Forschungszentrum Jülich GmbH,85747 Garching, Germany
[6]Department of Chemistry, University of California, Irvine, Irvine, California 92697, United States

Abstract



Protein amyloid fiber formation is the pathological hallmark in various neurodegenerative diseases such as Parkinson's or Alzheimer's. The physico-chemical origin of protein fibrillation, as well as the role that hydration-water might play remain elusive. We combined neutron spectroscopy and molecular dynamics simulations on hydrated powders of α-synuclein and tau to investigate both structural and dynamical properties of the protein-hydration water system. Hydration-water dynamics is enhanced in the fiber state of both α-synuclein and tau, as shown by increased water mean-square displacements and broader quasi-elastic neutron scattering spectra. Molecular dynamics simulations of hydrated α-synuclein powders evidence a compact monomeric state and a more extended fiber state, in which interaction between the NAC segment and the N-terminal and C-terminal regions is reduced. As a consequence, water around N-terminal and C-terminal is no more constrained by the hydrophobic residues in the NAC segment, resulting in increased dynamics and entropy. The increase in water dynamics upon fiber formation is larger for tau than for α-synuclein. Since the latter contains a much smaller fraction of disordered residues in the fiber state than the former, we conclude that residual fiber disorder correlates with hydration water dynamics. The entropic driving force that increased water dynamics present for fiber formation is suggested to be maximized in amyloids with an extensive fuzzy coat such as tau.

Introduction

Alongside their native three-dimensional structures, many proteins tend to adopt alternative configurations that can lead to aggregation and the loss of protein function[1]. Furthermore, proteins without a well-defined three-dimensional structure, so-called intrinsically disordered proteins often form highly organized aggregates such as amyloid fibrils[2]. Aggregates, whether amorphous or fibrilar, have deleterious if not fatal effects, as manifested in human pathologies, so-called protein aggregation diseases, such as Alzheimer's and Parkinson's diseases. In Alzheimer's, intracellular fibrils formed by the intrinsically disordered protein tau[3] and extracellular plaques by the amyloid-beta peptide[4] are the two major pathological hallmarks. In Parkinson's, the amyloid fibrils formed by intrinsically disordered α-synuclein are the molecular manifestation of the disease[5]. Major progress recently made in the fields of solid-state NMR and single-particle cryo electron-microscopy now allows studying the structure of amyloid fibrils at the atomic level of resolution, such as those formed from the full-length human α-synuclein[6] and from the human protein tau[7,8], respectively. Structural details of the core of amyloid fibrils, mimicked by crystalline peptide fragments, have been obtained by X-ray[9] and electron crystallography[10]. This core displays a so-called cross-β architecture, where interdigitating β-sheets form dry steric zippers that are devoid of water[9]. The hydrophobic effect thus contributes to amyloid stability[11], thereby hinting at the potentially important role that hydration water dynamics and entropy may play in protein aggregation.

The role of water in fiber formation has recently started to move into the focus of amyloid research[12]. Aggregation of the intrinsically disordered human protein tau, for instance, has been suggested to involve a transition from homogeneous to heterogeneous hydration[13]. Likewise, five water pools have been distinguished in Alzheimer's β-amyloid fibrils, including a water pore in their center[14]. An increase in water entropy by expulsion of water molecules during oligomer and fiber formation has been suggested to act as a driving force for fibril formation[15]. Indeed, a neutron scattering on monomers and fibers of tau provided evidence for an increase in water translational diffusion on the fiber surface[16]. To the contrary, femto- to picosecond time-resolved fluorescence measurements on bovine κ-casein indicated that hydration dynamics is slower in the case of amyloid compared to monomeric protein[17]. It has been concluded that further studies are required to decide whether the apparent discrepancy is due to hydration-water on different proteins behaving differently or to the different techniques used[12]. Regardless, a change in water mobility upon fiber formation might provide a diagnostic means for amyloidosis by *in vivo* magnetic resonance imaging[16].

The protein α-synuclein is involved in synucleinopathies, such as Parkinson's disease[5] and can also affect homeostasis and gene expression outside the central nervous system and at early aggregated states[18,19]. Intrinsically disordered in its native state, α-synuclein (140 amino acids, 14.4 kDa) contains an N-terminal domain (N-ter) mainly involved in lipid binding which extends from residue 1 to 60, a central hydrophobic domain (NAC) with several post-translational modification sites that is involved in amyloid-fiber formation which spans residues 61 to 95 and a mainly unstructured C-terminal tail (CTT) from residues 96 to 140[20,21]. The fibrillation is usually described as involving nucleation and seeding mechanisms, for which several monomers could oligomerize to form pre-fibrillar species[22,23]. The atomic-resolution structure of α-synuclein fibrils have been solved by solid-state NMR[6] and cryo-electron microscopy[24–26], showing a complex Greek-key topology stabilized by a salt bridge and hydrophobic packing. Forty residues surround the fiber core[26], similarly than does the so-called fuzzy coat in tau fibers[7]. Structural studies on mutants showed that long-range interactions that compact the native state contribute to prevent aggregation[27]. These mutants can affect fibril conformation and thermodynamic stability but can also make native and early aggregation stages less stable and more exposed to the solvent[28]. Moreover, it has been shown that exposure of the N-ter of the protein to the solvent is of particular importance for aggregation propensity[29]. Whereas the dynamical heterogeneity of hydration water has been mapped for native α-synuclein[30], it remains to be studied if changes in hydration-water occur during the fibrillation of α-synuclein[12].

Neutron spectroscopy represents an exquisite tool to study the dynamics of biological macromolecules and of their hydration water[31,32]. Incoherent neutron scattering is largely dominated by the signal from hydrogen atoms. Consequently, examining a perdeuterated protein hydrated in H_2O provides a means to mask the protein contribution and to focus on hydration-water dynamics[33,34]. A comparison with hydrogenated proteins hydrated in D_2O provides information on protein dynamics and the coupling with hydration water dynamics. In particular, elastic incoherent neutron scattering (ENS) experiments in a temperature range from 50 K to room temperature have been employed to study the so-called protein dynamical transition[35] at around 220 K and their relation to protein function and hydration-water dynamics[36]. Quasi-elastic incoherent neutron scattering (QENS), measured at a few distinct temperatures, informs about the nature of protein and hydration-water motions[37]. Both ENS and QENS have been applied to study the dynamics of protein fibers[38] and of their hydration water[16].

Here, we combine ENS and QENS on powders of deuterated α-synuclein and tau monomers and fibers with molecular dynamics (MD) simulations on α-synuclein powders to study changes in hydration water dynamics upon amyloid formation. Mean-square displacement of water molecules are larger in both α-synuclein and tau fibers than in monomers, and quasi-elastic spectra are broader, evidencing a larger water mobility in fibers than in monomers. The increase in water mobility upon fibrillation is larger in tau than in α-synuclein. MD simulations of α-synuclein powders reveal a compact monomeric and a more extended fiber state, suggesting that water is displaced from NAC to N-ter residues during fiber formation. Taken together, our results suggest that a gain in water entropy drives pathogenic protein fibrillation and that the extend of this gain is higher for protein fibers with a higher ratio between ordered and disordered regions.

Results

Characterization of α-synuclein fibers. The formation and integrity of α-synuclein amyloid fibrils was assessed using electron microscopy and X-ray diffraction (Figure 1), showing the typical elongated morphology of fibrils (Figure 1a) and patterns corresponding to amyloid structures[39] (Figure 1b), respectively, for both hydrogenated and deuterated samples. To the contrary, monomeric α-synuclein samples did neither show fibrils in electron micrographs (Figure S1a) nor amyloid-rings in X-ray diffraction patterns (Figure S1b).

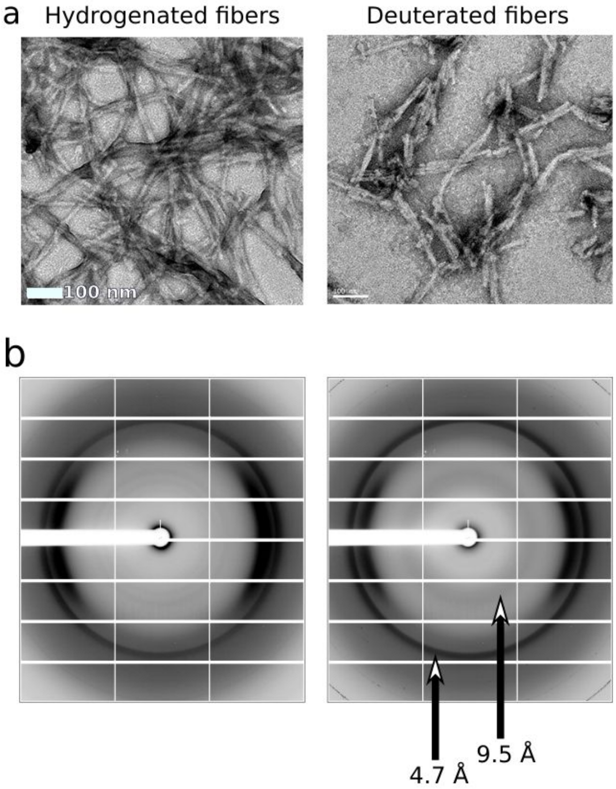


Figure 1. Characterization of α-synuclein fibrils. (a) Negative-staining electron micrographs of hydrogenated (left) and deuterated (right) fibrils. (b) X-ray fiber diffraction patterns of hydrogenated (left) and deuterated (right) fibrils show rings at 4.7 and 9.5 Å (highlighted by arrows) characteristic of cross-β structures in amyloid fibers. The two antipodal arcs arise from the Mylar capillary carrying the sample.

Elastic and quasi-elastic neutron scattering on α-synuclein monomers and fibers. Elastic and quasi-elastic incoherent neutron scattering was applied to study and compare the hydration-water dynamics of monomeric and fibrillar α-synuclein. The incoherent scattering signal from the deuterated α-synuclein samples hydrated in H_2O is dominated by hydrogen atoms in the hydration shell (73%, whereas 27% originates from the protein atoms - see Table S1), thus allowing to monitor mainly hydration water dynamics. Mean-square displacement (MSD) were obtained by fitting the angular dependence of the elastic incoherent neutron scattering (ENS) intensity within the Gaussian approximation[40] at each temperature (see Methods for details). MSDs of deuterated monomers in H_2O (D-mono-αsyn-H_2O) and fibrils in H_2O (D-fib-αsyn-H_2O) are compared in Figure 2a. Above the dynamical transition at 220 K, the MSDs are larger for D-fib-αsyn-H_2O than for

D-mono-αsyn-H_2O. In a control experiment, MSDs were determined for hydrogenated monomers in D_2O (H-mono-αsyn-D_2O) and hydrogenated fibrils in D_2O (H-fib-αsyn-D_2O), whose incoherent scattering signal originates to 98% from protein atoms (Figure S3a). The MSDs are slightly larger for H-fib-αsyn-D_2O than for H-mono-αsyn-D_2O, indicating that α-synuclein fibers are slightly more flexible than monomers. Since the difference in MSDs above 220 K of H-fib-αsyn-D_2O of H-mono-αsyn-D_2O are smaller than of D-fib-αsyn-H_2O and of D-mono-αsyn-H_2O, we conclude that the difference between D-fib-αsyn-H_2O and D-mono-αsyn-H_2O is dominated by dynamical differences in the hydration water, indicating an increase of the water dynamics upon fibrillation.

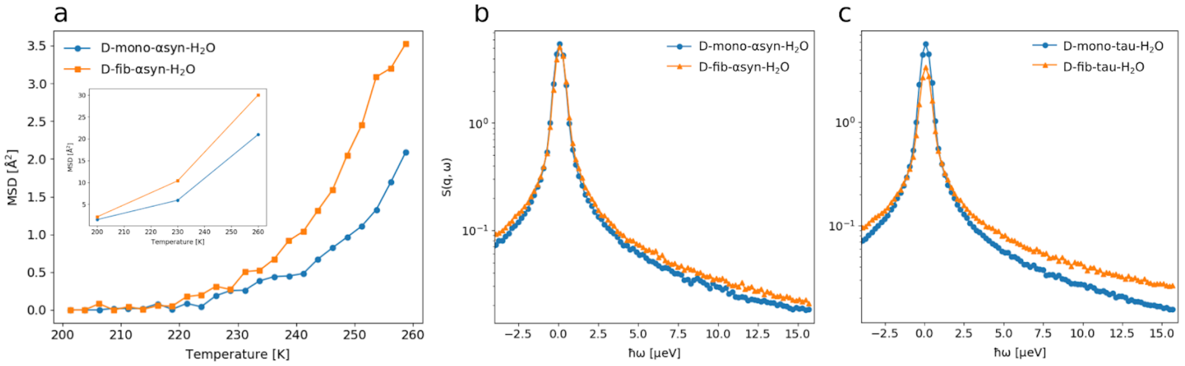


Figure 2. Hydration water dynamics is increased in fibrils compared to monomers. (a) Hydration water MSDs - obtained from a Gaussian fit in the momentum transfer range $0.22 < q < 0.96$ Å^{-1} - from elastic scans on the D-mono-αsyn-H_2O (blue circles) and D-fib-αsyn-H_2O (orange squares). The inset shows the corresponding MSDs computed directly from simulation trajectories as described in Methods (eq. 3). (b) QENS spectra summed over all momentum transfer values for D-mono-αsyn-H_2O (blue circles) and the D-fib-αsyn-H_2O (orange squares). (c) QENS spectra summed over all momentum transfer values for D-mono-tau-H_2O (blue circles) and D-mono-tau-H_2O (orange squares).

To get a more detailed information on the nature of the dynamical increase of hydration water in fibrils, full QENS spectra were measured at 260 K - that is before the water dynamics goes out of the time window probed by the instrument - for 12 hours in an energy range -15 μeV < E < 15 μeV. Qualitatively, the QENS spectra summed over all momentum transfer values are broader for the D-fib-αsyn-H_2O sample than for the D-mono-αsyn-H_2O one (Figure 2b). To determine the nature of the water motions broadening the QENS spectra in D-fib-αsyn-H_2O, we used a model (eq. 1 in Methods) that accounts for a fraction A_0 of apparently immobile water, a fraction A_r of water molecules undergoing rotational motions at a rate Γ_r, and a fraction A_t of water molecules undergoing translational motion with an apparent diffusion coefficient D_t. The fitting of D-mono-αsyn-H_2O and D-fib-αsyn-H_2O spectra (Figure S2a) with this model (in the momentum transfer range $0.6 < q < 1.8$ Å^{-1}) shows that Γ_r is increased by 87% and D_t by 48% in D-fib-αsyn-H_2O compared to D-mono-αsyn-H_2O (Figure S2b). In D-fib-αsyn-H_2O, the fraction A_0 decreases by about 1%, A_r decreases by 6% and A_t increases by 20% as compared to D-mono-αsyn-H_2O, thereby indicating that the increase in the water dynamics upon fibrillation arises mainly from a higher mobility of water molecules with a small contribution from an increased fraction of water undergoing rotational or translational motions.

To determine how the protein dynamics is affected by fiber formation, full QENS spectra were measured at 300 K for 8 hours in an energy range -15 µeV < E < 15 µeV. To describe the protein dynamics, the model used (eq. 2 in Methods) contains a fraction a_0 of apparently immobile atoms, and two fractions a_1 and a_2 of atoms undergoing Fickian-type diffusion with apparent diffusion coefficients D_1 and D_2. The QENS spectra of H-mono-αsyn-D_2O and H-fib-αsyn-D_2O were fitted according to this model in the momentum transfer range $0.6 < q < 1.8$ Å^{-1} (Figure S3). The parameters a_0 barely change after fiber formation (less than 5%), while a_1 decreases by 15% and a_2 increases by 26%. However, D_1 increases by 31% and D_2 increases by 29%. This analysis shows that the increased dynamics observed with the MSDs on H-mono-αsyn-D_2O and H-fib-αsyn-D_2O arises mainly from a higher mobility of protein hydrogen atoms.

Elastic and quasi-elastic neutron scattering on tau monomers and fibers. The ENS data were obtained from a previous work[16] and reprocessed using the same procedure as for α-synuclein (see Methods) and a QENS experiment was conducted to measure the hydration water dynamics on the monomers and the fibers of the protein tau. Similarly to α-synuclein, the incoherent neutron scattering arises from hydrogen atoms in the hydration shell (72% from water and 28% from protein atoms). The MSDs for the deuterated tau monomers hydrated with H_2O (D-mono-tau-H_2O) and the deuterated tau fibers hydrated with H_2O (D-fib-tau-H_2O) were determined by fitting the dependence of the ENS intensity on the momentum transfer q within a Gaussian approximation[40] at each temperature (see Methods). At 260 K, the MSDs are approximately 2.5 times larger for D-fib-tau-H_2O than D-mono-tau-H_2O (Figure S4). Conversely, the MSDs for the tau protein dynamics, measured in a previous work[16], are identical for hydrogenated tau monomers and fibers hydrated with D_2O (98% of the ENS intensity due to the protein). As for α-synuclein, we conclude that the observed increase in the MSDs is mainly due to the increased dynamics of hydration water.

To better understand the origin of the increased dynamics in water dynamics around tau fibers, QENS spectra were acquired on D-mono-tau-H_2O and D-fib-tau-H_2O for 12 hours in an energy range -15 µeV < E < 15 µeV at 260 K (Figure 2c), *i.e.* at the same temperature than QENS data on D-mono-αsyn-H_2O and D-fib-αsyn-H_2O (Figure 2b). Previously published QENS data on D-mono-tau-H_2O and D-fib-tau-H_2O were measured only at 280 and not at 260 K[16]. The dynamics of tau hydration water was modelled as for α-synuclein using eq. 1 (Methods). The QENS spectra of D-mono-tau-H_2O and D-fib-tau-H_2O were fitted in the momentum transfer range $0.6 < q < 1.8$ Å^{-1} using this model (Figure S5). The parameter Γ_r decreases by 49% and the parameter D_t by around 42%. In D-fib-tau-H_2O, the parameter A_0 decreases by 41%, A_r decreases by 22% and A_t increases by 98% as compared to D-mono-tau-H_2O. Hence, the increased water dynamics as monitored by the MSDs originates from two phenomena, *i.e.* an increase of water mobility (as in α-synuclein**)** and a substantial fraction of apparently immobile water becoming mobile (where the latter phenomenon is less important in α-synuclein).

Temperature dependent molecular dynamics simualtions of hydrated α-synuclein monomers and fibers. Molecular dynamics (MD) simulations were carried out to provide further, spatially-resolved insight into water dynamics measured by neutron scattering. To this end, two *in-silico* models were constructed (see Methods) that mimic the samples studied by neutron scattering, namely a hydrated powder of 20 α-synuclein monomers and a hydrated powder of two 10-mer α-synuclein fibers. Each model served as the starting point for 20-ns MD simulations at 50, 110, 200, 230, 260, and 300 K. The agreement of simulations and experimental data was assessed by computing the temperature dependence of the MSDs (Figure 2a inset and Figure S3a inset) and the Fourier transform of the intermediate scattering function for all non-exchangeable hydrogens of water molecules (Figures S6 and eq. 4 in Methods) from the simulation trajectories. Both calculated MSDs (Figure 2) and QENS spectra (Figures S6) show excellent agreement with experimental data, thus justifying further *in silico* analysis of the latter.

To identify the regions of α-synuclein that are responsible for the increased dynamics in fibers, we computed the continuous (describing mainly rotational motions) and the intermittent (describing mainly translational motions) autocorrelation functions of hydrogen bonds between water molecules that are within 3 Å of either the N-ter, the NAC or the CTT region (Figure 3). The continuous autocorrelation function decreases faster for fibers in the N-ter and the CTT regions, but no differences are observed for the NAC region. The intermittent autocorrelation function decreases faster for fibers for the three regions but the decay rate seems more pronounced in the N-ter and CTT regions. Hence, the increased dynamics of the hydration water observed with neutron scattering arises mainly from the water located around the N-ter and the CTT.

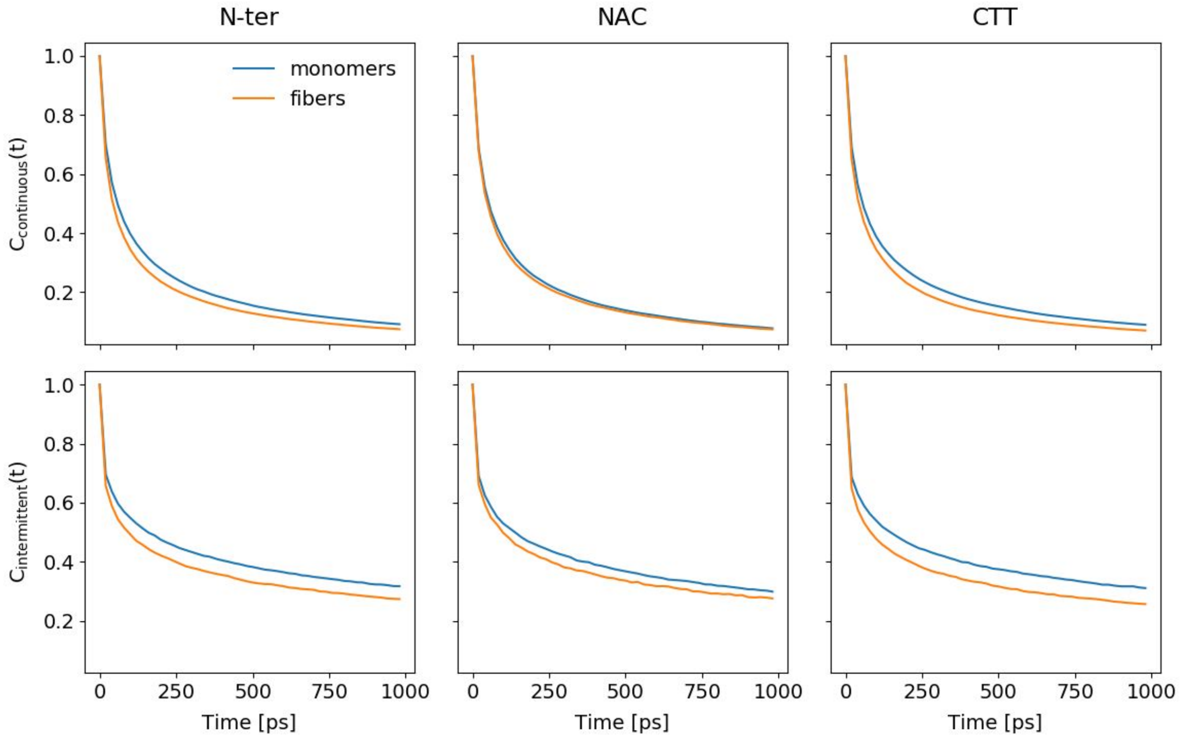


Figure 3. The increase in water-water hydrogen-bond reorganization dynamics is higher around N-ter and CTT regions than NAC region at 260 K. Simulation trajectories were analyzed using the NAMDAnalyzer - HydrogenBond (eq. 5 in Methods) routine at 260 K to calculate both the continuous (upper row) and intermittent (lower row) water-water hydrogen-bond autocorrelation functions. At time t=0 ps, the water molecules that are within 3 Å of either N-ter (left), NAC (middle) or CTT (right) region are selected. The autocorrelation functions were then computed for these selected water molecules using an increment of 20

ps between each point and averaged over 20 different time origins as described in Methods. The result for the monomeric α-synuclein is represented in blue, and the orange line is used for the fibers.

In addition, the radial density profile for water around monomers and fibers was computed (eq. 6 in Methods) to obtain a structural picture of the hydration shell. For a better comparison, the water density around monomer residues was divided by the density of water around fibril residues (Figure 4). The result reveals an excess of water density in the NAC region of monomers as compared to fibrils, while the density around N-ter and CTT remain constant. Therefore it indicates a restructuring of water molecules, where a fraction of water molecules, which was shared between the N-ter, the NAC and the CTT regions in monomers, pertains fully to the N-ter or the CTT neighborhood after fibers formation.

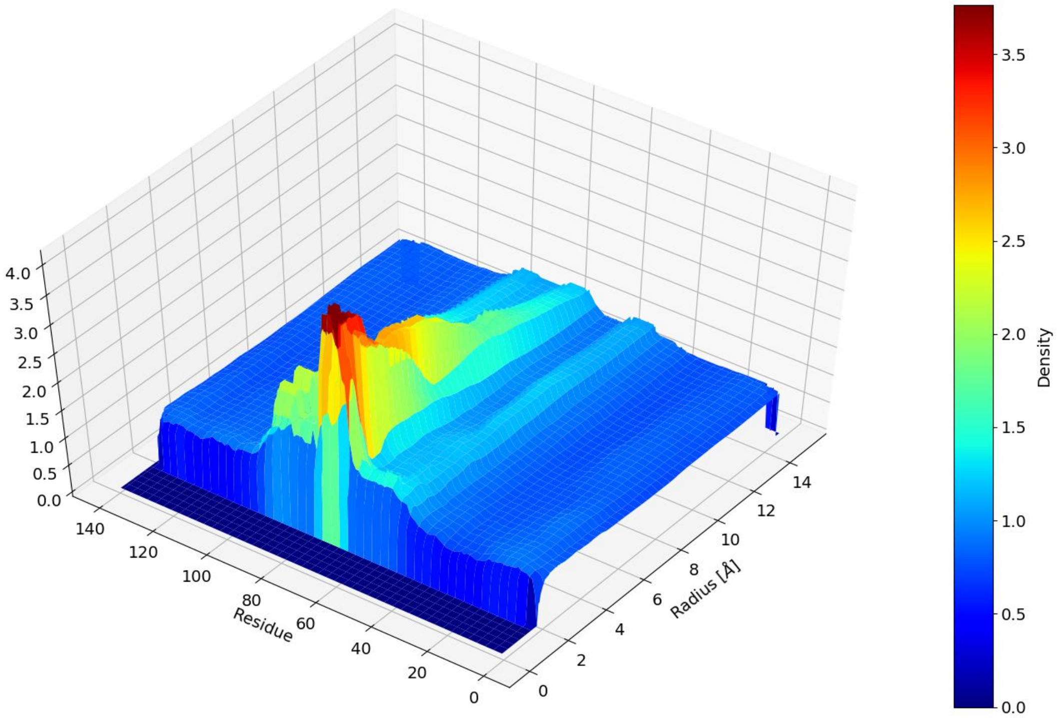


Figure 4. Density of water moecules at monomers and fibril surfaces. The simulation trajectories were analyzed using NAMDAnalyzer RadialDensity module (eq. 6 in Methods) at 260 K to compute the density of water molecules around each individual residue as a function of the minimum radial distance between water oxygen atoms and the residue atoms. The density profile for the monomers was divided by the profile for the fibrils. The three-dimensional surface plotted has been smoothed using the Scipy medfilt2D function with a filter size of [13 (residues),7 (radius)].

In order to provide a structural framework for interpreting changes in hydration-water dynamics upon fibrillation, we computed averaged distances between Cα in either N-ter, NAC or CTT regions and all the other Cα and plotted them on a chord diagram where only distances between 2 and 4 Å or 4 and 6 Å are represented (Figure S7). As already suggested on the basis of experimental data[28,41,42], α-synuclein monomers are highly compact and water interacting strongly with the N-ter and the CTT is close to the NAC. In the α-synuclein fibers, extensive contact breaking occurs, in particular between the N-ter and the NAC regions that move further apart from each other.

Discussion

We combined incoherent elastic (ENS) and quasi-elastic neutron scattering (QENS) and molecular dynamics (MD) simulations to assess hydration water dynamics in hydrated powders of deuterated α-synuclein and tau monomers and fibers. Increased water dynamics is experimentally and computationally observed in α-synuclein fibers compared to monomers. A larger increase in hydration water dynamics as a result of fiber formation is experimentally observed in tau compared to α-synuclein.

Hydration water dynamics is increased in α-synuclein fibers compared to monomers

Both MSD determined by ENS (Figure 2a) and QENS spectra (Figure 2b) indicate hydration water dynamics is enhanced in α-synuclein fibers. Fitting the QENS data with a model involving three water populations, namely immobile, translating and rotating molecules[37], allowed extracting the fraction of each population and the translational and rotational diffusion constants (Figure S2b). The main origin of enhanced water dynamics in fibrils is an increased diffusion constant of those water molecules that mainly undergo translational motions.

A calculation of water density from the MD trajectories reveals a dehydration of the NAC region when fibers form (Figure 4). A structural explanation of this water reorganization can be provided by the conformational changes that the protein undergoes when forming fibrils (Figure S7). The highly compact nature of the monomer contrains water molecules to be close to the hydrophobic residues of the NAC segment. Upon fibril formation, an extensive contact breaking occurs between the NAC segment and the N-ter and CTT regions, resulting in a more extended structure. Water interacting with both the N-terminal and the NAC fragments in the monomers is pulled away from the NAC fragment upon fibrillation, reducing water density of the latter as seen in figure 4. This water restructuring upon fibrillation results in reduced confinement of hydration water and thus in higher dynamics and entropy in the fibers.

Increase of water mobility correlates with fraction of disordered regions in tau and α-synuclein fibers

Interestingly, the increase in hydration water MSD (Figure S4) and the broadening of the QENS spectra (Figure 2) upon fibrillation is found to differ between α-synuclein and tau. The increase in MSD is twice as important for the latter than for the former. Both proteins are highly compact in their monomeric state[42,43]. In the fiber state, a fraction of the amino acid sequence remains disordered and extends from the well-ordered fiber state in both proteins[6–8,24,25]. This fraction is much smaller for α-synuclein fibers (28% disordered residues, 72% core-residues[26]) than for tau fibers (85% disordered, 15% core). Hence, assuming that the same water restructuring observed in α-synuclein occurs in tau, a larger fraction of hydration water is displaced from hydrophobic core upon fibrillation for tau than for the α-synuclein. This results in the observed larger increase of water dynamics around tau fibers than around α-synuclein fibers. Our data thus suggest that the larger the ratio between disordered and core residues is in amyloid fibrils, the greater the increase is in water dynamics and entropy upon fibrillation.

Materials and Methods

Expression and purification of hydrogenated and perdeuterated α-synuclein. The transformed *E. coli* BL21 (DE3) cells were incubated in H or D-Enfors medium with kanamycin at a concentration of 30 µg/mL. The 150 mL pre-culture was then transferred to a 1.2 L fermenter (Labfors HT Infors) using 120 g/L glycerol in H or D-Enfors as the feeding solution. The bacteria were grown at 30% O_2 level with 30 µg/mL of kanamycin at pH 6.9. Once the culture reached an optical density of 10 at a wavelength 600 nm, the induction was done with 0.5 mM IPTG as final concentration and the fermentation was continued until we reached an OD (600 nm) of 14-18. This resulted in 59.5 and 36.2 g of cell paste for the hydrogenated and the deuterated sample, respectively.

The cell lysis was carried out with 5 g of cells and 30 mL of lysis buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA + 1 µL benzonase). The cells were sonicated five times for 30 seconds at 60% power using a Vibracell 75115 Bioblock Scientific sonicator. The lysate was centrifuged at 18,000 rpm for 30 minutes, and then the supernatant was heated at 90°C for 10 minutes and centrifuged again under the same conditions. The supernatant was then filtered at 0.8 µm and loaded on a 5 mL DEAE Sephacel column pre-equilibrated in buffer A (10 mM Tris pH 7.5). The protein was eluded using buffer B (10 mM Tris pH 7.5, 2 M NaCl) with a step at 11% for 40 mL, and then a gradient was applied from 11 to 100% to remove the residuals. Eventually, the fractions of interest were loaded on a S200 16/600 column (GF - SEC) without any concentration step with buffer A. The fractions containing purified α-synuclein were pooled and dialyzed three times against pure water using a 10 kDa dialysis membrane. Subsequently the protein was flash-cooled in 1 mL volume using liquid nitrogen and lyophilized in a CHRIST alpha 2-4 LD plus lyophilizer with the temperature of the ice trap set at -50°C and the pressure at 0.06 mbar. Importantly, oligomers formed when the sample was slow cooled in a -20°C freezer instead of flash-cooling in liquid nitrogen, where less oligomers were present (Figure S1). The final amount of protein was calculated using a mass absorption coefficient of 0.412 (~ ε / Molar Mass = 5400 / 14000). In addition, the tubes were weighed before and after lyophilization of the protein to confirm the UV measurement. We thus obtain the dry powders of hydrogenated and deuterated monomeric α-synuclein, H-mono-αsyn and D-mono-αsyn, respectively.

α-synuclein fiber formation. To prepare the fibrils of α-synuclein, approximately 160 mg of the powders of hydrogenated and deuterated α-synuclein were dissolved at 10 mg/mL in a 20 mM sodium phosphate H_2O buffer at pH 7.4 with 50 (for hydrogenated α-synuclein) and 150 (for deuterated α-synuclein) mM NaCl. The protein solution is then filtered at 0.22 µm and incubated in 1 mL fractions in Eppendorf tubes at 37°C with cycles of 300 s of mixing at 900 rpm with 60 s of pause during 48 h using an Eppendorf ThermoMixer® Comfort. The fibrils were then pelleted by centrifugation at 16,500 rpm during 15 minutes and resuspended, in pure D_2O for hydrogenated α-synuclein fibers or in H_2O for deuterated α-synuclein fibers, three times to remove salts. Finally, the fibers of hydrogenated α-synuclein and deuterated α-synuclein were lyophilized using flash-cooling in liquid nitrogen. We thus obtain 100 mg of the deuterated and hydrogenated fibrillar α-synuclein powders, D-fib-αsyn and H-fib-αsyn respectively.

Characterization of α-synuclein monomers and fibers by electron microscopy and X-ray diffraction.

Electron microscopy

A fraction of each sample (D-mono-αsyn, H-mono-αsyn, D-fib-αsyn and H-mono-αsyn) was solubilized in H_2O right before imaging with EM. Samples were absorbed to the clean side of a carbon film, that is deposited on mica, stained with 2% (w/v) uranyle acetate (pH 4.5) and transferred to a 400-mesh copper grid. The images were taken under low dose conditions (<10 $e^-/Å^2$) with defocus values between 1.2 and 2.5 µm on a Tecnai 12 LaB6 electron microscope at 120 kV accelerating voltage using a Gatan Orius 1000 CCD camera. Electron microscope images still show small objects in D- and H-mono-αsyn samples that were flash-cooled in liquid nitrogen and lyophilized (Figure S1) that might be small aggregates or contaminants that we removed using 10 minutes centrifugation at 16000 rpm prior to sample lyophilization.

X-ray diffraction

A fraction of each sample (D-mono-αsyn, H-mono-αsyn, D-fib-αsyn and H-fib-αsyn) was analysed by X-ray diffraction. The experimental protocol was the same for all the samples: about 1 mg of sample was sealed in a Mylar capillary and diffraction patterns were recorded at room temperature at the beamline ID30A of the European Synchrotron Radiation Facility (France). A monochromatic X-ray beam (λ = 0.98 Å) with an average flux of 4×10^{12} ph/s was used to collect 20 diffraction images. Every image was obtained by exposing the sample to the X-ray beam during 1 s, then the sample was rotated by 1° before acquiring the subsequent image. The 20 images were then averaged to obtain the diffraction patterns presented in Figure 1 for H-fib-αsyn-D_2O and D-fib-αsyn-H_2O, and in Figure S3 for H-mono-αsyn-D_2O and D-mono-αsyn-H_2O.

α-synuclein sample preparation for neutron scattering experiments. A mass of about 100 mg of each sample after freeze-drying (*i.e.* D-mono-αsyn, H-mono-αsyn, D-fib-αsyn and H-fib-αsyn) was deposited on a 4 x 3 cm^2 flat aluminum sample holder base used for neutron scattering and left dehydrating in the presence of P_2O_5 in a closed desiccator for 24 h at room temperature. A similar drying procedure over P_2O_5 for lysozyme has been described to result in four bound water molecules per protein molecule[44]. After the 24 h incubation with P_2O_5 the mass of the samples was determined using an analytical scale with 0.1 mg precision and assumed to correspond to the dry protein mass (0 g water / g protein). The samples were then placed in the desiccator in the presence of H_2O (for D-mono-αsyn and D-fib-αsyn) or D_2O (for H-mono-αsyn and H-fib-αsyn) and their mass monitored until they reached the hydration level h=0.4 [g H_2O]/[g protein] (for D-mono-αsyn and D-fib-αsyn) or h=0.44 [g D_2O]/[g protein] (for H-mono-αsyn and H-fib-αsyn). The different hydration levels (i.e. h=0.4 and h=0.44) were chosen according to the different molecular weights of H_2O and D_2O to guarantee a similar number of water molecules per protein in all the samples. The sample holders were then sealed using a 1 mm indium O-ring and stored at - 20° C. We thus obtain the H_2O hydrated D-mono-αsyn-H_2O and D-fib-αsyn-H_2O powders and the D_2O hydrated H-mono-αsyn-D_2O and H-fib-αsyn-D_2O powders.

Tau samples for neutron scattering experiments

The D-mono-tau-H_2O[45] (h = 0.38) and D-fib-tau-H_2O[16] (h = 0.40) samples had been prepared for an earlier study and kept at -20°C since then without opening the neutron sample holders. Their weights were unchanged compared to the ones measured immediately after closing the sample holders, indicating that the sample hadn't lost hydration water over the past years. Sample preparation details have been published earlier and the contributions of water and protein atoms to the coherent and incoherent scattering signals are listed in table S1.

Neutron scattering. Elastic and quasi-elastic neutron scattering experiments were performed at the SPHERES backscattering spectrometer[46] of the Heinz Meier-Leibnitz Zentrum (MLZ) in Garching, Germany (proposal #13899 from the 28 May to the 2 June 2018 for ENS and QENS on H-mono-αsyn-D_2O and H-fib-αsyn-D_2O; proposal #15611 from 12 to 17 February 2020 for QENS on D-mono-αsyn-H_2O, D-mono-αsyn-H_2O, D-mono-tau-H_2O and D-fib-tau-H_2O. The ENS data for D-mono-tau-H_2O and D-fib-tau-H_2O have been measured earlier[16,45]). SPHERES is equipped with a mechanical Doppler drive to set the incident energy and a phase space transformation (PST) chopper to optimize the neutron flux at the sample position[47]. The sample holders were weighed before and after each neutron experiment to ensure that no dehydration due to possible leaks had occurred. The sample holder was mounted in a cryostat at room temperature and the temperature was subsequently lowered within 120 minutes to 10 K. For each of the four α-synuclein samples, an elastic neutron scattering scan was performed from 10 to 300 K at a heating rate of 0.2 K/min. Raw data were pre-processed using the SLAW software available on MLZ facility computers (http://apps.jcns.fz-juelich.de/slaw).

Atomic mean-square displacements of hydrogen atoms of H-mono-αsyn-D_2O, H-fib-αsyn-D_2O, D-mono-αsyn-H_2O and D-fib-αsyn-H_2O were determined using a Gaussian model[40] to fit the experimental elastic neutron scattering intensity in the momentum transfer range $0.22 < q < 0.96$ Å^{-1} with the Scipy python library 'curve_fit' routine. The water dynamics for D-mono-αsyn-H_2O and D-fib-αsyn-H_2O have left the experimental temporal and spatial resolution at temperatures above 260 K and the MSDs cannot be confidently determined anymore. Hence, all the data on hydration water dynamics were analyzed up to 260 K only and QENS data were acquired at that temperature (see below). To the contrary, motions in H-mono-αsyn-D_2O and H-fib-αsyn-D_2O samples remained within the experimental window up to 300 K.

The ENS data for D-mono-tau-H2O[45] and D-fib-tau-H2O[16] were reprocessed in the same way the the ENS data for D-mono-αsyn-H_2O and D-fib-αsyn-H_2O described above.

The Quasi-elastic neutron scattering (QENS) spectra for D-mono-αsyn-H_2O, D-fib-αsyn-H_2O, D-mono-tau-H_2O and D-fib-tau-H_2O were recorded for 12 hours using an energy range -15 μeV $< E <$ 15 μeV at 260 K. The QENS resolution function $R(q,\omega)$ for each sample was determined using the QENS spectra from the same sample at 10 K, fitted with a normalized pseudo-Voigt profile multiplied by a scaling factor. This scaling factor was subsequently used to normalize the data at 260 K, which were subsequently binned using a bin window of 5 points. The QENS spectra from D-mono-αsyn-H_2O and D-fib-αsyn-H_2O were fitted with the following model function[37]:

$$S(q,\omega) = DW(q)R(q,\omega) \otimes \left[A_0\delta(\omega) + \sum_{l}^{N} (2l+1) j_l^2(qd) \frac{A_r}{\pi} \frac{l(l+1)}{\omega^2 + [l(l+1)\Gamma_r^2]^2} + \frac{A_t}{\pi} \frac{\Gamma_t}{\omega^2 + \Gamma_t^2} \right] + b(q)$$

(eq. 1)

Here, $DW(q) = \exp^{-(MSD)q^2/3}$ is the Debye-Waller factor accounting for fast vibrational motions with (MSD) being the mean-squared displacement for these motions, $R(q,\omega)$ denotes the resolution function, A_0 is the elastic incoherent structure factor given by $A_0 = A_r j_0^2(qd) + A_i$, j_l denotes the spherical Bessel function of order l, *d* is the length of an O-H bond in a water molecule, A_i, A_r and A_t are scalars, Γ_r and Γ_t are Lorentzian widths for rotational and translational motions where $\Gamma_t = D_t q^\tau$, and *b*(*q*) is a *q*-dependent background.

The QENS spectra for H-mono-αsyn-D_2O, H-fib-αsyn-D_2O were recorded for 8 hours using an energy range -15 µeV < E < 15 µeV at 300 K. The QENS resolution function $R(q,\omega)$ for each sample was determined using the QENS spectra from the same sample at 10 K fitted with a normalized pseudo-Voigt profile multiplied by a scaling factor. This scaling factor was subsequently used to normalize the data at 300 K, which were subsequently binned using a bin window of 5 points.

The QENS spectra for H-mono-αsyn-D_2O and H-fib-αsyn-D_2O were analyzed using the following model containing two Lorentzians that accounts for two fractions of atoms, the motion of which obeys Fickian-type diffusion $\Gamma_i = D_i q^2,\ i \in \{1,2\}$:

$$S(q,\omega) = DW(q)R(q,\omega) \otimes [a_0\delta(\omega) + a_1\mathcal{L}_{\Gamma_1} + a_2\mathcal{L}_{\Gamma_2}] + b(q) \quad \text{(eq. 2)}$$

where a_0, a_1 and a_2 are scalars, $\mathcal{L}_{\Gamma_1}$ is a Lorentzian of width Γ_1, $\mathcal{L}_{\Gamma_2}$ is a Lorentzian of width Γ_2 and $b(q)$ a q-dependent background.

The data are fitted for all scattering vectors at once (global fits) using the Scipy *basinhopping* algorithm[48] with 100 iterations. All described analyses were performed by a custom-made python API called nPDyn (https://github.com/kpounot/nPDyn).

Relative contributions of protein and water to the neutron scattering signal. For all samples, the number of each element (H, D, C, N, O - D is used for ^{2}H) is obtained using the amino acids sequence and considering that all exchangeable hydrogens (amino-, hydroxy- and carboxy-bounded hydrogens) are exchanged against the isotope of the hydration water. The mass (MW) of the protein is then computed based on the element composition. Subsequently, the hydration level h that is used experimentally allows to determine the number of water molecules per protein monomer $N = h.MW/MW_{water}$ where MW_{water} is 18 Da for H_2O and 20 Da for D_2O. The neutron scattering cross-section for water is computed using $\sigma_X^{water} = \sum_{i\in\{H,D,C,N,O\}} n_i^{water}\sigma_X^i$ where X stands for coherent or incoherent cross-section, n_i^{water} is the number of element i in water and σ_X^i is the X (coherent or incoherent) cross-section of element i for thermal neutrons[49]. The neutron scattering cross-section for protein is computed the same way. In Table S1, the element composition, the protein mass, the number of water molecules per protein monomer and the coherent and incoherent cross-sections for water and protein are given for all the samples used in this work.

Generation of models for MD simulations. Initial ensemble for monomeric α-synuclein were obtained from a previous work50. Twenty monomers were randomly selected and placed into a box, hydrated at h=0.4 [g H_2O / g protein] with TIP4P[51,52] water, by deleting extra water molecules that were further than a given distance from protein surface until the desired hydration level is reached. Finally, charges were neutralized with Na^+ and Cl^- ions using the VMD autoionize plugin. This process was performed using VMD software53. For the fibrils, we used the structure of full-length α-synuclein published by Tuttle and coworkers6. Two of these partial fibrils were placed into a box, hydrated at h=0.4 [g H_2O / g protein] and neutralized in the same way as the monomers.

Molecular dynamics simulations and analysis. Using NAMD[54], a 2 ns simulation was run at a constant temperature of 300 K and a constant pressure of 1 atm, using harmonic restraints on the fibril NAC region (residues from 61 of 95) with a force constant of 2 kcal/(mol·Å^2). Pressure was controlled using the Nosé-Hoover Langevin-Piston algorithm[55,56], and the temperature using the stochastic velocity rescaling method[57]. The CHARMM36 force field[58] was used for the protein. Electrostatic interactions were computed using the particle-mesh Ewald summation[59,60]. The real space of electrostatic and Van der Waals forces was smoothly switched off between 12 and 14 Å, and the pairlist cutoff was 16 Å. The equations of motion were integrated using the Verlet-I/r-RESPA multiple time step algorithm[61,62] with time steps of 4 fs for the long-range non-bonded forces, 2 fs for the short-range non-bonded forces, and 2 fs for the bonded intramolecular forces. Lengths of bonds involving hydrogen atoms were held fixed using the SHAKE algorithm[62].

The initial constant pressure equilibration at 1 atm was followed by 2 ns simulation at 100 atm pressure, keeping the other parameters fixed, in order to relax bad contacts and eliminate gaps that could have been generated by the fast box collapse in the first step. The pressure was then set back to 1 atm, and a 100 ns equilibration was performed at each temperature considered (50, 110, 200, 230, 260, and 300 K). These equilibrated systems formed the finalized initial models for subsequent constant volume and energy production runs of 20 ns duration, during which configurations were saved at 20 ps intervals. The simulation trajectories were analyzed using the custom-made API, NAMDAnalyzer (https://github.com/kpounot/NAMDAnalyzer).

To compare the experimental ENS data with the simulations, the MSDs of either hydrogen atoms of water or non-exchangable hydrogens of the protein were computed using the following

$$MSD(t) = \frac{1}{N}\sum_{k=1}^{N}\langle|\mathbf{r}_k(t) - \mathbf{r}_k(0)|^2\rangle \qquad \text{(eq. 3)}$$

where N is the number of atoms considered, $r_k(t)$ is the kth atom position vector at time t, and the angular brackets denote average over the time origins.

To compare the experimental QENS data with the simulations result, the intermediate scattering function was computed using

$$I(q,t) = \frac{1}{N}\sum_{i=1}^{N}\langle \exp^{-i\mathbf{q}\cdot(\mathbf{r}_i(t)-\mathbf{r}_i(0))}\rangle \qquad \text{(eq. 4)}$$

where $\mathbf{r}_i(t)$ denotes the ith atom position vector at time t, $\mathbf{q}$ the momentum transfer vector (an average is performed over 20 randomly oriented vectors of magnitude q), the angular brackets the average over 50 time origins, and N the total number of atom used. The computed QENS spectra are obtained by computing the Fourier transform in the frequency domain of $I(q,t)$, the modulus of which is squared to obtain the power spectrum that is plotted.

The hydrogen bonds were identified based on geometrical criteria, that is, a maximum distance of 2.5 Å between the hydrogen of the donor and the acceptor and an angle of minimum 130° between the donor-hydrogen vector and the acceptor-donor vector. The hydrogen bond autocorrelation was computed using

$$C(t) = \left\langle \sum_{i,j} \frac{h_{ij}(t)h_{ij}(0)}{h_{ij}^2(0)} \right\rangle \quad \text{(eq. 5)}$$

where $h_{ij}(t)$ is 1 if a hydrogen bond can be formed between atom pair ij at time t, and 0 otherwise. In the continuous case, if the bond has been broken at any time t, $h_{ij}(t)$ is set to 0 for all subsequent time.

The radial number density is computed using

$$\rho(r) = \left\langle \frac{1}{N} \sum_{i}^{N} \frac{1}{4\pi r^2 dr} [n_i(r+dr) - n_i(r)] \right\rangle \quad \text{(eq. 6)}$$

where N is the number of atoms from which radial density is computed, r is the radius of the selection sphere from the center of atom i, dr is the increment of radius, and $n_i(r)$ is the number of atoms in the user selection that are located within the sphere of radius r centered on atom i. The angular brackets denote the average over multiple frames.

Acknowledgments

This work was partially carried out at the platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The electron microscope facility is supported by the Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER, the Centre National de la Recherche Scientifique (CNRS), the Commissariat à l'Energie Atomique (CEA), the University of Grenoble, EMBL, and the GIS-Infrastrutures en Biologie Sante et Agronomie (IBISA). The Heinz Maier-Leibnitz Zentrum (MLZ), Garching, Germany is acknowledged for the allocation of neutron scattering beam time. We thank Daphna Fenel, Christine Moriscot and Guy Schoehn for having recorded electron microscopy data. KP was supported by a fellowship co-funded by the ILL Graduate School and the CEA IRTELIS program and by a travel grant from the 'Initiative d'excellence' (IDEX) of the Grenoble University to spend a 3-month research period at UC Irvine.

Author contributions

K.P. and M.M. expressed and purified the α-synuclein, K.P. produced the monomeric samples and the fibers samples, G.S., K.P., M.Z. and D.N. performed the neutron scattering experiment, K.P., G.S. and T.S. analyzed the neutron scattering data, K.P. and G.S. collected the X-ray powder diffraction patterns, K.P. and D.T. constructed the powder models, performed the MD simulations and analyzed the data, K.P. and M.W. wrote the paper.

References

(1) Stefani, M. Protein Misfolding and Aggregation: New Examples in Medicine and Biology of the Dark Side of the Protein World. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2004**, *1739* (1), 5–25. https://doi.org/10.1016/j.bbadis.2004.08.004.

(2) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry* **2006**, *75* (1), 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901.

(3) Mandelkow, E.-M.; Mandelkow, E. Tau in Alzheimer's Disease. *Trends in Cell Biology* **1998**, *8* (11), 425–427. https://doi.org/10.1016/S0962-8924(98)01368-3.

(4) Hamley, I. W. The Amyloid Beta Peptide: A Chemist's Perspective. Role in Alzheimer's and Fibrillization. *Chem. Rev.* **2012**, *112* (10), 5147–5192. https://doi.org/10.1021/cr3000994.

(5) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. Alpha-Synuclein in Lewy Bodies. *Nature* **1997**, *388* (6645), 839–840. https://doi.org/10.1038/42166.

(6) Tuttle, M. D.; Comellas, G.; Nieuwkoop, A. J.; Covell, D. J.; Berthold, D. A.; Kloepper, K. D.; Courtney, J. M.; Kim, J. K.; Barclay, A. M.; Kendall, A.; et al. Solid-State NMR Structure of a Pathogenic Fibril of Full-Length Human α-Synuclein. *Nature Structural & Molecular Biology* **2016**, *23* (February), 1–9. https://doi.org/10.1038/nsmb.3194.

(7) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. W. Cryo-EM Structures of Tau Filaments from Alzheimer's Disease. *Nature Publishing Group* **2017**, *547*, 185–190. https://doi.org/10.1038/nature23002.

(8) Falcon, B.; Zivanov, J.; Zhang, W.; Murzin, A. G.; Garringer, H. J.; Vidal, R.; Crowther, R. A.; Newell, K. L.; Ghetti, B.; Goedert, M.; et al. Novel Tau Filament Fold in Chronic Traumatic Encephalopathy Encloses Hydrophobic Molecules. *Nature* **2019**, *568* (7752), 420–423. https://doi.org/10.1038/s41586-019-1026-5.

(9) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; et al. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447* (7143), 453–457. https://doi.org/10.1038/nature05695.

(10) Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; et al. Structure of the Toxic Core of α-Synuclein from Invisible Crystals. *Nature* **2015**, *525* (7570), 486–490. https://doi.org/10.1038/nature15368.

(11) Eisenberg, D.; Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **2012**, *148* (6), 1188–1203. https://doi.org/10.1016/j.cell.2012.02.022.

(12) Stephens, A. D.; Kaminski Schierle, G. S. The Role of Water in Amyloid Aggregation Kinetics. *Current Opinion in Structural Biology* **2019**, *58*, 115–123. https://doi.org/10.1016/j.sbi.2019.06.001.

(13) Pavlova, A.; Cheng, C.-Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S. Protein Structural and Surface Water Rearrangement Constitute Major Events in the Earliest Aggregation Stages of Tau. *PNAS* **2016**, *113* (2), E127–E136. https://doi.org/10.1073/pnas.1504415113.

(14) Wang, T.; Jo, H.; DeGrado, W. F.; Hong, M. Water Distribution, Dynamics, and Interactions with Alzheimer's β-Amyloid Fibrils Investigated by Solid-State NMR. *J. Am. Chem. Soc.* **2017**, *139* (17), 6242–6252. https://doi.org/10.1021/jacs.7b02089.

(15) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012**, *45* (1), 83–92. https://doi.org/10.1021/ar2000869.

(16) Fichou, Y.; Schirò, G.; Gallat, F.-X.; Laguri, C.; Moulin, M.; Combet, J.; Zamponi, M.; Härtlein, M.; Picart, C.; Mossou, E.; et al. Hydration Water Mobility Is Enhanced around Tau Amyloid Fibers. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112* (20), 6365–6370. https://doi.org/10.1073/pnas.1422824112.

(17) Arya, S.; Singh, A. K.; Khan, T.; Bhattacharya, M.; Datta, A.; Mukhopadhyay, S. Water Rearrangements upon Disorder-to-Order Amyloid Transition. *J. Phys. Chem. Lett.* **2016**, *7* (20), 4105–4110. https://doi.org/10.1021/acs.jpclett.6b02088.

(18) Liddle, R. A. Parkinson's Disease from the Gut. *Brain Research* **2018**, *1693*, 201–206. https://doi.org/10.1016/J.BRAINRES.2018.01.010.

(19) Decressac, M.; Volakakis, N.; Björklund, A.; Perlmann, T. NURR1 in Parkinson Disease—from Pathogenesis to Therapeutic Potential. *Nature Reviews Neurology* **2013**, *9* (11), 629–636. https://doi.org/10.1038/nrneurol.2013.209.

(20) Stephens, A. D.; Zacharopoulou, M.; Moons, R.; Fusco, G.; Seetaloo, N.; Chiki, A.; Hooper, P. J.; Mela, I.; Lashuel, H. A.; Philips, J. J.; et al. Extent of N-Terminus Exposure by Altered Long-Range Interactions of Monomeric Alpha-Synuclein Determines Its Aggregation Propensity. *bioRxiv* **2019**, 740241. https://doi.org/10.1101/740241.

(21) Emamzadeh, F. N. Alpha-Synuclein Structure, Functions, and Interactions. *J Res Med Sci* **2016**, *21*. https://doi.org/10.4103/1735-1995.181989.

(22) Ait-Bouziad, N.; Lv, G.; Mahul-Mellier, A.-L.; Xiao, S.; Zorludemir, G.; Eliezer, D.; Walz, T.; Lashuel, H. A. Discovery and Characterization of Stable and Toxic Tau/Phospholipid Oligomeric Complexes. *Nature Communications* **2017**, *8* (1), 1678. https://doi.org/10.1038/s41467-017-01575-4.

(23) Pinotsi, D.; Michel, C. H.; Buell, A. K.; Laine, R. F.; Mahou, P.; Dobson, C. M.; Kaminski, C. F.; Kaminski Schierle, G. S. Nanoscopic Insights into Seeding Mechanisms and Toxicity of α-Synuclein Species in Neurons. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, 1–5. https://doi.org/10.1073/pnas.1516546113.

