Analysis of turbidity progress curves from protein aggregation reactions

A thesis submitted for the degree of Master of Philosophy at The Australian National University

Ran Zhao July 2018

© Copyright by Ran Zhao 2018
All Rights Reserved

Index

Statement of Authorship

I, Ran Zhao confirm that the work presented in this thesis has been performed and interpreted solely by myself except where explicitly identified to the contrary.

My Master's thesis includes three published journal papers:

1. Protein aggregate turbidity: Simulation of turbidity profiles for mixed aggregation reactions. Damien Hall, **Ran Zhao**, Ian Dehlsen, Nathaniel Bloomfield, Steven Williams, Fumio Ariska, Yuji Goto and John Carver. Analytical Biochemistry. 2016, 498: 78-94.

The genesis of this paper was associated with my attempts at analysing a large amount of my own experimental data of protein aggregation reactions conducted using the turbidity procedure. The gradual realization that the turbidity assay was a non-linear descriptor of the underlying extent of aggregated protein led me to a post-facto search for robust methods for de-convoluting the turbidity signal. For this I first needed a formal means for relating aggregation to turbidity. I achieved this by collecting previously developed theoretical treatments of light scattering by protein aggregates and then used empirical interpolation techniques to provide an analytical description in the forms of a unified synthesis. In the publication described above I assisted in preparing the draft and final versions of the paper.

2. Recognizing and analyzing variability in amyloid formation kinetics: Simulation and statistical methods. Damien Hall, **Ran**

Zhao, Masatomo So, Masayuki Adachi, German Rivas, John Carver and Yuji Goto. Analytical Biochemistry. 2016, 510: 56-71.

The inspiration for this second paper came from my attempts to discern differences between my own aggregation assays conducted on a test and control basis. Conducted in parallel with the first study this second investigation started on the premise that there existed a way to relate the experimental assay measure to the underlying aggregation kinetics. Starting with simple methods based on average and standard error I soon realized that such approaches were not valid due to the highly asymmetric/ long-tailed distributions of parameters used to describe protein aggregation kinetics. Here I looked at what improvements could be made in terms of optimal sample size and choice of statistical methods used in the analysis of parameters gained from a conventional data-reduction approach. I assisted in preparing the draft and final versions of the paper.

3. Measurement of Amyloid Formation by Turbidity Assay – Seeing Through the Cloud. **Ran Zhao**, Masatomo So, Hendrik Maat, Nicholas Ray, Fumio Arisaka, Yuji Goto, John Carver and Damien Hall. Biophysics Review, 8: 445-71.

This review paper represented a summary of the work described in papers 1 and 2. The longer review format allowed me to go into a lot more detail regarding the non-linear relationship between the turbidity signal and the underlying chemical kinetics associated with protein aggregation process. In regard to this point I examined subtleties associated with signal produced by the bundling of amyloid fibres, fibre breakage and end-to-end fibre

joining. I wrote the draft and assisted in preparing the final version of the paper.

I confirm that this work is submitted in partial fulfilment for the degree of Master of Science in Biochemistry and has not been submitted elsewhere in any other form for the fulfilment of any other degree or qualification.

Word count of “Motivation” and “Future Studies”: 3,036.

Signature:

Dated: 09/07/2018

Acknowledgements

I would like to thank my two supervisors – Dr Damien Hall and Professor John Carver at the Australian National University. There would not be a thesis without them.

I would also like to thank the following professional scientists for advice along the way: Dr Joanna Woodcock at the Centre for Cancer Biology (Adelaide), Professor Grant Booker, the University of Adelaide and Professor Gottfried Otting and Dr Ruhu Qi of the Australian National University. Thanks also for numerous pieces of advice and kind words from fellow students and researchers at the University of Adelaide (David, Yanqin, Katy, Danielle, Manjeet, Nick, Bless, Matthew, Al, Anita, Dawei, Weiwei, Yun, Anton and Steven) and also those at the Australian National University (Daniel, Yao, Anitha, Chris, Ian, Nathaniel, Aidan and Elmira). Their kind words and actions helped me to (eventually) finish my research.

I would also like to express my sincere thanks to the other authors of the publications included in this thesis.

I would also like to thank the China Scholarship Council, the University of Adelaide and the Australian National University for the combined scholarships they provided.

Finally, I must thank my parents and my friends for their unconditional support. This accomplishment would not have been possible without you.

Author: Ran Zhao

Abbreviations

ADH: alcohol dehydrogenase

NMR: nuclear magnetic resonance

RCM: reduced and carboxy-methylated

sHsps: small heat shock proteins

TEM: transmission electron microscopy

WT: wild-type

Motivation

1. My research history

Having obtained an undergraduate and a Masters degree from Tsinghua University in China I moved to Australia to further my postgraduate study at the University of Adelaide at the beginning of 2010. My initial research project was concerned with the association of the 14-3-3 zeta protein with small heat shock proteins (sHsps) and their potential synergistic effects in promoting molecular chaperone ability. The aim of the project was to find out whether 14-3-3 proteins associate with sHsps, especially phosphorylated sHsp forms, and if they do, whether this is a synergistic interaction that can enhance the combined chaperone ability of the two proteins.

To carry out these aims I used recombinant DNA technology to produce wild-type 14-3-3zeta, wild-type alpha B-crystallin, S59D alpha B-crystallin and S19,45,59D alpha B-crystallin from a plasmid transformed into E. coli using previously published methods (Williams *et al*., 2011; Horwitz *et al.*, 1998). I successfully expressed and purified these products as demonstrated by the positive ion mass spectra obtained for each of the four proteins (Figure 1). To investigate the chaperone abilities of these proteins I used ADH, insulin and alpha-lactalbumin as substrates in a turbidity based light scattering assay (Figure 2). In addition to the light scattering turbidity assays I also performed Thioflavin T dye binding fluorescence assays to examine the amyloid aggregation reaction (Ecroyd *et al.*, 2007). I used circular dichroism spectropolarimetry (CD) to investigate the conformational change of the proteins upon aggregation (Williams *et al*., 2011). I prepared N15 labeled WT 14-3-3zeta for

multidimensional NMR research (Treweek *et al.,* 2010) and used TEM imaging to observe the formation of amyloid fibrils and amorphous aggregates (Ecroyd and Carver, 2009).

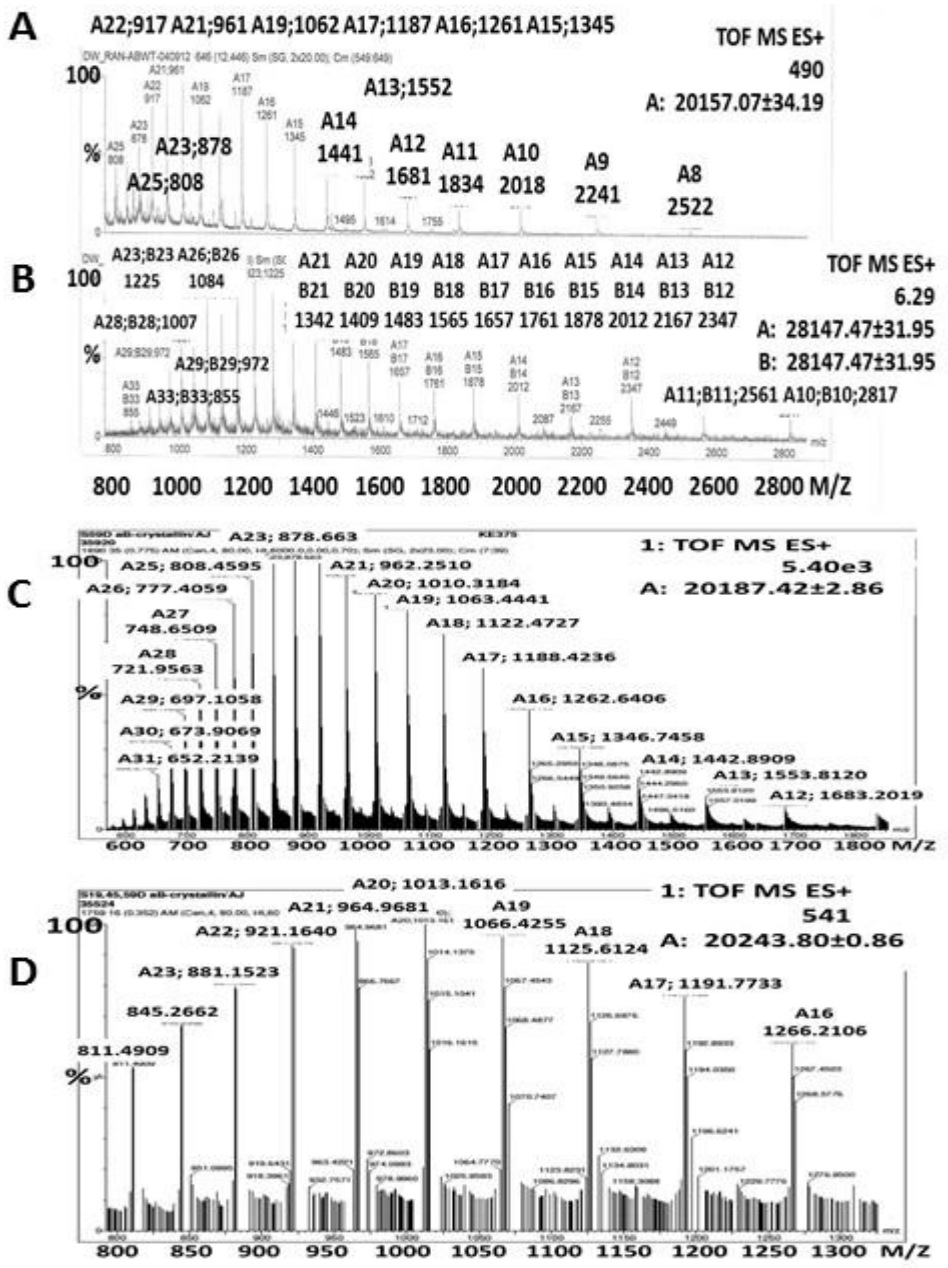


Figure 1. (A) WT alpha B-crystallin mass spectrometry spectrum. Measured mass 20157Da vs theoretical mass 20159Da. (B) WT 14-3-3zeta (with 5 amino acid residues from TEV protease) positive ion electrospray ionization mass spectrometry spectrum. Measured mass 28148Da vs theoretical mass 28174Da. (C) S59D alpha B-crystallin mass spectrometry spectrum. Measured mass 20187Da vs theoretical mass 20187Da. (D) S19, 45, 59D alpha B-crystallin mass spectrometry spectrum. Measured 20243Da vs theoretical mass 20243Da.

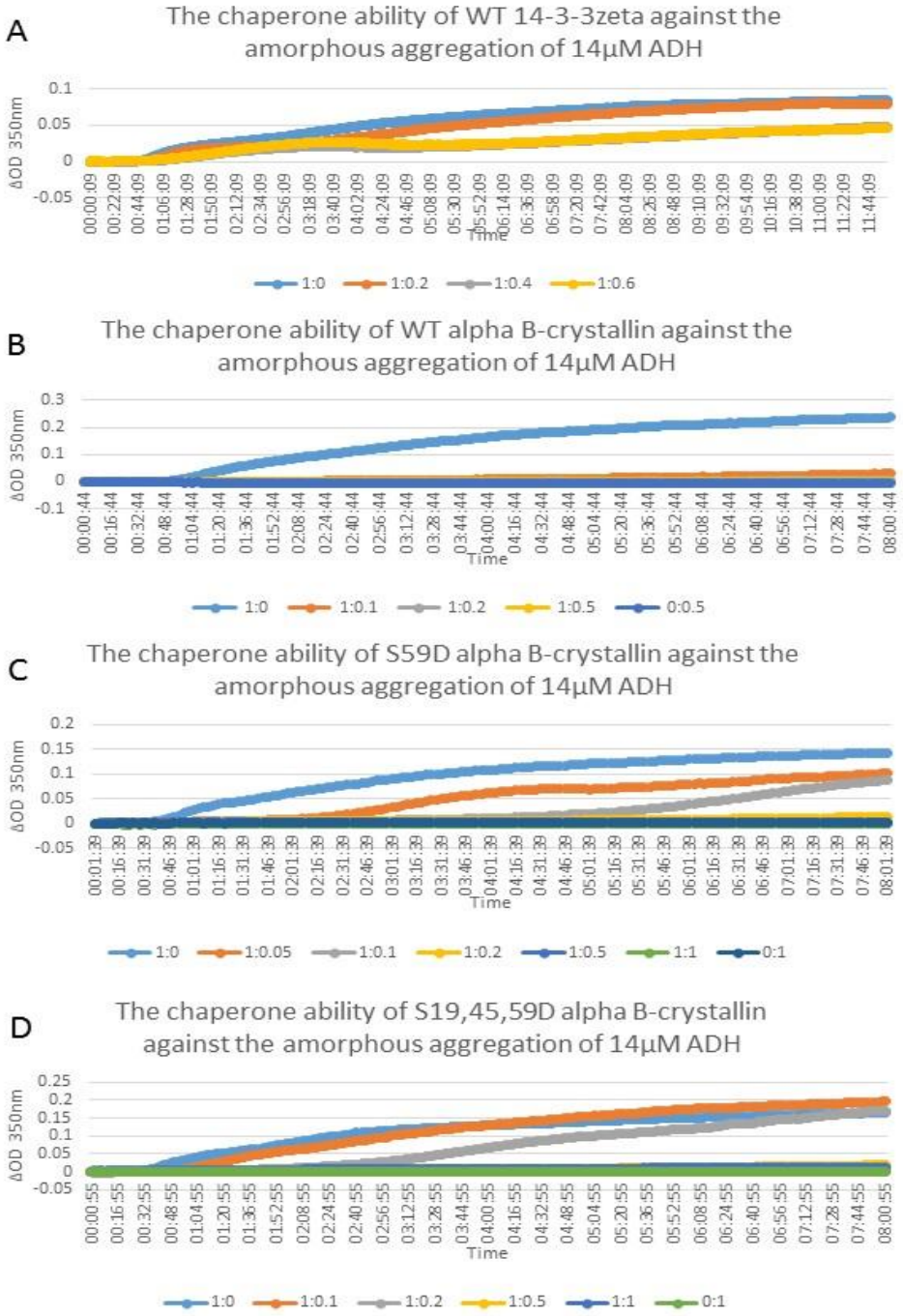


Figure 2. Turbidity assay based investigation of (A) Chaperone ability of WT 14-3-3zeta against the amorphous aggregation of 14µM ADH at different ADH/14-3-3 molar ratios. (B) Chaperone ability of WT alpha B-crystallin against the amorphous aggregation of 14µM ADH at different ADH/crystallin molar ratios. (C) Chaperone ability of S59D alpha B-crystallin against the amorphous aggregation of 14µM ADH at different ADH/crystallin molar ratios. (D) Chaperone ability of S19,45,59D alpha B-crystallin against the amorphous aggregation of 14µM ADH at different ADH/crystallin molar ratios.

