UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Diagnostik

Institut für Neuropathologie

Prof. Dr. Markus Glatzel

Automated SAXS data analysis of the amyloid fibril formation process

Dissertation

zur Erlangung des Doktorgrades Dr. rer. biol. hum.
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Taisiia Cheremnykh
aus Abakan

Hamburg 2022

(wird von der Medizinischen Fakultät ausgefüllt)

Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 05.10.2022

Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Markus Glatzel

Prüfungsausschuss, zweite/r Gutachter/in: Dr. Jan Kosinski

This thesis describes the work performed under the joint supervision of Dr. Dmitri Svergun at the European Molecular Biology Laboratory (EMBL) in Hamburg, Germany, and Prof. Dr. Markus Glatzel at the Universitätsklinikum Hamburg-Eppendorf (UKE).

Contents

List of Figures

List of Tables

1 Introduction

Recent fascinating developments in neuroscience, medicine and structural biology made neurodegenerative diseases and their key molecular mechanisms critical topics in neuropathology and structure-driven drug design (Ciccone et al. (2020); Mouchlis et al. (2020)). Here, amyloid fibrils whose building blocks are peptides or proteins, as monomers or in oligomeric form, are particularly intriguing because they are naturally soluble proteins forming insoluble fibres and can thus accompany a degenerative disease (Pepys (2001); Rambaran and Serpell (2008)).

Time-resolved biophysical studies are of interest in fundamental molecular and structural biology as well as their applications in medicine and pharmaceutical industry. Many diseases are underpinned by biochemical reactions at the molecular level (Liu et al. (2012)). Inhibiting these reactions could prevent the development of the disease and this is widely utilised in molecular medicine (Sausville (2002); Stein and Waterfield (2000)). Revealing structural peculiarities of real-time biochemical reactions and elucidation of intermediate states thereof allows for more targeted drug design (Sun and Scott (2010); Shukla et al. (2014)).

In silico methods such as molecular dynamics simulations (Karplus and Mc-Cammon (2002)) and integrative structural modelling (Koukos and Bonvin (2020)) provide valuable information for testing experimental conditions and investigating drug candidates. Experimental methods are employed for structural and functional characterisation of biomacromolecular systems. Biophysical and biochemical methods are utilised in order to obtain relevant macroscopic parameters and characteristic times of the reactions *in vitro* (Likhtenshtein et al. (2000)). Atomic structures of biological macromolecules and in particular proteins can be determined with macromolecular crystallography (MX), nuclear magnetic resonance (NMR) and cryogenic electron microscopy (cryoEM) (Cowieson et al. (2008)). The latter also allows for single-particle shape reconstruction (Cheng (2015)). Although single-particle methods studying the kinetics of conformational changes revealed a significant progress in recent years (Fischer et al. (2010); Spence (2017); Mäeots

et al. (2020)), investigating dynamics of ensembles of native randomly oriented biomolecules (e.g. in solution) serves as a good approximation of the processes, which could happen in living cells.

The present thesis describes the application of time-resolved solution scattering of X-rays to the analysis of the fibril formation process. The approach is based on the developments of synchrotron radiation facilities and it can further enhance relevant functional studies in molecular biology. By performing structural studies in solution, the sample environment acts as a medium to support and capture interactions and reactions, which provide the path for creating different bioregulatory agents.

1.1 Amyloid beta peptide: misfolding and neurodegeneration

1.1.1 The human brain

The human brain is an organ of the central nervous system. Its major parts are the medulla, cerebellum, midbrain and two cerebral hemispheres also called cerebrum (Carter (2019)). The cerebrum has an inner layer of white matter mainly consisting of myelinated axons and an outer layer is the grey matter of the cerebral cortex. The cerebral cortex amounts to 80% of the human brain's mass. It plays a key role in the cognitive functions of a human (Carter (2019)). The grey matter mostly consists of nerve cells – neurons organised into neural circuits. Together, the neural circuits are forming a neuron network allowing signals to travel through. The nerve cells are traversed by signals that create memories and thoughts in the form of a tiny electrical charge. At synapses, nerve cells are linked to one another. A charge may cause the release of small proportions of chemicals known as neurotransmitters when it reaches a synapse (Alzheimer's Association (2022)). These processes are essential for thinking and cognition.

1.1.2 Alzheimer's disease and role of amyloid beta peptide

Alzheimer's disease (AD) is a neurodegenerative disease which is characterised by memory loss and other cognitive misfunctioning. AD is the most common form of dementia which is the seventh leading cause of death among all diseases (WHO (2021)). AD is associated with loss of neuronal connections in the brain, neurofibrillary tangles and amyloid plaques. Plaques, that are unnatural protein fragment clusters, accumulate between the nerve cells (Alzheimer's Association (2022)).

Tangles, comprised of twisted strands of another protein – amyloid precursor protein (APP), are present in dead and dying nerve cells. APP is a membrane protein abundantly expressed in the brain. Its isoforms range from 365 to 770 amino acid residues. APP is processed by several proteinases resulting in shedding the APP fragments (Zheng and Koo (2006)). A fragment of APP derived by proteolytic cleavage and called amyloid beta (Aβ) peptide aggregates in plaques to form them. The larger protein that makes up the Aβ is present in the fatty membrane that surrounds nerve cells. Plaques are progressively developed as a result of Aβ "stickiness": Aβ undergoes a conformational transition to a predominantly β-sheet structure accompanied by peptide aggregation (Tew et al. (2008)) Rather than just plaques themselves, Aβ clusters of a few Aβ monomers could be the most harmful kind of species. Cell-to-cell interaction at synapses may be obstructed by the small aggregates. They may also stimulate immunological cells, which cause inflammation and destroy diseased cells (Alzheimer's Association (2022)).

Although cognitive impairment could be caused by numerous factors, AD research is mostly focused on the Aβ hypothesis. This hypothesis was suggested 30 years ago (Hardy and Selkoe (2002)). Its notion is that Aβ accumulation and the imbalance between Aβ production and clearance in the brain are driving AD pathogenesis. Overproduced Aβ oligomerises and it is being deposited as diffuse plaques. These oligomers take part in further processes leading to neuritic dysfunction.

1.1.3   Structural kinetics of the amyloid beta peptide

There are two main Aβ species that terminate at Val-40 and Ala-42, and it is now believed that the latter, called Aβ42 is the most significant and earliest species to have been deposited in the cortex (Funato et al. (1998)). Aβ42 is a 42 amino acid peptide with molecular mass about 4,5 kDa (Kuo et al. (1999); Pedersen et al. (1996)). This peptide has propensity to aggregate and extraordinary poor solubility (Snyder et al. (1994)). A model of Aβ fibril based on electron microscopy studies was first suggested by Shirahama and Cohen (1967). This study revealed cross-sectional diameters characteristic of the fibrils and their compounds such as filaments, protofibrils and subprotofibrils. More reliable parameters and structures were made available with the development of cryoEM. The study by Schmidt et al. (2015) describes the dimer interfaces of the Aβ42 in cross-βstructure, that are of 7-10 nm in size. Synthetic protofibrils are found to be up to 150 nm in length and ∼5 nm in width and they have β-sheet structure (Haass and Selkoe (2007)).

Some structural features, as well as kinetics, of the Aβ fibrillogenesis, were

comprehensively reviewed by Teplow (1998).

The review already suggested possible kinetics models for the process of amyloid fibrils formation. Those models, however, are linear pathways and suggest that there might be different ways of conversion of Aβ monomers to fibrils, but not the other way around. Later, Bartolini et al. (2007) discovered that the Aβ fibrillogenesis proceeds in three phases: a lag phase devoid of notable conformation changes, an exponential phase with a rapid increase in β-sheet content, and a plateau phase in which Aβ assumed a predominat β-sheet secondary structure. The authors also stated that during this third phase, aggregation nuclei, amyloid protofibrils, and fibrils were formed as well. Multi-methodological study of the Aβ aggregation kinetics published later by the same group (Bartolini et al. (2011)) revealed the kinetics of conformational changes and described the conversion with the following intermediate species: (i) small transient species; (ii) soluble oligomers of β-sheet conformation and low molecular weight; (iii) protofibrils and oligomers with high molecular weight.

Fodera et al. (2013) presented a theoretical model for the description of a generic aggregate formed from an ensemble of charged proteins. Within this model, the authors explain the formation of multifractal structures emerging in protein aggregation reactions. It also justified large-scale polymorphism of amyloid aggregates controlled by electrostatic interactions.

Recent studies of Aβ42 dynamics showed that along with the Aβ aggregation the oligomers are also dissociating into monomers. Indeed, oligomeric species are heterogeneous and only the minority converts into mature fibrils. This conversion occurs after oligomers undergo formation-dissociation cycles (Michaels et al. (2020)).

The timeline of the oligomerisation kinetics depends on the initial monomer concentration. As suggested in Schmidt et al. (2015), for the 1 mg/ml solution of Aβ42, incubation at room temperature for a minimum of 12 hours is recommended to observe mature fibrils. Aggregation end state might be reached in up to 5 days with a monomer concentration of 350 µM ($\sim$1.6 mg/ml), where the end state amyloid fibrils have a cross-section diameter of around 10 nm as shown for the small-angle X-ray and neutron scattering by Lattanzi et al. (2021). Overall, despite numerous studies, structural description of the process of Aβ aggregation and the characterisation of its intermediates is still a matter of debate.

1.2 Time-resolved small-angle X-ray scattering for structural studies

One of the techniques to characterise the behaviour of macromolecular systems in solution especially powerful for Aβ aggregation studies is small-angle X-ray scattering (SAXS) (Svergun et al. (2013)). The SAXS data can be rapidly collected allowing for time-resolved studies to characterise macromolecular ensembles in solution including heterogeneous evolving systems. Time-resolved small-angle X-ray scattering (TR-SAXS) is one of the most efficient structural methods to study biological processes and complicated systems in solution. It allows one to quantitatively analyse polydisperse mixtures, capture conformational changes and measure the kinetics of the macromolecular transitions and complex formation in near-native conditions (Svergun et al. (2013)).

1.2.1 Small-angle scattering fundamentals

SAXS is a low-resolution diffraction technique. The use of elastic scattering at small angles is employed in various scientific areas such as condensed matter physics, polymer science, molecular biology and biophysics. These methods are based on the use of radiation with a wavelength about a few angstroms, which is comparable with the interatomic distances. To determine the supra-atomic structure of such systems, it is useful to examine the diffraction pattern in the region of small scattering angles. For heterogeneous systems, small-angle scattering is one of the few methods revealing their internal organisation. For solutions of biopolymers, interpretation of the scattering profile requires *a priori* knowledge about the systems under study to obtain comprehensive models (Svergun et al. (2013)). Small-angle neutron scattering (SANS) is also employed in the research, including biological macromolecules, and most of the concepts introduced for SAXS, are also valid for SAXS.

The scattering patterns from dilute solutions of macromolecules are typically isotropic because of the random orientation of particles in solution. They are represented as azimuthally averaged one-dimensional curves of the scattering intensity as a function of the angle. The scattered radiation originates, other than from the macromolecules themselves, also from the buffer and the surrounding container and these unwanted contributions are removed by background subtraction. For monodisperse solutions of identical particles, the subtracted SAXS curves are proportional to the scattering from a single macromolecule averaged over all orientations.

Using SAXS, one can determine the radius of gyration of a particle, which is related to its size, as well as assess the intensity of scattering at zero angle (forward scattering), related to the molecular weight of a macromolecule. SAXS was first applied by André Guinier in 1939 (Guinier (1939)) to study zones of coherent scattering in metal alloys. The scattering intensities are functions defined in the so-called reciprocal (Fourier-transformed) space measured in inverted length units. Analysis of the distribution of radiation intensity near the incident beam (scattered at small angles) provides information about the large distances, i.e. about size and shape of the scattering particles. For monodisperse solutions, after subtracting the signal from the independently measured solvent scattering, the X-ray intensity I close to the primary beam can be described as:

$$I(s) = I_0 \exp\left(-\frac{s^2 R_g^2}{3}\right),$$

(1.1)

where I_0 is the intensity of forward scattering, $s = \frac{4\pi \sin\theta}{\lambda}$ is the momentum transfer, 2θ is the scattering angle, R_g is the radius of gyration of electron density distribution in the particle with respect to the centre of mass.

Equation (1.1) is the so-called Guinier approximation, which is valid for the smallest angles and is utilised to determine the overall dimensions of the particles. The R_g^2 is calculated from the Guinier plot by representing $\ln I(s)$ *vs* s^2 and calculating the tangent of the slope $R_g^2 \sim (\ln I(s))/s^2$. This equation is valid for biological macromolecules in solution as well as for many other substances.

Besides the Guinier approximation and determination of radius of gyration (R_g), other parameters can be directly assessed from the small-angle scattering data. Porod invariant

$$Q = \int\limits_0^\infty s^2 I(s)\, ds$$

(1.2)

is one of the most important integral characteristics of the scattering intensity. It gives the full scattering capacity of the object and, for uniform particles, depends only on the particle volume. Definition of the forward scattering intensity $I(0)$ (cannot be determined experimentally and is obtained by extrapolating $I(s)$ to $s = 0$) allows one to estimate the molecular mass of the particle. Interpretation of the SAXS data could be assisted by computational modelling of the scattering from a set of simple shapes like spheres or cylinders (Svergun et al. (2013)). More detailed shape determination could be done by employing, for instance, modelling tools of the ATSAS software suite (Manalastas-Cantos et al. (2021)).

SAXS arises due to the interactions of the electrons in the sample with the incoming X-rays, and the scattering by an atom is proportional to its molecular mass. In SANS, neutrons interact with matter either via nuclear interactions with nuclei or via interaction with magnetic momentum of unpaired electrons (Feigin and Svergun (1987)). For biological objects, only contribution of nuclear interactions is usually considered. Since the scattering amplitudes of atoms are irregular the method allows for the detection of hydrogens in biological samples in presence of heavier elements. Further, the neutrons scattering amplitudes of hydrogen and deuterium are significantly different such that deuteration can be used for contrast variation, which is widely used for biological solutions. The high penetration depth of electrically neutral neutrons allows for lower intensities to be used and the absence of radiation damage is advantageous for biological samples. In many cases, SANS may provide valuable complementary information to SAXS.

Compared to SANS, SAXS uses higher intensities which allow faster measurements and needs a lower quantity of samples, although at a risk of radiation damage. SAXS is an elastic scattering of X-rays (photons emitted by accelerated electrons) on heterogeneities of matter close to the incident X-ray beam. SAXS studies of the structure of matter rely on the interaction of X-rays with the electrons of the matter. Such interactions depend on the presence and features of heterogeneities of electron density if the sizes thereof are much larger than the wavelength of the X-ray beam. In SAXS, an X-ray beam interacts with the electrons of the sample and, then, scatters. The scattering by an atom is proportional to the number of electrons, such that heavier atoms are better "seen" by this method. Scattered radiation is recorded by the X-ray detector.

Various small-angle X-ray facilities are available. The sources of X-rays are either X-ray tubes (radiation wavelength $\lambda = 0.07 \div 0.23$ nm) or synchrotron radiation ($\lambda = 0.06 \div 0.325$ nm). The latter sources yield much higher beam intensity and providing advantages especially for biological objects in solution, as the latter consist of lighter atoms and yield relatively weak signals (Feigin and Svergun (1987)).

1.2.2   Solution SAXS for polydisperse and flexible systems

Flexible systems are polydisperse (or heterogeneous) systems containing multiple types of particles in solution. Generally speaking, polydispersity makes SAXS analysis of molecular arrangement more difficult compared to monodisperse systems. The observed intensity from solution is no longer proportional to scattering from a single type of particle, as it is in monodisperse systems. Each particle in the

population contributes to the scattering pattern, resulting in an averaged intensity. The SAXS intensity $I(s)$ is therefore a linear combination of the contributions of K different components in the mixture expressed as

$$I(s) = \sum_{k=0}^{K} v_k I_k(s).$$
(1.3)

Here, v_k and $I_k(s)$ denote the volume fraction and scattering intensity from the k-th component, respectively. If either volume fractions or scattering intensities of the contributing components are known, the other parameters can be recovered from the resulting intensity of a mixture (Svergun et al. (2013); Kikhney and Svergun (2015)).

In the case of intrinsically disordered and flexible proteins, the scattering patterns could be different for the same proteins depending on their conformations. This makes it difficult to predict the scattering curve of each conformation. For that, an ensemble optimisation method may be applied (Bernadó et al. (2007); Tria et al. (2015)), where the scattering from multiple conformations is simulated and utilised to fit the experimental data with the linear combination(s) thereof.

1.2.3 Time-resolved SAXS for structural kinetics

TR-SAXS investigations became popular in the last years, thanks to the availability of specific time-resolved setups like stopped-flow devices at dedicated SAXS beamlines (Blanchet et al. (2015)). The temporal resolution of such measurements is dependent on the optical characteristics of the beamline and - in case of rapid mixing with stopped-flow - on the time of mixing and delay line (Narayanan et al. (2001)). These setups allow for the temporal resolution down to milliseconds which is sensible at the molecular level Blanchet et al. (2015). Slow processes could also be explored with SAXS in a time-resolved manner by collecting the data at different time points (Kirby and Cowieson (2014)). The advantage of the slower mode of data collection is the possibility to acquire a better statistics of the measurements.

Unlike static SAXS measurements, TR-SAXS data also contain information about structural changes and this requires the data to be analysed in series. Usually, TR data capture mixtures of different states i.e. provide scattering from polydisperse systems and need relevant approaches for data analysis.

1.2.4 SAXS and TR-SAXS data analysis

SAXS data processing and analysis procedures start with an azimuthal averaging of two-dimensional (2D) scattering patterns from the X-ray detector producing an one-dimensional (1D) curve. Given that the buffer (in which the protein is dissolved) contributes to the scattering signal, the protein's contribution is normally calculated by subtracting the buffer's contribution. These initial steps of the data processing can be done automatically, using, for instance, pipelines as described in Franke et al. (2012). The automated analysis also allows for the preliminary assessment of structural parameters such as R_g, maximum dimension of the particle (D_{max}), its molecular weight and flexibility.

Further and more comprehensive analysis may be performed utilising program packages like SASfit (Breßler et al. (2015)), SasView (SasView (2022)) or ATSAS (Manalastas-Cantos et al. (2021)). The latter offers a broad range of tools to process, analyse, manipulate and simulate the SAXS data. Simulating SAXS data from available crystal structures or atomistic models allows one to compare the scattering data with those of a reference model. The program called CRYSOL was developed by Svergun et al. (1995) and is available in ATSAS (Franke et al. (2017)). Using indirect Fourier transform in GNOM (Svergun (1992)), it is possible to obtain real-space information about the sample which could be helpful for further three-dimensional (3D) modelling (e.g. programs DAMMIN (Svergun (1999)) and DAMMIF (Franke and Svergun (2009)) reconstructing low-resolution bead models).

To assess the statistical similarity between measured SAXS intensities $I_{exp}(s)$, and those calculated from a model, $I_{calc}(s)$, the reduced χ^2 statistic (Pearson (1900)) is used:

$$\chi^2 = \frac{1}{n-1} \sum_{k=1}^{n} \left[\frac{I_{exp}(q_k) - I_{calc}(q_k)}{\sigma(I_{exp}(q_k))} \right]^2, \qquad (1.4)$$

where n is the number of data points in the SAXS curve and $\sigma(I_{exp}(s_k))$ are the experimental errors. The model adequately describes the experimental data if $\chi^2 \in [0.9, 1.1]$ (Franke et al. (2015)).

There are also tools available to analyse multiple data sets and time-resolved data. CHROMIXS (Panjkovich and Svergun (2018)) is aimed to process SEC coupled SAXS, by selecting the signal from the consequently measured fractions of a sample passing through a chromatography column. EFAMIX (Konarev et al. (2022)) can decompose partially overlapping components of chromatography SAXS. Shape determination of unknown intermediate species for time-resolved data is implemented in DAMMIX (Konarev and Svergun (2018)). Calculating the vol-

ume fractions of known components in a mixture is possible within OLIGOMER that fits the experimental data with linear combinations of these components (Konarev et al. (2003)). The number of components in the TR-SAXS data may be determined using, for example, SVDplot component of ATSAS or other implementations of the singular value decomposition (SVD) (Svergun et al. (2013)) (see next section about SVD).

Although various software suites are aimed at TR-SAXS data analysis (Meis-burger et al. (2021); Hopkins et al. (2017); Sagar et al. (2021)), *ab initio* shape reconstruction at specific time points remains an important step. *Ab initio* reconstruction of particles' shapes in solution is implemented in DAMMIN (Svergun (1999)), DAMMIF (Franke and Svergun (2009)), DAMMIX (Konarev and Svergun (2018)) and DATMIF of ATSAS suite (Manalastas-Cantos et al. (2021)), DENSS (Grant (2018)), DENFERT (Koutsioubas and Perez (2013)) and others.

Singular value decomposition

Modern research methods involve large sets of experimental (or statistical) data, and it is in many cases important to use dimensionality reduction to optimise the analysis. To find hidden dependencies or factors, and to determine which data can be neglected and which are important, relevant mathematical approaches are often helpful. In our case, multiple TR data sets can be considered a matrix of data with columns representing the scattering curves, and a formal description and reduction of this matrix is achieved by the so-called SVD.

As early as the 1960s, Voevodin proposed a method for solving the complete eigenvalue problem based on a generalisation of the method for determining principal values (Voevodin (1968)). Here, the principal values are non-negative square roots of the eigenvalues of an original data matrix multiplied by its conjugate. Here, the problem of finding eigenvalues of an arbitrary matrix can be reduced to a similar one for a quasi-diagonal matrix.