(24) Li, Y.; Zhao, C.; Luo, F.; Liu, Z.; Gui, X.; Luo, Z.; Zhang, X.; Li, D.; Liu, C.; Li, X. Amyloid Fibril Structure of α-Synuclein Determined by Cryo-Electron Microscopy. *Cell Research* **2018**, *28* (9), 897–903. https://doi.org/10.1038/s41422-018-0075-x.

(25) Li, B.; Ge, P.; Murray, K. A.; Sheth, P.; Zhang, M.; Nair, G.; Sawaya, M. R.; Shin, W. S.; Boyer, D. R.; Ye, S.; et al. Cryo-EM of Full-Length α-Synuclein Reveals Fibril Polymorphs with a Common Structural Kernel. *Nature communications* **2018**, *9* (1), 3609. https://doi.org/10.1038/s41467-018-05971-2.

(26) Guerrero-Ferreira, R.; Taylor, N. M.; Mona, D.; Ringler, P.; Lauer, M. E.; Riek, R.; Britschgi, M.; Stahlberg, H. Cryo-EM Structure of Alpha-Synuclein Fibrils. *eLife 7*. https://doi.org/10.7554/eLife.36402.

(27) Bertoncini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. Release of Long-Range Tertiary Interactions Potentiates Aggregation of Natively Unstructured α-Synuclein. *Proceedings of the National Academy of Sciences* **2005**, *102* (5), 1430–1435. https://doi.org/10.1073/pnas.0407146102.

(28) Xu, L.; Ma, B.; Nussinov, R.; Thompson, D. Familial Mutations May Switch Conformational Preferences in α-Synuclein Fibrils. *ACS Chem. Neurosci.* **2017**, *8* (4), 837–849. https://doi.org/10.1021/acschemneuro.6b00406.

(29) Stephens, A. D.; Zacharopoulou, M.; Kaminski Schierle, G. S. The Cellular Environment Affects Monomeric α-Synuclein Structure. *Trends in Biochemical Sciences* **2019**, *44* (5), 453–466. https://doi.org/10.1016/j.tibs.2018.11.005.

(30) Arya, S.; Singh, A. K.; Bhasne, K.; Dogra, P.; Datta, A.; Das, P.; Mukhopadhyay, S. Femtosecond Hydration Map of Intrinsically Disordered α-Synuclein. *Biophysical Journal* **2018**, *114* (11), 2540–2551. https://doi.org/10.1016/j.bpj.2018.04.028.

(31) Gabel, F.; Bicout, D.; Lehnert, U.; Tehei, M.; Weik, M.; Zaccai, G. Protein Dynamics Studied by Neutron Scattering. *Q Rev Biophys* **2002**, *35* (4), 327–367. https://doi.org/10.1017/S0033583502003840.

(32) Grimaldo, M.; Roosen-Runge, F.; Zhang, F.; Schreiber, F.; Seydel, T. Dynamics of Proteins in Solution. *Quarterly Reviews of Biophysics* **2019**, *52*. https://doi.org/10.1017/S0033583519000027.

(33) Haertlein, M.; Moulin, M.; Devos, J. M.; Laux, V.; Dunne, O.; Trevor Forsyth, V. Chapter Five - Biomolecular Deuteration for Neutron Structural Biology and Dynamics. In *Methods in Enzymology*; Kelman, Z., Ed.; Isotope Labeling of Biomolecules - Applications; Academic Press, 2016; Vol. 566, pp 113–157. https://doi.org/10.1016/bs.mie.2015.11.001.

(34) Wood, K.; Plazanet, M.; Gabel, F.; Kessler, B.; Oesterhelt, D.; Tobias, D. J.; Zaccai, G.; Weik, M. Coupling of Protein and Hydration-Water Dynamics in Biological Membranes. *Proceedings of the National Academy of Sciences* **2007**, *104* (46), 18049–18054. https://doi.org/10.1073/pnas.0706566104.

(35) Doster, W.; Cusack, S.; Petry, W. Dynamical Transition of Myoglobin Revealed by Inelastic Neutron Scattering. *Nature* **1989**, *337*, 754–756. https://doi.org/10.1038/337754a0.

(36) Schirò, G.; Weik, M. Role of Hydration Water in the Onset of Protein Structural Dynamics. *J. Phys.: Condens. Matter* **2019**, *31* (46), 463002. https://doi.org/10.1088/1361-648X/ab388a.

(37) Schirò, G.; Fichou, Y.; Gallat, F.-X.; Wood, K.; Gabel, F.; Moulin, M.; Härtlein, M.; Heyden, M.; Colletier, J.-P.; Orecchini, A.; et al. Translational Diffusion of Hydration Water Correlates with Functional Motions in Folded and Intrinsically Disordered Proteins. *Nature communications* **2015**, *6*, 6490. https://doi.org/10.1038/ncomms7490.

(38) Schiró, G.; Vetri, V.; Frick, B.; Militello, V.; Leone, M.; Cupane, A. Neutron Scattering Reveals Enhanced Protein Dynamics in Concanavalin a Amyloid Fibrils. *Journal of Physical Chemistry Letters* **2012**, *3* (8), 992–996. https://doi.org/10.1021/jz300082x.

(39) Eanes, E. D.; Glenner, G. G. X-Ray Diffraction Studies on Amyloid Filaments. *J Histochem Cytochem.* **1968**, *16* (11), 673–677. https://doi.org/10.1177/16.11.673.

(40) Yi, Z.; Miao, Y.; Baudry, J.; Jain, N.; Smith, J. C. Derivation of Mean-Square Displacements for Protein Dynamics from Elastic Incoherent Neutron Scattering. *Journal of Physical Chemistry B* **2012**, *116* (16), 5028–5036. https://doi.org/10.1021/jp2102868.

(41) Theillet, F.-X.; Binolfi, A.; Bekei, B.; Martorana, A.; Rose, H. M.; Stuiver, M.; Verzini, S.; Lorenz, D.; van Rossum, M.; Goldfarb, D.; et al. Structural Disorder of Monomeric α-Synuclein Persists in Mammalian Cells. *Nature* **2016**, *530* (7588), 45–50. https://doi.org/10.1038/nature16531.

(42) Brodie, N. I.; Popov, K. I.; Petrotchenko, E. V.; Dokholyan, N. V.; Borchers, C. H. Conformational Ensemble of Native α-Synuclein in Solution as Determined by Short-Distance Crosslinking Constraint-Guided Discrete Molecular Dynamics Simulations. *PLOS Computational Biology* **2019**, *15* (3), e1006859. https://doi.org/10.1371/journal.pcbi.1006859.

(43) Popov, K. I.; Makepeace, K. A. T.; Petrotchenko, E. V.; Dokholyan, N. V.; Borchers, C. H. Insight into the Structure of the “Unstructured” Tau Protein. *Structure* **2019**, *27* (11), 1710-1715.e4. https://doi.org/10.1016/j.str.2019.09.003.

(44) Dolman, M.; Halling, P. J.; Moore, B. D.; Waldron, S. How Dry Are Anhydrous Enzymes? Measurement of Residual and Buried 18O-Labeled Water Molecules Using Mass Spectrometry. *Biopolymers* **1997**, *41* (3), 313–321. https://doi.org/10.1002/(SICI)1097-0282(199703)41:3<313::AID-BIP6>3.0.CO;2-V.

(45) Gallat, F.-X.; Laganowsky, A.; Wood, K.; Gabel, F.; van Eijck, L.; Wuttke, J.; Moulin, M.; Härtlein, M.; Eisenberg, D.; Colletier, J.-P.; et al. Dynamical Coupling of Intrinsically Disordered Proteins and Their Hydration Water: Comparison with Folded Soluble and Membrane Proteins. *Biophysical Journal* **2012**, *103* (1), 129–136. https://doi.org/10.1016/j.bpj.2012.05.027.

(46) Wuttke, J.; Budwig, A.; Drochner, M.; Kämmerling, H.; Kayser, F. J.; Kleines, H.; Ossovyi, V.; Pardo, L. C.; Prager, M.; Richter, D.; et al. SPHERES, Jülichs High-Flux Neutron Backscattering Spectrometer at FRM II. *Review of Scientific Instruments* **2012**, *83* (7). https://doi.org/10.1063/1.4732806.

(47) Kirstein, O.; Prager, M.; Kozielewski, T.; Richter, D. Phase Space Transformation Used at the FRM II Backscattering Spectrometer: Concepts and Technical Realization. *Physica B: Condensed Matter* **2000**, *283* (4), 361–364. https://doi.org/10.1016/S0921-4526(00)00340-9.

(48) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. https://doi.org/10.1021/jp970984n.

(49) Sears, V. F. Neutron Scattering Lengths and Cross Sections. *Neutron News* **1992**, *3* (3), 26–37. https://doi.org/10.1080/10448639208218770.

(50) Schwalbe, M.; Ozenne, V.; Bibow, S.; Jaremko, M.; Jaremko, L.; Gajda, M.; Jensen, M. R.; Biernat, J.; Becker, S.; Mandelkow, E.; et al. Predictive Atomic Resolution Descriptions of Intrinsically Disordered HTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. *Structure* **2014**, *22* (2), 238–249. https://doi.org/10.1016/j.str.2013.10.020.

(51) Jorgensen, W. L.; Madura, J. D. Temperature and Size Dependence for Monte Carlo Simulations of TIP4P Water. *Molecular Physics* **2006**. https://doi.org/10.1080/00268978500103111.

(52) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(53) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(54) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26* (16), 1781–1802. https://doi.org/10.1002/jcc.20289.

(55) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *The Journal of Chemical Physics* **1994**, *101* (5), 4177–4189. https://doi.org/10.1063/1.467468.

(56) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *The Journal of Chemical Physics* **1995**, *103* (11), 4613–4621. https://doi.org/10.1063/1.470648.

(57) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *The Journal of Chemical Physics* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.

(58) Huang, J.; MacKerell, A. D.; Jr. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *Journal of computational chemistry* **2013**, *34* (25), 2135–45. https://doi.org/10.1002/jcc.23354.

(59) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *The Journal of Chemical Physics* **1995**, *103* (19), 8577–8593. https://doi.org/10.1063/1.470117.

(60) Wells, B. A.; Chaffee, A. L. Ewald Summation for Molecular Simulations. *Journal of Chemical Theory and Computation* **2015**, *11* (8), 3684–3695. https://doi.org/10.1021/acs.jctc.5b00093.

(61) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-Range Interactions. *Molecular Simulation* **1991**, *6* (1–3), 121–142. https://doi.org/10.1080/08927029108022142.

(62) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal of Computational Physics* **1977**, *23* (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

SLAW software:

Wuttke, J.: SLAW – a neutron histogram to scattering law converter, http://apps.jcns.fz-juelich.de/slaw

Supporting figures and tables

Water restructuring upon α-synuclein fibril formation induces an increase of dynamics and entropy

Kevin Pounot[1,2], Giorgio Schiro[1], Martine Moulin[3], Carlotta Marasini[4], Daria Noferini[5], Michaela Zamponi[5], Michael Haertlein[3], Trevor Forsyth[3], Martin Blackledge[1], Bente Vestergaard[4], Tilo Seydel[2], Annette E. Langkilde[4], Douglas J. Tobias[6], Martin Weik[1]

[1] Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

[2] Institut Max von Laue – Paul Langevin (ILL), CS 20156, F-38042 Grenoble, France

[3] Life Sciences Group, Institut Laue–Langevin, 6 Rue Jules Horowitz, 38042 Grenoble, France

[4]Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark

[5]Jülich Centre for Neutron Science, outstation at Heinz Maier-Leibnitz Zentrum, Forschungszentrum Jülich GmbH,85747 Garching, Germany

[6] Department of Chemistry, University of California, Irvine, Irvine, California 92697, United States

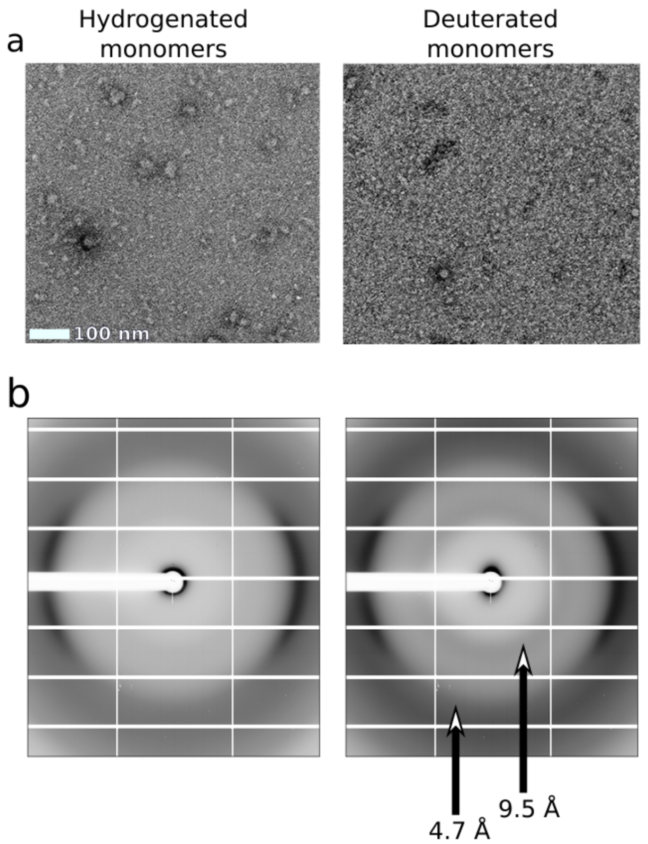


Figure S1. Characterization of α-synuclein monomers. (a) Negative-staining electron micrographs of hydrogenated (left) and deuterated (right) monomers. (b) X-ray diffraction patterns of hydrogenated (left) and deuterated (right) monomers. The position of the rings usually observed for the cross-β structures in amyloid fibers at 4.7 and 9.5 Å, but absent here, are highlighted by arrows). The two antipodal arcs arise from the Mylar capillary carrying the sample.

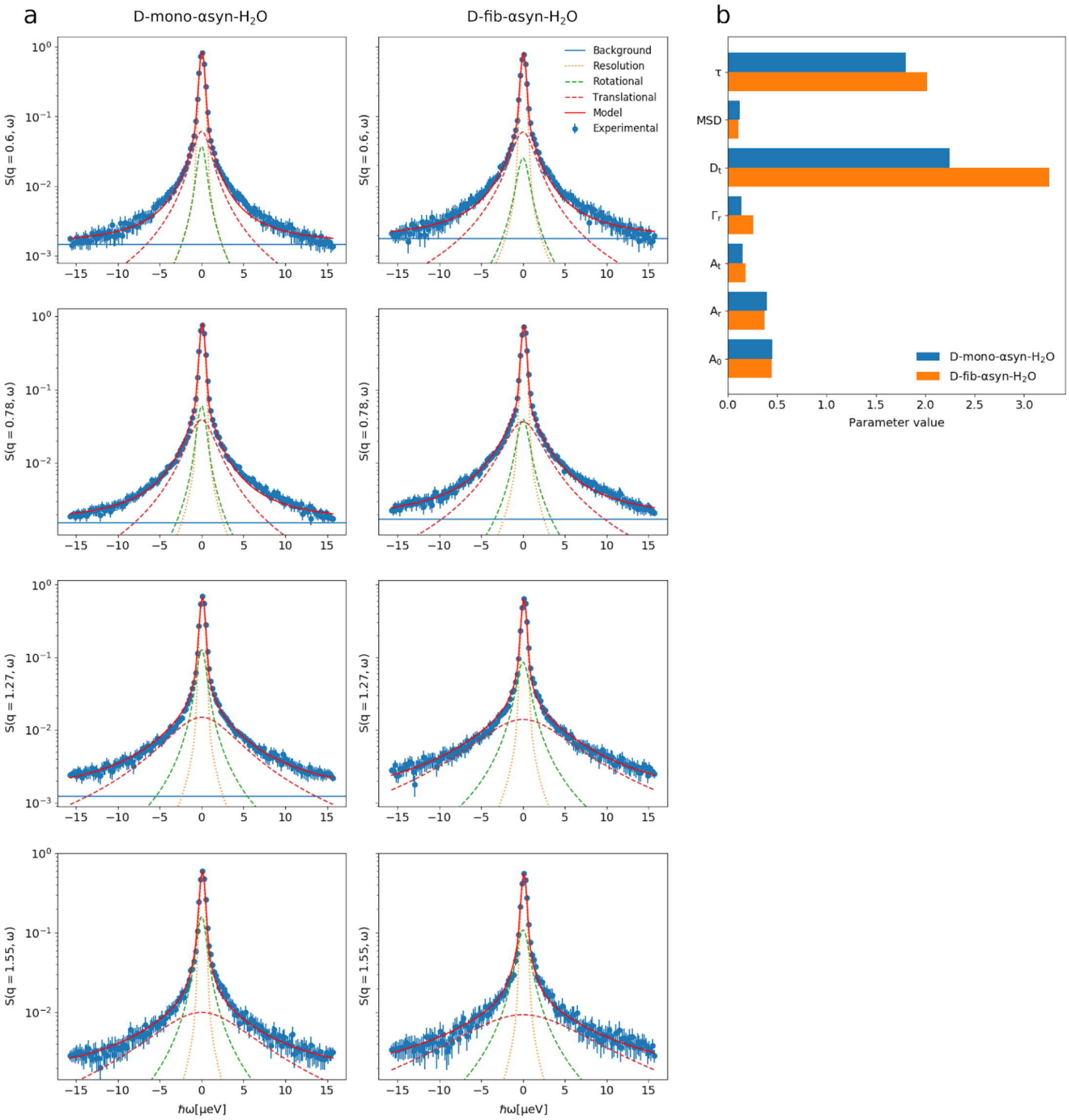


Figure S2. α-synuclein hydration water shows higher average dynamics around fibrils than around monomers. (a) The QENS spectra for D-mono-αsyn-H_2O and D-fib-αsyn-H_2O were fitted according to eq. 1 (Methods) in the momentum transfer range $0.6 < q < 1.8$ Å^{-1} . The experimental data are represented by the blue circles with error bars. The resolution function is represented by the orange dotted line, the Lorentzian for rotations by the green dashed line, the Lorentzian for translation by the red dashed line, the background by the blue solid line, and the resulting fit by the red solid line. (b) The fitted parameters from eq. 1 (where A_0, A_r and A_t are divided by $(A_0+A_r+A_t)$) are represented as a bar chart for the D-mono-αsyn-H_2O in blue and the D-fib-αsyn-H_2O in orange.

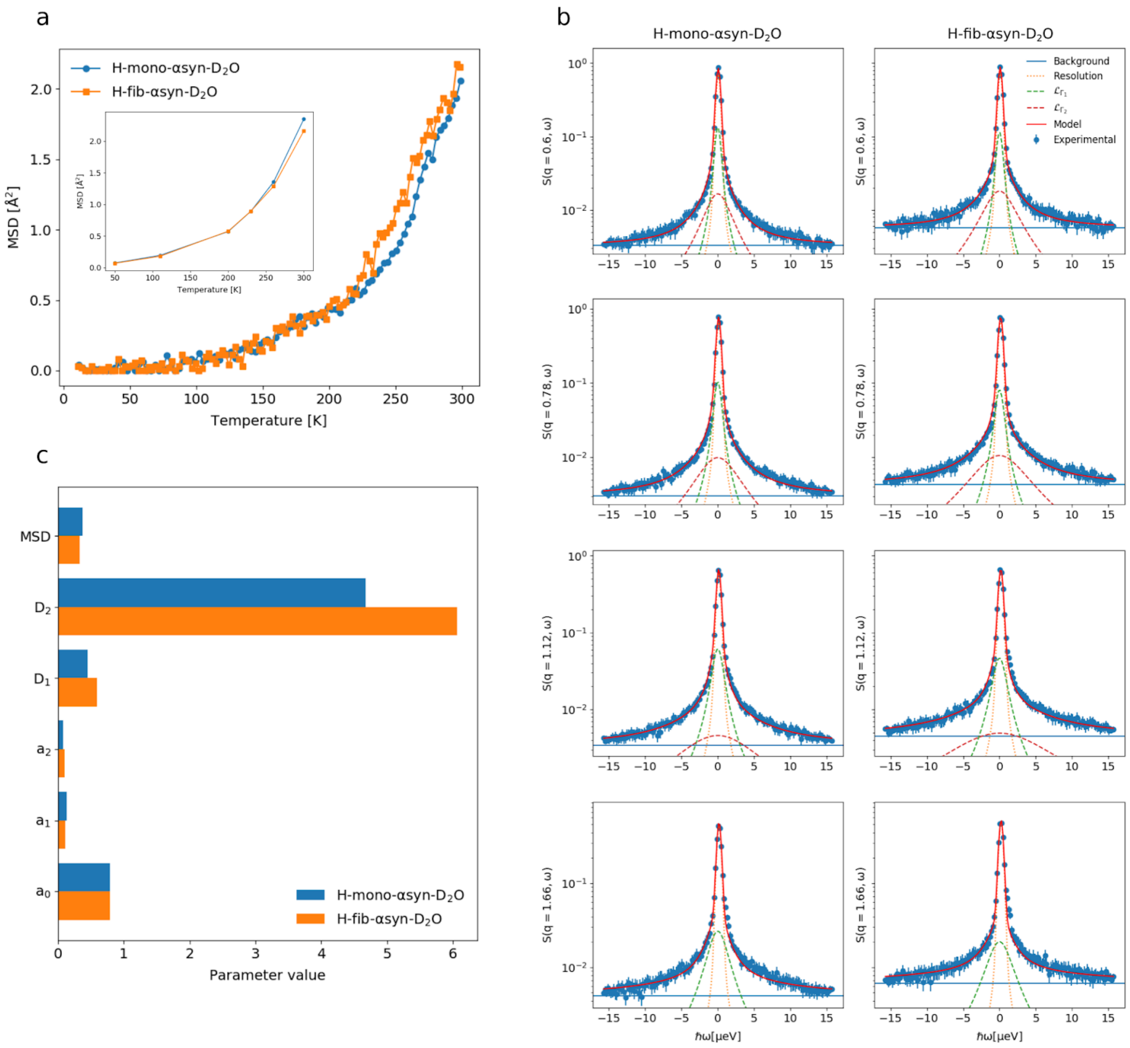


Figure S3. α-synuclein internal dynamics is slightly increased in fibrils compared to monomers. (a) Protein MSDs - obtained from a Gaussian fit in the momentum transfer range $0.22 < q < 0.96$ Å^{-1} - of H-mono-αsyn-D_2O (blue circles) and H-fib-αsyn-D_2O (orange squares). The inset represents the MSDs computed from simulations according to eq. 3 (Methods). (b) The QENS spectra for H-mono-αsyn-D_2O and H-fib-αsyn-D_2O were fitted according to eq. 2 (Methods) in the momentum transfer range $0.6 < q < 1.8$ Å^{-1}. The experimental data are represented by the blue circles with error bars. The resolution function is represented by the orange dotted line, the two Lorentzians by the green and red dashed line, the background by the blue solid line, and the resulting fit by the red solid line. (c) The fitted parameters from eq. 2 (where a_0, a_1 and a_2 are divided by $(a_0+a_1+a_2)$) are represented as a bar chart for H-mono-αsyn-D_2O in blue and H-fib-αsyn-D_2O in orange.

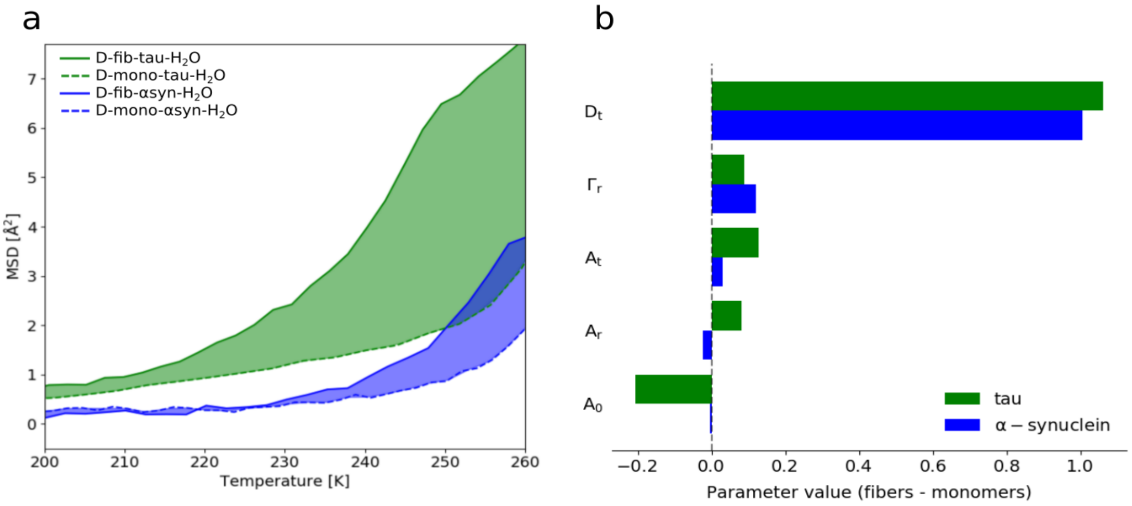


Figure S4. The hydration water around tau presents an increase in MSDs upon fibrillation that is twice as important as for the α-synuclein, explained by a larger fraction of water being converted from the immobile to the mobile state. (a) The MSDs for D-mono-αsyn-H_2O are represented by a blue dashed line and the solid blue line represents D-fib-αsyn-H_2O. The MSDs for tau were obtained by reprocessing data recorded in a previous study[16,45] and are represented by a green dashed line for D-mono-tau-H_2O and the solid green line for D-fib-tau-H_2O. (b) The parameters A_0, A_r, A_t and Γ_t from eq. 1 (Methods) were extracted from the analysis of the QENS spectra of D-mono-αsyn-H_2O, D-fib-αsyn-H_2O, D-mono-tau-H_2O and D-fib-tau-H_2O. The parameters A_0 (immobile), A_r (rotational), and A_t (translational) were rescaled by the sum (A_0 + A_r + A_t) for each sample. Subsequently each parameter value for monomers was subtracted from the value for fibers. The result is represented as a vertical bar chart with tau in green and α-synuclein in blue.

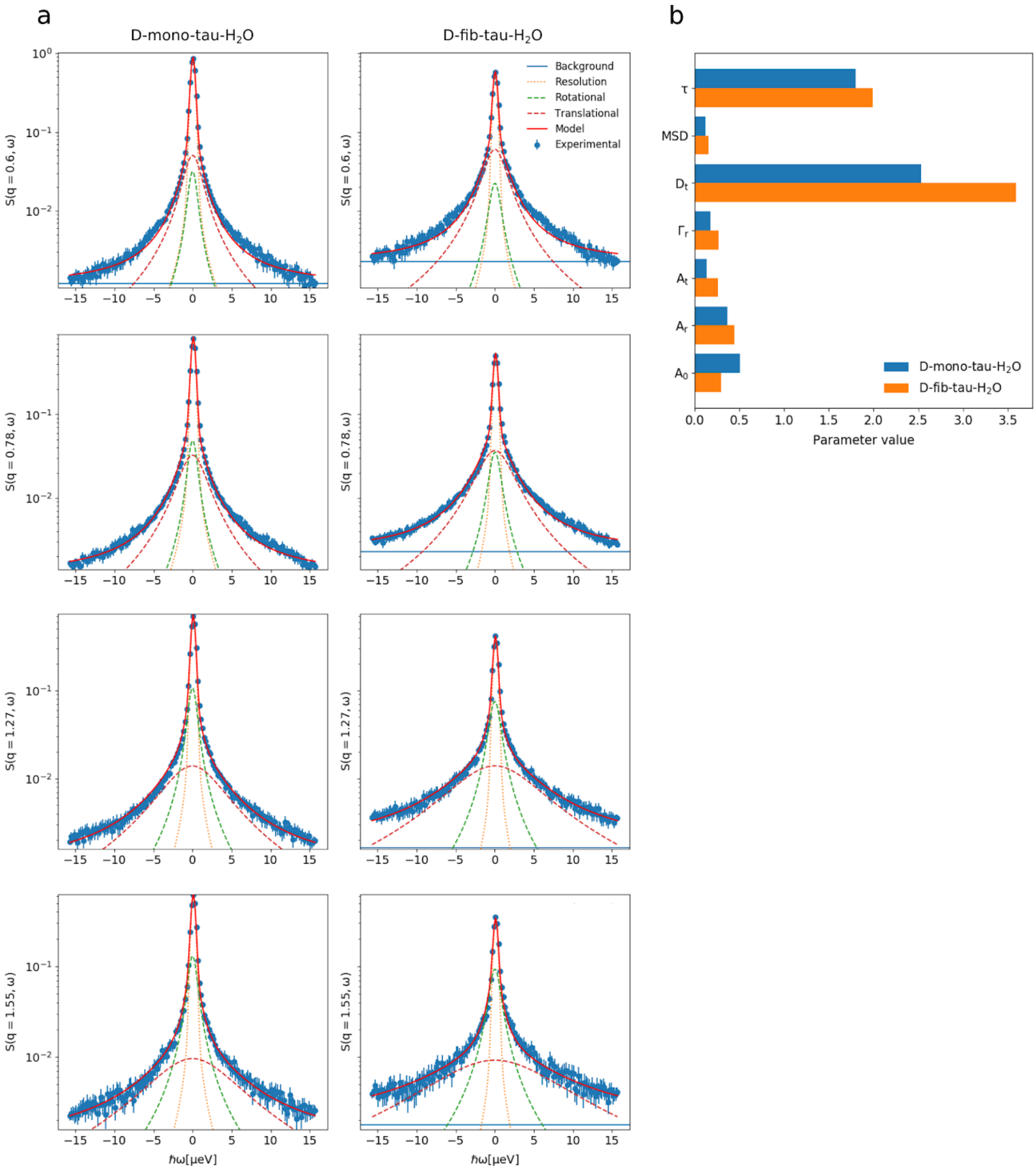


Figure S5. α-synuclein hydration water shows higher average dynamics around tau fibrils than around tau monomers. (a) The QENS data at 260 Kfor D-mono-tau-H_2O and D-fib-tau-H_2O were fitted using eq. 1 (Methods) in the momentum transfer range 0.6 < q < 1.8 Å^{-1} . The experimental data, are represented by blue circles with error bars. The resolution function is represented by the orange dotted line, the Lorentzian for rotations by the green dashed line, the Lorentzian for translation by the red dashed line, the background by the blue solid line, and the resulting model by the red solid line. (b) The fitted parameters from eq. 1 (where A_0, A_r and A_t are divided by $(A_0+A_r+A_t)$) are represented as a bar chart for the D-mono-tau-H_2O in blue and the D-fib-tau-H_2O in orange.

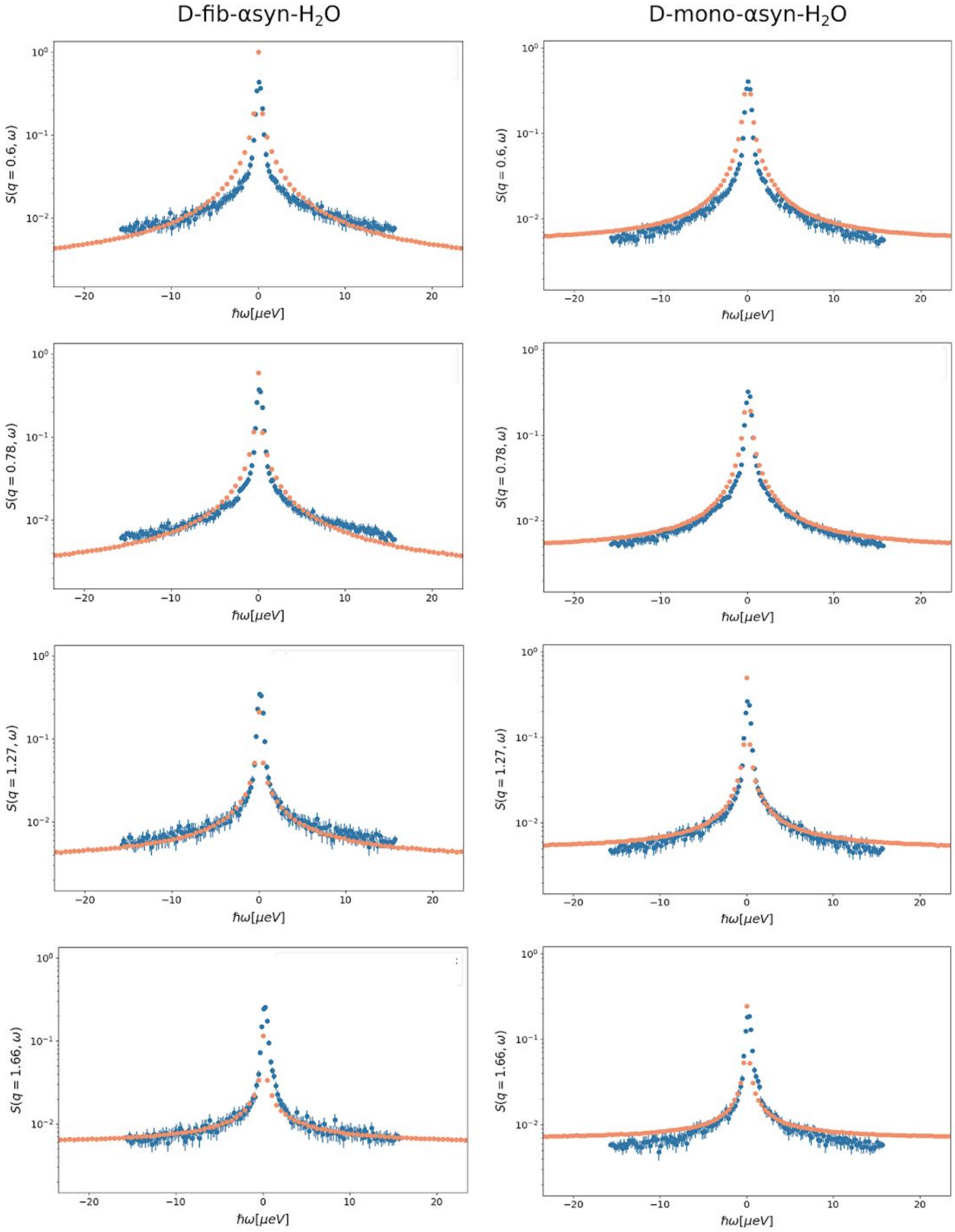


Figure S6. Experimental and simulated QENS data of α-synuclein water. Data from simulation at 260 K were processed using the NAMDAnalyzer - backscatteringDataConvert module to compute the QENS spectra for α-synuclein water (eq. 4 in Methods) and adding the background fitted from experimental data with eq. 1 (Methods). Here, the simulation results (orange circles) are directly compared with the experimental QENS data (blue circles) for different momentum transfer (*q*) values. The simulations qualitatively reproduce the *q*-dependent broadening evident in the experimental QENS data.

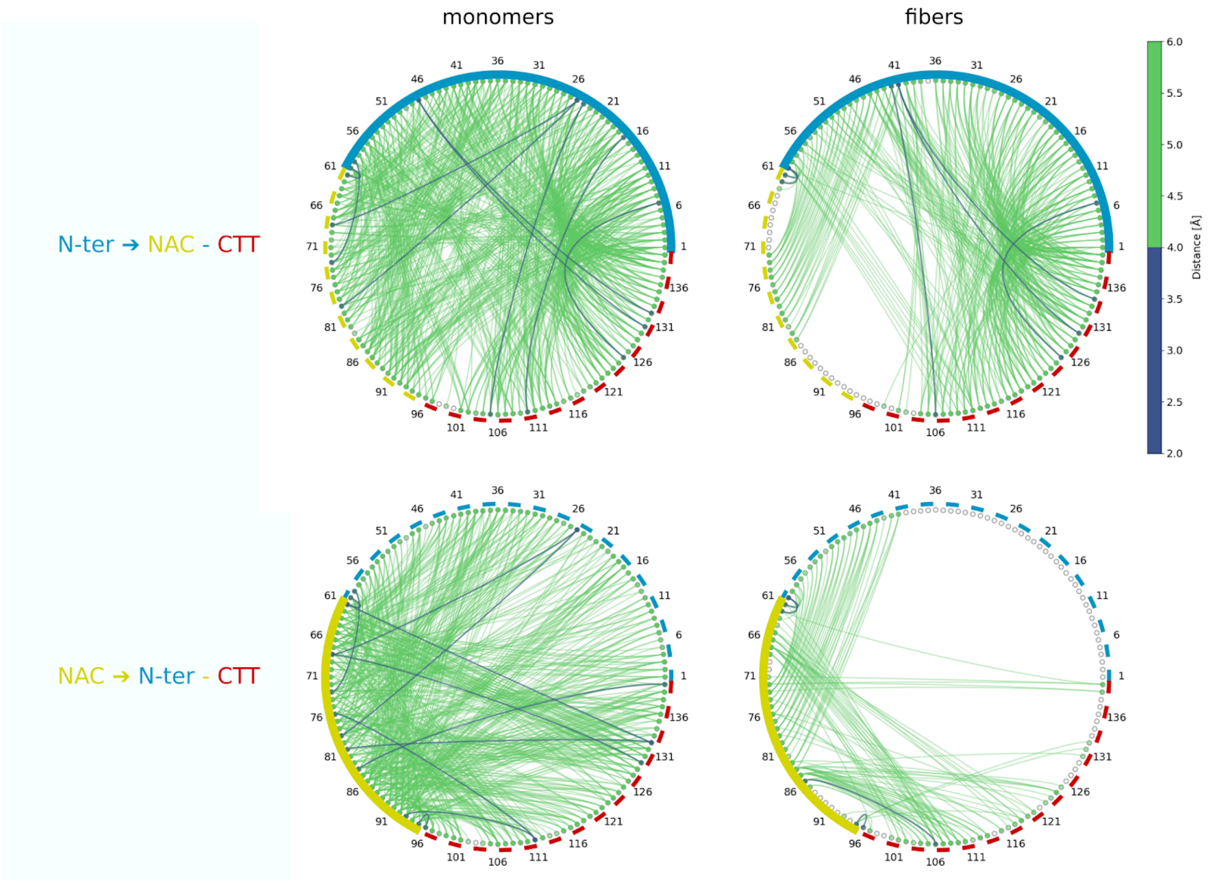


Figure S7. The α-synuclein structure changes from a compact to a extended state upon the formation of fibers. Simulation trajectories were analyzed using the NAMDAnalyzer - plotAveragedDistances_chordDiagram routine to compute the pair distances between Cα atoms of N-ter residues (upper row) or NAC residues (lower row) and all other Cα. The pair distances are averaged over all frames in α-synuclein monomers (left) and fibers (right). For all pair distances, the two residues they belong to were identified, and linked on a chord diagram with a bin for distances between 2 and 4 Å (blue Bezier curves) and one for distances between 4 and 6 Å (green Bezier curves). The N-ter region is represented with the blue arc, the NAC with the yellow one, and the CTT with the red one. A solid line is used for the region in which the Cα are selected to compute the pair distances with all other Cα in the regions represented as a dashed line. For instance, we see in the upper row that the Cα of residue 25 in the N-ter region is at a distance between 2 and 4 Å from the Cα of residues 68 and 79 in the NAC region for monomers (left). These distances are lost in fibers (right), indicating that the residues are now at least at a distance of 6 Å or more.

	D-X-αsyn-H_2O h = 0.4 MW = 15226 Da $H_{226}D_{784}C_{627}N_{166}O_{215}$	H-X-αsyn-D_2O h = 0.44 MW = 14668 Da $H_{784}D_{226}C_{627}N_{166}O_{215}$	D-mono-tau-H_2O (*) h = 0.38 MW = 48269 Da $H_{757}D_{2438}C_{1959}N_{581}O_{668}$	D-fib-tau-H_2O (#) h = 0.4 MW = 48269 Da $H_{757}D_{2438}C_{1959}N_{581}O_{668}$
exchangeable H	226	226	757	757
water / monomer	338	323	1019	1072
σ^{water}_{coh} [barns]	2621 (19%)	4975 (36%)	7892 (18%)	8308 (19%)
σ^{water}_{inc} [barns]	54313 (73%)	1323 (2%)	163573 (72%)	172183 (72%)
$\sigma^{protein}_{coh}$ [barns]	11011 (81%)	8871 (64%)	35096 (82%)	35096 (81%)
$\sigma^{protein}_{inc}$ [barns]	19829 (27%)	63470 (98%)	66045 (28%)	66045 (28%)
f_{coh}	16%	18%	16%	15%
f_{inc}	84%	82%	84%	85%

Table S1. Contributions of water and protein atoms to the coherent and incoherent signals. For all samples (X stands for 'mono' or 'fib'), the number of each atom type (H, D, C, N, O), the mass of the protein (MW), the number of exchangeable hydrogens (exchangeable H), the number of water molecules per protein monomer (water / monomer), the coherent neutron scattering cross-section for water and protein (σ^{water}_{coh} and $\sigma^{protein}_{coh}$, respectively), the incoherent cross-section for water and protein (σ^{water}_{coh} and $\sigma^{protein}_{coh}$, respectively), and the relative fractions of coherent and incoherent scattering (f_{coh} and f_{inc}, respectively) are computed as described in Methods, using the provided Python script. The percentages that are given alongside the cross-sections represent the relative contribution to the coherent or incoherent neutron scattering signal. h corresponds to the hydration level expressed in g water (H_2O or D_2O) per g protein (hydrogenated or perdeuterated). (*) The sample was prepared in 2012 for a previous study[45]. (#) The sample was prepared in 2015 for a previous study[16].

3.1.1 A possible outlook for simulations

Our work on α-synuclein showed that hydration water mobility is increased around fibrils as compared to the monomers, as a result of the decreased interaction of water with the central and C-terminal parts of the protein in fibrils. As we showed that a fraction of water molecules goes from an apparent immobile state to a mobile state, we can suppose that hydration water is becoming more bulk-like upon fibrillation and gains new degrees of freedom. We would like to link our work with a recent one that makes use of simulations to study the angular distribution of water molecules relative to a definite surface as a function of the surface type and distance to it [Shin and Willard, 2018b] [Shin and Willard, 2018a]. In this work by Shin and Willard, it appears that water presents strong orientation preferences close to a hydrophobic surface. When the distance of water from the protein surface increases, they observed that the water molecules are less oriented and become more bulk-like. This change in orientation preference can be viewed as a transition from a highly anisotropic environment at the protein surface to an isotropic environment in the bulk water. Therefore, the increased entropy upon α-synuclein fibrillation can be interpreted as a globally decreased anisotropy. In the following, we develop some considerations on the link between entropy and anisotropy and we present an initial idea for a way to measure anisotropy during simulations. Due to time limitations, this idea was not pushed further yet. In particular we would need to check its validity, and how it could affect, or bias, the thermodynamic ensemble. Also, the idea would need to be integrated in a MD simulation package for initial testing and optimisation. A link between entropy and anisotropy can be drawn and qualitatively stated as:

$$\text{entropy} \propto \frac{1}{\text{anisotropy}} \tag{3.67}$$

If we think of the transition from an anisotropic environment to an isotropic one as resulting in a gain of new degrees of freedom, the link with entropy can be formally stated using statistical mechanics. Indeed, following standard methods from statistical mechanics [Tong, 2020], for a number N of indistinguishable free particles with the associated Hamiltonian $\mathcal{H} = \frac{\mathbf{p}^2}{2m}$, the partition function for $\mathrm{D^F}$ degrees of freedom is given by:

$$\mathcal{Z} = \frac{1}{(2\pi\hbar)^{\mathrm{D^F}}} \int \prod_{i}^{N} d^{\mathrm{D^F}}\mathbf{x}_i \; d^{\mathrm{D^F}}\mathbf{p}_i \; \frac{1}{\mathrm{N!}} \exp^{-\beta \frac{\mathbf{p}_i}{2m}} \tag{3.68}$$

where $\beta = \frac{1}{k_B \mathrm{T}}$. This results in:

$$\mathcal{Z} = \frac{V^N}{N!\lambda^{\mathrm{D^F N}}} \tag{3.69}$$

By inserting this partition function in the formula for the entropy $\mathrm{S} = \frac{\partial}{\partial \mathrm{T}} \left(k_B \mathrm{T} \log \mathcal{Z}\right)$, we obtain:

$$\mathrm{S} = k_B \mathrm{N} \left[\log \frac{\mathrm{V}}{\mathrm{N}\lambda^{\mathrm{D^F}}} + \frac{\mathrm{D^F}}{2} + 1 \right] \tag{3.70}$$

where V is the volume of integration and $\lambda = \sqrt{\frac{2\beta\pi\hbar^2}{m}}$ is the De Broglie wavelength. We can readily see on eq. 3.70 that for $V = l^{\mathrm{D^F}}$ with l the length of the hypercube enclosing the volume V, and $a = (\frac{l}{\lambda})^{\mathrm{D^F}}$, we have $\frac{\partial}{\partial D} \log\left(a^{\mathrm{D^F}}/N\right) = \frac{\mathrm{D^F}}{a}$ and the entropy increases with the number of degrees of freedom $\mathrm{D^F}$.

Another way to see how anisotropy is linked to entropy is to introduce a potential term $U(r)$ in the Hamiltonian. We make use of the Mayer f function given by $f(r) = \exp^{-\beta U(r)} - 1$, and the partition function reads:

$$Z = Z_{\text{ideal}} \left(1 + \frac{N}{2V} \int dr f(r)\right)^N \tag{3.71}$$

where Z_{ideal} refers to the partition function corresponding to the free particle (eq. 3.69). The entropy is obtained as before:

$$S = S_{ideal} + N k_B \log\left(1 + \frac{N}{2V} \int dr f(r)\right) + N k_B T \frac{\partial}{\partial T} \log\left(1 + \frac{N}{2V} \int dr f(r)\right) \tag{3.72}$$

We use a hard-sphere potential applied in one dimension, which would correspond to a particle for which the distance to the plane is restricted by the potential:

$$U(r) = \begin{cases} \infty & \text{if } r \leq r_0 \\ -U_0 (r_0/r)^6 & \text{if } r > r_0 \end{cases} \tag{3.73}$$

In this case, the particle is free to move in the two dimensions parallel to the plane, but not along the normal to the plane, thereby resulting in a strong anisotropy in the potential and a degree of freedom that is lost. We consider $r^6 = (\mathbf{r}\cdot\mathbf{n})^6 = (\|r\|\|n\|\cos(\theta))^6$ in three dimensions with $\mathbf{n}$ being the unit vector normal the the plane. The integration over ϕ and θ in spherical coordinates simply yields a factor of

$\frac{4\pi}{5}$, which is neglected below. We thus obtain $\int dr f(r) = r_0 \left(\frac{\beta U_0}{5} - 1\right)$. Hence, using $C = 1 + \frac{N}{2V} r_0 \left(\frac{\beta U_0}{5} - 1\right)$, the entropy reads:

$$S = S_{\text{ideal}} + N k_B \log C - \frac{N^2 r_0 U_0}{2VTC} \tag{3.74}$$

For a high number of particles, that is $N \simeq 10^{23}$ for one mole, the third term with N^2 completely dominates the second and implies a strong reduction of the entropy as expected.
The entropy cannot be directly computed during the simulation. However, the potential energy is computed very often. We propose to compute the anisotropy of the potential energy $U^{(k)}(r)$ of the k^{th} atom using the tensor:

$$T_{ij}^{(k)} = \frac{\partial}{\partial i} \frac{\partial}{\partial j} U^{(k)}(r) \tag{3.75}$$

Taking only the principal axes, we have $A^{(k)} = \frac{1}{3} T_{ij}^{(k)} \delta_{ij}$. This quantity can be used with a gradient descent method, where a force $F = -\frac{\partial A^{(k)}}{\partial \mathbf{r}^{(k)}}$ (where $\mathbf{r}^{(k)}$ is the position of the k^{th} atom) would be applied frequently during the simulation to drive the system toward a lower anisotropy, and thus to a maximum entropy state.

3.2 The internal dynamics of γS-crystallin is decreased in the G18V mutant

Some proteins can be highly stable, even at very high concentration. That is the case of tau, lysozyme [Hankiewicz and Swierczek, 1974] and protein of the crystallin family [Vendra et al., 2016]. Yet, the three of them are found to be able to form amyloid fibrils *in-vivo* and are related to neuropathic - for tau - and non-neuropathic diseases. In particular, crystallin proteins are involved in cataract, for which several genes in the CRYG cluster were found to be often associated with the disease [Héon et al., 1999]. Similarly to other amyloid disorders, various types of aggregates can be found, resulting in different types of cataract. How such a stable system can undergo fibrillation remains unclear.

Protein of the crystallin family are divided in two broad classes, α and $\beta\gamma$, found in very high concentration in the eye lens, usually several hundreds of milligrams per milliliters. They provide the transparency and refractive index of the lens. Among the different crystallins, the γ-crystallins are attracting particular attention. They are small - 21 kDa - proteins, with at least six variants, each expressed at particular stage during life [Vendra et al., 2016]. Structurally, they are characterized by a Greek key motif, that results in a compact, globular shape. The γ-crystallins seem to be weakly interacting with each other or with the solvent. Their equilibrium as a monodisperse solution is highly sensitive to changes in concentration and the type of short-range interactions that occur between proteins. Indeed, α and γB-crystallins show a very different short-time self-diffusion dependence on the volume fraction [Bucciarelli et al., 2016]. The former behave like hard spheres, while the latter are better described by anisotropic, patchy attractions between proteins that make the short-time self-diffusion to decrease faster with the volume fraction and lead to a glass-like state and possibly to aggregation. In addition, a liquid-liquid phase separation can occur and lead to cataract [Cinar et al., 2019]. The transition to a homogeneous solution to phase separated droplets of γ-crystallin is very sensitive on temperature, pressure and on the presence of solutes.

Interestingly, the nature of the protein surface can strongly influence the liquid-liquid phase separation [Cinar et al., 2019], the aggregation kinetics, and even determine the type of aggregates that are formed [Roskamp et al., 2017]. Indeed, wild-type γD-crystallin shows an increased turbidity after 2 days incubation at pH 7 and pH 2 and 55°C [Wu et al., 2014]. However, the amyloid signal from the ThT fluorescence is observed only for species formed at pH 2.0. A similar observation can be made with the cataract associated P23T γD-crystallin mutant [Boatz et al., 2017], where molecules in aggregates formed at physiological pH retain their native conformation.

Hence, the charges present on the protein surface can affect the aggregation pathway. In addition, it was also demonstrated [Huang et al., 2016] that wild-type γS-crystallin is able to retain a robust hydration shell, with local water dynamics that barely differs in a protein-concentration range from 5 to 500 mg/mL. On the contrary, the aggregation-prone G18V mutant, which has a bigger hydrophobic surface exposed to the solvent than wild-type γS-crystallin [Khago et al., 2016],shows a much more fragile shell, as the protein dehydrates quickly with concentration change, or environment perturbations [Huang et al., 2016]. The amino acid composition of the protein can thus strongly influence the hydration shell and can affect the aggregation propensity as well as the aggregation pathway [**?**].

In the beginning of my PhD thesis, a collaboration with Rachel Martin's group (UC Irvine, USA) started, aiming at studying, where our role is to provide information on the ensemble-averaged dynamics of γS-crystallin and its hydration water. To this end, we used incoherent neutron scattering spectroscopy on hydrated protein powders, thus complementing the results on the structure and the local dynamics obtained earlier in Rachel Martin's group by NMR [Huang et al., 2016][**?**]. The incoherent neutron scattering allows to probe the ensemble-averaged ps-ns dynamics of biological samples. The method is mostly sensitive to hydrogen atoms, allowing to measure specifically protein dynamics when a hydrogenated protein is hydrated with D_2O. Moreover, the use of hydrated protein powders permits to suppress the center-of-mass diffusion and thus provides access to the protein internal dynamics. Conversely, when a deuterated protein powder is hydrated with H_2O, the dynamics of the hydration water can be studied.

We aimed at complementing the data obtained by Rachel Martin's group by information on the ensemble averaged dynamics of γS-crystallin wild-type and G18V mutant to better understand the effect of the amino acid substitution on protein internal and hydration water dynamics. Unfortunately, unforeseen shutdowns of the neutron scattering facilities at the ILL and he MLZ caused important delays in our allocated beamtime. Hence, only a subset of samples were measured thus far. More specifically, we could measure the elastic neutron scattering signal for temperatures ranging from 20 to 300 K for two hydrogenated wild-type γS-crystallin incubated at pH 7 and pH 2 for 48 hours (WT-pH7 and WT-pH2, respectively), and two G18V γS-crystallin samples incubated at pH 7 and pH 2 for 48 hours (G18V-pH7 and G18V-pH2, respectively). Only the WT-pH7 sample is free of aggregated species, the WT-pH2 and G18V-pH2 samples contain amyloid aggregates and the G18V-pH7 sample contains non-amyloid aggregates [**?**].

By analyzing dependence of the elastic signal on the momentum transfer q, we observed similar protein internal dynamics for the two wild-type samples, namely the monomeric WT-pH7 and the aggregated WT-pH2. Also, the two G18V mutant samples, the amyloid G18V-pH7 and the non-amyloid G18V-pH2, present similar dynamics. However, the wild-type presents increased dynamics compared to the G18V mutant above 250K. Our results suggest that the internal dynamics of γS-crystallin depends essentially on the amino acid composition and weakly on the aggregation state.

3.2.1 Results and Discussion

The dependence of the elastic neutron scattering (ENS) signal on the momentum transfer q provides access to the mean-squared displacement (MSD) of the hydrogen atoms in the sample. The MSD at each temperature was extracted by fitting the ENS dependence on q using the Gaussian approximation [Yi et al., 2012] as described in Methods. The WT-pH7 and WT-pH2 samples show identical MSDs from 20 to 300 K (Figure 3.7). Similarly, the G18V-pH2 and G18V-pH7 samples do not differ significantly. However, the WT-pH2 and WT-pH7 samples present significantly higher MSDs after 250 K than the G18V-pH2 and G18V-pH7 samples. The G18V-pH2 and G18V-pH7 samples appear less hydrated than the WT-pH2 and WT-pH7 samples as the protein dynamical transition (PDT) at 230 K [Doster et al., 2010] is barely visible. It suggests that the substitution of a single amino acid is sufficient to dramatically affect the MSDs that reflects backbone and side-chain motions and hydration of the protein in the time-scale probed in this study. The interpretation of the result as arising from a decreased hydration of the protein in the G18V mutant corroborates previous results by Rachel Martin's group showing that the hydration shell of the G18V mutant is less stable than for the wild-type [Huang et al., 2016].