Upon Professor Carver's acquisition of a new job at the Australian National University, in July of 2013 I transferred with him to continue my postgraduate research program. Soon after this time (February 2014) I have mostly worked under the supervision of Dr. Damien Hall, an academic who had just joined the ANU Research School of Chemistry. Under his supervision I examined some of the fundamental aspects of the turbidity aggregation assay.

With regards to this assay technique, solution turbidity can be taken as reflecting the total extinction of light passing through solution by non-absorptive, scattering based mechanisms. In general, throughout the protein aggregation/chaperone literature, turbidity is commonly considered as a linear marker of the extent of the protein aggregation reaction, however I noted that this assumption is frequently made without any experimental or theoretical support. In my Masters course I sought to examine the basic veracity of this assumption in much greater detail. My research involved the following general sequence,

(i) Investigating solution conditions that cause proteins to unfold and aggregate.
(ii) Monitoring the protein aggregation reaction by the turbidimetric method.
(iii) Examining the differential effects of chaperones on the protein aggregation reaction as monitored by the recording of turbidity.

As these three steps constituted the core of my scientific approach I briefly introduce the pertinent aspects of each in the following sections.

2. Protein aggregation and protein conformational diseases

2.1 Protein folding and unfolding

All proteins are amino acid polymers. Amino acid molecules are linked together by peptide bonds to form polypeptides. After peptide synthesis, most proteins need to correctly fold from a randomly structured polypeptide coil into a well-defined three-dimensional globular structure to function (Alberts *et al.*, 2002). The protein folding process is generally reversible[1]. In some circumstances, such as mutation or the presence of environmental stresses, natively folded proteins may unfold from their functional globular formation. This process is called protein unfolding or protein denaturation. Denatured proteins can often fold back to their native states when the environmental stress reduces or vanishes.

2.2 Protein aggregation and protein conformational diseases

Sometimes unfolded polypeptides cannot fold into their native three-dimensional structure due to errors in the folding process (Figure 3). For example, they could enter the off-folding pathways via the association of intermediately folded molten globules (Ecroyd and Carver, 2008). Misfolded proteins often form insoluble amorphous or amyloid fibril aggregates (Ecroyd and Carver, 2008). Amyloid fibrils are characterised by highly ordered, cross-beta sheet fibrillar structures that are from nanometre to micrometer in length (Treweek *et al.*, 2003).

---

[1] Reversible within the context of the IEME approximation i.e. the instantaneous establishment of microscopic equilibrium. The IEME approximation defines equilibrium time periods as an interval sufficiently long for satisfying ergodicity requirements of the particular parameter being measured [Hlavacek *et al.*, 1999].

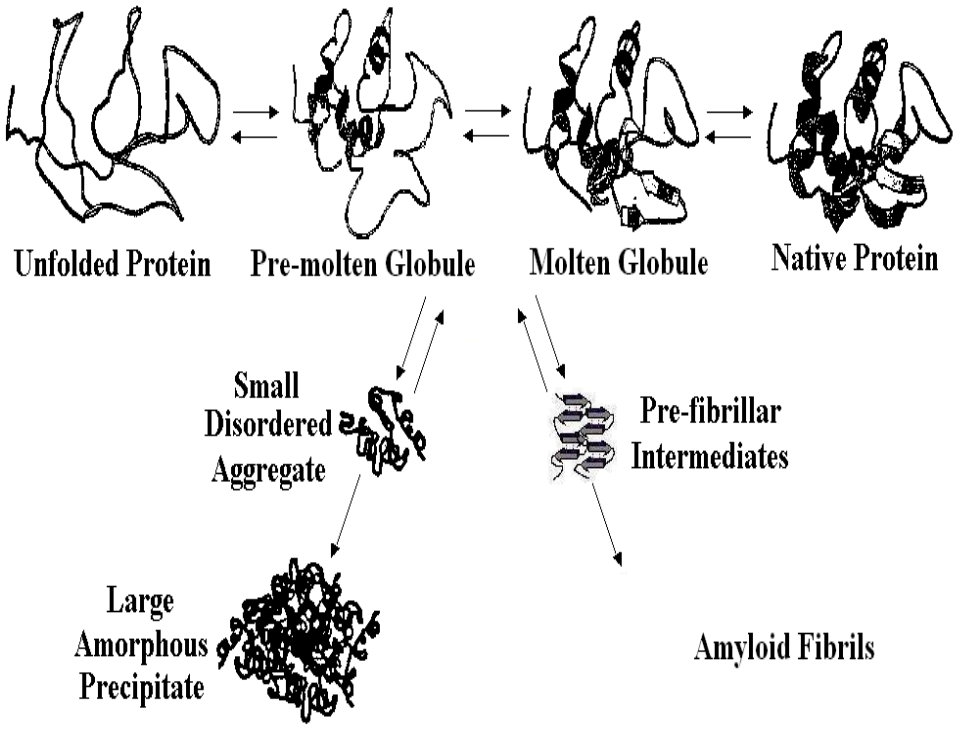


Figure 3: Basic pathways of protein aggregation. In the first pathway, proteins form small disordered aggregates, developing into large amorphous precipitate. In the second pathway, the protein will form pre-fibrillar intermediates, and then amyloid fibrils (taken from Treweek *et al.*, 2003).

Many diseases, especially neurodegenerative diseases, are caused by misfolded proteins. These type of diseases are called protein conformational diseases or proteopathies (Ecroyd and Carver, 2008). Misfolded proteins are usually degraded by the ubiquitin proteasome system and lose their normal functions, hence cause protein conformational diseases. Toxic species can form during protein misfolding and aggregation, and disrupt the functions of cells, tissues and organs, causing protein conformational diseases (Cohen and Kelly, 2003). The risk of protein conformational diseases increases with age (Walker and LeVine, 2000). It is partly because the ability of cells to degrade or refold abnormal proteins becomes increasingly poorer as the cells grow older (Walker and LeVine, 2000). For example, prions

arise from a type of misfolded protein able to form amyloid fibrils which then cause Creutzfeldt-Jakob disease (Jarrett and Lansbury, 1993). Alzheimer's and Parkinson's disease are also characterized by the formation of plaques/deposits in the brain that contain amyloid fibrils (Huang and Mucke, 2012; Jankovic, 2008). Protein misfolding and aggregation are also related with cataract, some types of cancers, some types of heart diseases and multiple sclerosis (Kelly, 1998; Dobson, 2001; Stefani, 2004; Hall and Edskes, 2009; Hall and Edskes, 2012).

3. Molecular Chaperones and their role in preventing aggregation

3.1 Molecular Chaperones

Molecular chaperones are proteins which can assist with other proteins' folding. They are able to identify protein molecules which are not correctly folded, and interact with them through different mechanisms. Many proteins often cannot fold into their native states by themselves and require the assistance of molecular chaperones to correctly fold and function (Chaudhuri and Paul, 2006). Molecular chaperones are considered to be the first protective mechanism of cells against various stresses, such as elevated temperature, pH shift, oxidative stress and possibly changes in the levels of macromolecular crowding (Hall, 2002; Hall and Dobson, 2006). In the following sections I provide a short summary of the chaperone proteins that formed the basis of the experimental systems studied during my postgraduate research.

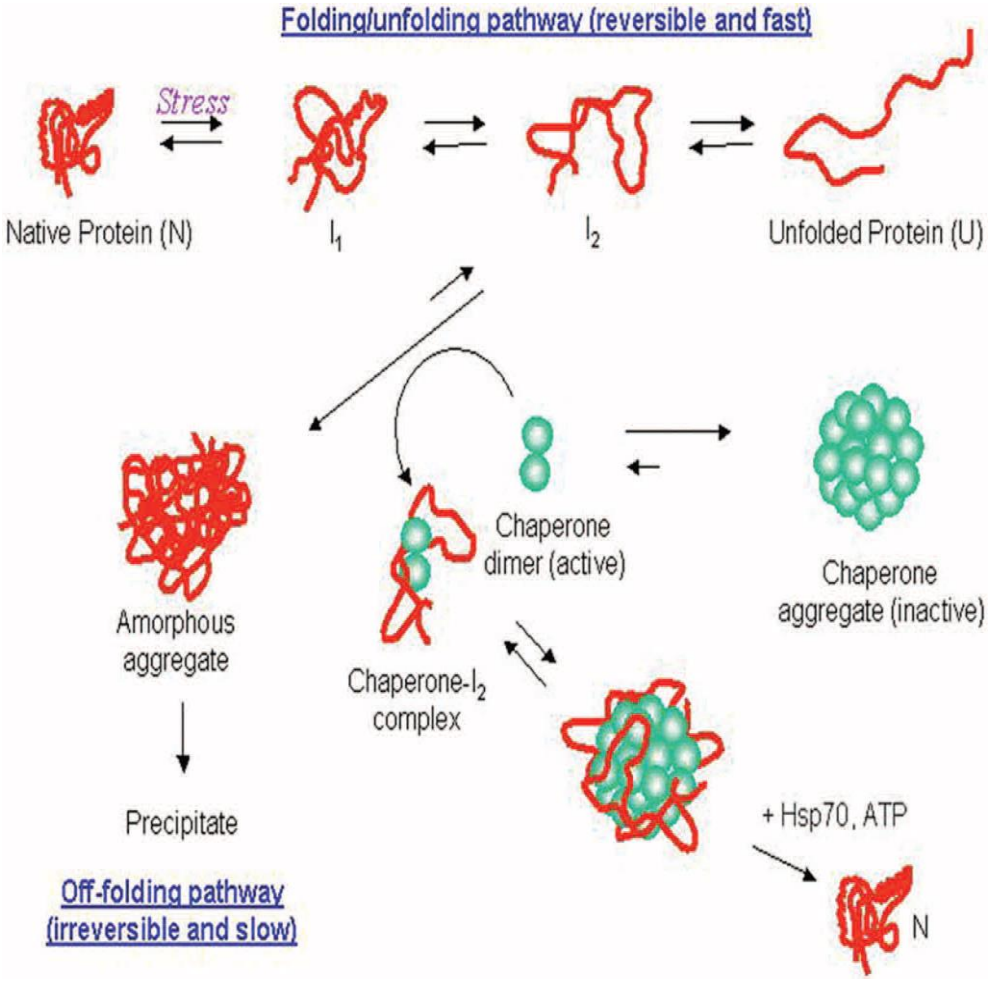


Figure 4: A schematic mechanism of chaperone action of sHsps, which selectively interact with target proteins on their off-folding pathways to prevent protein aggregation (figure taken from Carver *et al.*, 2003).

3.2 Alpha B-crystallin

The crystallins are a protein family which are most prominent in the eye lens and cornea. The main function of crystallins is maintaining the transparency of said organs by inducing a well-ordered 'crystal-like' arrangement (Andley, 2007). Alpha crystallin is the predominant eye lens protein. Aside from their structural function, alpha crystallin also functions as a molecular chaperone, preventing protein aggregation in the eye lens.

Alpha crystallin is composed of two basic subunit types - alpha A and alpha B (Ecroyd and Carver, 2009). Alpha B-crystallin is also expressed in abundance in other tissues besides the eye lens,

such as brain, lung and muscle (Kato *et al*., 1991). Both alpha A and alpha B-crystallin contain flexible C-terminal extensions (Carver *et al.*, 1992) and are members of the small heat-shock protein (sHsp) family of molecular chaperones (Figure 4, Carver *et al.* 2003).

3.3 14-3-3 proteins

The 14-3-3 protein family is highly conserved and expresses in abundance in all eukaryotic cells, and also it is one of the most abundant cell proteins (Moore and Perez, 1967; Martin *et al.*, 1994). They are able to bind to and interact with more than 200 different proteins, suggesting their possible importance as a hub interacting protein or chaperone protein (Sun *et al*., 2009). Muslin *et al*. first discovered 14-3-3 proteins' phosphoserine-binding ability (Muslin *et al*., 1995). Most 14-3-3 protein binding partners are phosphorylated proteins that are capable of binding to its amphipathic central groove (Fu *et al.,* 2000).

The 14-3-3 proteins also exhibit a small heat shock protein like chaperone ability, by inhibiting the thermally and chemically induced aggregation of several target proteins in vitro (Yano *et al*., 2006; Williams *et al*., 2011). However, this chaperone ability has still not been extensively investigated.

4. Protein aggregation turbidity assay

To investigate the individual and combined effects of the molecular chaperone abilities of both 14-3-3zeta and alpha B-crystallin on the aggregation of proteins a vital first step is the development of a robust and simple assay for examining the aggregation of proteins. By far the most common in vitro method to monitor protein aggregation is the turbidity assay (Dolado *et al*.,

2005; Sant'Anna *et al*., 2016) (Figure 5).