Later, Golub suggested using such a reduction in the problem of finding a pseudo-inverse matrix (Golub and Kahan (1965)). This algorithm is valid for any initial matrix A and allows finding such unitary matrices U and V, which transform matrix A to a diagonal matrix whose elements are called singular numbers: first, matrix A is reduced to bidiagonal form and then it is diagonalised. The resulting singular numbers coincide with the principal values of the original matrix. A simple implementation of this algorithm for a complex matrix was proposed by Businger and Golub (1969).

The application of the singular number decomposition to the linear approxima-

tion problem is described in detail in Golub and Reinsch (1971). It is in this form that this transformation is called SVD (adapted from Golub and Reinsch (1971)): «Let A be a real matrix $m \times n$, where $m \geqslant n$. Then

$$A = UWV^T, \tag{1.5}$$

where $U^T U = V^T V = V V^T = I_n$ and $W = \text{diag}(w_1, ..., w_n)$. Matrix U consists of n orthonormalised eigenvectors corresponding to n the largest eigenvalues of matrix AA^T, and matrix V — of orthonormalised eigenvectors $A^T A$. The diagonal elements of matrix W are– non-negative square roots of the eigenvalues of matrix AA^T (or $A^T A$), they are called the singular values of matrix A. Moreover, $w_1 \geqslant w_2 \geqslant ... \geqslant w_n \geqslant 0$. Thus, if $rank(A) = r$, then $w_{r+1} = w_{r+2} = ... = w_n = 0$. The decomposition (1.5) is called a singular value decomposition. »

In practice the reduced SVD is usually used:

$$A = \sum_{i=1}^{r} \overrightarrow{u_i} w_i \overrightarrow{v_i}^T, \tag{1.6}$$

where only non-zero singular values are considered in the reconstruction of the original matrix. Since the components of the decomposition are ranked, depending on the conditions of the problem, one can choose, starting from which value i it makes sense to put w_i equal to zero (rounded).

The SVD is often used in the kinetics analysis of biological and chemical processes being among the most popular practical approaches to dimensionality reduction. SVD is based on a non-parametric kinetics, which uses a decomposition of the measurement matrix (A) of a thermodynamic process to estimate the dependence on changing conditions. The algorithm records in U the values of the transformation function or kinetic model, and in V the changes in the measured parameters as a function of temperature (Ferrer et al. (2017)).

SVD has been successfully used in data analysis and for searching of intermediates in time-resolved X-ray crystallography (Schmidt et al. (2002)). In this case, the experimental data of the sample optical absorption or electron density is considered as a function of the emission wavelength and as a function of time. SVD represents the data set by two sets of vectors which are weighted by their respective singular values. In time-resolved spectroscopy, for example, the "left" set of singular vectors (lSVs) constitutes a time-independent and orthonormalised basis from which all time-dependent difference spectra in the data matrix are constructed. "Right" singular vectors (rSVs) describe the time dependence of lSVs (Schmidt et al. (2002)).

In SAXS, SVD was applied for more than 30 years to obtain useful information on multicomponent systems. Here, the number of components and their pure scattering patterns are unknown, but in a series of experiments, the volume fractions of the components are changed (Fowler et al. (1983)). For the SAXS data analysis, n is the number of SAXS curves and m is the number of points in each curve.

Physically, the number of significant singular vectors (non-random curves with significant singular values) represents a minimum number of independent curves needed to describe the whole data set by their linear combinations. Non-random curves can be identified by suitable non-parametric statistical tests and their number gives an estimate of the minimum number of independent components in equilibrium or non-equilibrium mixtures (Svergun and Koch (2003)).

In the last years, the SVD has become the most used method for interpretation of the TR-SAXS experiments. This approach provided important results in structural biology, e.g. studies of nucleosome dynamics (Chen et al. (2014)); spontaneous reassembly of denatured cytochrome c (Mitić et al. (2017)) as well as other biochemical reactions (e.g. Oang et al. (2017)).

1.2.5   Amyloid beta fibrillation studies with SAXS

With SAXS, kinetics of Aβ42 aggregation in solution was earlier investigated by Messa et al. (2014). It was reported that the aggregation rate could be controlled by changing the temperature; the oligomer composition of the sample after 3 hours of incubation at room temperature was similar to that after 1 day at 4°C. This study also reported simultaneous presence of monomers as well as rod-like structures in solution.

To clarify structural details of amyloid aggregation using SAXS, additional information about the system is required. The types of Aβ42 aggregation kinetics could be determined only if specific features in the data are simultaneously monitored. For instance, in Meisl et al. (2016), an algorithm determining types of aggregation kinetics was described for fluorescence studies of amyloids. In this case, a specific kinetics model of the process was utilised, whereas for SAXS data different phenomena are observed at different angular ranges (i.e. length scales in the system). Additional information is therefore required to detect and adequately characterise certain types of kinetics with SAXS. In particular, some studies demonstrated the applicability of SVD for exploring the Aβ aggregation (Ikeda et al. (2011)) , but the kinetics of the SVD components recovered from SAXS was not matching the ones from the studies by circular dichroism.

1.3   Aim of the work

This thesis will introduce a general data-analysis procedure for biological TR-SAXS data and its application to Aβ42 aggregation process. The first part of the thesis describes the collection of TR-SAXS data from Aβ42 exploring the process of amyloid fibrils formation and developing a protocol to characterise these data. In the experimental part of the thesis, the process of Aβ42 aggregation is studied by TR-SAXS, and the data are analysed using the appropriately adapted tools suitable for evolving polydisperse systems.

The second part of the thesis will be devoted to the automation of TR-SAXS data analysis procedure implemented as a pipeline. The pipeline is aimed at streamlining the analysis of TR-SAXS data and providing an easy-to-use tool for the beamline users. The procedure makes TR-SAXS data analysis easier and more understandable for inexperienced users while being still sufficiently comprehensive to interpret the scattering data in the biophysical context.

2 Materials and Methods

2.1 Materials and experimental method

2.1.1 Biological samples

The manufacturer supplied HFIP (Hexafluoroisopropanol) dried film of recombinant Aβ42 was solubilised in 1mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and added NH$_4$OH (pH = 12.0) at a concentration of 2mg/mL, followed by 15 minutes of sonication in an ice bath. The concentration of NH$_4$OH was varying for different experiments. The following buffer compositions with the HEPES stock solution pH 7.0÷7.6 were utilised:

- 1 mM HEPES, 0.1% NH$_4$OH (resulting pH 10.7);

- 10 mM HEPES, 0.001% NH$_4$OH, 150 mM NaCl, resulting pH$\sim$10.7.

Prepared Aβ42 samples were then immediately infused with the corresponding buffers and were proceeded for T$_0$ SAXS measurement after 15 minutes of sonication at room temperature. All the samples were exposed to X-rays at 20°C. Before being exposed, the samples were stored and aged at room temperature. Several measurements were conducted at the 10°C temperature of the sample storage.

2.1.2 SAXS measurements

The TR-SAXS data collected in batch mode at the P12 beamline of PETRA III (Blanchet et al. (2015)). For the TR-SAXS experiments a solution of Aβ peptide was placed on the sample storage plate of the robotic sample changer in PCR tubes; then the sample was loaded into the capillary, which was illuminated by a monochromatic X-ray beam; the procedure was repeated at multiple time points (within several hours up to 40 hours) following the process of the formation of the fibrils during sample incubation. The scattering intensity $I(s)$ recorded by an X-ray

detector is presented as a function of momentum transfer s (Svergun et al. 2013; Blanchet et al. 2015). The TR-SAXS data were used to study the process of Aβ fibril formation and also to showcase the performance of the data analysis pipeline.

P12 hardware

The beamline setup is schematically shown in Figure 2.1 (adapted from Blanchet et al. (2015)). The X-ray beam goes from the undulator to the detector (right-to-left direction in this case). Petra III U29 undulator consisting of magnets creates a magnetic field perpendicular to the electron beam converting it to the beam of photons. A system of slits is used to set the beam size. Then, the beam traverses the double crystal Si (111) monochromator, which selects the wavelength (typically of 0.124 nm) used for the SAXS measurements. Slits 1 are employed for beam shaping after the monochromator. The use of a so-called KB geometry (Kirkpatrick and Baez (1948)) combining horizontal and vertical mirrors creates a highly fo-cused beam with the flux over 10^{13} ph/s. Three-paired-slits collimator system (slits 2, 3 and 4 in 2.1) is employed to reduce the instrument background. Before reaching the sample, the X-ray beam can be attenuated by the beam attenuator which means inserting a foil of a certain thickness absorbing extra radiation to prevent radiation damage to a biological sample. The shutter opens during the measurement exposing the sample to the X-rays. The X-ray beam reaches the sample cell, the photons scatter upon interacting with the sample's electrons and the scattered beam passes through an evacuated flight tube reaching the detector. Dectris 2D photon counting detector Pilatus 6M registering scattered photons is positioned typically at 3.1 meters away from the sample cell.
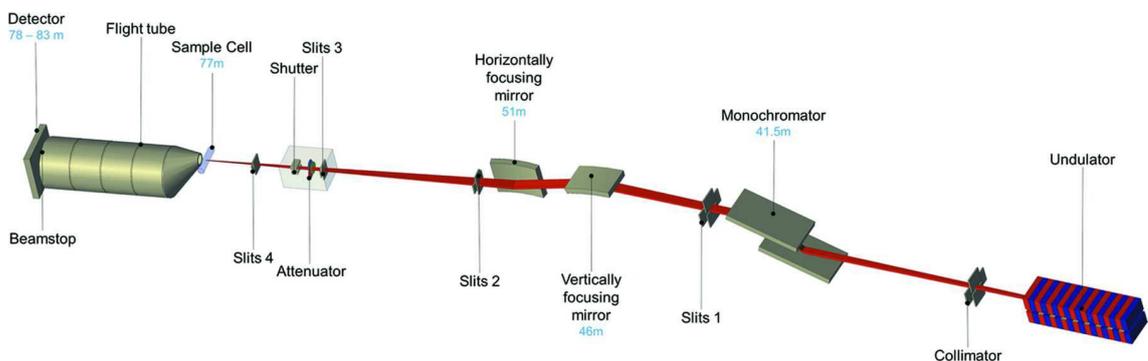
Figure 2.1: Optical scheme of the P12 BioSAXS beamline adapted from Blanchet et al. (2015).

2.2 Computational methods

2.2.1 Data analysis tools

For the data analyses and pipeline implementation the following ATSAS (3.1.0) components were employed:

1. PRIMUS was used for SAXS data visualisations.

2. DATADJUST and DATCROP were used to select the most representative data ranges and to allow for a better visual comparison of the data.

3. AUTORG was used for the R_g assessment.

4. GNOM was used for performing indirect Fourier transform and evaluating the particle distance distribution function P(r) and the maximum dimension of a particle.

5. DATAVER and DATREGRID were used to prepare the experimental data for further processing.

6. SVDPLOT and SVDSAS (currently Windows only) were used to evaluate the Python implementation of SVD tailored for the TR-SAXS data analysis.

7. DAMMIN and DAMMIX were used to create *ab initio* dummy atom models based on the experimental data and accounting for P(r) function.

8. SUPCOMB, DAMAVER and DAMFILT were used to refine the *ab initio* models.

9. FFMAKER was used to prepare form-factor files of the *ab initio* models.

10. OLIGOMER was used to fit the experimental data with linear combinations of form-factors of the respective models.

The following modules of Python 3 programming language (Van Rossum and Drake (2009)) were utilised for the data analysis pipeline implementation:

- `os, glob, shutil, io` are the modules for performing file operations and interacting with the system's interfaces from the Python code;

- `functools, operator` are the modules used to simplify performing mathematical operations for multiple arguments;

- `multiprocessing, subprocess` are the modules employed to spawn the process to optimise the performance;

- `tabulate` is a module utilised to create portable tabular outputs;

- `numpy` is a module aimed to perform mathematical operations (Harris et al. (2020));

- `pandas` is a module, which allows working with data as data frames making the analysis more convenient (Reback et al. (2022));

- `plotly` is a module that was used for making plots and data visualisations (Plotly Technologies Inc. (2015));

- `pysaxsdocument` is an ATSAS component allowing for easier integration with Python (Manalastas-Cantos et al. (2021)).

3 Results

3.1 Experimental results

To explore potential heterogeneity and variability of the amyloid fibrils formation process, several experimental series were followed by TR-SAXS. The six TR-SAXS data sets representing the structural kinetics of the Aβ fibril formation at different experimental conditions are presented in Table 3.1.

	Storage temperature, °C	Buffer	Preparation details	Data collected
Set1	10			November 2018
Set2	20	1 mM HEPES, 0.1% NH_4OH (pH$\sim$10.7)	HFIP dried film resuspended in buffer, sonicated for 15 minutes on ice before ageing at room temperature	April 2019
Set3	20			December 2019
Set4	20			October 2020
Set5	10			October 2020
Set6	20	10 mM HEPES, 0.001% NH_4OH, 150 mM NaCl (pH$\sim$10.7)		June 2021

Table 3.1: Collected TR-SAXS data and experimental conditions.

A pilot experiment was conducted at the P12 beamline in November 2018 and will be further referred to in this thesis as "Set 1". This set (Figure 3.1) represents the kinetics of amyloid fibrils formation at low temperature ($10°$C). The scattering

intensities recorded immediately after sample preparation are labelled as "0 min" for a minimum possible time lag in minutes between the sample placement and the measurement. The time points with the following lags were collected: 5 minutes, 10 minutes, 15 minutes, 30 minutes, 1 hour, 2 hours, 10 hours and 24 hours of incubation. Importantly, for Set 1 the samples have been prepared separately and incubated until the measurement. An approach of mixing fractions at different ageing times was attempted in parallel.

For Set 1, significant changes were detected after 60 minutes of incubation with the most pronounced variation between 120 and 1440 minutes (from 2 to 24 hours). The changes are seen in the SAXS curves at smaller angles while the higher angle portions reveal relatively low variations as illustrated in Figure 3.1. Such behaviour of the SAXS data series suggests that the internal organisation of the particles remains unchanged whereas the overall size and shape of the peptide aggregates demonstrate significant alterations over time.
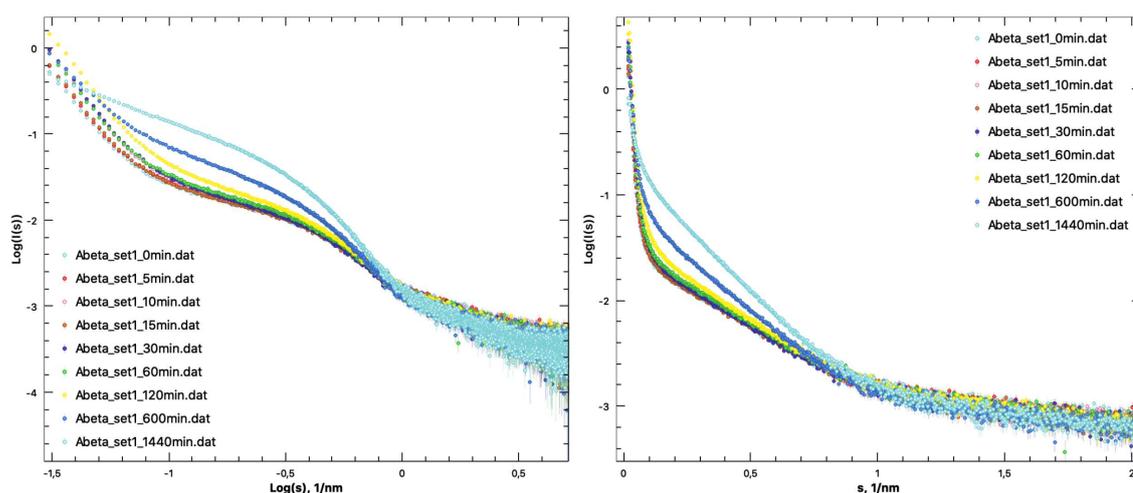
Figure 3.1: Set 1 data displayed in double logarithmic scale on the left and in logarithmic scale on the right. Here and below, the two scales are utilised; the double log scale is often employed in representation of the scattering from fibrillar systems to highlight the changes at the smallest angles.

In Figure 3.2, a comparison of the lower angle portion of the data is displayed in a semi-log plot. All data of Set 1 show a noticeable upturn of the intensity at the very small angles, which indicates that large particles (oligomers or protofibrils) are present in solutions at all time points recorded. The scattering intensity at the smallest angles increases with time indicating an increase in the amount of the aggregates. The steepest upturn at the smallest angles was observed at the very beginning of the process pointing out that the largest particles were most significantly contributing to the scattering. At further time points, the slope at the smallest angles becomes less steep suggesting that the sizes of the largest parti-
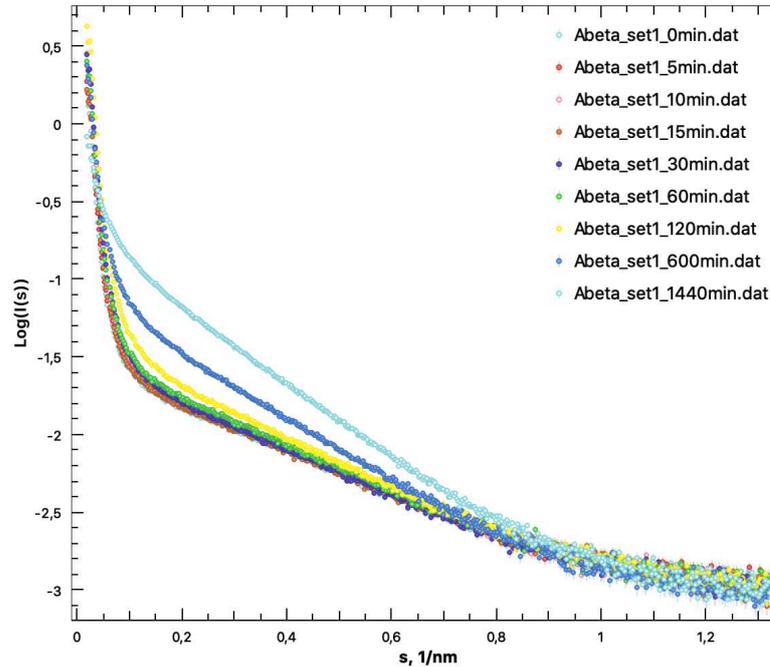
Figure 3.2: Set 1 data smaller angles range

cles are decreasing. These observations indicate that large (fibrillar) aggregates are present in solution from the very beginning of the process and, interestingly, these aggregates may display an exchange with the developing Aβ42 fibrils during the process of formation of the latter.

As the major differences in the data are observed in the range of momentum transfer $s \in (0.2 \div 2)$ nm^{-1}, this range (corresponding to the real space are distances about $3 \div 30$ nm) was further considered for quantitative analyses. The smallest angles corresponding to the largest aggregates and wider angles corresponding to atomic scale packing distances were excluded. Real-space sizes of the intermediate states of the amyloid fibrils formation process are expected to be $5 \div 30$ nm (larger than the monomer but smaller than a protofibril). The possibility of conformational changes within the monomeric blocks ($3 \div 5$ nm) could also be covered by considered within the s-range of $(1.2 \div 2)$ nm^{-1}.

The processes investigated in the present thesis were expected to go through similar pathways for all samples such that the range of sizes of possible intermediates ($5 \div 30$ nm) was sufficient to describe the fibril formation. The same angular range ($0.2 \div 2$) nm^{-1}) was therefore employed in the analyses of all collected sets.

The kinetic data from Set 2 (collected in April 2019) go from 0 to 2403 minutes (0 to ∼40 hours). From Figure 3.3, conformational changes are observed during the entire time of incubation. Similarly to Set 1, the changes are concentrated at smaller angles. Set 2 process starts from extended oligomeric species with further

increase of their sizes over time. Interestingly though, after 35 hours of incubation, the scattering intensities decrease (time points after 2084 minutes) suggesting that the largest oligomers dissociate reaching an equilibrium state by 40 hours (the latest time point).
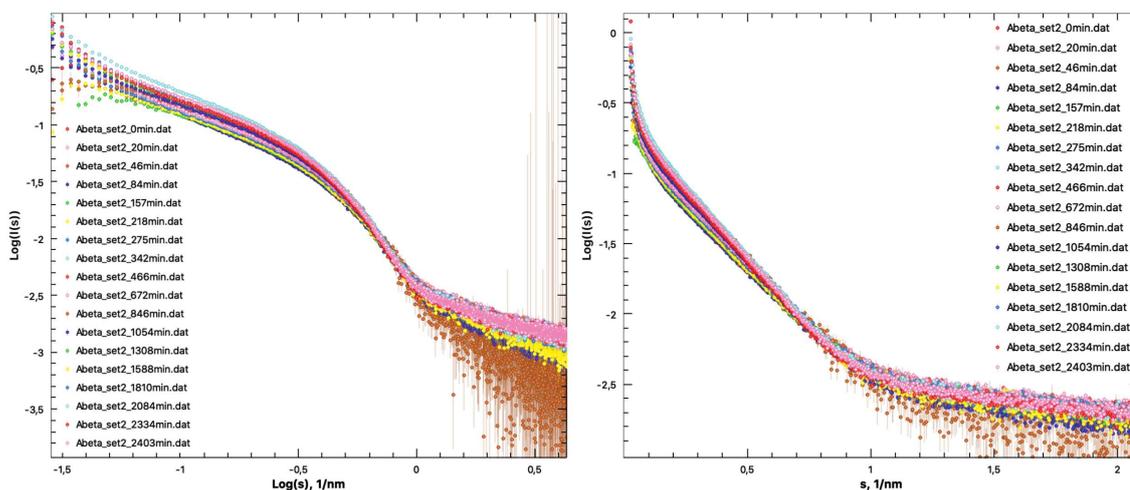
Figure 3.3: Set 2 data displayed in double logarithmic scale on the left and in logarithmic scale on the right.