To confirm that the protein MSDs depends more importantly on the amino acid composition than on the protein aggregation state, we compared the MSDs of several proteins in monomeric and aggregated form, namely tau, α-synuclein, γS-crystallin wild-type and γS-crystallin G18V. The elastic neutron scattering data of tau monomer and fiber were obtained from previous work [Gallat et al., 2012b] [Fichou et al., 2015b], and the data on α-synuclein monomer and fiber were obtained from the article 1 in this thesis. By summing the elastic signal over all momentum transfer values q, it can be readily observed that the elastic signal decreases differently for each protein, but the differences between the monomeric and aggregated forms are systematically much smaller than the differences between the proteins. (Figure 3.8). All these elastic data were reprocessed using the Gaussian approximation within the same momentum transfer range $0.3 < q < 1$ Å^{-1}. The protein tau has the highest MSDs with a value of around 2.1 $\AA^2$ at 300 K for both the monomeric

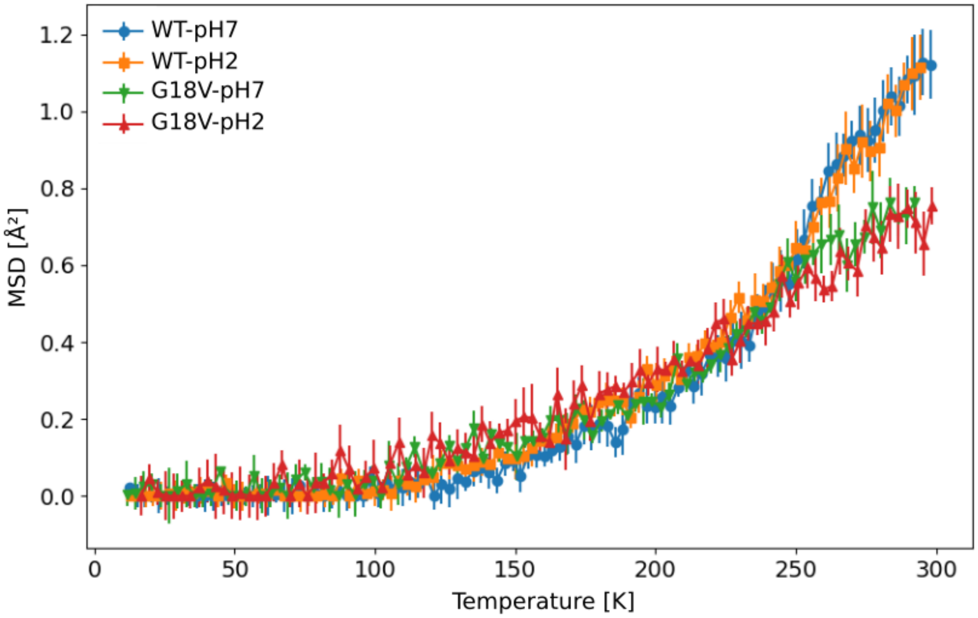


Figure 3.7: γS-crystallin dynamics depends on primary sequence
The ENS data were fitted in the Gaussian approximation using the momentum transfer range $0.35 < q < 1.12$ $\AA^{-1}$ at each temperature. The obtained MSDs are plotted against temperature for WT-pH7 (blue circles), WT-pH2 (orange squares), G18V-pH7 (green triangles) and G18V-pH2 (red triangles).

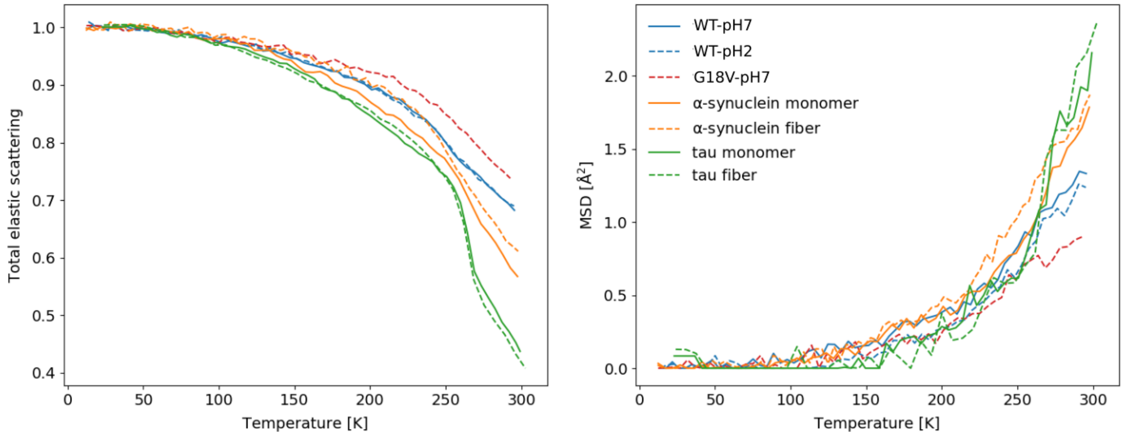


Figure 3.8: MSDs of tau fibers and monomers, α-synuclein fibers and monomers, γS-crystallin wild-type monomers and aggregates and G18V amyloid aggregates. The ENS data of tau monomer and fiber (obtained from previous work [Gallat et al., 2012b] [Fichou et al., 2015b]), α-synuclein monomer and fiber (obtained from article 1 in this thesis), WT-pH2, WT-pH7 and G18V-pH7 samples were summed over all momentum transfer q and normalized to the first 8 points starting from the lowest temperature **(left panel)**. The ENS data were fitted using the Gaussian approximation in the range $0.3 < q < 1$ Å^{-1} and the obtained MSDs were plotted against temperature **(right panel)**.

and the fibril forms. The α-synuclein shows a MSD of 1.75 Å^{-1} at 300 K for both monomer and fiber, the WT-pH2 and WT-pH7 has a MSD of around 1.25 Å^{-1} at 300 K and the G18V-pH7 sample of 0.75 Å^{-1}.

These results have an important implication, which is that the protein energy landscape depends weakly on the changes in protein aggregation or folding state - in the ps-ns time scale. However, a different result was obtained with the concanavilin A, for which the aggregation state affects the protein-internal dynamics in a significant manner [Schiró et al., 2012]. A more thorough study is needed to better understand how the amino acid composition affects the protein dynamics and the aggregation propensity. In particular, dimensionality reduction algorithms such as principal component analysis [Andersson et al., 2017] or support vector machine could be used with information on the protein charge, dynamics, hydrophobicity, disorder, aggregation propensity and any other parameter that can be computed to determine how these parameters correlate with each other.

Importantly, the effect of point mutations on the aggregation propensity and dynamics is of major importance, as the recent progress in adeno-associated virus (AAV) based genetic therapies [Kuranda and Mingozzi, 2017] bears the great promise to allow to rescue a disorder by inserting a new gene in specific cells. Hence, the insertion of well-selected single point mutations in the γS-crystallin sequence might help in significantly improving the stability of the protein *in-vivo* and prevent the cataract or other related diseases.

3.2.2 Methods

Protein production and purification. All the proteins used in this work were provided by our collaborator Rachel Martin from the University of California, Irvine (see [Khago et al., 2016]). The proteins were received already purified and lyophilized, ready to use for the neutron experiment.

Neutron scattering experiment and data analysis. The ENS data on tau monomers were obtained from a previous study [Gallat et al., 2012b] as well as the data on tau fibers [Fichou et al., 2015b]. The ENS data on α-synuclein are the same as for the article 1 in this thesis. The data for WT-pH2, WT-pH7, G18V-pH2, and G18V-pH7 samples were recorded on the SPHERES backscattering spectrometer [Wuttke et al., 2012] operated by JCNS at the Heinz Meier-Leibnitz Zentrum (MLZ) in Garching, Germany (proposal number 13777, from the 18^{th} to the 24^{th} September 2018). A mechanical Doppler drive was used to set the incident energy, and a phase space transformation (PST) chopper was employed to optimize the neutron flux at the sample position [Kirstein et al., 2000]. Freeze-dried protein powders were incubated for 24h with P_2O_5 in a desiccator to completely remove water [Dolman et al., 1997]. The dry powders were then weighed and re-hydrated by vapor-exchange with pure D_2O to reach a hydration level h=0.44 (g D_2O/g dry protein). Hydrated protein powders were then sealed in flat 4x3 cm^2 aluminum sample holders using 1 mm indium wire. Samples were weighed before and after the experiment to assure the hydration remained constant throughout the neutron experiment. The aluminum cells were mounted in a cryofurnace at room temperature and the temperature was subsequently lowered to 10 K in approximately 2 hours. For each sample, the elastic peak - Doppler drive stopped - was recorded during a 24 h temperature ramp from 10 to 300 K (0.2 K/min). Between each data point, the temperature was increased by 0.05 K. The raw data were pre-processed using the SLAW software available on facility computers (http://apps.jcns.fz-juelich.de/slaw). All the data on tau and α-synuclein were reprocessed and the determination of the mean-squared displacements was performed using of standard a Gaussian model [Yi et al., 2012] fitted with the *curve_fit* routine from Python Scipy library.

3.3 Zinc affects insulin dynamics and aggregation

Insulin is a small proteic hormone involved generally in anabolism, and specifically in glucose homeostasis in the organism [Sonksen and Sonksen, 2000]. The protein is produced by β cells in pancreatic islets. Initially, a prepropeptide is translated, which will be subsequently cleaved - first in endoplasmic reticulum and then in the Golgi transport system - to form mature insulin, with 51 amino acids, ready to be released in blood vessels to regulate glucose concentration.
Insulin is mainly known for its essential role to treat diabetes. Initially, the only source of insulin were animals, but this resulted in poorly purified samples. The vast majority of insulin produced today is recombinant, and expressed in yeast. Injection of insulin lead to decreased release of sugar in the blood and thus, helps the organism in maintaining a stable and healthy level of glucose in the blood. An important concern for the pharmaceutical industry is the protein stability, as it affects the storage, the delivery (methodology for the injection) and the duration of action in patients.
Moreover, amyloid deposits were found at local injection site in diabetes patients [Swift et al., 2002] [Shikama et al., 2010]. The protein was found to be able to form amyloid fibrils and super-aggregates, known as particulates - spherical particles in the μm range - and spherulites - spherical particles up to hundreds of μm in diameter [Vetri and Foderà, 2015]. It is intriguing to see that type II diabetes is often associated with neurodegenerative disorders. Especially, it seems to play a significant role in Alzheimer 's [Biessels and Kappelle, 2005] and Parkinson's [Aviles-Olmos et al., 2013] [Hogg et al., 2018] diseases. How insulin is involved is not clear yet, such as the form - native or aggregated - that can impair cellular functions and favor neurodegenerative disorders.

To better understand the insulin aggregation, and the factors influencing the final morphology - that is, either fibril, particulates or spherulites - we initiated, in collaboration with Vito Foderà at University of Copenhagen, the study of this amyloid system. To this purpose, we firstly worked on the conditions influencing the aggregation pathway.
It is important to mention that heterogeneity exists also for the insulin aggregation, even in the same tube where both fibrils and spherulites can be formed at the same time [Foderá and Donald, 2010]. In consequence, we investigated the influence of various parameters such as protein and salt concentration, pH or shaking on the aggregation process. This was further motivated by the need to isolate these different forms. It appeared that pH and salt concentration play a huge role in the process, but both fibrils and spherulites were always formed at the same time. Only

shaking allowed us to completely prevent the formation of the latter species. Concerning spherulites, their size make it easy to isolate them by simple centrifugation. Subsequent measurement of the neutron scattering signal from our samples showed an unexpected result. Indeed, the native insulin exhibits a standard mean-squared displacement (MSD), with a well defined protein dynamical transition around 220 K. However, both fibrils and spherulites show very linear MSDs, without any detectable transition, a behavior that is usually seen with dry proteins. This effect could be attributed to real dynamical suppression upon aggregation, but we did not take into account all factors that can affect the dynamics.

Indeed, insulin is intimately linked to zinc. The metal is essential in several biological aspects of the protein, including its release and stabilization [Li, 2014]. Moreover, metal ions in general are known to play major role in biology, but also in protein aggregation where the role of calcium in α-synuclein fibrillation has been investigated [Lautenschläger et al., 2018]. Hence, we produced new samples, for which zinc content and interaction was characterized by proton induced X-Ray emission (PIXE) in collaboration with Elspeth Garman in Oxford and X-Ray absorption fluorescence spectroscopy (XAFS) in collaboration with Alessandro Longo at the ESRF, respectively. Five samples were produced in total, namely the native insulin, containing zinc, a 'washed' insulin for which the powder was simply suspended in pure D_2O, centrifuged and re-suspended several times. Also, an EDTA-treated sample was produced, for which we expect no zinc leftover, and that was washed as well to remove EDTA. And finally, the two aggregated samples, fibrils and spherulites that were washed the same way the native 'washed', and EDTA-treated samples was.

Using PIXE, we found that zinc was completely removed from EDTA-treated, and aggregated samples. However, the washed sample showed a slightly lower amount of zinc than native insulin, indicating it is loosely bound to it, and can be partly removed in pure D_2O. Moreover, the XAFS measurements showed that zinc can still interact with native insulin, even at pH 2, while this interaction is completely lost in aggregated samples. Together with the neutron backscattering result, where the protein dynamical transition is visible only for native and washed insulin, this suggests that zinc contributes to hydrate the highly hydrophobic protein by bringing water from its solvation shell with it. This effect probably contributes significantly to insulin function as well.
In addition, if removal of zinc with EDTA does not affect the secondary structure, as shown by circular dichroism, it does increase the hydration radius, as verified by dynamic light scattering, which can be interpreted as an increase of the insulin ag-

gregation propensity. Indeed, using thioflavin T (ThT) fluorescence, we could show that insulin aggregates in a two-sigmoidal way, where a first onset on fluorescence increase corresponds to fibril formation, which is barely sensitive to zinc concentration. While a second onset, highly sensitive to zinc, seems to correspond to spherulites formation. Whether spherulites are formed from fibrils directly or by another pathway is not clear yet, but zinc might play an important role in preventing it, either by facilitating protein solvation or by screening electrostatic interactions that make insulin to aggregate into spherulites. These results constitute the following paper that was submitted to Biophysical Journal.

Article 2

Zinc determines dynamical properties and aggregation kinetics of human insulin

Zinc Determines Dynamical Properties and Aggregation Kinetics of Human Insulin

Kevin Pounot* [1,2], Geoff Grime [3], Alessandro Longo [4], Michaela Zamponi [5], Daria Noferini [5], Viviana Cristiglio [1], Tilo Seydel [1], Elspeth Garman [6], Martin Weik [2], Vito Foderà [7]*, Giorgio Schirò [2]*

1 Institut Max von Laue – Paul Langevin (ILL), CS 20156, F-38042 Grenoble, France
2 Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France
3 Ion Beam Centre, Advanced Technology Institute, University of Surrey, Guildford GU2 7XH, Surrey, UK
4 ISMN-CNR, UOS Palermo, Via Ugo La Malfa, 153, 90146 Palermo, Italy
5 Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), D-85748 Garching, Germany
6 Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK
7 Department of Pharmacy, Universitetsparken 2, University of Copenhagen, 2100 Copenhagen, Denmark
* kevin.pounot@ibs.fr, vito.fodera@sund.ku.dk, giorgio.schiro@ibs.fr

Abstract

Protein aggregation is a widespread process leading to deleterious consequences in the organism, with amyloid aggregates being important not only in biology but also for drug design and biomaterial production. Insulin is a protein largely used in diabetes treatment and its amyloid aggregation is at the basis of the so-called insulin-derived amyloidosis. Here we uncover the major role of zinc in both insulin dynamics and aggregation kinetics at low pH, where the formation of different amyloid superstructures (fibrils and spherulites) can be thermally induced. Amyloid aggregation is accompanied by zinc release and the suppression of water-sustained insulin dynamics, as shown by particle-induced X-ray emission and X-ray absorption spectroscopy and by neutron spectroscopy, respectively. Our study shows that zinc binding stabilizes the native form of insulin by facilitating hydration of this hydrophobic protein and suggests that introducing new binding sites for zinc can improve insulin stability and tune its aggregation propensity.

Statement of Significance

Localized amyloidosis has been found near insulin injection sites for diabetes treatment, that leads to deleterious repeated inflammations known as the insulin-derived amyloidosis. Neutron scattering, particle-induced X-ray emission, X-ray absorption spectroscopy, and fluorescence spectroscopy are used here to study human insulin aggregation and reveal that zinc, coordinated to insulin in the native form, is released upon amyloid aggregation, when insulin forms different superstructures known as fibrils and spherulites. Zinc release leads to a full suppression of functionally essential internal protein dynamics through a profound modification of the protein's hydration properties. Furthermore, zinc removal completely modifies insulin amyloid kinetics. The results suggest that changes in protein hydration upon zinc binding/release contribute to modify both stability and dynamics of insulin. Changes in metal-modulated protein hydration might then be a general strategy to control protein stability and tune protein aggregation into amorphous and ordered superstructures. Protein superstructures are also promising candidates in the field of biomaterials: our findings on the control of kinetics and stability by means of zinc addition/removal helps defining chemical conditions for tailoring insulin superstructure properties.

Introduction

In vitro under destabilizing conditions, a large number of proteins aggregate and form specific ordered superstructures (1). These can be amyloid fibrils, characterized by a typical cross-beta pattern (2, 3) or larger, micrometric assemblies of amyloid nature, like spherulites (4, 5). The variety of amyloid species poses a challenge for pharmaceutical drug development, where protein-based products have to be optimized while controlling and characterizing any by-product particles. Indeed, due to the simultaneous occurrence of different species, identifying and isolating each individual species is a highly demanding task, still remaining a *conditio sine qua non* for the quality control of the final product (6). On the other hand, having access to a number of markedly different protein self-assembled species offers a unique opportunity for the development of a rational platform for the design of bio-inspired materials for applications ranging from bio-sensing and tissue engineering to drug delivery (7). Finally yet importantly, a large-scale morphological variability of aggregates is also observed in the context of amyloid-related pathologies. This variability potentially determines different etiological subtypes of diseases (8–10), making it even more challenging to univocally establish a connection between aggregate structure/morphology and the onset/progression of the diseases. In all the above fields, it is then crucial to disentangle the concurring aggregation pathways and isolate the different species. Indeed, identifying the different aggregates would allow the mechanisms of formation of the different species to be mapped, as well as unraveling how the structure-function-dynamics relationship is perturbed by amyloid formation for each of the different aggregation states.

Insulin, a 5.7 kDa hormone consisting of two mainly alpha-helical chains (A and B) linked by disulfide bonds, can form amyloid fibrils (11, 12) and spherulites (13). The two types of aggregate coexist at low pH when human insulin samples are thermally destabilized (14, 15). Subcutaneous insulin injections are largely used during insulin therapy for diabetic patients. Near injection sites, localized amyloidosis has been found that can lead to deleterious repeated inflammations, known as the insulin-derived amyloidosis (IDA) (16–19) or also "insulin ball" (20) or amyloidoma (21). The number of IDA reported cases has increased significantly in the last years (22). When IDA occurs, it leads to poor glycemic control and also to catheter occlusion in the case of continuous infusion. Recent *in vitro* studies on the stability of insulin formulations used to treat type 1 diabetes have shown that amyloid formation occurs either at neutral pH or below the insulin isoelectric point ($pH < 5.2$) (23). An alternative way recently explored for insulin delivery to diabetic patients is the administration by the oral route, that could potentially overcome the drawbacks of insulin injection reducing the number of injections needed and the risk of side effects (24). However, when administered orally, insulin first arrives at the stomach, where pH is between 1.2 and 3.0, which is the range where, *in vitro*, insulin forms both amyloid fibrils (11) and spherulites (13). It is also known that diabetes is commonly associated with neurodegenerative diseases (25) and insulin signaling impairment - for which insulin aggregation might play a role - was shown to promote neurodegeneration (26, 27). Furthermore, insulin stability is a major concern in the pharmaceutical industry for production and storage (8–10). As a consequence, understanding the key factors involved in insulin self-assembly is highly desired for the optimization of downstream product processing, as well as for a more exhaustive mapping of the multiplicity of different structures occurring in its aggregation reaction. The insulin aggregation pathway starts from a mixture of oligomeric forms. The oligomerization state depends on the solution conditions, with mainly a dimer-monomer mixture occurring at low pH (28) and a hexamer species at higher pH. However, insulin amyloid species can be formed independent of the details of the early oligomerization process (29) and such aggregates are Thioflavin T (ThT)-positive, indicating the presence of a cross-beta fibrillar structure as is found for toxic amyloid proteins such as tau or α-synuclein (2, 3). Several factors are reported to affect the aggregation process of insulin. Approaches involving sedimentation kinetics measurements (30), optical microscopy imaging (14), and small angle X-ray scattering (31) have revealed the role of both insulin and salt concentration and of pH in the aggregation process. In the context of amyloid diseases, the pathological impact of metal ions in the alteration of amyloidogenesis is now accepted (32). In the specific case of insulin, zinc ions (Zn^{2+}) are essential components for protein expression and activity (33). It has been shown that Zn^{2+} slows down fibrillation of monomeric insulin at physiological pH: this result has been used to propose that Zn^{2+} co-secreted from pancreatic β-cells protects the organism from the formation of non-native insulin oligomers and aggregates (34). However, Zn^{2+} has been found to be coordinated in both functional and aggregated bovine insulin at physiological pH (35).

More generally, the significant influence of metal ions on aggregation kinetics, pathway, and aggregate morphology has been reported for several amyloid systems (36–38). As a matter of example, zinc and copper ions increase the stability of Aβ oligomers and reduce the stability of Aβ fibrils, the effect on fibrils being specific to the type of ion (39). Ions can also dramatically affect the balance between the numbers of native multimers and monomers (40), which are prone to form amyloid-like fibrils in systems containing GAPR-1 and heparin (41). Notwithstanding pronounced metal ion-related effects are clearly detected in other amyloid-forming systems (37, 42–46), the mechanism at play remains elusive. Metal coordination at the level of single protein molecules certainly changes the overall charge, protein structure and/or hydration, eventually influencing the dynamics of the single protein molecule (47, 48). How these changes modify the physico-chemical properties of protein aggregates, however, remains unclear.

Recently, several pieces of evidence have been reported on the relationship between amyloid aggregation and the dynamical properties of aggregating proteins. Measurements of intramolecular reconfiguration dynamics in different pathological amyloid proteins suggested a direct correlation between internal dynamics and aggregation propensity and kinetics (49, 50). Neutron scattering results obtained on different proteins showed that upon amyloid aggregation either the protein (51) or its hydration water (52) can show an increased mobility at temperatures above the so-called dynamical transition (53).

Here, we used neutron scattering to characterize the dynamics of human insulin in its native state and in both its fibril and spherulite amyloid forms. In order to characterize the internal dynamics of both fibrils and spherulites, we first screened and selected the conditions for isolating samples containing either predominantly fibrils or predominantly spherulites. Neutron scattering revealed a complete suppression of functionally relevant internal protein dynamics in both insulin fibrils and spherulites. An analysis based on microbeam Particle Induced X-ray Emission (μPIXE) and X-ray absorption spectroscopy (XAS), combined with a protein conformational characterization, unexpectedly revealed that the protein rigidification is due to the release, upon amyloid aggregation, of the Zn^{2+} coordinated to insulin in the native state (33). A crucial role of Zn^{2+} in determining insulin aggregation kinetics was then unveiled by fluorescence spectroscopy.

Materials and Methods

Sample preparation: native and EDTA treated insulin. Crystalline powder of human insulin was obtained from Novo Nordisk and stored at -20°C before use. In order to dissolve crystals to obtain monomeric samples, the powder was dissolved by adding 5 µl of HCl to a suspension in H_2O until the solution was clear. After flash-cooling in liquid nitrogen and freeze-drying, electron microscopy and X-ray powder diffraction experiments were used to check for the presence of any crystals – protein or salt – and for aggregates. For the washed and EDTA treated samples, the freeze-dried solutions were dissolved with and without EDTA, respectively. The pH was then set to the insulin isoelectric point (pI = 5.2) to perform washing steps using several cycles of centrifugation at 18000 rpm for 20 minutes and resuspension in pure D_2O. Finally, flash-cooling and freeze-drying were performed prior to D_2O hydration for the neutron experiments.

Sample preparation: insulin fibrils and spherulites. The conditions for preparing pure fibril and pure spherulite samples are described in the main text, as they were obtained and optimized in the present work. Prior to neutron experiments, fibrils and spherulites were washed using several cycles of centrifugation at 18000 rpm for 20 minutes and resuspension in pure D_2O to remove any salt and contaminants. They were then flash-cooled and freeze-dried as for the monomeric samples.

Incoherent elastic neutron scattering. All data were recorded on the SPHERES backscattering spectrometer (proposal P14013, 14-20th June, 2018) (54, 55) operated by JCNS at the Heinz Maier-Leibnitz Zentrum (MLZ) in Garching, Germany . Freeze-dried insulin powders were incubated with P_2O_5 for 24 hours in a desiccator to further remove water. In the case of lysozyme, such a procedure resulted in a residual presence of four waters per protein molecule. The insulin powders were then weighed and re-hydrated by vapor-exchange with pure D_2O to reach a hydration level h = 0.44 D_2O/protein w/w. Hydrated protein powders were then sealed in aluminum flat neutron cells (sample thickness=0.3 mm) using 1 mm indium wire from Alfa Aesar. Samples were weighed before and after the neutron experiment to verify that they were correctly sealed and no water was lost. The cell was mounted in a cryostat at room temperature, and the temperature was subsequently lowered to 10 K within approximately 2 hrs. The elastic scattering signal was recorded at an energy resolution of 0.66 µeV (FWHM), employing Si(111) monochromator and analyzer crystals in exact backscattering geometry (for large scattering angles), corresponding to an incident wavelength of $\lambda = 6.27$ Å. The elastic signal was recorded while continuously increasing the sample temperature from 10 to 300 K at a rate of 0.2 K/min and binned together to 0.05 K steps. Raw data were pre-processed - i.e. normalized to the detector efficiency and to the incident beam intensity recorded by a so-called monitor device - using the SLAW software available on facility computers (http://apps.jcns.fz-juelich.de/slaw). With the elastic signal arising mainly from the incoherent scattering of the hydrogen atoms, the mean square displacements (MSD) of hydrogen atoms can be obtained using the well established Gaussian approximation (56, 57). We performed this analysis using custom-made python scripts (http://github.com/kpounot/nPDyn), with the Gaussian approximation verified for the momentum transfer ($q = 4\pi\lambda^{-1}\sin(\theta/2)$, with θ being the scattering angle) range 0.6-1.2 $Å^{-1}$. Using higher terms in cumulant expansion leading to a corrected Gaussian model (57) or using a gamma distribution based model (58) gave similar results but with higher numerical instability during the fitting procedure.

Neutron diffraction. Data was recorded on the D16 instrument at the ILL in Grenoble, France. The same samples within the sealed aluminum cells as used for measuring the incoherent elastic neutron scattering were mounted in an Orange cryostat to measure the data at two temperatures, 300 K and 200 K. The diffracted beam was measured over an angular range 12 - 112.5° which corresponds to a q-range 0.05 - 2.5 $Å^{-1}$. Neutron data were corrected for the empty cell scattering, the ambient room background and the nonuniform detector response. The transmission and the thickness of the sample were also taken into account. The 2D scattering intensities were normalized in absolute units with a standard calibration and radially integrated to obtain 1D diffraction patterns.

Microbeam Particle Induced X-ray Emission (µPIXE). Zinc stoichiometries in the different insulin samples were measured using µPIXE analysis in combination with simultaneous Rutherford backscattering analysis (RBS) to allow correction for sample matrix effects (59, 60). For µPIXE measurements of proteins, sulfur acts as an internal standard. The measurements were carried out at the Ion Beam Centre, University of Surrey, UK (61). A 2.5 MeV proton beam of 1.5 µm in diameter was used to induce characteristic X-ray emission from dried insulin sample droplets (volume per droplet ~0.1 µl) under vacuum. The X-rays were detected using a solid state lithium drifted silicon detector and backscattered protons were detected using a silicon particle detector. By scanning the proton beam in x and y over the dried sample, spatial maps were obtained of all elements heavier than magnesium present in the sample. Quantitative information was obtained by collecting 3 or 4 point spectra from each droplet. PIXE spectra were analyzed with GUPIX (62) using the matrix composition derived from the simultaneous RBS spectrum. Data processing is carried out using the data acquisition software OMDAQ-3 (Oxford Microbeams Ltd, UK) to extract the relative amount of each element in the sample.

X-ray absorption spectroscopy.

Data collection. X-ray absorption spectra were collected at the zinc K-edge (9660.75 eV) on the EXAFS station (BM26A) of the Dutch-Belgian beamline (DUBBLE) (63) at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. The energy of the X-ray beam was tuned by a double-crystal monochromator operating in fixed-exit mode using a Si(111) crystal pair. Three different samples were put in separate glass capillaries and measured in fluorescence mode at ambient temperature and pressure: i) a zinc solution prepared by dissolving 4 mg of $ZnSO_4$ in 10 ml of 0.25 M NaCl solution at pH 1.8, ii) a native insulin solution prepared by dissolving 10 mg of lyophilized insulin powder (Novo Nordisk) in 1 ml of 0.25 M NaCl solution at pH 1.8, and iii) the same solution as in ii) incubated at 60° C for about 24 h to induce amyloid aggregation. The EXAFS spectra, three scans per sample with a new solution for each scan, were energy-calibrated, averaged and further analyzed using GNXAS (64, 65).

Data analysis. In the GNXAS approach, the local atomic arrangement around the absorbing atom is decomposed into model atomic configurations containing 2, ..., n atoms. The theoretical EXAFS signal $\boldsymbol{\chi}(k)$ is given by the sum of the n-body contributions $\boldsymbol{\gamma}^2$, $\boldsymbol{\gamma}^3$, …, $\boldsymbol{\gamma}^n$, which take into account all the possible single and multiple scattering (MS) paths between the n atoms. The fitting of $\boldsymbol{\chi}(k)$ to the experimental EXAFS signal allows refinement of the relevant structural parameters of the different coordination shells; the suitability of the model is also evaluated by comparison of the experimental EXAFS signal Fourier transform (FT) with the FT of the calculated $\boldsymbol{\chi}(k)$ function. The coordination numbers and the global fit parameters that were allowed to vary during the fitting procedure were the distance R(Å), Debye-Waller factor ($\boldsymbol{\sigma}^2$) and the angles of the $\boldsymbol{\gamma}^n$ contributions which were defined according to atomic structural models constructed using the VMD software(66). Zinc was solvated based on a published refined structure (67). For the model containing histidine, the residue was positioned by replacing one coordinating water molecule with it. Weak distance constraints were used, so that water and histidine could compete to coordinate the zinc atom. Simulations were then performed using NAMD 2.13 (68) with the TIP3P model for water (69) and the CHARMM36 force field (70). The Nose-Hoover-Langevin piston algorithm maintained a constant pressure (71, 72), and the stochastic velocity rescaling algorithm controlled the temperature (73). Bonds with hydrogen atoms were constrained using the SHAKE algorithm (74) with a force constant k set to 2, and the Verlet-I/r-RESPA multiple-time step scheme (74–76) integrated the equations of motion, with time steps of 2 fs for long-range nonbonded forces, and 1 fs for short-range and bonded forces. Electrostatic interactions were computed using the smooth Particle Mesh-Ewald (PME) sum (77), with a cutoff set to 10 Å , a switching function starting at 8 Å, and a pairlist distance of 14 Å was used. A snapshot of the simulation run was then used to extract coordinates that were processed by GNXAS. According to the GNXAS approach, we used a single two body configuration $\boldsymbol{\gamma}^2$ for the Zn-L (L=O or N) distance. In order to fit the higher shell contributions two $\boldsymbol{\eta}^3$ corresponding to the Zn-O-C and Zn-O-N three-body configurations, were used (see inset in Figure 5c). Importantly, due to the high value of the vertex angle in the Zn-O-C and Zn-O-N configurations, which are connected to the aromatic ring present in the histidine, the signals of theses shells are enhanced because of the multiple scattering effect, which is considered in the analysis. For these shells, whose Zn-O distance is the first shell, the only free parameter needed was the vertex angle $\boldsymbol{\theta}$.

Dynamic Light Scattering (DLS). DLS measurements were performed using a DynaPro Nanostar (Wyatt Technology) detector in Wyatt cuvettes with 10 µL of protein solution. Each measurement consisted of 10 times five seconds of integration with auto-attenuation enabled. Insulin was dissolved in 0.25 M NaCl, pH 1.8, H_2O solution with or without EDTA, then centrifuged at 16000 rpm for 30 minutes. Samples were measured three times and data were analyzed using DYNAMICS software (Wyatt Technology).

Circular dichroïsm (CD). Insulin was dissolved in pure H_2O by lowering the pH with sulfuric acid to avoid chlorine for CD measurement. Stock insulin solution at 2 mg/mL was filtered, then diluted to reach the desired concentration with or without the appropriate amount of EDTA. Measurements were performed on a JASCO J-810 spectropolarimeter at 20 nm/s scanning speed with six accumulations.

X-ray powder diffraction. A fraction of the insulin-sample powders was used for X-ray diffraction measurements. A small amount was sealed in a Mylar capillary and diffraction patterns were recorded on beamline ID30B at the European Synchrotron Radiation Facility (Grenoble, France) using the following parameters: 20 images with oscillation step of 1° and 0.5 s exposure time for each image (100 % transmission for 4×10^{12} ph/s at $\lambda = 0.98$ Å).

Aggregation kinetics. Insulin was dissolved in 0.25 M NaCl, pH 1.8, H_2O solution, with or without EDTA, then filtered. 100 µL were pipetted into wells of a 96-well plate and 2 µL of 1 mM Thioflavin T (ThT) was added to each. Each well was sealed using vacuum grease and a glass slide to prevent any evaporation. One measurement was then performed every 5 min at 60°C for 24 h using 450 nm excitation light and a 490 nm detection filter for ThT. The measurements were performed on a BioTek Synergy H4 plate-reader using the bottom configuration and a manual gain of 75.

Electron microscopy. Samples were absorbed on the clean side of a carbon film, that was then deposited on mica, stained with 2% (w/v) uranyl acetate (pH 4.5) and transferred to a 400-mesh copper grid. The images were acquired under low dose conditions (<10 e^-/Å^2) with defocus values between 1.2 and 2.5 µm on a Tecnai 12 LaB6 electron microscope at 120 kV accelerating voltage using a Gatan Orius 1000 CCD camera.

Optical microscopy. For each sample observed, Thioflavin T was added to a final concentration of 20 µM. Subsequently, 8 µL of the sample were deposited between a glass slide and a cover slip. The samples were observed by an Olympus IX81 microscope equipped with GFP filter cube set and differential interference contrast (DIC) and a sCMOS

Hamamatsu Orca Flash4 camera. The control of the microscope and the image acquisition were performed using the software Volocity.

Results

Screening and selection of homogeneous aggregated samples. The first step of the present work was to screen and identify the experimental conditions for producing homogeneous samples, i.e. predominantly containing either fibrils or spherulites. This step was necessary to probe any potential difference in protein dynamics depending on the specific aggregate formed. The effect of salt, protein concentration, temperature (which has been shown to affect spherulite size and concentration (14)) and mechanical perturbation on insulin amyloid aggregation was investigated.

Effect of salt and protein concentration. Insulin was incubated for 24 h at 60° C in different solutions with salt and protein concentration ranging from 0.1 to 0.5 M and from 0.5 mg/mL up to 5 mg/mL, respectively. Whereas protein concentration has no relevant effect on the final aggregate formation (i.e. both spherulites and fibrils were found at all the concentrations explored), salt concentration strongly influences it: at high salt concentration (□ 0.5 mM), neither spherulites nor fibrils were observed (Fig. S1 in the Supporting Material).

Effect of temperature. Insulin was then incubated for 24 h at either 50, 60 or 70° C and at salt and protein concentrations of 0.25 M and 2 mg/ml, respectively. Amyloid aggregation in quiescent conditions (i.e. without shaking) occurs only at or above 60° C (Fig. S2 in the Supporting Material).

Effect of mechanical perturbation. Having defined the conditions for maximizing amyloid aggregation in quiescent conditions, mechanical perturbation was used to select either fibrils or spherulites. It was found that shaking the sample at 600 or 900 rpm during incubation produces a sample in which a large majority of the initial native protein is converted into elongated amyloid-like fibrils, even at 50° C. However, even though shaking allows the suppression of spherulites formation, it has a noteworthy effect on the morphology of fibrils. Electron microscopy revealed the presence of well-defined fibrils at 600 rpm and strongly aggregated ThT-positive material at 900 rpm (Fig. S3 in the Supporting Material). A pure spherulite sample could be obtained by heating the protein in 0.25 M NaCl solution at 60° C or 70° C without shaking, and then including a post-production treatment based on a series of vortexing/centrifugations/supernatant removal steps (see Methods for details). After this treatment, no fibrils could be found in the samples (Fig. S4 in the Supporting Material).

The screening described above enabled us to define conditions for obtaining homogeneous fibril and spherulite samples. Fibrils were produced by incubating at 50° C for 24 h a human insulin solution at 5 mg/ml, pH 1.8, 0.25 mM NaCl under shaking at 600 rpm. Spherulites were produced by incubating at 60° C for 24 h a human insulin solution at 5 mg/ml, pH 1.8, 0.25 mM NaCl in quiescent conditions and then isolated by the post-production treatment mentioned above and described in Methods. Both fibril and spherulite samples were then washed in pure D_2O (see Methods for details) and subsequently freeze-dried. As expected, the macroscopic morphology of both fibrils and spherulites appears slightly damaged after freeze-drying, yet the main features of both fibrils and spherulites remain clearly visible (Figure 1). Moreover, the presence of cross-β structure is revealed in both amyloid samples by X-ray powder diffraction, showing marked differences if compared with the alpha-helix rich native state (Fig. S5 in the Supporting Material).

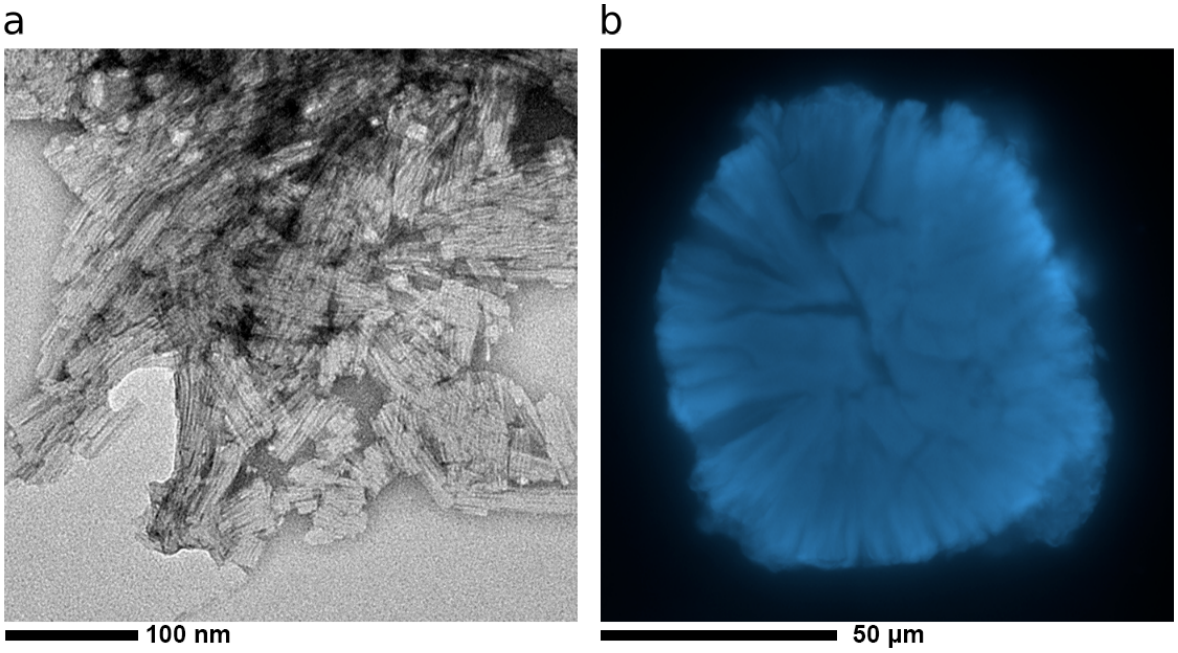


Figure 1. Morphology of insulin amyloid aggregates. a Electron microscopy of freeze-dried insulin fibrils. **b** Confocal microscopy of a freeze-dried insulin spherulite.

Protein dynamical transition is suppressed in insulin fibrils and spherulites. Incoherent neutron scattering of native insulin and of insulin fibrils and spherulites was then employed to measure the elastic incoherent signal as a function of temperature and the momentum transfer *q*. The dependence of the elastic incoherent signal on *q* provides information on the average motions of individual atoms in the system. From the *q*-dependence of incoherent elastic scattering from a D_2O-hydrated protein powder, one can obtain the ensemble-averaged apparent mean square displacements (MSD) of protein hydrogen atoms, reflecting sub-nanosecond signature of vibrations as well as backbone and side chain relaxation (56). The scattering signal from hydrated insulin powders (hydration level h = 0.4 [D_2O mass]/[protein mass]) was acquired during a heating scan at 0.2 K/min from 10 to 300 K, yielding MSDs as a function of temperature (Figure 2). Native insulin shows the typical temperature dependent MSDs of hydrated proteins (53), with harmonic behaviour at cryogenic temperature, a first anharmonic onset at ~100 K, attributed to methyl group rotations entering the experimental window (78–80), and a second onset of water-sustained (81, 82) anharmonic motions above ~240 K (the so-called protein dynamical transition (53)). In both amyloid insulin species, i.e. fibrils and spherulites, the protein dynamical transition is completely suppressed, as in a dry protein (78), proving that this property is independent of the type of amyloid aggregates. Interestingly, the MSDs of native insulin treated with EDTA, known to be a highly efficient zinc-chelating agent, display a similar temperature dependence to that of aggregated samples and also lack the protein dynamical transition (Figure 2). Intriguingly, the dynamical transition can be partially suppressed by successive washing steps in pure D_2O (Fig. S6 in the Supporting Material). As the formation of crystalline ice within the sample might also lead to dewetting of the protein at sub-zero temperatures, and hence to the observed dry-like MSD temperature dependence, we measured the neutron diffraction signal of insulin fibrils and spherulites to determine if crystalline ice had formed at low temperature. As shown in Fig. S7 in the Supporting Material, no Bragg peaks of hexagonal ice are present in the diffraction pattern of fibers at 200 K. A broad peak centered at ~1.65 $Å^{-1}$ and a peak at ~1.35 $Å^{-1}$ indicate the presence of amorphous water and of amyloid cross-β structure, respectively.

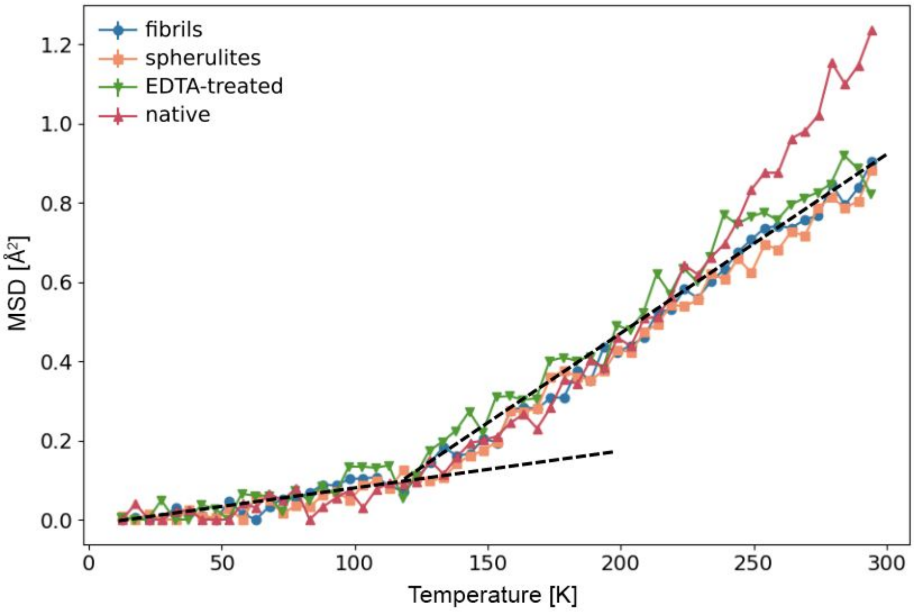


Figure 2. Amyloid aggregation suppresses the protein dynamical transition of hydrated insulin. Mean square displacements (MSD) measured as a function of temperature by elastic incoherent neutron scattering on the backscattering spectrometer SPHERES (MLZ, Garching) for native insulin (red triangles), amyloid fibrils (blue circles) and spherulites (orange squares) and for insulin treated with EDTA (green triangles).

Zinc release affects the kinetics of insulin amyloid aggregation. As revealed by neutron scattering, insulin amyloid species (both fibrils and spherulites) and EDTA-treated insulin (where zinc is expected to be chelated by EDTA and no amyloid aggregation is induced as shown by X-ray diffraction data in Fig. S5) are characterized by the same dynamical behavior. This suggests that the presence/absence of zinc may affect the aggregation kinetics. To verify this, we monitored insulin aggregation kinetics at 60° C as a function of EDTA concentration. Thioflavin T (ThT, a probe whose fluorescence correlates with amyloid formation (83)) fluorescence was used as a probe of aggregate formation. As shown in Figure 3, the kinetics in the absence of EDTA and without mechanical shaking is characterized by a double sigmoidal curve as previously observed (84): a lag phase is followed by a first increase of the fluorescence signal, then by a plateau prior to a second increase. Adding EDTA clearly reduces the double-sigmoidal trend, which is essentially suppressed at high EDTA concentration. As expected, optical and electron microscopy observations at different time points along the aggregation kinetics showed that fibrils and spherulites are formed roughly at the same time, with more numerous and larger spherulites appearing in the presence of EDTA (Figure 3). As shown above, shaking the insulin solution during incubation prevents spherulite formation. The kinetics experiments were then repeated as for those shown in Figure 3 but in the presence of mechanical shaking, which inhibits the formation of spherulites, and results in pure fibril samples. The results are shown in Fig. S8 in the Supporting Material and reveal that shaking suppresses the first plateau phase, while the addition of EDTA in the presence of mechanical perturbation does not produce any relevant effect. In addition, no spherulites were found with or without EDTA with mechanical shaking. In order to explore any possible structural/conformational perturbation directly induced by EDTA, we also verified that in the range of EDTA concentration used both the oligomeric state and secondary structure of native insulin at room temperature are largely unaffected. In fact, a dynamic light scattering (DLS) analysis of size distribution showed a monodisperse signal for native insulin, with a single peak at a radius of approximately 2.8 nm (suggestive of the presence of an insulin dimer), barely affected upon addition of EDTA up to a concentration of 5 mM (Fig. S9 in the Supporting Material). Circular dichroism (CD) in the 200-260 nm spectral region indicated that insulin secondary structure content is unchanged upon EDTA addition even at high protein:EDTA ratios (Fig. S10 in the Supporting Material).

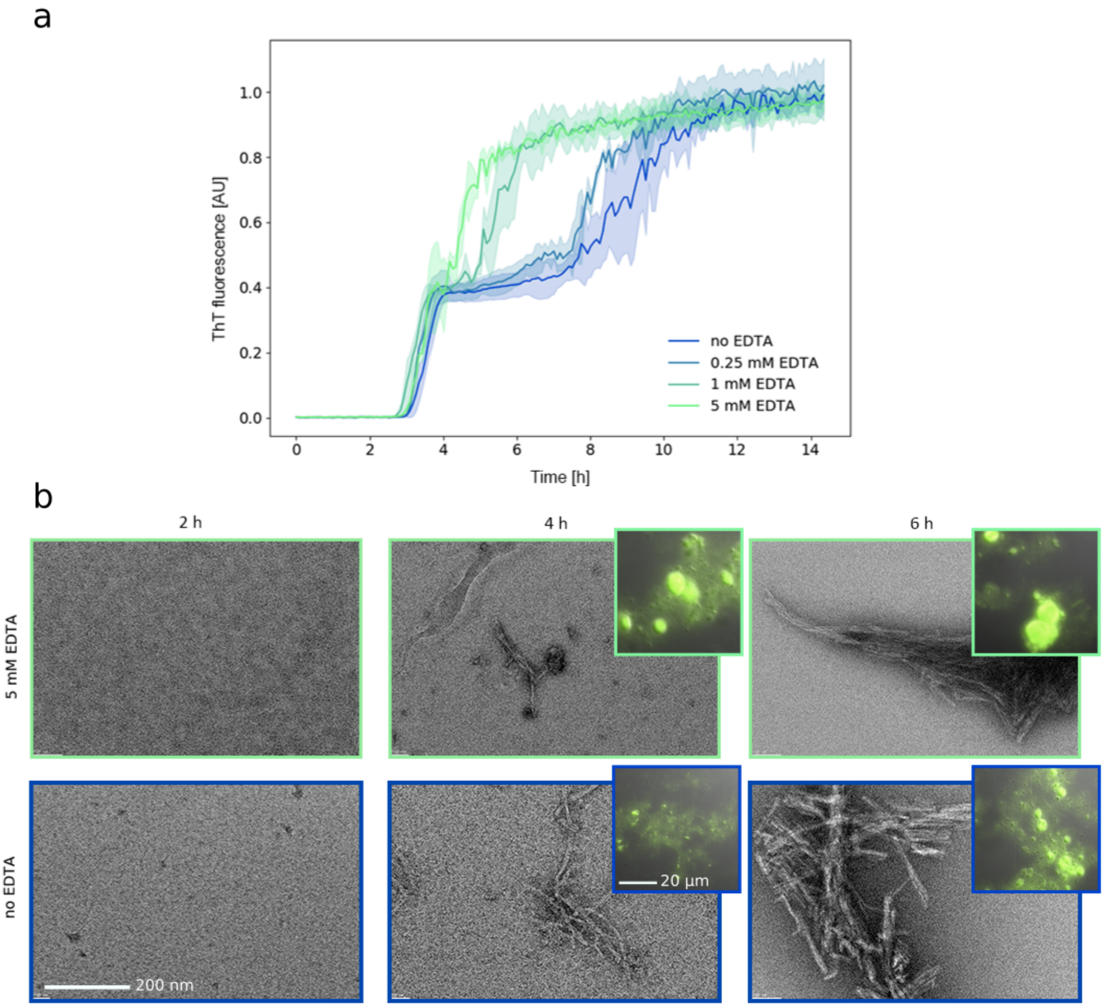


Figure 3. EDTA modifies insulin aggregation kinetics. a Insulin aggregation kinetics at 60° C as probed by Thioflavin T fluorescence as a function of EDTA concentration. The color shaded areas indicate the standard deviation for three independent aggregation processes. **b** Transmission electron microscopy (fibrils) and fluorescence microscopy (spherulites, insets) images acquired during the kinetics without EDTA (blue curve in **a,** blue contoured images in **b**) and at 5 mM EDTA (light green in **a**, light green contoured images in **b**). Micrographs were acquired after 2 h (left), 4 h (middle) and 6 h (right) of incubation at 60° C and showed that fibrils and spherulites were formed roughly at the same time, with more numerous and larger spherulites appearing in the presence of EDTA.

Insulin releases zinc upon aggregation. In the light of the effect of EDTA on both insulin dynamics and aggregation kinetics, our hypothesis is then that, upon aggregation, insulin molecules release the zinc ions. To test this hypothesis, the presence of zinc in the different samples was explored. μPIXE (59, 60) analysis was first used to determine the zinc concentration in the insulin samples investigated by neutron scattering. The results are shown in Figure 4 and stoichiometric ratios are summarized in Tab. S1 in the Supporting Material. μPIXE analysis revealed that native insulin contained 1.5 zinc atoms per insulin monomer, while zinc was below the detection limit (<0.001 atoms of zinc/monomer) in both fibrils and spherulites and, as expected, most of the zinc was removed in the EDTA-treated sample (0.02 zinc atoms per insulin monomer). However, the washed insulin sample still contains zinc with a stoichiometry of 0.116 zinc atoms per insulin monomer.

To get insights into the zinc-insulin interaction in native and amyloid conditions in solution we used XAS spectroscopy. XAS can be divided into X-ray absorption near-edge spectroscopy (XANES) and extended X-ray absorption fine structure (EXAFS), with XANES providing information on the oxidation state, geometry and electronic configuration of the absorber atom. In contrast, EXAFS allows solving an atom's local environment in terms of the number of neighboring atoms, their distance from the absorber atom and their disorder on a short length scale. XAS experiments were performed at the zinc K-edge (9660.75 eV) on both a native insulin solution and an insulin solution incubated at 60° C to induce the formation of amyloid species (since in quiescent conditions both fibrils and spherulites are always formed, as described above, this sample was named “aggregates”), and on a $ZnSO_4$ water solution. The XANES spectra in Figure S11 in the Supporting Material do not show any significant difference between the various samples, indicating that Zn^{2+} ions are solvated in all the samples. Figure 5 (panel a) shows k^3-weighted EXAFS transmission spectra and their k^3-weighted Fourier transforms (panel b) for the three samples. The experimental curves in real space (Figure 5b) are comparable up to about 3 Å, indicating that the first shell is very similar in all three samples. At larger distances, the curve for native insulin shows small but significant differences in the 3-5 Å range compared with the other two samples (in particular the peak at ~4.2 Å present only in native insulin). Small but significant differences are also evident in the reciprocal space (Figure 5a), namely a shoulder at 4 $Å^{-1}$ and a broadening of the peak around 7.6 $Å^{-1}$ visible only in native insulin. Figure 5a shows fitted curves in terms of structural models (see Methods). The structural models contain either only a solvated Zn^{2+} ion, used to analyze data of both the $ZnSO_4$ and the amyloid insulin solution samples, or a solvated Zn^{2+} ion in the vicinity of a histidine residue (Figure 5c), known to coordinate zinc in insulin, used to analyze data of the native insulin solution. Both models were obtained from constrained molecular dynamics simulations, based on refined geometrical parameters (85) and used to extract coordinates that were processed by GNXAS, as described in the Methods. Structural parameters obtained from this analysis are reported in Tab. S2 in the Supporting Material. A good agreement was found between experimental and calculated curves (Figure 5a). The Fourier-transformed calculated curves also reproduced the peak positions of the Fourier-transformed experimental data up to ~4.5 Å, in particular the differences in the 3-5 Å range (Figure 5b). XAS results indicate that Zn^{2+} is coordinated with a histidine residue in the native insulin solution while it is released in solution upon amyloid formation.

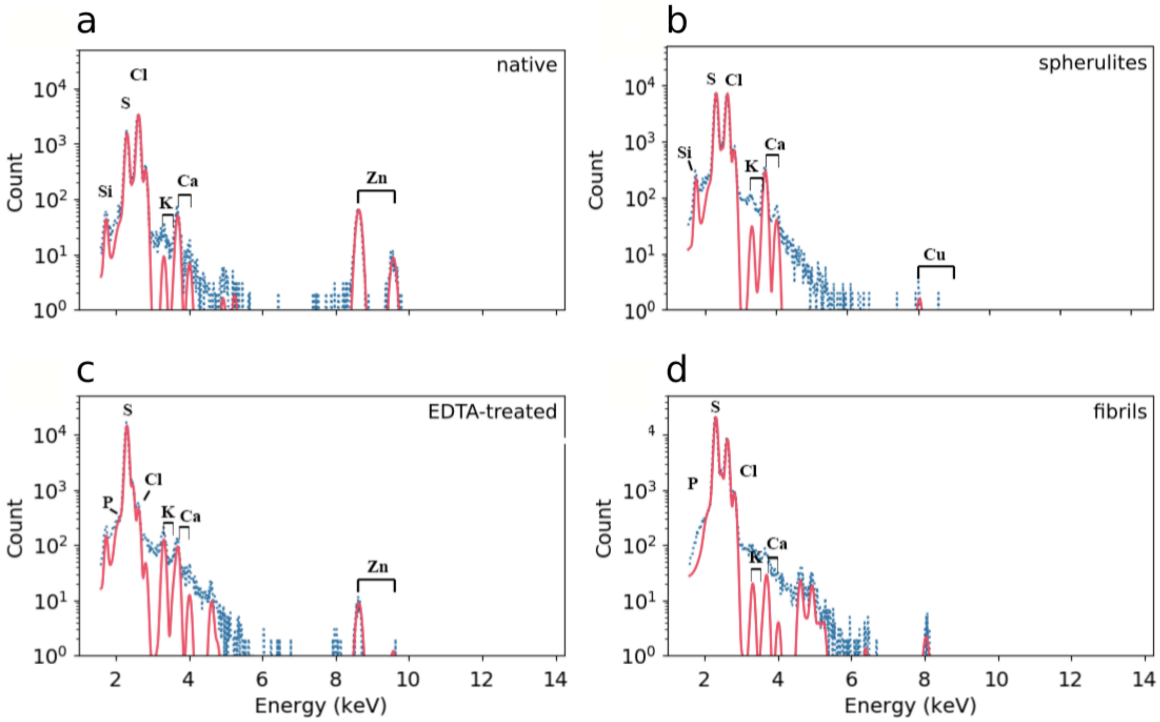


Figure 4. Insulin loses coordinated zinc upon amyloid aggregation. Proton Induced X-ray Emission (PIXE) spectra (blue points) and fit to the characteristic peaks (red lines) for (**a**) native insulin, (**b**) insulin spherulites, (**c**) EDTA treated insulin, and (**d**) insulin fibrils. The theoretical positions for X-ray emission maxima are indicated in all cases.