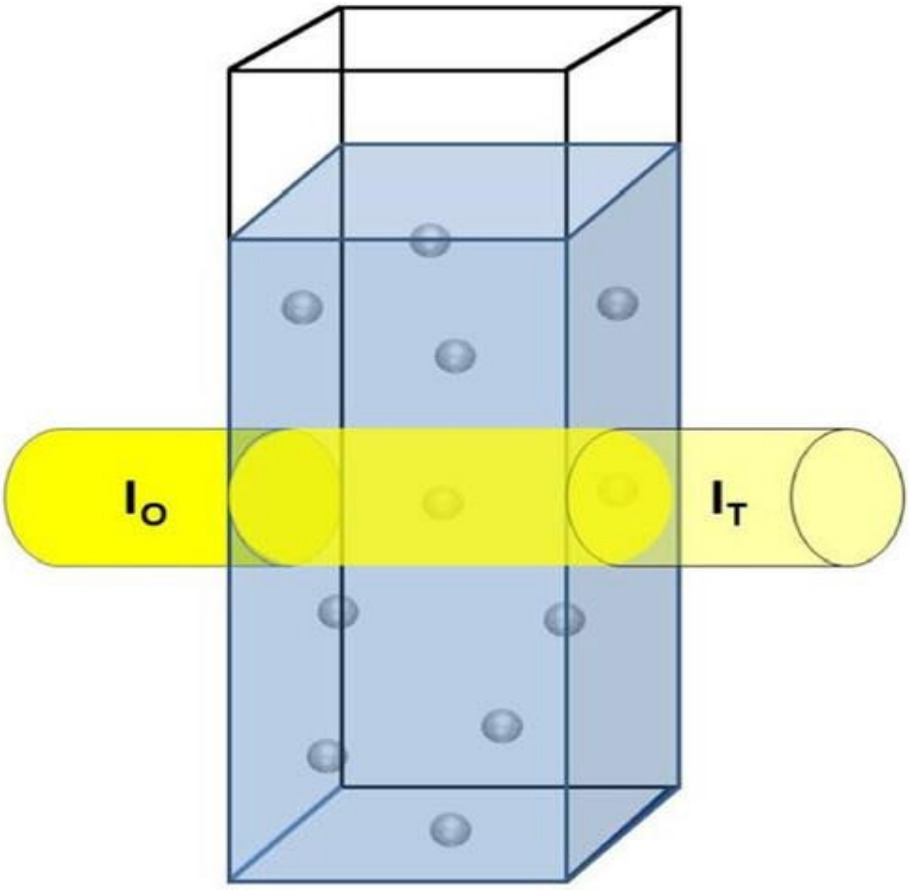


Figure 5: Schematic describing the transmission-based measurement of excess solution turbidity of protein aggregates in which the transmitted light intensity (IT) is measured in relation to the incident light intensity (I0) using a standard spectrophotometer or plate reader (Zhao *et al.,* 2016).

All protein aggregates scatter light in the visible wavelength region since their size ranges from nanometer to micrometers (Bohren and Huffman, 2008; Doty and Steiner, 1950; Invernizzi *et al*., 2012; Stoppini and Bellotti, 2015). This characteristic combined with a lack of absorption in the visible wavelength region makes the low-cost turbidity assay a particularly attractive method for monitoring protein aggregation (Andreu and Timasheff, 1986; Mahler *et al*., 2009). Colloidal solution turbidity is generally considered to exhibit a linear relationship with the aggregation

reaction. However, this assumption is usually not based on convincing supporting experiments or theory. The turbidity of a colloidal solution is not only determined by the size, but also the shape of the particles. As a result, analyzing the relationship between solution turbidity and protein aggregation can be quite challenging (Camerini-Otero *et al*., 1978; Garcia-Lopez and Garcia-Rubio, 2008; Garcia-Lopez *et al*., 2006; Hall and Minton, 2005; Korolevskaya and Khlebtsov, 2010; Mahler *et al*., 2009; Moody *et al*., 1996; Silver and Birk, 1983). In my postgraduate research I examined and developed improved methods for simulating and analyzing turbidity profiles of mixed protein aggregation reactions (Hall *et al.*, 2016a; Hall *et al.*, 2016b; Zhao *et al*., 2016), which will greatly facilitate the understanding of protein aggregation and the effect of molecular chaperone reactions.

5. Questions relating to turbidity assay and its quantitative interpretation

5.1 Simulating turbidity of protein aggregates of different sizes

In my first paper (Hall *et al.*, 2016a), I contributed to developing a hybrid method for simulating turbidity of protein aggregates of different sizes in the low concentration limit. This simulation utilises a combination of the Rayleigh, the Rayleigh-Gans-Debye (RGD) and approximate forms of the Mie scattering equations (Bohren and Huffman, 2008; Elimlech *et al.*, 1995; Garcia-Lopez, 2005; Mezzenga and Fischer, 2013; Wickner *et al.*, 2007). This hybrid approach was used to generate empirical interpolating functions, which may be used for both simulation and analysis of turbidity profiles.

5.2 Quantifying the variability in the amyloid aggregation assay

In my second paper (Hall *et al.*, 2016b), I helped to develop a method for quantifying the variability in the amyloid aggregation assay. We investigated the variability in the amyloid aggregation kinetics, and developed methods for its simulation, identification and analysis. Rather unexpectedly, such an analysis had not been previously developed despite it being the fundamental cornerstone of all differential analyses of drug and condition effects upon the protein aggregation reaction.

5.3 Physical chemistry of the turbidity signal

In my third paper (Zhao *et al.,* 2016), I reviewed the physical chemistry of the turbidimetric assay methodology, investigating the reviewed information with a series of pedagogical kinetic simulations. We particularly focused upon recent literature relating to ultra-microscope image analysis (Hall, 2012; Usov and Messenga, 2015), light scattering and turbidity development by protein aggregates (Garcia-Lopez *et al.,* 2006, 2008; Hall *et al.,* 2016a), and computer simulation of the kinetics of amyloid and other aggregate types (Adachi *et al.,* 2015; Hall *et al.,* 2015; Hall *et al.,* 2016a).

Paper 1: Protein aggregate turbidity: Simulation of turbidity profiles for mixed aggregation reactions

The first of the three papers published during my studies at the Australian National University Research School of Chemistry (Hall et al. 2016a) looked at the turbidity signal – which was the basis of my assay procedure for looking at the effect of chaperone action. In this paper I re-examined the already published theoretical relations for describing the turbidity of protein aggregates. In this paper I contributed to developing a hybrid method for simulating turbidity of protein aggregates of different sizes in the low concentration limit. This simulation utilizes a combination of the Rayleigh, the Rayleigh-Gans-Debye (RGD) and approximate forms of the Mie scattering equations (Bohren and Huffman, 2008; Elimlech *et al.*, 1995; Garcia-Lopez, 2005; Mezzenga and Fischer, 2013; Wickner *et al.*, 2007). This hybrid approach was used to generate empirical interpolating functions, which may be used for both simulation and analysis of turbidity profiles.

Analytical Biochemistry 498 (2016) 78–94

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Protein aggregate turbidity: Simulation of turbidity profiles for mixed-aggregation reactions

Damien Hall [a, b, *], Ran Zhao [a], Ian Dehlsen [a], Nathaniel Bloomfield [a], Steven R. Williams [a], Fumio Arisaka [c, d], Yuji Goto [b], John A. Carver [a]

[a] Research School of Chemistry, Australian National University, Acton, ACT 2601, Australia
[b] Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan
[c] Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 226-8501, Japan
[d] Advanced Materials Science Center, College of Bioresource Sciences, Nihon University, Chiyoda-ku, Tokyo 102-8275, Japan

ARTICLE INFO

Article history:
Received 20 September 2015
Received in revised form
27 November 2015
Accepted 30 November 2015
Available online 4 January 2016

Keywords:
Protein aggregation
Turbidity
Mixed aggregation
Light scattering

ABSTRACT

Due to their colloidal nature, all protein aggregates scatter light in the visible wavelength region when formed in aqueous solution. This phenomenon makes solution turbidity, a quantity proportional to the relative loss in forward intensity of scattered light, a convenient method for monitoring protein aggregation in biochemical assays. Although turbidity is often taken to be a linear descriptor of the progress of aggregation reactions, this assumption is usually made without performing the necessary checks to provide it with a firm underlying basis. In this article, we outline utilitarian methods for simulating the turbidity generated by homogeneous and mixed-protein aggregation reactions containing fibrous, amorphous, and crystalline structures. The approach is based on a combination of Rayleigh–Gans–Debye theory and approximate forms of the Mie scattering equations.

Crown Copyright © 2015 Published by Elsevier Inc. All rights reserved.

Protein aggregation is a complex phenomenon of major importance to the fields of medicine [1–3], biology [4,5], and industry [6,7]. Because protein aggregates span the nanometer to micrometer size range, they tend to scatter an appreciable amount of light in the visible region of the spectrum[1] [8] (see also chaps. 1 and 5 in Ref. [9]). Their scattering potential, coupled with their general lack of absorption over this region, has meant that visible wavelength solution turbidity has been a cornerstone method for monitoring protein aggregation reaction kinetics [10,11]. However, because the extent of light scattered by an aggregate is a function of both the size and shape of the particle, the unambiguous analysis of the turbidity signal for comparative studies of protein aggregation can be challenging [11–18]. In the following series of articles, we describe methods for simulating and analyzing turbidity profiles of mixed protein aggregation reactions. In this first article of the series, we develop robust methods for simulating the turbidity generated by different classes of protein aggregates in the low concentration limit.[2] The simulation approach makes use of a combination of Rayleigh, Rayleigh–Gans–Debye (RGD), and approximate forms of the Mie scattering equations [19,20] (see also chaps. 4–6 in Ref. [9]) as a basis for generating empirical interpolating functions that can be used in a straightforward manner for both simulation and analysis of turbidity profiles.

Utilitarian approaches to simulating turbidity

Although the topic of aggregate turbidity has been explored previously [10–18], the range of theoretical approaches used, prior to the current one, have been either restricted to a single size regime (e.g., the Rayleigh–Gans limit [9–13]) or lacking of a format that can be easily adopted by those wanting to translate their experimentally accessible parameters into useful theory-based

Abbreviations used: RGD, Rayleigh–Gans–Debye; AD, anomalous diffraction.

* Corresponding author. Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan.
E-mail addresses: damien.hall@anu.edu.au, damienhall30@gmail.com (D. Hall).

[1] When formed in a medium having a substantially different refractive index from them.

[2] By low concentration limit, we mean free of significant external interference effects that allow the system to comport to the single incoherent scattering criterion described in chap. 3 of Ref. [9].

http://dx.doi.org/10.1016/j.ab.2015.11.021
0003-2697/Crown Copyright © 2015 Published by Elsevier Inc. All rights reserved.

simulations [14–18]. In regard to this point, the major obstacles have been the following:

1. Lack of a nomenclature for reducing protein aggregates to a limited number of structural models that correctly relate mass, refractive index, and shape.
2. Lack of easily calculable turbidity models that are applicable over a range of aggregate sizes and shapes.

Our primary goal in this article was to develop utilitarian methods for simulating turbidity profiles of mixed-protein aggregation reactions. By *utilitarian*, we mean analytical expressions having relatively few independent variables that can be used for quantitative/semi-quantitative prediction and analysis of experimental results.

Mesoscopic models of protein aggregates

Depending on the nature of the aggregation pathway, a number of different classes of protein aggregate structures may be formed [6,21–23] (Fig. 1). Amorphous aggregates [6,21–23] for which the internal arrangement of monomers exhibits no long-range structural ordering [21,24] may be produced. Alternatively, semi-crystalline structures, such as one-dimensional fibers [6,21–23] and three-dimensional crystals [6,25], that do display significant long-range ordering within the aggregate configuration (cf. Refs. [24] and [26]) may be produced. In this section, we describe a general system for categorizing the range of possible protein aggregate structures into a limited subset, with the intention of using this subset as the basis of quantitative simulation and characterization procedures. Toward this goal, a protein aggregate

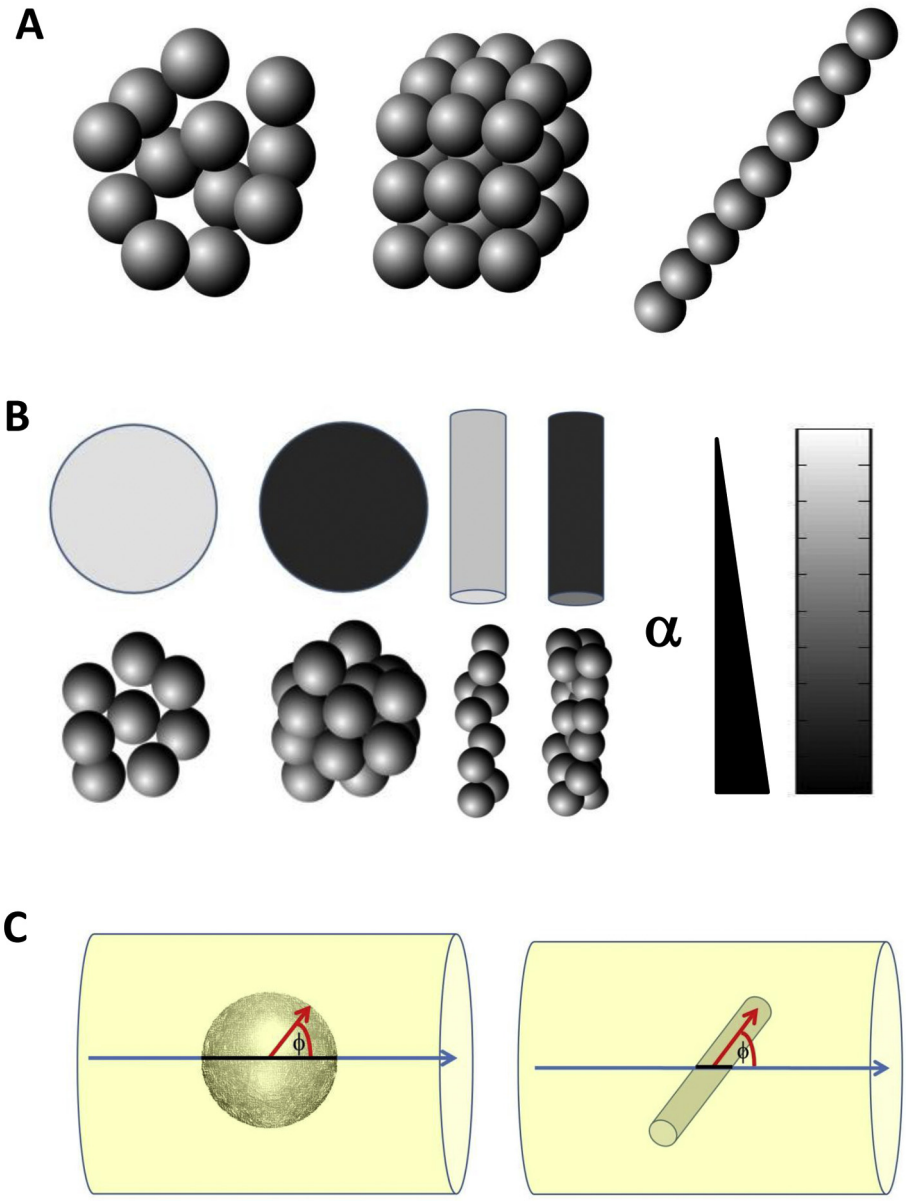


Fig.1. (A) Various types of protein aggregate—amorphous, crystalline, and fibrous. (B) Schematic of mesoscopic structural approximations for rod-like and spherical-shaped aggregates in the scattering experiment. Aggregate packing density is defined by the parameter α. (C) Scattering cross-sectional area as a function of angle of presentation, ϕ, for rods and spheres.