The SAXS experiment on Set 3 aimed to repeat the measurements of Set 2 was conducted in December 2019 at the same experimental conditions as Set 2 (Table 3.1). However, from Figure 3.4, the recorded curves are somewhat different from the ones of Set 2 (Figure 3.3) revealing the largest particles already after 16 hours of incubation (timepoint: 960 minutes). Kinetics of Set 3 is dominated by the large oligomers observed from the very beginning of the measurements.
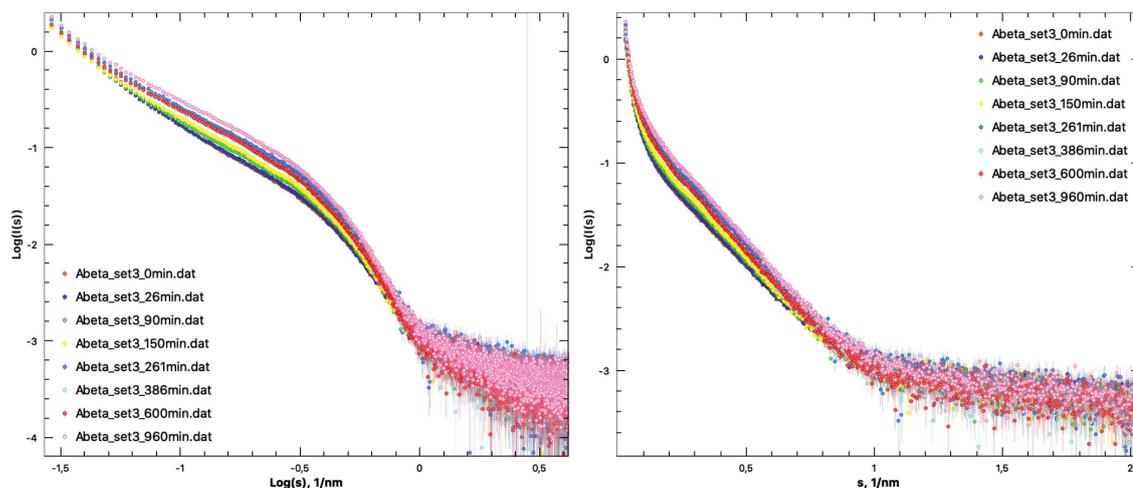
Figure 3.4: Set 3 data displayed in double logarithmic scale on the left and in logarithmic scale on the right.

A significantly different behaviour was observed for Set 4 structural kinetics measured in the buffer with the same concentration of NH_4OH (0.1%) in October 2020. As presented in Figure 3.5, although some aggregates are present in solution (possibly, due to the stickiness of the peptide), the particles detected by SAXS remain rather small indicating no formation of protofibrils. Further incubation increases the contribution of slightly larger particles to the scattering. The largest species observed for Set 4 are still relatively small oligomers detected after about 15 hours of incubation (time point: 897 minutes).
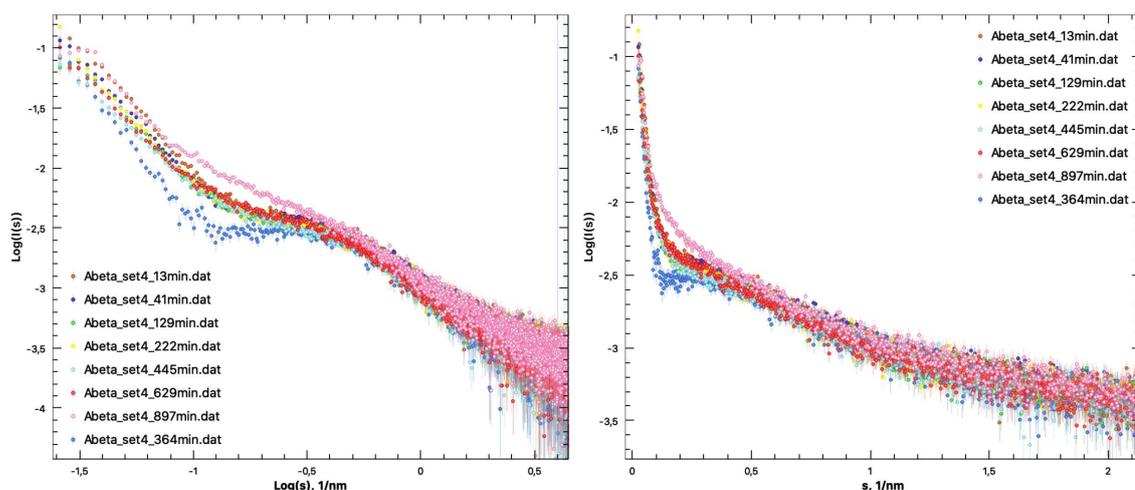
Figure 3.5: Set 4 data displayed in double logarithmic scale on the left and in logarithmic scale on the right.

For Set 5 measured in the same buffer as Set 4 but at a lower temperature in October 2020, the initial oligomerization state of the solute showed very little time-dependent alteration. Indeed, the scattering curves only displayed some change after 2 hours of incubation (in Figure 3.6, corresponding time point: 106 minutes). Such behaviour indicates that the cross-section of the oligomers remains the same for later time points of Set 5. Similar to Set 1 measured at the same temperature, there is evidence of some dissociation for Set 5, too. The largest aggregates were observed at the very first time point and partially dissociated after 1 hour of incubation.
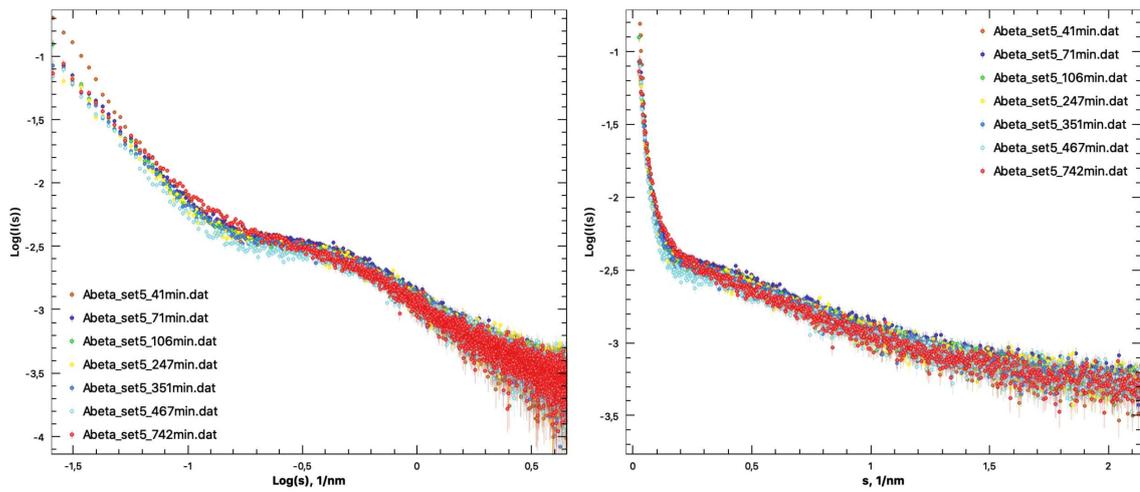
Figure 3.6: Set 5 data displayed in double logarithmic scale on the left and in logarithmic scale on the right.
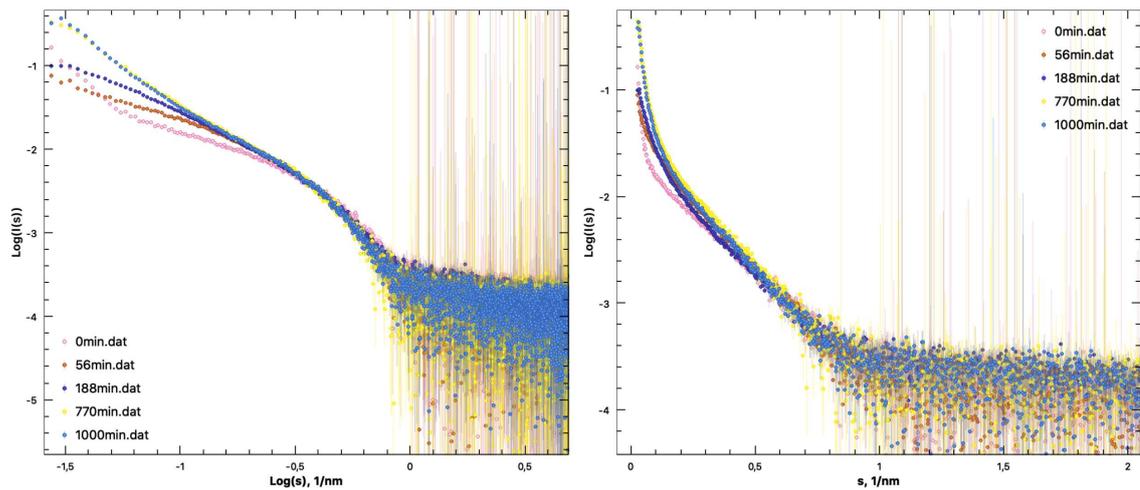
Figure 3.7: Set 6 data displayed in double logarithmic scale on the left and in logarithmic scale on the right.

For Set 6 (Figure 3.7, data collected in June 2021), it was possible to observe a clear formation of larger fibrils again, and the collected SAXS curves could be positioned between the ones of Set 2 and Set 3 (Figure 3.8).
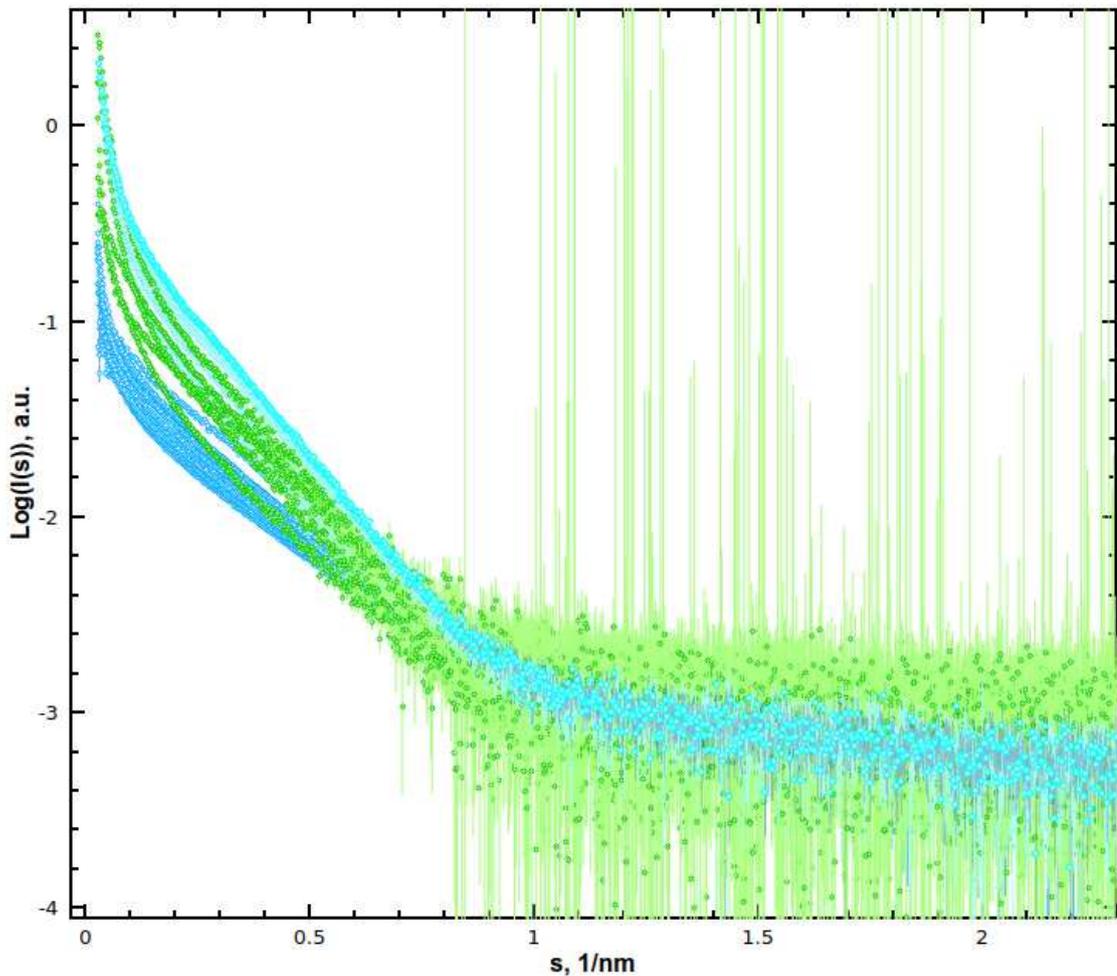
Figure 3.8: Sets 2, 3 and 6. The data of Set 6 displayed in green are between the ones of Set 2 (blue) and Set 3 (cyan) at the smallest angles.

Comparing different sets to each other (Figure 3.9) all SAXS data collected for sets 1, 2, 3, and 6 are located between the first curve of Set 1 and the last curve of Set 3 (the latter corresponds to the largest species observed amongst all six sets).

From Figure 3.10, the final states for the Aβ42 aggregation process in different sets appear to be different. The measurements of Set 3 revealed the largest particles, although sets 1,2 and 6 were incubated for a longer time. Set 4 and Set 5 showed very little increase in size but demonstrated the presence of smaller particles in solution. Sets 2, 3 and 6 revealed only elongated growing species. Together, the recorded data may fully represent some critical stages of the Aβ42 fibril formation process, where the solution is populated by varying oligomeric species. No homogenous solution of Aβ42 monomers was observed, probably due to their intrinsic stickiness and due to the presence of the Aβ42 aggregates as suggested by the steep upturns of the SAXS curves at the smallest angles.
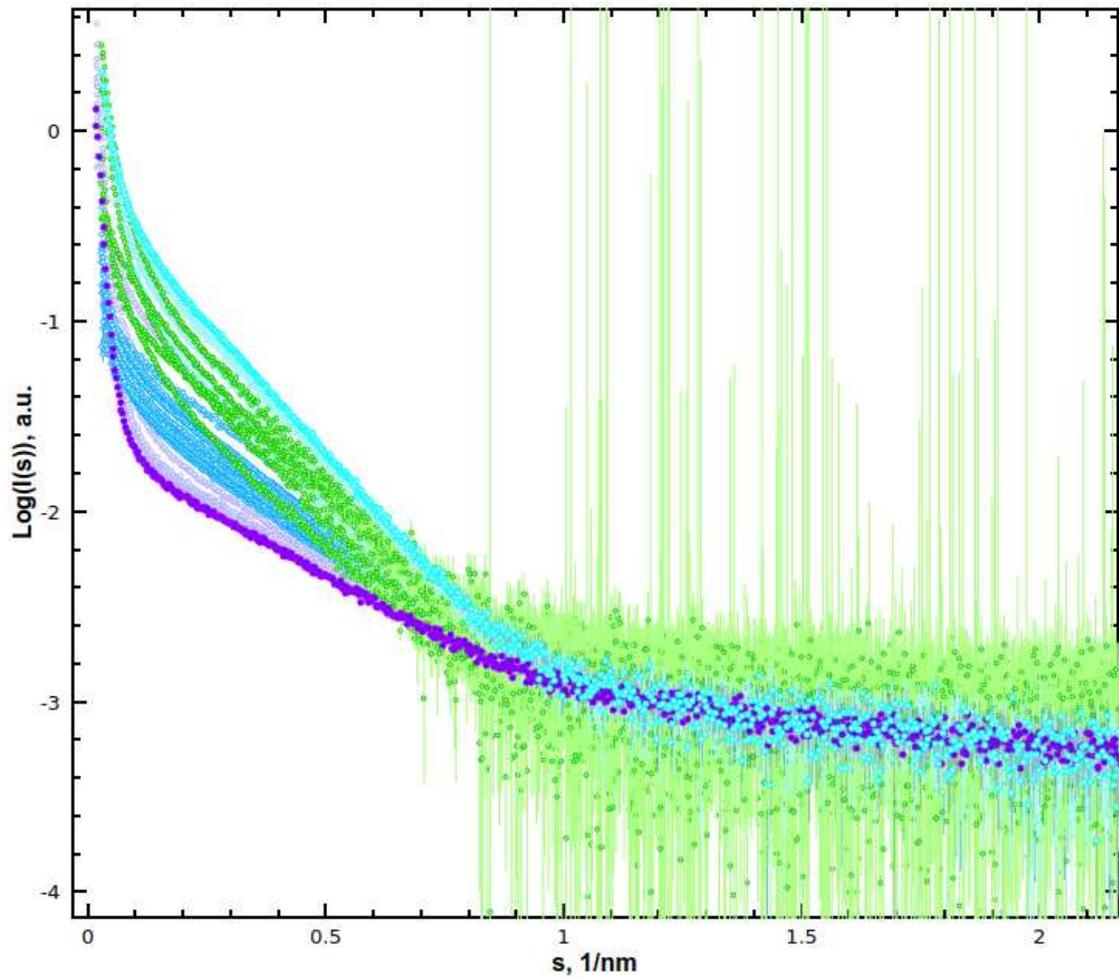
Figure 3.9: Sets 1, 2, 3 and 6. The first curve of Set 1 is displayed in purple and followed by lilac curves of the rest of Set 1. Blue curves represent Set 2, green ones are for Set 6, light cyan for Set 3, with the bright cyan indicating the last curve of Set 3.
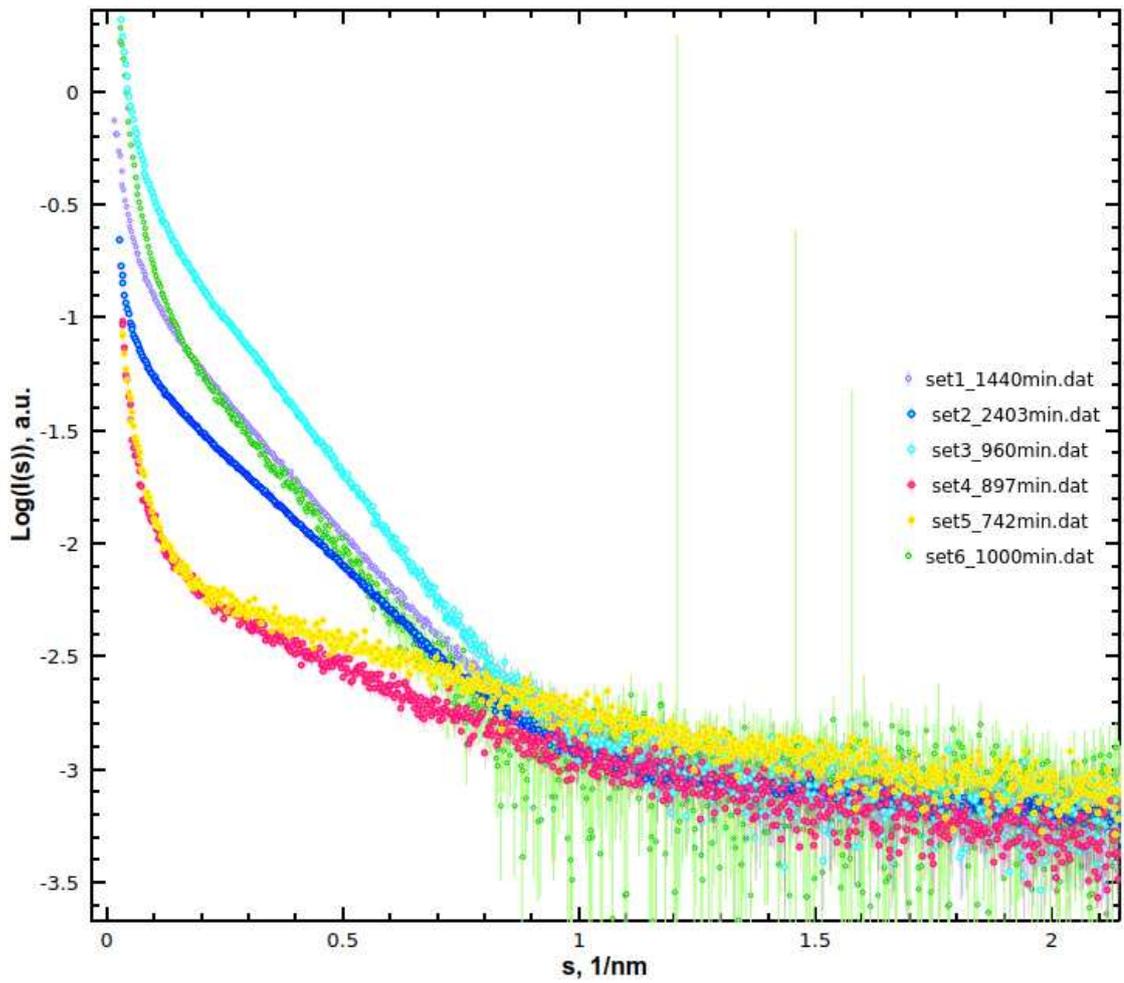
Figure 3.10: Comparison of the final states for the different experiments on $A\beta42$ aggregation.

	Storage temperature, °C	Buffer	Preparation details	Data collected
Set1	10	1 mM Hepes, 0.1% NH$_4$OH (pH$\sim$10.7)	HFIP dried film resuspended in buffer, sonicated for 15 minutes on ice before ageing at room temperature	Clear increase in size
Set2	20			Very little increase in size
Set3				Moderate increase in size
Set4				Smaller particles, little increase in size
Set5	10			Smaller particles, very little increase in size
Set6	20	10 mM Hepes, 0.001% NH$_4$OH, 150 mM NaCl (pH$\sim$10.7)		Rather moderate increase in size (between set2 and set3)

Table 3.2: Conformational changes observed for the collected TR-SAXS data.