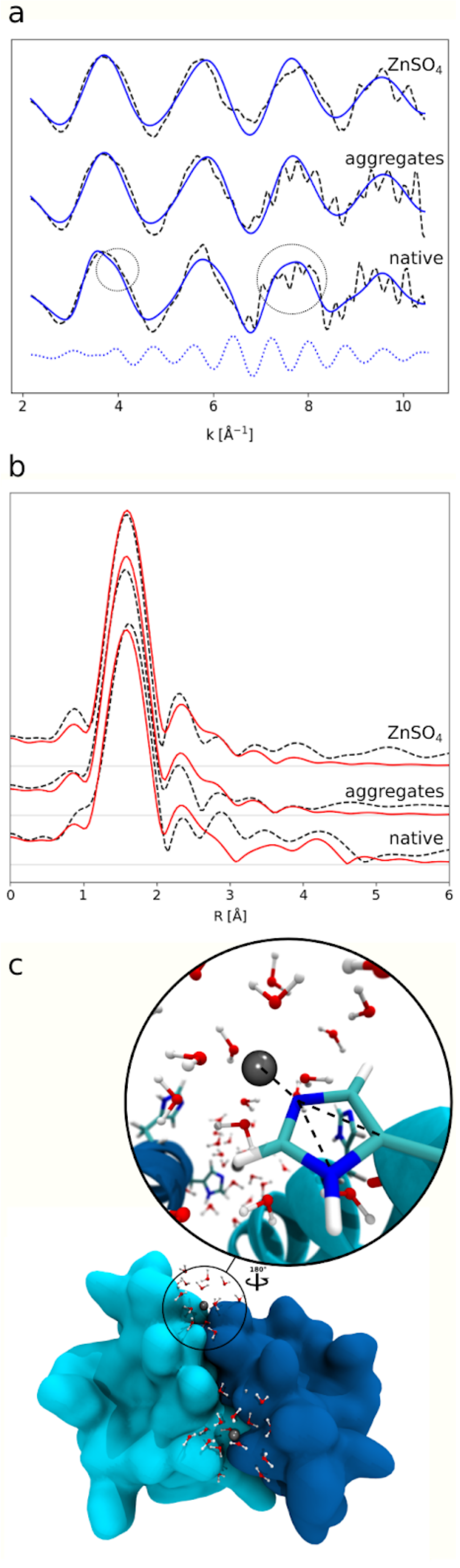


Figure 5. X-ray absorption spectroscopy sheds light on the zinc environment in native insulin and insulin aggregates in solution. a k^3-weighted EXAFS transmission spectra (dashed lines) of a native insulin solution, an insulin solution treated at 60° C for about 24 h (aggregates), and a $ZnSO_4$ solution. Blue continuous lines represent the signals simulated on the basis of the structure of a fully solvated Zn^{2+} ion (for the $ZnSO_4$ solution and for the aggregates) and of the structure of a solvated Zn^{2+} ion in the presence of a histidine residue (native insulin, see panel **c**). Details on the structural models are given in the Methods. Blue dotted line represents the component of the simulated signal that is responsible for the features present only in native insulin and indicated by the circles (i.e. a shoulder at 4 Å^{-1} and a broadening of the peak around 7.6 Å^{-1}). **b** k^3-weighted Fourier-transformed EXAFS transmission spectra (dashed lines) of the same samples as in the top panel (**a**). Red continuous lines represent the signals simulated with the same structural models as in the top panel (**a**). **c** The structure of the dimeric human insulin (PDB 5CNY) is represented with a zoom on the structure of the zinc environment used to analyze the EXAFS data. Zinc is represented by a grey bead surrounded by water molecules and an histidine residue. Dashed black lines indicate the multiple scattering contribution from the histidine to the EXAFS signal.

Discussion

Our study aimed at comparing the molecular dynamics of native and amyloid insulin species by neutron scattering and complementary biophysical methods. Zinc binding proved to be pivotal in controlling not only molecular dynamics at mesoscopic scale but also the aggregation kinetics. To enable the comparison of various insulin species and monitor potential differences in the protein dynamics depending on the type of aggregates, we first established protocols to produce batches containing either native insulin, insulin fibrils or insulin spherulites at a high level of homogeneity. We found that while fine-tuning temperature and salt concentration allows amyloid aggregation into both fibrils and spherulites to be observed, shaking the vial during incubation leads to fibril formation and full inhibition of spherulite formation. We can argue that this inhibition of spherulite formation originates from mechanical perturbation preventing the formation of sufficiently stable spherulite-specific precursors. Spherulites are always produced along with fibrils in quiescent conditions. However, taking advantage of the much larger sedimentation rate of spherulites compared to the other species, they could be isolated from fibrils using gentle centrifugation.

Neutron backscattering spectroscopy showed a drastic change in the dynamics between native insulin and insulin in aggregated form above ~240 K, but no significant difference between the two types of aggregates (spherulites and fibers) were observed (Figure 2). Indeed large-amplitude motions associated with the protein dynamical transition are fully suppressed in insulin amyloid fibers and spherulites, so that the temperature-dependence of the mean square displacements resembles that of completely dry, non-functional proteins. In the native-insulin control measurement, a dynamical transition takes place (Figure 2). These results are in contrast with our previous results on tau protein (52), where no differences between native-state and fibril dynamics were observed on the same timescale, and on concanavalin A, where internal dynamics were found to be larger in the amyloid than in the native species (51). The tau results were suggestive of the protein internal energy depending only on the amino acid composition (86) and not on the structure or aggregation state, as also supported by solution studies of protein denaturation showing that the internal dynamics remain unchanged upon denaturation (87).

In order to rationalize the differential dynamical behavior of native and amyloid insulin, we suspected that interaction with Zn^{2+} ions might play a role, since zinc is known to be essential for insulin expression and activity (33). We investigated how zinc binding affects the kinetics of insulin amyloid aggregation. Upon thermal treatment of the insulin solutions, ThT fluorescence profiles first showed an increase in the signal barely dependent on zinc presence, which has been suggested to be related to the formation of prefibrillar aggregates or early formation of fibrils (84) small enough to leave turbidity and temperature dependence unaffected. We attributed the intermediate plateau phase to spherulite formation as shown by fluorescence and electron microscopies (Figure 3): both fibrils and spherulites are present, with spherulites being larger and more numerous in the presence of EDTA, i.e. when zinc is removed. After the first plateau a second increase occurs, followed by a final plateau, where spherulites are densely present and significantly larger. Increasing EDTA concentration gradually suppresses the double-sigmoidal behavior and enhances spherulite formation. This indicates that the presence of zinc is the limiting step for the more efficient conversion of insulin native molecules into spherulites. When the zinc is fully removed by adding a sufficient amount of EDTA, the intermediate phase disappears. If the loss of zinc occurs at the level of single molecules prior to aggregation or happens once the molecules are embedded in high MW species remains to be clarified. Importantly, shaking also suppresses the double-sigmoidal behavior as well as the spherulite formation. From the kinetic results we can conclude that: 1) the first plateau can be attributed to spherulite formation; 2) zinc binding inhibits spherulite formation, probably through electrostatic repulsion.

To obtain quantitative information on zinc content in our samples, we performed measurements with the µPIXE technique, the results of which unambiguously show that zinc was below the limit of detection (<0.001 zinc atom/monomer, Tab. S1) in amyloid aggregates, where the dynamical transition is suppressed. Furthermore, an EDTA-treated control sample is shown by µPIXE to have a zinc content reduced by a factor of 75 compared with the native protein (Figure 4), and lacks the protein dynamical transition at 240 K in neutron scattering experiments (Figure 2). We note that the EDTA-treated sample and the native form present the same secondary structure according to CD data (Fig. S10) and are predominantly dimeric according to DLS (Fig. S9). This means that both the conformational and colloidal stability of the solution are not significantly influenced by the presence of EDTA and the loss of zinc is mainly affecting the dynamics of the single protein molecule. Indeed, an insulin sample with a zinc content intermediate between those of native and amyloid insulin (as confirmed by μPIXE, Tab. S1), obtained by successive washing steps of a native sample, shows a dynamical transition but with reduced MSD if compared to the native sample (Fig. S6). Zinc binding to insulin thus modulates molecular dynamics and its absence suppresses functionally relevant motions activated at the dynamical transition in amyloid insulin, without modifying secondary, tertiary or quaternary structure. XAS data in solution (Figure 5) confirmed that zinc is coordinated with histidine residues in native insulin, while it is released upon amyloid aggregation. The XAS results thus established that zinc loss is not an effect of the sample treatment (i.e. buffer exchange and freeze-drying) preceding neutron scattering experiments.

Translating our *in vitro* results to a physiological context, they imply that amyloid aggregation of insulin *in vivo* could induce an accumulation of free zinc ions at the injection sites, where IDA is observed. Zinc is anyway released from monomeric insulin during its function (33), yet our finding suggests that a substantially larger quantity of zinc may be released and could accompany IDA. Zinc overload is known to be a pro-oxidant condition leading to oxidative damage of biomolecules. When free zinc ion concentrations increase, zinc can bind to proteins that otherwise would not interact with zinc and affect their functions, thus inducing the production of free radicals by inhibiting antioxidant enzymes and the mitochondrial respiratory chain (88). On the basis of our results we are not able to quantitatively assess the potential cytotoxicity of zinc release during IDA, because the threshold at which zinc ion concentration becomes high enough to produce pro-oxidant/cytotoxic effects depends on the zinc buffering capacity of cells (33, 88). Future *in vivo* investigations could clarify the physiological impact of the zinc accumulation revealed by our study.

An important information provided by the XAS results is that even in native insulin, which coordinates zinc, the metal ion environment is largely composed of water molecules. The presence of the high charge on the Zn^{2+} ion is known to perturb water molecules in the first and second solvation shells, which correspond to at least 20 water molecules interacting with one Zn^{2+} ion (89). Considering the 1.5 Zn/protein ratio indicated by the μPIXE analysis, the presence of zinc on the protein surface corresponds to a minimum hydration level $h = [20 \times 1.5 \times \text{water MW}]/[\text{insulin MW}] \simeq 0.1$. In other words, the presence of Zn^{2+} favors the hydration of the highly hydrophobic insulin surface, thus influencing the dynamical behaviour as revealed by neutron scattering.

Taken together, the results from neutron scattering, μPIXE, and XAS clearly show that zinc binding is the key parameter which allows hydrated insulin to undergo the protein dynamical transition, while the dramatic dynamical effect on fibrils and spherulites has to be attributed to loss of zinc. Removal of zinc results in a complete loss of water-sustained backbone and side-chain motions with the MSD profile corresponding to that of a dry protein, thus suggesting a dehydration effect upon zinc release. As previously suggested (52, 90, 91), the changes in hydration water entropy may be one of the 'driving forces' leading proteins to form amyloid aggregates. Restructuring of hydration water is expected to be involved both in the initial steps of unfolding/crossβ-amyloid formation and in the growth of amyloid aggregates, thereby implying a possible change in the water configurational entropy (90). Protein dehydration has also been shown to modify amyloid aggregation kinetics (92). The presence of metal ions integrated within the protein structure, which generally influences protein solvation and electrostatics of the protein surface (93), can add a key parameter to the energetic balance between native and amyloid state. In the case of human insulin, which is highly hydrophobic(94), we postulate that zinc binding favors protein hydration, thus modifying both protein dynamics, as revealed by neutron scattering, and water entropy. Indeed, dewetting of hydrophobic groups (which are known to reduce water mobility (95) and entropy (96)) during aggregation, and subsequent release of water molecules in the bulk phase, corresponds to an increase of water entropy (91). The presence of zinc coordinated to the protein reduces the hydrophobicity of the insulin surface and then the propensity for aggregation, whereas zinc removal increases the entropy difference between native and amyloid state and then promotes amyloid aggregation.

It has been shown that a Thr to His mutation in the chain A of human insulin leads to a significant increase in the lag phase of amyloid aggregation (97). More generally, studies of protein rational design (48, 98) have shown how the introduction of His binding sites for Zn^{2+} into hydrophobic peptides can modify both their stability and dynamics. Our results suggest that protein hydration changes upon Zn^{2+} binding/release can contribute to such modifications, as revealed by our finding that both dynamics and aggregation kinetics of human insulin change as an effect of Zn^{2+} release from His binding sites. Changes in metal-modulated protein hydration might be a general strategy to control protein stability and tune protein aggregation into amorphous and ordered super-structures.

Author contributions

K.P., V.F., M.W., and G.S. designed the research; K.P., G.S., D.N., and M.Z. performed neutron backscattering spectroscopy and K.P. analyzed data; K.P. and V.C. performed neutron diffraction and analyzed data; K.P. and V.F. defined sample preparation conditions and K.P. prepared and characterized the samples; G. W. G. and E. F. G. performed μPIXE and analyzed data; A.L., K.P. and G.S. performed XAS and A.L. analyzed data based on structures extracted from MD simulations performed by K.P.; K.P. performed CD, DLS and fluorescence spectroscopy; K.P. and G.S. performed X-ray powder diffraction; K.P., V.F., and G.S. wrote the paper with input from all authors.
The authors declare no competing interests.

Acknowledgments

This work used the platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). V.F. acknowledges VILLUM FONDEN for the Villum Young Investigator Grant *"Protein Superstructures as Smart Biomaterials (ProSmart)*" 2018-2023 (project number: 19175).
The electron microscope facility is supported by the Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER, the Centre National de la Recherche Scientifique (CNRS), the CEA, the University of Grenoble, EMBL, and the GIS-Infrastructures en Biologie Santé et Agronomie (IBISA).
The financial support provided by JCNS to perform neutron scattering measurements at the Heinz Maier-Leibnitz Zentrum (MLZ), Garching, Germany is gratefully acknowledged. The access to ILL and ESRF beamlines to perform diffraction characterization is also acknowledged.
We thank Daphna Fenel, Christine Moriscot and Guy Schoehn from the Electron Microscopy platform, Caroline Mas and Marc Jamin for assistance and access to the Biophysical platform, Aline Le Roy, Michel Thépaut and Christine Ebel for assistance and access to the Protein Analysis On Line (PAOL) platform.

References

1. Stefani, M. 2004. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta*. 1739:5–25.

2. Fitzpatrick, A.W.P., B. Falcon, S. He, A.G. Murzin, G. Murshudov, H.J. Garringer, R.A. Crowther, B. Ghetti, M. Goedert, and S.H.W. Scheres. 2017. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature*. 547:185–190.

3. Tuttle, M.D., G. Comellas, A.J. Nieuwkoop, D.J. Covell, D.A. Berthold, K.D. Kloepper, J.M. Courtney, J.K. Kim, A.M. Barclay, A. Kendall, W. Wan, G. Stubbs, C.D. Schwieters, V.M.Y. Lee, J.M. George, and C.M. Rienstra. 2016. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nat. Struct. Mol. Biol.* 23:409–415.

4. Vetri, V., and V. Foderà. 2015. The route to protein aggregate superstructures: Particulates and amyloid-like spherulites. *FEBS Lett.* 589:2448–2463.

5. Exley, C., E. House, J.F. Collingwood, M.R. Davidson, D. Cannon, and A.M. Donald.

2010. Spherulites of Amyloid-β42 In Vitro and in Alzheimer's Disease. *Journal of Alzheimer's Disease*. 20:1159–1165.

6. Schack, M.M., E.H. Møller, A.V. Friderichsen, J.F. Carpenter, T. Rades, and M. Groenning. 2019. Optimization of Infrared Microscopy to Assess Secondary Structure of Insulin Molecules Within Individual Subvisible Particles in Aqueous Formulations. *J. Pharm. Sci.* 108:1117–1129.

7. Knowles, T.P.J., and R. Mezzenga. 2016. Amyloid Fibrils as Building Blocks for Natural and Artificial Functional Materials. *Adv. Mater.* 28:6546–6561.

8. Fändrich, M., S. Nyström, K.P.R. Nilsson, A. Böckmann, H. LeVine 3rd, and P. Hammarström. 2018. Amyloid fibril polymorphism: a challenge for molecular imaging and therapy. *J. Intern. Med.* 283:218–237.

9. Condello, C., T. Lemmin, J. Stöhr, M. Nick, Y. Wu, A.M. Maxwell, J.C. Watts, C.D. Caro, A. Oehler, C.D. Keene, T.D. Bird, S.G. van Duinen, L. Lannfelt, M. Ingelsson, C. Graff, K. Giles, W.F. DeGrado, and S.B. Prusiner. 2018. Structural heterogeneity and intersubject variability of Aβ in familial and sporadic Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 115:E782–E791.

10. Rasmussen, J., J. Mahler, N. Beschorner, S.A. Kaeser, L.M. Häsler, F. Baumann, S. Nyström, E. Portelius, K. Blennow, T. Lashley, N.C. Fox, D. Sepulveda-Falla, M. Glatzel, A.L. Oblak, B. Ghetti, K.P.R. Nilsson, P. Hammarström, M. Staufenbiel, L.C. Walker, and M. Jucker. 2017. Amyloid polymorphisms constitute distinct clouds of conformational variants in different etiological subtypes of Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 114:13018–13023.

11. Jiménez, J.L., E.J. Nettleton, M. Bouchard, C.V. Robinson, C.M. Dobson, and H.R. Saibil. 2002. The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 99:9196–9201.

12. Brange, J., L. Andersen, E.D. Laursen, G. Meyn, and E. Rasmussen. 1997. Toward understanding insulin fibrillation. *J. Pharm. Sci.* 86:517–525.

13. Krebs, M.R.H., E.H.C. Bromley, S.S. Rogers, and A.M. Donald. 2005. The mechanism of amyloid spherulite formation by bovine insulin. *Biophys. J.* 88:2013–2021.

14. Smith, M.I., V. Foderà, J.S. Sharp, C.J. Roberts, and A.M. Donald. 2012. Factors affecting the formation of insulin amyloid spherulites. *Colloids Surf. B Biointerfaces*. 89:216–222.

15. Vetri, V., F. Piccirilli, J. Krausser, G. Buscarino, U. Łapińska, B. Vestergaard, A. Zaccone, and V. Foderà. 2018. Ethanol Controls the Self-Assembly and Mesoscopic Properties of Human Insulin Amyloid Spherulites. *The Journal of Physical Chemistry B*. 122:3101–3112.

16. Shikama, Y., J.-I. Kitazawa, N. Yagihashi, O. Uehara, Y. Murata, N. Yajima, R. Wada, and S. Yagihashi. 2010. Localized Amyloidosis at the Site of Repeated Insulin Injection in a Diabetic Patient. *Intern. Med.* 49:397–401.

17. Swift, B., P.N. Hawkins, C. Richards, and R. Gregory. 2002. Examination of insulin injection sites: an unexpected finding of localized amyloidosis. *Diabet. Med.* 19:881–882.

18. Gupta, Y., G. Singla, and R. Singla. 2015. Insulin-derived amyloidosis. *Indian J. Endocrinol. Metab.* 19:174–177.

19. Dische, F.E., C. Wernstedt, G.T. Westermark, P. Westermark, M.B. Pepys, J.A. Rennie, S.G. Gilbey, and P.J. Watkins. 1988. Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia*. 31:158–161.

20. Nagase, T., Y. Katsura, Y. Iwaki, K. Nemoto, H. Sekine, K. Miwa, T. Oh-I, K. Kou, K. Iwaya, M. Noritake, and T. Matsuoka. 2009. The insulin ball. *Lancet*. 373:184.

21. Samlaska, C., S. Reber, and T. Murry. 2020. Insulin-derived amyloidosis: The insulin ball, amyloidoma. *JAAD Case Rep*. 6:351–353.

22. Nilsson, M.R. 2016. Insulin amyloid at injection sites of patients with diabetes. *Amyloid*. 23:139–147.

23. Ohno, Y., T. Seki, Y. Kojima, R. Miki, Y. Egawa, O. Hosoya, K. Kasono, and T. Seki. 2019. Investigation of factors that cause insulin precipitation and/or amyloid formation in insulin formulations. *J Pharm Health Care Sci*. 5:22.

24. Xie, J., A. Li, and J. Li. 2017. Advances in pH-Sensitive Polymers for Smart Insulin Delivery. *Macromol. Rapid Commun.* 38:1700413.

25. Ristow, M. 2004. Neurodegenerative disorders associated with diabetes mellitus. *Journal of Molecular Medicine*. 82.

26. Craft, S., and G.S. Watson. 2004. Insulin and neurodegenerative disease: shared and specific mechanisms. *Lancet Neurol.* 3:169–178.

27. Schubert, M., D. Gautam, D. Surjo, K. Ueki, S. Baudler, D. Schubert, J. Alber, C.R. Kahn, W. Krone, and J.C. Brüning. 2004. The role for neuronal insulin resistance in neurodegenerative diseases. *Experimental and Clinical Endocrinology & Diabetes*. 112.

28. Kupke, D.W., and K. Linderstrøm-Lang. 1954. On the size of the monomer of insulin. *Biochim. Biophys. Acta*. 13:153–154.

29. Pease, L.F., 3rd, M. Sorci, S. Guha, D.-H. Tsai, M.R. Zachariah, M.J. Tarlov, and G. Belfort. 2010. Probing the nucleus model for oligomer formation during insulin amyloid fibrillogenesis. *Biophys. J.* 99:3979–3985.

30. Foderà, V., and A.M. Donald. 2010. Tracking the heterogeneous distribution of amyloid spherulites and their population balance with free fibrils. *Eur. Phys. J. E Soft Matter*. 33:273–282.

31. Vestergaard, B., M. Groenning, M. Roessle, J.S. Kastrup, M. van de Weert, J.M. Flink, S. Frokjaer, M. Gajhede, and D.I. Svergun. 2007. A helical structural nucleus is the

primary elongating unit of insulin amyloid fibrils. *PLoS Biol.* 5:e134.

32. Kim, A.C., S. Lim, and Y.K. Kim. 2018. Metal Ion Effects on Aβ and Tau Aggregation. *Int. J. Mol. Sci.* 19.

33. Li, Y.V. 2014. Zinc and insulin in pancreatic beta-cells. *Endocrine*. 45:178–189.

34. Noormägi, A., J. Gavrilova, J. Smirnova, V. Tõugu, and P. Palumaa. 2010. Zn(II) ions co-secreted with insulin suppress inherent amyloidogenic properties of monomeric insulin. *Biochem. J.* 430:511–518.

35. Frankær, C.G., P. Sønderby, M.B. Bang, R.V. Mateiu, M. Groenning, J. Bukrinski, and P. Harris. 2017. Insulin fibrillation: The influence and coordination of Zn. *J. Struct. Biol.* 199:27–38.

36. Domínguez-Calva, J.A., M.L. Pérez-Vázquez, E. Serebryany, J.A. King, and L. Quintanar. 2018. Mercury-induced aggregation of human lens γ-crystallins reveals a potential role in cataract disease. *J. Biol. Inorg. Chem.* 23:1105–1118.

37. Ahmadi, S., S. Zhu, R. Sharma, D.J. Wilson, and H.-B. Kraatz. 2019. Interaction of metal ions with tau protein. The case for a metal-mediated tau aggregation. *J. Inorg. Biochem.* 194:44–51.

38. Su, X.-C., B.-B. Pan, and Y. Yang. 2019. Coordination of platinum to α-synuclein inhibits filamentous aggregation in solution. *ChemBioChem*.

39. Lee, M., J.I. Kim, S. Na, and K. Eom. 2018. Metal ions affect the formation and stability of amyloid β aggregates at multiple length scales. *Phys. Chem. Chem. Phys.* 20:8951–8961.

40. Roosen-Runge, F., F. Zhang, F. Schreiber, and R. Roth. 2014. Ion-activated attractive patches as a mechanism for controlled protein interactions. *Sci. Rep.* 4:7016.

41. Sheng, J., N.K. Olrichs, W.J. Geerts, D.V. Kaloyanova, and J.B. Helms. 2019. Metal ions and redox balance regulate distinct amyloid-like aggregation pathways of GAPR-1. *Sci. Rep.* 9:15048.

42. Lautenschläger, J., A.D. Stephens, G. Fusco, F. Ströhl, N. Curry, M. Zacharopoulou, C.H. Michel, R. Laine, N. Nespovitaya, M. Fantham, D. Pinotsi, W. Zago, P. Fraser, A. Tandon, P. St George-Hyslop, E. Rees, J.J. Phillips, A. De Simone, C.F. Kaminski, and G.S.K. Schierle. 2018. C-terminal calcium binding of α-synuclein modulates synaptic vesicle interaction. *Nat. Commun.* 9:712.

43. Deas, E., N. Cremades, P.R. Angelova, M.H.R. Ludtmann, Z. Yao, S. Chen, M.H. Horrocks, B. Banushi, D. Little, M.J. Devine, P. Gissen, D. Klenerman, C.M. Dobson, N.W. Wood, S. Gandhi, and A.Y. Abramov. 2016. Alpha-Synuclein Oligomers Interact with Metal Ions to Induce Oxidative Stress and Neuronal Death in Parkinson's Disease. *Antioxid. Redox Signal.* 24:376–391.

44. Mantyh, P.W., J.R. Ghilardi, S. Rogers, E. DeMaster, C.J. Allen, E.R. Stimson, and J.E. Maggio. 1993. Aluminum, iron, and zinc ions promote aggregation of physiological

concentrations of β-amyloid peptide. *J. Neurochem.* 61:1171–1174.

45. Huang, X., C.S. Atwood, R.D. Moir, M.A. Hartshorn, R.E. Tanzi, and A.I. Bush. 2004. Trace metal contamination initiates the apparent auto-aggregation, amyloidosis, and oligomerization of Alzheimer's Aβ peptides. *J. Biol. Inorg. Chem.* 9:954–960.
46. Moir, Moir, Atwood, Huang, Tanzi, and Bush. 1999. Mounting evidence for the involvement of zinc and copper in Alzheimer's disease. *European Journal of Clinical Investigation*. 29:569–570.
47. Ariöz, C., and P. Wittung-Stafshede. 2018. Folding of copper proteins: role of the metal? *Q. Rev. Biophys.* 51:e4.
48. Handel, T.M., S.A. Williams, and W.F. DeGrado. 1993. Metal ion-dependent modulation of the dynamics of a designed protein. *Science*. 261:879–885.
49. Acharya, S., B. Ahmad, and L. Lapidus. 2013. Aggregation of Alpha-Synuclein is Kinetically Controlled by Intramolecular Diffusion. *Biophysical Journal*. 104:51a.
50. Srivastava, K.R., and L.J. Lapidus. 2017. Prion protein dynamics before aggregation. *Proc. Natl. Acad. Sci. U. S. A.* 114:3572–3577.
51. Schirò, G., V. Vetri, B. Frick, V. Militello, M. Leone, and A. Cupane. 2012. Neutron Scattering Reveals Enhanced Protein Dynamics in Concanavalin A Amyloid Fibrils. *J. Phys. Chem. Lett.* 3:992–996.
52. Fichou, Y., G. Schirò, F.-X. Gallat, C. Laguri, M. Moulin, J. Combet, M. Zamponi, M. Härtlein, C. Picart, E. Mossou, H. Lortat-Jacob, J.-P. Colletier, D.J. Tobias, and M. Weik. 2015. Hydration water mobility is enhanced around tau amyloid fibers. *Proc. Natl. Acad. Sci. U. S. A.* 112:6365–6370.
53. Doster, W., S. Cusack, and W. Petry. 1989. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature*. 337:754–756.
54. Wuttke, J., A. Budwig, M. Drochner, H. Kämmerling, F.-J. Kayser, H. Kleines, V. Ossovyi, L.C. Pardo, M. Prager, D. Richter, G.J. Schneider, H. Schneider, and S. Staringer. 2012. SPHERES, Jülich's high-flux neutron backscattering spectrometer at FRM II. *Rev. Sci. Instrum.* 83:075109.
55. Zamponi, M., and M. Khaneft. 2015. SPHERES: Backscattering spectrometer. *Journal of large-scale research facilities JLSRF*. 1:30.
56. Gabel, F., D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccai. 2002. Protein dynamics studied by neutron scattering. *Q. Rev. Biophys.* 35:327–367.
57. Yi, Z., Y. Miao, J. Baudry, N. Jain, and J.C. Smith. 2012. Derivation of mean-square displacements for protein dynamics from elastic incoherent neutron scattering. *J. Phys. Chem. B*. 116:5028–5036.
58. Kneller, G.R., and K. Hinsen. 2009. Quantitative model for the heterogeneity of atomic

position fluctuations in proteins: a simulation study. *J. Chem. Phys.* 131:045104.

59. Garman, E.F., and G.W. Grime. 2005. Elemental analysis of proteins by microPIXE. *Prog. Biophys. Mol. Biol.* 89:173–205.

60. Grime, G.W., O.B. Zeldin, M.E. Snell, E.D. Lowe, J.F. Hunt, G.T. Montelione, L. Tong, E.H. Snell, and E.F. Garman. 2020. High-Throughput PIXE as an Essential Quantitative Assay for Accurate Metalloprotein Structural Analysis: Development and Application. *J. Am. Chem. Soc.* 142:185–197.

61. Grime, G.W., M. Dawson, M. Marsh, I.C. McArthur, and F. Watt. 1991. The Oxford submicron nuclear microscopy facility. *Nucl. Instrum. Methods Phys. Res. B*. 54:52–63.

62. Maxwell, J.A., W.J. Teesdale, and J.L. Campbell. 1995. The Guelph PIXE software package II. *Nucl. Instrum. Methods Phys. Res. B*. 95:407–421.

63. Nikitenko, S., A.M. Beale, A.M.J. van der Eerden, S.D.M. Jacques, O. Leynaud, M.G. O'Brien, D. Detollenaere, R. Kaptein, B.M. Weckhuysen, and W. Bras. 2008. Implementation of a combined SAXS/WAXS/QEXAFS set-up for time-resolvedin situexperiments. *Journal of Synchrotron Radiation*. 15:632–640.

64. Filipponi, A., Di Cicco A, and C.R. Natoli. 1995. X-ray-absorption spectroscopy and n-body distribution functions in condensed matter. I. Theory. *Phys. Rev. B Condens. Matter*. 52:15122–15134.

65. Filipponi, A., and Di Cicco A. 1995. X-ray-absorption spectroscopy and n-body distribution functions in condensed matter. II. Data analysis and applications. *Phys. Rev. B Condens. Matter*. 52:15135–15149.

66. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–8, 27–8.

67. Obst, S., and H. Bradaczek. 1997. Molecular Dynamics Simulations of Zinc Ions in Water Using CHARMM. *Molecular modeling annual*. 3:224–232.

68. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.

69. Mark, P., and L. Nilsson. 2001. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A*. 105:9954–9960.

70. Huang, J., and A.D. MacKerell Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.

71. Martyna, G.J., D.J. Tobias, and M.L. Klein. 1994. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*. 101:4177–4189.

72. Feller, S.E., Y. Zhang, R.W. Pastor, and B.R. Brooks. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*.

103:4613–4621.

73. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*. 126:014101.

74. Ryckaert, J.-P., G. Ciccotti, and H.J.C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*. 23:327–341.

75. Grubmüller, H., H. Heller, A. Windemuth, and K. Schulten. 1991. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Molecular Simulation*. 6:121–142.

76. Tuckerman, M., B.J. Berne, and G.J. Martyna. 1992. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*. 97:1990–2001.

77. Essmann, U., L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. 1995. A smooth particle mesh Ewald method. *The Journal of Chemical Physics*. 103:8577–8593.

78. Roh, J.H., V.N. Novikov, R.B. Gregory, J.E. Curtis, Z. Chowdhuri, and A.P. Sokolov. 2005. Onsets of anharmonicity in protein dynamics. *Phys. Rev. Lett.* 95:038101.

79. Schirò, G., C. Caronna, F. Natali, and A. Cupane. 2010. Direct evidence of the amino acid side chain and backbone contributions to protein anharmonicity. *J. Am. Chem. Soc.* 132:1371–1376.

80. Wood, K., D.J. Tobias, B. Kessler, F. Gabel, D. Oesterhelt, F.A.A. Mulder, G. Zaccai, and M. Weik. 2010. The low-temperature inflection observed in neutron scattering measurements of proteins is due to methyl rotation: direct evidence using isotope labeling and molecular dynamics simulations. *J. Am. Chem. Soc.* 132:4990–4991.

81. Wood, K., A. Frölich, A. Paciaroni, M. Moulin, M. Härtlein, G. Zaccai, D.J. Tobias, and M. Weik. 2008. Coincidence of dynamical transitions in a soluble protein and its hydration water: direct measurements by neutron scattering and MD simulations. *J. Am. Chem. Soc.* 130:4586–4587.

82. Schirò, G., Y. Fichou, F.-X. Gallat, K. Wood, F. Gabel, M. Moulin, M. Härtlein, M. Heyden, J.-P. Colletier, A. Orecchini, A. Paciaroni, J. Wuttke, D.J. Tobias, and M. Weik. 2015. Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nat. Commun.* 6:6490.

83. LeVine, H. 1999. [18] Quantification of β-sheet amyloid fibril structures with thioflavin T. In: Methods in Enzymology. Academic Press. pp. 274–284.

84. Foderà, V., S. Cataldo, F. Librizzi, B. Pignataro, P. Spiccia, and M. Leone. 2009. Self-organization pathways and spatial heterogeneity in insulin amyloid fibril formation. *J. Phys. Chem. B*. 113:10830–10837.

85. Cauët, E., S. Bogatko, J.H. Weare, J.L. Fulton, G.K. Schenter, and E.J. Bylaska. 2010. Structure and dynamics of the hydration shells of the Zn(2+) ion from ab initio molecular dynamics and combined ab initio and classical molecular dynamics simulations. *J.*

Chem. Phys. 132:194502.

86. Schirò, G., C. Caronna, F. Natali, M.M. Koza, and A. Cupane. 2011. The "Protein Dynamical Transition" Does Not Require the Protein Polypeptide Chain. *J. Phys. Chem. Lett.* 2:2275–2279.

87. Grimaldo, M., F. Roosen-Runge, N. Jalarvo, M. Zamponi, F. Zanini, M. Hennig, F. Zhang, F. Schreiber, and T. Seydel. 2015. High-resolution neutron spectroscopy on protein solution samples. *EPJ Web of Conferences*. 83:02005.

88. Maret, W. 2017. Zinc in Pancreatic Islet Biology, Insulin Sensitivity, and Diabetes. *Prev Nutr Food Sci*. 22:1–8.

89. Fatmi, M.Q., T.S. Hofer, B.R. Randolf, and B.M. Rode. 2005. An extended ab initio QM/MM MD approach to structure and dynamics of Zn(II) in aqueous solution. *J. Chem. Phys.* 123:054514.

90. Thirumalai, D., G. Reddy, and J.E. Straub. 2012. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* 45:83–92.

91. Krone, M.G., L. Hua, P. Soto, R. Zhou, B.J. Berne, and J.-E. Shea. 2008. Role of water in mediating the assembly of Alzheimer amyloid-β Aβ16- 22 protofilaments. *J. Am. Chem. Soc.* 130:11066–11072.

92. Mukherjee, S., P. Chowdhury, and F. Gai. 2009. Effect of dehydration on the aggregation kinetics of two amyloid peptides. *J. Phys. Chem. B*. 113:531–535.

93. Capdevila, D.A., K.A. Edmonds, G.C. Campanello, H. Wu, G. Gonzalez-Gutierrez, and D.P. Giedroc. 2018. Functional Role of Solvent Entropy and Conformational Entropy of Metal Binding in a Dynamically Driven Allosteric System. *J. Am. Chem. Soc.* 140:9108–9119.

94. Bigelow, C.C. 1967. On the average hydrophobicity of proteins and the relation between it and protein structure. *J. Theor. Biol.* 16:187–211.

95. Rezus, Y.L.A., and H.J. Bakker. 2007. Observation of immobilized water molecules around hydrophobic groups. *Phys. Rev. Lett.* 99:148301.

96. Dahanayake, J.N., and K.R. Mitchell-Koch. 2018. Entropy connects water structure and dynamics in protein hydration layer. *Phys. Chem. Chem. Phys.* 20:14765–14777.

97. Nielsen, L., S. Frokjaer, J. Brange, V.N. Uversky, and A.L. Fink. 2001. Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry*. 40:8397–8409.

98. Pessi, A., E. Bianchi, A. Crameri, S. Venturini, A. Tramontano, and M. Sollazzo. 1993. A designed metal-binding protein with a novel fold. *Nature*. 362:367–369.

SUPPLEMENTARY MATERIAL

Zinc determines dynamical properties and aggregation kinetics of human insulin

Kevin Pounot*[1,2], Geoff Grime[3], Alessandro Longo[4], Michaela Zamponi[5], Daria Noferini[5], Viviana Cristiglio[1], Tilo Seydel[1], Elspeth Garman[6], Martin Weik[2], Vito Foderà[7]*, Giorgio Schirò[2]*

[1] Institut Max von Laue – Paul Langevin (ILL), CS 20156, F-38042 Grenoble, France

[2] Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

[3] Ion Beam Centre, Advanced Technology Institute, University of Surrey, Guildford GU2 7XH, Surrey, UK

[4] ISMN-CNR, UOS Palermo, Via Ugo La Malfa, 153, 90146 Palermo, Italy

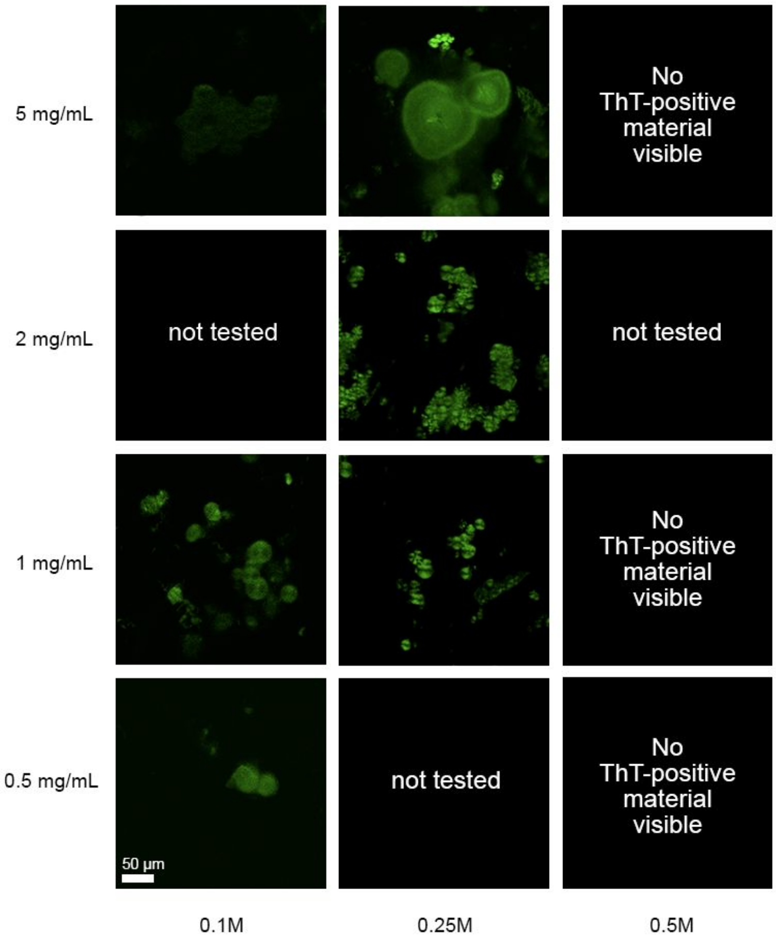
[5] Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), D-85748 Garching, Germany

[6] Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

[7] Department of Pharmacy, Universitetsparken 2, University of Copenhagen, 2100 Copenhagen, Denmark

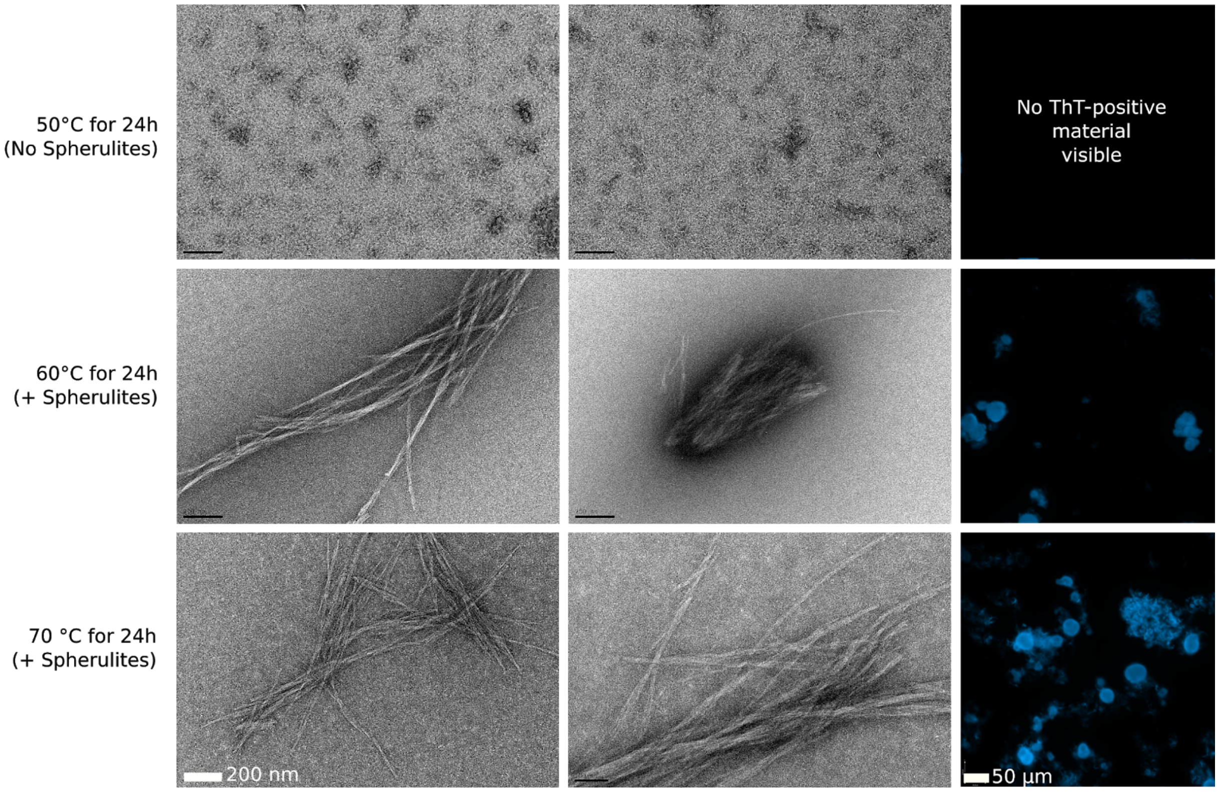
* kevin.pounot@ibs.fr, vito.fodera@sund.ku.dk, giorgio.schiro@ibs.fr

Supplementary Figures

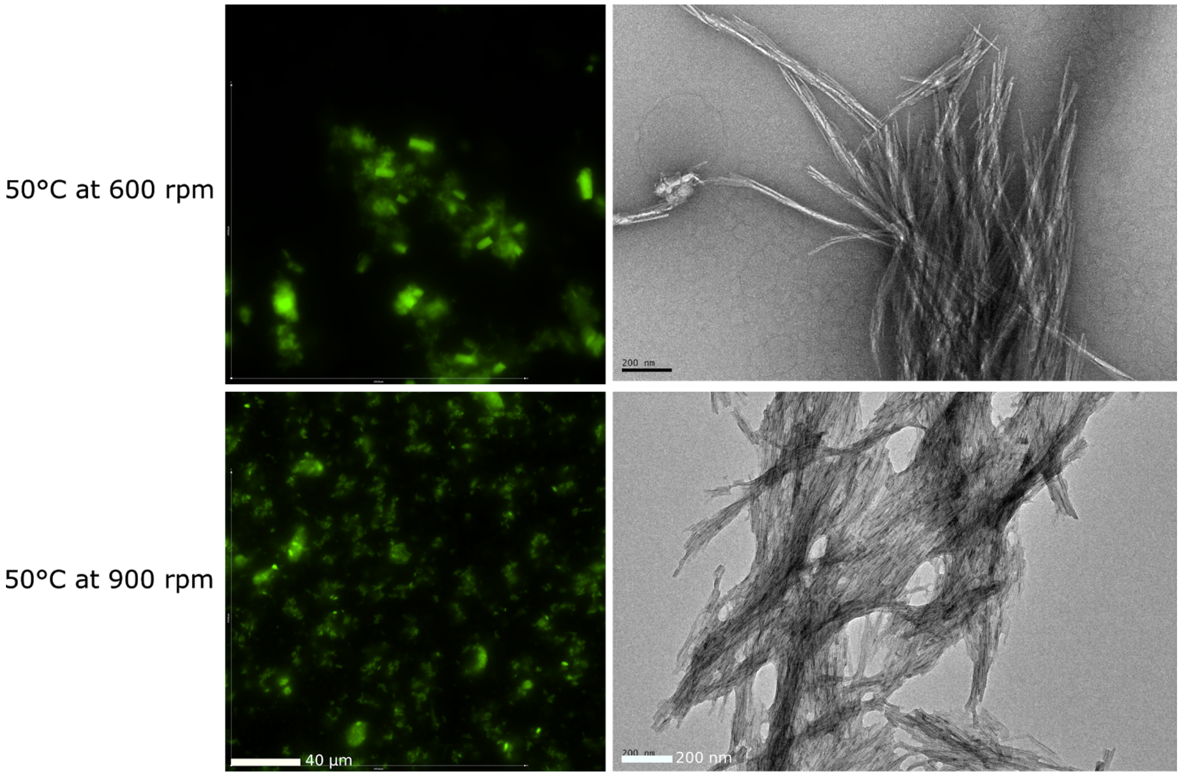


Supplementary Figure 1: Increasing the salt concentration to 0.5 M prevents the formation of spherulites and fibrils

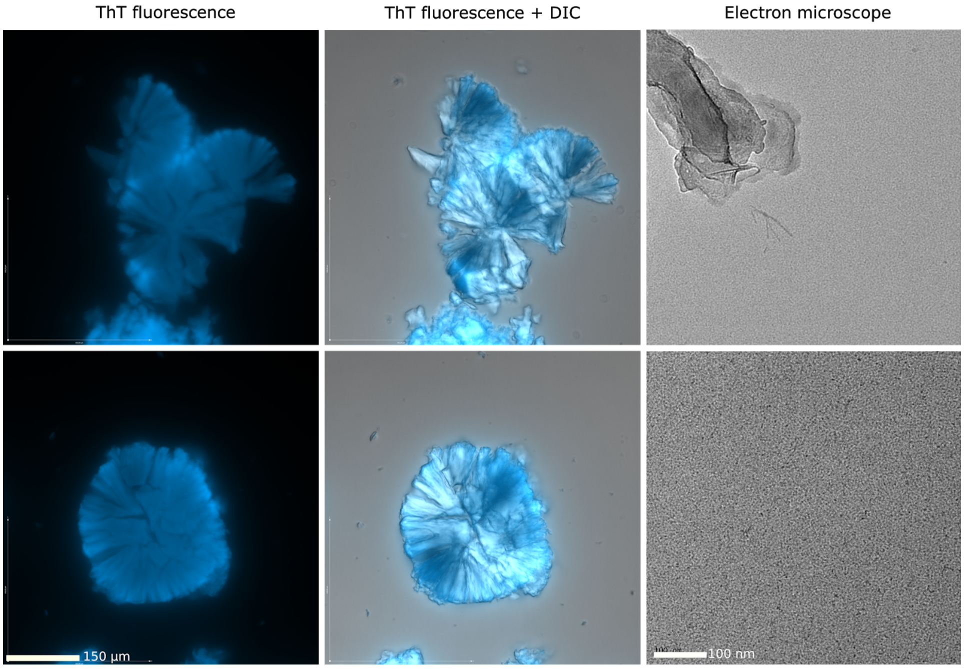
The proteins were dissolved in the corresponding buffer and incubated at 60°C for 24h without shaking. The samples were then observed with optical and electron (data not shown) microscopes after addition of ThT. Protein concentration increases along the rows and NaCl concentration decreases down the columns.



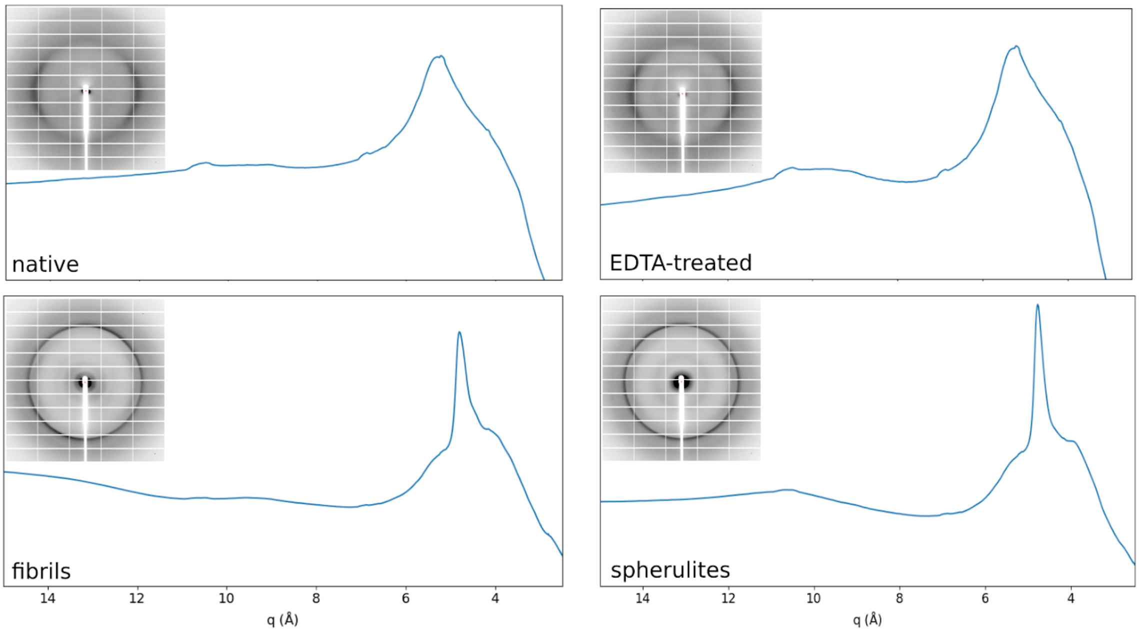
Supplementary Figure 2: No aggregation is visible at 50°C without shaking but it is at higher temperatures. The proteins were dissolved in the corresponding buffer and incubated at 50, 60 or 70°C for 24h without shaking. The samples were then observed with optical and electron microscopes after addition of ThT.



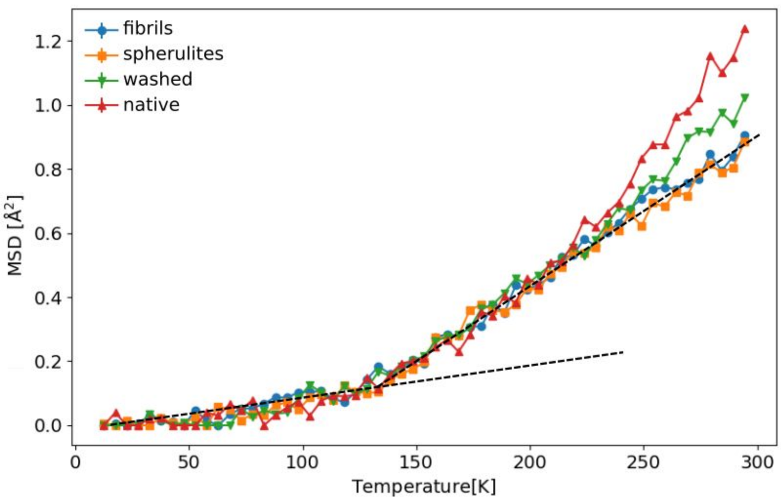
Supplementary Figure 3: Shaking inhibits formation of spherulites but can affect the fibril morphology.
The proteins were dissolved in the corresponding buffer and incubated at 50°C for 48h with the indicated shaking speed. The samples were then observed with optical and electron microscopes after addition of ThT.



Supplementary Figure 4: Pure spherulite sample can be obtained after centrifugation and lyophilization. The proteins were dissolved in the corresponding buffer and incubated at 60 or 70°C for 24h without shaking. The samples were then observed with optical and electron microscopes after addition of ThT.

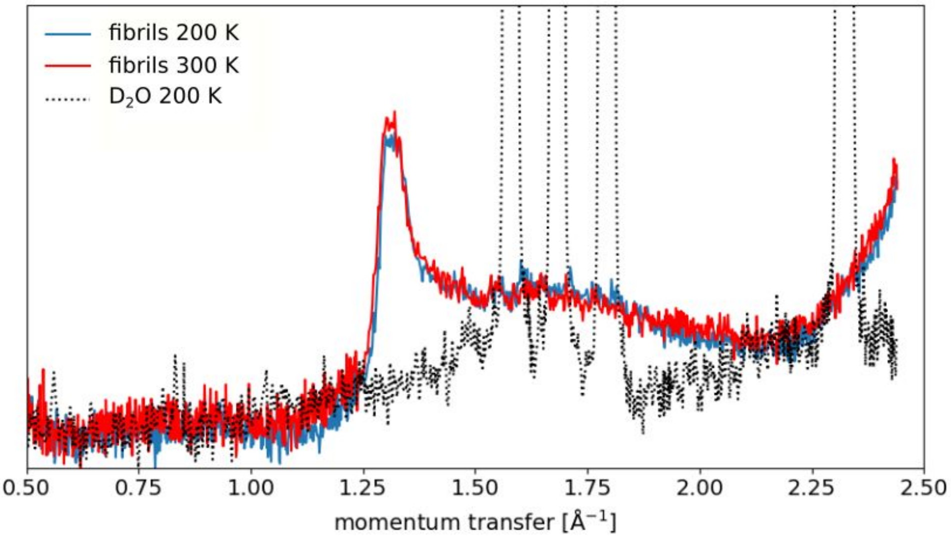


Supplementary Figure 5: Insulin aggregated species present the cross-beta pattern.
A small fraction of the insulin sample powder that was used for neutron scattering was sealed in Mylar capillary for X-ray diffraction on ID30B at the ESRF. An average of 20 images were collected for each sample, then the azimuthal integration was performed using pyFAI library to generate the lower panel plots. The signal at 5.2 Å corresponds to the Mylar capillary.

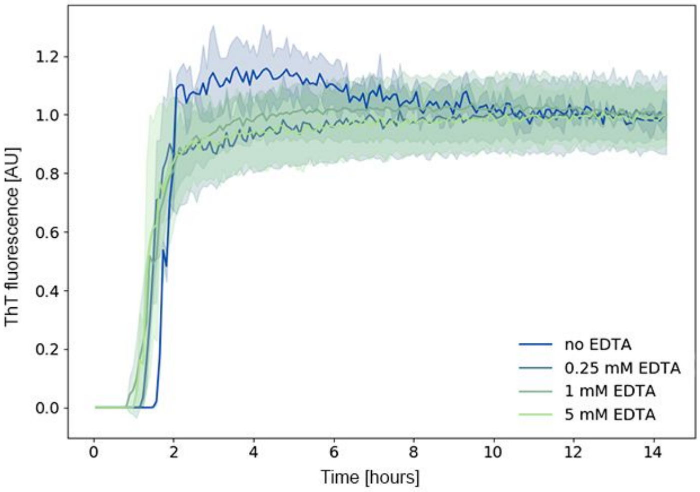


Supplementary Figure 6: Only zinc affects insulin protein dynamics.