composed of i monomers may be usefully defined in terms of three experimentally realizable properties: the molecular weight, M_i, the shape, S_i, and the volume trace, $(V_i)_{TRACE}$ [27,28]. Quantitative descriptions of these three parameters are given in Eq. (1). Two limiting approximations of the protein aggregate shape [28,29] are considered: an arbitrarily diffuse rod, defined by a trace length L_i and a trace radius R_i, and an arbitrarily diffuse sphere, singly characterized by R_i (Eq. (1b) and Fig. 1):

$$M_i = iM_1 \tag{1a}$$

$$(V_i)_{TRACE} = \frac{iM_1 \upsilon}{\alpha_i N_A} \tag{1b}$$

$$S_i = \begin{cases} \text{rod; } L_i = (V_i)_{TRACE} \Big/ \left(\pi R_i^2\right) \\ \text{sphere; } R_i = \sqrt[3]{3(V_i)_{TRACE}/(4\pi)} \end{cases}. \tag{1c}$$

In Eq. (1b), N_A is Avogadro's number and $(V_i)_{TRACE}$ refers to the volume of the aggregate approximated by tracing a large body over the surface to produce an approximately regular shape [27,28]. The term α_i describes the effective packing density of the trace volume (i.e., the fraction of $(V_i)_{TRACE}$ actually occupied by protein), and υ refers to the protein partial specific volume that, to a first approximation, is taken as having a constant value of approximately 0.73 ml/g [30]. Defining α in the manner outlined by Eq. (1b) allows it to be used to parameterize the transition between compact and diffuse states ($\alpha_{DIFFUSE} < \alpha_{COMPACT} \leq 1$), with the intention that a higher value of α would be more appropriate for crystal-like aggregates, whereas a lower value would describe a less dense amorphous aggregate [6,21–23,31].

As particles get larger and denser, the physics of scattering becomes less concerned with particle volume and more concerned with the relative orientation of the particle, ϕ_i, and the average area that it presents to the incident light, $\langle A_i \rangle$ [9,20]. For the case of a sphere of uniform composition, the scattering relationship is isotropic, thereby making the average physical cross-sectional area equal to that defined by its trace volume radius (Eqs. (2a) and (2b)). For the case of a rod, an estimate of

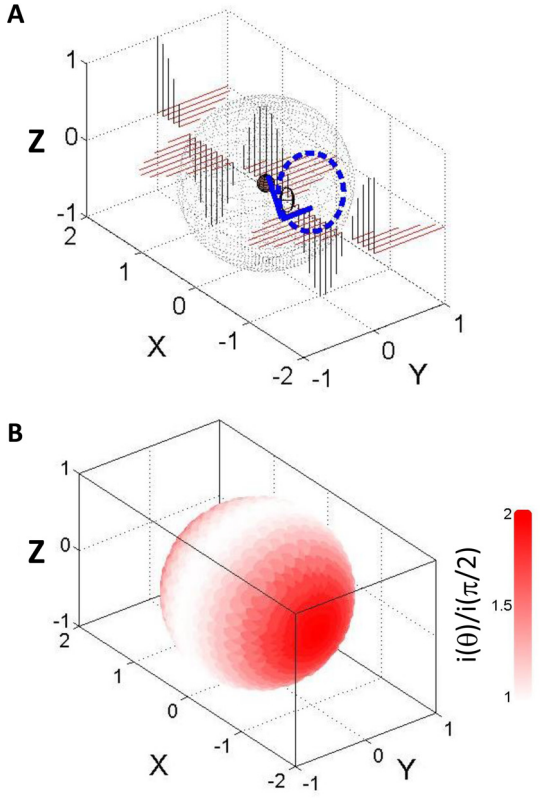


Fig.2. Schematic of (A) general scattering experiment showing the interaction of non-polarized light with a molecule located at the origin with an outer sphere (not to scale) representing all possible points of detection of the scattered light at a set radial distance and solid angle. (B) The intensity envelope surrounding the scattering molecule showing the scattering intensity as a function of the sub-apex of the spherical solid angle.

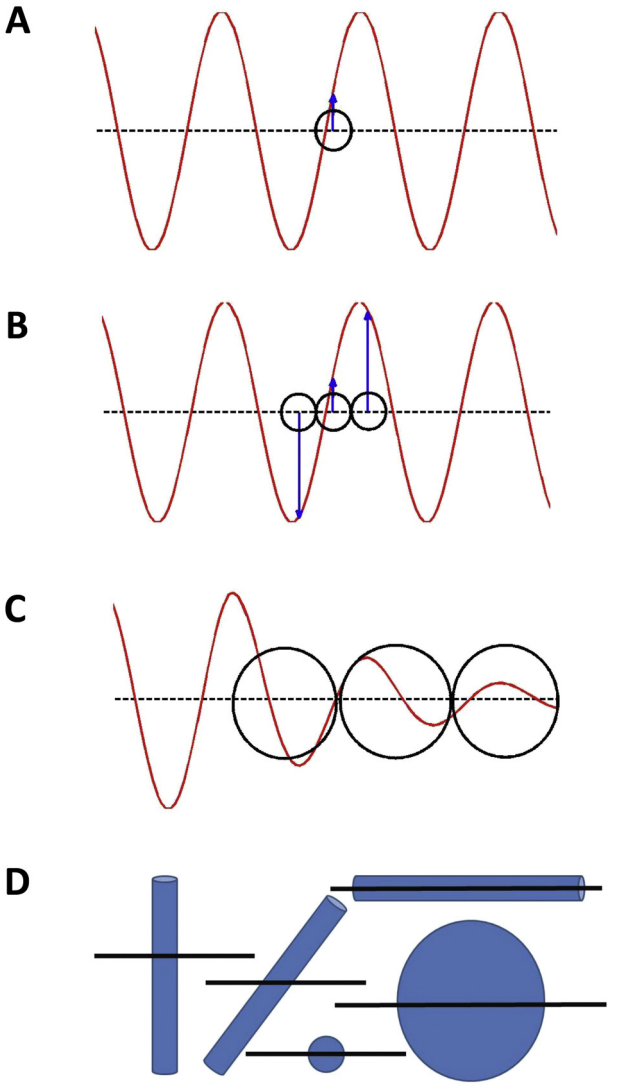


Fig.3. Schematic of the three general scattering regimes considered in this work. (A) Rayleigh scattering, where the scattering particle is small in relation to the wavelength of light (i.e., $\langle R_i \rangle < \lambda/20$) (blue arrow represents the magnitude of the electric field). (B) Rayleigh–Gans–Debye scattering, where the particle can be relatively large in relation to the wavelength of light sufficient to experience out of phase scattering effects but not so large as to significantly decrease the intensity of incident light (i.e., $\lambda/20 < \langle R_i \rangle < \lambda/2$) (blue arrow represents the magnitude of the electric field). (C) Mie scattering, where the particle is sufficiently large to experience out of phase scattering and to diminish the intensity of incident light moving through the particle. (D) Schematic describing orientation effects on the intra-particle path length of the incident light in the Mie scattering regime for a variety of shapes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

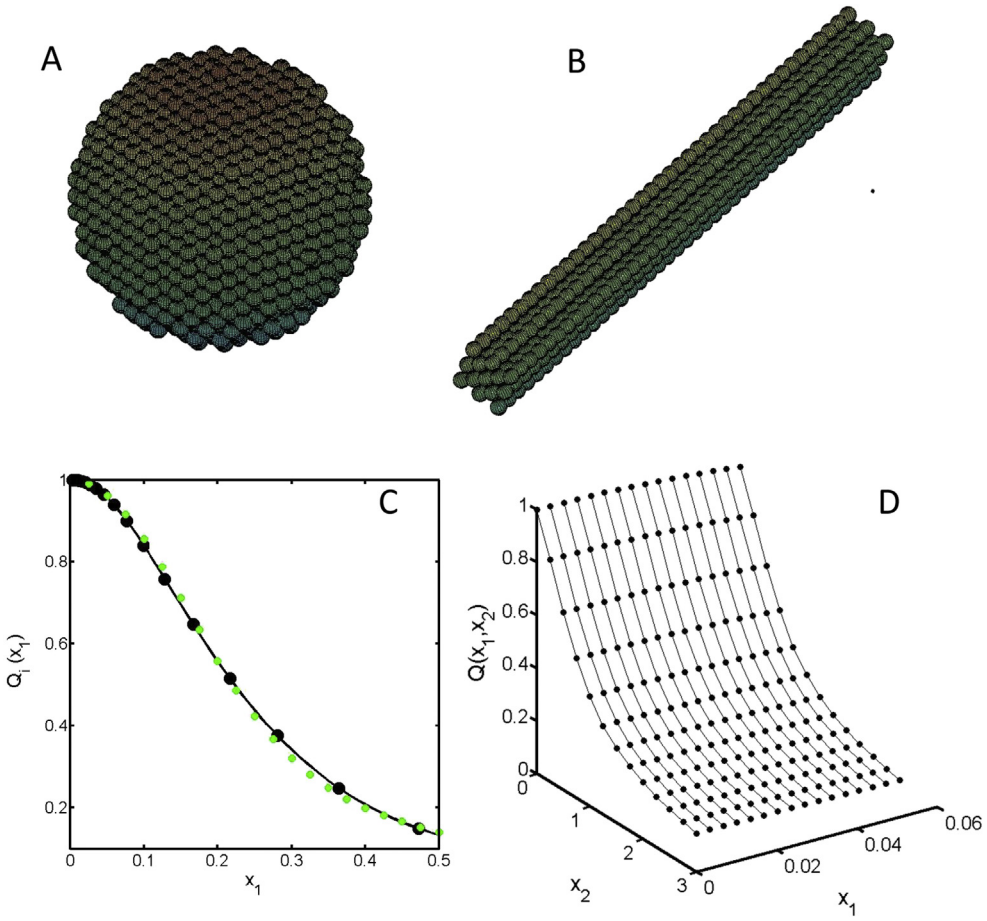


Fig.4. Schematic of the RGD approximation for (A) spheres and (B) rods of arbitrary size. (C) Results of RGD calculations (black symbols, ●) for different sizes of spheres scaled in relation to the wavelength of light and the solvent refractive index ($x_1 = R_i\, n/\lambda$). Solid line represents best fit to empirical equation defined by Eq. (11a), and green symbols (●) represent overlaid data of Camerini-Otero and Day [12]. (D) RGD output for different lengths and widths of rods relative to solution refractive index and the wavelength of light ($x_2 = L_i\, n/\lambda$). Lines represent best fits to empirical expression shown as Eq. (11b). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$\langle A_i \rangle$ in terms of L_i and R_i can be obtained by first performing spherical integration to find the area of projection (Eq. (2b)) (e.g., Ref. [32]):

$$A_i(\phi_i) = \pi R_i^2 \;\; \therefore \langle A_i \rangle = \pi R_i^2 \quad \text{for a sphere} \tag{2a}$$

$$A_i(\phi_i) = \pi R_i^2 \cos\phi_i + 2L_i R_i \sin\phi_i \quad \text{for a rod.}$$
$$\therefore \langle A_i \rangle \approx 2R_i^2 + 2/\pi L_i R_i \tag{2b}$$

With coarse-grained models and formalism suitable for describing both the physical dimensions and the area of projection for rod-like fibers, crystals, and amorphous aggregates (conforming to spherical shape), we now introduce methods for calculating the wavelength-specific turbidity.

Light scattering and turbidity transformations

The wave model pictures light as a pair of coincident, yet perpendicular, traveling electric and magnetic vectors oscillating at a frequency, f, over a wavelength, λ [9,33]. Each time a light wave train encounters some element of matter, the electric field vector displaces the normal charge distribution, causing it to oscillate at the characteristic frequency of the light. Conservation laws require that this accelerated charge re-emit energy in the form of another light wave of the same wavelength, producing the phenomenon known as elastic light scattering[3] [9,33]. Lord Rayleigh quantified the relationship between the measured intensity of the scattering (at a point $p = [r,\theta]$) from a single particle (located at the origin) for the particular case[4] where the particle is small in relation to the wavelength of light (Eq. (3a)) (Fig. 2) [34]. Under these Rayleigh-limiting conditions, the total amount of light scattered away from the forward direction, I_S, can be calculated by spherical integration of the angle-specific scattering intensity, $i(r,\theta)$ (Eq. (3b)):

$$i(r,\theta) = \frac{i_0}{r^2}\left[\frac{9\pi^2 (V_i)^2_{TRACE}}{2\lambda^4}\left(\frac{m_i^2 - 1}{m_i^2 + 2}\right)^2 \left(1 + \cos^2\theta\right)\right] \tag{3a}$$

$$I_S = 2\pi r^2 \int_0^{\pi} i(r,\theta)\sin(\theta)\, d\theta \tag{3b}$$

[3] Elastic light scattering refers to the case where the scattered and incident radiation has the same wavelength. Inelastic light scattering refers to the case where some energy is absorbed, donated, or dissipated by the molecule with subsequent change in wavelength between incident and scattered radiation.