The conformational changes observed for the six experimental series are qualitatively summarised in Table 3.2. Despite the apparently different pathways observed for different preparations of Aβ, the collected data may be meaningfully grouped for post-processing and analysis. As shown in Figure 3.11, the experimental data are forming two major clusters. Here, sets 1, 2, 3 and 6 appear to belong to the same structural group and sets 4 and 5 may describe another kind of the Aβ kinetics, presumably, earlier stages of the process observed for the other four sets. The systematic analysis of the collected data is presented in the forthcoming sections.
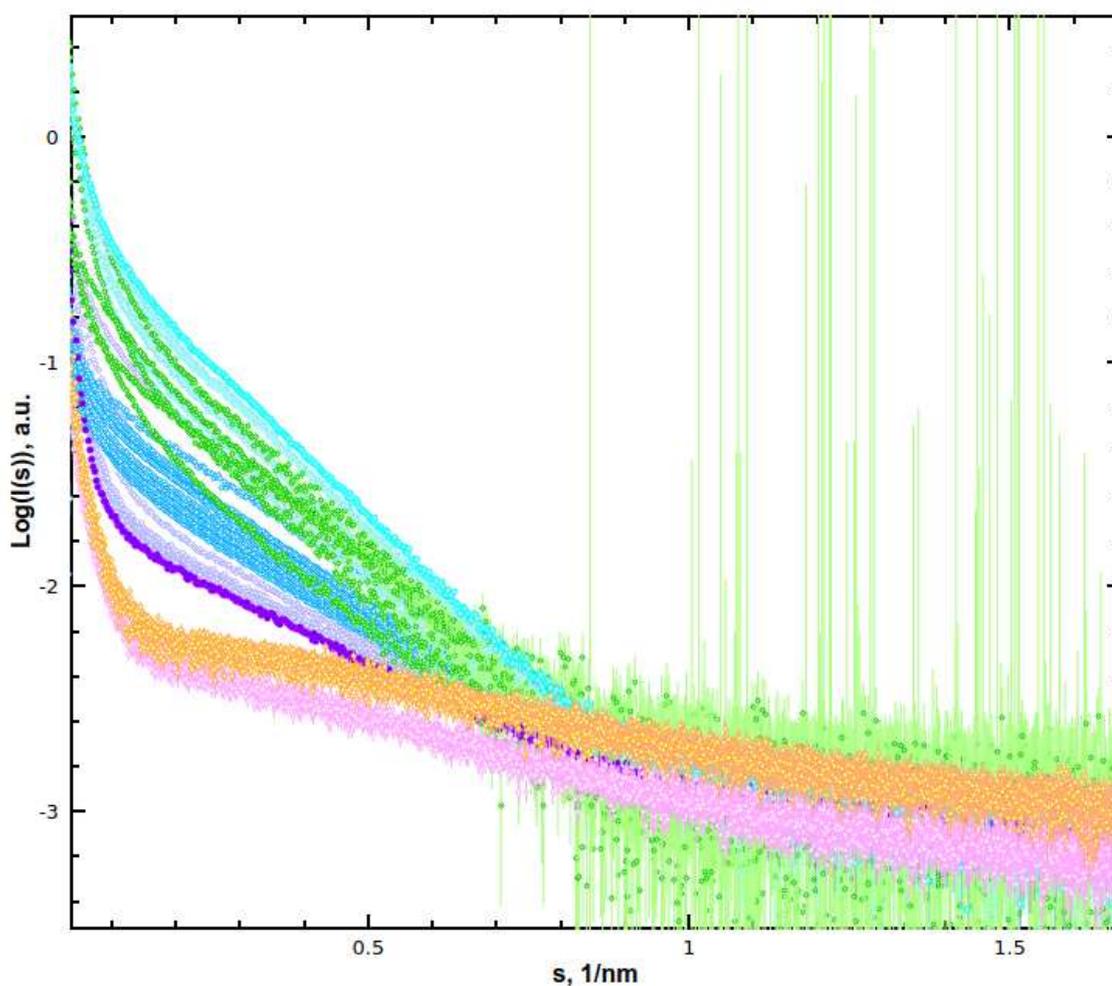
Figure 3.11: Six data sets representing structural kinetics of the amyloid beta fibril formation process. The colour codes are: purple and lilac for Set 1, blue for Set 2, cyan for Set 3, pink for Set 4, orange for Set 5 and green for Set 6.

3.2 Data analysis pipeline implementation

To automate the analysis of the collected Aβ fibril data, a pipeline was developed that executes ATSAS programs in a proper order and provides structural interpretation of the multiple TR-SAXS data sets. The pipeline was designed in a general way such that it can be further employed to simplify the TR-SAXS data analysis procedures for regular users of P12 dealing with kinetic data.

The pipeline is implemented as a script written in Python 3 aimed to process SAXS data studying the Kinetics of Conformational chanGes (KinConG). The code is available from https://git.embl.de upon request.

Several Python modules are employed as tools for data engineering and visualization. The pipeline utilises programs of the ATSAS package(Manalastas-Cantos et al. (2021)) relevant for the time series analysis. The major working block of

KinConG is a single TR-SAXS experiment but the analysis of repeated experiments was implemented as well in the pipeline. KinConG is capable to run in parallel on multiple folders. This mode is suitable for different data series describing the same process. Performing such kind of analysis allows one to generalise the outputs of several TR-SAXS experiments.

For the single experiment mode, a pre-processed series of subtracted 1D SAXS profiles (e.g. processed with SASFLOW (Franke et al. (2012)) is placed in a dedicated folder(s) to implement the following workflow (Figure 3.12):

1. Determination of the first and the last SAXS curves in the series from the available metadata

2. Creating P(r) functions of the initial and final curves using GNOM

3. *Ab initio* shape reconstruction of the initial and final states using DAMMIN

4. Creation of a template to unify the angular axis for the series using DATAVER module

5. Unification of the angular axis with DATREGRID module

6. Building a form-factor matrix using the FFMAKER program

7. SVD of the data matrix (Golub and Reinsch (1971)) visualised with Plotly(Plotly Technologies Inc. (2015))

8. Assessment of the principal components contributing to the signal based on different criteria Durbin-Watson (Durbin and Watson (1950)), Wald-Wolfowitz (Wald and Wolfowitz (1943)), elbow method (Thorndike (1953))

9. *Ab initio* shape reconstruction of an unknown intermediate (if expected from SVD analysis) with DAMMIX(Konarev and Svergun (2018))

10. Building reduced form factor matrix based on the *ab initio* models with FFMAKER

11. Fitting the mixtures of the *ab initio* models to the experimental data using OLIGOMER (Konarev et al. (2003))

12. Parsing OLIGOMER output files and visualisation of the fit quality and volume fractions of the selected models.

The experimental data as the series of 1D scattering curves are taken as an input

set1 set2 set3

The first and the last curves of a set are defined from the metadata.

ab initio modelling of initial and final states of the system

Preparing data for further analysis using DATtools of ATSAS

The experimental data should be on the same angular axis and have equal numbers of points. This is necessary to correctly perform the analysis. This part of the pipeline makes experimental data ready to employ the ATSAS components which are supposed to deal with multiple data sets.

Assessment of the components contributing to the scattering with SVD

Singular Value Decomposition is used to showcase the linearly independent components of the data set. This provides an idea of what may contribute to the signal.

Are the intermediates expected?

Fitting the experimental data with *ab initio* models and calculating contributions thereof as volume fractions

ab initio modelling of an unknown intermediate using DAMMIX of ATSAS

Characterizing the conformational changes occurring during the process

If the SVD analysis and/or additional information from other experiments indicates the presence of an intermediate, it is being also modelled based on the experimental data. In case of several intermediates, the most probable one would be determined.

Observing how the contributions of the components or volume fractions of different species are changing over time, one may employ this information to characterize the process in biological context.

Figure 3.12: Data analysis pipeline.

The pipeline could be run as a command-line Python application and only requires entering a path to the directory containing the time series of the TR-SAXS measurements. Passing multiple directories and processing them in parallel is supported allowing for a simultaneous fitting of the repetitive measurements of the same process. The workflow automates the data analysis procedures conducted manually (step by step) for time-resolved data using the relevant ATSAS components.

The multiple experiments mode requires the files to be placed into the subfolders dedicated to each of the experiments, whereas the upper-level folder contains all of them. Dealing with multiple experiments requires a preliminary assessment of the number of clusters in the datasets. If the number of clusters is fewer than the number of data sets, the sets are grouped by clusters and processed separately, although still in the multiple experiments mode. In this mode, the

workflow described in Figure 3.12 is followed by trying all combinations of the generated models employed to fit all the sets together. The fitting starts with two models and is considered to be satisfactory if additional models do not significantly improve the quality of the fit for all data sets.

3.3   Computational post-analysis results

The experimental data were processed with the KinConG pipeline where both, single-experiment and multiple-experiment modes were applied. The scattering curves from the initial and final states were utilised to construct *ab initio* shapes for each series.

3.3.1   Singular value decomposition and search for hidden intermediates

An important question to answer was, whether the recorded data can be described by a mixture of initial state (monomers or oligomers) and mature fibrils, or there are other components (intermediates) present. Here, we call "hidden intermediates" the species yielding the scattering curves that would complete the dataset such that a linear combination of the 3 components: initial state+final state+hidden intermediate fits the experimental data (see Equation (1.3)). The question concerning hidden intermediates can be assessed with a singular value decomposition SVD of the TR-SAXS data. In SVD, a set of scattering curves is decomposed into the so-called singular vectors, and the vectors with systematic contributions are considered meaningful components contributing to the signal. The SVD analysis has shown that two components are usually not sufficient to describe the processes. The approach is illustrated for Set 1 and also reveals more than two components for the rest of the sets. The corresponding singular values displayed in logarithmic scale along with their derivatives demonstrate non-zero contributions up to the 5th component (Figure 3.13) as determined by the elbow method (Thorndike (1953)). Therefore an attempt to determine at least one (hidden) intermediate was taken for each of the six data sets recorded.
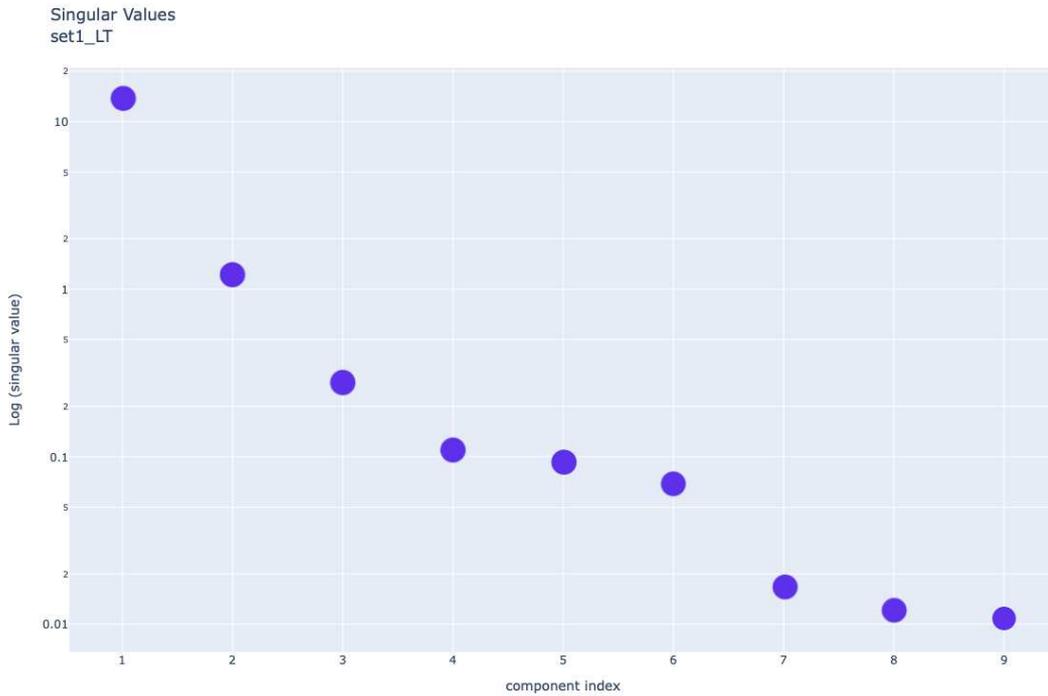
(a)
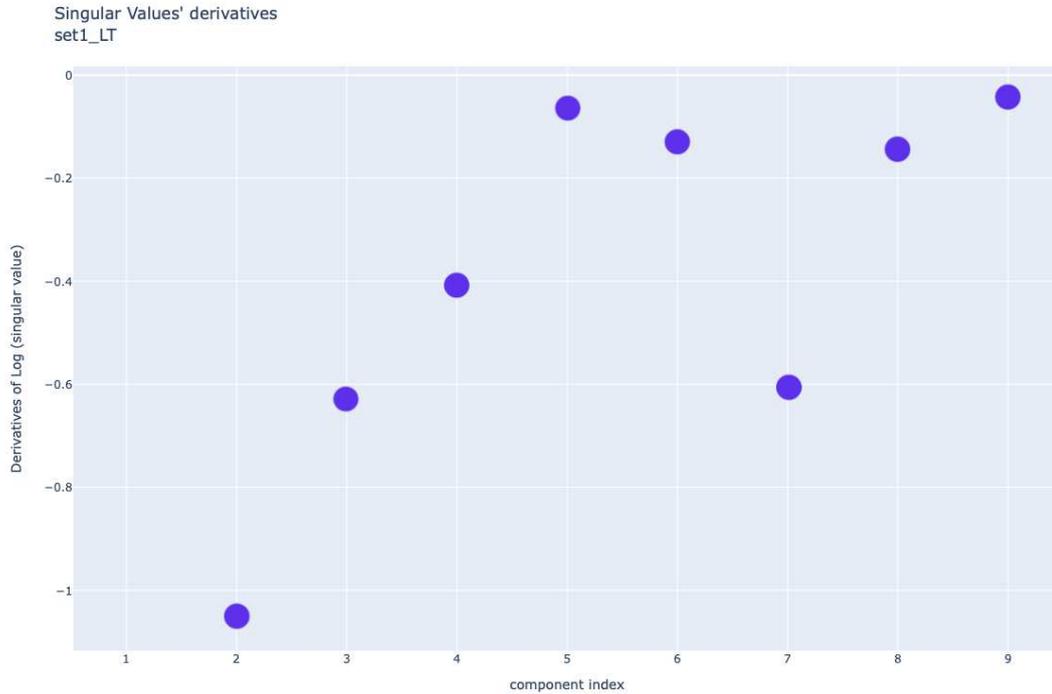
(b)

Figure 3.13: Singular values for Set 1 in logarithmic scale (a) and their derivatives (b)

The first and the last curves of each dataset were taken as processed by KinConG. The shapes of the potential intermediate states were reconstructed for each dataset by DAMMIX. Together with the initial and final states of the peptide in each experiment, these intermediates fitted all the data within each set collected on the process of Amyloid-beta fibril formation with mean discrepancy $\chi^2 = 2.16$. Radii of gyration for the intermediates recovered from the DAMMIX, from 1.5 to 5 nm, are in agreement with the sizes of the smallest Aβ42 oligomers reported before based on SANS data and molecular dynamics (MD) simulations (Wolff et al. (2017)).

The results of the procedures to depict the hidden intermediates in the experimental datasets are illustrated in Figures 3.14 – 3.19. The curves are plotted in a double logarithmic scale to emphasise the most representative angular range. Although the red curves (those from the intermediates) do complete the datasets, they look different – both, from each other and also from the experimental data. For Set 1 (Figure 3.14), the red curve is much closer to the initial state curve suggesting that Set 1 could be well fitted with a linear combination of initial and final states except for the smallest angles. Adding a scattering curve from larger particles than those found in Set 1 to the linear combination would improve the fit at the smallest angles.
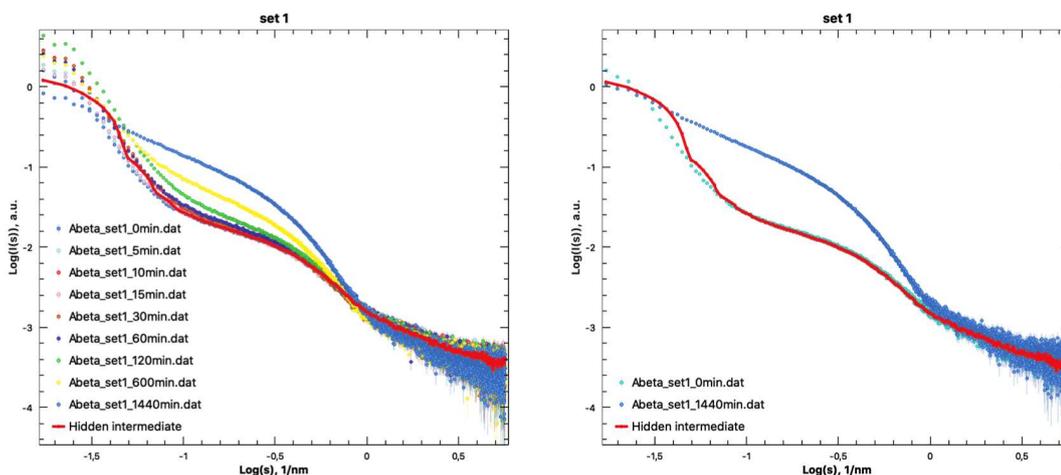
Figure 3.14: Set 1 and three curves sufficient to fit it with linear combination including a hidden intermediate.

A somewhat different situation is observed for Set 2 (Figure 3.15); here, the curve representing the intermediate state is quite close to the initial state. The slope at the smallest angles may indicate a contribution of the small particles. Here, the hidden intermediate would be smaller than the species observed for the

initial state of Set 2.
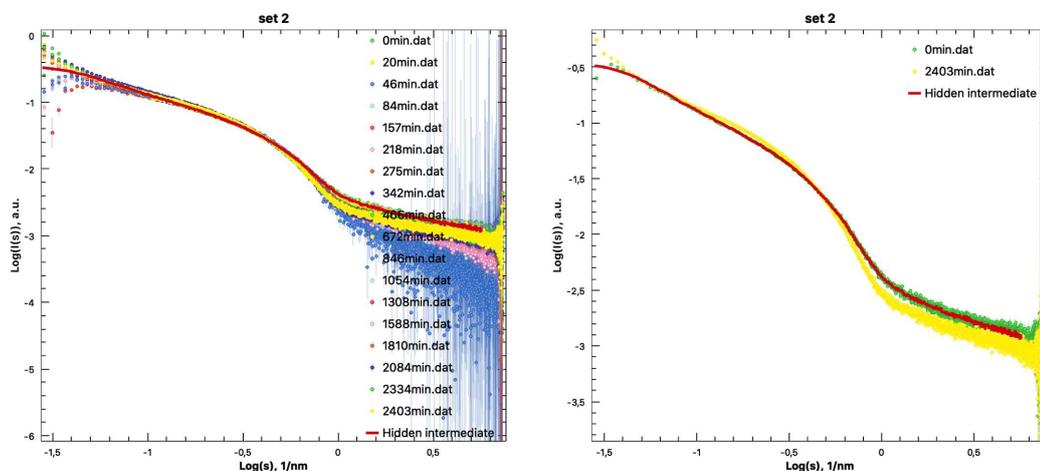
Figure 3.15: Set 2 and three curves sufficient to fit it with linear combination including a hidden intermediate.