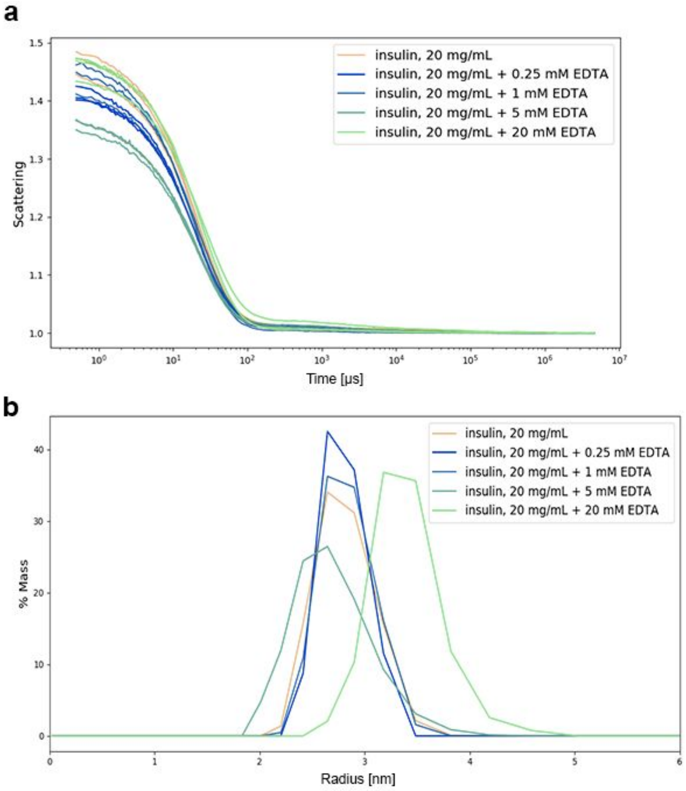
Mean square displacements (MSD), measured as a function of temperature by elastic incoherent neutron scattering with the backscattering spectrometer SPHERES (MLZ, Garching) for native insulin (red triangles), amyloid fibrils (blue circles) and spherulites (orange squares) and for washed insulin sample containing 0.12 zinc atom per insulin molecule (green triangles).



Supplementary Figure 7: Insulin samples do not show any sign of crystalline ice.
The samples, contained in aluminum plate holders, were mounted on the D16 instrument at the ILL. The diffraction patterns were recorded over an angle between 12° and 112.5°, with a corresponding q-range 0.05-2.5 $Å^{-1}$. Each scan was obtained by integrating the 2D signal over 20 minutes.

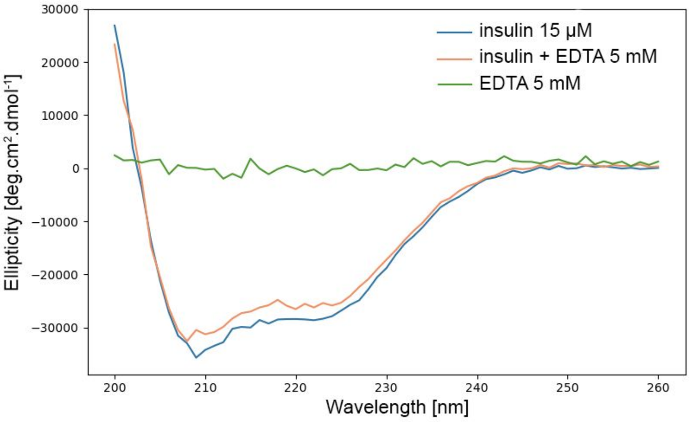


Supplementary Figure 8: Insulin fibril formation under shaking is weakly affected by EDTA.
Insulin aggregation kinetics at 60° C with 300 rpm shaking as probed by Thioflavin T fluorescence as a function of EDTA concentration. The color shaded areas indicates the standard deviation for three independent aggregation processes.

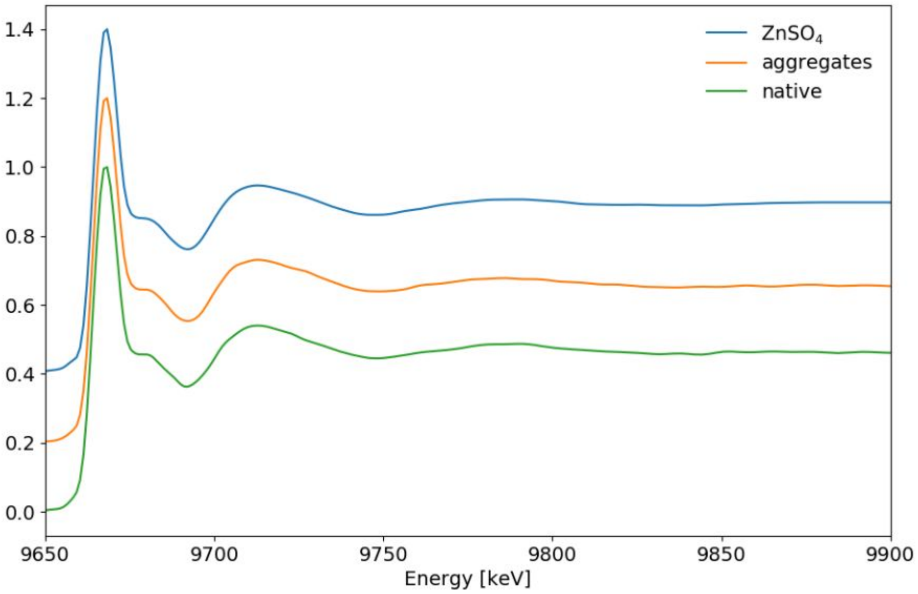


Supplementary Figure 9: Removal of zinc barely affects insulin quaternary structure

Insulin powder was dissolved in 0.25 M NaCl, pH 1.8 then an appropriate amount of filtered 0.26 M EDTA was added for each measurement of dynamic light scattering performed as described in methods. **a** Static scattering signal versus time for sample (measured 3 times, each consisting of 10 five seconds average). **b** Computed radius distribution, only one peak is present with an area accounting for 99-100 % of the signal.



Supplementary Figure 10: Zinc removal has no effect on insulin secondary structure content
Insulin was dissolved in pure water to a final concentration of 2 mg/mL by lowering the pH using sulfuric acid. It was then filtered and diluted again with or without EDTA to reach the indicated concentrations. Data were acquired at 20 nm/s with 6 accumulations on a JASCO spectrophotometer.



Supplementary Figure 11: XANES spectra of $ZnSO_4$, native insulin and insulin aggregates
Zinc sulfate was dissolved in a 0.25 M NaCl, pH 1.8 solution. Similarly, the native insulin was dissolved in the same buffer, and the aggregates were formed by incubating half of the native insulin solution at 60°C for 24 hours. The XANES spectra were recorded as described in methods on DUBBLE beamline at the ESRF in Grenoble, France.

Supplementary Tables

Table 1

	Sulfur		Zinc	
	No. of atoms per protein molecule	LOD†	No. of atoms per protein molecule	LOD†
Native	6*	0.05	1.56	0.03
	6*	0.07	1.46	0.07
Average		**0.06**	**1.50 (±0.17)**	**0.05**
Spherulites	6*	0.01	<0.001	0.001
	6*	0.01	<0.001	0.001
	6*	0.01	<0.001	0.001
Average		**0.01**	**<0.001 (±0.001)**	**0.001**
Washed	6*	0.01	0.101	0.004
	6*	0.01	0.145	0.003
	6*	0.01	0.101	0.003
Average		**0.01**	**0.116 (±0.032)**	**0.003**
EDTA treated	6*	0.013	0.020	0.003
	6*	0.012	0.020	0.003
	6*	0.010	0.023	0.003
Average		**0.012**	**0.021 (±0.005)**	**0.003**
Fibrils	6*	0.01	<0.001	0.001
	6*	0.01	<0.001	0.001
Average		**0.01**	**<0.001 (±0.001)**	**0.001**

* Assumed values from protein sequence
† LOD = Limit of detection (atoms/protein molecule)

Supplementary Table 1: Zinc is hardly detected in aggregated samples
Measurement on µPIXE beamline was performed as described in the text. The table reports the values obtained for the individual measurements carried out on each sample, as well as the averages.

Table 2

	Native insulin	**Aggregated insulin**	**$ZnSO_4$ solution**
N_1	6 ± 0.5	6	6
R_1 [Å]	2.091 ± 0.005	2.083 ± 0.005	2.091 ± 0.005
σ^2_1 [$Å^2$]	5 e-3	3.9e-3	4e-3
ϑ_1 [deg]	150 ± 10 (R_2 = 4.042; σ^2_2 = 1e-2)	-	-
ϑ_2 [deg]	147 ± 5 (R_2 = 4.40; σ^2_2 = 9.7e-3)	-	-

Supplementary Table 2: Parameters of EXAFS data analysis

3.4 Lysozyme aggregation followed by E/IFWS

The amyloid fibrillation is a complex process that consists in multiple protein-protein interactions as well as protein-solvent interactions. Usually, several types of aggregates are formed at the same time [Pansieri et al., 2018] [Skamris et al., 2019], and intermediate species have challenged scientists regarding their isolation for proper studies. However, this is a very needed objective, as these early species are thought to be the ones toxicity originates from [Winner et al., 2011].
To circumvent this problem, time-resolved methods have been employed with good success. For instance, small-angle X-Ray scattering (SAXS) has been used with α-synuclein, along with chemometric data decomposition to obtain pairwise distribution and volume fraction of intermediate oligomers during the course of fibrillation [Herranz-Trillo et al., 2016]. Also, mass spectrometry methods have proven to be well suited to separate different aggregated species [Pansieri et al., 2018]. This method can be applied in a time-resolved manner as well.
Incoherent neutron backscattering spectroscopy provides also a way to monitor fibrillation in real-time. Indeed, the method allows to measure ensemble-averaged self-dynamics, where both center-of-mass motion and internal dynamics can be followed simultaneously. In the pico-second to nano-second time scales probed by the instrument, and the large momentum transfer q involved, a Fickian type diffusion is represented by a simple Lorentzian line shape of width given by $\Gamma = D_s q^2$, where D is an apparent diffusion coefficient containing contributions from rotational and translational diffusion [Roosen-Runge et al., 2011]. The internal motions are usually described by a more complicated dependence on q where $\Gamma(q)$ was chosen here to follow the jump diffusion model [Singwi and Sjölander, 1960]. The strong deviation from Fickian diffusion of internal motions allows to separate both contributions and provides an accurate result in liquid state, as shown recently in work by Beck et al. [Beck et al., 2018]. Recording a full set of QENS spectra usually requires at least two hours, a time that is not in good adequacy with usual amyloid fibrillation time scale. To obtain a sufficient resolution in time, we took advantage of the unique capability of the IN16B backscattering spectrometer at the ILL to perform fixed- window scans (FWS) [Frick et al., 2012]. This provides a way to measure the neutron scattering signal at specific energy offsets, thereby reducing the acquisition time to approximately 20 minutes with 4 different offsets.

Prior to the actual experiment run, several systems were tested to obtain a well characterized and optimized fibrillation protocol that fulfills the measurement requirements. More specifically, the use of shaking - which can be required for some proteins to aggregate or to better control the type of aggregates - is not possible yet on the instrument, the sample concentration should be rather high - higher than 50 mg/mL - and the fibrillation kinetics should be sufficiently slow. An initial test, using insulin, γ-crystallin and lysozyme, showed that the latter was the most promising for the experiment. In addition, the protein is of particular interest for studying amyloid fibrillation. Indeed, the protein is invoved in hereditary lysozyme amyloidosis, a disorder that can affect liver, spleen and kidney and for which no efficient treatment was found yet [Fazili et al., 2016]. Similarly to other systems, several aggregates morphology are usually formed and can have different effects on cells [Booth et al., 1997]. In our case, due to time limitations, we could not test the three known global morphologies of lyzozyme, namely fibrils, spherulites and particulates. Only particulates and fibrils were tested and gave promising results prior to experiment. However, only the particulates measurement could be exploited. The results are presented in the following paper - to be submitted as a communications in JACS - and will be briefly summarized here.

The measurement on the lysozyme sample was performed as follows, the protein powder was dissolved in deuterated buffer and sealed in a cylindrical aluminum neutron cell. A first full QENS spectrum was obtained to characterize the initial state, by maintaining the system at low temperature to prevent aggregation. This shows a stable signal, for which analysis provides a diffusion coefficient in agreement with the presence of small, transient clusters as suggested also by dynamic light scattering (DLS) and molecular dynamics (MD) simulations. Then, the temperature was increased to 363 K to start the formation of particulates. The sparse QENS data resulting from the set of FWS can be analyzed the same way as for initial QENS. Qualitatively, the obtained signal already shows signs of ongoing aggregation, with the signal at low energy offset - slowly diffusing objects - and low q-value - large distances - increasing monotonically during the entire measurement. The fitted model parameters exhibited the expected momentum transfer dependence, and the initial center-of-mass diffusion coefficient is in agreement with the presence of small clusters again. Then, internal dynamics appears steady until the end, while global diffusion decreases with a simple exponential line shape until it reaches a steady value lower than 1 Å/ns corresponding to spherical objects of at least 80 nm. Eventually, a final QENS measurement at 363 K after fibrillation confirmed the previously obtained values, showing that the analysis of FWS data provides sufficiently accurate values. This result shows that lysozyme particulates formation occurs in one step on the time-scale probed in this study, in which no fibril formation appears to be involved. The aggregation process is shown to be controlled first, by a coalescence phenomenon, which results from the pH being close to the isoelectric point of the protein, thereby

decreasing the protein net charge and the colloidal stability. Second, the structural change could also play a role through the energy barrier to overcome to form the cross-β pattern. The single exponential behavior indicates that the structural change occur either much faster than the coalescence, or on the same time-scale.
The following article - under revision with the Journal of Physical Chemistry Letters - presents these results.

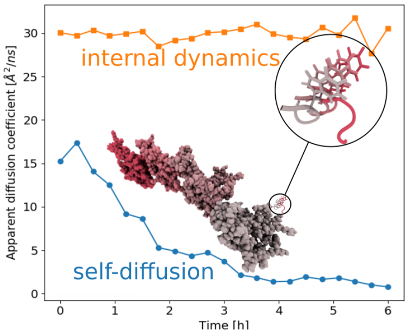
Article 3

Tracking internal and global diffusive dynamics during protein aggregation by high- resolution neutron spectroscopy

Tracking internal and global diffusive dynamics during protein aggregation by high-resolution neutron spectroscopy

Kevin Pounot, Hussein Chaaban, Vito Foderà*, Giorgio Schirò, Martin Weik*, Tilo Seydel*

Abstract



Proteins can misfold and form either amorphous or organized aggregates with different morphologies and features. Aggregates of amyloid nature are pathological hallmarks in so-called protein conformational diseases, including Alzheimer's and Parkinson's. Evidence prevails that the transient early phases of the reaction determine the aggregate morphology and toxicity. As a consequence, real-time monitoring of protein aggregation is of utmost importance. Here, we employed time-resolved neutron backscattering spectroscopy to follow center-of-mass self-diffusion and nano- to picosecond internal dynamics of lysozyme during aggregation into a specific β-sheet rich superstructure, called particulates, formed at the isoelectric point of the protein. Particulate formation is found to be a one-step process and protein internal dynamics to remain unchanged during the entire aggregation process. The time-resolved neutron backscattering spectroscopy approach developed here, in combination with standard kinetics assays, provides a unifying framework in which dynamics and conformational transitions can be related to the different aggregation pathways.

Proteins can convert from their soluble, native form into amorphous, fibril-like aggregates or more complex superstructures[1]. In humans, these aggregates can have deleterious if not fatal effects, as manifested in so-called protein aggregation diseases. These include eye lens cataract caused mainly by amorphous aggregates of lens proteins[2], and Alzheimer's and Parkinson's diseases associated with highly ordered amyloid fibrils formed by the intrinsically disordered proteins tau and α-synuclein, respectively[3,4] or core-shell structures (spherulites)[5]. Such a variety of aggregates poses fundamental questions on the different mechanisms underlying the formation of each species and on how the dynamics-structure-function relationship is altered throughout aggregation. To extend our understanding beyond the initial and end states of aggregation, real-time analysis methods are required. These methods shed light on changes that a protein undergoes in both structure and, equally important for biological activity, dynamics on different length and time scales during the aggregation process.

Several biophysical methods have been developed and applied to follow the time evolution of protein aggregation processes. The most ubiquitous and readily applicable method is monitoring Thioflavin T (ThT) fluorescence, the increase during amyloid-fibril aggregation of which[6] can be kinetically modelled[7]. Fluorescence self-quenching in dye-labelled peptides[8] and fluorescence correlation spectroscopy in combination with amyloid-binding fluorophores[9] have also been employed. In addition, fluorescence-based methods can also be used to monitor aggregation *in vivo*[10]. Furthermore, small-angle X-ray scattering (SAXS) has been used to infer the average size and volume fraction of intermediate oligomers of α-synuclein during fibrillation[11]. Similarly, small-angle neutron scattering (SANS[12]) can be used to study protein amyloidogenesis[13] and aggregation[14]. Electric impedance spectroscopy[15] and dielectric relaxation spectroscopy[16] are also sensitive to amyloid aggregation. Nuclear magnetic resonance (NMR), in contrast, provides exquisite structural and dynamic information of amyloid fibrils[17,18], but cannot follow the entire aggregation process in real-time. Here, we adapted neutron backscattering spectroscopy to provide a complementary and unique time-resolved approach that provides access to the center-of-mass self-diffusion (D_s) of a protein in solution and simultaneously to its internal diffusive dynamics (D_i)[19] during aggregation.

Incoherent neutron spectroscopy accesses the ensemble-averaged single-particle self-dynamics and, when applied to biological samples, is mainly sensitive to the scattering from hydrogen atoms[20,21]. This technique simultaneously probes the center-of-mass self-diffusion of the assembly and its internal diffusive dynamics[22,23] without requiring the labeling mandatory for the aforementioned fluorescence-based methods. Moreover, high-resolution neutron backscattering spectroscopy accesses the dynamics on the pico- to nanosecond time and Ångström length scales, thereby permitting to obtain a center-of-mass short-time self-diffusion coefficient[21] that is only sensitive to the hydrodynamic size of the protein as opposed to NMR, SAXS and dynamic light scattering (DLS), for which particle-particle interactions contribute to the signal as well. We employed elastic/inelastic fixed-window scans (E/IFWS) to reduce data acquisition to 20 min per scan as compared to ~2 hours for quasi-elastic neutron scattering (QENS) spectra (see Methods in the Supporting Information (SI)), providing an adequate temporal resolution to follow protein aggregation.

Here, we studied lysozyme aggregation into spherical assemblies called particulates[24]. The formation of these superstructures can by triggered by raising the temperature of a lysozyme solution at a pH close to the protein's isoelectric point (pI) to 90°C[1,24]. A D_2O lysozyme solution (50 mg/mL) at pD 10.5 and 0.1 M NaCl was used. The aggregation process was first monitored in the aluminum cell used for the neutron experiment by ThT fluorescence and its time-dependence showed excellent reproducibility (Figure 1). The particulates are of a similar morphology as the ones formed in a plastic container (figure S1). ThT kinetics indicates that the entire aggregation process reaches a plateau within 6 hours (Figure 1a), not showing any significant lag phase that is generally observed during the formation of amyloid-like spherulites or fibrils[25,26]. Fluorescence imaging (Figure 1a, inset) and SEM-based analysis (Figure 1b) confirm the presence of spherical particles with a diameter distribution peaked at around 600-700 nm (Figure 1c). Structural features in the protein before and after aggregation were investigated by Fourier-transformed infrared spectroscopy (FTIR, Figure S2) and X-ray powder diffraction (XPD, Figure S3). The FTIR spectrum shows a peak centered at 1650 cm^{-1} for native lysozyme, while the particulates are characterized by a peak at around 1625 cm^{-1} (Figure S2), as expected for a β-sheet rich sample in D_2O[27]. In addition, XPD on lysozyme particulates showed amyloid-like features[28], with a powder ring at 4.7 Å and a faint, broad signal at around 9.5 Å (Figure S3), in agreement with a previous report[29]. These ThT fluorescence, FTIR and XPD results confirm that lysozyme particulates are enriched in β-sheets, yet it remains unclear whether the latter are of purely amyloid nature or not[15].

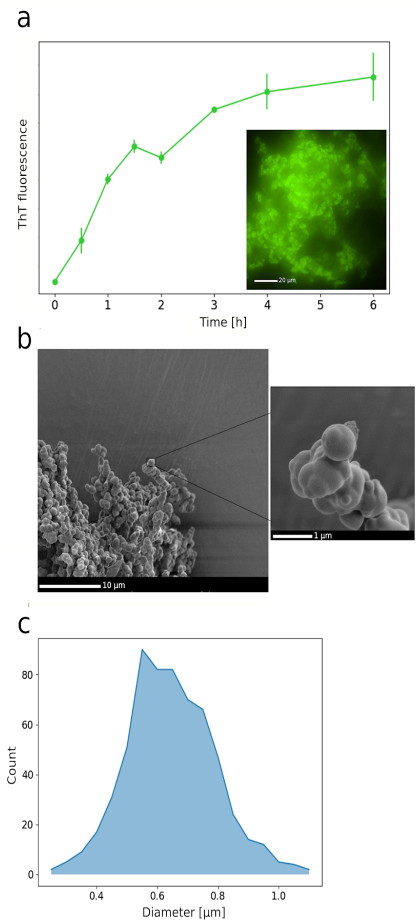


Figure 1. Lysozyme particulates reproducibly form within 6 hrs. a The average of three independent ThT fluorescence kinetics is plotted as green circles. The error bars represent the standard deviation. The inset shows the ThT fluorescent particulates, extracted from the aluminum sample holder after the neutron scattering experiment. **b** SEM micrographs of particulates formed in D_2O under the same conditions as for ThT kinetics at 90°C in the presence of 0.1 M NaCl. The SEM images confirmed the spherical morphology that was observed with fluorescence microscopy. **c** Size distribution of lysozyme particulates as measured in SEM micrographs.

Neutron spectroscopy experiments were carried out at the IN16B backscattering spectrometer[30,31] of the Institut Laue Langevin (ILL) in Grenoble, France. To characterize the initial state, we measured quasi-elastic neutron scattering (QENS) spectra at 7°C (see Methods in SI). Fitting scattering intensities (see Methods in SI) in a global-fit approach yielded an experimental apparent center-of-mass self-diffusion coefficient $D_s^{(exp)}$ = 3 Å^2/ns (Figure S4) to which both translational and rotational diffusive motions contribute. This experimental $D_s^{(exp)}$ is smaller than the corresponding theoretical $D_s^{(theo)}$ = 5.6 Å/ns of monomeric lysozyme estimated as described in Methods in SI at the same protein concentration (50 mg/mL). This difference indicates the presence of oligomers at 7°C prior to triggering aggregation (Figure S5a), consistent with earlier work on lysozyme[32,33] and with the absence of electrostatic repulsion among single protein molecules under our experimental conditions. The latter naturally favor oligomerization, which is the early step for particulate formation The presence of oligomers was also confirmed using DLS (see Methods and Figure S6), with faster translational motions observed with DLS being consistent with a picture of transient clusters reported earlier[33] (see SI). The DLS data contains a larger contribution from small oligomers and monomers being exchanged between oligomers on the µs time scale.

After the QENS measurement at 7°C, aggregation was triggered by increasing the temperature to 90°C within approximately 30 min (Figure S8). Based on ThT kinetics data (Figure 1a), a sequence of E/IFWS was acquired for four energy offsets within 20 min and repeated during 6 hrs to cover the entire aggregation process (see SI/Methods). After data processing (see SI/Methods), we focused on the momentum transfer interval ($0.4 < q < 1.6$ Å^{-1}) where the signal/noise ratio was acceptable (Figure 2). Qualitatively, the signal changes with time as follows. At low q, the scattering signal increases at energy offsets of 0 and 0.6 µeV, remains constant at 1.5 µeV and decreases at 3 µeV. These model-free observations suggest a decrease of D_s in the sample as a function of time, which we attribute to a growth of the average size of lysozyme particles (cf. SI section 1.10 and figure S11). To obtain quantitative insight into D_s and into internal diffusive dynamics (diffusion coefficient D_i) of lysozyme, we used the same fitting procedure and model (equation 5 in Methods) as for the full QENS dataset recorded at 7°C. The fit yields D_i = 30 Å^2/ns for protein-internal diffusive motions on the ns – ps timescale that is constant with time over the entire aggregation process (Figure 3) and in reasonable agreement with values found for other proteins[34]. To the contrary, $D_s^{(exp)}$ decreases from ~16 Å^2/ns at the beginning of the aggregation process to ~1 Å^2/ns after 6 hrs (Figure 3). To determine whether the oligomers observed at 7° C are still present at 90° C at the beginning of the aggregation process, we estimated $D_s^{(theo)}$ of monomeric lysozyme (cf. SI/Methods). We find $D_s^{(theo)}$ = 30 Å^2/ns for unfolded monomeric lysozyme (Figure S5b), whilst we obtain $D_s^{(exp)}$ = 16 Å^2/ns from QENS at the beginning of the aggregation process (Figure 3). This difference shows that the oligomers observed at 7°C are likely still present when aggregation starts at 90°C. Six hours after aggregation was triggered, full QENS spectra were again recorded during 120 min at 90° C. Analysis of these spectra yielded $D_s^{(exp)}$ = 1 Å^2/ns (Figure S7), confirming the reliability of the results obtained with the E/IFWS sequence recorded in only 20 min. Interestingly, the temporal evolution of $D_s^{(exp)}$ well correlates with the inverse of the ThT signal, indicating that the change in the $D_s^{(exp)}$ is entirely determined by the occurrence of massive aggregation events (Figure 3).

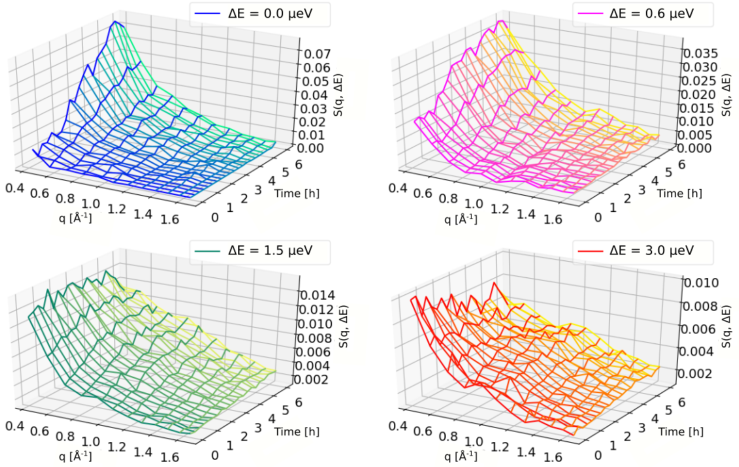


Figure 2. The elastic/inelastic fixed window (E/IFWS) data show qualitative signs of lysozyme aggregation. The neutron scattering signal $S(q,\omega)$ (dots) is shown as a function of time and momentum transfer q for the four energy offsets ω=0 (blue), 0.6 (pink), 1.5 (green), 3.0 (red) µeV measured, in one panel each. The lines connecting the dots are guides to the eye.

The center-of-mass diffusion coefficient $D_s^{(exp)}$ obtained from neutron scattering follows a single exponential decay (Figure 3), implying lysozyme particulate formation is a one-step process. In computational studies on amyloid β-protein[35] and β-amyloid (16-22) peptide[36] aggregation, it was found that cross-β formation is substantially slower than oligomer formation during the aggregation process. In the case of lysozyme, oligomers are already present in solution before aggregation, thereby suggesting that association between oligomers would give rise to the immediate formation of the particulates as detected by the increase of the ThT signal (Figure 1). Moreover, the single-exponential decrease of $D_s^{(exp)}$ and its correlation with the inverse of the ThT signal (reflecting particulate growth), suggests that cross-β formation of lysozyme occurs either substantially faster or on the same time-scale than particulate growth. For a nucleation-dependent[37] process, or if oligomeric species persisted during a plateau phase, one would expect $D_s^{(exp)}$ to exhibit a more complex behavior[38]. The observation of the single exponential time-dependence of $D_s^{(exp)}$ is not surprising if one considers that, under our experimental conditions, *i.e.* the pD being close to the pI of the protein, the absence of a net charge on the protein reduces the protein-protein repulsion, thereby favoring coalescence phenomena between single protein molecules as the driving mechanism for aggregation[39,40].

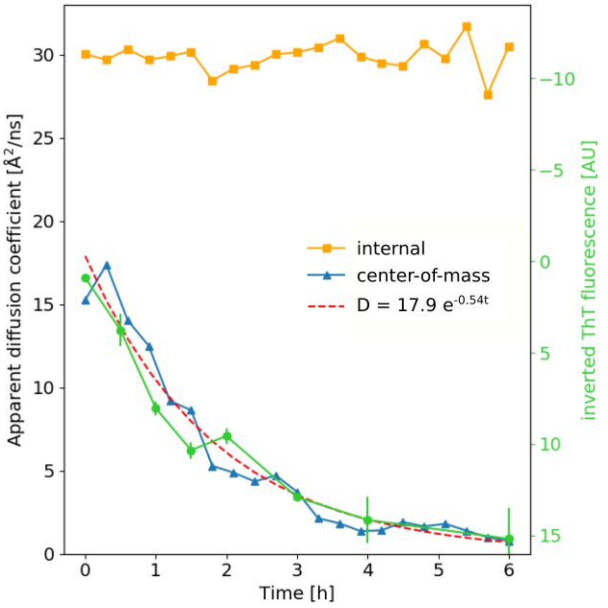


Figure 3. A single exponential law describes both ThT fluorescence and center-of-mass diffusion kinetics. The time evolution of the center-of-mass diffusion coefficient determined from neutron scattering ($D_s^{(exp)}$, blue triangles) follows the one of ThT fluorescence (green circles, inverted with respect to figure 1). The time dependence of $D_s^{(exp)}$ was fitted (red dashed line) using a single exponential. The orange squares represent the protein internal diffusion D_i.

Besides providing access to the center-of-mass diffusion coefficient, time-resolved E/IFWS allowed us to determine D_i. Surprisingly, D_i - corresponding mainly to protein backbone fluctuations in the range captured by our experiment[34] - remains constant throughout the aggregation process (Figure 3), suggesting that lysozyme does not experience significant changes in backbone fluctuations or hydration properties[20] when it is integrated into the growing particulate, consistent with observations made on denatured proteins compared to the native state[34], or with tau amyloid fibrils compared to the monomeric state[41], but in contrast to results reported for Concanavalin A[42] and α-synuclein[43]. Possibly, time-of-flight neutron spectroscopy that probes a broader energy range than neutron backscattering might uncover dynamical differences between the initial monomeric protein and the final particulates outside the time window accessed in the present study.

We hereby established the application of high-resolution neutron backscattering spectroscopy, in particular the E/IFWS to simultaneously follow the time evolution of center-of-mass diffusion and of protein-internal dynamics of lysozyme during its aggregation into particulate superstructures. We demonstrate that lysozyme internal dynamics remain unchanged during particulates formation.

E/IFWS constitute a crucial addition to the toolbox of experimental methods that monitor dynamical changes of proteins during the formation of amorphous and fibrillar aggregates. Indirectly, these dynamical changes also inform on structural changes, since $D_s^{(exp)}$ reflects the hydrodynamic size of the aggregates. As a future outlook, Raman scattering and dielectric spectroscopy can potentially be measured simultaneously during E/IFWS acquisition[44,45]. E/IFWS can simultaneously provide unique insight into internal dynamical changes during the aggregation. The approach presented here will be valuable to study aggregation of intrinsically disordered proteins involved in neurodegenerative diseases, such as alpha-synuclein[26] and tau, provided aggregation can be triggered *in situ* on the neutron spectrometer. In particular, the internal dynamics of supposedly toxic intermediate species that transiently form along the aggregation pathway can be determined[46,47].

Acknowledgments. This work was partially carried out at platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The electron microscopy facility is supported by the Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER, the Centre National de la Recherche Scientifique (CNRS), the CEA, the University of Grenoble, EMBL, and the GIS-Infrastructures en Biologie Sante et Agronomie (IBISA). V.F. and H.C. acknowledge VILLUM FONDEN for the Villum Young Investigator Grant "Protein Superstructures as Smart Biomaterials (ProSmart)" 2018-2023 (project number: 19175). We thank Daphna Fenel, Christine Moriscot and Guy Schoehn for having recorded electron microscopy data. The ILL is acknowledged for the allocation of neutron scattering beam time. We acknowledge the Core Facility for Integrated Microscopy, Faculty of Health and Medical Sciences, University of Copenhagen. KP was supported by a fellowship co-funded by the ILL Graduate School and the CEA IRTELIS program.

Supporting Information: Experimental methods as well as FTIR, XPD, initial and final state characterization and fitting procedure robustness assessment.
The neutron data are permanently curated by the ILL and accessible via
http://dx.doi.org/10.5291/ILL-DATA.8-04-811 .

AUTHOR INFORMATION
Corresponding Authors
Tilo Seydel - Email: seydel@ill.eu
Martin Weik - Email: weik@ibs.fr
Vito Foderà - Email: vito.fodera@sund.ku.dk

Authors
Kevin Pounot - *Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France; Institut Laue Langevin, 71, avenue des Martyrs - CS 20156, 38042 Grenoble cedex 9 – France;*
Hussein Chaaban - *Department of Pharmacy, Universitetsparken 2, University of Copenhagen, 2100 Copenhagen, Denmark;*
Vito Foderà - *Department of Pharmacy, Universitetsparken 2, University of Copenhagen, 2100 Copenhagen, Denmark;*
Giorgio Schirò - *Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France;*
Martin Weik - *Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France*
Tilo Seydel - *Institut Max von Laue – Paul Langevin, 71, avenue des Martyrs - CS 20156, 38042 Grenoble cedex 9 – France*

Authors contribution

H.C., V.F. defined and optimized lysozyme particulates formation. T.S. and K.P. designed the neutron scattering experiment. H.C., G.S., V.F. and K.P. performed the sample characterization experiments using ThT fluorescence kinetics, optical and electron microscopies, FTIR, DLS, XPD and MD simulations. T.S. and K.P. analyzed the neutron scattering data. M.W., T.S., and K.P. wrote the paper with input from V.F. and G.S.

Note

The authors declare no competing financial interest.

REFERENCES

(1) Vetri, V.; Foderà, V. The Route to Protein Aggregate Superstructures: Particulates and Amyloid-like Spherulites. *FEBS Lett.* **2015**, *589* (19), 2448–2463. https://doi.org/10.1016/j.febslet.2015.07.006.

(2) Clark, R.; Zigman, S.; Lerman, S. Studies on the Structural Proteins of the Human Lens. *Exp. Eye Res.* **1969**, *8* (2), 172–182. https://doi.org/10.1016/S0014-4835(69)80029-1.

(3) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75* (1), 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901.

(4) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27–68. https://doi.org/10.1146/annurev-biochem-061516-045115.

(5) Exley, C.; House, E.; Collingwood, J.; Davidson, M.; Cannon, D.; Donald, A. Spherulites of Aβ 42 in Vitro and in Alzheimer's Disease. *Nat. Preced.* **2009**. https://doi.org/10.1038/npre.2009.3982.1.

(6) Collins, S. R.; Douglass, A.; Vale, R. D.; Weissman, J. S. Mechanism of Prion Propagation: Amyloid Growth Occurs by Monomer Addition. *PLoS Biol.* **2004**, *2* (10). https://doi.org/10.1371/journal.pbio.0020321.

(7) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **2009**, *326* (5959), 1533–1537. https://doi.org/10.1126/science.1178250.

(8) D. Quinn, S.; A. Dalgarno, P.; T. Cameron, R.; J. Hedley, G.; Hacker, C.; M. Lucocq, J.; S. Baillie, G.; W. Samuel, I. D.; Carlos Penedo, J. Real-Time Probing of β-Amyloid Self-Assembly and Inhibition Using Fluorescence Self-Quenching between Neighbouring Dyes. *Mol. Biosyst.* **2014**, *10* (1), 34–44. https://doi.org/10.1039/C3MB70272C.

(9) Guan, Y.; Cao, K. J.; Cantlon, A.; Elbel, K.; Theodorakis, E. A.; Walsh, D. M.; Yang, J.; Shah, J. V. Real-Time Monitoring of Alzheimer's-Related Amyloid Aggregation via Probe Enhancement–Fluorescence Correlation Spectroscopy. *ACS Chem. Neurosci.* **2015**, *6* (9), 1503–1508. https://doi.org/10.1021/acschemneuro.5b00176.

(10) Villar-Piqué, A.; Espargaró, A.; Ventura, S.; Sabate, R. In Vivo Amyloid Aggregation Kinetics Tracked by Time-Lapse Confocal Microscopy in Real-Time. *Biotechnol. J.* **2016**, *11* (1), 172–177. https://doi.org/10.1002/biot.201500252.

(11) Herranz-Trillo, F.; Groenning, M.; van Maarschalkerweerd, A.; Tauler, R.; Vestergaard, B.; Bernadó, P. Structural Analysis of Multi-Component Amyloid Systems by Chemometric SAXS Data Decomposition. *Structure* **2016**, 1–11. https://doi.org/10.1016/j.str.2016.10.013.

(12) Mahieu, E.; Gabel, F. Biological Small-Angle Neutron Scattering: Recent Results and Development. *Acta Crystallogr. Sect. D* **2018**, *74* (8), 715–726. https://doi.org/10.1107/S2059798318005016.

(13) Ricci, C.; Spinozzi, F.; Ortore, P. M. and M. G. Protein Amyloidogenesis Investigated by Small Angle Scattering http://www.eurekaselect.com/142331/article (accessed Jun 19, 2020).

(14) Anunciado, D.; Rai, D. K.; Qian, S.; Urban, V.; O'Neill, H. Small-Angle Neutron Scattering

Reveals the Assembly of Alpha-Synuclein in Lipid Membranes. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2015**, *1854* (12), 1881–1889. https://doi.org/10.1016/j.bbapap.2015.08.009.

(15) da Silva, R. R.; de Lima, S. V.; de Oliveira, H. P.; de Melo, C. P.; Frías, I. A. M.; Oliveira, M. D. L.; Andrade, C. A. S. Real-Time Monitoring of Amyloid Fibrillation by Electrical Impedance Spectroscopy. *Colloids Surf. B Biointerfaces* **2017**, *160*, 724–731. https://doi.org/10.1016/j.colsurfb.2017.10.010.

(16) Choi, Y.; Hong, S.; Kang, T.; Lee, L. P. Noninvasive Real-Time Monitoring of Amyloid-β Fibrillization via Simultaneous Label-Free Dielectric Relaxation Spectroscopy and Dark-Field Imaging. *J. Phys. Chem. C* **2009**, *113* (33), 14587–14590. https://doi.org/10.1021/jp904157j.

(17) Tuttle, M. D.; Comellas, G.; Nieuwkoop, A. J.; Covell, D. J.; Berthold, D. A.; Kloepper, K. D.; Courtney, J. M.; Kim, J. K.; Barclay, A. M.; Kendall, A.; Wan, W.; Stubbs, G.; Schwieters, C. D.; Lee, V. M. Y.; George, J. M.; Rienstra, C. M. Solid-State NMR Structure of a Pathogenic Fibril of Full-Length Human α-Synuclein. *Nat. Struct. Mol. Biol.* **2016**, *23* (February), 1–9. https://doi.org/10.1038/nsmb.3194.

(18) Iwata, K.; Fujiwara, T.; Matsuki, Y.; Akutsu, H.; Takahashi, S.; Naiki, H.; Goto, Y. 3D Structure of Amyloid Protofilaments of Beta2-Microglobulin Fragment Probed by Solid-State NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (48), 18119–18124. https://doi.org/10.1073/pnas.0607180103.

(19) Beck, C.; Grimaldo, M.; Roosen-Runge, F.; Braun, M. K.; Zhang, F.; Schreiber, F.; Seydel, T. Nanosecond Tracer Diffusion as a Probe of the Solution Structure and Molecular Mobility of Protein Assemblies: The Case of Ovalbumin. *J. Phys. Chem. B* **2018**, acs.jpcb.8b04349. https://doi.org/10.1021/acs.jpcb.8b04349.

(20) Gabel, F.; Bicout, D.; Lehnert, U.; Tehei, M.; Weik, M.; Zaccai, G. Protein Dynamics Studied by Neutron Scattering. *Q Rev Biophys* **2002**, *35* (4), 327–367. https://doi.org/10.1017/S0033583502003840.

(21) Grimaldo, M.; Roosen-Runge, F.; Zhang, F.; Schreiber, F.; Seydel, T. Dynamics of Proteins in Solution. *Q. Rev. Biophys.* **2019**, *52*. https://doi.org/10.1017/S0033583519000027.

(22) Braun, M. K.; Grimaldo, M.; Roosen-Runge, F.; Hoffmann, I.; Czakkel, O.; Sztucki, M.; Zhang, F.; Schreiber, F.; Seydel, T. Crowding-Controlled Cluster Size in Concentrated Aqueous Protein Solutions: Structure, Self- and Collective Diffusion. *J. Phys. Chem. Lett.* **2017**, *8* (12), 2590–2596. https://doi.org/10.1021/acs.jpclett.7b00658.

(23) Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R. M. J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. Protein Self-Diffusion in Crowded Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (29), 11815–20. https://doi.org/10.1073/pnas.1107287108.

(24) Krebs, M. R. H.; Devlin, G. L.; Donald, A. M. Protein Particulates: Another Generic Form of Protein Aggregation? *Biophys. J.* **2007**, *92* (4), 1336–1342.

(25) Vetri, V.; Piccirilli, F.; Krausser, J.; Buscarino, G.; Łapińska, U.; Vestergaard, B.; Zaccone, A.; Foderà, V. Ethanol Controls the Self-Assembly and Mesoscopic Properties of Human Insulin Amyloid Spherulites. *J. Phys. Chem. B* **2018**, *122* (12), 3101–3112. https://doi.org/10.1021/acs.jpcb.8b01779.

(26) Skamris, T.; Marasini, C.; Madsen, K. L.; Foderà, V.; Vestergaard, B. Early Stage Alpha-Synuclein Amyloid Fibrils Are Reservoirs of Membrane-Binding Species. *Sci. Rep.* **2019**, *9* (1), 1733. https://doi.org/10.1038/s41598-018-38271-2.

(27) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta BBA - Bioenerg.* **2007**, *1767* (9), 1073–1101. https://doi.org/10.1016/j.bbabio.2007.06.004.

(28) Eanes, E. D.; Glenner, G. G. X-Ray Diffraction Studies on Amyloid Filaments. *J. Histochem. Cytochem.* **1968**, *16* (11), 673–677. https://doi.org/10.1177/16.11.673.

(29) Foderà, V.; Vetri, V.; Wind, T. S.; Noppe, W.; Cornett, C.; Donald, A. M.; Morozova-Roche, L. A.; Vestergaard, B. Observation of the Early Structural Changes Leading to the Formation of Protein Superstructures. *J. Phys. Chem. Lett.* **2014**, *5* (18), 3254–3258.

https://doi.org/10.1021/jz501614e.

(30) Frick, B. The Neutron Backscattering Spectrometer IN16 at ILL—High Energy Resolution with High Intensity and Excellent Signal-to-Noise Ratio. *Neutron News* **2002**, *13* (2), 15–22. https://doi.org/10.1080/10448630208218478.

(31) Frick, B.; Mamontov, E.; Eijck, L. van; Seydel, T. Recent Backscattering Instrument Developments at the ILL and SNS. *Z. Für Phys. Chem.* **2010**, *224* (1–2), 33–60. https://doi.org/10.1524/zpch.2010.6091.

(32) Stradner, A.; Sedgwick, H.; Cardinaux, F.; Poon, W. C. K.; Egelhaaf, S. U.; Schurtenberger, P. Equilibrium Cluster Formation in Concentrated Protein Solutions and Colloids. *Nature* **2004**, *432* (7016), 492–495. https://doi.org/10.1038/nature03109.

(33) Porcar, L.; Falus, P.; Chen, W.-R.; Faraone, A.; Fratini, E.; Hong, K.; Baglioni, P.; Liu, Y. Formation of the Dynamic Clusters in Concentrated Lysozyme Protein Solutions. *J. Phys. Chem. Lett.* **2010**, *1* (1), 126–129. https://doi.org/10.1021/jz900127c.

(34) Grimaldo, M.; Roosen-Runge, F.; Hennig, M.; Zanini, F.; Zhang, F.; Jalarvo, N.; Zamponi, M.; Schreiber, F.; Seydel, T. Hierarchical Molecular Dynamics of Bovine Serum Albumin in Concentrated Aqueous Solution below and above Thermal Denaturation. *Phys. Chem. Chem. Phys.* **2015**, *17* (6), 4645–4655. https://doi.org/10.1039/C4CP04944F.

(35) Barz, B.; Olubiyi, O. O.; Strodel, B. Early Amyloid β-Protein Aggregation Precedes Conformational Change. *Chem. Commun.* **2014**, *50* (40), 5373–5375. https://doi.org/10.1039/C3CC48704K.

(36) Chiricotto, M.; Melchionna, S.; Derreumaux, P.; Sterpone, F. Hydrodynamic Effects on β-Amyloid (16-22) Peptide Aggregation. *J. Chem. Phys.* **2016**, *145* (3), 035102. https://doi.org/10.1063/1.4958323.

(37) Giehm, L.; Svergun, D. I.; Otzen, D. E.; Vestergaard, B. Low-Resolution Structure of a Vesicle Disrupting α-Synuclein Oligomer That Accumulates during Fibrillation. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (8), 3246–51. https://doi.org/10.1073/pnas.1013225108.

(38) Cohen, S. I. A.; Vendruscolo, M.; Welland, M. E.; Dobson, C. M.; Terentjev, E. M.; Knowles, T. P. J. Nucleated Polymerization with Secondary Pathways I. Time Evolution of the Principal Moments. *J. Chem. Phys.* **2011**, *135* (6), 065105. https://doi.org/10.1063/1.3608916.

(39) Chi, E. Y.; Krishnan, S.; Kendrick, B. S.; Chang, B. S.; Carpenter, J. F.; Randolph, T. W. Roles of Conformational Stability and Colloidal Stability in the Aggregation of Recombinant Human Granulocyte Colony-Stimulating Factor. *Protein Sci. Publ. Protein Soc.* **2003**, *12* (5), 903–913.

(40) Nicoud, L.; Owczarz, M.; Arosio, P.; Morbidelli, M. A Multiscale View of Therapeutic Protein Aggregation: A Colloid Science Perspective. *Biotechnol. J.* **2015**, *10* (3), 367–378. https://doi.org/10.1002/biot.201400858.

(41) Fichou, Y.; Schirò, G.; Gallat, F.-X.; Laguri, C.; Moulin, M.; Combet, J.; Zamponi, M.; Härtlein, M.; Picart, C.; Mossou, E.; Lortat-Jacob, H.; Colletier, J.-P.; Tobias, D. J.; Weik, M. Hydration Water Mobility Is Enhanced around Tau Amyloid Fibers. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (20), 6365–6370. https://doi.org/10.1073/pnas.1422824112.

(42) Schiró, G.; Vetri, V.; Frick, B.; Militello, V.; Leone, M.; Cupane, A. Neutron Scattering Reveals Enhanced Protein Dynamics in Concanavalin a Amyloid Fibrils. *J. Phys. Chem. Lett.* **2012**, *3* (8), 992–996. https://doi.org/10.1021/jz300082x.

(43) Fujiwara, S.; Araki, K.; Matsuo, T.; Yagi, H.; Yamada, T.; Shibata, K.; Mochizuki, H. Dynamical Behavior of Human Alpha-Synuclein Studied by Quasielastic Neutron Scattering. *PLoS ONE* **2016**, *11* (4), 1–17. https://doi.org/10.1371/journal.pone.0151447.

(44) Sanz, A.; Hansen, H. W.; Jakobsen, B.; Pedersen, I. H.; Capaccioli, S.; Adrjanowicz, K.; Paluch, M.; Gonthier, J.; Frick, B.; Lelièvre-Berna, E.; Peters, J.; Niss, K. High-Pressure Cell for Simultaneous Dielectric and Neutron Spectroscopy. *Rev. Sci. Instrum.* **2018**, *89* (2), 023904. https://doi.org/10.1063/1.5007021.

(45) Adams, M. A.; Parker, S. F.; Fernandez-Alonso, F.; Cutler, D. J.; Hodges, C.; King, A. Simultaneous Neutron Scattering and Raman Scattering. *Appl. Spectrosc.* **2009**, *63* (7),

727–732. https://doi.org/10.1366/000370209788701107.

(46) Fusco, G.; Chen, S. W.; Williamson, P. T. F.; Cascella, R.; Perni, M.; Jarvis, J. A.; Cecchi, C.; Vendruscolo, M.; Chiti, F.; Cremades, N.; Ying, L.; Dobson, C. M.; Simone, A. D. Structural Basis of Membrane Disruption and Cellular Toxicity by α-Synuclein Oligomers. *Science* **2017**, *358* (6369), 1440–1443. https://doi.org/10.1126/science.aan6160.

(47) Ghag, G.; Bhatt, N.; Cantu, D. V.; Guerrero-Munoz, M. J.; Ellsworth, A.; Sengupta, U.; Kayed, R. Soluble Tau Aggregates, Not Large Fibrils, Are the Toxic Species That Display Seeding and Cross-Seeding Behavior. *Protein Sci.* **2018**, *27* (11), 1901–1909. https://doi.org/10.1002/pro.3499.

Supporting information

Tracking internal and global diffusive dynamics during protein aggregation by high-resolution neutron spectroscopy

Kevin Pounot[1,2], Hussein Chaaban[3], Vito Foderà[3]*, Giorgio Schirò[1], Martin Weik[1]*, Tilo Seydel[2]*

[1]Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France.
[2]Institut Max von Laue - Paul Langevin, 71, avenue des Martyrs, CS 20156, 38042 Grenoble cedex 9, France.
[3]Department of Pharmacy, Universitetsparken 2, University of Copenhagen, 2100 Copenhagen, Denmark

1. Materials and methods

1.1. Lysozyme particulate formation followed by thioflavin T fluorescence kinetics

Lysozyme from chicken egg white (batch: 62970-5G-F, from by Sigma-Aldrich (Germany)) powder was kept at 4°C for storage. The lysozyme powder was dissolved in a D_2O (99.9%) solution at pD 10.5 and 0.1 M NaCl to a final concentration of 50 mg/mL. Subsequently, Thioflavine T (ThT) was added to a final concentration of 20 µM, the solution was sealed in a double-walled aluminum cylindrical sample container (outer diameter 22 mm, gap Δr = 0.15mm) with a 1 mm indium wire and incubated in a water bath at 90°C for either 0.5, 1, 1.5, 2, 3, 4 or 6 hours. After incubation, 300 µL of the solution was transferred in three wells (100 µL in each well) of a 96-well plate and the ThT fluorescence was measured on a BioTek Synergy H4 Hybrid plate-reader. ThT fluorescence was measured using 450 nm light for excitation and 490 nm filter for detection employing bottom optics and a manual gain of 65. The measurements in the three wells were averaged for each incubation time. The result is reported in figure 1b. In addition, the morphology of the particulates formed either in a plastic container or in the aluminum sample holder during the neutron scattering experiment (section 1.9) was observed to be similar as shown in figure S1.

1.2. Scanning Electron Microscopy

The morphology of the lysozyme particulates was studied by Scanning Electron Microscopy (SEM) using a Quanta™ 3D FEG scanning electron microscope (Thermo Fisher Scientific, Hillsboro, OR, US). The lysozyme powder was dissolved in a deuterated solution (D_2O, 99.9%) with 0.1 M NaCl at pD 10.5 to a final concentration of 50 mg/mL. The solution was incubated at 90°C for 90 minutes. Afterwards, the obtained lysozyme particulates were centrifuged at 10.000 rpm, 10 °C for 10 min. The samples were then washed in 1 ml deionized water (H_2O) and re-suspended. The centrifuge step was repeated. Subsequently, the samples were mounted on carbon tapes and sputter-coated with 0.2 nm gold using a Leica EM ACE200 (Leica Microsystems GmbH, Wetzlar, Germany) prior to imaging. The size distribution of the particulates was investigated by quantitative measurement of the size of 500 particulates from different representative SEM images (Main text Figure 1b).

1.3. Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy

Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) was used on native lysozyme and lysozyme particulates in transmission mode to obtain information on the secondary structure. The solid samples were run on a Bomem IR spectrometer (Bomem, Quebec, Canada). The spectra were collected during 256 scans at 2 cm^{-1} resolution in the 4000-600 cm^{-1} region. A double subtraction method was used tailed by the second derivative of 12 points, and a 2-3 point baseline correction in the amide I region. Finally the obtained spectra were area-normalized (Figure S2).

1.4. X-ray powder diffraction

The lysozyme powder was dissolved in deuterated solution (D_2O, 99.9%) and 0.1 M NaCl at pD 10.5 (100 µL) to a final concentration of 50 mg/mL in an Eppendorf tube. The solution was incubated at 90°C for 6 hours. Afterwards, the obtained lysozyme particulates were lyophilized by freezing them at -80°C and then using a Christ Martin lyophilizer operating at 0.01 mbar. A fraction of the powder (0.1 mg) was sealed in a Mylar capillary and mounted on the ID29 beamline of the ESRF synchrotron in Grenoble, France. Diffraction data were collected with a PILATUS 6M detector using the following scheme: 20 images with a total oscillation range of 10° (0.5° per image) and 1 s exposure time per image (100% transmission for 5.42×10^{12} ph/s at λ = 0.98 Å). The obtained diffraction images were averaged and processed using the pyFAI library[1] to perform the azimuthal integration (Figure S3).

1.5.Theoretical apparent center-of-mass diffusion coefficient of monomeric lysozyme

The theoretical diffusion coefficient $D_s^{(theo)}$ of lysozyme (50mg/mL) in D_2O was computed at 7°C and 90°C (Figure S5) for comparison with the experimental coefficient $D_s^{(exp)}$ (Figure S4) obtained from neutron scattering data.

This calculation involves (1) the theoretical diffusion coefficient in the limit of infinite dilution D_0, and (2) concepts from colloid physics to calculate the diffusion coefficient at the given finite protein concentration.

To obtain D_0, HYDROPRO[2] is employed, using the protein structure in the pdb format as input. The pdb structure for lysozyme is published only for ambient temperature (PDB entry code 3IJV used here). For the HYDROPRO calculation at 7°C*, i.e.* at the temperature at which full QENS spectra (Figure S4) were recorded prior to triggering aggregation, this native lysozyme structure (PDB entry code 3IJV) was used directly.

To compute the theoretical apparent center-of-mass diffusion coefficient for the unfolded[3] lysozyme monomer at 90°C, *i.e*. the temperature at which particulate-formation was followed in the neutron spectroscopy experiment, molecular dynamics simulations were employed to unfold the native structure. A starting model was constructed using the same published structure of lysozyme (PDB entry code 3IJV). This structure was solvated employing the VMD solvate plugin with a padding of 15 Å and charges were neutralized using the VMD autoionize plugin with Na and Cl ions[4]. The equilibration was performed using NAMD 2.13[5] with the TIP3 model for water[6] and the CHARMM36 force field[7]. The Nose-Hoover-Langevin piston algorithm maintained a constant pressure[8,9] and the stochastic velocity rescaling algorithm controlled the temperature[10]. Bonds with hydrogen atoms were constrained using the SHAKE algorithm[11], and the Verlet-I/r-RESPA multiple-time step scheme[12,13] was used to integrate the equations of motion, with time steps of 2 fs for long-range nonbonded forces, and 1 fs for short-range bonded forces. Electrostatic interactions were computed using the smooth Particle Mesh-Ewald (PME) sum[14], with a cutoff set to 12 Å , a switching function starting at 10 Å, and a pair list distance of 16 Å was used.