[4] Where θ represents the forward scattering apical sub-angle of the solid angle.

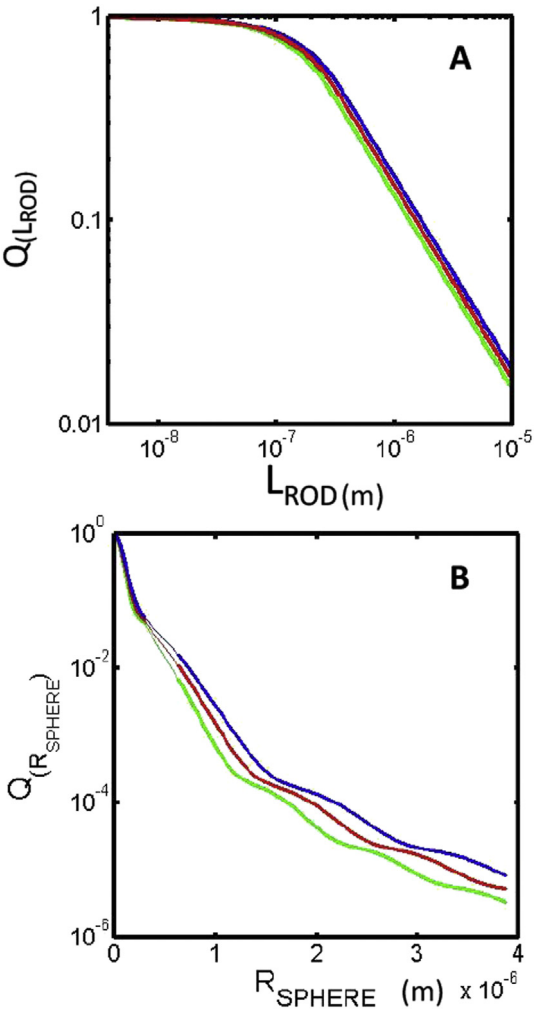


Fig.5. (A) RGD and AD estimation of Q for sphere-like aggregates for three different wavelengths (400 nm [blue], 450 nm [red], and 500 nm [green]). Note the smooth one-dimensional polynomial fitting of Q across the theory juncture acting to fill missing data points. (B) Results of RGD approximations of Q for rod-like aggregates. Note the region of extension for Q beyond the RGD theory conforming to its earlier characteristic dependence. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$I_0 = A_0 i_0 \tag{3c}$$

$$\frac{I_S}{I_0} = \frac{24\pi^3 (V_i)^2_{TRACE}}{\lambda^4}\left(\frac{m_i^2 - 1}{m_i^2 + 2}\right)^2 \left(\frac{1}{A_o}\right) \tag{3d}$$

$$\frac{I_S}{I_0} = LC_i \frac{24\pi^3 (V_i)^2_{TRACE}}{\lambda^4}\left(\frac{m_i^2 - 1}{m_i^2 + 2}\right)^2. \tag{3e}$$

In Eq. (3), i_0 refers to the initial forward intensity of light and A_0 is the area of illumination by the light beam such that the area integrated power relation for the incident light is Eq. (3c). For a single particle, the ratio of power lost to that incident is given by Eq. (3d). For a solution in which there are multiple particles in the incident beam, the total relative power loss due to scattering is shown in Eq. (3e), where C_i is the number density of particle i in solution and L is the optical path length.

The m_i term in Eq. (3) refers to the relative refractive index of the aggregate (relative to the solvent). A numerical value of m_i can be calculated on the basis of the solvent refractive index, n_o, and the incremental change in solution refractive index with incremental change in aggregate weight concentration of the particle, dn/dc_i (Eq. (4a)). Empirical formulas developed from tabulated data can be used to account for the wavelength dependence of n_o and dn/dc_i [35] (Eqs. (4b) and (4c)):

$$m_i = 1 + \alpha_i[(dn/dc)(1/v)]/n_0 \tag{4a}$$

$$n(\lambda) = 1.3403\left[0.9922 + 2.31 \times 10^{-15}/\lambda^2\right] \tag{4b}$$

$$dn/dc_i(\lambda) = 0.19 \times 10^{-3}\left[0.925 + 2.2 \times 10^{-14}/\lambda^2\right]. \tag{4c}$$

Eq. (3) is valid for particles much smaller than the wavelength of light ($R_i < \lambda/20$) when illuminated by an unpolarized source. In a standard spectrophotometer arrangement, the scattering of the light beam by the particles in solution decreases the power of the transmitted light reaching the detector, $I_D = I_0 - I_S$, in a manner proportional to the length of the solution light path. The path length-corrected excess turbidity[5] is defined as the natural logarithm of the fractional reduction of the incident power, I_0, due to cumulative scattering outside of the forward direction, I_S, per path length, dL, traveled (Eq. (5a)).[6] This logarithmic term can be expressed as an infinite series (Eq. (5b)) that can be adequately described by the leading term for values of $I_S/I_0 < 0.1$, that is, for transmittance of greater than 90% (Eq. (5c))[7] [9,36,37]:

$$\tau = -\left(\frac{1}{I}\right)\frac{dI}{dL} = -\frac{1}{L}\log_e\left(1 - \frac{I_S}{I_0}\right) = 2.303(O.D.) \tag{5a}$$

$$\tau = \frac{1}{L}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{I_S}{I_0}\right)^n \tag{5b}$$

$$\underset{(\lim \tau \to 0)}{\tau} \approx \frac{1}{L}\left(\frac{I_S}{I_0}\right). \tag{5c}$$

Inserting Eq. (3e) into Eq. (5b) and truncating the expansion at the first term (Eq. (5c)) produces an expression that accurately describes turbidity values in the low concentration and small particle size limits (i.e., the Rayleigh limit ~ $R_i < \lambda/20$, $\tau_i \to 0$). As the calculated extent of scattered light I_S/I_0 increases beyond ~0.1, the correction described in Eq. (5a) must be performed to correct for errors introduced by premature truncation of the series. In the dilute limit,[8] the turbidity for a solution of differently sized and shaped aggregates is equal to the sum of the contributions to turbidity from each particle (Eq. (6)):

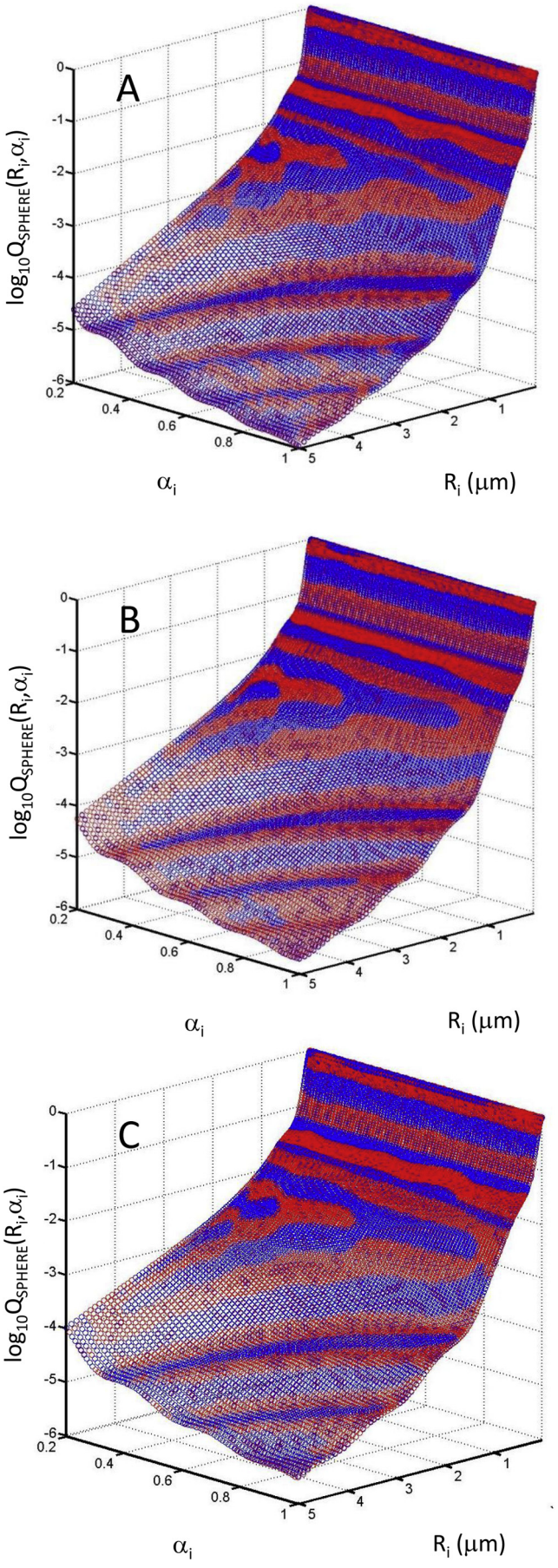
$$\underset{(\lim \tau \to 0)}{\tau} = \sum_{i=1}^{N} \tau_i. \tag{6}$$

[5] By excess turbidity, we mean solvent baseline subtracted.

[6] Thereby making it equal to $\log_e(10) \approx 2.303$ multiplied by the solvent-corrected solution optical density (OD).

[7] With a value of $I_S/I_0 < 0.1$ corresponding to an error in the turbidity of less than 5%.

[8] The dilute limit refers to the concentration range at which we may neglect significant re-scattering of the scattered light, by other aggregate particles in solution, back into the collimated detection pathway. This re-scattering effect is sometimes called external interference. The dilute limit is not influenced by internal interference due to the particles being large in relation to the wavelength of light; therefore, Eq. (7) can retain its validity across the range of sizes.



Eq. (6) is an important relation, the veracity of which we depend on throughout this work and the consecutive articles in this series.

Calculating turbidity for different particle sizes

Eq. (5) provides us with a general means for understanding the relationship between the total scattered light and the solution turbidity. However, as discussed in this section, there are a range of different theoretical models, of increasing complexity, for calculating the aggregate turbidity across different particle size regimes [9,19,20,36,37]. Because a major aim of the current work was to integrate these different approaches to produce a useful empirical interpolation, it will prove helpful to recast the turbidity equation, shown by Eq. (5), in the form outlined by Eq. (7):

$$\underset{(\lim\tau\to 0)}{\tau_i} = C_i F_i Q_i. \tag{7}$$

As formulated by Eq. (7), τ_i, the turbidity per meter at limiting dilution, is composed of three terms: C_i, F_i, and Q_i. C_i is the scattering particle concentration (units of molecules m^{-3}), F_i is the idealized turbidity per unit molecular concentration that would be generated if the particle scattered light as if it were a point mass (units of $m^2 \cdot molecule^{-1}$), and Q_i is the unitless transmittance form factor that modulates the ideal turbidity to account for the effects of the shape and size of the particle. In what follows, we consider the appropriate functionalization of Eq. (7) for three particle size regimes of scattering.

Rayleigh approximation

For particle sizes smaller than $\lambda/20$ (Fig. 3A), we may describe the functional form of Eq. (7) in terms of the trace volume radius described in Eq. (1c):

$$(F_i)_{RAYLEIGH} = \frac{24(V_i)^2_{TRACE}\pi^3}{\lambda^4}\left(\frac{m_i^2-1}{m_i^2+2}\right)^2 \tag{8a}$$

$$(Q_i)_{RAYLEIGH} = 1. \tag{8b}$$

An important distinguishing feature of Rayleigh-type scattering is that it exhibits an approximately inverse fourth-power dependence of scattering and turbidity on wavelength [19,37] (see also chap. 5 of Ref. [9]).

RGD approximation

Predicting the extent of light removed from the forward direction due to scattering becomes progressively harder as the aggregate size increases (in relation to the wavelength of light) for two reasons (Fig. 3):

1. In large aggregates, different regions of the particle will sample different phases of the incident light's oscillating electric field vector, resulting in a complex super-positioning of scattered light [16,19,37,38] (see also chap. 6 of Ref. [9]) (Fig. 3B).

Fig.6. Smooth two-dimensional polynomial fitting of $\log_{10}Q_i$ for spherical aggregates as a function of R_i and α_i for three different values of λ: (A) $\lambda = 400$ nm; (B) $\lambda = 450$ nm; (C) $\lambda = 500$ nm. Two-dimensional polynomial was of the form shown by Eq. (14), with values of the coefficients given in Appendix A.