Set 3 (Figure 3.16) is similar to Set 2 (they were collected under the same experimental conditions). Although Set 3 contains mostly extended oligomers, some smaller species are contributing, too. Similarly to Set 2, the presence of smaller particles is indicated by the more gradual slope of the red curve (representing the hidden intermediate) at the smallest angles.
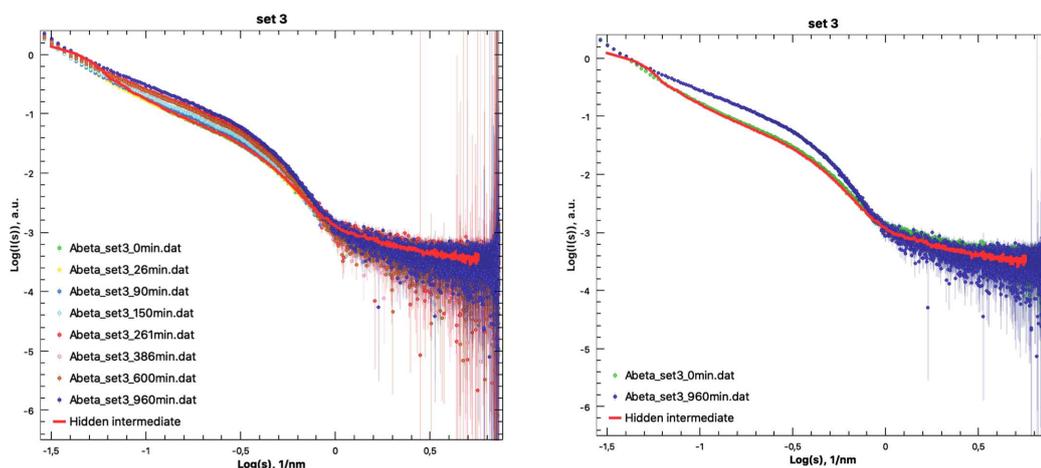
Figure 3.16: Set 3 and three curves sufficient to fit it with linear combination including a hidden intermediate

For Set 4, the scattering data from initial and final states represented in blue and yellow in Figure 3.17 have a minor downturn in the smallest angles. This could indicate that the missing component (represented by the red curve) has a larger size whereas the initial and final states are represented by more compact species.
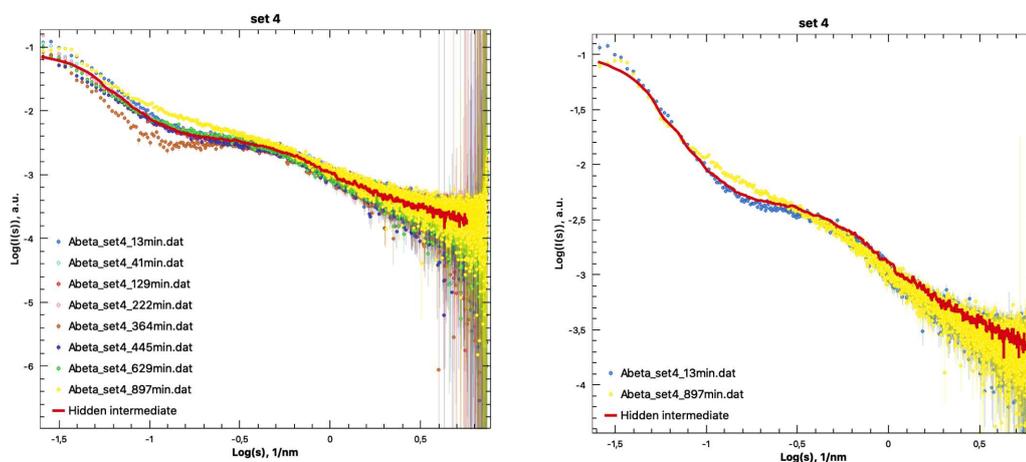
Figure 3.17: Set 4 and three curves sufficient to fit it with linear combination including a hidden intermediate

The intermediate state of Set 5, according to Figure 3.18 (red curve), is smaller than its initial state (blue). This indicates that there can be emerging and disappearing small intermediate during this process.
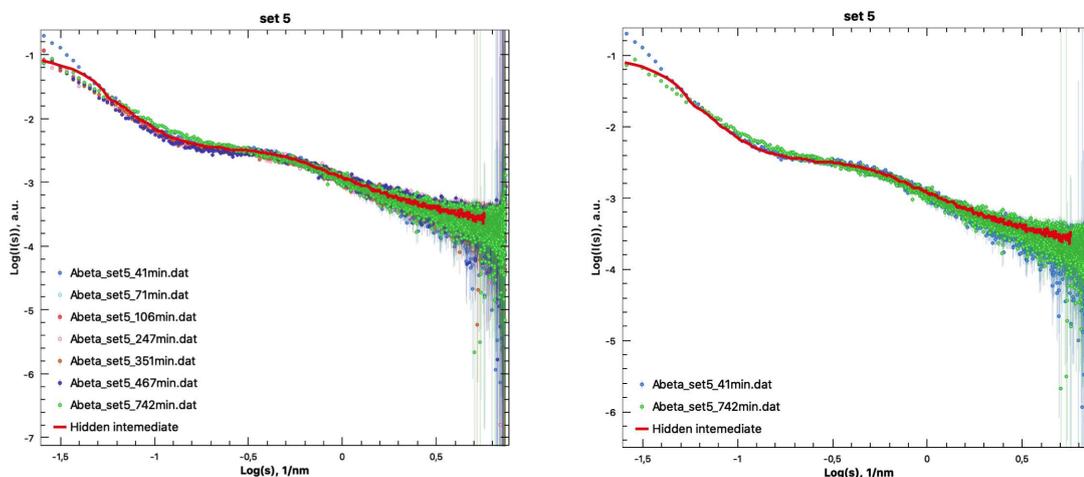
Figure 3.18: Set 5 and three curves sufficient to fit it with linear combination including a hidden intermediate

Set 6 is similar to sets 1, 2 and 3. As shown in Figure 3.19, the computed scattering from the intermediate state is also rather similar to the initial curve. However, at the smallest angles, the scattering curve from the intermediate is significantly different from the experimental SAXS curves of Set 6. More gradual slope suggests that the intermediate can be more compact than the particles observed for the initial state of Set 6. Addition of the SAXS curve from a smaller particle to the linear combination would improve the fit.

35
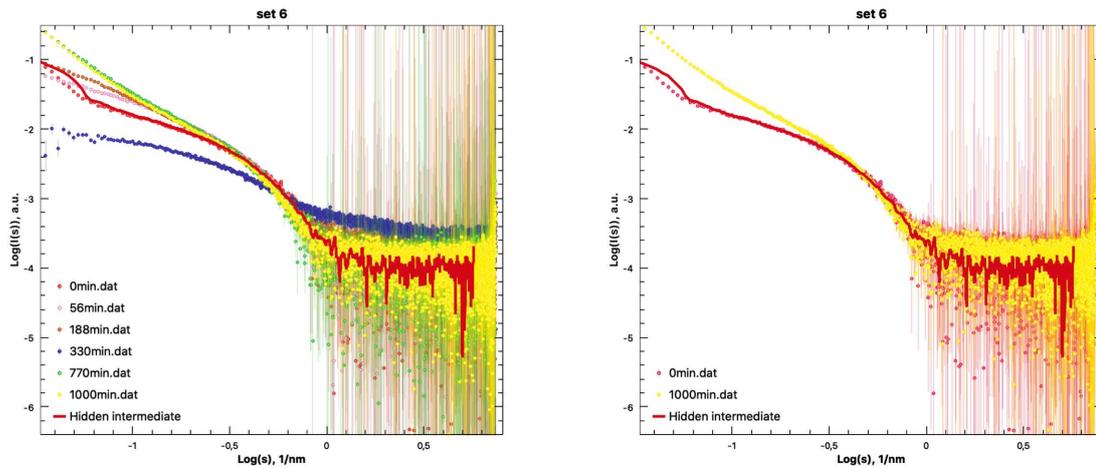
Figure 3.19: Set 6 and three curves sufficient to fit it with linear combination including a hidden intermediate

Although the hidden intermediates allow one to complete the set of models and better fit the experimental data, this approach can rather be considered as an initial guess to analyse the data within the single experimental series. As shown in Figure 3.20, the models of the hidden intermediates created within a single-experiment mode of data analysis, look rather different in shape from each other. This may either point to a limitation of the approach or suggest that the processes do indeed go through different intermediates.
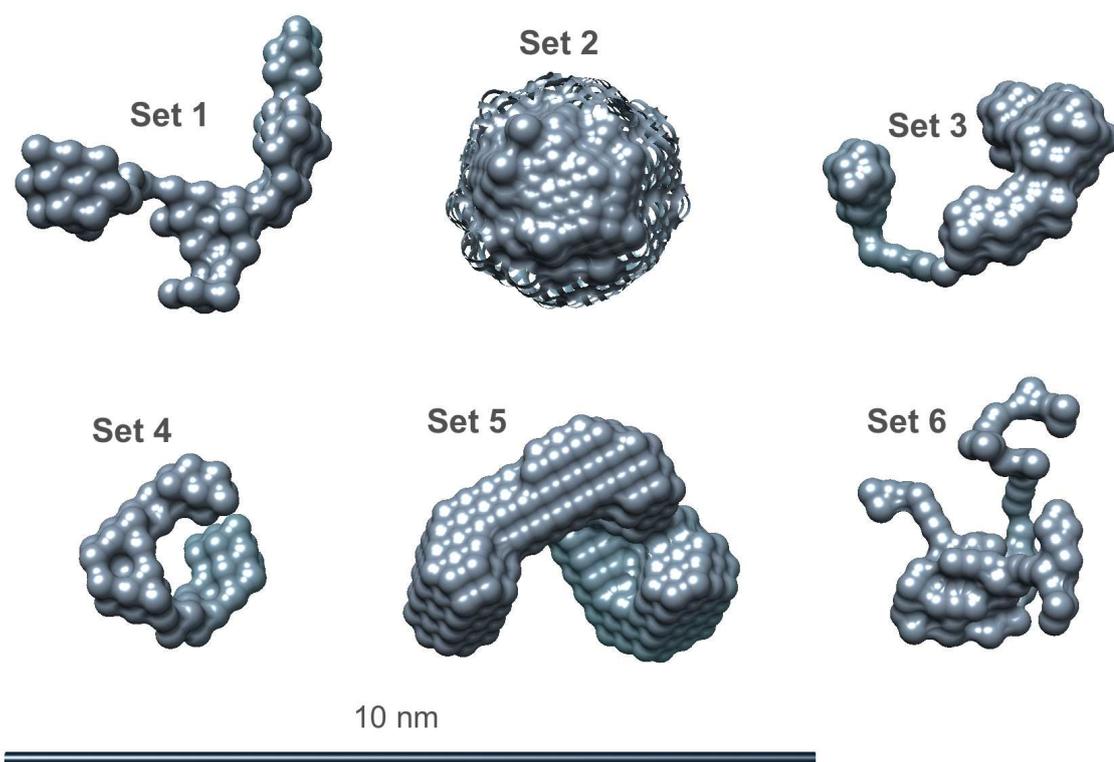
Figure 3.20: Hidden intermediates recovered with DAMMIX for the six data sets.

3.3.2   Data analysis in multiple-experiments mode

As evident from the previous section, the hypothesis of a common hidden intermediate accounting for different initial and final states within individual experimental series did not provide a consistent description of the entire data collected in several experimental sessions. The intermediates in Figure 3.20 displayed varying shapes suggesting that a multiple-experiments analysis mode is probably required to obtain a more adequate picture.

The number of components in solution suggested by the SVD and elbow method points that multiple intermediate component(s) may be needed to describe each TR data set. As seen from Figure 3.11 and Table 3.2, the data are forming two separate clusters and these were analysed by KinConG separately.

Growing oligomeric species of the sets 1, 2, 3 and 6

To analyse the process of growing oligomeric species in sets 1, 2, 3 and 6 multiple-experiments mode of KinConG was applied. Given that the fibrillation processes were beginning and ending at different conformations it was feasible to assume that the starting and finishing points of each process could be utilised to fit all the experimental data. The data sets were therefore fitted simulta-neously

by the linear combinations of *ab initio* models from the initial and final states. The multiple-experiments mode allows for the selection of the number of models to be used in linear combinations. The mode automatically probes all possible combinations and selects the most appropriate one if no *a priori* information is available. For this particular system it was possible to limit the number of models. Linear combinations of 3, 4 and 5 models were tried in the global fitting of the 40 curves representing sets 1, 2, 3 and 6 (note that in the fitting, the time point of 330 minutes was excluded from Set 6 due to an incorrectly subtracted background). The best three conformers yielding a good fit to the data are shown in Figure 3.21 and an example of the fit to Set 1 is given in Figure 3.22. In the selected models, that of the initial state of Set 1 is the smallest one significantly contributing to all four data sets. The largest model is the oligomer observed for the final state of Set 3. The corresponding curve has the steepest slope representing the largest species amongst all the data in the four sets. These two models, however, are not sufficient to adequately describe the entire growing process of the oligomers. The model of the final state of Set 2 was necessary and plays a role of an intermediate state of the entire process for sets 1,2,3 and 6. Given that the smallest model (initial state of Set 1) is already oligomeric, it can be considered as yet another representative intermediate state. Note that the initial monomeric state was not observed in these TR-SAXS series.

Linear combinations of the scattering curves from the selected three models fit all the 40 curves of sets 1, 2, 3 and 6 well as displayed in Figures 3.22, 3.23 (a-c) for sets 1, 2, 3 and 6, respectively. In the fitting, the three curves representing SAXS from the selected models were multiplied by their contributions (volume fractions) and added to each other the way they would best fit the experimental SAXS curve (see Equation (1.3)). Importantly, addition of other components to the fitting by using scattering curves from other initial and final states did not significantly improve the fits. The three selected models can therefore be considered the most representative components appearing for the four data sets.

The change of the volume fractions of the different oligomeric species over time is displayed in Figures 3.24–3.27, and the observed dependencies are discussed below.

The clearest increase in size was observed for Set 1 (Figures 3.1, 3.22; Table 3.2), where all three models significantly contribute to the structural kinetics (Figure 3.24). Starting with the small oligomers, the particles of larger size emerge at the beginning of the process reaching their maximum fraction at about 10 hours of incubation. The contribution of the small oligomers gradually decreases but a small fraction of them is still present after one day of incubation. The largest

species (those from the final state of Set 3) are also present almost all the time. Their contribution increases nearly linearly with time for Set 1.

Very little increase in size was observed for Set 2 (Figures 3.3, 3.23a, Table 3.2). Indeed, the model representing the final state of Set 2 contributes most significantly to the structural kinetics of this set (Figure 3.25). There is a small fraction of the small oligomers in the very beginning which almost disappears after the time point 672 minutes ($\sim$ 11 hours). A small contribution of the largest oligomers is also seen but this contribution is rather insignificant.

Although a moderate increase in size characterises the structural kinetics of Set 3 (Figures 3.4, 3.23b, Table 3.2), the solution is dominated by the largest oligomers (Figure 3.26). Smaller oligomers (those from the initial state of Set 1 and final state of Set 2) could be found during the first six hours of incubation (the corresponding time point is 382 minutes).

The time course for Set 6 is similar to those of sets 2 and 3 (Figure 3.8, Table 3.2). The kinetics in Figure 3.27 shows that there is a very significant fraction of large oligomers similar to Set 3. Contribution of the oligomers characteristic to Set 2 is also present here and completed by a fraction of small oligomeric species (initial state of Set 1).