Subsequent to the above steps, the native solvated lysozyme was unfolded using a 2 ns simulation at 327°C in the isothermal-isobaric (NPT) ensemble. Subsequently, a 100 ns equilibration at 90°C was performed, again in the NPT ensemble. After the simulation, the coordinates from six different frames along the simulation trajectory were used to produce pdb files. These files were used as input in the HYDROPRO software[2] to compute rotational and translational diffusion coefficients in the dilute limit.

In the procedure described above, the protonation state of the protein corresponds to a pH of 7 as obtained from CHARMM topology files. A calculation with PROPKA[15,16] gives mean pKa values of 10.1 for tyrosines and 11 for lysines for the unfolded lysozyme, thereby suggesting that less than half of tyrosines and a tenth of lysines are oxidized at pH 10, and that the structure at pH 7 is a reasonable representation of the monomeric lysozyme. The conformational ensemble probed during the 100 ns equilibration is sketched in Figure S5a where 25 structures obtained along the trajectory of 2500 frames are superposed. The radius of gyration of the unfolded lysozyme oscillates between 15 and 20 Å but remains larger than the monomer (Figure S5b). The radius of gyration is in agreement with a previous study on thermal unfolding of lysozyme[17], even though the authors used a pH lower than 7. Six conformers along the trajectory are used to compute the theoretical apparent diffusion coefficient, which yields 30 ± 1 Å²/ns.

1.6. Calculation of the diffusion coefficient at finite volume fraction ($\Phi > 0$)

To obtain the apparent diffusion coefficient observable in QENS (cf. section 1.10), we used an effective hydrodynamic protein volume fraction computed using:

$$\Phi_{\text{eff}} = \Phi \left(\frac{R_h}{R} \right)^3 \tag{1}$$

where $\Phi$is the dry volume fraction computed from the lysozyme specific volume ν = 0.726 mL/g[18] R_h is the hydrodynamic radius obtained from HYDROPRO by inserting the translational diffusion coefficient into the Stokes-Einstein relation, and R the dry protein radius obtained with

$$R = \sqrt[3]{\frac{3}{4\pi} \frac{\nu M_w}{\mathcal{N}_A}} \tag{2}$$

with M_w being the molar weight (14.3kDa) and $\mathcal{N}_A$ the Avogadro number.

HYDROPRO returns the translational and rotational diffusion coefficients in the dilute limit. Subsequently, molecular crowding can be considered using established analytical expressions of scaling factors derived from colloid models for translational motions[19] (eq. 11,12) and rotational motions[20] (eq. 21). Finally, the apparent diffusion coefficient D=D(D_r, D_t) is computed from the thus obtained rotational D_r and translational D_t components using the root finding algorithm scipy.optimize.root to solve the implicit relation[21,22]:

$$\sum_{l=0}^{n} B_l(q) \frac{D_r l(l+1) + (D_t - D)q^2}{[D_r l(l+1) + (D_t + D)q^2]^2} = 0 \tag{3}$$

with n being set to 100 to obtain good convergence and $B_l(q)$ being obtained using the radial hydrogen number density function (relative to the protein center of mass) $\rho(r,n)$:

$$B_l(q) = (2l+1)\int_0^\infty dr\rho(r,n)j_l^2(qr) \tag{4}$$

The Python/C scripts that were used to perform the analysis are available on github (github.com/kpounot/NAMDAnalyzer – ScatDiffusion module).

The resulting theoretical apparent diffusion coefficients as a function of volume fraction for native and unfolded lysozyme, respectively, are reported in figure S5c.

1.7. Dynamic Light Scattering

Lysozyme was prepared at final concentration of 20 mg/mL in 100 µL of a deuterated solution (D_2O, 99.9%) and 0.1 M NaCl, pD 10.5. The solution was filtered using a 20 µm filter and centrifuged at 17000 rpm for 15 minutes prior to measurements. All measurements were carried out on a Wyatt DynaPro NanoStar instrument, using standard cuvettes containing 10 µL of sample. The data were acquired and analyzed using the Dynamics software from Wyatt technology. Each sample showed an integrated intensity of about 700 000 counts and a sum of squared errors lower than 400 for cumulant fit. Instrument limitations, due to the lack of a Peltier element, restricted the accessible temperature range from 25 to 60°C. Dynamic Light Scattering (DLS) experiments were carried out at these two boundary temperature values (Figure S6).

1.8. Comparison of the long-time translational diffusion coefficient D_t obtained from DLS with the apparent short-time diffusion coefficient $D_s^{(exp)}$ obtained from neutron scattering, and the calculated diffusion coefficient $D_s^{(theo)}$

The diffusion coefficient D_t that is obtained from DLS (Figure S6a, by fitting the signal corresponding to the first peak in Figure S6b) reflects the long-time collective translational diffusion, where direct particle-particle interactions (collisions) influence the diffusion. Conversely, the coefficient $D_s^{(exp)}$ extracted from the analysis of the neutron scattering data reflects the short-time self-diffusion, where the time and length scales considered are too small for direct particle-particle interactions to affect the result. In the dilute limit, *i.e.* for a single protein in solution, long-time collective and short-time self diffusion coefficients are equal. The experimental values we obtained for both coefficients at finite volume fractions can be extrapolated to the dilute limit using eq. 20 and eq. 24 in the work by Banchio and Nägele[20] for $D_s^{(exp)}$ and D_t, respectively. In addition, assuming that the lysozyme - monomer or oligomer - is spherical, the result from the DLS measurement at 25°C can be extrapolated to 7°C (temperature of neutron scattering experiment prior to temperature raise) using the Stokes-Einstein relation. We thus extrapolate the DLS result, D_t at 7°C to the dilute limit to obtain $D_{0,t}$ = 3.52 Å²/ns, and we extrapolate the neutron scattering result $D_s^{(exp)}$ to the dilute limit at 7°C to obtain $D_{0,s}^{(exp)}$ = 3.2 Å²/ns. The theoretical diffusion coefficient in the dilute limit can be obtained by taking the value where the blue curve intersects the ordinate in Figure S5a (cf. section 1.6), which corresponds to a zero protein volume fraction. We obtain $D_{0,s}^{(theo)}$ = 6.3 Å²/ns, which indicates that both $D_{0,s}^{(exp)}$ and $D_{0,t}$ are approximately by a factor of two lower than the theoretical expectation for monomers. From this observation, we conclude that our experimental QENS and DLS results indicate the presence of oligomers prior to triggering the aggregation. We note again that the extrapolated experimental coefficient $D_{0,s}^{(exp)}$ from QENS contains contributions from both translational and rotational motions[19] (cf. section 1.6). We also note that, since the aggregates are presumably transient, the different observation time scales of DLS (µs) and QENS (ns), respectively, do not necessarily give fully consistent results on the cluster size (cf. discussion and references in the main part).

An approximation of the diffusion coefficient for the lysozyme solution at 90°C can be obtained also from the experimental value at 7°C. Using the Stokes-Einstein relation with tabled values for the viscosity of D_2O[23], we obtain $363\eta_{280\mathrm{K}}/280\eta_{363\mathrm{K}} = 6.6$ where η_{280K} is the viscosity of D_2O at 280K, which gives $D_s^{(\mathrm{exp\ approx})} = 19.8$ Å²/ns for the folded lysozyme at 90°C. This value is slightly higher than the one obtained experimentally, which can be explained by the fact that lysozyme is unfolded at 90°C, while this value is approximated from a folded structure. Hence, we can assume a similar distribution of oligomers at 7°C and 90°C.

1.9. Neutron scattering experiments

We have used the cold neutron backscattering spectrometer IN16B[24,25] with Si(111) analyzer crystals, corresponding to an elastic wavelength of 6.27 Å. A mechanical Doppler drive was used to set the incident energy, and a phase space transformation (PST) chopper was employed to optimize the neutron flux at the sample position[26]. The lysozyme was dissolved at 50 mg/mL in 1 mL of D_2O (99.9%) and 0.1 M NaCl at pD 10.5 and was sealed in a double-walled cylindrical aluminum sample container (outer diameter 22 mm, gap $\Delta r = 0.15$ mm) and introduced at 7°C into a standard Orange cryofurnace that allowed changing the sample temperature during the experiment. The choice of the instrument was motivated by the high flux available and its unique capability to perform elastic and inelastic fixed-window scans (E/IFWS) that allowed us to acquire data with a reasonable time resolution - around 20 minutes for a full cycle of elastic and three inelastic data points (Figure 2)[26]. A schematic illustration of the concept of E/IFWS is depicted in figure S11.

Quasi-elastic neutron scattering (QENS) spectra were first recorded at 7°C (Figure S4), with the momentum transfer q range $0.19 < q < 1.90$ Å^{-1} explored. Subsequently, the temperature was increased to 90°C (heating time approximately 30 min; temperature vs. time displayed in Figure S8) to trigger the aggregation process, and E/IFWS were acquired at 90°C using the following scheme: 1 min counting time at 0 μeV, 2 min at 0.6 μeV, 6 min at 1.5 μeV, and 6 min at 3 μeV. This sequence was repeated continuously for 6.5 hours. Finally, QENS spectra were recorded at 90°C (Figure S7) to obtain a more accurate measurement of the final state as a control.

No visible corrosion of the employed Al sample container for the lysozyme solution was detected at the end of the neutron experiment. This absence of corrosion is in agreement with published work on the behavior of Al surfaces[27].

The neutron data are permanently curated by the ILL and accessible at http://dx.doi.org/10.5291/ILL-DATA.8-04-811

1.10. Neutron scattering: Basic concepts and details of modeling and data analysis

Basic concepts: Crucially, the analysis of the relatively scarce fixed-window neutron scattering data (E/IFWS. section 1.9 and schematic figure S11) is achieved by a so-called global fit for the data at all momentum transfer values q and energy transfer values ΔE simultaneously. The model contains two Lorentzian contributions (one accounting for the center-of-mass diffusion and one for the protein-internal dynamics), the q-dependence of which is fixed in the model. Moreover, the aqueous solvent contribution in the fits is fixed by results from separate fits and tabled values of corresponding measurements on pure water (D_2O). In this way, the number of free parameters is drastically reduced at the expense of imposing the *q*-dependence, which involves assumptions on the type of diffusion. However, the validity of these assumptions was shown previously in stable-state QENS where full spectra could be recorded[28,29].

It is important to note that the neutron backscattering experiment with its energy resolution of approximately 0.8 µeV FWHM corresponds to an observation time scale of the diffusion in the sample of a few nanoseconds (cf. sections 1.6 and 1.8). On this short observation time scale, the proteins diffuse by less than their radius, and protein-protein collisions are negligible[30]. Therefore, the observed center-of-mass diffusion coefficient of the proteins ${D_s}^{(exp)}$ can be interpreted in terms of a colloidal short-time coefficient. Since the scattering from the proteins with its prevailing hydrogen atoms is mostly incoherent, this coefficient ${D_s}^{(exp)}$ can be identified with the self-diffusion or, synonymously, tracer diffusion coefficient. Due to the large momentum transfers involved in the backscattering experiment, ${D_s}^{(exp)} = {D_s}^{(exp)}(D_{r,} D_t)$ is an implicit function that contains contributions from both rotational ("tumbling") D_r and translational D_t diffusion[30]. Therefore, ${D_s}^{(exp)}$ is also denoted *apparent* or *observable* diffusion coefficient (cf. sections 1.5 and 1.6). The internal diffusive dynamics arising mainly from protein backbone is convoluted with the center-of-mass diffusion in the measured signal. On the observation scale of our experiment, this internal diffusion can be well modeled by the inclusion of an additional Lorentzian function in the model (equation 5 below), as shown in previous work[30]. This additional Lorentzian accounts mainly for the protein backbone fluctuations, while the protein side chains diffuse too fast to be detected within the range of our E/IFWS experiment[28]. The resulting combined model function for our neutron spectra will be explained in the following.

Details of modeling and data analysis: To analyze the neutron scattering data, we assumed that the center-of-mass diffusion follows Fick's law and is, thus, reflected in the incoherent QENS signal by a Lorentzian contribution with the width[21,30] $\Gamma(q) = D_s^{(exp)} q^2$. The validity of this relation was shown in previous work for various protein solutions both at physiological and elevated temperatures[28,29]. The superimposed contribution to the scattering function from the backbone and side-chain motions - that is, internal dynamics - within the protein was assumed to follow a more complicated functional dependence, the so-called jump diffusion model[30]. In brief, it is approximated by an additional Lorentzian contribution, the *q*-dependence of which reads $\gamma(q) = \dfrac{D_i q^2}{1 + D_i q^2 \tau_0}$ where *q* is the momentum transfer,

D_i is the diffusion coefficient associated with internal dynamics and τ_0 is the residence time for atoms between diffusive jumps. The jump-diffusion model has been shown to capture the backbone contribution of the protein internal dynamics reasonably well[30]. In contrast, it does not describe the fast side-chain motions that are beyond the energy range probed by the present experiment.

The additional aqueous solvent contribution to the signal was analyzed by measuring the pure solvent separately. The solvent was then fitted separately using a single Lorentzian lineshape and the result was added to the model used for the protein solution[22]. This approach of fixing the solvent contribution further reduced the number of free parameters in the fit of the protein solution data.

Data reduction, *i.e*. normalization to the incident beam intensity recorded by a so-called monitor device, integration over the regions of interest of the vertically position-sensitive detector tubes, calculation of the energy axis, and centering of the elastic line positions – using a separate vanadium measurement - was performed using Mantid's data reduction module for IN16B[31].

Empty cell subtraction, as well as the fitting of the resolution function, D_2O signal, normalization of the detector efficiency employing the Vanadium signal, and fitting of the fully reduced data were performed using an custom-made python API, with the source code available on github (github.com/kpounot/nPDyn). The resolution function was fitted using both a pseudo-Voigt profile and a double Gaussian model – presented results used pseudo-Voigt profile. The D_2O spectrum was fitted using published linewidths already available from measurements on the time-of-flight spectrometer IN6, following the procedure described previously[21,30].

Hence, the final model reads:

$$S(q, \Delta E) = R(q, \Delta E) \otimes \{\beta[a_0 \mathcal{L}_\Gamma(q, \Delta E) + (1 - a_0)\mathcal{L}_{\Gamma+\gamma}(q, \Delta E)] + \beta_{\mathrm{D_2O}}\mathcal{L}_{\mathrm{D_2O}}\} \tag{5}$$

where $R(q, \Delta E)$ is the resolution function, β accounts for the overall amplitude, a_0 can be identified with the Elastic Incoherent Structure Factor EISF[21,30] and represents the relative contribution of the center-of-mass diffusion Lorentzian, $\mathcal{L}_\Gamma$ is the center-of-mass diffusion Lorentzian, the Lorentzian with the width $\mathcal{L}_{\Gamma+\gamma}$ is the convolution between the Lorentzians for center-of-mass diffusion and internal dynamics, and $\beta_{\mathrm{D_2O}}\mathcal{L}_{\mathrm{D_2O}}$ is the D_2O signal, which is fixed in the fits as explained above. The convolution with the resolution function is performed analytically. The convolution of the Lorentzians representing the resolution and the model is straightforward and simply gives a Lorentzian of linewidth being the sum of the two convoluted Lorentzians. The convolution of the Gaussian representing the resolution function and the Lorentzians in the model is analytically given by the Voigt function[32].

To find the global minimum of the fit, *i.e.* to remove the dependence on the starting parameters ("initial guess") during the fitting procedure, we used the *scipy basinhopping* algorithm, which performs multiple iterations of the following scheme: (1) random shift of the starting parameter values, (2) use of a minimization algorithm to find the local minimum, (3) compare the local minimum with the previously found minimum by the Metropolis criterion to accept or reject the parameters fitted in the iteration[33]. The basinhopping procedure thus allows to better explore the solution landscape and maximizes the chance to find the true global minimum. In the case of the global fit, *i.e.* fitting all momentum transfer values at once, the aforementioned formula for the Lorentzian widths Γ and γ were used, with an explicit dependence on momentum transfer *q*. For momentum transfer-wise fits, where each momentum transfer value is fitted individually, the two Lorentzian widths become scalar parameters.

In addition, the robustness of the fitting procedure for E/IFWS was assessed by performing momentum transfer-wise fitting of data (Figure S9), which shows that the information about center-of-mass diffusion and protein-internal dynamics is already present at each momentum-transfer. Moreover, we observed that the subtraction of the empty cell signal suppresses most of the elastic peak in the data, and might lead to a bias of the result when the elastic point is lower than the others. We thus performed a fit of a subset of the data which does not contain the EFWS (Figure S10), yielding again the same result. The errors at the energy transfer $\Delta E = 0.6$ µeV are larger than at higher energy transfers, such that the fitting procedure favors the latter due to the error weighting.

2. Supplementary figures

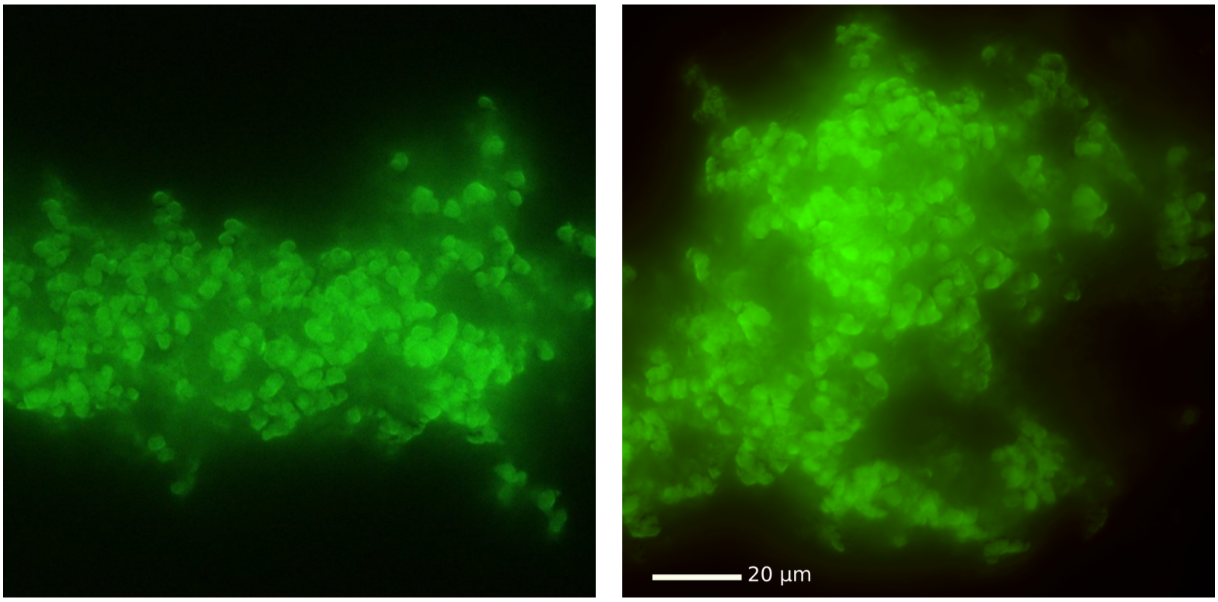


Figure S1. Lysozyme particulates formed in plastic and aluminum containers are similar. Observation of particulates by ThT fluorescence microscopy. Particulates were formed either in the aluminum sample holder during the neutron scattering experiment **(right),** or, as a control, in a plastic container **(left),** showing the same morphology. The scale bar holds for both images. We note that the samples for the experiments in the figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

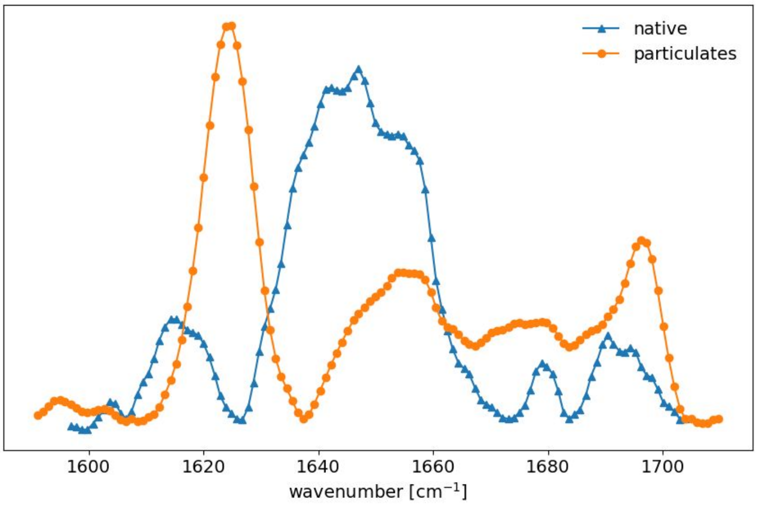


Figure S2. Lysozyme undergoes substantial secondary structure change upon particulates formation. ATR-FTIR spectra (section 1.3) for native lysozyme (blue) and lysozyme particulates (orange), respectively. Native lysozyme is mainly alpha-helical in accordance with other studies[34]. However a transition to β-sheet is observed in the particulate samples, which was also shown by Krebs et al.[35]. We note that the samples for the experiments in the figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

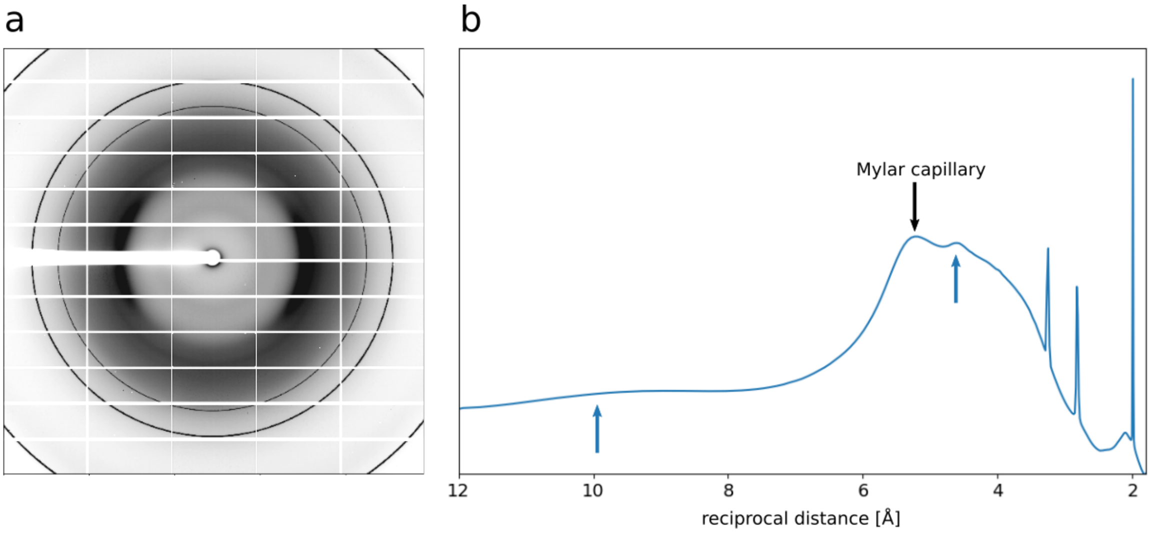


Figure S3. Lysozyme particulates exhibit an amyloid-like signal in X-ray powder diffraction patterns. X-ray powder diffraction pattern for lysozyme particulates (section 1.4); **a** Average over the 20 measurements of the diffraction patterns of the lysozyme particulates powder. **b** The azimuthal integration yields a curve that presents three sharp peaks at small reciprocal distance, corresponding to salt crystals. The blue arrows indicate amyloid-related rings[36] and the black arrow the position of the Mylar capillary peak. We note that the samples for the experiments in figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

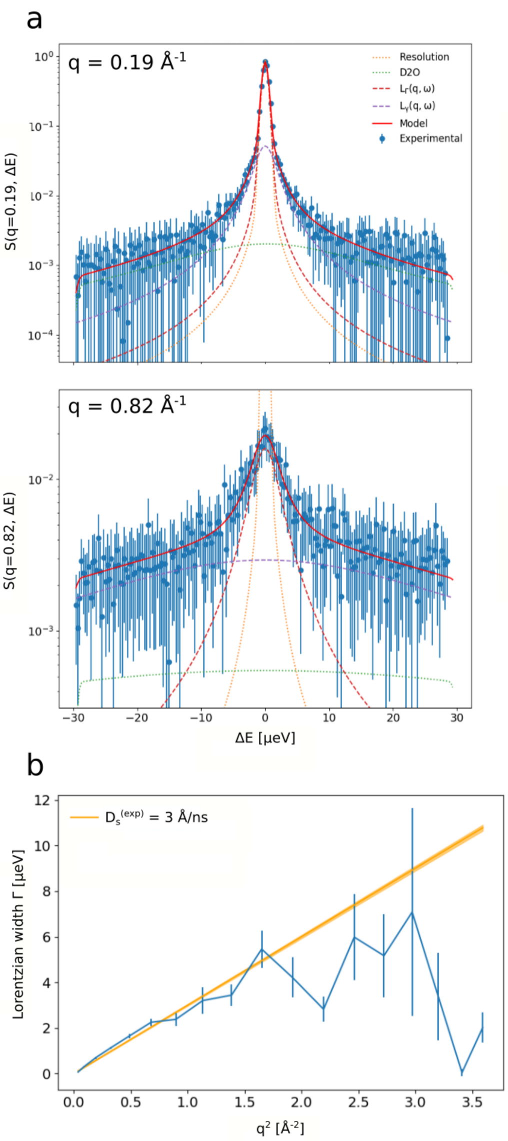


Figure S4. Determination of the center-of-mass diffusion coefficient $D_s^{(exp)}$ of lysozyme with QENS prior to aggregation. a QENS spectra at four different momentum transfer (*q*) values measured at 7°C prior to aggregation (blue circles, errors represented by blue vertical bars) and fit results using the momentum transfer-wise procedure (sections 1.9 and 1.10). The dotted orange line represents the resolution function, the green dotted line the D_2O signal, the dashed red line the Lorentzian describing center-of-mass diffusion, the dashed purple line the Lorentzian describing the protein-internal dynamics, and the red solid line the resulting model. **b** The width of the Lorentzian describing center-of-mass diffusion (blue line with error bars) is plotted against q^2 and fitted using the relation $\Gamma(q) = D_s^{(exp)} q^2$ (orange line with shaded area to represent the standard deviation) to obtain the indicated apparent diffusion coefficient $D_s^{(exp)}$.

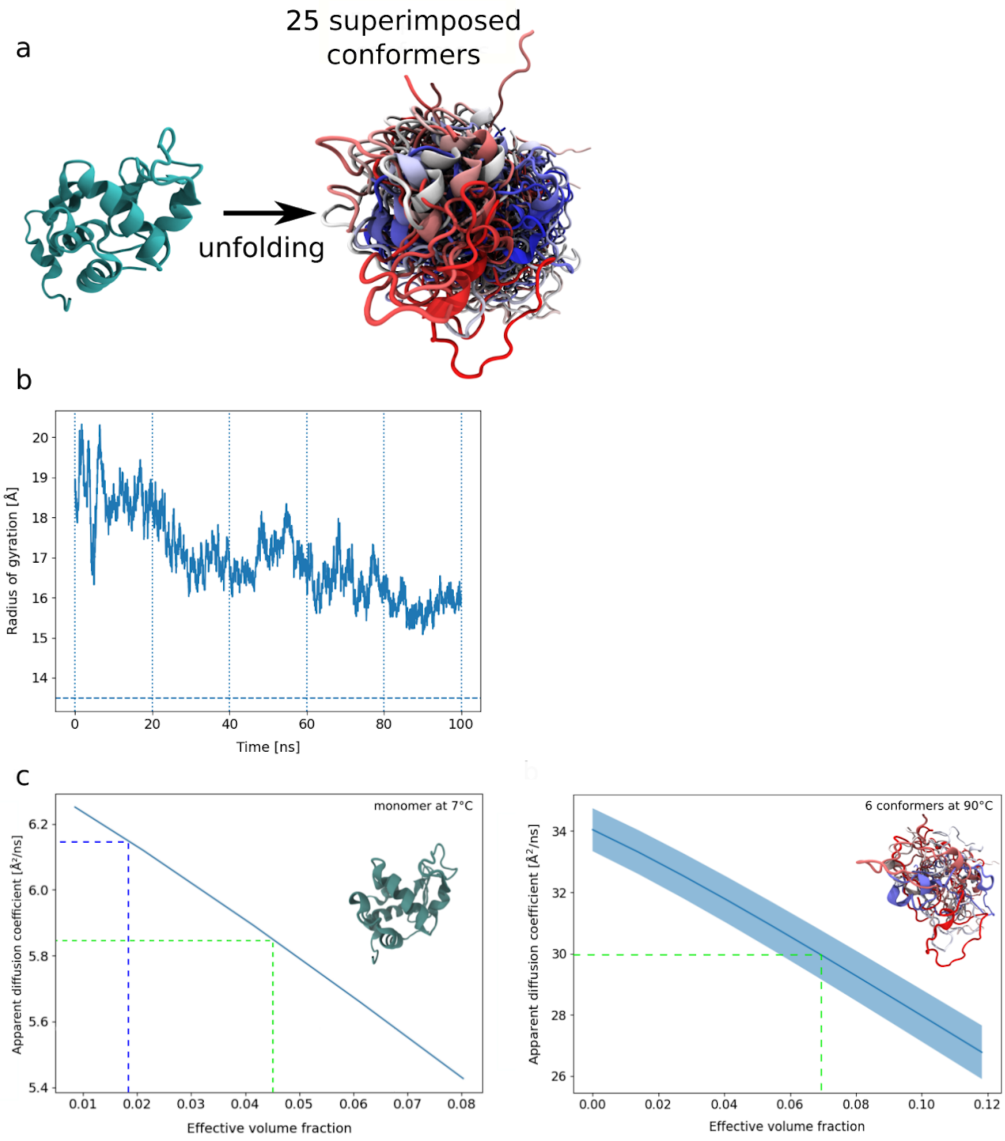


Figure S5. Determination of theoretical apparent center-of-mass diffusion coefficients for monomeric lysozyme at 7 a 90°C. a The native lysozyme structure (PDB 3IJV) was used to compute the apparent center-of-mass diffusion coefficient (as described in Methods, sections 1.5 and 1.6) for the monomer as a function of the protein effective volume fraction at 7° and was subsequently unfolded as explained in section 1.5 during MD simulation. Twenty-five superimposed conformers, extracted at a definite interval of 100 frames along the 2500 frames simulation trajectory, are shown.. **b** Radius of gyration for each conformer during the 100 ns equilibration. The horizontal blue line gives the radius of gyration for the native lysozyme. The vertical dotted lines represent the frames that were used to compute the theoretical diffusion coefficient (as described in Methods, sections 1.5 and 1.6). **c - left** Theoretical center-of-mass diffusion coefficient for the native monomeric lysozyme as a function of the volume fraction. The green dashed line indicates the diffusion coefficient corresponding to the volume fraction that was used in the neutron scattering experiment and the blue dashed line indicates the diffusion coefficient we would obtain for the lysozyme monomer at the volume fraction used in the DLS experiment. **c - right** Theoretical apparent diffusion coefficient as a function of the effective volume fraction for the unfolded lysozyme monomer at 90°C averaged over six conformers (blue line). The shaded blue area represents the standard deviation from the average value. The green dashed line indicates the diffusion coefficient corresponding to the volume fraction that was used in the neutron experiment. Due to the unfolding, the volume fraction occupied by the protein is increased relative to the native protein.

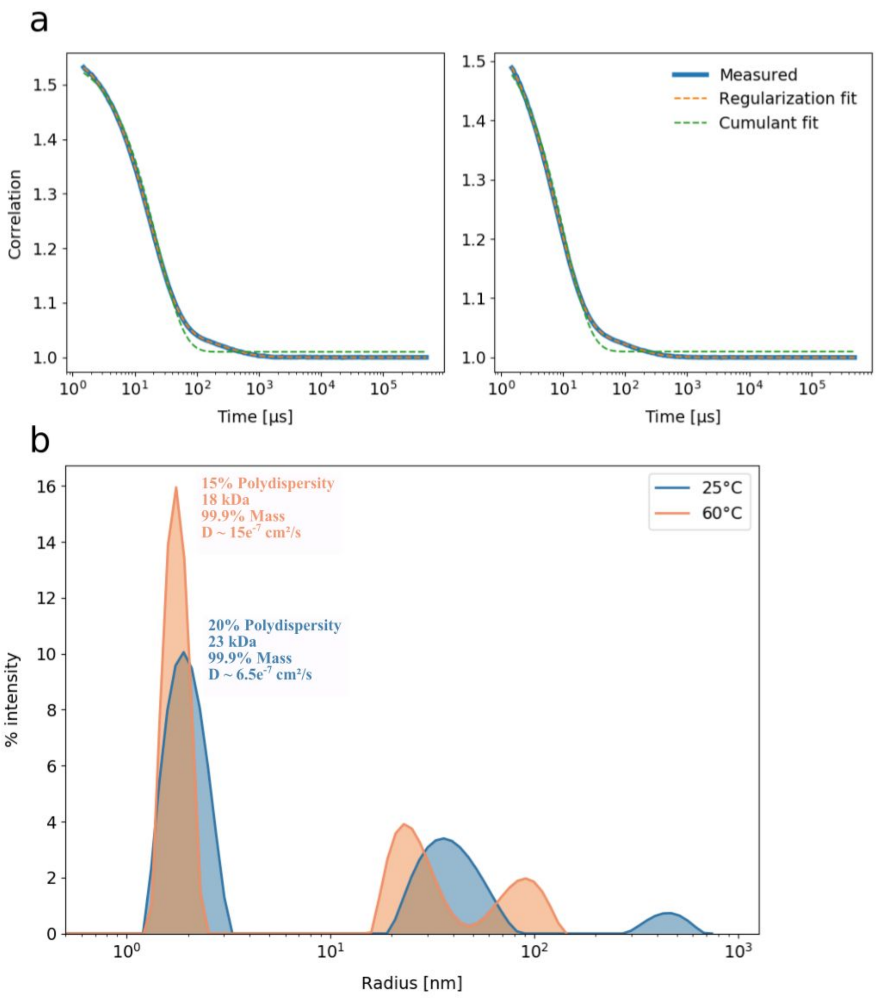


Figure S6. Dynamic light scattering measurement of lysozyme solution suggests the presence of oligomers. a Measured DLS signal (blue line) and fit result - using cumulant fit in green dashed line or regularization fit in orange dashed line as provided in the software (see section 1.7) - for a lysozyme solution at 25 °C (left panel) and 60 °C (right panel). **b** Radius distribution for both temperatures - 25°C in blue and 60°C in orange - represented using the percentage of intensity. Parameters obtained from the software analysis are given for the first peak at each temperature.

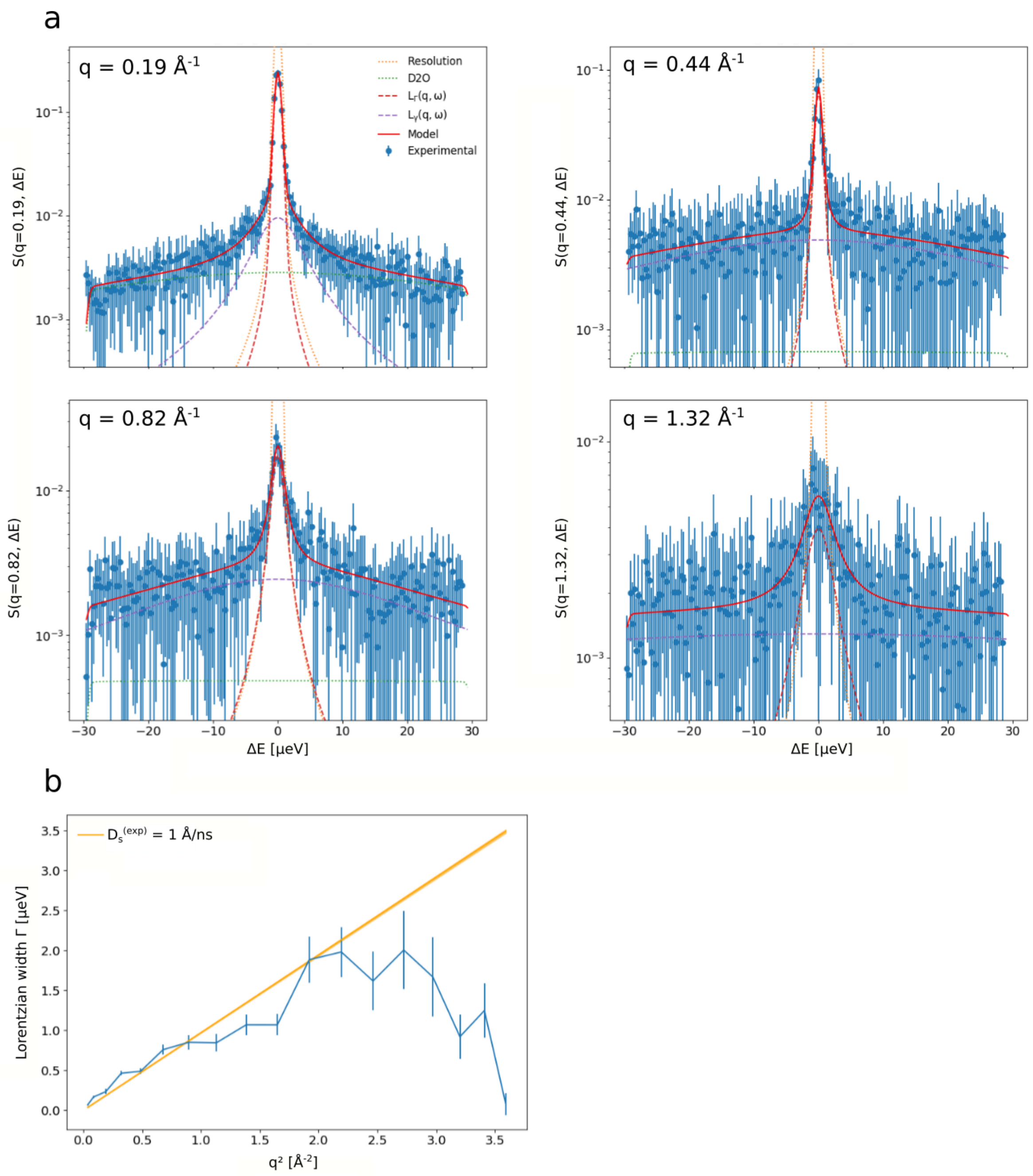


Figure S7. Determination of the center-of-mass diffusion coefficient after aggregation with full QENS. a QENS spectra at four different momentum transfer q values measured at 90°C 6.5 hours after aggregation was triggered (blue circles, errors represented with blue vertical bars) and fit results using the momentum transfer-wise procedure (sections 1.9 and 1.10). The dotted orange line represents the resolution function the green dotted line the D_2O signal, the dashed red line the Lorentzian describing center-of-mass diffusion, the dashed purple line the Lorentzian describing the protein-internal dynamics, and the red solid line the resulting model. **b** The width of the Lorentzian describing center-of-mass diffusion (blue line with error bars) is plotted against q^2 and fitted using the relation $\Gamma(q) = D_s^{(exp)} q^2$ (orange line with shaded area to represent the standard deviation) to obtain the indicated apparent diffusion coefficient $D_s^{(exp)}$.

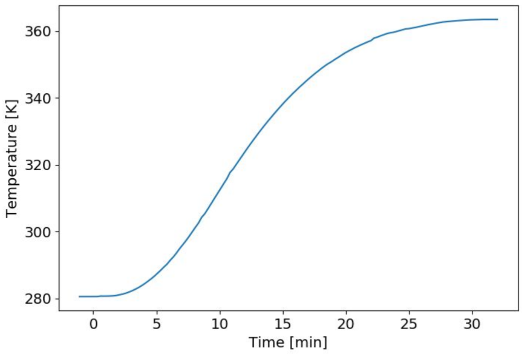


Figure S8. Temperature increase - measured on the sample temperature sensor of the instrument - triggering the aggregation process used during the neutron scattering experiment on IN16B at the ILL (section 1.9).

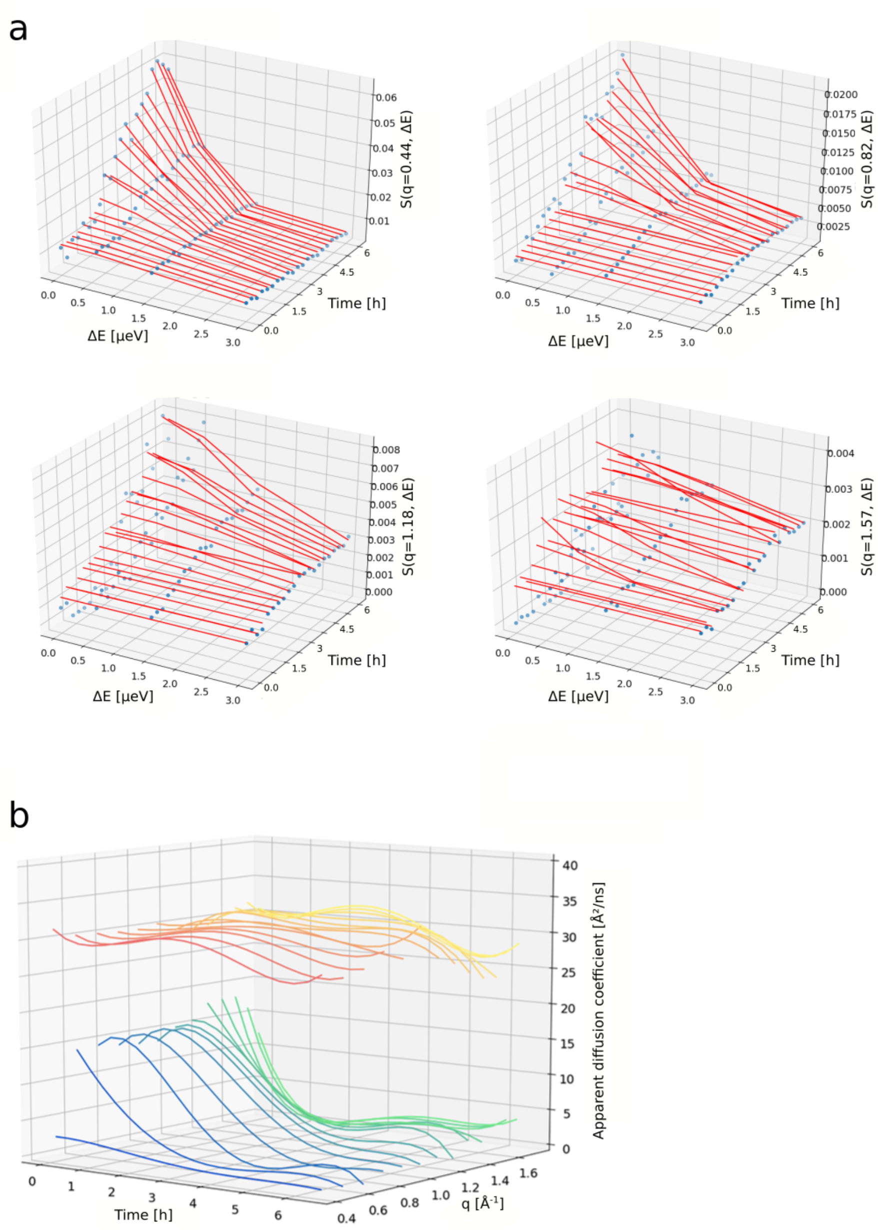


Figure S9. E/IFWS data fitted using a momentum transfer-wise procedure. a Normalized experimental E/IFWS spectra for each scan recorded during aggregation (blues circles) and fitted model (eq. 5 in section 1.10; red lines). **b** Apparent diffusion coefficient from E/IFWS data fitting for center-of-mass (cold colors) and protein-internal dynamics (warm colors) as a function of time and momentum transfer *q* values, as obtained from individual fits of eq. 5 in section 1.10. A Savitzki-Golay filter was used in the time dimension for better visualization (window length of 1 hour and polynomial order of 3).

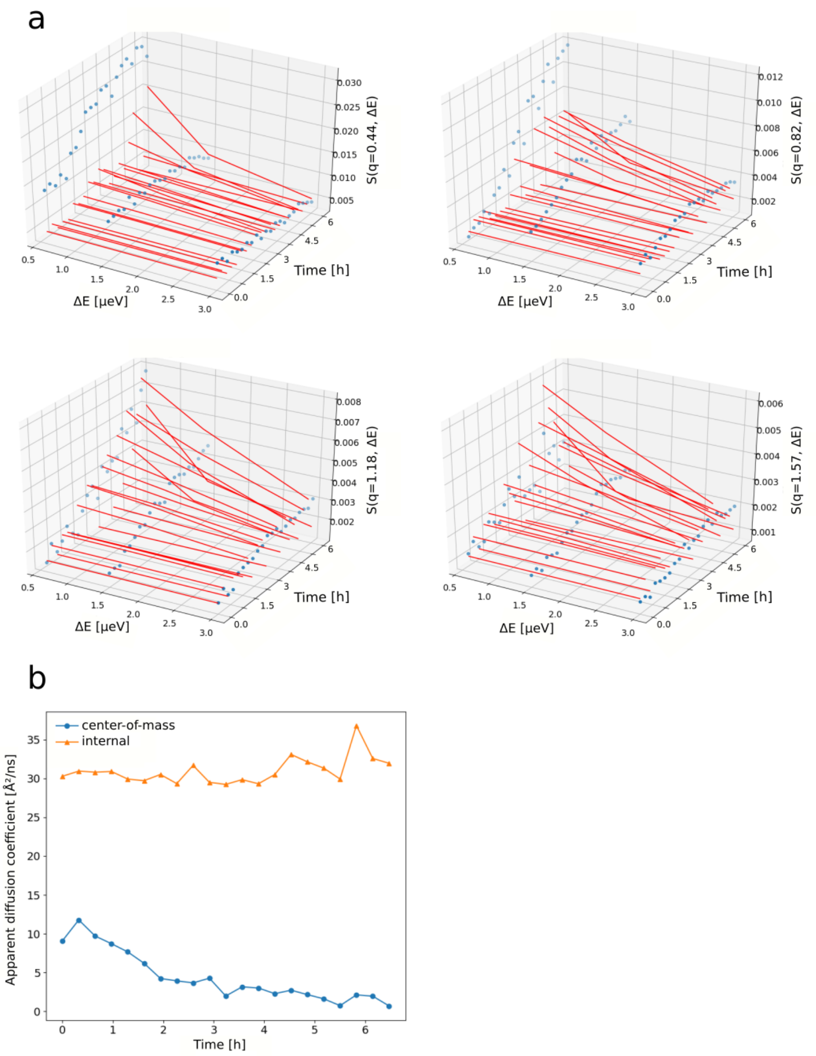


Figure S10. IFWS data fitted using a global fitting procedure. a Normalized experimental IFWS (elastic data excluded) spectra for each scan recorded during aggregation (blues circles) and fitted model (eq. 5 in section 1.10; red lines). **b** Apparent diffusion coefficients from IFWS data (i.e. excluding the EFWS data) fitting for center-of-mass (blue circles) and protein-internal (orange triangles) dynamics as a function of time and momentum transfer q values, as obtained from eq. 5 in section 1.10. The apparent diffusion coefficients obtained using E/IFWS are in reasonable agreement (Figure 3).

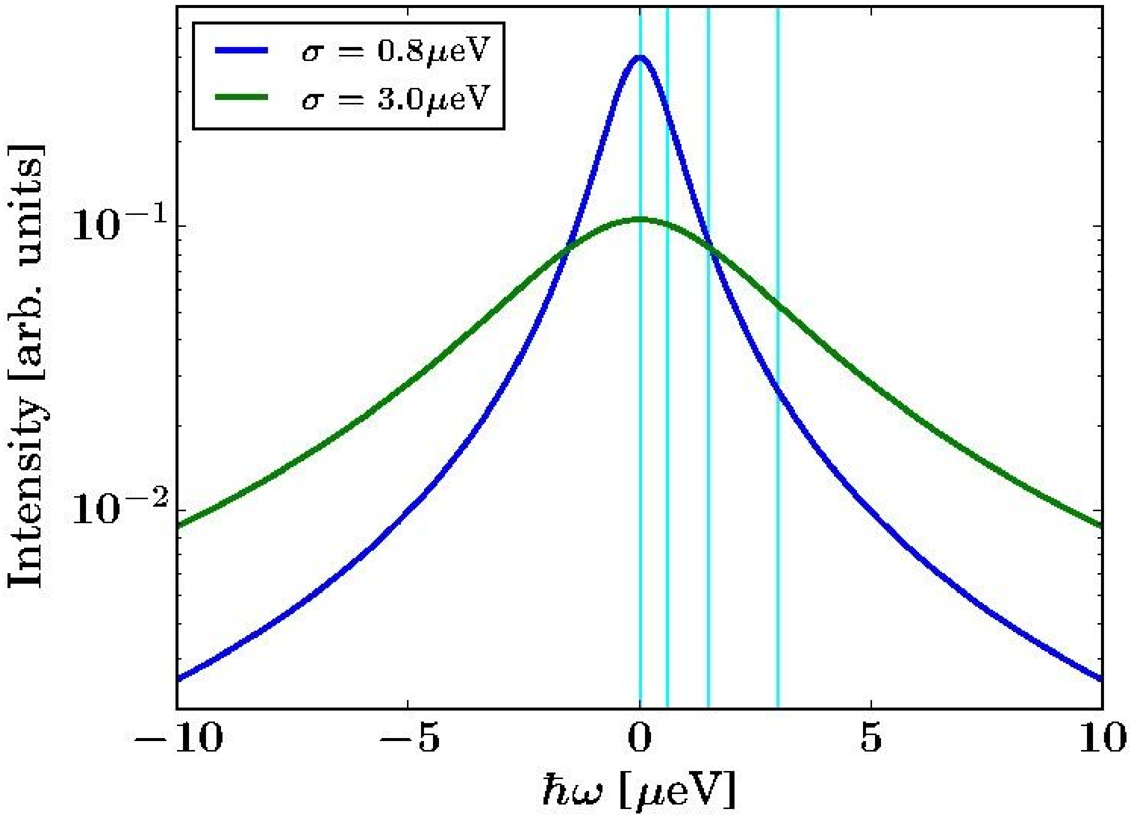


Figure S11: Schematic illustration of the concept of elastic and inelastic fixed window scans (E/IFWS): Two Lorentzian functions with linewidths (HWHW) of σ=0.8 and 3.0 μeV, respectively, are displayed as intensity versus energy transfer, corresponding to simple full QENS spectra (without convolution with the energy resolution function). The vertical lines mark the elastic (EFWS) and three inelastic positions (IFWS) where the intensities were recorded during the neutron backscattering experiment. The intercepts of the vertical lines with the two Lorentzians help to illustrate how the measured intensities evolve at different energy offsets as the sample evolves dynamically: For instance, the formation of protein aggregates corresponds to a narrowing of the Lorentzians. (Cf. section 1.9)

REFERENCES

(1) Kieffer, J.; Wright, J. P. PyFAI: A Python Library for High Performance Azimuthal Integration on GPU. *Powder Diffr.* **2013**, *28* (S2), S339–S350. https://doi.org/10.1017/S0885715613000924.

(2) Ortega, A.; Amorós, D.; García de la Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophys. J.* **2011**, *101* (4), 892–8. https://doi.org/10.1016/j.bpj.2011.06.046.

(3) Booth, D. R.; Sunde, M.; Bellotti, V.; Robinson, C. V.; Hutchinson, W. L.; Fraser, P. E.; Hawkins, P. N.; Dobson, C. M.; Radford, S. E.; Blake, C. C. F.; Pepys, M. B. Instability, Unfolding and Aggregation of Human Lysozyme Variants Underlying Amyloid Fibrillogenesis. *Nature* **1997**, *385* (6619), 787–793. https://doi.org/10.1038/385787a0.

(4) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(5) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802. https://doi.org/10.1002/jcc.20289.

(6) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960. https://doi.org/10.1021/jp003020w.

(7) Huang, J.; MacKerell, A. D.; Jr. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–45. https://doi.org/10.1002/jcc.23354.

(8) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189. https://doi.org/10.1063/1.467468.

(9) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621. https://doi.org/10.1063/1.470648.

(10) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.

(11) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

(12) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-Range Interactions. *Mol. Simul.* **1991**, *6* (1–3), 121–142. https://doi.org/10.1080/08927029108022142.

(13) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992**, *97* (3), 1990–2001. https://doi.org/10.1063/1.463137.

(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593. https://doi.org/10.1063/1.470117.

(15) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of PKa Values. *J. Chem. Theory Comput.* **2011**, *7* (7), 2284–2295. https://doi.org/10.1021/ct200133y.

(16) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537. https://doi.org/10.1021/ct100578z.

(17) Meersman, F.; Atilgan, C.; Miles, A. J.; Bader, R.; Shang, W.; Matagne, A.; Wallace, B. A.; Koch, M. H. J. Consistent Picture of the Reversible Thermal Unfolding of Hen Egg-White Lysozyme from Experiment and Molecular Dynamics. *Biophys. J.* **2010**, *99* (7), 2255–2263. https://doi.org/10.1016/j.bpj.2010.07.060.

(18) Millero, F. J.; Ward, G. K.; Chetirkin, P. Partial Specific Volume, Expansibility, Compressibility, and Heat Capacity of Aqueous Lysozyme Solutions. *J. Biol. Chem.* **1976**, *251* (13), 4001–4004.

(19) Tokuyama, M.; Oppenheim, I. Dynamics of Hard Sphere Suspensions. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **1994**, *50* (1), 1–16. https://doi.org/10.1016/0378-4371(90)90213-C.

(20) Banchio, A. J.; Nägele, G. Short-Time Transport Properties in Dense Suspensions: From Neutral to Charge-Stabilized Colloidal Spheres. *J. Chem. Phys.* **2008**, *128* (10). https://doi.org/10.1063/1.2868773.

(21) Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R. M. J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. Protein Self-Diffusion in Crowded Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (29), 11815–20. https://doi.org/10.1073/pnas.1107287108.

(22) Grimaldo, M.; Roosen-Runge, F.; Jalarvo, N.; Zamponi, M.; Zanini, F.; Hennig, M.; Zhang, F.; Schreiber, F.; Seydel, T. High-Resolution Neutron Spectroscopy on Protein Solution Samples. *EPJ Web Conf.* **2015**, *83*, 02005. https://doi.org/10.1051/epjconf/20158302005.

(23) Cho, C. H.; Urquidi, J.; Singh, S.; Robinson, G. W. Thermal Offset Viscosities of Liquid H2O, D2O, and T2O. *J. Phys. Chem. B* **1999**, *103* (11), 1991–1994. https://doi.org/10.1021/jp9842953.

(24) Frick, B. The Neutron Backscattering Spectrometer IN16 at ILL—High Energy Resolution with High Intensity and Excellent Signal-to-Noise Ratio. *Neutron News* **2002**, *13* (2), 15–22. https://doi.org/10.1080/10448630208218478.

(25) Frick, B.; Mamontov, E.; Eijck, L. van; Seydel, T. Recent Backscattering Instrument Developments at the ILL and SNS. *Z. Für Phys. Chem.* **2010**, *224* (1–2), 33–60. https://doi.org/10.1524/zpch.2010.6091.

(26) Frick, B.; Combet, J.; van Eijck, L. New Possibilities with Inelastic Fixed Window Scans and Linear Motor Doppler Drives on High Resolution Neutron Backscattering Spectrometers. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **2012**, *669*, 7–13. https://doi.org/10.1016/J.NIMA.2011.11.090.

(27) Uehara, K.; Takeshita, H.; Kotaka, H. Hydrogen Gas Generation in the Wet Cutting of Aluminum and Its Alloys. *J. Mater. Process. Technol.* **2002**, *127* (2), 174–177. https://doi.org/10.1016/S0924-0136(02)00121-8.