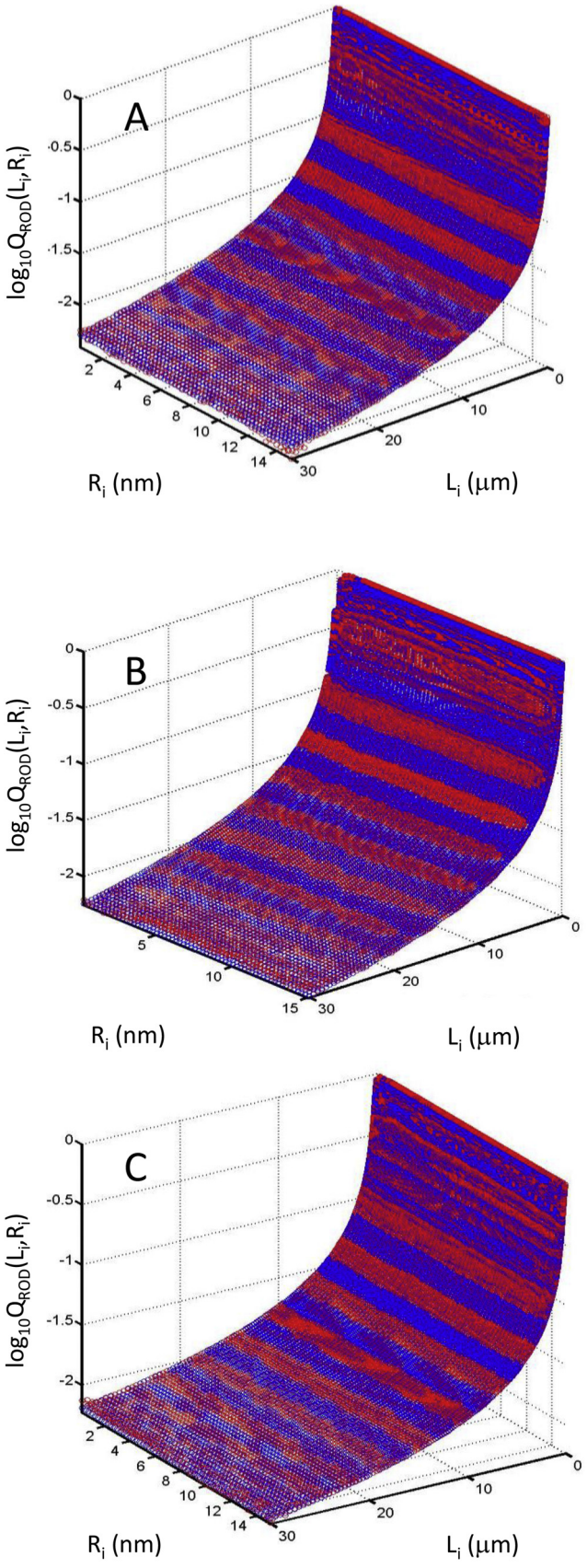


Fig.7. Smooth two-dimensional polynomial fitting of $\log_{10}Q_i$ for rod-like aggregates as a function of L_i and $\log_{10}R_i$ for three different values of λ: (A) $\lambda = 400$ nm; (B) $\lambda = 450$ nm; (C) $\lambda = 500$ nm. Two-dimensional polynomial was of the form shown by Eq. (15), with values of the coefficients given in Appendix A.

2. As aggregates grow even larger, there exists a significant diminution of the intensity of the incident light as it moves through the aggregate [17,19,20,37] in addition to the phase difference (Fig. 3C and D).

The RGD formalism addresses the first difficulty but not the second and so is appropriate for particles of rotationally averaged cross-sectional radius, $\langle R_i \rangle$, approaching half the wavelength of light, that is, $2\langle R_i \rangle < \lambda/n_o(\lambda)$ [38] (see also chap. 6 of Ref. [9]). In the RGD approach, the net scattering intensity produced by the aggregate is treated as the superposition of the scattering from N discretized centers within the aggregate (without correcting for diminution of incident light). The fractional reduction in the scattered light intensity due to intraparticle interference (i.e., $i(r,\theta)_{\text{real}}/i(r,\theta)_{\text{ideal}}$) is termed the particle form factor and is denoted by $P(\theta,\lambda)$. It can be calculated based on knowledge of the internal distances separating the discretized scattering centers, d_{ij}, and the magnitude of the scattering vector, h, through use of the Debye equation (Eq. (9)) [33,36,37]:

$$P(\theta,\lambda) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sin(hd_{ij})}{hd_{ij}} \tag{9a}$$

$$h = 4\pi n \sin(\theta/2)/\lambda \tag{9b}$$

$$d_{ij} = \sqrt{(x_i - x_j)^2 + \left(y_i - y_j\right)^2 + (z_i - z_j)^2}. \tag{9c}$$

Insertion of Eq. (9a) into Eq. (3a) along with its subsequent spherical integration, due to an unpolarized light source, yields Q_i, the transmittance form factor (Eq. (10b)):

$$(F_i)_{RGD} = (F_i)_{RAYLEIGH} \tag{10a}$$

$$(Q_i)_{RGD} = \int_0^{\pi} P(\theta,\lambda)\left(1+\cos^2\theta\right)\sin\theta d\theta. \tag{10b}$$

Eq. (10b) was solved numerically to generate values of $Q_i(\lambda)$ for discrete approximations of rods and spheres over the RGD regime [8,12,14,39] (Fig. 4). These data were then empirically fitted using either a sigmoidal-like function (spheres, Eq. (11a)) or a hybrid sigmoidal-type inverse function (rods, Eq. (11b)) to re-parameterize it in terms of the characteristic dimensions of radius or length (or both). The resultant best-fit equations are shown as Eq. (11a) (spheres) and (11b) (rods):

Spheres:

$$(Q_i)_{RGD}(x_1) = 1 - 0.955\left(1 - e^{-6.48x_1}\right)^{2.40} \quad \text{for } x_1 < 0.5, \tag{11a}$$

where $x_1 = R_i n(\lambda)/\lambda$ and Q_i are defined for $x_1 = [0, 0.5]$.

Rods:

$$(Q_i)_{RGD}(x_1,x_2)=\begin{cases}1-0.955\left(1-e^{-6.48x_1}\right)^{2.40} & \text{for } x_1<x_R\\ (Q_i)_{RGD}(x_1)\left[1-0.995\left(1-e^{-1.08x_2}\right)^{1.275}\right]^{1} & \text{for } x_1=x_R, x_2<1\\ (Q_i)_{RGD}(x_1)\Big/\left(2.40x_2^{0.95}\right) & \text{for } x_1=x_R, x_2\geq 1\end{cases}, \tag{11b}$$

where $x_2=L_i n(\lambda)/\lambda$, x_1, and Q_i are defined for $x_1=[0,x_R]$; $x_2=[0,3]$.

Mie scattering calculations for spheres

The Mie scattering equation [40] is an exact theoretical solution to the scattering problem for particles of arbitrary size in relation to the wavelength of light. Generally speaking, the procedure (Fig. 3C and D) accounts for change in positional intensity of the incident light as it travels through the particle in addition to calculation of the multiple internal scattering superposition effect [9,40]. For nonpolarized light, the first aspect of this statement can be simply formalized as Eq. (12), for which the terms have been defined previously:

$$I(x,y,z)=f(x,y,z,A_i(\phi),S_i,m_i,\lambda). \tag{12}$$

The anomalous diffraction (AD) approximation constitutes a very accurate and simple closed-form representation of the Mie scattering equations for spherical aggregate geometry [20,37,40] valid for particles ranging in size from approximately two times the wavelength to 15 times the wavelength of light (i.e., $2\lambda \leq \langle R_i\rangle \leq \sim 15\lambda$) that also have a relative refractive index, m_i, less than 1.3 [20]:

$$(F_i)_{AD}=(F_i)_{RAYLEIGH} \tag{13a}$$

$$(Q_i)_{AD}=\left[2-\left(\frac{4}{\rho}\right)\sin\rho+\left(\frac{4}{\rho^2}\right)(1-\cos\rho)\right]\Bigg/\left[\frac{(F_i)_{RAYLEIGH}}{\pi\left\langle R_i^2\right\rangle}\right] \tag{13b}$$

$$\rho=\frac{4\pi\langle R_i\rangle(m_i-1)}{\lambda}. \tag{13c}$$

Two notable features of the Mie/AD description Eq. (13) are (i) the damped decaying sinusoidal nature of the turbidity as the aggregate size increases and (ii) the lack of uniqueness in the value of the predicted turbidity against particle size. In relation to aggregate growth, an interesting extension of these two points is that the turbidity of spherical aggregates should, after an initial rise, undergo a decrease as the average aggregate radius exceeds the magnitude of 2λ.

For the case of stiff rods of high aspect ratio, the loss of intensity of the incident light through the rod is significant when the rods lie parallel to the incident beam, that is, $\phi=0$ [41,42] (see Figs. 1C and 3D). This angular dependence leads to small deviations from the RGD case for randomly oriented fibers; however, large changes for nematically ordered fibers can occur, depending on the angle of fiber alignment relative to the incident light [41,42]. As such, a suitable approximation for estimating $Q_i(\lambda)$ for low concentrations of randomly oriented rod-like fibers beyond the RGD domain, where $L_i > 3\lambda/n_o(\lambda)$, involves assuming that the limiting extrapolation over the domain $[\lambda/n_o(\lambda),3\lambda/n_o(\lambda)]$ extends indefinitely. Experimental and theoretical support for this contention was originally provided by Gaskin and coworkers [43] and Berne [44], who showed that turbidity generated by solutions composed of long microtubule rods remained approximately unchanged after surpassing a size corresponding to $L_i \geq 3\lambda/n_o(\lambda)$.[9] The basic correctness of this approximation is also borne out via more complex simulations based on numerical evaluations [9,40,42].

Single functions for Q for spheres/rods

We previously outlined three methods for calculating Q over different and noncontiguous sets of wavelength domains for rod-like and spherical aggregates of varying internal density. In this section, we first outline strategies for estimating values of Q in regions not covered by either theoretical approach (Fig. 5). We then use these interpolated/padded data sets to develop single empirical functions that approximate Q over the entire range of aggregate dimensions, nanometers to tens of micrometers, in the visible wavelength region.

Spheres — padding data/continuous description

To develop a continuous description of Q versus sphere radius, a one-dimensional interpolating polynomial was first used to fill values of Q over the intervening region between the RGD and AD results (Fig. 5A). It can be noted from Eqs. (13b) and (13c) that in the AD description, Q is dependent on the effective density of protein [20,41]. As such, a set of data describing the dependence of Q on sphere radius and aggregate density α was simulated using Eqs. (11a) and (13), and the resultant collected data set was fitted to a high-order two-dimensional polynomial. For reasons relating to improving the goodness of the achieved fit, $\log_{10}Q$ was fitted against α and R (Fig. 6) with the form of the characteristic empirical equations shown as Eq. (14) (coefficients given in Appendix A for λ equal to 400, 450, and 500 nm):

$$\log_{10}Q_i(R_i,\alpha_i)=\sum_{j=0}^{N}\sum_{k=0}^{N}u_{j,k}(R_i)^j(\alpha_i)^k. \tag{14}$$

Rods — padding data/continuous description

Based on the arguments advanced in the preceding section, values for Q covering a range of amyloid sizes with regard to length ($0 < L_i < 30$ μm) and radius (0.25 nm $< R_i <$ 15 nm) were calculated using Eqs. (11a) and (11b). Again a high-order two-dimensional polynomial was used to find an appropriate interpolation, this time for $\log_{10}Q$ against R and L (Fig. 7). As can be noted from the associated residual plot, this fitting procedure provides a suitable description of the data set within the limits of the specified

[9] Such a finding can be rationalized mathematically by inserting relevant values of Q from Eq. (10b) for $L > 3\lambda/n_o(\lambda)$ into the general turbidity expression developed here (see Appendix A).

domain.[10] The form of the polynomial function is shown as Eq. (15), with the exact values for these characteristic empirical equations for λ equal to 400, 450, and 500 nm given in Appendix A:

$$\log_{10} Q_i(L_i, R_i) = \sum_{j=0}^{N} \sum_{k=0}^{N} w_{j,k} (L_i)^j (R_i)^k. \tag{15}$$

Simulation of total aggregate turbidity

In the preceding sections, we outlined utilitarian approaches for (i) assigning protein aggregate dimensions in terms of the degree of polymerization, i, a presumed shape, S_i (which we limit to a cylindrical rod or a sphere), and an effective packing density, α_i, and (ii) an empirical interpolation of three forms of light scattering theory that cover the range of sizes from effective point particles through to a 10-μm diameter sphere and a 30-μm length, 30-nm width cylindrical rod. In combination, these approaches can be used to calculate the turbidity of a solution of mixed protein aggregates using Eqs. (16) and (17) (used with Eq. (5a) for the logarithmic correction):

Spheres: Eq. (16)

$$\tau_i(i, M_1, C_i, \alpha_i, \lambda) = \frac{24 C_i \left(M_1^2 i^2 \upsilon^2 \Big/ \left[\alpha^2 N_A^2 \right] \right)}{\lambda^4} \left(\frac{m_i^2 - 1}{m_i^2 + 2} \right)^2 (Q_i)_{sphere}\{R_i, \alpha_i\}, \tag{16}$$

where

$$R_i = \sqrt[3]{\frac{3}{4\pi} \left(\frac{i M_1 \upsilon}{\alpha_i N_A} \right)}$$

Rods: Eq. (17)

$$\tau_i(i, M_1, C_i, \alpha_i, R_{rod}, \lambda) = \frac{24 C_i \left(M_1^2 i^2 \upsilon^2 \Big/ \left[\alpha^2 N_A^2 \right] \right)}{\lambda^4} \left(\frac{m_i^2 - 1}{m_i^2 + 2} \right)^2 (Q_i)_{rod}\{L_i, R_i\}, \tag{17}$$

where

$$L_i = 0; R_i = \sqrt[3]{\frac{3}{4\pi} \left(\frac{i M_1 \upsilon}{\alpha_i N_A} \right)} \quad \text{when } \mathrm{R_i} < \mathrm{R_{rod}}$$

$$L_i = \left(\frac{i M_1 \upsilon}{\alpha_i N_A} \right) \Big/ \pi R_{rod}^2 \quad \text{when } \mathrm{R_i} = \mathrm{R_{rod}}.$$

In conjunction with the empirical interpolations for Q developed in Eqs. (14) and (15), the above formulation allows for straightforward transformation of protein aggregate distributions into turbidity traces over a wide range of sizes for a limited number of protein aggregate shapes. Figs. 8 and 9 demonstrate a set of practical numerical examples of turbidity generated by a 1-mg/ml solution of a protein of monomer molecular weight $M_1 = 5000$ g/mol for a range of aggregate densities, α, at three different wavelengths: 400, 450, and 500 nm.