Summarising, fitting the sets 1, 2, 3 and 6 by the mixtures of three *ab initio* models allowed us to describe the structural kinetics of the growing oligomers during the amyloid fibril formation. Notably, a total 40 SAXS curves were adequately fitted by the scattering from only three models.
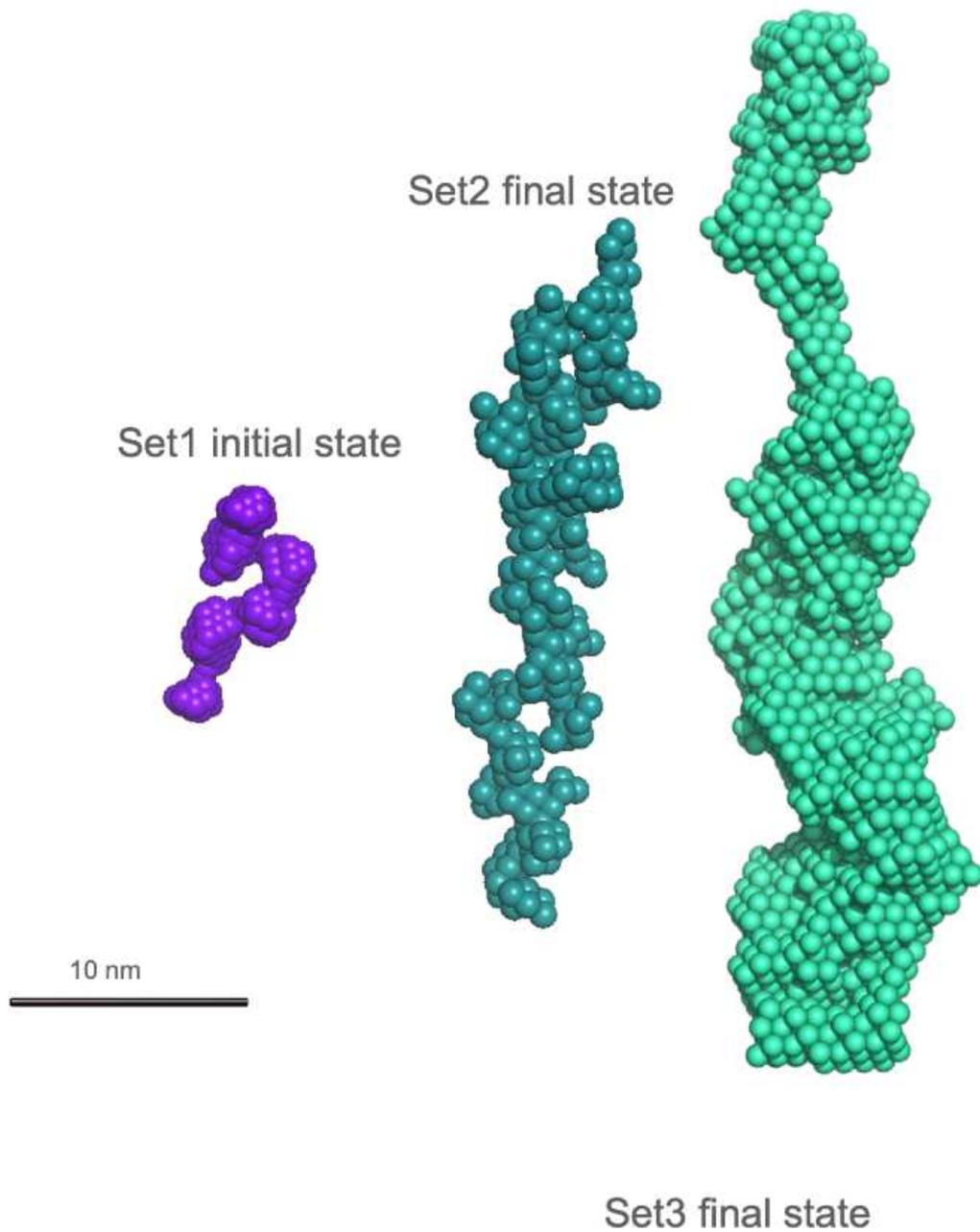
Figure 3.21: Three selected models that represent the sets 1, 2, 3 and 6 the best
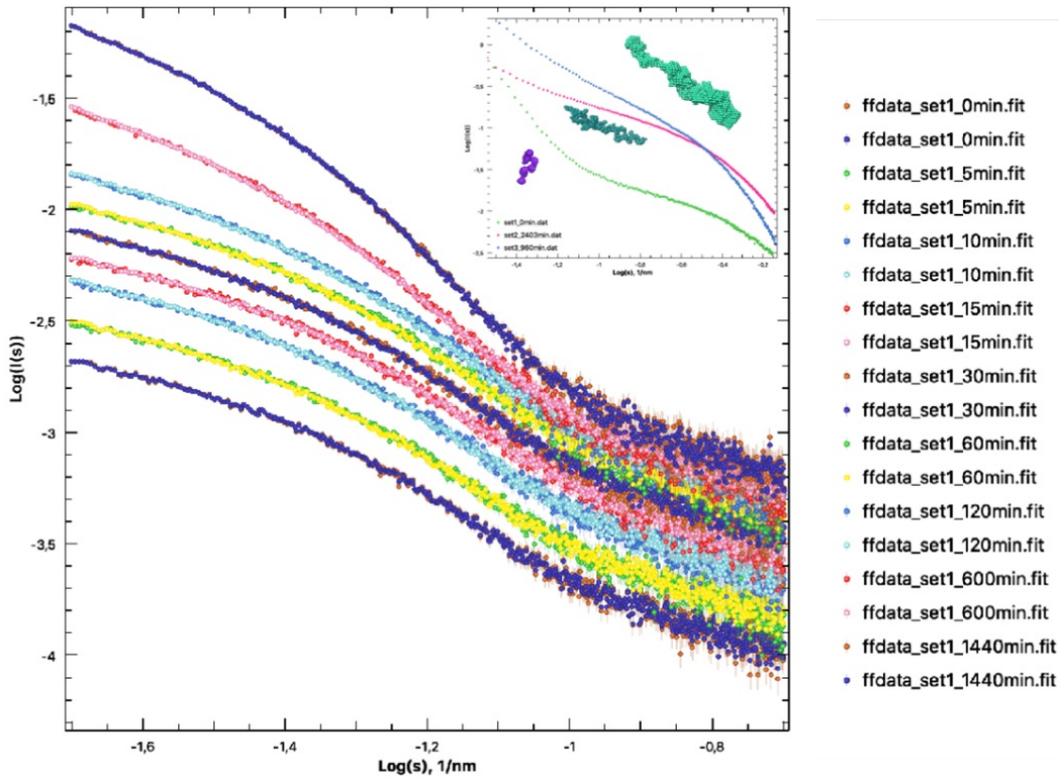
Figure 3.22: Example of fitting Set 1 data with three models
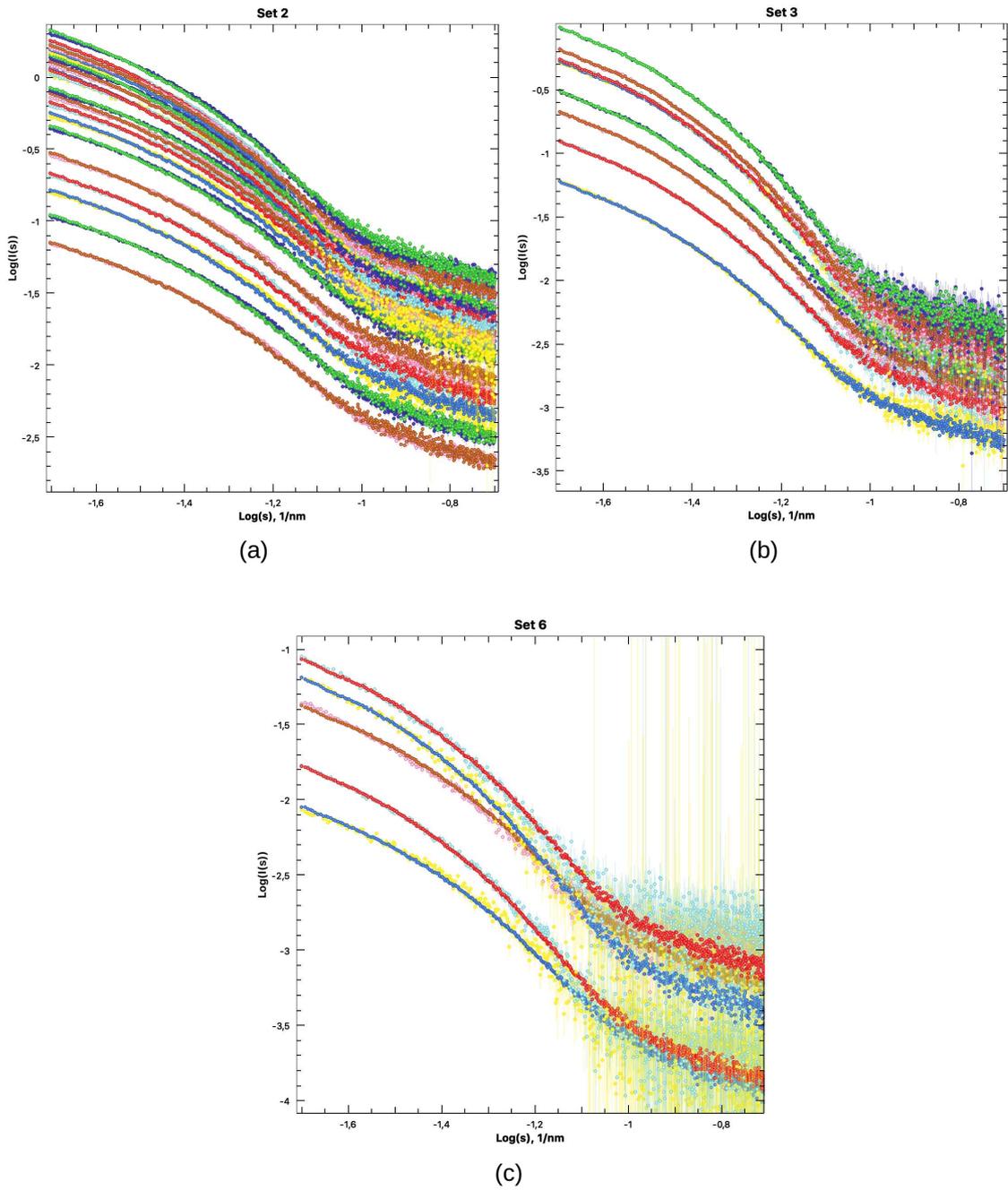
Figure 3.23: Fitting Set 2 (a), 3 (b), 6 (c) with three models
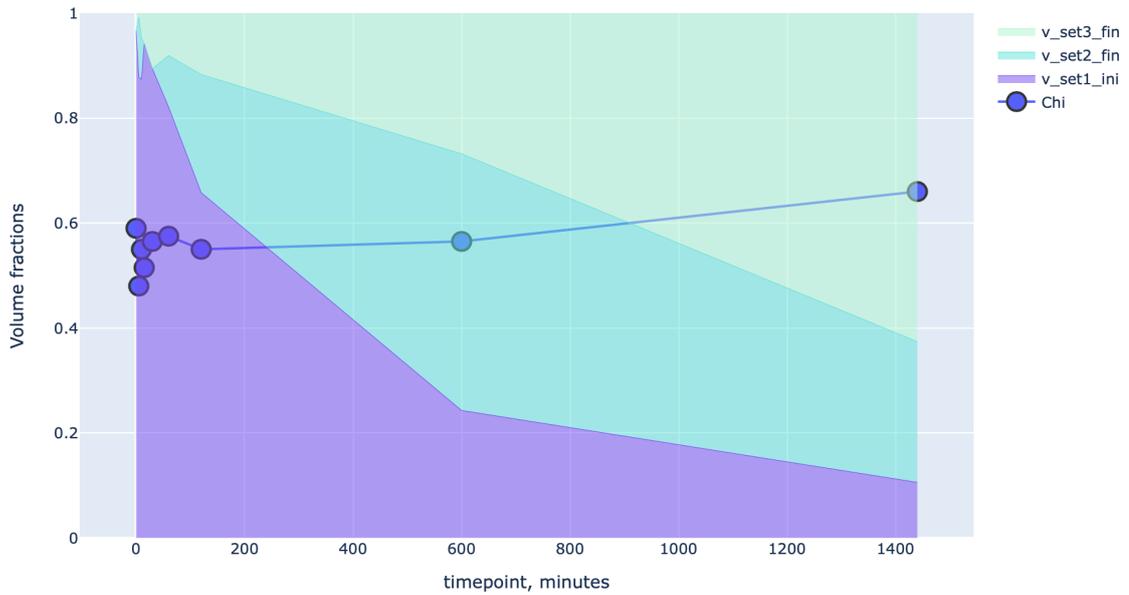
Figure 3.24: Scattering data of Set 1 fitted by three scattering patterns of the selected models. Volume fractions and goodness of fits are plotted against the time points from 0 to 1440 minutes (24 hours).
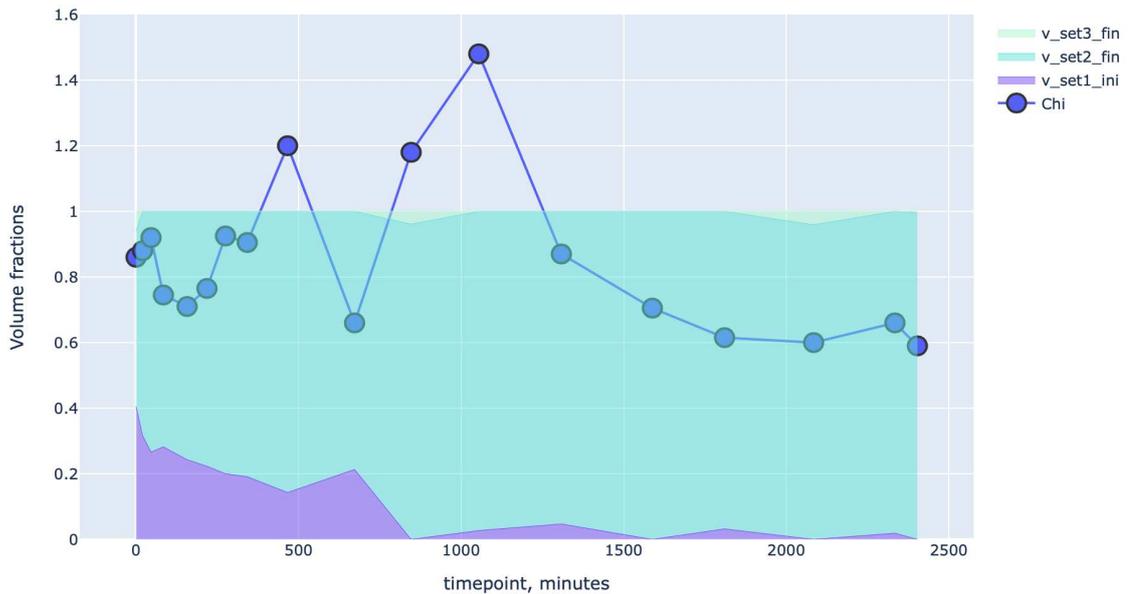
Figure 3.25: Scattering data of Set 2 fitted by three scattering patterns of the selected models. Volume fractions and goodness of fits are plotted against the time points from 0 to 2403 minutes ($\sim$40 hours).
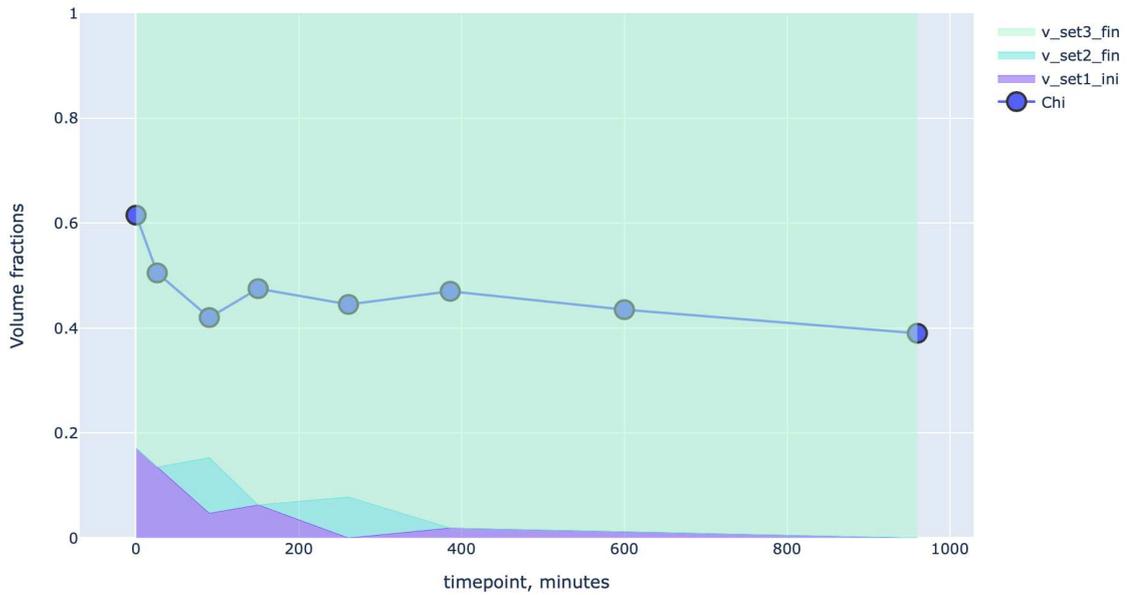
Figure 3.26: Scattering data of Set 3 fitted by three scattering patterns of the selected models. Volume fractions and goodness of fits are plotted against the time points from 0 to 960 minutes (16 hours).
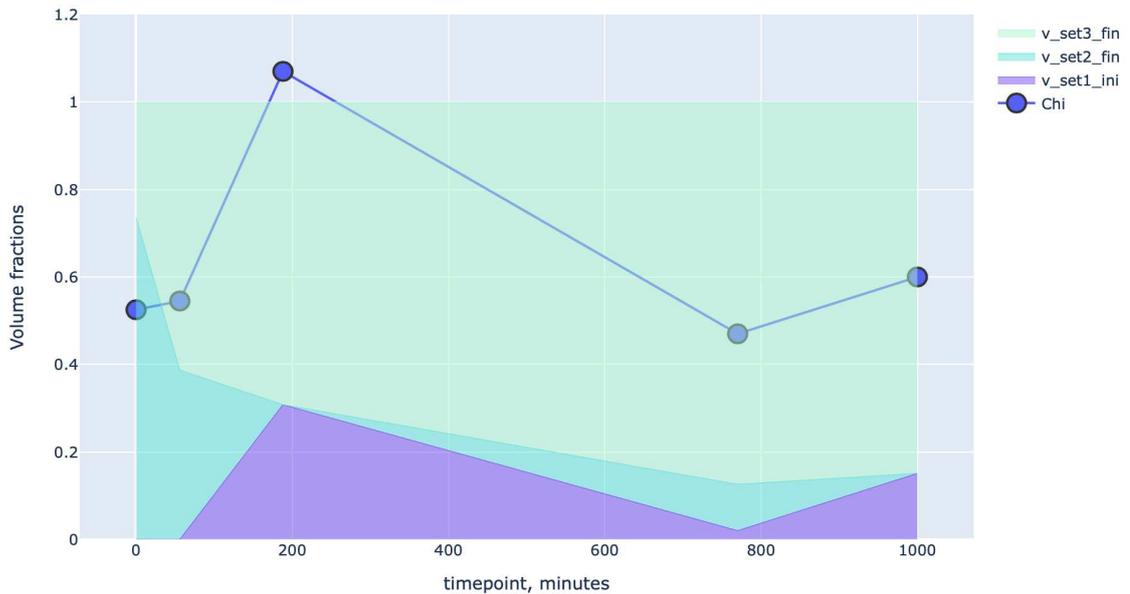
Figure 3.27: Scattering data of Set 6 fitted by three scattering patterns of the selected models. Volume fractions and goodness of fits are plotted against the time points from 0 to 1000 minutes ($\sim$17 hours).

Structural kinetics of the smaller particles for sets 4 and 5

Sets 4 and 5 were analysed together as they represent an evolution of smaller particles, actually, the smallest oligomers detected in our TR-SAXS experiments. The number of possible models to try (2 initial, 2 final and 2 intermediate states) was lower than the number of curves in each of these sets. Exploring combinations of two, three, and four models revealed the combination that describes the experimental data the best. This combination consists of three *ab initio* models, namely the initial and final states of Set 4 and the final state of Set 5 (Figure 3.28).

The models describing sets 4 and 5 are all much smaller than those for sets 1,2,3 6; and the smallest model has ~8 nm in length. The largest model for sets 4 and 5 is the one that of the final state of Set 5 reaching the size ~15 nm. Similarly to sets 1, 2, 3 and 6, the middle size model (final state of Set 4, depicted in red in Figure 3.28) could act as an intermediate in the fibrillation process.
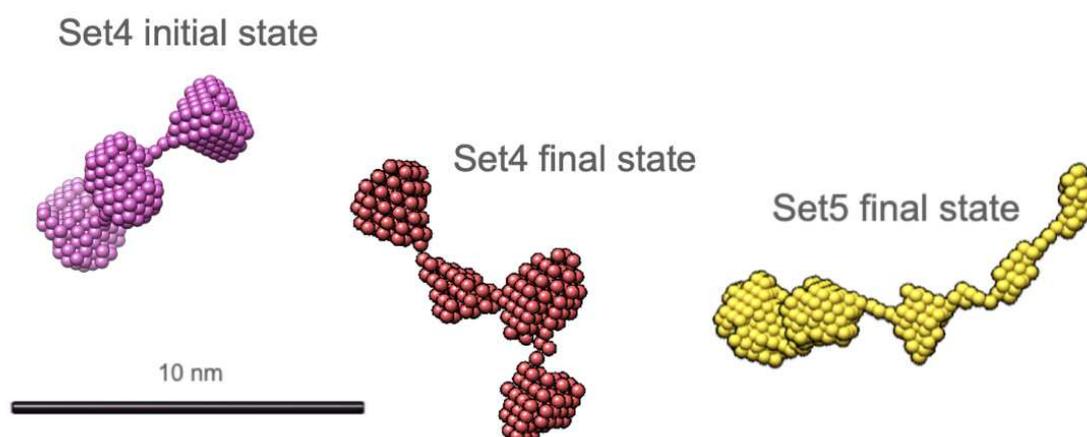
Figure 3.28: Three selected models that represent the sets 4 and 5 the best

Linear combinations of the scattering curves (Figure 3.29) used for building the selected three *ab initio* models fit all the curves of sets 4 and 5 as shown in Figure 3.30.

As shown in Figure 3.31 (a), the fraction of the smallest particles decreases over time and the contribution of the final state model of Set 4 increases. However, there is also a significant fraction of somewhat yet larger species representing the final state of Set 5 (yellow). Its contribution is more prominent between 4 and 6 hours of incubation.

For Set 5 (Figure 3.6, Table 3.2), very little increase in size of the particles could be observed. From Figure 3.31 (b), the final state of Set 5 is reached during the very first hours of incubation. In the beginning, though, the mixture contained about 30% of the smallest particles (magenta), whereas middle-size oligomers
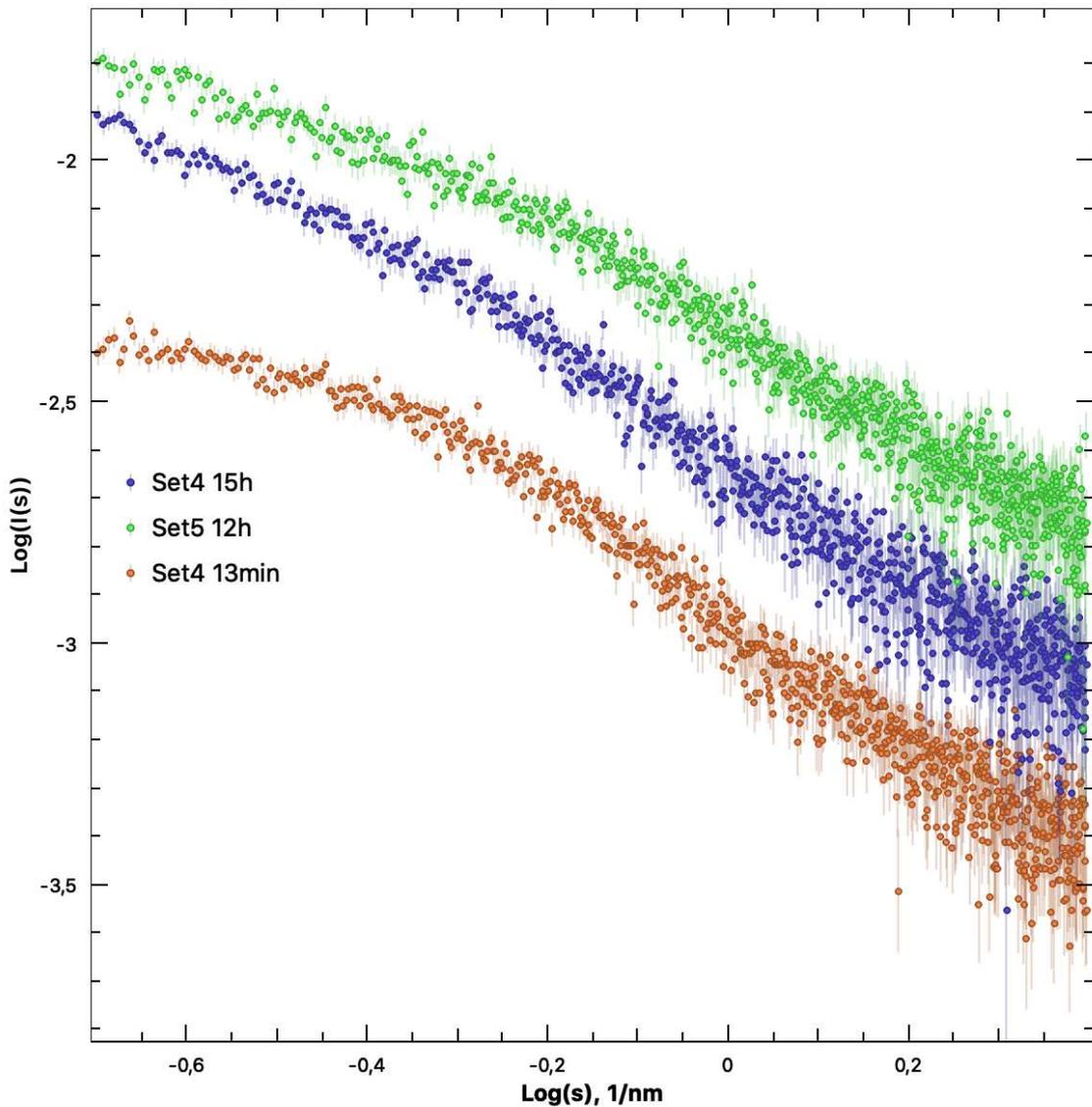
Figure 3.29: Experimental SAXS curves used for building the models that represent sets 4 and 5

(red line close to zero) were present all the time in very insignificant quantities.