(28) Grimaldo, M.; Roosen-Runge, F.; Hennig, M.; Zanini, F.; Zhang, F.; Jalarvo, N.; Zamponi, M.; Schreiber, F.; Seydel, T. Hierarchical Molecular Dynamics of Bovine Serum Albumin in Concentrated Aqueous Solution below and above Thermal Denaturation. *Phys. Chem. Chem. Phys.* **2015**, *17* (6), 4645–4655. https://doi.org/10.1039/C4CP04944F.

(29) Grimaldo, M.; Lopez, H.; Beck, C.; Roosen-Runge, F.; Moulin, M.; Devos, J. M.; Laux, V.; Härtlein, M.; Da Vela, S.; Schweins, R.; Mariani, A.; Zhang, F.; Barrat, J.-L.; Oettel, M.; Forsyth, V. T.; Seydel, T.; Schreiber, F. Protein Short-Time Diffusion in a Naturally Crowded Environment. *J. Phys. Chem. Lett.* **2019**, *10* (8), 1709–1715. https://doi.org/10.1021/acs.jpclett.9b00345.

(30) Grimaldo, M.; Roosen-Runge, F.; Zhang, F.; Schreiber, F.; Seydel, T. Dynamics of Proteins in Solution. *Q. Rev. Biophys.* **2019**, *52*. https://doi.org/10.1017/S0033583519000027.

(31) Arnold, O.; Bilheux, J. C.; Borreguero, J. M.; Buts, A.; Campbell, S. I.; Chapon, L.; Doucet, M.; Draper, N.; Ferraz Leal, R.; Gigg, M. A.; Lynch, V. E.; Markvardsen, A.; Mikkelson, D. J.; Mikkelson, R. L.; Miller, R.; Palmen, K.; Parker, P.; Passos, G.; Perring, T. G.; Peterson, P. F.; Ren, S.; Reuter, M. A.; Savici, A. T.; Taylor, J. W.; Taylor, R. J.; Tolchenov, R.; Zhou, W.; Zikovsky, J. Mantid—Data Analysis and Visualization Package for Neutron Scattering and μ SR Experiments. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **2014**, *764*, 156–166. https://doi.org/10.1016/J.NIMA.2014.07.029.

(32) *NIST Handbook of Mathematical Functions*; Olver, F. W. J., National Institute of Standards and Technology (U.S.), Eds.; Cambridge University Press : NIST: Cambridge ; New York, 2010.

(33) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy

Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. https://doi.org/10.1021/jp970984n.

(34) van de Weert, M.; Andersen, M. B.; Frokjaer, S. Complex Coacervation of Lysozyme and Heparin: Complex Characterization and Protein Stability. *Pharm. Res.* **2004**, *21* (12), 2354–2359. https://doi.org/10.1007/s11095-004-7689-z.

(35) Krebs, M. R. H.; Devlin, G. L.; Donald, A. M. Protein Particulates: Another Generic Form of Protein Aggregation? *Biophys. J.* **2007**, *92* (4), 1336–1342.

(36) Eanes, E. D.; Glenner, G. G. X-Ray Diffraction Studies on Amyloid Filaments. *J. Histochem. Cytochem.* **1968**, *16* (11), 673–677. https://doi.org/10.1177/16.11.673.

3.5 Tau fibrillation on aluminum surfaces - preliminary results

The protein tau is a member of the microtubule-associated proteins family. They are mostly expressed in the brain. The alternative splicing results in six possible isoforms, all of which can be found in amyloid deposits in the brain of Alzheimer's patients [Braak et al., 1989] [Buée et al., 2000]. It was found that the proteins in these deposits exhibit a specific structural pattern, called cross-β, which consists in a stack of β-sheets that eventually result an elongated fibril [Kirschner et al., 1986] [Fitzpatrick et al., 2017]. The formation of these elongated fibrils is known as the amyloid fibrillation process. Various evidences show that the toxicity of the amyloid aggregation originates from early-stage oligomeric species or growing fibrils that can interact with lipid membranes and damage them upon fibril growth [Winner et al., 2011] [Skamris et al., 2019] [Martel et al., 2017] [Jones et al., 2012] [Ait-Bouziad et al., 2017]. Whilst the Alzheimer's disease is a complex pathology for which amyloid aggregation is certainly not the sole responsible [Canter et al., 2016], tau remains an important target for the development of new treatments, yet it is still poorly understood.
Several risk factors were identified for Alzheimer's disease, among which we can find the exposition to aluminum [Mantyh et al., 1993]. It appears that a lifetime exposition to the metal can lead to its accumulation in the brain [Walton, 2010] and a small amount of aluminum are sufficient to trigger amyloid aggregation *in-vitro* [Tomljenovic, 2011]. In mice, it has been shown that the chronic administration of aluminum ions promotes tau aggregation and neurotoxicity [Oshima et al., 2013]. However, aluminum ions do not induce a substantial structural change in tau [Madhav et al., 1996] but do promote phosphate-independent aggregation into non-fibrillar, possibly oligomeric species [Scott et al., 1993]. Interestingly, aluminum oxide nanoparticles can induce detectable structural change in tau upon binding to the protein hydrophilic residues, resulting in cytotoxicity [Kermani et al., 2018]. Hence, aluminum ions and aluminum oxide surfaces seem to have distinct effects on tau aggregation. We used incoherent neutron scattering, combined with thioflavin T (ThT) fluorescence measurement to monitor the center-of-mass self-diffusion coefficient and the protein internal dynamics during the amyloid aggregation process.
The incoherent neutron scattering allows to probe the ensemble-averaged dynamics, with a ps-ns time scale for the quasi-elastic neutron scattering. Moreover, the incoherent scattering signal in biological samples is dominated by the hydrogen atoms, a property that permits to employ deuteration strategy to monitor either the dynamics of protein or hydration water independently [Gabel et al., 2002] [Grimaldo et al., 19ed]. The technique was previously employed with hydrated pow-

ers of tau monomers [Gallat et al., 2012a] and mature fibrils [Fichou et al., 2015b] [Schirò et al., 2015]. It has been shown that the hydration water dynamics is increased in fibrils as compared to monomers, while the protein-internal dynamics - backbone and side-chain motions essentially - is not affected. The result suggests that hydration water plays a crucial role for amyloid fibrillation of tau, but it tells nothing on the aggregation pathway itself, and whether or not transient oligomeric species are formed.

The oligomeric species that appear on the tau aggregation pathway attract more and more attention as they exhibit cytotoxicity, and can also spread in a prion-like way through endo and exo-cytosis [Shafiei et al., 2017]. Unfortunately, their transient nature makes them challenging to isolate and study. Recently, methods were found to form stable oligomers. One makes use of chemical labeling of the protein cysteines to prevent the cysteine-dependent aggregation [Karikari et al., 2019]. The other makes use of a cross-seeding mechanism, that is, oligomers of Aβ-peptide are used as seeds to induce the aggregation of tau into stable oligomers that exhibit toxicity [Lasagna-Reeves et al., 2010]. However, no method was reported yet to isolate oligomers formed on the heparin-triggered tau aggregation pathway, nor in presence of aluminum, which motivates the use of time-resolved studies to obtain information about these elusive oligomers.

In a previous section (3.4, we demonstrated that the formation of lysozyme particulates - a spherical, micrometer-sized aggregate of lysozyme - can be followed using incoherent neutron scattering, by recording the so-called elastic and inelastic fixed-window scans (E/IFWS). The method allowed to monitor simultaneously the center-of-mass (self-)diffusion and the protein-internal dynamics of lysozyme during the aggregation process with a time resolution of approximately 20 minutes. The incoherent nature of the scattering and the ns time correlation make the neutron scattering unique in that it probes the so-called self-diffusion, which depends only on the particle effective hydrodynamic radius and electro/hydrodynamic effects, and thus provide a very accurate measure of protein aggregation [Beck et al., 2018] [Grimaldo et al., 2019].

We hereby present the time-resolved elastic and inelastic fixed-window scans (E/IFWS) study of tau amyloid fibrillation triggered by heparin (measurements performed at the ILL, end of January 2020). We used the tau isoform 2, of 441 amino acid long, which showed a good reproducibility for the fibrillation kinetics. The heparin was added to a solution of tau and directly measured using E/IFWS. We observed that the aluminum surface induces a fast decrease of the center-of-mass diffusion and an increase of the protein-internal dynamics. Moreover, the aluminum surface and aluminum ions in solution do not have the same effects on tau aggregation, as the aluminum surface promotes the formation of ThT-positive (cross-β rich?) oligomeric species, while no aggregates, nor ThT-positive species were detected with aluminum ions.

Our preliminary results suggest that the aluminum surface inhibits the formation of heparin-triggered tau fibrils, but promotes the formation of oligomeric species that might pertain to a different aggregation pathway than fibrils. Furthermore, it is also clear that the nature and oxidation state of the aluminum - either oxidized surface or trivalent Al^{3+} ions (aluminum chloride) - has distinct effect on the aggregation process. The effect of aluminum on the dynamics of tau and fibrillation kinetics was not expected originally as we wanted to repeat the experiment we conducted on lysozyme (section 3.4) on a more medically relevant system.

This work is presented in details in the following and constitute a preliminary study for which several results need to be reproduced. The effect of aluminum on the fibrillation and the dynamics of tau - and potentially on its hydration water - needs to be further investigated as discussed in section 3.5.2.

3.5.1 Results

Tau fibrillation monitoring. The conditions to obtain fibrils for tau at 60 mg/mL within 24 h were investigated using ThT fluorescence monitoring in a plate-reader as described in methods. We found that the following buffer - 20 mM HEPES, 150 mM NaCl, 5 mM DTT, 5 mM EDTA at pD 7.2, D_2O - allows for the formation of fibrils within 24 h with a good reproducibility (Figure 3.9).

Time-resolved E/IFWS. The protein solution was prepared as described in methods with a concentration of tau of 60 mg/ mL. Two independent samples were measured using the E/IFWS procedure described in methods and we observed a similar behavior for each sample, that is the elastic signal decreases for approximately 2 hours and then remains constant throughout the experiment (Figure 3.10a). The observation of the samples using electron microscopy, after radioprotection clearance, revealed that no aggregated species is present (Figure 3.10b). Also, 80 μL of the same sample preparation was used to monitor the ThT fluorescence signal in a plate-reader. It appeared that the ThT signal increases as expected for the amyloid fibrillation (Figure 3.10c) and the observation with electron microscopy after 16 hours showed the presence of fibrillated species in the sample incubated in the plate-reader (Figure 3.10d). The inelastic fixed-window scans show a more erratic behavior that is difficult to interpret qualitatively, except for energy offset $\Delta E = 0.6$

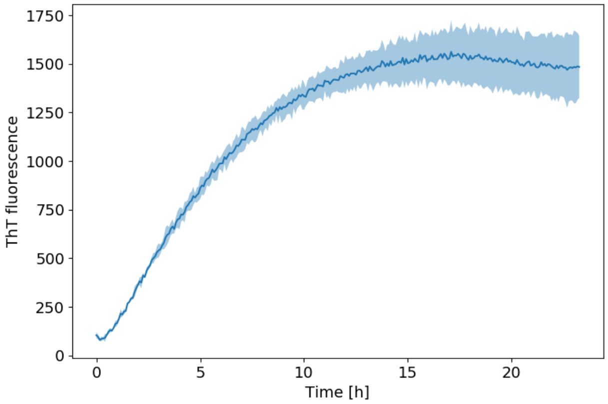


Figure 3.9: The fibrillation of tau is highly reproducible in the plate-reader. The protein solution was prepared and the ThT fluorescence was monitored as described in methods. The solid blue line represents the average of 3 independent sample preparations, and the blue shaded area represents the standard deviation.

μeV where the signal evolves as for the elastic signal (Figure 3.11). The neutron scattering data were analyzed as described in methods. The fitted model shows that the sample can be interpreted with two populations of tau, one pertains to the apparent immobile fraction which represents about 15 % of the measured signal at the beginning of the experiment and decreases to 7 % on average after 2 hours (Figure 3.12a). The other fraction is described by both center-of-mass self-diffusion $\mathrm{D_s^{(exp)}}$ and protein-internal dynamics $\mathrm{D_i}$. The coefficient $\mathrm{D_s^{(exp)}}$ is initially of 3 $\mathrm{Å^2/ns}$ and decreases to 0.8 $\mathrm{Å^2/ns}$ in 2 hours. The coefficient $\mathrm{D_i}$ steadily increases from 15 to 28 $\mathrm{Å^2/ns}$ in approximately 5 hours. The weighed sum of both the immobile and mobile fractions allows to obtain a good fit to the data (Figure 3.12b). To understand what $\mathrm{D_s^{(exp)}} = 3\ \mathrm{Å^2/ns}$ corresponds to, a theoretical center-of-mass self-diffusion coefficient was computed for the monomeric tau at 60 mg/mL and at 37°C (see methods for details). Five structures were used and an average value $\mathrm{D_s^{(theo)}} = 7.8\ \mathrm{Å^2/ns}$ was obtained (Figure 3.13). However, the standard deviation is significant as it can be seen on Figure 3.13b, where, depending on the tau structure, $\mathrm{D_s^{(theo)}}$ ranges from 5 to 13 $\mathrm{Å^2/ns}$.

The effect of aluminum on fibrillation kinetics. To determine if and how the aluminum surface of the neutron cell can affect the tau amyloid fibrillation, a solution

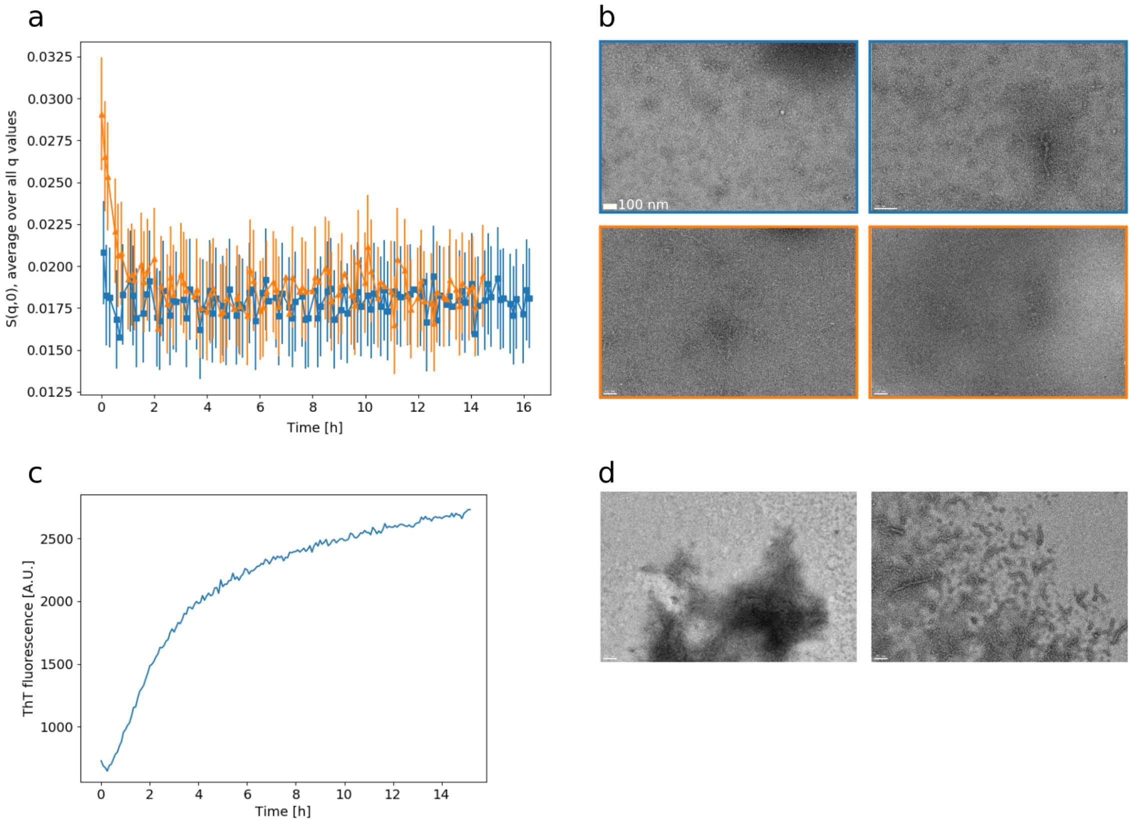


Figure 3.10: The neutron scattering changes over time but no aggregates are formed in the aluminum cell.

The protein solution was prepared for the E/IFWS experiment as described in methods. Two independent measurements were performed as described in methods **(a)** represented by the blue squares and orange triangles with associated errors. **b** After radioprotection clearance, the samples were recovered from the cell and observed with an electron microscope, where 2 images from different places on the grid are shown for each sample. The blue bordered images correspond to an observation made 50 h after triggering the fibrillation with heparin, and the orange to 26 h after triggering the fibrillation. **c** A volume of 80 μL was pipetted in the sample used in **(a)** (orange triangles) before running the second E/IFWS experiment, ThT was then added to a final concentration of 20 μM and the fibrillation kinetics was followed in parallel to the E/IFWS measurement **(a)** in a 96-well plastic plate as described in methods. After 16 h, the sample was recovered from the plate-reader and observed with an electron microscope, where 2 images from different places on the grid are shown for each sample **(d)**.

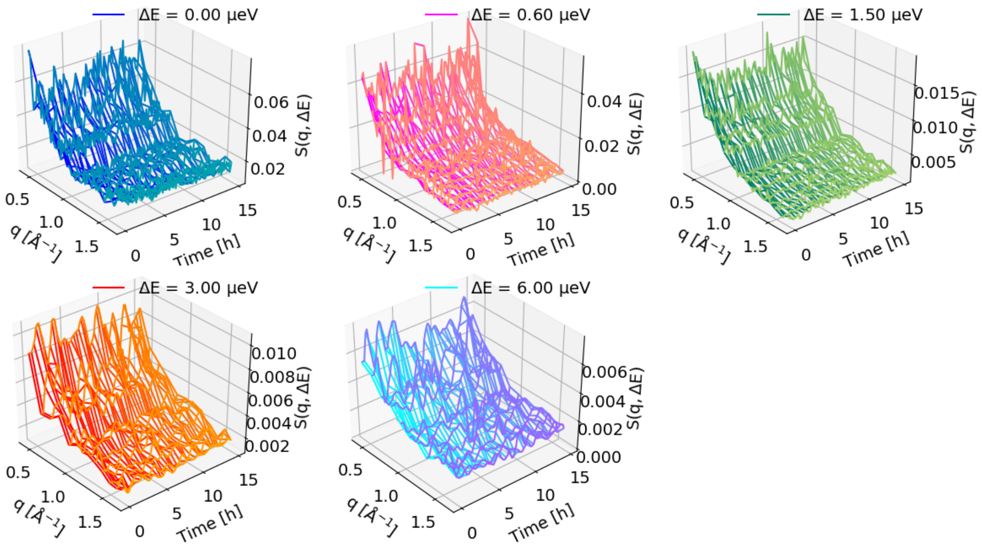


Figure 3.11: Qualitative overview of the E/IFWS experimental data.
The processed data (normalization and empty cell subtraction, see methods) are plotted against momentum transfer q values and time for each energy transfer ΔE.

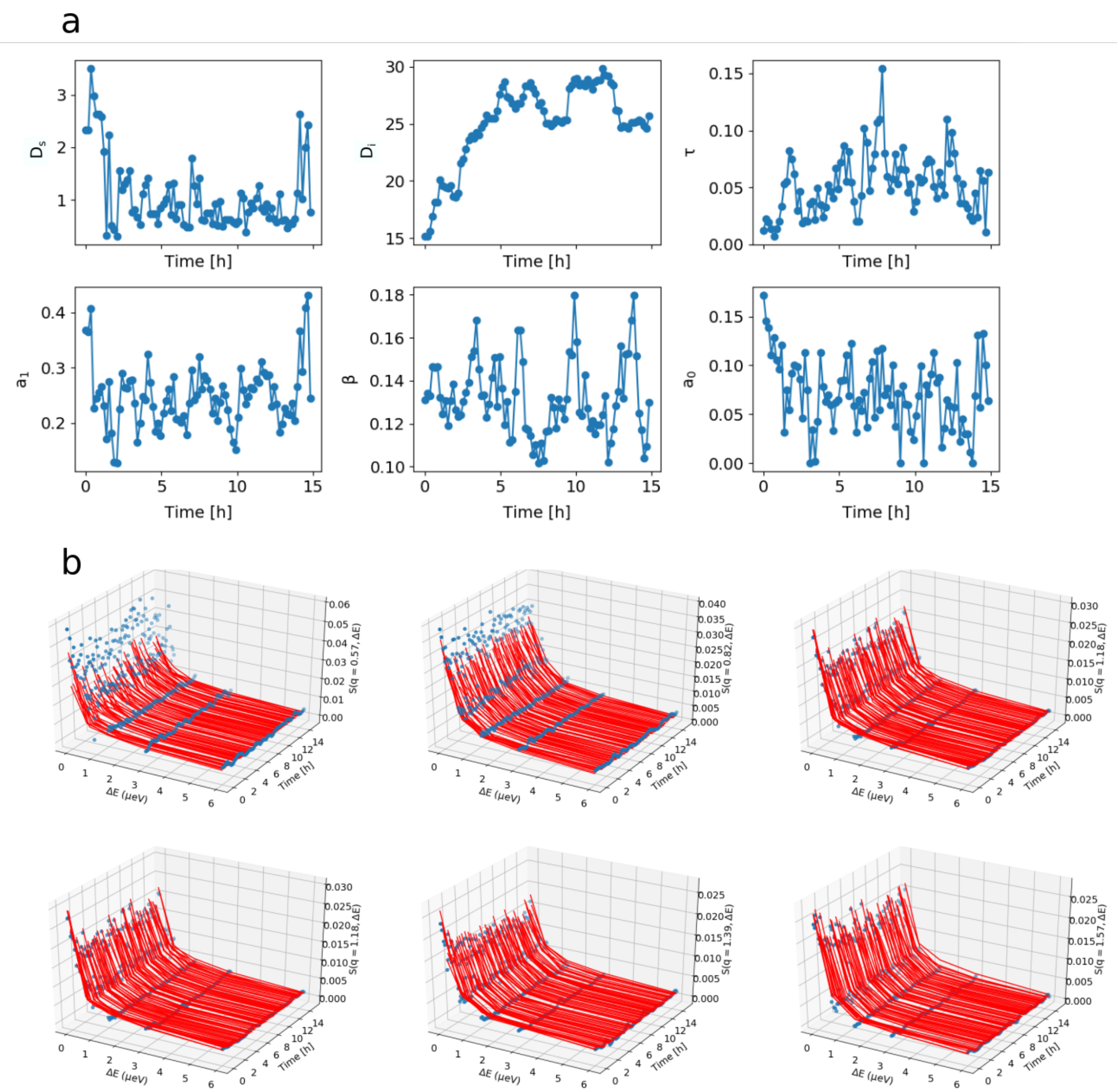


Figure 3.12: The center-of-mass diffusion decreases with time, while the internal dynamics increases.

The E/IFWS data were fitted using equation 3.81 in methods. **a** The fitted parameters are plotted against time where the symbols on the y-axes corresponds to the ones used in equation 3.81. **b** The experimental data are plotted using blue circles as a function of energy transfer ΔE and time. The fitted model is represented by the solid red line.

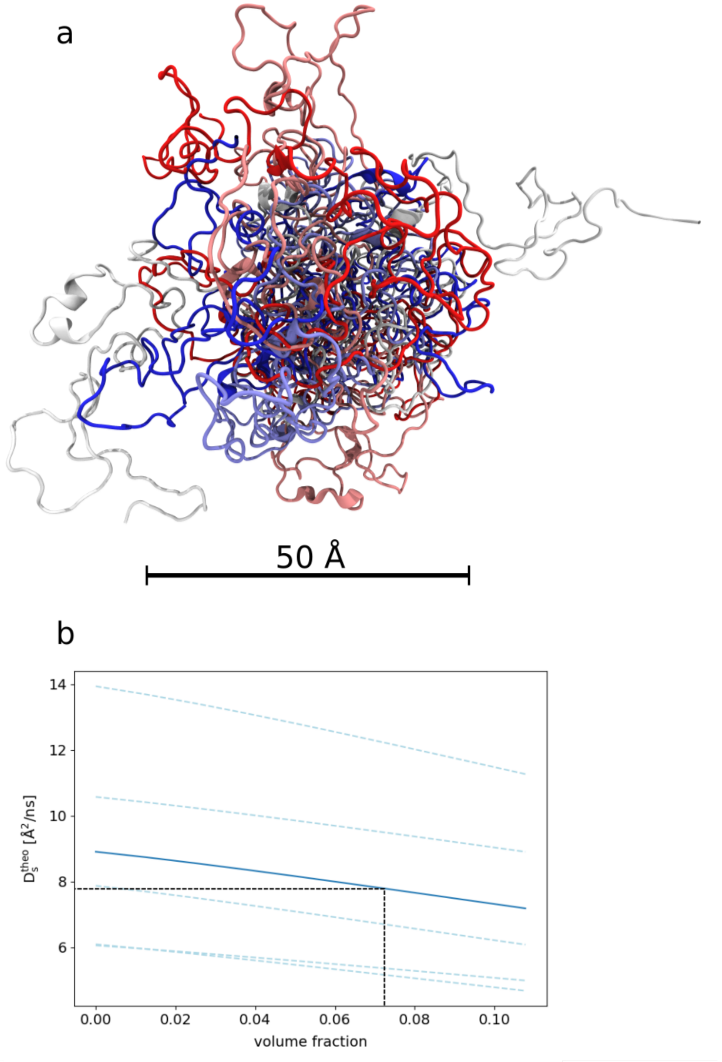


Figure 3.13: The protein tau exhibits important variation of its center-of-mass diffusion coefficient between different conformers.

The theoretical center-of-mass diffusion coefficient $D_s^{(theo)}$ was computed as described in method using the five structure represented in **(a)** with different colors. **b** The parameter $D_s^{(theo)}$ is plotted as a function of the volume fraction, where the average of the five structures is a the blue solid line and individual structure as the lightblue dashed line.

of tau at a concentration of 4 mg/mL was prepared and the ThT fluorescence was monitored in three wells containing either an aluminum foil, aluminum chloride (Al^{3+} ions) or no aluminum (see methods). In the well without aluminum, we observed an increase of the fluorescence, followed by a plateau phase reached 1 hour after the fibrillation was triggered with heparin (Figure 3.14a). With the aluminum foil, the plateau is reached slower but the amplitude of the increase of ThT signal is 7 times bigger than without aluminum foil. The well containing aluminum ions does not show any increase of the ThT signal. The observation with an electron microscope revealed that only the well devoid of aluminum contains tau fibrils (Figure 3.14b). The aluminum ions seem to completely inhibit the fibrillation, while the aluminum foil sample presents several round- shaped particles of 50 nm or less in dimension. The results from figure 3.14 represent only one experiment that needs to be repeated and extended with more protein and concentrations of aluminum ions.

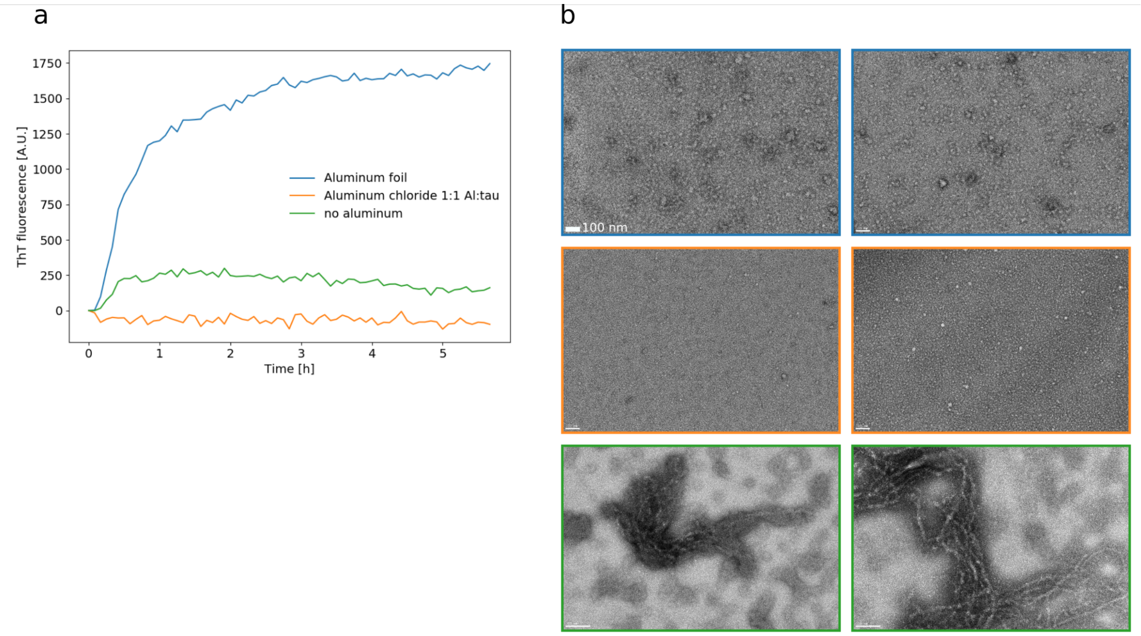


Figure 3.14: Aluminum ions prevent tau fibrillation, but the aluminum surface promotes the formation of a ThT-positive species.
The protein solution was prepared as described in methods with the concentration of tau set at 4 mg/mL and the fibrillation kinetics was followed using a plate- reader **(a)**, for the well for which the walls were covered by an aluminum foil (blue line), the well containing aluminum chloride (orange line) and the well that does not contain aluminum (green line). **b** After 6 hours, the samples were recovered and observed with electron microscopy, where 2 images from different places on the grid are shown for each sample. The border colors correspond to the panel **(a)**.

3.5.2 Discussion

In this preliminary work, we show that the presence of aluminum likely has a profound effect on the tau amyloid fibrillation by combining time-resolved neutron scattering, ThT fluorescence and electron microscopy. Our results suggest that presence of an aluminum surface induces a fast formation of cross-β structure as the ThT signal increases substantially with aluminum foil. This conformation change is accompanied by a decrease of the center-of-mass diffusion coefficient, which indicates that the structure of tau is either becoming more elongated, or that tau is forming oligomers. In addition, the protein-internal dynamics increases significantly, which was unexpected considering a previous result obtain on hydrated powders of tau [Fichou et al., 2015b], where the internal dynamics do not change between monomers and fibrils. The round particles observed with electron microscopy in Figure 3.14 might correspond to oligomers that pertain to a different aggregation pathway than fibrils observed without aluminum.

Additional work is still needed to better understand the effect of aluminum, in particular, X- ray powder diffraction will help to determine whether a cross-β pattern is present. In addition, the stability and the size of the oligomers formed in presence of the aluminum foil have to be measured using dynamic light scattering, native and SDS acrylamide gel, mass spectrometry, and longer ThT kinetics (at least 72 h). If stable enough, the oligomers might even be used to study the hydration water dynamics using hydrated powders of deuterated tau samples with incoherent neutron scattering. Also, additional kinetics measurements are needed, especially, we need to determine if the heparin is necessary to induce aggregation when the aluminum foil is present, which would help in establishing whether the oligomers observed pertain the same pathway of heparin-triggered aggregation or not. Eventually, molecular dynamics simulations might be envisaged to obtain atomic-scale information of the interaction between tau and aluminum surface or Al^{3+} ions.

3.5.3 Methods

Production of tau. The human tau isoform 2 - of 441 amino acids, with a histidine-tag - was expressed in *E. coli* (BL21 DE3) cells using a pET28h vector with a kanamycine resistance gene. A volume of 30 mL of a pre-culture was transferred in 2 L of LB medium and incubated at 37°C and 160 rpm mixing, with kanamycine at 30 μg/mL until an optical density of 0.6 at 600 nm is obtained. The expression of tau was then triggered by adding 1mM IPTG and the cells were harvested 2 h later. For each 2 L of culture, the lysis was performed in 35 mL of a 20 mM Tris, pH 8, 500 mM NaCl, 10 mM imidazole, 5% glycerol, 5 mM β-mercaptoethanol, 5 mM $MgSO_4$, and 0.05 mg/mL DNAse buffer with one tablet of Complete EDTA-free protease inhibitor. The lysis was performed by sonication with a Branson digital sonifier 450 using 8 cycles of 15 s sonication at 90% power followed by 45 s pause. The sonicated cells were then centrifuged at 30000 g for 30 minutes and the supernatant was loaded on a 10 mL Ni-NTA resin equilibrated with 40 mL of the lysis buffer. The resin was washed using 40 mL of a 20 mM Tris, pH 8, 10 mM imidazole, 5% glycerol buffer with 1 M NaCl for the first wash, and 250 mM NaCl for the second. The elution of the protein tau was performed with 30 mL of a 20 mM Tris, pH 8, 250 mM NaCl, 250 mM imidazole, 5% glycerol buffer. The fractions containing the protein tau were identified using a 12% SDS-PAGE gel and were pooled for concentration in Amicon Ultra-15 10 kDa concentrators such that a final volume of slightly less than 2 mL was obtained. Subsequently, 1 mL of the protein tau was loaded on a HiLoad 16/60 S200 size exclusion column with a 0.5 M ammonium acetate solution. The purity of the fractions containing the protein was assessed using a 12% SDS-PAGE gel and the protein concentration was measured on a ThermoScientific NanoDrop One spectrometer using an molecular extinction coefficient $\epsilon = 7450\ \mathrm{M^{-1}cm^{-1}}$. Finally, fractions of 20 mL of the purified tau solution was frozen in a -80°C freezer (the flash freezing used in section 3.1 is not necessary here as we work in liquid state) in 50 mL Falcon tubes and lyophilized using a Christ alpha 1-2 LD plus lyophilizer working at 0.05 mbar with the condenser set at -54°C.

Fibrillation of tau. The heparin used is polydisperse with an average size of 15 kDa, and was ordered at Sigma (ref. H3393-25KU, lot #SLCC8011). The fibrillation protocol was established as follows, tau was dissolved in D_2O with 20 mM HEPES, 150 mM NaCl, 5 mM DTT, 5 mM EDTA at pD 7 at either 5 mg/mL for the reproducibility tests, or at 60 mg/mL for the neutron experiment. The fibrillation was triggered by adding heparin at a molar ratio of 1:4, heparin:tau. To determine the effect of the aluminum, the protein was dissolved in the same buffer as above at a final concentration of 4 mg/mL. Using a 96-well plate, aluminum chloride was added in a well with a molar ratio 1:1, aluminum:tau. In a second well, a piece of aluminum foil was inserted such that all surfaces are covered with aluminum, and a third well without aluminum was used as a control. This last measurement was performed only once and it needs to be reproduced with various protein and aluminum ions concentrations.

Thioflavin T fluorescence assays. The fluorescence of thioflavin T (ThT) was used to monitor the fibrillation of tau with light either prior to the neutron experiment to establish the fibrillation condition (Figure 3.9), in parallel to the neutron experiment (Figure 3.10), or after the neutron experiment to determine the effect of aluminum (Figure 3.14). The protein solution for fibrillation was prepared as described above in a volume of 100 μL and ThT was added to a final concentration of 20 μM. A volume of 80 μL of the solution was pipetted into a 96-well plate. The fluorescence was then monitored at 37°C in a BioTek H4 hybrid plate-reader with optics in bottom position. The excitation wavelength was set to 450 nm and a 490 nm emission filter was used. The gain was set manually to 60. For the tests with aluminum, the optics were in top position, with a distance set to 8 mm, thereby explaining the differences in intensity observed with other measurements.

Theoretical apparent center-of-mass diffusion coefficient for tau monomer. The structures of monomeric tau were extracted from a powder model, equilibrated during a molecular dynamics simulation in a previous study [Fichou et al., 2015a]. Five structures were extracted from the molecular dynamics simulation at random frames along the trajectory. These structures were then used as input in the HYDROPRO software [Ortega et al., 2011] to compute the rotational and translational diffusion coefficient of tau in the dilute limit and at the temperature used in the neutron scattering experiment. To obtain the apparent diffusion coefficient that can be compared with our result from neutron scattering, we used an effective hydrodynamic volume fraction computed using:

$$\Phi_{\text{eff}} = \Phi \left(\frac{R_h}{R} \right)^3 \tag{3.76}$$

where Φ is the dry volume fraction computed from a specific volume $\nu = 0.73$ mL/g (approximation computed using NAMDAnalyzer - ScatDiffusion module), R_h is the hydrodynamic radius obtained from HYDROPRO by plugging the translational diffusion coefficient into the Stokes-Einstein relation, and R the dry protein radius obtained with

$$R = \sqrt[3]{\frac{3}{4\pi} \frac{\nu M_w}{\mathcal{N}_A}} \tag{3.77}$$

with M_w being the molar mass and $\mathcal{N}_A$ the Avogadro number. HYDROPRO returns the translational and rotational diffusion coefficients in the dilute limit, but molecular crowding can be considered using established analytical expressions of scaling factors for translational motions [Tokuyama and Oppenheim, 1994]$^{(eq.11,12)}$ and rotational motions [Banchio and Nägele, 2008]$^{(eq.21)}$. Finally, the apparent diffusion coefficient is computed using the root finding algorithm scipy.optimize.root to solve the implicit relation [Grimaldo et al., 2015b] [Roosen-Runge et al., 2011]:

$$\sum_{l=0}^{n} = B_l(q)\frac{D_r l(l+1) + (D_t - D)^2}{[D_r l(l+1) - (D_t - D)^2]^2} = 0 \tag{3.78}$$

With n being set to 100 to obtain good convergence and $B_l(q)$ being obtained using the radial number density function (relative to the protein center of mass) $\rho(r,n)$ and spherical Bessel function :

$$B_l(q) = (2l+1)\int_0^{\infty} dr \rho(r,n) j_l^2(qr) \tag{3.79}$$

The Python/C scripts that were used to perform the analysis are available on github (github.com/ kpounot/NAMDAnalyzer – ScatDiffusion module).

Electron microscopy. For all samples observed, the protein powder was dissolved in H_2O and 5 μL was absorbed to the clean side of a carbon film that is placed on mica. The adsorbed sample was then negatively stained with 2 % (wt/vol) uranyle acetate and transferred to a 400-mesh copper grid. The images were taken under low dose conditions ($< 10\ e^-/\AA^2$) with defocus values between 1.2 and 2.5 μm on a Tecnai 12 LaB6 electron microscope at 120 kV accelerating voltage using a Gatan Orius 1000 CCD camera.

Neutron scattering. All the measurements were performed at the ILL, on the IN16B spectrometer [Frick, 2002] [Frick et al., 2010]. A mechanical Doppler drive was used to set the incident energy, and a phase space transformation (PST) chopper was employed to optimize the neutron flux at the sample position [Kirstein et al., 2000]. The protein solution for fibrillation was prepared on ice by dissolving the tau powder at 60 mg/mL in D_2O with 20 mM HEPES, 150 mM NaCl, 5 mM DTT, 5 mM EDTA at pD 7. The fibrillation was triggered by adding heparin at a molar ratio of 1:4, heparin:tau. For the second run, 80 μL was pipetted to perform a parallel ThT kinetics in the plate-reader. The solution was then quickly pipetted into a cylindrical aluminum cell with a sample thickness of approximately 0.3 mm, and sealed with 1 mm indium wire. The cell was placed in an Orange cryofurnace with the temperature set at 37°C. The data acquisition scheme for E/IFWS went as follows: (a) - the measurement of the energy transfers at 0, 0.6 and 1.5 μeV - for 0.5, 1 and 3 minutes respectively - was repeated three times. Then (b) - one measurement of the energy transfers at 3 and 6 μeV - for 3 and 5 minutes respectively - is performed. This scheme was repeated until the experiment is stopped. Taking into account the time for the Doppler drive to change its velocity profile, approximately 5 minutes for each measurement of (a), are required resulting in a total time of 15 minutes, which is followed by one measurement of approximately 9 minutes for (b). In addition, QENS data were acquired at 37°C during 2 hours for the vanadium, the empty cell, and the buffer, as well as for the tau solution after 15 h E/IFWS using a energy transfer range of $-30 < \Delta E < 30$ μeV. The data are acquired simultaneously for a range of momentum transfer $0.19 < q < 1.89$ Å^{-1}.

Neutron data analysis. The neutron data were analyzed using a custom-made written python API (nPDyn - github.com/kpounot/nPDyn). The data were normalized to the incoming neutron beam intensity recorded by the so-called monitor device, integrated over region of interest of the position-sensitive detector tubes. For QENS data, the energy axis was computed using the maximum energy transfer measured and the number of channels, and the elastic peaks of the mirrored data were aligned by locating the elastic peaks using a Gaussian profile that was fitted to the data. The empty cell signal and the vanadium signal (resolution function) were fitted using a pseudo- Voigt profile given by:

$$R(q, \Delta E) = A \left[c \frac{\Gamma_R}{\pi(\Delta E^2 + \Gamma_R^2)} + \frac{(1-c)}{\sqrt{2\pi}\gamma_R} \exp^{-\Delta E^2/2\gamma_R^2} + bkgd \right] \tag{3.80}$$

where q is the momentum-transfer value, ΔE the energy exchange value, A a scaling factor that is used for subsequent normalization of data, c , Γ_R and γ_R are scalars, and *bkgd* is a background term accounting for fast vibrational modes. The data were normalized using the parameter A in the resolution function, and the signal from the empty-cell was subtracted. The model describing the dynamics takes into account the possibility of adsorption of tau on the aluminum surface. Hence, we make the assumption that there exists a fraction of protein that diffuses freely in solution and a fraction that is adsorbed, for which an elastic factor was added to the model, which reads:

$$\begin{aligned} S(q, \Delta E) = R(q, \Delta E) \ \otimes \ \beta(q)\{ \ & a_0(q)\delta(\Delta E) \\ & + (1 - a_0(q))[a_1\mathcal{L}_\Gamma + (1 - a_1)\mathcal{L}_{\Gamma+\gamma}] \ \} \\ & + \beta_{\mathrm{D_2O}}\mathcal{L}_{\mathrm{D_2O}} \end{aligned} \tag{3.81}$$

where β is a momentum transfer q dependent scalar, a_0 accounts for the q-dependent elastic incoherent structure factor, $\delta(\Delta E)$ is the Dirac delta, a_1 is a scalar between

0 and 1, $\mathcal{L}_\Gamma$ and $\mathcal{L}_{\Gamma+\gamma}$ are two Lorentzians, with $\Gamma = D_s q^2$, D_s being the center-of-mass diffusion coefficient, and $\gamma = \frac{D_i q^2}{1+D_i q^2 \tau}$ where D_i is the apparent diffusion coefficient for protein-internal dynamics and τ the average time the atom spend in translational motion (from the jump diffusion model [Singwi and Sjölander, 1960]). Finally, $\beta_{\mathrm{D_2O}} \mathcal{L}_{\mathrm{D_2O}}$ represents the signal from the buffer (a scaled Lorentzian), which was measured separately. All the fits were performed using the nPDyn Python package. All the momentum transfer q were fitted at once using the basinhopping algorithm[Wales and Doye, 1997] using momentum transfer values in the range $0.4 < q < 1.7$ Å^{-1}. The global fitting was performed individually for each time point along the experiment.

Conclusion and perspectives

4.1 The protein amyloid aggregation

During this thesis, we obtained unprecedented detailed information on the hydration water dynamics and its coupling to the protein dynamics for the α-synuclein amyloid fibrillation. This was made possible by the high complementarity between incoherent neutron scattering and MD simulations. In addition, we have shown how zinc cations can dramatically affect the insulin dynamics and aggregation kinetics using a combination of neutron scattering experiments, proton-induced X-ray emission, EXAFS, and ThT kinetics. Furthermore, we took advantage of the unique capability of IN16B to perform elastic and inelastic fixed window scans - E/IFWS - to perform the first time-resolved neutron backscattering experiment on an amyloid system, the lysozyme particulates.

The work on α-synuclein benefited from the availability of native and fibril structures for the full-length protein recently solved by NMR and cryo-EM microscopy. Hence, neutron scattering data could be complemented by molecular dynamics simulation that I carried out under the guidance of Pr. Douglas Tobias at UC Irvine. Our results show a systematic increase in hydration water dynamics in fibrils forms as compared to monomers. The simulations informed us that water is gaining in dynamics upon fibrillation, and also that water molecules interact less with fibrils than with monomers, resulting in a larger fraction pertaining to a mobile state. The change in hydration water dynamics, or entropy, is smaller for α-synuclein than previously observed for tau. We preclude that this effect originated from the proportion of the fuzzy coat, which upon fibrillation displace more water molecules into a bulk-like state in tau than in α-synuclein. Indeed, the native forms of both proteins are highly compact [Brodie et al., 2019] [Popov et al., 2019], whilst the fibril forms are more extended, with a central region exhibiting a Greek-key motif that barely interacts with water. Hence, the water molecules are close to the hydrophobic core in the native form, but during the fibrillation they are pulled away from the core, which results in a gain of dynamics and entropy. Interestingly, protein fibrillation barely affects the backbone and side-chain dynamics. However, the protein amino acid sequence plays an important role in determining the internal dynamics, as it was shown with γS-crystallin (section 3.2).

The analysis of the elastic incoherent neutron scattering spectra of γS-crystallin shows that the internal dynamics of the two wild-type forms - native and aggregated - is identical. Similarly for the G18V pH2 and pH7 samples, which however show a systematic lower MSD at high temperature than the wild-type form. This result shows that the γS-crystallin internal dynamics is strongly influenced by its amino acid composition, and not its aggregation state. Not only the protein intrinsic properties - such as amino acid sequence - affect the dynamics of proteins, but also a change in the extrinsic properties can play an important role, as we showed for insulin.

Insulin is a hormone, the expression and activity of which is intimately associated with zinc. It is thus an appropriate system to study the effect of divalent metal ions on protein dynamics and aggregation. In this work, I have performed measurements of insulin dynamics and aggregation kinetics in presence or absence of zinc to decipher its effect on protein dynamics. Our results show that zinc can contribute to hydrate the highly hydrophobic insulin. The initial aggregation step that corresponds to fibril formation is slightly affected by its presence. Hence, while the metal contributes to regulating protein dynamics and activity, it barely affects the amyloid fibril formation of insulin. This observation is in contrast to the effect of metal ions on other systems such as α-synuclein, the amyloid fibrillation kinetics of which is substantially faster in presence of calcium ions [Lautenschläger et al., 2018]. This might be explained by the already high stability of the native folding of insulin. Zinc might screen electrostatic interactions that favor fibrils and spherulites formation, but it contributes also to protein stability by favoring hydration and the water-sustained protein dynamics. Hence, metal ions are of particular interest in the frame of amyloid aggregates formation. As their intra-cellular concentration is well controlled within the cells, we can presume that any age-related change in metal homeostasis can lead to the formation of aggregation-prone protein species in the cell. These early, aggregation-prone, species are often elusive and thus difficult to isolate and study. But their critical role in amyloid-associated toxicity makes it necessary to decipher their formation. To this end, many time-resolved methods were developed. An extensive list of time-resolved methodologies applied to amyloid aggregation are cited in section 3.4 included in this text. In this thesis, we propose a new approach, using elastic and inelastic fixed-window scans (E/IFWS) on IN16B at the ILL, which was applied to a model system, namely the lysozyme aggregation into particulates.

The measurement of E/IFWS in a time-resolved manner allows to unambiguously access both the center-of-mass (self-)diffusion of the protein, and its internal dynamics. By analyzing the obtained center-of-mass diffusion complemented with X-ray powder diffraction and Fourier-transformed infrared spectroscopy, we showed that lysozyme undergoes unfolding and aggregates in a one-step process into amyloid-like particulates. Concurrently, we showed that lysozyme oligomers are already present when the aggregation is triggered. These oligomers are likely to be directly involved in the aggregation process by providing fast forming nuclei from which particulates can grow. Notwithstanding the dramatic change in structure and size of the system as particulates are formed, we observed that the internal dynamics remained constant throughout the experiment. This is in agreement with above results for α-synuclein and γS-crystallin for which backbone and side-chains motions were not severely affected by fibrils formation.

The E/IFWS experiment can be adapted to the system under study by, for instance, reducing the integration time if the concentration can be increased, or recording only the elastic signal in a minute or less if the quasi-elastic signal is not considered that necessary. Moreover, additional measurements can be performed simultaneously within the neutron spectrometer. That is, Raman spectroscopy can be performed [Adams et al., 2009] as well as dielectric relaxation spectroscopy [Sanz et al., 2018]. The E/IFWS experiment was repeated with the protein tau and an unexpected effect of the aluminum is strongly suggested by our preliminary results.

In conclusion, we have at hand plenty of tools to obtain valuable information on the protein amyloid aggregation, ranging from atomistic picture of protein and hydration water dynamics to μm-scale structural features. However, addressing the neurodegenerative diseases problem requires to extend all the biophysical information we can obtain to the biological field, in order to build a systemic approach that is going to help us to link our biophysical results to protein activity and toxicity in cells.

4.2 Outlook

In this work, a great deal of dynamics and mechanistic details of the protein amyloid aggregation could be obtained, but the link to cell biology and neurodegeneration was not explored. This link is of prime importance to address, because the development of new diagnosis methods or treatments precludes knowledge of which aggregated species, on the amyloid aggregation pathway, are toxic as well as when and where the conditions for the formation of these species are met in the organism.

To this purpose, experts from different fields need to interact, including those working on protein dynamics, structural, cell biological and neuroscientific aspects. In the DYNAMOP group at the Institut de Biologie Structurale, we are currently setting up new collaborations, to add to the molecular projects a cellular dimension. This endeavor is facilitated by the great environment in Grenoble, with the presence of the CEA, where promising methodologies for new diagnostic tools are being developed [Pansieri et al., 2019] [Pansieri et al., 2018] and the presence of the Grenoble Institute of Neurosciences, where the biological effects of the amyloid proteins mentioned in this thesis can be thouroughly studied [Decressac et al., 2013a] [Decressac et al., 2013b].

This collaborative work will aim at linking the biophysical characterization (dynamics, aggregation kinetics, structure) of the amyloid proteins to their toxicity in cells and possibly in mice, by studying various α-synuclein and tau mutants and by varying the metal ion concentration.

List of Figures

List of Tables

List of Abbreviations and Symbols

$\langle \cdots \rangle$	Denotes average over indicated quantity in the text (usually time origins)
API	Application programming interface
DLS	Dynamic light scattering
(E/I)FWS	(Elastic/Inelastic)Fixed-window scan
EISF	Elastic incoherent structure factor
EM	Electron microscopy
ENS	Elastic neutron scattering
FRM	Forschungs-Neutronenquelle (Garching - Germany)
FWHM	Full width at half maximum
IDP	Intrinsically disordered protein
ILL	Institut Laue-Langevin (Grenoble - France)
k_B	Boltzmann constant, $k_\mathrm{B} = 8.6173303\ eV/K$
MLZ	Maier-Leibnitz Zentrum (Garching - Germany)
MSD	Mean squared displacement
NMR	Nuclear magnetic resonance
PDT	Protein dynamical transition
PST	Phase transformation chopper

QENS	Quasi-elastic neutron scattering
ThT	Thioflavin T, amyloid specific fluorescent probe
XAFS	X-ray absorption fine structure

References

[Adams et al., 2009] Adams, M. A., Parker, S. F., Fernandez-Alonso, F., Cutler, D. J., Hodges, C., and King, A. (2009). Simultaneous Neutron Scattering and Raman Scattering. *Applied Spectroscopy*, 63(7):727–732.

[Ait-Bouziad et al., 2017] Ait-Bouziad, N., Lv, G., Mahul-Mellier, A.-L., Xiao, S., Zorludemir, G., Eliezer, D., Walz, T., and Lashuel, H. A. (2017). Discovery and characterization of stable and toxic Tau/phospholipid oligomeric complexes. *Nature Communications*, 8(1):1678.

[Albe et al., 1990] Albe, K. R., Butler, M. H., and Wright, B. E. (1990). Cellular concentrations of enzymes and their substrates. *Journal of Theoretical Biology*, 143(2):163–195.

[Andersson et al., 2017] Andersson, C. D., Martinez, N., Zeller, D., Rondahl, S. H., Koza, M. M., Frick, B., Ekström, F., Peters, J., and Linusson, A. (2017). Changes in dynamics of α-chymotrypsin due to covalent inhibitors investigated by elastic incoherent neutron scattering. *Physical Chemistry Chemical Physics*, 19(37):25369–25379.

[Araya-Secchi et al., 2014] Araya-Secchi, R., Perez-Acle, T., Kang, S.-g., Huynh, T., Bernardin, A., Escalona, Y., Garate, J.-A., Martínez, A. D., García, I. E., Sáez, J. C., and Zhou, R. (2014). Characterization of a Novel Water Pocket Inside the Human Cx26 Hemichannel Structure. *Biophysical Journal*, 107(3):599–612.

[Arnold et al., 2014] Arnold, O., Bilheux, J., Borreguero, J., Buts, A., Campbell, S., Chapon, L., Doucet, M., Draper, N., Ferraz Leal, R., Gigg, M., Lynch, V., Markvardsen, A., Mikkelson, D., Mikkelson, R., Miller, R., Palmen, K., Parker, P., Passos, G., Perring, T., Peterson, P., Ren, S., Reuter, M., Savici, A., Taylor, J., Taylor, R., Tolchenov, R., Zhou, W., and Zikovsky, J. (2014). Mantid—Data analysis and visualization package for neutron scattering and μ SR experiments.

Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 764:156–166.

[Aviles et al., 1978] Aviles, F. J., Chapman, G. E., Kneale, G. G., Crane-Robinson, C., and Bradbury, E. M. (1978). The Conformation of Histone H5. Isolation and Characterisation of the Globular Segment. *European Journal of Biochemistry*, 88(2):363–371.

[Aviles-Olmos et al., 2013] Aviles-Olmos, I., Limousin, P., Lees, A., and Foltynie, T. (2013). Parkinson's disease, insulin resistance and novel agents of neuroprotection. *Brain*, 136(2):374–384.

[Badger, 1993] Badger, J. (1993). Multiple hydration layers in cubic insulin crystals. *Biophysical Journal*, 65(4):1656–1659.

[Bahar et al., 1998] Bahar, I., Atilgan, A. R., Demirel, M. C., and Erman, B. (1998). Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability. *Physical Review Letters*, 80(12):2733–2736.

[Ball, 2008] Ball, P. (2008). Water as an Active Constituent in Cell Biology. *Chemical Reviews*, 108(1):74–108.

[Banchio and Nägele, 2008] Banchio, A. J. and Nägele, G. (2008). Short-time transport properties in dense suspensions: From neutral to charge-stabilized colloidal spheres. *Journal of Chemical Physics*, 128(10).

[Beck et al., 2018] Beck, C., Grimaldo, M., Roosen-Runge, F., Braun, M. K., Zhang, F., Schreiber, F., and Seydel, T. (2018). Nanosecond Tracer Diffusion as a Probe of the Solution Structure and Molecular Mobility of Protein Assemblies: The Case of Ovalbumin. *The Journal of Physical Chemistry B*, page acs.jpcb.8b04349.

[Bee, 1988] Bee, M. (1988). Quasielastic neutron scattering.

[Berendsen et al., 1984] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690.

[Biessels and Kappelle, 2005] Biessels, G. J. and Kappelle, L. J. (2005). Increased risk of Alzheimer's disease in Type II diabetes: Insulin resistance of the brain or insulin-induced amyloid pathology? *Biochemical Society Transactions*, 33(5):1041–1044.

REFERENCES

[Boatz et al., 2017] Boatz, J. C., Whitley, M. J., Li, M., Gronenborn, A. M., and Van Der Wel, P. C. (2017). Cataract-associated P23T γ 3D-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH. *Nature Communications*, 8(May):1–10.

[Bonar et al., 1969] Bonar, L., Cohen, A. S., and Skinner, M. M. (1969). Characterization of the Amyloid Fibril as a Cross-β Protein. *Proceedings of the Society for Experimental Biology and Medicine*, 131(4):1373–1375.

[Booth et al., 1997] Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature*, 385(6619):787–793.