A test simulation was set up to demonstrate the incorporation of the turbidity transformation procedure into a kinetic simulation of aggregate growth (Fig. 10). In a similar vein to the work of Yoshimura and coworkers [22], Hall and coworkers [31], and Adachi and coworkers [45], two types of protein aggregation reaction, amorphous and amyloid, were considered to be in competition for a common monomer pool. A simplified set of equations was used to describe the average properties of the two competing polymer types, designated A for amyloid and G for globular amorphous (Eqs. (18–20)):

Eq. Set (18)

$$\begin{aligned}
&nM \xrightarrow{k_1} N \ \text{ or } \ bM \xrightarrow{k_3} B \\
&N + M \xrightarrow{k_2} A_{n+1} \ \text{ or } \ B + M \xrightarrow{k_4} G_{b+1} \\
&A_{n+1} + M \xrightarrow{k_2} A_{n+2} \ \text{ or } \ G_{b+1} + M \xrightarrow{k_4} G_{b+2} \\
&\cdots \\
&A_{i-1} + M \xrightarrow{k_2} A_i \ \text{ or } \ G_{j-1} + M \xrightarrow{k_4} G_j
\end{aligned} \tag{18}$$

$$\frac{dC_N}{dt} = k_1 (C_M)^n - k_2 C_N C_M$$

$$\frac{d \sum C_{A_i}}{dt} = k_2 C_N C_M$$

$$\frac{d \sum i C_{A_i}}{dt} = (n+1) k_2 C_N C_M + k_2 \sum C_{A_i} C_M \tag{19}$$

$$\frac{dC_B}{dt} = k_3 (C_M)^b - k_4 C_B C_M$$

$$\frac{d \sum C_{G_j}}{dt} = k_4 C_B C_M$$

$$\frac{d \sum j C_{G_j}}{dt} = (b+1) k_4 C_B C_M + k_4 \sum C_{G_j} C_M \tag{20}$$

Each mechanistic pathway is regulated by a set of rate

[10] Although the residuals are not random and exhibit slight oscillatory behavior, the associated error is very small.

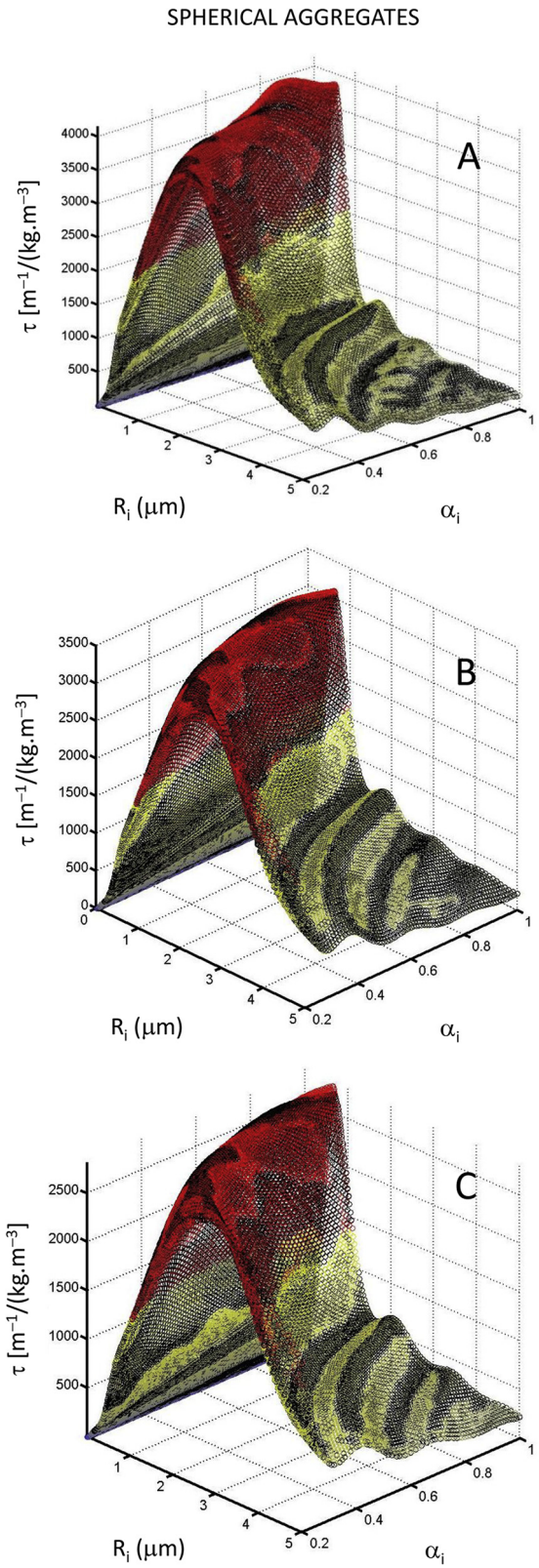


Fig.8. Calculation of specific turbidity (turbidity per kg/m^3 of aggregate) for protein existing as spherical aggregates. Shown is turbidity for the case of 1 mg/ml total protein existing as spherical aggregate of a certain radii and internal packing density for three different wavelengths: (A) $\lambda = 400$ nm; (B) $\lambda = 450$ nm; (C) $\lambda = 500$ nm. Black symbols represent calculation based on derived empirical relations (Eqs. (14) and (16)), whereas colored symbols represent calculations based on Eq. (11a). Protein monomer molecular weight M_1 was set at 5000 g/mol. (For interpretation of the reference to color in this figure legend, the reader is referred to the Web version of this article.)

constants and a critical nucleus species [A; k_1, k_2, N], [G; k_3, k_4, B]. Fig. 10 describes the results of simulations carried out by numerical integration of Eq. Sets (19) and (20), followed by subsequent transformation of the simulated aggregate product into excess solution turbidity by Eqs. (16) and (17). The left-hand panels of Fig. 10 show the kinetics of the effective weight concentration of each aggregate type along with the transformed turbidity values derived from this time course. The right-hand panels describe the evolution of the size characteristics of the aggregates. For the size regime of aggregates described here, the simulations reveal that the turbidity developed by the spherical aggregate is much greater than that developed by the same molecular weight rod-like aggregate, dwarfing it by a roughly 10-fold ratio.

Discussion

Due to its simplicity, and lack of requirement for specialist equipment (aside from a spectrophotometer), visible wavelength turbidity has become the principal assay technique for studying differential changes in protein aggregation state across a large number of scientific fields [10,11,20,36]. This ubiquity speaks to the need for development of a strong understanding[11] of the turbidity assay so as to gain scientific insight into areas as diverse as the biomedical investigation of proteopathic diseases,[12] cell biology research into the cytoskeleton,[13] food science formulation chemistry,[14] pharmaceutical science,[15] and colloid chemistry.[16]

Our motivation in the current work was to gain a deeper understanding of the turbidimetric principle. In the current work, we set ourselves two goals. The first goal was to develop functional theory capable of approximating the differential turbidity generated by a heterogeneous protein aggregate distribution of defined composition. The second goal was to tailor the theory to a level of complexity sufficient to form the basis of a future analytical routine. There are a number of competing aspects of these two goals that constrained the development of our approach.

The general nature of the first goal (composition → turbidity) makes its solution, in principle, relatively straightforward. However, in practice there exist a number of complications that prevent such a straightforward solution from being attained. The first difficulty is that polymer heterogeneity in protein aggregation reactions often takes on extreme forms with a very large potential range of aggregate structures, sizes, and densities [6,14,20,27,39,41]. This wide variation in structural composition makes judicious assignment of a reduced number of representative scattering objects (i.e., a relatively few basic structural models of the scattering components) a nontrivial challenge (e.g., see Ref. [15]). Similarly, the range of sizes across which the protein aggregates can span (nm to μm) necessitates the adoption of a number of different levels of noncontiguous approximations to the scattering theory [9,20]. Finally, complications induced by interparticle scattering effects [9,37,39,56] (seen at high concentrations of scattering particles) restricts any descriptive/analytical efforts to the dilute limit. The opposite nature of the second goal

[11] Or at least the gaining of an empirical handle.
[12] For example, amyloidosis [1,3,46], clotting [47], and sickle cell disease [48].
[13] For example, polymerization of tubulin into microtubules [14,49,50].
[14] For example, dairy protein aggregation and stabilization [51], muscle protein aggregation [13], and efficacy of protein emulsification agents [52].
[15] For example, stability of peptide drug formulations [7,11,53].
[16] For example, soft matter/colloidal investigations of aggregation kinetics and non-ideal solution theory [54,55].

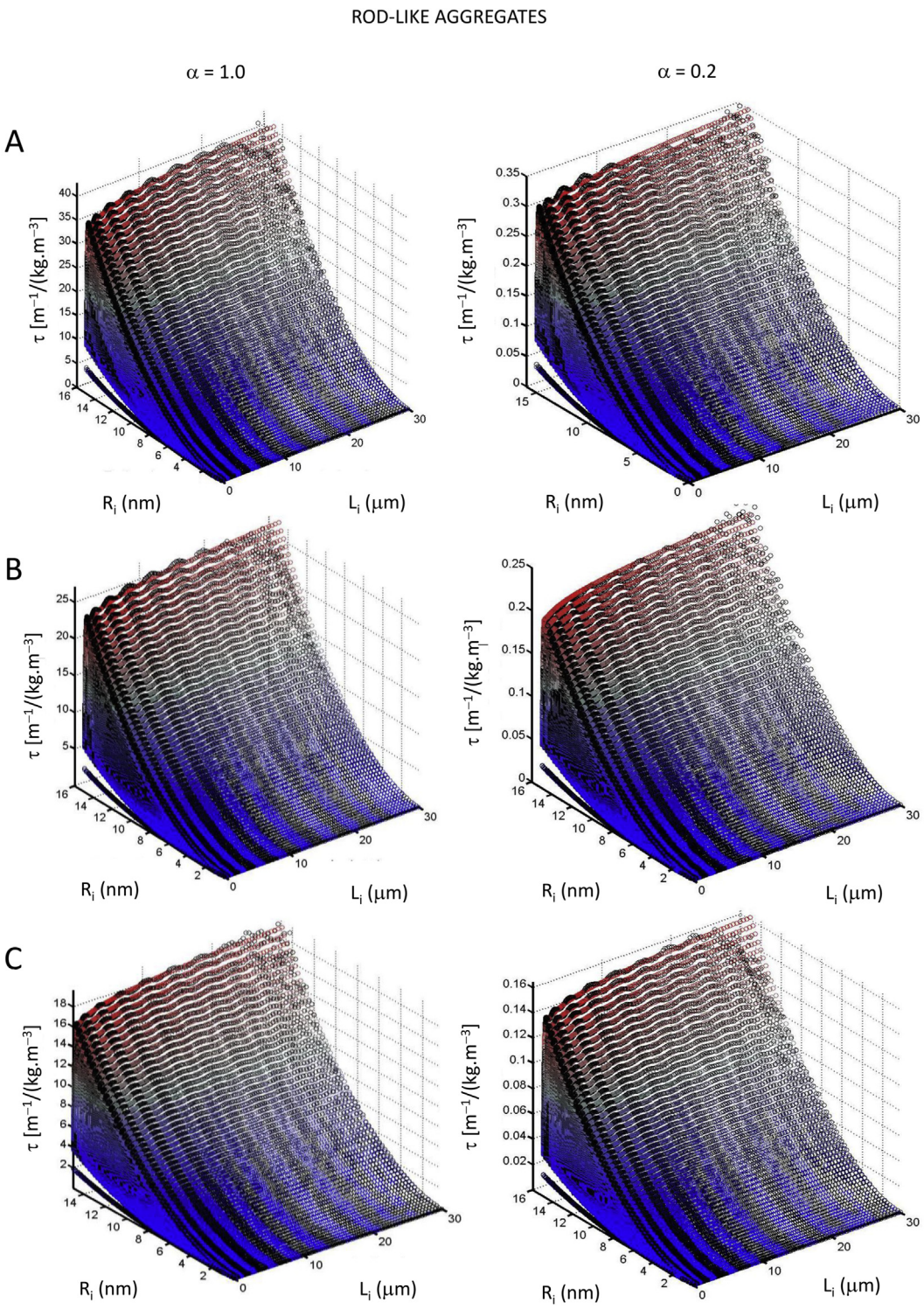


Fig.9. Calculation of specific turbidity (turbidity per kg/m^3 of aggregate) for protein existing as rod-like aggregates. Shown is turbidity for the case of 1 mg/ml total protein existing as rod-like aggregates of variable length L and radius R for three different values of λ: (A) $\lambda = 400$ nm; (B) $\lambda = 450$ nm; (C) $\lambda = 500$ nm. Total packing density was varied such that $\alpha = 1$ for left panels and $\alpha = 0.2$ for right panels. Black symbols represent calculation based on derived empirical relations (Eqs. (15) and (17)), whereas colored symbols represent calculations based on Eq. (11b). Protein monomer molecular weight M_1 was set at 5000 g/mol. (For interpretation of the reference to color in this figure legend, the reader is referred to the Web version of this article.)