In summary, the structural kinetics of sets 4 and 5 was drastically different from the other sets. Here, only very small Aβ oligomers were forming in solution, and 14 experimental SAXS curves were fitted by linear combination from the scattering by three selected models (Figures 3.29, 3.30). These models may be representing potential intermediates of a more complex process of amyloid fibril formation. The formation of yet larger fibrils was apparently not observed for the two sets 4 and 5.
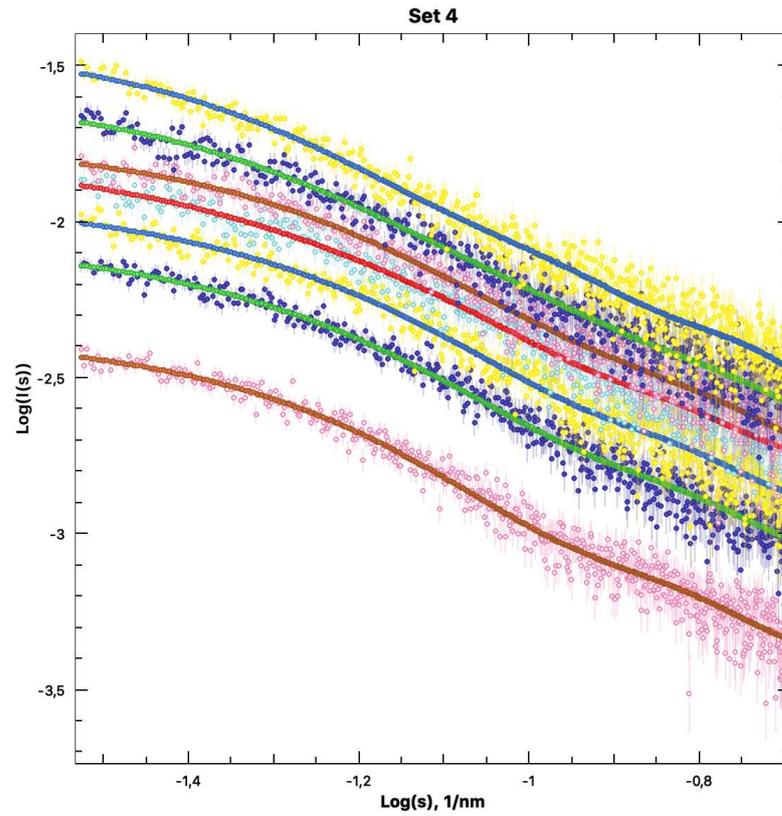
(a)
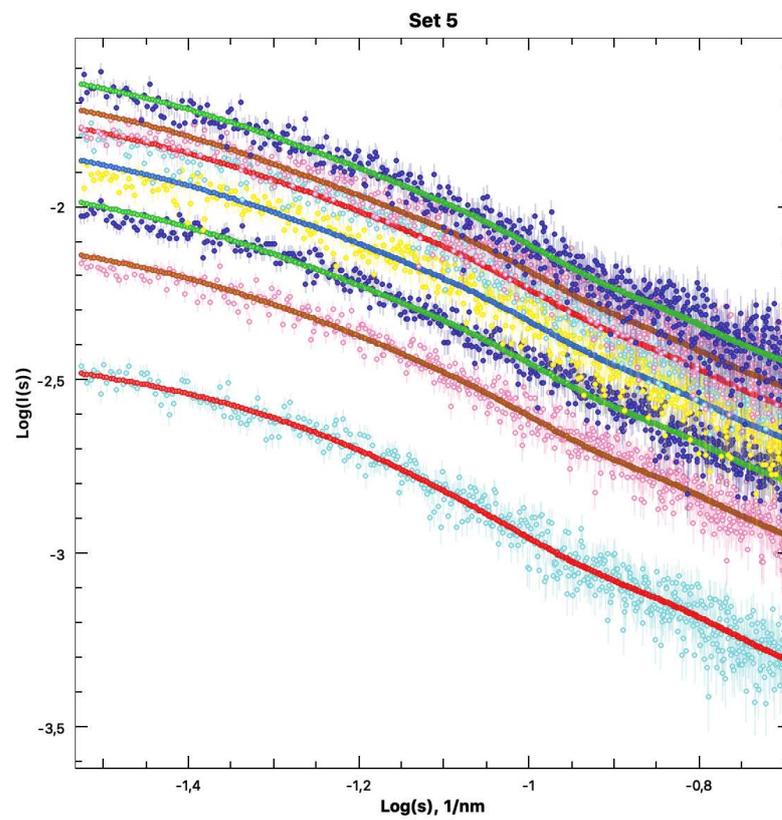
(b)

Figure 3.30: Fitting Set 4 (a) and Set 5 (b) with three models
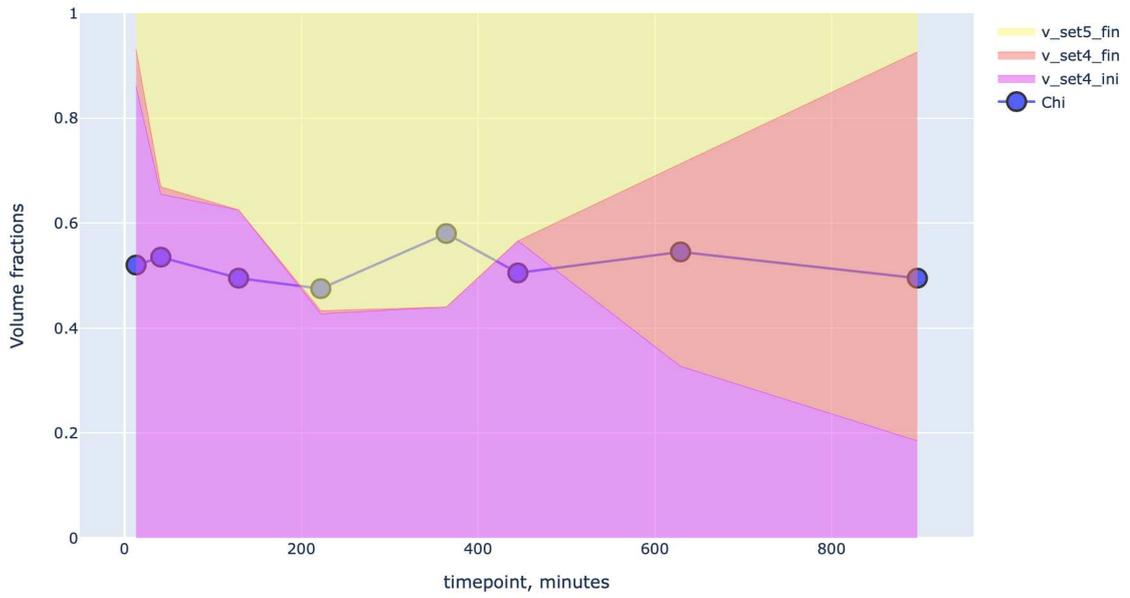
(a)
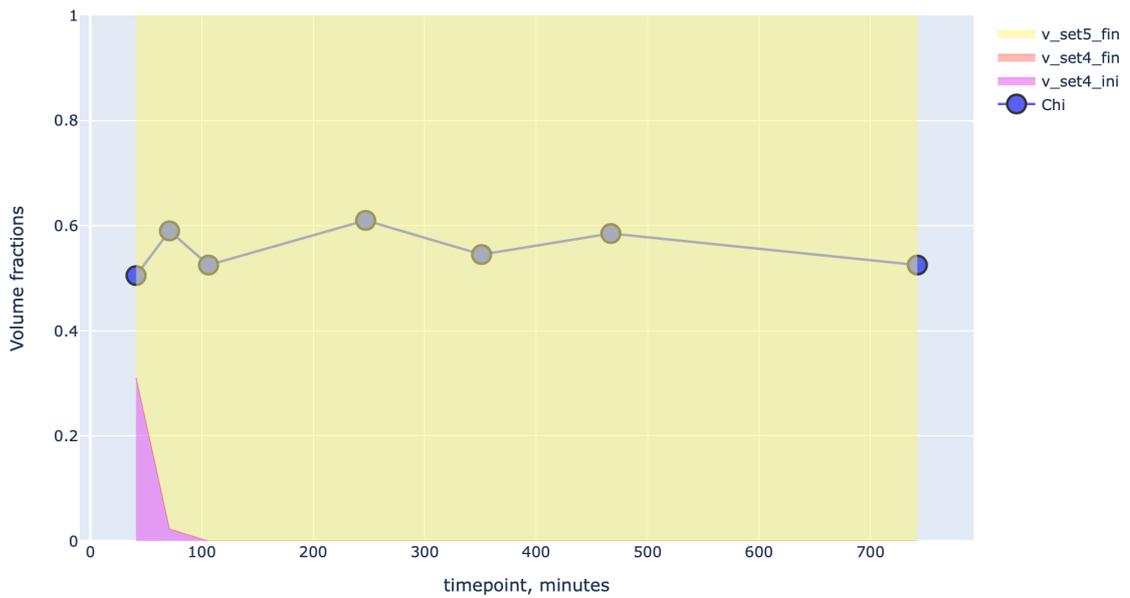
(b)

Figure 3.31: Scattering data of Set 4 (a) and Set 5 (b) fitted by three scattering patterns of the selected models. Volume fractions and discrepancy of fits are plotted against the time points

4 Discussion

4.1 Comparison with recent results on Aβ42

Our data were compared with the recent SAXS data from what was called "final state fibrils" by Lattanzi et al. (2021). In the latter publication, amyloid fibrils of Aβ42 were measured at 37°C after 5 days of incubation in 20 mM phosphate buffer (no time-resolved data were collected). A comparison of Lattanzi et al. (2021) data with those from the largest species observed in our TR-SAXS measurements (final state of Set 3) is given in Figure 4.1. Interestingly, the cross-sectional R_g of Set 3 data is smaller than that of Lattanzi et al. (2021) data, whereas for the overall R_g the opposite result is observed. Quantitatively, the particles as detected for Set 3 are thinner (cross-section $\sim$ 5 nm) compared to those reported by Lattanzi et al. (2021)) (cross-section $\sim$ 9 nm). However, Set 3 provides clearly longer aggregates (with the length over 150 nm) than the ones from Lattanzi et al. (2021) (about 120 nm). This comparison indicates that the results by Lattanzi et al. (2021) do not contradict to our observations, and on the contrary, they add yet another possible polymorphic state belonging to the pathway of Aβ fibrils formation.

Figure 4.1: Scattering from the final state of Set 3 (cyan) scaled to the curve reported by Lattanzi et al. (red)

4.2 Selection of the experimental conditions

Although solution TR-SAXS studies often allow for the use of nearly physiological environments, isolated *in vitro* biomacromolecular systems usually require the conditions that would better stabilise the system or permit to capture specific features of the sample. Indeed, the solvent conditions of the TR-SAXS experiments on the Aβ kinetics described in the thesis were deliberately selected to be rather artificial in order to observe specific phenomena in the kinetics.

The following buffer conditions were also tried:

- 10 mM HEPES pH=7.2 150 mM NaCl;

- 20 mM Acetate pH=5;

- 10 mM Hepes pH=7.5 150 mM NaCl;

- 1 mM Hepes pH=7.5, 0.001% NH_4OH 150 mM NaCl.

These conditions are closer to the physiological ones (pH 7 and 150 mM NaCl), however, they did not easily allow one to properly monitor the kinetics of conformational changes.

Actually, structural and biophysical studies of the Aβ peptide often utilise less physiological conditions. For instance, in Gremer et al. (2017) the fibrils are grown at pH=2, with 30% acetontirile and trifluoroacetic acid. Low-ionic strength conditions were reported in the infrared study by Cerf et al. (2009) for Aβ42 incubations. For

the NMR studies by Colvin et al. (2015), fibrils were prepared in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% NaN3, pH 8.0. The use of true physiological conditions could be more valuable for *ex vivo* studies but sub-optimal for the kinetic analysis. The choice of the particular experimental conditions in the present work was dictated by the necessity to balance between physiological environments and the need to observe different conformations and oligomeric states of Aβ.

4.3 Interpretation of the results in biological context

The observed conformational changes indicate that the process of the amyloid fibrils formation is non-linear and complex. The TR-SAXS experiments conducted for the current thesis revealed two possible pathways of structural kinetics of the Aβ oligomerisation. The described types of structural kinetics may possibly represent different stages of the amyloid fibrils formation present regardless of the conditions.

Different oligomers detected for each repetition of the same process experimentally confirm the concept of the structural instability of these neurotoxic oligomeric intermediates (Michaels et al. (2020)). Moreover, the irreproducibility of the measurements under the same starting conditions in solution demonstrated the polymorphic nature of Aβ fibrils growing in the same region of the human brain, suggesting that different forms of a fibril could be present at the same time (Colletier et al. (2011), Rasmussen et al. (2017)).

4.3.1 The smallest oligomers detected in solution

The smallest species detected for the initial state of Set 4 are slightly larger ($\sim$8 nm) than the monomer of the Aβ peptide ($\sim$5-6 nm). Given that SAXS provides an average shape of the solute, the smallest species should be composed of two or three monomers of the peptide. The largest model employed to fit sets 4 and 5 is still smaller than 20 nm in length suggesting the maximum number of monomers in such an oligomer of about five. Although such small species were revealed, very large aggregates (possibly composed of these species) were present in solution for all the experiments. The presence of such aggregates can be related to the intrinsic attractive interactions between Aβ peptides.

Structural kinetics of the smaller oligomers explored for sets 4 and 5 may represent a cycle of oligomerisation/dissociation at the first stage of the amyloid fibrils formation process if the oligomerisation rate is sufficient. It seems that either the experimental conditions or the specific batch of the commercial peptide could be a reason for preventing the oligomers from further conversion to protofibrils.

Obviously, for these sets, the dissociation rate is higher than the oligomerisation rate. Therefore, it was possible to keep some fraction of the Aβ in its soluble state meaning that the neurotoxicity thereof might be higher than the neurotoxicity of larger oligomeric species that can convert to protofibrils and form plaques destroying neuronal connections in the brain.

4.3.2 Larger oligomers have the potential to form a protofibril

The formation of larger oligomers was detected for sets 1, 2, 3, and 6. These measurements, unlike sets 4 and 5, predominantly revealed the next stage of the process of fibrils formation. At this stage, the unstable oligomers tend to grow attaching the monomers, and the dissociation is less prominent here. Starting with disordered oligomers for Set 1, the peptide develops sub-fibrillar elongated particles that are ~ 40 nm in length. The kinetics of this part of the process showed that the elongation may occur differently. This fact could indicate that the elongation is accompanied by a dissociation occurring in the system at the same time.

The system appears to have not only two stages of the ongoing process, but also two levels of structural organisation and conformational changes. Whereas the oligomers tend to grow, the large fibrils, observed for sets 1, 2, 3 and 6 appear smaller at the later time points. Supposedly, there are Aβ monomers laterally attached to the fibrils. Those monomers could then be donated to the forming sub-fibrillar oligomers while possibly, undergoing a restructuring process, which might be intrinsic for the Aβ peptide (Figure 4.2).

Figure 4.2: Schematic representation of restructuring of the amyloid fibrils

The observed results and the above mechanism would fit the hypotheses discussed in other studies:

- the process of the amyloid fibrils formation involves oligomerisation, dissociation, nucleation and elongation (Michaels et al. (2020));

- conformational polymorphism of Aβ: the fibrils may form along different pathways independently of the conditions (Colletier et al. (2011), Rasmussen et al. (2017)).

Merging these two concepts, an important outcome of the present TR-SAXS study is that the structural kinetics of Aβ fibrils formation reveals a polymorphism of the process itself. TR-SAXS results reveal non-linear structural kinetics and conformational polymorphism of the amyloid fibrils formation *in vitro*. The analysis of the SAXS data allowed us to construct the structural models of the oligomeric aggregates formed during the aggregation process.

4.4   Automation and applicability of the pipeline to further experiments

The KinConG pipeline was tailored to analyse the data obtained from TR-SAXS experiments on the Aβ peptide. However, the pipeline has many adjustable

parameters and can be applied to other series of successively recorded SAXS data. For other than time-resolved series, the user would explicitly tell KinConG which SAXS curves should be considered as the initial and final states. For example, measurements of evolving samples can be conducted upon gradually changing physical parameters (e.g. temperature or pH).

By default, the pipeline may analyse the data series in the full measured angular range. Depending on the sample, different conformational changes may be expected in its structural organisation, and, therefore, specific angular ranges could be monitored. The user can tell the KinConG which levels of the structural organisation should be explored.

The analysis starts with calculating the P(r) function. A default computation provides the P(r) function based on the full angular range of a SAXS curve, mainly to assess the largest distance in the sample. This works best for monodisperse solutions of compact proteins; for more complicated systems, *a priori* information about the samples is useful to better compute the P(r) function. Here, the user may select the desired s-range and pass the limits to the KinConG. s_{min} would be considered as a starting point and refer to the resolution d (not the maximum size of a particle but the maximum size explored) in real space ($d = 2\pi/s$). s_{max} would be considered the last point to analyse and defines how detailed the P(r) function would describe the inner distances in the sample. An expert user may also pass pre-made P(r) function(s) to the pipeline.

Also, the *ab initio* modelling can be guided by a user. The default search volume is defined by a sphere filled with dummy atoms. For the extended fibrils, the modelling of large oligomers (> 30 nm) was performed here in a cylindrical search volume. Using the default search volume is much less biased but could provide artefacts in the reconstructed shapes. Generally, the usage of the pipeline requires more knowledge of the samples but less expertise in the SAXS data analysis.

4.5   Conclusions

In the thesis, we have characterised the conformational polymorphism of the Aβ fibrils formation process. In particular, we have analysed how and why are the fibrils forming differently at two levels of their structural organisation (as growing oligomers or thinning fibrils) depending on their initial and final states and the environment. We further described the mechanism of oligomers' formation and structural rearrangements in these systems. The collected TR-SAXS data from the Aβ samples in solution and the models are presently being deposited into the

Small Angle Scattering Biological Data Bank (Kikhney et al. (2020)) to be available for the community (the project link: https://www.sasbdb.org/project/1842/).

The analysis of complex systems is still far from being fully automated and requires a human's creative approach and a tremendous amount of effort and time to be invested in the identification of structure-property relationships. Even the application of modern dimensionality reduction and visualisation techniques should be considered only as a first step that can give inspiration for further analysis (Maksimov (2022)). For the heterogenous $A\beta$ systems, we found that it is advantageous to start from a comprehensive sampling of the conformational states of the possible intermediates of the process using a less constrained modelling (in our case, *ab initio* dummy atom modelling in the spherical search volume). Indeed, in this modelling, the observed intermediate conformers lie among structural conformations that were already present in the solution of $A\beta$ as the initial and final states, albeit usually not sufficient to fully describe the system. In the general case, the solutions may also contain other conformational states than just the initial and final ones as can be seen from the sampling of the hidden intermediate $A\beta$ structures, the variability of which increasing for the sparser data sets.

Although the possible hidden oligomeric intermediates of $A\beta$ could be found by this approach, based on the single-experiment analysis we were not able to capture consistent intermediate states for all the experimental series. The hidden intermediates may serve as an initial guess for further studies. Data such as those we collected provide a source of knowledge of the structural kinetics of the $A\beta$ oligomers and their agglomerates. TR-SAXS results for the sets 1,2,3 and 6 indicate that the oligomerisation of $A\beta$ peptide leads to the formation of long fibrils that may also dissociate into monomers or other smaller oligomers. This is different for the data collected on sets 4 and 5. The dissociation rate kinetically constrains the conformational changes and oligomerisation of the $A\beta42$ peptide, thus decreasing the fractions of the observed large oligomeric structures with respect to the fractions of the small species for these sets.

When aggregated, $A\beta$ donates monomers to the forming oligomers, which are becoming somewhat thinner. Focusing on the neurotoxic oligomers of the $A\beta$, we do not observe fully extended structures in the selected angular range, and most oligomers exhibit intramolecular conformational changes.

Time-resolved SAXS experiments on biological processes are rather challenging to analyse. Interpretation of the results in the biological context requires an advanced level of expertise and competence on the samples and on the processes. Automated and easy-to-use data analysis tools are valuable, especially for novice users on large-scale facilities like biological SAXS beamlines. Nowadays, data

engineering procedures are crucial to get the information correctly processed and analysed. Automation of these procedures is very important for reducing human error and unifying the experimental data. Specific SAXS data analysis tools usually require a certain format of input files and they should be executed in a pre-defined order before the results are ready to be interpreted. The pipeline presented in this work is a convenient tool, which allows the user to focus more on the biological interpretation of the experimental results.

The pipeline is capable of handling multiple datasets in parallel and the suggested workflow allows for rapid processing of the TR-SAXS data. Initially devised for the analysis of amyloid fibrils formation, it can be tailored to other studies of structural kinetics with TR-SAXS.

Abstract

Structural and functional biophysical studies often require temporal component to explore the kinetics of processes in macromolecular systems. The processes like amyloid fibrils formation and protein aggregation involve consequent chemical reactions occurring under native conditions. Small-angle X-ray scattering (SAXS) is a structural method allowing one to capture conformational changes and measure the kinetics of the macromolecules and complexes in near-native solutions. For functional biological complexes, it is important not only to observe structural changes but also to recognise their biological implications. As the analysis of one-dimensional SAXS data in terms of three-dimensional (3D) models is an ill-posed problem, and the analysis of kinetics needs the detection of time-dependent changes, characteristic times of the structural changes need to be defined to analyse large amounts of time-resolved data. Here, dimensionality reduction techniques were applied to structurally characterise the amyloid fibrils formation process. The computational approach utilized ATSAS software, an established set of tools for small-angle scattering data analysis. ATSAS is capable of extracting structural information from the experimental SAXS data utilising, if available, models provided by other methods. This allows one to combine the structural information into a wealth of biophysical and biochemical evidence. Although all ATSAS tools are straightforward to use, the data analysis still requires a significant level of expertise to interactively utilise the tools when dealing with time-resolved studies. To optimise and simplify the data analysis procedures for the analysis of processes occurring in biomacromolecular systems, a new pipeline has been developed. Its capacity is illustrated by the application to the time-resolved data on amyloid fibrils formation in solution. The SAXS data processed with the pipeline revealed conformational polymorphism of the amyloid fibrils formation and allowed one to determine shapes of the starting/final states and of the possible intermediates in the process.

Zusammenfassung

Strukturelle und funktionelle biophysikalische Studien erfordern häufig eine zeitliche Komponente um die Kinetik von Prozessen in makromolekularen Systemen zu untersuchen. Die Prozesse, wie die Bildung von Amyloidfibrillen und die Proteinaggregation, umfassen nachfolgende chemische Reaktionen, die unter natürlichen Bedingungen stattfinden. Die Röntgenkleinwinkelstreuung (KWS) ist eine Strukturmethode, die es ermöglicht, Konformationsänderungen zu erfassen und die Kinetik von Makromolekülen und Komplexen in nahezu nativen Lösungen zu messen. Für funktionelle biologische Komplexe ist es wichtig, nicht nur strukturelle Veränderungen zu beobachten, sondern auch ihre biologischen Implikationen zu erkennen. Da die Analyse eindimensionaler KWS-Daten im Form von dreidimensionalen (3D) Modellen ein schwieriges Problem darstellt und die Analyse der Kinetik die Detektion von zeitabhängigen Änderungen erfordert, müssen charakteristische Zeitpunkte der strukturellen Änderungen definiert werden, um große Mengen von zeitaufgelösten Daten zu analysieren. Hier wurden Dimensionsreduktiontechniken angewendet, um den Bildungsprozess der Amyloidfibrillen strukturell zu charakterisieren. Der entwickelte Ansatz basiert auf dem ATSAS-Software, einem bewährten Werkzeug für die Analyse von Kleinwinkelstreuungsdaten. ATSAS ist in der Lage, Strukturinformationen aus den experimentellen KWS-Daten zu extrahieren, wobei, falls verfügbar, Modelle verwendet werden, die von anderen Methoden bereitgestellt werden. Dies ermöglicht es, die Strukturinformationen zu einer Vielzahl von biophysikalischen und biochemischen Beweisen zu kombinieren. Obwohl alle ATSAS-Tools einfach zu bedienen sind, erfordert die Datenanalyse dennoch ein erhebliches Maß an Fachwissen, um die Tools interaktiv zu nutzen, wenn es um zeitaufgelöste Studien geht. Zur Optimierung und Vereinfachung der Verfahren zur Analyse von Prozessen in biomakromolekularen Systemen wurde eine neue Pipeline entwickelt. Seine Leistungsfähigkeit wird durch die Anwendung auf die zeitaufgelösten Daten zur Bildung von Amyloidfibrillen in Lösung veranschaulicht. Die mit der Pipeline verarbeiteten KWS-Daten offenbaren einen Konformationspolymorphismus der Amyloidfibrillenbildung und ermöglichten die Bestimmung der Formen der Start-/Endzustände und der möglichen Zwischenprodukte im Prozess.