[Braak et al., 1989] Braak, H., Braak, E., Bohl, J., and Lang, W. (1989). Alzheimer's disease: Amyloid plaques in the cerebellum. *Journal of the Neurological Sciences*, 93(2):277–287.

[Braun et al., 2017] Braun, M. K., Grimaldo, M., Roosen-Runge, F., Hoffmann, I., Czakkel, O., Sztucki, M., Zhang, F., Schreiber, F., and Seydel, T. (2017). Crowding-Controlled Cluster Size in Concentrated Aqueous Protein Solutions: Structure, Self- and Collective Diffusion. *The Journal of Physical Chemistry Letters*, 8(12):2590–2596.

[Breydo et al., 2012] Breydo, L., Wu, J. W., and Uversky, V. N. (2012). α-Synuclein misfolding and Parkinson's disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1822(2):261–285.

[Brodie et al., 2019] Brodie, N. I., Popov, K. I., Petrotchenko, E. V., Dokholyan, N. V., and Borchers, C. H. (2019). Conformational ensemble of native α-synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations. *PLOS Computational Biology*, 15(3):e1006859.

[Brown et al., 2005] Brown, R. C., Lockwood, A. H., and Sonawane, B. R. (2005). Neurodegenerative Diseases: An Overview of Environmental Risk Factors. *Environmental Health Perspectives*, 113(9):1250–1256.

[Brück et al., 2016] Brück, D., Wenning, G. K., Stefanova, N., and Fellner, L. (2016). Glia and alpha-synuclein in neurodegeneration: A complex interaction. *Neurobiology of Disease*, 85:262–274.

REFERENCES

[Bucciarelli et al., 2016] Bucciarelli, S., Myung, J. S., Farago, B., Das, S., Vliegenthart, G. A., Holderer, O., Winkler, R. G., Schurtenberger, P., Gompper, G., and Stradner, A. (2016). Dramatic influence of patchy attractions on short-time protein diffusion under crowded conditions. *Science Advances*, 2(December):e1601432.

[Buée et al., 2000] Buée, L., Bussière, T., Buée-Scherrer, V., Delacourte, A., and Hof, P. R. (2000). Tau protein isoforms, phosphorylation and role in neurodegenerative disorders11These authors contributed equally to this work. *Brain Research Reviews*, 33(1):95–130.

[Bussi et al., 2007] Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101.

[Canter et al., 2016] Canter, R. G., Penney, J., and Tsai, L.-H. (2016). The road to restoring neural circuits for the treatment of Alzheimer's disease. *Nature*, 539(7628):187–196.

[Cerajewska et al., 2015] Cerajewska, T. L., Davies, M., and West, N. X. (2015). Periodontitis: A potential risk factor for Alzheimer's disease. *British Dental Journal*, 218(1):29–34.

[Chandra et al., 2017] Chandra, R., Hiniker, A., Kuo, Y.-M., Nussbaum, R. L., and Liddle, R. A. (2017). α-Synuclein in gut endocrine cells and its implications for Parkinson's disease. *JCI insight*, 2(12):e92295.

[Chin and Means, 2000] Chin, D. and Means, A. R. (2000). Calmodulin: A prototypical calcium sensor. *Trends in Cell Biology*, 10(8):322–328.

[Chiricotto et al., 2016] Chiricotto, M., Melchionna, S., Derreumaux, P., and Sterpone, F. (2016). Hydrodynamic effects on β-amyloid (16-22) peptide aggregation. *The Journal of Chemical Physics*, 145(3):035102.

[Chiti and Dobson, 2006] Chiti, F. and Dobson, C. M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry*, 75(1):333–366.

[Cinar et al., 2019] Cinar, S., Cinar, H., Chan, H. S., and Winter, R. (2019). Pressure-Sensitive and Osmolyte-Modulated Liquid–Liquid Phase Separation of Eye-Lens γ-Crystallins. *Journal of the American Chemical Society*, 141(18):7347–7354.

REFERENCES

[Das and Pappu, 2013] Das, R. K. and Pappu, R. V. (2013). Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences*, 110(33):13392–13397.

[De Young et al., 1993] De Young, L. R., Fink, A. L., and Dill, K. A. (1993). Aggregation of globular proteins. *Accounts of Chemical Research*, 26(12):614–620.

[Decressac et al., 2012] Decressac, M., Kadkhodaei, B., Mattsson, B., Laguna, A., Perlmann, T., and Björklund, A. (2012). α-Synuclein–Induced Down-Regulation of Nurr1 Disrupts GDNF Signaling in Nigral Dopamine Neurons. *Science Translational Medicine*, 4(163):163ra156–163ra156.

[Decressac et al., 2013a] Decressac, M., Mattsson, B., Weikop, P., Lundblad, M., Jakobsson, J., and Björklund, A. (2013a). TFEB-mediated autophagy rescues midbrain dopamine neurons from α-synuclein toxicity. *Proceedings of the National Academy of Sciences*, 110(19):E1817–E1826.

[Decressac et al., 2013b] Decressac, M., Volakakis, N., Björklund, A., and Perlmann, T. (2013b). NURR1 in Parkinson disease—from pathogenesis to therapeutic potential. *Nature Reviews Neurology*, 9(11):629–636.

[Dixit et al., 2002] Dixit, S., Soper, A. K., Finney, J. L., and Crain, J. (2002). Water structure and solute association in dilute aqueous methanol. *EPL (Europhysics Letters)*, 59(3):377.

[Dobson, 2003] Dobson, C. M. (2003). Protein folding and misfolding. 426:7.

[Dobson, 2015] Dobson, C. M. (2015). Alzheimer's disease: Addressing a twenty-first century plague. *Rendiconti Lincei*, 26(3):251–262.

[Dolman et al., 1997] Dolman, M., Halling, P. J., Moore, B. D., and Waldron, S. (1997). How dry are anhydrous enzymes? Measurement of residual and buried 18O-labeled water molecules using mass spectrometry. *Biopolymers*, 41(3):313–321.

[Dominy et al., 2019] Dominy, S. S., Lynch, C., Ermini, F., Benedyk, M., Marczyk, A., Konradi, A., Nguyen, M., Haditsch, U., Raha, D., Griffin, C., Holsinger, L. J., Arastu-Kapur, S., Kaba, S., Lee, A., Ryder, M. I., Potempa, B., Mydel, P., Hellvard, A., Adamowicz, K., Hasturk, H., Walker, G. D., Reynolds, E. C., Faull, R. L. M., Curtis, M. A., Dragunow, M., and Potempa, J. (2019). *Porphyromonas*

Gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Science Advances*, 5(1):eaau3333.

[Doster et al., 2010] Doster, W., Busch, S., Gaspar, A. M., Appavou, M.-S., Wuttke, J., and Scheer, H. (2010). Dynamical Transition of Protein-Hydration Water. *Physical Review Letters*, 104(9):098101.

[Doster et al., 1989] Doster, W., Cusack, S., and Petry, W. (1989). Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature*, 337:754–756.

[Eanes and Glenner, 1968] Eanes, E. D. and Glenner, G. G. (1968). X-ray diffraction studies on amyloid filaments. *Journal of Histochemistry & Cytochemistry*, 16(11):673–677.

[Exley et al., 2009] Exley, C., House, E., Collingwood, J., Davidson, M., Cannon, D., and Donald, A. (2009). Spherulites of Aβ 42 in vitro and in Alzheimer's disease. *Nature Precedings*.

[Fazili et al., 2016] Fazili, N. A., Bhat, I. A., Bhat, W. F., and Naeem, A. (2016). Anti-fibrillation propensity of a flavonoid baicalein against the fibrils of hen egg white lysozyme: Potential therapeutics for lysozyme amyloidosis. *Journal of Biomolecular Structure and Dynamics*, 34(10):2102–2114.

[Fee and Phillips, 1975] Fee, J. A. and Phillips, W. (1975). The behavior of holo- and apo-forms of bovine superoxide dismutase at low pH. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 412(1):26–38.

[Feller et al., 1995] Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. (1995). Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, 103(11):4613–4621.

[Ferrand et al., 1993] Ferrand, M., Dianoux, A. J., Petry, W., and Zaccaï, G. (1993). Thermal motions and function of bacteriorhodopsin in purple membranes: Effects of temperature and hydration studied by neutron scattering. *Proceedings of the National Academy of Sciences*, 90(20):9668–9672.

[Fichou et al., 2015a] Fichou, Y., Heyden, M., Zaccai, G., Weik, M., and Tobias, D. J. (2015a). Molecular Dynamics Simulations of a Powder Model of the Intrinsically Disordered Protein Tau. *The Journal of Physical Chemistry B*, 119(39):12580–12589.

[Fichou et al., 2015b] Fichou, Y., Schirò, G., Gallat, F.-X., Laguri, C., Moulin, M., Combet, J., Zamponi, M., Härtlein, M., Picart, C., Mossou, E., Lortat-Jacob, H., Colletier, J.-P., Tobias, D. J., and Weik, M. (2015b). Hydration water mobility is enhanced around tau amyloid fibers. *Proceedings of the National Academy of Sciences of the United States of America*, 112(20):6365–6370.

[Fick, 1855] Fick, A. (1855). On liquid diffusion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(63):30–39.

[Filipponi and Cicco, 2000] Filipponi, A. and Cicco, A. D. (2000). GNXAS: A software package for advanced EXAFS multiple-scattering calculations and data analysis. *Task Quaterly 4*, 4:575–669.

[Finney et al., 2003] Finney, J., Bowron, D., Daniel, R., Timmins, P., and Roberts, M. (2003). Molecular and mesoscale structures in hydrophobically driven aqueous solutions. *Biophysical Chemistry*, 105(2-3):391–409.

[Fitzpatrick et al., 2017] Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M., and Scheres, S. H. W. (2017). Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature Publishing Group*, 547:185–190.

[Foderá and Donald, 2010] Foderá, V. and Donald, A. M. (2010). Tracking the heterogeneous distribution of amyloid spherulites and their population balance with free fibrils. *European Physical Journal E*, 33(4):273–282.

[Frederickson, 1989] Frederickson, C. J. (1989). Neurobiology of Zinc and Zinc-Containing Neurons. In Smythies, J. R. and Bradley, R. J., editors, *International Review of Neurobiology*, volume 31, pages 145–238. Academic Press.

[Frick, 2002] Frick, B. (2002). The Neutron backscattering spectrometer IN16 at ILL—high energy resolution with high intensity and excellent signal-to-noise ratio. *Neutron News*, 13(2):15–22.

[Frick et al., 2012] Frick, B., Combet, J., and Van Eijck, L. (2012). New possibilities with inelastic fixed window scans and linear motor Doppler drives on high resolution neutron backscattering spectrometers. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 669:7–13.

[Frick et al., 2010] Frick, B., Mamontov, E., van Eijck, L., and Seydel, T. (2010). Recent Backscattering Instrument Developments at the ILL and SNS. *Zeitschrift für Physikalische Chemie*, 224(1-2):33–60.

[Frölich et al., 2009] Frölich, A., Gabel, F., Jasnin, M., Lehnert, U., Oesterhelt, D., Stadler, A. M., Tehei, M., Weik, M., Wood, K., and Zaccai, G. (2009). From shell to cell: Neutron scattering studies of biological water dynamics and coupling to activity. *Faraday Discuss.*, 141:117–130.

[Fujiwara et al., 2016] Fujiwara, S., Araki, K., Matsuo, T., Yagi, H., Yamada, T., Shibata, K., and Mochizuki, H. (2016). Dynamical behavior of human alpha-synuclein studied by quasielastic neutron scattering. *PLoS ONE*, 11(4):1–17.

[Fujiwara et al., 2013] Fujiwara, S., Yamada, T., Matsuo, T., Takahashi, N., Kamazawa, K., Kawakita, Y., and Shibata, K. (2013). Internal Dynamics of a Protein That Forms the Amyloid Fibrils Observed by Neutron Scattering. *Journal of the Physical Society of Japan*, 82(Suppl.A):SA019.

[Gabel et al., 2002] Gabel, F., Bicout, D., Lehnert, U., Tehei, M., Weik, M., and Zaccai, G. (2002). Protein dynamics studied by neutron scattering. *Q Rev Biophys*, 35(4):327–367.

[Gallat et al., 2012a] Gallat, F.-X., Brogan, A. P. S., Fichou, Y., McGrath, N., Moulin, M., Härtlein, M., Combet, J., Wuttke, J., Mann, S., Zaccai, G., Jackson, C. J., Perriman, A. W., and Weik, M. (2012a). A Polymer Surfactant Corona Dynamically Replaces Water in Solvent-Free Protein Liquids and Ensures Macromolecular Flexibility and Activity. *Journal of the American Chemical Society*, 134(32):13168–13171.

[Gallat et al., 2012b] Gallat, F. X., Laganowsky, A., Wood, K., Gabel, F., van Eijck, L., Wuttke, J., Moulin, M., Härtlein, M., Eisenberg, D., Colletier, J. P., Zaccai, G., and Weik, M. (2012b). Dynamical Coupling of Intrinsically Disordered Proteins and Their Hydration Water: Comparison with Folded Soluble and Membrane Proteins. *Biophysical Journal*, 103(1):129–136.

[Garman and Grime, 2005] Garman, E. F. and Grime, G. W. (2005). Elemental analysis of proteins by microPIXE. *Progress in Biophysics and Molecular Biology*, 89(2):173–205.

[Gasteiger and Marsili, 1980] Gasteiger, J. and Marsili, M. (1980). Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*, 36(22):3219–3228.

REFERENCES

[Glenner et al., 1988] Glenner, G. G., David Eanes, E., and Wiley, C. A. (1988). Amyloid fibrils formed from a segment of the pancreatic islet amyloid protein. *Biochemical and Biophysical Research Communications*, 155(2):608–614.

[Gowers et al., 2016] Gowers, R. J., Linke, M., Barnoud, J., Reddy, T. J. E., Melo, M. N., Seyler, S. L., Domański, J., Dotson, D. L., Buchoux, S., Kenney, I. M., and Beckstein, O. (2016). MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference*, pages 98–105.

[Grimaldo et al., 2019] Grimaldo, M., Lopez, H., Beck, C., Roosen-Runge, F., Moulin, M., Devos, J. M., Laux, V., Härtlein, M., Da Vela, S., Schweins, R., Mariani, A., Zhang, F., Barrat, J.-L., Oettel, M., Forsyth, V. T., Seydel, T., and Schreiber, F. (2019). Protein Short-Time Diffusion in a Naturally Crowded Environment. *The Journal of Physical Chemistry Letters*, 10(8):1709–1715.

[Grimaldo et al., 2015a] Grimaldo, M., Roosen-Runge, F., Hennig, M., Zanini, F., Zhang, F., Zamponi, M., Jalarvo, N., Schreiber, F., and Seydel, T. (2015a). Salt-Induced Universal Slowing Down of the Short-Time Self-Diffusion of a Globular Protein in Aqueous Solution. *The Journal of Physical Chemistry Letters*, 6(13):2577–2582.

[Grimaldo et al., 2015b] Grimaldo, M., Roosen-Runge, F., Jalarvo, N., Zamponi, M., Zanini, F., Hennig, M., Zhang, F., Schreiber, F., and Seydel, T. (2015b). High-resolution neutron spectroscopy on protein solution samples. *EPJ Web of Conferences*, 83:02005.

[Grimaldo et al., 19ed] Grimaldo, M., Roosen-Runge, F., Zhang, F., Schreiber, F., and Seydel, T. (2019/ed). Dynamics of proteins in solution. *Quarterly Reviews of Biophysics*, 52.

[Grimaldo et al., 2014] Grimaldo, M., Roosen-Runge, F., Zhang, F., Seydel, T., and Schreiber, F. (2014). Diffusion and Dynamics of γ-Globulin in Crowded Aqueous Solutions. *The Journal of Physical Chemistry B*, 118(25):7203–7209.

[Hankiewicz and Swierczek, 1974] Hankiewicz, J. and Swierczek, E. (1974). Lysozyme in human body fluids. *Clinica Chimica Acta*, 57(3):205–209.

[Hartmann et al., 1982] Hartmann, H., Parak, F., Steigemann, W., Petsko, G. A., Ponzi, D. R., and Frauenfelder, H. (1982). Conformational substates in a protein:

Structure and dynamics of metmyoglobin at 80 K. *Proceedings of the National Academy of Sciences*, 79(16):4967–4971.

[Hayter and Mook, 1989] Hayter, J. B. and Mook, H. A. (1989). Discrete thin-film multilayer design for X-ray and neutron supermirrors. *Journal of Applied Crystallography*, 22(1):35–41.

[Hellstrand et al., 2013] Hellstrand, E., Nowacka, A., Topgaard, D., Linse, S., and Sparr, E. (2013). Membrane Lipid Co-Aggregation with α-Synuclein Fibrils. *PLoS ONE*, 8(10):e77235.

[Heneka et al., 2014] Heneka, M. T., Kummer, M. P., and Latz, E. (2014). Innate immune activation in neurodegenerative disease. *Nature reviews. Immunology*, 14(July):463–477.

[Hennig et al., 2011] Hennig, M., Frick, B., Seydel, T., and IUCr (2011). Optimum velocity of a phase-space transformer for cold-neutron backscattering spectroscopy. *Journal of Applied Crystallography*, 44(3):467–472.

[Héon et al., 1999] Héon, E., Priston, M., Schorderet, D. F., Billingsley, G. D., Girard, P. O., Lubsen, N., and Munier, F. L. (1999). The gamma-crystallins and human cataracts: A puzzle made clearer. *American journal of human genetics*, 65(5):1261–1267.

[Herranz-Trillo et al., 2016] Herranz-Trillo, F., Groenning, M., van Maarschalkerweerd, A., Tauler, R., Vestergaard, B., and Bernadó, P. (2016). Structural Analysis of Multi-component Amyloid Systems by Chemometric SAXS Data Decomposition. *Structure*, pages 1–11.

[Hogg et al., 2018] Hogg, E., Athreya, K., Basile, C., Tan, E. E., Kaminski, J., and Tagliati, M. (2018). High Prevalence of Undiagnosed Insulin Resistance in Non-Diabetic Subjects with Parkinson's Disease. *Journal of Parkinson's Disease*, 8(2):259–265.

[Holm et al., 1996] Holm, R. H., Kennepohl, P., and Solomon, E. I. (1996). Structural and Functional Aspects of Metal Sites in Biology. *Chemical Reviews*, 96(7):2239–2314.

[Hong et al., 2013] Hong, L., Glass, D. C., Nickels, J. D., Perticaroli, S., Yi, Z., Tyagi, M., O'Neill, H., Zhang, Q., Sokolov, A. P., and Smith, J. C. (2013). Elastic and Conformational Softness of a Globular Protein. *Physical Review Letters*, 110(2):028104.

REFERENCES

[Hoover, 1985] Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697.

[Huang et al., 2013] Huang, J., MacKerell, A. D., and Jr. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry*, 34(25):2135–45.

[Huang et al., 2016] Huang, K.-Y., Kingsley, C. N., Sheil, R., Cheng, C.-Y., Bierma, J. C., Roskamp, K. W., Khago, D., Martin, R. W., and Han, S. (2016). Stability of Protein-Specific Hydration Shell on Crowding. *Journal of the American Chemical Society*, 138(16):5392–5402.

[Humphrey et al., 1996] Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38.

[Iqbal et al., 2010] Iqbal, K., Liu, F., Gong, C.-X., and Grundke-Iqbal, I. (2010). Tau in Alzheimer disease and related tauopathies. *Current Alzheimer research*, 7(8):656–64.

[Jasnin et al., 2010] Jasnin, M., Stadler, A., Tehei, M., and Zaccai, G. (2010). Specific cellular water dynamics observed in vivo by neutron scattering and NMR. *Physical Chemistry Chemical Physics*, 12(35):10154–10160.

[Jiang et al., 2011] Jiang, Y., Shi, K., Xia, D., Wang, S., Song, T., and Cui, F. (2011). Protein Spherulites for Sustained Release of Interferon: Preparation, Characterization and in vivo Evaluation. *Journal of Pharmaceutical Sciences*, 100(5):1913–1922.

[Johansson et al., 1970] Johansson, T. B., Akselsson, R., and Johansson, S. A. E. (1970). X-ray analysis: Elemental trace analysis at the 10-12 g level. *Nuclear Instruments and Methods*, 84(1):141–143.

[Jones et al., 2012] Jones, E. M., Dubey, M., Camp, P. J., Vernon, B. C., Biernat, J., Mandelkow, E., Majewski, J., and Chi, E. Y. (2012). Interaction of Tau protein with model lipid membranes induces tau structural compaction and membrane disruption. *Biochemistry*, 51(12):2539–2550.

[Jorgensen et al., 1983] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935.

[Karikari et al., 2019] Karikari, T. K., Nagel, D. A., Grainger, A., Clarke-Bland, C., Hill, E. J., and Moffat, K. G. (2019). Preparation of stable tau oligomers for cellular and biochemical studies. *Analytical Biochemistry*, 566:67–74.

[Karplus and Kuriyan, 2005] Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences*, 102(19):6679–6685.

[Kauzmann, 1959] Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. In *Advances in Protein Chemistry*, volume 14, pages 1–63. Elsevier.

[Kermani et al., 2018] Kermani, Z. R., Haghighi, S. S., Hajihosseinali, S., Fashami, A. Z., Akbaritouch, T., Akhtari, K., Shahpasand, K., and Falahati, M. (2018). Aluminium oxide nanoparticles induce structural changes in tau and cytotoxicity of the neuroblastoma cell line. *International Journal of Biological Macromolecules*, 120:1140–1148.

[Khago et al., 2016] Khago, D., Wong, E. K., Kingsley, C. N., Alfredo Freites, J., Tobias, D. J., and Martin, R. W. (2016). Increased hydrophobic surface exposure in the cataract-related G18V variant of human γS-crystallin. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1860(1):325–332.

[Kieffer and Wright, 2013] Kieffer, J. and Wright, J. P. (2013). PyFAI: A Python library for high performance azimuthal integration on GPU. *Powder Diffraction*, 28(S2):S339–S350.

[Kirschner et al., 1986] Kirschner, D. A., Abraham, C., and Selkoe, D. J. (1986). X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-beta conformation. *Proceedings of the National Academy of Sciences*, 83(2):503–507.

[Kirstein et al., 2000] Kirstein, O., Prager, M., Kozielewski, T., and Richter, D. (2000). Phase space transformation used at the FRM II backscattering spectrometer: Concepts and technical realization. *Physica B: Condensed Matter*, 283(4):361–364.

[Kneller, 2018] Kneller, G. R. (2018). Franck–Condon picture of incoherent neutron scattering. *Proceedings of the National Academy of Sciences*, 115(38):9450–9455.

REFERENCES

[Kneller and Calandrini, 2007] Kneller, G. R. and Calandrini, V. (2007). Estimating the influence of finite instrumental resolution on elastic neutron scattering intensities from proteins. *Journal of Chemical Physics*, 126(12).

[Kneller and Chevrot, 2012] Kneller, G. R. and Chevrot, G. (2012). Impact of anisotropic atomic motions in proteins on powder-averaged incoherent neutron scattering intensities. *The Journal of Chemical Physics*, 137(22):225101.

[Kneller and Hinsen, 2009] Kneller, G. R. and Hinsen, K. (2009). Quantitative model for the heterogeneity of atomic position fluctuations in proteins: A simulation study. *Journal of Chemical Physics*, 131(4):1–6.

[Koppel, 1972] Koppel, D. E. (1972). Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants. *The Journal of Chemical Physics*, 57(11):4814–4820.

[Krebs et al., 2005] Krebs, M., Bromley, E., Rogers, S., and Donald, A. (2005). The Mechanism of Amyloid Spherulite Formation by Bovine Insulin. *Biophysical Journal*, 88(3):2013–2021.

[Krebs et al., 2004] Krebs, M. R. H., MacPhee, C. E., Miller, A. F., Dunlop, I. E., Dobson, C. M., and Donald, A. M. (2004). The formation of spherulites by amyloid fibrils of bovine insulin. *Proceedings of the National Academy of Sciences*, 101(40):14420–14424.

[Kuranda and Mingozzi, 2017] Kuranda, K. and Mingozzi, F. (2017). AAV Vector-Based Gene Therapy, Progress and Current Challenges. In Brunetti-Pierri, N., editor, *Safety and Efficacy of Gene-Based Therapeutics for Inherited Disorders*, pages 77–112. Springer International Publishing, Cham.

[Laganowsky et al., 2012] Laganowsky, A., Liu, C., Sawaya, M. R., Whitelegge, J. P., Park, J., Zhao, M., Pensalfini, A., Soriaga, A. B., Landau, M., Teng, P. K., Cascio, D., Glabe, C., and Eisenberg, D. (2012). Atomic View of a Toxic Amyloid Small Oligomer. *Science*, 335(6073):1228–1231.

[Landau et al., 2016] Landau, S., Harris, V., Burn, D. J., Hindle, J. V., Hurt, C. S., Samuel, M., Wilson, K. C., and Brown, R. G. (2016). Anxiety and anxious-depression in Parkinson's disease over a 4-year period: A latent transition analysis. *Psychological Medicine*, 46(3):657–667.

[Langmuir, 1939] Langmuir, I. (1939). The structure of proteins. *Proceedings of the Physical Society*, 51(4):592–612.

[Lasagna-Reeves et al., 2010] Lasagna-Reeves, C. A., Castillo-Carranza, D. L., Guerrero-Muñoz, M. J., Jackson, G. R., and Kayed, R. (2010). Preparation and Characterization of Neurotoxic Tau Oligomers. *Biochemistry*, 49(47):10039–10041.

[Lautenschläger et al., 2018] Lautenschläger, J., Stephens, A. D., Fusco, G., Ströhl, F., Curry, N., Zacharopoulou, M., Michel, C. H., Laine, R., Nespovitaya, N., Fantham, M., Pinotsi, D., Zago, W., Fraser, P., Tandon, A., St George-Hyslop, P., Rees, E., Phillips, J. J., De Simone, A., Kaminski, C. F., and Schierle, G. S. K. (2018). C-terminal calcium binding of α-synuclein modulates synaptic vesicle interaction. *Nature Communications*, 9(1):712.

[Lázaro et al., 2014] Lázaro, D. F., Rodrigues, E. F., Langohr, R., Shahpasandzadeh, H., Ribeiro, T., Guerreiro, P., Gerhardt, E., Kröhnert, K., Klucken, J., Pereira, M. D., Popova, B., Kruse, N., Mollenhauer, B., Rizzoli, S. O., Braus, G. H., Danzer, K. M., and Outeiro, T. F. (2014). Systematic Comparison of the Effects of Alpha-synuclein Mutations on Its Oligomerization and Aggregation. *PLoS Genetics*, 10(11):e1004741.

[Leach, 2001] Leach, A. R. (2001). *Molecular Modelling: Principles and Applications, 2nd Edition.* Pearson.

[Levitt and Sharon, 1988] Levitt, M. and Sharon, R. (1988). Accurate simulation of protein dynamics in solution. *Proceedings of the National Academy of Sciences*, 85(20):7557–7561.

[Li et al., 2018a] Li, B., Ge, P., Murray, K. A., Sheth, P., Zhang, M., Nair, G., Sawaya, M. R., Shin, W. S., Boyer, D. R., Ye, S., Eisenberg, D. S., Zhou, Z. H., and Jiang, L. (2018a). Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. *Nature communications*, 9(1):3609.

[Li et al., 2018b] Li, Y., Zhao, C., Luo, F., Liu, Z., Gui, X., Luo, Z., Zhang, X., Li, D., Liu, C., and Li, X. (2018b). Amyloid fibril structure of α-synuclein determined by cryo-electron microscopy. *Cell Research*, 28(9):897–903.

[Li, 2014] Li, Y. V. (2014). Zinc and insulin in pancreatic beta-cells. *Endocrine*, 45(2):178–189.

[Liddle, 2018] Liddle, R. A. (2018). Parkinson's disease from the gut. *Brain Research*, 1693:201–206.

REFERENCES

[Likić and Prendergast, 2001] Likić, V. A. and Prendergast, F. G. (2001). Dynamics of internal water in fatty acid binding protein: Computer simulations and comparison with experiments. *Proteins: Structure, Function, and Bioinformatics*, 43(1):65–72.

[Lim and Lang, 2010] Lim, S.-Y. and Lang, A. E. (2010). The nonmotor symptoms of Parkinson's disease—An overview. *Movement Disorders*, 25(S1):S123–S130.

[Liu et al., 2017] Liu, Z., Huang, J., Tyagi, M., O'Neill, H., Zhang, Q., Mamontov, E., Jain, N., Wang, Y., Zhang, J., Smith, J. C., and Hong, L. (2017). Dynamical Transition of Collective Motions in Dry Proteins. *Physical Review Letters*, 119(4):048101.

[M. Stadler et al., 2016] M. Stadler, A., Demmel, F., Ollivier, J., and Seydel, T. (2016). Picosecond to nanosecond dynamics provide a source of conformational entropy for protein folding. *Physical Chemistry Chemical Physics*, 18(31):21527–21538.

[Madhav et al., 1996] Madhav, T. R., Vatsala, S., Ramakrishna, T., Ramesh, J., and Easwaran, K. R. (1996). Preservation of native conformation during aluminium-induced aggregation of tau protein. *Neuroreport*, 7(5):1072–1076.

[Makarov et al., 1998] Makarov, V. A., Feig, M., Andrews, B. K., and Pettitt, B. M. (1998). Diffusion of Solvent around Biomolecular Solutes: A Molecular Dynamics Simulation Study. *Biophysical Journal*, 75(1):150–158.

[Mantyh et al., 1993] Mantyh, P. W., Ghilardi, J. R., Rogers, S., DeMaster, E., Allen, C. J., Stimson, E. R., and Maggio, J. E. (1993). Aluminum, Iron, and Zinc Ions Promote Aggregation of Physiological Concentrations of ?-Amyloid Peptide. *Journal of Neurochemistry*, 61(3):1171–1174.

[Martel et al., 2017] Martel, A., Antony, L., Gerelli, Y., Porcar, L., Fluitt, A., Hoffmann, K., Kiesel, I., Vivaudou, M., Fragneto, G., and De Pablo, J. J. (2017). Membrane Permeation versus Amyloidogenicity: A Multitechnique Study of Islet Amyloid Polypeptide Interaction with Model Membranes. *Journal of the American Chemical Society*, 139(1):137–148.

[Martinez et al., 2016] Martinez, N., Michoud, G., Cario, A., Ollivier, J., Franzetti, B., Jebbar, M., Oger, P., and Peters, J. (2016). High protein flexibility and reduced hydration water dynamics are key pressure adaptive strategies in prokaryotes. *Scientific Reports*, 6(1):32816.

REFERENCES

[Martyna et al., 1994] Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5):4177–4189.

[Matsumoto et al., 2002] Matsumoto, M., Saito, S., and Ohmine, I. (2002). Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing. *Nature*, 416(6879):409–413.

[Michaud-Agrawal et al., 2011] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. (2011). MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, 32(10):2319–2327.

[Mirsky and Pauling, 1936] Mirsky, A. E. and Pauling, L. (1936). On the Structure of Native, Denatured, and Coagulated Proteins. *Proceedings of the National Academy of Sciences*, 22(7):439–447.

[Moore et al., 2010] Moore, A. H., Bigbee, M. J., Boynton, G. E., Wakeham, C. M., Rosenheim, H. M., Staral, C. J., Morrissey, J. L., and Hund, A. K. (2010). Non-Steroidal Anti-Inflammatory Drugs in Alzheimer's Disease and Parkinson's Disease: Reconsidering the Role of Neuroinflammation. *Pharmaceuticals (Basel, Switzerland)*, 3(6):1812–1841.

[Mukherjee et al., 2017] Mukherjee, S., Mondal, S., and Bagchi, B. (2017). Distinguishing dynamical features of water inside protein hydration layer: Distribution reveals what is hidden behind the average. *The Journal of Chemical Physics*, 147(2):024901.

[Nilsson, 2004] Nilsson, M. (2004). Techniques to study amyloid fibril formation in vitro. *Methods*, 34(1):151–160.

[Nosé, 1984] Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519.

[Oertel and Schulz, 2016] Oertel, W. and Schulz, J. B. (2016). Current and experimental treatments of Parkinson disease: A guide for neuroscientists. *Journal of Neurochemistry*, 139(S1):325–337.

[Olivares et al., 2012] Olivares, D., K. Deshpande, V., Shi, Y., K. Lahiri, D., H. Greig, N., T. Rogers, J., and Huang, X. (2012). N-Methyl D-Aspartate (NMDA) Receptor Antagonists and Memantine Treatment for Alzheimer's Disease, Vascular Dementia and Parkinson's Disease. *Current Alzheimer Research*, 9(6):746–758.

[Omta, 2003] Omta, A. W. (2003). Negligible Effect of Ions on the Hydrogen-Bond Structure in Liquid Water. *Science*, 301(5631):347–349.

[Ortega et al., 2011] Ortega, A., Amorós, D., and García de la Torre, J. (2011). Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal*, 101(4):892–898.

[Oshima et al., 2013] Oshima, E., Ishihara, T., Yokota, O., Nakashima-Yasuda, H., Nagao, S., Ikeda, C., Naohara, J., Terada, S., and Uchitomi, Y. (2013). Accelerated Tau Aggregation, Apoptosis and Neurological Dysfunction Caused by Chronic Oral Administration of Aluminum in a Mouse Model of Tauopathies. *Brain Pathology*, 23(6):633–644.

[Pansieri et al., 2018] Pansieri, J., A. Halim, M., Vendrely, C., Dumoulin, M., Legrand, F., Moulin Sallanon, M., Chierici, S., Denti, S., Dagany, X., Dugourd, P., Marquette, C., Antoine, R., and Forge, V. (2018). Mass and charge distributions of amyloid fibers involved in neurodegenerative diseases: Mapping heterogeneity and polymorphism. *Chemical Science*, 9(10):2791–2796.

[Pansieri et al., 2019] Pansieri, J., Josserand, V., Lee, S.-J., Rongier, A., Imbert, D., Sallanon, M. M., Kövari, E., Dane, T. G., Vendrely, C., Chaix-Pluchery, O., Guidetti, M., Vollaire, J., Fertin, A., Usson, Y., Rannou, P., Coll, J.-L., Marquette, C., and Forge, V. (2019). Ultraviolet–visible–near-infrared optical properties of amyloid fibrils shed light on amyloidogenesis. *Nature Photonics*, 13:473–479.

[Parak and Frauenfelder, 1993] Parak, F. and Frauenfelder, H. (1993). Protein dynamics. *Physica A: Statistical Mechanics and its Applications*, 201(1):332–345.

[Perticaroli et al., 2017] Perticaroli, S., Ehlers, G., Stanley, C. B., Mamontov, E., O'Neill, H., Zhang, Q., Cheng, X., Myles, D. A. A., Katsaras, J., and Nickels, J. D. (2017). Description of Hydration Water in Protein (Green Fluorescent Protein) Solution. *Journal of the American Chemical Society*, 139(3):1098–1105.

[Perutz et al., 1960] Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. T. (1960). Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature*, 185(4711):416.

[Peters and Kneller, 2013] Peters, J. and Kneller, G. R. (2013). Motional heterogeneity in human acetylcholinesterase revealed by a non-Gaussian model for elastic incoherent neutron scattering. *The Journal of Chemical Physics*, 139(16):165102.

[Phillips et al., 2005] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802.

[Popov et al., 2019] Popov, K. I., Makepeace, K. A. T., Petrotchenko, E. V., Dokholyan, N. V., and Borchers, C. H. (2019). Insight into the Structure of the “Unstructured” Tau Protein. *Structure*, 27(11):1710–1715.e4.

[Press et al., 1992] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing. Second Edition.*

[Rahaman et al., 2017] Rahaman, O., Kalimeri, M., Katava, M., Paciaroni, A., and Sterpone, F. (2017). Configurational Disorder of Water Hydrogen-Bond Network at the Protein Dynamical Transition. *The Journal of Physical Chemistry B*, 121(28):6792–6798.

[Rehr and Albers, 2000] Rehr, J. J. and Albers, R. C. (2000). Theoretical approaches to x-ray absorption fine structure. *Reviews of Modern Physics*, 72(3):621–654.

[Ristow, 2004] Ristow, M. (2004). Neurodegenerative disorders associated with diabetes mellitus. *Journal of Molecular Medicine*, 82(8):510–529.

[Roh et al., 2005] Roh, J. H., Novikov, V. N., Gregory, R. B., Curtis, J. E., Chowdhuri, Z., and Sokolov, A. P. (2005). Onsets of Anharmonicity in Protein Dynamics. *Physical Review Letters*, 95(3):038101.

[Roosen-Runge et al., 2011] Roosen-Runge, F., Hennig, M., Zhang, F., Jacobs, R. M. J., Sztucki, M., Schober, H., Seydel, T., and Schreiber, F. (2011). Protein self-diffusion in crowded solutions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29):11815–11820.

[Roosen-Runge and Seydel, 2015] Roosen-Runge, F. and Seydel, T. (2015). A generalized mean-squared displacement from inelastic fixed window scans of incoherent neutron scattering as a model-free indicator of anomalous diffusion confinement. *EPJ Web of Conferences*, 83:02015.

[Roosen-Runge et al., 2014] Roosen-Runge, F., Zhang, F., Schreiber, F., and Roth, R. (2014). Ion-activated attractive patches as a mechanism for controlled protein interactions. *Scientific Reports*, 4(1):1–5.

[Roskamp et al., 2017] Roskamp, K. W., Montelongo, D. M., Anorma, C. D., Bandak, D. N., Chua, J. A., Malecha, K. T., and Martin, R. W. (2017). Multiple Aggregation Pathways in Human γS-Crystallin and Its Aggregation-Prone G18V Variant. *Investigative Ophthalmology & Visual Science*, 58(4):2397–2405.

[Rupley and Careri, 1991] Rupley, J. A. and Careri, G. (1991). Protein Hydration and Function. In Anfinsen, C. B., Richards, F. M., Edsall, J. T., and Eisenberg, D. S., editors, *Advances in Protein Chemistry*, volume 41, pages 37–172. Academic Press.

[Sanz et al., 2018] Sanz, A., Hansen, H. W., Jakobsen, B., Pedersen, I. H., Capaccioli, S., Adrjanowicz, K., Paluch, M., Gonthier, J., Frick, B., Lelièvre-Berna, E., Peters, J., and Niss, K. (2018). High-pressure cell for simultaneous dielectric and neutron spectroscopy. *Review of Scientific Instruments*, 89(2):023904.

[Sanz et al., 2013] Sanz, E., Vega, C., Espinosa, J. R., Caballero-Bernal, R., Abascal, J. L. F., and Valeriani, C. (2013). Homogeneous Ice Nucleation at Moderate Supercooling from Molecular Simulation. *Journal of the American Chemical Society*, 135(40):15008–15017.

[Sayers, 1971] Sayers, D. E. (1971). New Technique for Investigating Noncrystalline Structures: Fourier Analysis of the Extended X-Ray—Absorption Fine Structure. *Physical Review Letters*, 27(18):1204–1207.

[Schelten and Alefeld, 1984] Schelten, J. and Alefeld, B. (1984). Backscattering Spectrometer with Adapted Q-Resolution at the Pulsed Neutron Source. *Proceedings of the Workshop on Neutron Scattering Instrumentation for SNQ*, pages 378–389.

[Schiró et al., 2010] Schiró, G., Caronna, C., Natali, F., and Cupane, A. (2010). Direct evidence of the amino acid side chain and backbone contributions to protein anharmonicity. *Journal of the American Chemical Society*, 132(4):1371–1376.

[Schirò et al., 2011] Schirò, G., Caronna, C., Natali, F., Koza, M. M., and Cupane, A. (2011). The "protein dynamical transition" does not require the protein polypeptide chain. *Journal of Physical Chemistry Letters*, 2(18):2275–2279.

[Schirò et al., 2015] Schirò, G., Fichou, Y., Gallat, F.-X., Wood, K., Gabel, F., Moulin, M., Härtlein, M., Heyden, M., Colletier, J.-P., Orecchini, A., Paciaroni, A., Wuttke, J., Tobias, D. J., and Weik, M. (2015). Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nature communications*, 6:6490.

[Schiró et al., 2012] Schiró, G., Vetri, V., Frick, B., Militello, V., Leone, M., and Cupane, A. (2012). Neutron scattering reveals enhanced protein dynamics in concanavalin a amyloid fibrils. *Journal of Physical Chemistry Letters*, 3(8):992–996.

[Sciortino et al., 1991] Sciortino, F., Geiger, A., and Stanley, H. E. (1991). Effect of defects on molecular mobility in liquid water. *Nature*, 354(6350):218.

[Sciortino et al., 1992] Sciortino, F., Geiger, A., and Stanley, H. E. (1992). Network defects and molecular mobility in liquid water. *The Journal of Chemical Physics*, 96(5):3857–3865.

[Scott et al., 1993] Scott, C. W., Fieles, A., Sygowski, L. A., and Caputo, C. B. (1993). Aggregation of tau protein by aluminum. *Brain Research*, 628(1):77–84.

[Sears, 1966a] Sears, V. F. (1966a). Theory of cold neutron scattering by homonuclear diatomic liquids: I. free rotation. *Canadian Journal of Physics*, 44(6):1279–1297.

[Sears, 1966b] Sears, V. F. (1966b). Theory of cold neutron scattering by homonuclear liquid: Ii. hindered rotation. *Canadian Journal of Physics*, 44(6):1299–1311.

[Sears, 1967] Sears, V. F. (1967). Cold neutron scattering by molecular liquids: Iii. methane. *Canadian Journal of Physics*, 45(2):237–254.

[Shafiei et al., 2017] Shafiei, S. S., Guerrero-Muñoz, M. J., and Castillo-Carranza, D. L. (2017). Tau Oligomers: Cytotoxicity, Propagation, and Mitochondrial Damage. *Frontiers in Aging Neuroscience*, 9:83.

[Shikama et al., 2010] Shikama, Y., Kitazawa, J.-i., Yagihashi, N., Uehara, O., Murata, Y., Yajima, N., Wada, R., and Yagihashi, S. (2010). Localized Amyloidosis at the Site of Repeated Insulin Injection in a Diabetic Patient. *Internal Medicine*, 49(5):397–401.

[Shin and Willard, 2018a] Shin, S. and Willard, A. P. (2018a). Characterizing Hydration Properties Based on the Orientational Structure of Interfacial Water Molecules. *Journal of Chemical Theory and Computation*, 14(2):461–465.

[Shin and Willard, 2018b] Shin, S. and Willard, A. P. (2018b). Three-Body Hydrogen Bond Defects Contribute Significantly to the Dielectric Properties of the Liquid Water–Vapor Interface. *The Journal of Physical Chemistry Letters*, 9(7):1649–1654.

[Singwi and Sjölander, 1960] Singwi, K. S. and Sjölander, A. (1960). Diffusive Motions in Water and Cold Neutron Scattering. *Physical Review*, 119(3):863–871.

[Sivia et al., 1992] Sivia, D., Carlile, C., and Howells, W. (1992). *Bayesian Analysis of Quasielastic Neutron Scattering Data*.

[Skamris et al., 2019] Skamris, T., Marasini, C., Madsen, K. L., Foderà, V., and Vestergaard, B. (2019). Early Stage Alpha-Synuclein Amyloid Fibrils are Reservoirs of Membrane-Binding Species. *Scientific Reports*, 9(1):1733.

[Smith et al., 2012] Smith, M. I., Foderà, V., Sharp, J. S., Roberts, C. J., and Donald, A. M. (2012). Factors affecting the formation of insulin amyloid spherulites. *Colloids and Surfaces B: Biointerfaces*, 89(1):216–222.

[Snow et al., 2017] Snow, W. M., Dale, R., Buist, R., Peirson, D., Martin, M., and Albensi, B. C. (2017). In Vivo. 58:841–853.

[Sonksen and Sonksen, 2000] Sonksen, P. and Sonksen, J. (2000). Insulin: Understanding its action in health and disease. *British Journal of Anaesthesia*, 85(1):69–79.

[Spillantini et al., 1997] Spillantini, M. G., Schmidt, M. L., Lee, V. M.-Y., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997). α-Synuclein in Lewy bodies. *Nature*, 388(6645):839–840.

[Springer, 1977] Springer, T. (1977). Molecular Rotations and Diffusion in Solids, in Particular Hydrogen in Metals. In Lovesey, S. W. and Springer, T., editors, *Dynamics of Solids and Liquids by Neutron Scattering*, Topics in Current Physics, pages 255–300. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Squires, 1996] Squires, G. L. (1996). *Introduction to the Theory of Thermal Neutron Scattering*. N.Y. : Dover publ., Mineola.

[Stefanis et al., 2019] Stefanis, L., Emmanouilidou, E., Pantazopoulou, M., Kirik, D., Vekrellis, K., and Tofaris, G. K. (2019). How is alpha-synuclein cleared from the cell? *Journal of Neurochemistry*, 150(5):577–590.

[Steinbach and Brooks, 1993] Steinbach, P. J. and Brooks, B. R. (1993). Protein hydration elucidated by molecular dynamics simulation. *Proceedings of the National Academy of Sciences*, 90(19):9135–9139.

[Stetefeld et al., 2016] Stetefeld, J., McKenna, S. A., and Patel, T. R. (2016). Dynamic light scattering: A practical guide and applications in biomedical sciences. *Biophysical Reviews*, 8(4):409–427.

[Stoeckli et al., 1986] Stoeckli, A., Furrer, A., Schoenenberger, C., Meier, B. H., Ernst, R. R., and Anderson, I. (1986). Dynamics of hydrogen bonds in carboxylic acids. *Physica B+C*, 136(1):161–164.

[Stradner et al., 2004] Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C. K., Egelhaaf, S. U., and Schurtenberger, P. (2004). Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432(7016):492–495.

[Svergun et al., 1998] Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, Z., Kuprin, S., and Zaccai, G. (1998). Protein hydration in solution: Experimental observation by x-ray and neutron scattering. *Proceedings of the National Academy of Sciences*, 95(5):2267–2272.

[Swift et al., 2002] Swift, B., Hawkins, P. N., Richards, C., and Gregory, R. (2002). Examination of insulin injection sites: An unexpected finding of localized amyloidosis [1]. *Diabetic Medicine*, 19(10):881–882.

[Tamburrino and Decressac, 2016] Tamburrino, A. and Decressac, M. (2016). Aged and Diseased Neurons Get Lost in Transport. *Trends in Neurosciences*, 39(4):199–201.

[Tarek and Tobias, 2000] Tarek, M. and Tobias, D. J. (2000). The Dynamics of Protein Hydration Water: A Quantitative Comparison of Molecular Dynamics Simulations and Neutron-scattering Experiments. *Biophysical Journal*, 79(6):3244–3257.

[Tarek and Tobias, 2002] Tarek, M. and Tobias, D. J. (2002). Single-Particle and Collective Dynamics of Protein Hydration Water: A Molecular Dynamics Study. *Physical Review Letters*, 89(27):275501.

[Tarek and Tobias, 2008] Tarek, M. and Tobias, D. J. (2008). The role of protein-solvent hydrogen bond dynamics in the structural relaxation of a protein in glycerol versus water. *European Biophysics Journal*, 37(5):701–709.

REFERENCES

[Tayeb-Fligelman et al., 2017] Tayeb-Fligelman, E., Tabachnikov, O., Moshe, A., Goldshmidt-Tran, O., Sawaya, M. R., Coquelle, N., Colletier, J.-P., and Landau, M. (2017). The cytotoxic Staphylococcus aureus PSMα3 reveals a cross-α amyloid-like fibril. *Science*, 355(6327):831–833.

[Theillet et al., 2016] Theillet, F.-x., Binolfi, A., Bekei, B., Martorana, A., Rose, H. M., Stuiver, M., Verzini, S., Lorenz, D., Rossum, M. V., Goldfarb, D., and Selenko, P. (2016). Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature*, 530(7588):45–50.

[Tokuyama and Oppenheim, 1994] Tokuyama, M. and Oppenheim, I. (1994). Dynamics of hard sphere suspensions. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 50(1):1–16.

[Tomljenovic, 2011] Tomljenovic, L. (2011). Aluminum and Alzheimer's Disease: After a Century of Controversy, Is there a Plausible Link? *Journal of Alzheimer's Disease*, 23(4):567–598.

[Tong, 2020] Tong, D. (2020). University of Cambridge - David Tong's lectures - http://www.damtp.cam.ac.uk/user/tong/statphys.html - last access on 02/2020.

[Tournier et al., 2003] Tournier, A. L., Xu, J., and Smith, J. C. (2003). Translational Hydration Water Dynamics Drives the Protein Glass Transition. *Biophysical Journal*, 85(3):1871–1875.

[Tuckerman et al., 1992] Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992). Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, 97(3):1990–2001.

[Tuckerman and Martyna, 2000] Tuckerman, M. E. and Martyna, G. J. (2000). Understanding Modern Molecular Dynamics: Techniques and Applications. *The Journal of Physical Chemistry B*, 104(2):159–178.

[Tuttle et al., 2016] Tuttle, M. D., Comellas, G., Nieuwkoop, A. J., Covell, D. J., Berthold, D. A., Kloepper, K. D., Courtney, J. M., Kim, J. K., Barclay, A. M., Kendall, A., Wan, W., Stubbs, G., Schwieters, C. D., Lee, V. M. Y., George, J. M., and Rienstra, C. M. (2016). Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nature Structural & Molecular Biology*, 23(February):1–9.

[Uversky, 2011] Uversky, V. N. (2011). Intrinsically disordered proteins from A to Z. *The International Journal of Biochemistry & Cell Biology*, 43(8):1090–1103.

[Vendra et al., 2016] Vendra, V. P. R., Khan, I., Chandani, S., Muniyandi, A., and Balasubramanian, D. (2016). Gamma crystallins of the human eye lens. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1860(1):333–343.

[Verlet, 1967] Verlet, L. (1967). Computer ”Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review*, 159(1):98–103.

[Vetri and Foderà, 2015] Vetri, V. and Foderà, V. (2015). The route to protein aggregate superstructures: Particulates and amyloid-like spherulites. *FEBS Letters*, 589(19):2448–2463.

[Villaverde and Carrió, 2003] Villaverde, A. and Carrió, M. M. (2003). Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. (25):1385.

[Volino and Dianoux, 1980] Volino, F. and Dianoux, A. (1980). Neutron incoherent scattering law for diffusion in a potential of spherical symmetry: General formalism and application to diffusion inside a sphere. *Molecular Physics*, 41(2):271–279.

[Wales and Doye, 1997] Wales, D. J. and Doye, J. P. K. (1997). Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116.

[Walton et al., 2011] Walton, J. H., Berry, R. S., and Despa, F. (2011). Amyloid Oligomer Formation Probed by Water Proton Magnetic Resonance Spectroscopy. *Biophysical Journal*, 100(9):2302–2308.

[Walton, 2010] Walton, J. R. (2010). Evidence for Participation of Aluminum in Neurofibrillary Tangle Formation and Growth in Alzheimer’s Disease. *Journal of Alzheimer’s Disease*, 22(1):65–72.

[Welch et al., 2002] Welch, K. D., Davis, T. Z., Van Eden, M. E., and Aust, S. D. (2002). Deleterious iron-mediated oxidation of biomolecules11This article is part of a series of reviews on “Iron and Cellular Redox Status.” The full list of papers may be found on the homepage of the journal.66Guest Editor: Mario Comporti. *Free Radical Biology and Medicine*, 32(7):577–583.

[Winner et al., 2011] Winner, B., Jappelli, R., Maji, S. K., Desplats, P. A., Boyer, L., Aigner, S., Hetzer, C., Loher, T., Vilar, M., Campioni, S., Tzitzilonis, C., Soragni, A., Jessberger, S., Mira, H., Consiglio, A., Pham, E., Masliah, E., Gage,

F. H., and Riek, R. (2011). In vivo demonstration that α-synuclein oligomers are toxic. *Proceedings of the National Academy of Sciences*, 108(10):4194–4199.

[Wood et al., 2013] Wood, K., Gallat, F. X., Otten, R., Van Heel, A. J., Lethier, M., Van Eijck, L., Moulin, M., Haertlein, M., Weik, M., and Mulder, F. A. (2013). Protein surface and core dynamics show concerted hydration-dependent activation. *Angewandte Chemie - International Edition*, 52(2):665–668.

[Wood et al., 2008] Wood, K., Paciaroni, A., Moulin, M., Tobias, D. J., and Weik, M. (2008). Coincidence of dynamical transitions in a soluble protein and its hydration-water: Direct measurements by neutron scattering and MD simulations. *J. Mol. Biol*, 130:4586–4587.

[Wood et al., 2007] Wood, K., Plazanet, M., Gabel, F., Kessler, B., Oesterhelt, D., Tobias, D. J., Zaccai, G., and Weik, M. (2007). Coupling of protein and hydration-water dynamics in biological membranes. *Proceedings of the National Academy of Sciences*, 104(46):18049–18054.

[Wood et al., 2010] Wood, K., Tobias, D. J., Kessler, B., Gabel, F., Oesterhelt, D., Mulder, F. A. A., Zaccai, G., and Weik, M. (2010). The low-temperature inflection observed in neutron scattering measurements of proteins is due to methyl rotation: Direct evidence using isotope labeling and molecular dynamics simulations. *Journal of the American Chemical Society*, 132(14):4990–4991.

[Wright and Dyson, 2009] Wright, P. E. and Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, 19(1):31–38.

[Wu et al., 2014] Wu, J. W., Chen, M. E., Wen, W. S., Chen, W. A., Li, C. T., Chang, C. K., Lo, C. H., Liu, H. S., and Wang, S. S. (2014). Comparative analysis of human γD-crystallin aggregation under physiological and low pH conditions. *PLoS ONE*, 9(11):e112309.

[Wuttke, 2012] Wuttke, J. (2012). Absorption-Correction Factors for Scattering from Flat or Tubular Samples: Open-Source Implementation libabsco, and Why it Should be Used with Caution.

[Wuttke et al., 2012] Wuttke, J., Budwig, A., Drochner, M., Kämmerling, H., Kayser, F. J., Kleines, H., Ossovyi, V., Pardo, L. C., Prager, M., Richter, D., Schneider, G. J., Schneider, H., and Staringer, S. (2012). SPHERES, Jülichs high-flux neutron backscattering spectrometer at FRM II. *Review of Scientific Instruments*, 83(7).

[Yi et al., 2012] Yi, Z., Miao, Y., Baudry, J., Jain, N., and Smith, J. C. (2012). Derivation of mean-square displacements for protein dynamics from elastic incoherent neutron scattering. *Journal of Physical Chemistry B*, 116(16):5028–5036.

[Yoshimura et al., 2012] Yoshimura, Y., Lin, Y., Yagi, H., Lee, Y.-H., Kitayama, H., Sakurai, K., So, M., Ogi, H., Naiki, H., and Goto, Y. (2012). Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proceedings of the National Academy of Sciences*, 109(36):14446–14451.

[Zeller et al., 2020] Zeller, D., Tan, P., Hong, L., Di Bari, D., Garcia Sakai, V., and Peters, J. (2020). Differences between calcium rich and depleted alpha-lactalbumin investigated by molecular dynamics simulations and incoherent neutron scattering. *Physical Review E*, 101(3):032415.

[Zeller et al., 2018] Zeller, D., Telling, M. T. F., Zamponi, M., García Sakai, V., and Peters, J. (2018). Analysis of elastic incoherent neutron scattering data beyond the Gaussian approximation. *The Journal of Chemical Physics*, 149(23):234908.

[Zhang et al., 2015] Zhang, H., Rochet, J.-C., and Stanciu, L. A. (2015). Cu(II) promotes amyloid pore formation. *Biochemical and Biophysical Research Communications*, 464(1):342–347.

[Zou et al., 2002] Zou, K., Gong, J.-S., Yanagisawa, K., and Michikawa, M. (2002). A Novel Function of Monomeric Amyloid β-Protein Serving as an Antioxidant Molecule against Metal-Induced Oxidative Damage. *Journal of Neuroscience*, 22(12):4833–4841.