Abbreviations

1D one-dimensional. 9, 29

2D two-dimensional. 9, 15

3D three-dimensional. 9

Aβ amyloid beta. 3–5, 12–15, 18, 20, 24, 27, 28, 33, 46, 49–55

AD Alzheimer's disease. 2, 3

APP amyloid precursor protein. 3

D$_{max}$ maximum dimension of the particle. 9

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid. 14, 50

KinConG aimed to process SAXS data studying the Kinetics of Conformational chanGes. 28, 29, 31, 33, 37, 53, 54

MD molecular dynamics. 33

NMR nuclear magnetic resonance. 1, 51

R$_g$ radius of gyration. 6, 9, 16, 49

SANS small-angle neutron scattering. 5, 7, 33

SAXS small-angle X-ray scattering. 5–9, 12, 14–16, 19, 21–24, 29, 35, 38, 39, 46, 49, 51, 53–56

SVD singular value decomposition. 10–12, 16, 29, 31, 37

TR-SAXS time-resolved small-angle X-ray scattering. 5, 8, 10, 12–16, 18, 28–31, 38, 45, 49–51, 53–56

Bibliography

Alzheimer's Association (2022). Inside the brain. https://www.alz.org/alzheimers-dementia/what-is-alzheimers/brain_tour.

Bartolini, M., Bertucci, C., Bolognesi, M. L., Cavalli, A., Melchiorre, C., and Andrisano, V. (2007). Insight into the kinetic of amyloid β (1–42) peptide self-aggregation: Elucidation of inhibitors' mechanism of action. *ChemBioChem*, 8(17):2152–2161.

Bartolini, M., Naldi, M., Fiori, J., Valle, F., Biscarini, F., Nicolau, D. V., and Andrisano, V. (2011). Kinetic characterization of amyloid-beta 1–42 aggregation with a multimethodological approach. *Analytical biochemistry*, 414(2):215–225.

Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007). Structural characterization of flexible proteins using small-angle x-ray scattering. *Journal of the American Chemical Society*, 129(17):5656–5664.

Blanchet, C. E., Spilotros, A., Schwemmer, F., Graewert, M. A., Kikhney, A., Jeffries, C. M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., et al. (2015). Versatile sample environments and automation for biological solution x-ray scattering experiments at the p12 beamline (petra iii, desy). *Journal of applied crystallography*, 48(2):431–443.

Breßler, I., Kohlbrecher, J., and Thünemann, A. F. (2015). Sasfit: a tool for small-angle scattering data analysis using a library of analytical expressions. *Journal of applied crystallography*, 48(5):1587–1598.

Businger, P. A. and Golub, G. H. (1969). Algorithm 358: Singular value decomposition of a complex matrix [f1, 4, 5]. *Communications of the ACM*, 12(10):564–565.

Carter, R. (2019). *The brain book: An illustrated guide to its structure, functions, and disorders*. Dorling Kindersley Ltd.

Cerf, E., Sarroukh, R., Tamamizu-Kato, S., Breydo, L., Derclaye, S., Dufrêne, Y. F., Narayanaswami, V., Goormaghtigh, E., Ruysschaert, J.-M., and Raussens, V.

(2009). Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. *Biochemical Journal*, 421(3):415–423.

Chen, Y., Tokuda, J. M., Topping, T., Sutton, J. L., Meisburger, S. P., Pabit, S. A., Gloss, L. M., and Pollack, L. (2014). Revealing transient structures of nucleosomes as dna unwinds. *Nucleic acids research*, 42(13):8767–8776.

Cheng, Y. (2015). Single-particle cryo-em at crystallographic resolution. *Cell*, 161(3):450–457.

Ciccone, L., Tonali, N., Nencetti, S., and Orlandini, E. (2020). Natural compounds as inhibitors of transthyretin amyloidosis and neuroprotective agents: Analysis of structural data for future drug design. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 35(1):1145–1162.

Colletier, J.-P., Laganowsky, A., Landau, M., Zhao, M., Soriaga, A. B., Goldschmidt, L., Flot, D., Cascio, D., Sawaya, M. R., and Eisenberg, D. (2011). Molecular basis for amyloid-β polymorphism. *Proceedings of the National Academy of Sciences*, 108(41):16938–16943.

Colvin, M. T., Silvers, R., Frohm, B., Su, Y., Linse, S., and Griffin, R. G. (2015). High resolution structural characterization of aβ42 amyloid fibrils by magic angle spinning nmr. *Journal of the American Chemical Society*, 137(23):7509–7518.

Cowieson, N. P., Kobe, B., and Martin, J. L. (2008). United we stand: combining structural methods. *Current opinion in structural biology*, 18(5):617–622.

Durbin, J. and Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, 37(3/4):409–428.

Feigin, L. and Svergun, D. (1987). *Structure analysis by small-angle X-ray and neutron scattering*, volume 1. Springer.

Ferrer, N., Serra, E., Sempere, J., and Nomen, R. (2017). Non-parametric kinetic analysis of autocatalytic reactions. *Journal of Loss Prevention in the Process Industries*, 49:357–366.

Fischer, N., Konevega, A. L., Wintermeyer, W., Rodnina, M. V., and Stark, H. (2010). Ribosome dynamics and trna movement by time-resolved electron cryomicroscopy. *Nature*, 466(7304):329–333.

Fodera, V., Zaccone, A., Lattuada, M., and Donald, A. M. (2013). Electrostatics controls the formation of amyloid superstructures in protein aggregation. *Physical review letters*, 111(10):108105.

Fowler, A., Foote, A., Moody, M., Vachette, P., Prrovencher, S., Gabriel, A., Bordas, J., and Koch, M. (1983). Stopped-flow solution scattering using synchrotron radiation: apparatus, data collection and data analysis. *Journal of Biochemical and Biophysical Methods*, 7(4):317–329.

Franke, D., Jeffries, C. M., and Svergun, D. I. (2015). Correlation map, a goodness-of-fit test for one-dimensional x-ray scattering spectra. *Nature methods*, 12(5):419–422.

Franke, D., Kikhney, A. G., and Svergun, D. I. (2012). Automated acquisition and analysis of small angle x-ray scattering data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 689:52–59.

Franke, D., Petoukhov, M., Konarev, P., Panjkovich, A., Tuukkanen, A., Mertens, H., Kikhney, A., Hajizadeh, N., Franklin, J., Jeffries, C., et al. (2017). Atsas 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of applied crystallography*, 50(4):1212–1225.

Franke, D. and Svergun, D. I. (2009). Dammif, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of applied crystallography*, 42(2):342–346.

Funato, H., Yoshimura, M., Yamazaki, T., Saido, T. C., Ito, Y., Yokofujita, J., Okeda, R., and Ihara, Y. (1998). Astrocytes containing amyloid beta-protein (abeta)-positive granules are associated with abeta40-positive diffuse plaques in the aged human brain. *The American journal of pathology*, 152(4):983.

Golub, G. and Kahan, W. (1965). Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, 2(2):205–224.

Golub, G. H. and Reinsch, C. (1971). Singular value decomposition and least squares solutions. In *Linear algebra*, pages 134–151. Springer.

Grant, T. D. (2018). Ab initio electron density determination directly from solution scattering data. *Nature methods*, 15(3):191–193.

Gremer, L., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R. B., Tusche, M., Lopez-Iglesias, C., Hoyer, W., Heise, H., et al. (2017). Fibril structure of amyloid-β (1–42) by cryo–electron microscopy. *Science*, 358(6359):116–119.

Guinier, A. (1939). La diffraction des rayons x aux très petits angles : application à l'étude de phénomènes ultramicroscopiques. *Ann. Phys.*, 11(12):161–237.

Haass, C. and Selkoe, D. J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the alzheimer's amyloid β-peptide. *Nature reviews Molecular cell biology*, 8(2):101–112.

Hardy, J. and Selkoe, D. J. (2002). The amyloid hypothesis of alzheimer's disease: progress and problems on the road to therapeutics. *science*, 297(5580):353–356.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Hopkins, J. B., Gillilan, R. E., and Skou, S. (2017). Bioxtas raw: improvements to a free open-source program for small-angle x-ray scattering data reduction and analysis. *Journal of applied crystallography*, 50(5):1545–1553.

Ikeda, K., Yamaguchi, T., Fukunaga, S., Hoshino, M., and Matsuzaki, K. (2011). Mechanism of amyloid β-protein aggregation mediated by gm1 ganglioside clusters. *Biochemistry*, 50(29):6433–6440.

Karplus, M. and McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature structural biology*, 9(9):646–652.

Kikhney, A. G., Borges, C. R., Molodenskiy, D. S., Jeffries, C. M., and Svergun, D. I. (2020). Sasbdb: Towards an automatically curated and validated repository for biological scattering data. *Protein Science*, 29(1):66–75.

Kikhney, A. G. and Svergun, D. I. (2015). A practical guide to small angle x-ray scattering (saxs) of flexible and intrinsically disordered proteins. *FEBS letters*, 589(19):2570–2577.

Kirby, N. M. and Cowieson, N. P. (2014). Time-resolved studies of dynamic biomolecules using small angle x-ray scattering. *Current Opinion in Structural Biology*, 28:41–46.

Kirkpatrick, P. and Baez, A. V. (1948). Formation of optical images by x-rays. *Journal of the Optical Society of America*, 38 9:766–74.

Konarev, P. V., Graewert, M. A., Jeffries, C. M., Fukuda, M., Cheremnykh, T. A., Volkov, V. V., and Svergun, D. I. (2022). Efamix, a tool to decompose inline chromatography saxs data from partially overlapping components. *Protein Science*, 31(1):269–282.

Konarev, P. V. and Svergun, D. I. (2018). Direct shape determination of intermediates in evolving macromolecular solutions from small-angle scattering data. *IUCrJ*, 5(4):402–409.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., and Svergun, D. I. (2003). Primus: a windows pc-based system for small-angle scattering data analysis. *Journal of applied crystallography*, 36(5):1277–1282.

Koukos, P. and Bonvin, A. (2020). Integrative modelling of biomolecular complexes. *Journal of molecular biology*, 432(9):2861–2881.

Koutsioubas, A. and Perez, J. (2013). Incorporation of a hydration layer in the-dummy atom'ab initio structural modelling of biological macromolecules. *Journal of Applied Crystallography*, 46(6):1884–1888.

Kuo, Y.-M., Emmerling, M. R., Lampert, H. C., Hempelman, S. R., Kokjohn, T. A., Woods, A. S., Cotter, R. J., and Roher, A. E. (1999). High levels of circulating aβ42 are sequestered by plasma proteins in alzheimer's disease. *Biochemical and biophysical research communications*, 257(3):787–791.

Lattanzi, V., André, I., Gasser, U., Dubackic, M., Olsson, U., and Linse, S. (2021). Amyloid β 42 fibril structure based on small-angle scattering. *Proceedings of the National Academy of Sciences*, 118(48):e2112783118.

Likhtenshtein, G. I., Febbraio, F., and Nucci, R. (2000). Intramolecular dynamics and conformational transition in proteins studied by biophysical labelling methods. common and specific features of proteins from thermophylic microorganisms. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 56(10):2011–2031.

Liu, R., Li, M., Liu, Z.-P., Wu, J., Chen, L., and Aihara, K. (2012). Identifying critical transitions and their leading biomolecular networks in complex diseases. *Scientific reports*, 2(1):1–9.

Mäeots, M.-E., Lee, B., Nans, A., Jeong, S.-G., Esfahani, M. M., Ding, S., Smith, D. J., Lee, C.-S., Lee, S. S., Peter, M., et al. (2020). Modular microfluidics enables kinetic insight from time-resolved cryo-em. *Nature communications*, 11(1):1–14.

Maksimov, D. (2022). Characterization and prediction of peptide structures on inorganic surfaces. Technical report, EPFL.

Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D., Gruzinov, A., Borges, C., et al. (2021). Atsas 3.0: expanded functionality and new tools for small-angle scattering data analysis. *Journal of Applied Crystallography*, 54(1).

Meisburger, S. P., Xu, D., and Ando, N. (2021). Regals: a general method to deconvolve x-ray scattering data from evolving mixtures. *IUCrJ*, 8(2).

Meisl, G., Kirkegaard, J. B., Arosio, P., Michaels, T. C., Vendruscolo, M., Dobson, C. M., Linse, S., and Knowles, T. P. (2016). Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nature protocols*, 11(2):252.

Messa, M., Colombo, L., Del Favero, E., Cantù, L., Stoilova, T., Cagnotto, A., Rossi, A., Morbin, M., Di Fede, G., Tagliavini, F., et al. (2014). The peculiar role of the a2v mutation in amyloid-β (aβ) 1–42 molecular assembly. *Journal of Biological Chemistry*, 289(35):24143–24152.

Michaels, T. C., Šarić, A., Curk, S., Bernfur, K., Arosio, P., Meisl, G., Dear, A. J., Cohen, S. I., Dobson, C. M., Vendruscolo, M., et al. (2020). Dynamics of oligomer populations formed during the aggregation of alzheimer's aβ42 peptide. *Nature chemistry*, 12(5):445–451.

Mitić, S., Strampraad, M. J., Hagen, W. R., and de Vries, S. (2017). Microsecond time-scale kinetics of transient biochemical reactions. *PloS one*, 12(10):e0185888.

Mouchlis, V. D., Melagraki, G., Zacharia, L. C., and Afantitis, A. (2020). Computer-aided drug design of β-secretase, γ-secretase and anti-tau inhibitors for the discovery of novel alzheimer's therapeutics. *International Journal of Molecular Sciences*, 21(3):703.

Narayanan, T., Diat, O., and Bösecke, P. (2001). Saxs and usaxs on the high brilliance beamline at the esrf. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 467:1005–1009.

Oang, K. Y., Yang, C., Muniyappan, S., Kim, J., and Ihee, H. (2017). Svd-aided pseudo principal-component analysis: A new method to speed up and improve determination of the optimum kinetic model from time-resolved data. *Structural Dynamics*, 4(4):044013.

Panjkovich, A. and Svergun, D. I. (2018). Chromixs: automatic and interactive analysis of chromatography-coupled small-angle x-ray scattering data. *Bioinformatics*, 34(11):1944–1946.

Pearson, K. (1900). X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

Pedersen, W. A., Kloczewiak, M. A., and Blusztajn, J. K. (1996). Amyloid beta-protein reduces acetylcholine synthesis in a cell line derived from cholinergic neurons of the basal forebrain. *Proceedings of the National Academy of Sciences*, 93(15):8068–8071.

Pepys, M. (2001). Pathogenesis, diagnosis and treatment of systemic amyloidosis. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1406):203–211.

Plotly Technologies Inc. (2015). Collaborative data science.

Rambaran, R. N. and Serpell, L. C. (2008). Amyloid fibrils: abnormal protein assembly. *Prion*, 2(3):112–117.

Rasmussen, J., Mahler, J., Beschorner, N., Kaeser, S. A., Häsler, L. M., Baumann, F., Nyström, S., Portelius, E., Blennow, K., Lashley, T., Fox, N. C., Sepulveda-Falla, D., Glatzel, M., Oblak, A. L., Ghetti, B., Nilsson, K. P. R., Hammarström, P., Staufenbiel, M., Walker, L. C., and Jucker, M. (2017). Amyloid polymorphisms constitute distinct clouds of conformational variants in different etiological subtypes of alzheimer's disease. *Proceedings of the National Academy of Sciences*, 114(49):13018–13023.

Reback, J., jbrockmendel, McKinney, W., den Bossche, J. V., Roeschke, M., Augspurger, T., Hawkins, S., Cloud, P., gfyoung, Sinhrks, Hoefler, P., Klein, A., Petersen, T., Tratner, J., She, C., Ayd, W., Naveh, S., Darbyshire, J., Shadrach, R., Garcia, M., Schendel, J., Hayden, A., Saxton, D., Gorelli, M. E., Li, F., Wörtwein, T., Zeitlin, M., Jancauskas, V., McMaster, A., and Li, T. (2022). pandas-dev/pandas: Pandas 1.4.3.

Sagar, A., Herranz-Trillo, F., Langkilde, A. E., Vestergaard, B., and Bernadó, P. (2021). Structure and thermodynamics of transient protein-protein complexes by chemometric decomposition of saxs datasets. *Structure*.

SasView (2022). Sasview.

Sausville, E. A. (2002). Complexities in the development of cyclin-dependent kinase inhibitor drugs. *Trends in molecular medicine*, 8(4):S32–S37.

Schmidt, M., Rajagopal, S., Ren, Z., and Moffat, K. (2002). Application of singular value decomposition to the analysis of time-resolved macromolecular x-ray data. *Biophysical journal*, 84(3):2112–2129.

Schmidt, M., Rohou, A., Lasker, K., Yadav, J. K., Schiene-Fischer, C., Fändrich, M., and Grigorieff, N. (2015). Peptide dimer structure in an aβ (1–42) fibril visualized with cryo-em. *Proceedings of the National Academy of Sciences*, 112(38):11858–11863.

Shirahama, T. and Cohen, A. S. (1967). High-resolution electron microscopic analysis of the amyloid fibril. *The Journal of cell biology*, 33(3):679–708.

Shukla, D., Meng, Y., Roux, B., and Pande, V. S. (2014). Activation pathway of src kinase reveals intermediate states as targets for drug design. *Nature communications*, 5(1):1–11.

Snyder, S. W., Ladror, U. S., Wade, W. S., Wang, G. T., Barrett, L. W., Matayoshi, E. D., Huffaker, H. J., Krafft, G. A., and Holzman, T. F. (1994). Amyloid-beta aggregation: selective inhibition of aggregation in mixtures of amyloid with different chain lengths. *Biophysical journal*, 67(3):1216–1228.

Spence, J. (2017). Xfels for structure and dynamics in biology. *IUCrJ*, 4(4):322–339.

Stein, R. C. and Waterfield, M. D. (2000). Pi3-kinase inhibition: a target for drug development? *Molecular medicine today*, 6(9):347–358.

Sun, H. and Scott, D. O. (2010). Structure-based drug metabolism predictions for drug design. *Chemical biology & drug design*, 75(1):3–17.

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of applied crystallography*, 25(4):495–503.

Svergun, D., Barberato, C., and Koch, M. H. (1995). Crysol–a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of applied crystallography*, 28(6):768–773.

Svergun, D. I. (1999). Restoring low resolution structure of biological macro-molecules from solution scattering using simulated annealing. *Biophysical journal*, 76(6):2879–2886.

Svergun, D. I. and Koch, M. H. (2003). Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, 66(10):1735.

Svergun, D. I., Koch, M. H., Timmins, P. A., and May, R. P. (2013). *Small angle X-ray and neutron scattering from solutions of biological macromolecules*, volume 19. Oxford University Press.

Teplow, D. B. (1998). Structural and kinetic features of amyloid β-protein fibrilloge-nesis. *Amyloid*, 5(2):121–142.

Tew, D. J., Bottomley, S. P., Smith, D. P., Ciccotosto, G. D., Babon, J., Hinds, M. G., Masters, C. L., Cappai, R., and Barnham, K. J. (2008). Stabilization of neurotoxic soluble β-sheet-rich conformations of the alzheimer's disease amyloid-β peptide. *Biophysical journal*, 94(7):2752–2766.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, 18(4):267–276.

Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015). Advanced ensemble modelling of flexible macromolecules using x-ray solution scattering. *IUCrJ*, 2(2):207–217.

Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateS-pace, Scotts Valley, CA.

Voevodin, V. (1968). Accuracy of the solution of a system of equations by di-rect methods. *USSR Computational Mathematics and Mathematical Physics*, 8(5):208–211.

Wald, A. and Wolfowitz, J. (1943). An exact test for randomness in the non-parametric case based on serial correlation. *The Annals of Mathematical Statis-tics*, 14(4):378–388.

WHO (2021). Dementia. https://www.who.int/en/news-room/fact-sheets/detail/dementia.

Wolff, M., Zhang-Haagen, B., Decker, C., Barz, B., Schneider, M., Biehl, R., Radulescu, A., Strodel, B., Willbold, D., and Nagel-Steger, L. (2017). Aβ42 pentamers/hexamers are the smallest detectable oligomers in solution. *Scientific reports*, 7(1):1–13.

Zheng, H. and Koo, E. H. (2006). The amyloid precursor protein: beyond amyloid. *Molecular neurodegeneration*, 1(1):1–12.

Contributions

The work presented in this thesis was entirely performed by myself with the following exceptions:

1. The samples for SAXS were prepared by Mohsin Shafiq, Ladan Amin, Behnam Mohammadi, Alexander Hartmann

2. The collection of the SAXS data was performed in collaboration with Stefano Da Vela and assisted by Mohsin Shafiq, Ladan Amin, Behnam Mohammadi, Alexander Hartmann

Acknowledgements

Curriculum vitae

Entfällt aus datenschutzrechtlichen Gründen

Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: ....