(turbidity → composition) makes for a more complex challenge, representing an inverse problem typically soluble only upon application of a number of restraining conditions such as the assumption of the form of a size distribution [38,57]. Because the use of such equations in the analysis of experimental data ultimately requires parameter estimation through nonlinear

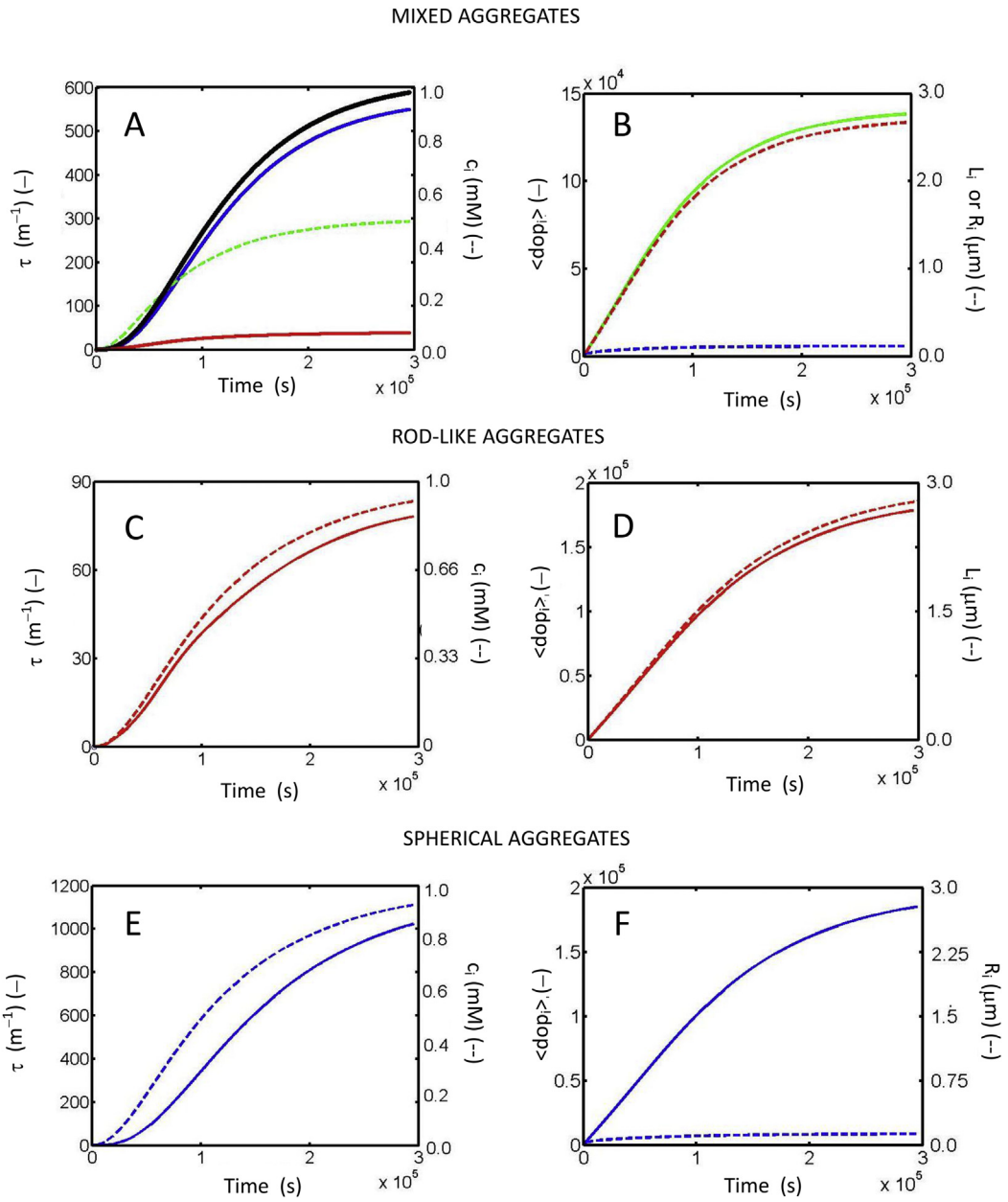


Fig.10. Application of the turbidity transform to simulations of competitive growth between rod-like and spherical aggregates. The left panels (A,C,E) describe the reaction progress curves simulated by numerical integration of Eqs. (18–20), c_i (dashed lines), and their associated transformation via Eqs. (14–17) into excess solution turbidity, τ_i (solid lines), whereas the right panels (B,D,F) describe the kinetic evolution of the average degree of polymerization <dop> (solid lines) and the average dimensions of each reaction component for a number of cases of competitive growth, L_i or R_i (dashed lines). The top, middle, and bottom panels reflect various cases of aggregate competition between rod-like (red lines) and spherical (blue lines) aggregate growth. (A,B) Mixed aggregates: Equivalent mixed rod-like and spherical aggregate competitive growth {$n = 4$; $k_1 = 0.1$ M^{-3} s^{-1}, $k_2 = 2000$ M^{-1} s^{-1}, $b = 4$; $k_3 = 0.1$ M^{-3} s^{-1}, $k_4 = 2000$ M^{-1} s^{-1}}. Green lines indicate equal rate of growth and development of polymer dop for rod and spherical aggregates, and black lines describe total turbidity. (C,D) Rod-like aggregates: Exclusive rod-like aggregate formation {$n = 4$; $k_1 = 0.1$ M^{-3} s^{-1}, $k_2 = 2000$ M^{-1} s^{-1}, $b = 4$; $k_3 = 0$ M^{-3} s^{-1}, $k_4 = 0$ M^{-1} s^{-1}}. Red lines indicate rod turbidity, blue lines indicate sphere-like turbidity, black line describes total turbidity, and green dotted line indicates concentration of monomers as either rods or spheres. (E,F) Spherical aggregates: Exclusive spherical aggregate formation {$n = 4$; $k_1 = 0$ M^{-3} s^{-1}, $k_2 = 0$ M^{-1} s^{-1}, $b = 4$; $k_3 = 0.1$ M^{-3} s^{-1}, $k_4 = 2000$ M^{-1} s^{-1}}. Note that the maximum average radius in the top panels is 64.5 nm, whereas that in the bottom panels is 71.1 nm. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

regression procedures, the chance of affecting a solution is significantly improved through reduction of the overall number of search parameters. This future requirement necessarily limited the complexity of any developed theoretical approach to that involving a low number of variable parameters.

In the current work, we have taken a utilitarian approach to achieving the necessary compromise between the two outlined goals—one that affords us semi-quantitative understanding, via a simulation capability, along with the potential for theory-grounded data reduction. With regard to defining a range of scattering models capable of accounting for heterogeneous protein aggregation,

we selected just two scattering models: isotropic spheres and isotropic cylindrical rods possessing a continuum of potential packing densities from $\alpha = 0.2$ to 1.0 (Eqs. (1) and (2)). The justification for adopting this approach was that (i) spheres and rods represent the extreme cases of scattering behavior (rod—least and sphere—most scattering per unit volume [9,39,41]), (ii) spheres and rods comport to the type of shape information easily observable from nanoscopic microscope data such as atomic force microscopy (AFM) and transmission electron microscopy (TEM) [27], and (iii) the assumption of isotropic character[17] for the aggregate (representing a mean field-type approximation allows for easy factoring into the model of the relative trace volume, relative packing density, and relative refractive index of the aggregate $[(V_i)_{TRACE}, \alpha_i, dn/dc_i]$.

To deal with the large range of sizes encountered in protein aggregation reactions, we employed three different ranges of already developed light scattering theory corresponding to the Rayleigh, RGD, and Mie scattering equations [9,20,36,39,40] to produce a continuous estimate of Q, the particle transmission dissipation factor, over a defined range of aggregate sizes. We joined these noncontiguous data sets together using a high-order one-dimensional polynomial as an interpolating function to fill in missing regions (Fig. 5). We then described these completed data sets in the form of high-order two-dimensional polynomials (spheres: Eq. (15); rods: Eq. (14); see also Appendix A).

The theoretical accounting of turbidity gets significantly more difficult at high aggregate concentrations due to the twin requirements of using higher order terms from the infinite series in the evaluation of the turbidity (Eqs. (5a—c)) and calculation of the contribution of inter-particle scattering in addition to the intra-particle scattering contribution computed using Eq. (5) [9,59]. To cope with the problem associated with premature truncation of the infinite series (Eq. (5b)), we instituted a procedure whereby any estimate of the turbidity was calculated over a sufficiently short path length to ensure that $I_S/I_O < 5\%$. The logarithmic correction was then applied to this value prior to reevaluation of the turbidity per path length. To cope with the problem associated with inter-particle scattering (occurring at high concentrations), we arbitrarily imposed the low concentration limit into our theoretical formulation [9,20]. Such a simplifying assumption can be experimentally reconciled in practice by serial dilution of the aggregate, prior to its measurement by turbidity, so as to identify that the experiment is being conducted in the linear single scattering region [9].

The competing goal, which tempered the form of our approach, was to develop a simulation routine containing the least number of independent variables. This was done with the reason that such a simplified analytical description could provide the basis of a future analysis routine. Rapid simulation of postulated turbidity profiles provides a basis for matching of the parameters to a turbidity profile derived from an unknown experimental distribution. In addition, the ready capacity to analytically differentiate a function to determine maxima and minima criteria is highly desirable in any prospective analytical function. The ease of simulating and manipulating explicit two-dimensional polynomials of Q_i based on terms reflecting α_i, R_i (for spheres), and L_i and R_i (for rods) holds real possibilities for a semi-empirical analysis of the turbidity derived from mixed aggregate profiles—something we will explore in the next article of this series.

Although not presenting any new theory in this article, we have combined and recast the general turbidity relations applicable to three different size regimes of scattering particles into an exceptionally utilitarian form (i.e., a polynomial) that allows for rapid and straightforward accounting of solution turbidity. The approach adopted suggests a path forward for development of a rapid analytical procedure for the interrogation of turbidity profiles. In this way, our work extends the aims of earlier approaches [8,12–19,38] that were designed to analyze and predict the size and molecular weight of homogeneous (of size and shape or both) protein aggregates. Our work should prove helpful to researchers performing kinetic studies of protein aggregation reactions that are entirely, or in part, dependent on a better understanding of the interplay between turbidity and aggregate composition[18] (see, e.g., Refs. [13,21–23,45,48,53,61]).

Acknowledgments

D.H. thanks Nami Hirota, Thomas Huber, Akira Kinjo, Junichi Higo, Haruki Nakamura, and John White for both helpful discussions and comments on an earlier version of the manuscript. The work of D.H. was jointly funded by an Australian National University Senior Research Fellowship and an Osaka University Cross-Appointment as an Associate Professor. R.Z. thanks the Chinese government and Australian National University for the award of the CSC—ANU Scholarship. The work of S.R.W. was funded by the Australian Research Council. The work of F.A. was funded by Funds in Aid provided by Nihon University. The research of Y.G. was supported by the Japanese Ministry of Education, Culture, Sports, Science, and Technology. The research of J.A.C. is supported by a Project Grant from the National Health and Medical Research Council of Australia.

Appendix A

Coefficients of the approximate interpolating polynomial functions for $\log_{10}Q$ for spheres described in terms of radius and packing density (Eq. (14)) and $\log_{10}Q$ for rods described in terms of length and radius (Eq. (15)) at 400 nm (corresponding coefficients at 450 and 500 nm are provided in the online supplementary material)

Coefficients for $\log_{10}\{Q_{SPHERE}(R_i,\alpha_i)\}$ *at* $\lambda = 400$ *nm*

[−8.9612226e+142,25,0], [−6.2178673e+137,24,1],
[4.3219257e+138,24,0], [1.9681305e+132,23,2],
[2.9802927e+133,23,1], [−9.1344692e+133,23,0],
[−2.0872587e+127,22,3], [−6.1403387e+127,22,2],
[−6.5503778e+128,22,1], [1.0754194e+129,22,0],
[1.4326259e+122,21,4], [7.8057605e+122,21,3],
[5.1240706e+122,21,2], [8.717724e+123,21,1],
[−7.0390456e+123,21,0], [−5.600057e+116,20,5],
[−6.3051042e+117,20,4], [−1.102367e+118,20,3],
[5.7949299e+117,20,2], [−7.7743739e+118,20,1],
[1.4478287e+118,20,0], [−1.6487249e+111,19,6],
[3.8066915e+112,19,5], [1.0375059e+113,19,4],
[5.6493303e+112,19,3], [−1.8106257e+113,19,2],
[4.7990097e+113,19,1], [1.67219e+113,19,0],
[6.8043805e+106,18,7], [−1.9191226e+107,18,6],

[17] In certain cases, such as for vesicles, the assumption of an isotropic dispersed scattering medium has been shown to be a poor approximation [58]. However, for more evenly dispersed aggregates such as amorphous protein arrangements, this approximation will have greater validity [54].

[18] For a recent semi-historical review of amyloid aggregation kinetics, see Ref. [60].

[−6.7824018e+107,18,5], [−8.5880733e+107,18,4],
[3.8113954e+107,18,3], [2.0993804e+108,18,2],
[−1.971506e+108,18,1], [−1.7030808e+108,18,0],
[−4.3711064e+101,17,8], [−9.3828793e+101,17,7],
[6.050508e+102,17,6], [4.5497109e+102,17,5],
[3.7342016e+102,17,4], [−9.1514938e+102,17,3],
[−1.4618403e+103,17,2], [3.9203551e+102,17,1],
[6.1809811e+102,17,0], [1.0031557e+096,16,9],
[1.2301516e+097,16,8], [−1.0656749e+097,16,7],
[−6.4852961e+097,16,6], [−6.98s42277e+096,16,5],
[3.2673965e+095,16,4], [8.2423695e+097,16,3],
[6.5029863e+097,16,2], [1.2771041e+097,16,1],
[8.2402616e+096,16,0], [−4.7593381e+089,15,10],
[−4.1968359e+091,15,9], [−9.2156469e+091,15,8],
[2.5809873e+092,15,7], [3.1752629e+092,15,6],
[−2.7138467e+091,15,5], [−1.9311451e+092,15,4],
[−4.4672793e+092,15,3], [−1.6441195e+092,15,2],
[−1.631586e+092,15,1], [−2.19639e+092,15,0],
[−2.2654694e+085,14,11], [1.9210622e+086,14,10],
[2.1232382e+086,14,9], [1.6506105e+086,14,8],
[−1.6619091e+087,14,7], [−1.0248018e+087,14,6],
[−3.1239481e+085,14,5], [2.1066011e+087,14,4],
[1.4494944e+087,14,3], [1.4612809e+085,14,2],
[8.9780795e+086,14,1], [1.3446042e+087,14,0],
[2.6000854e+080,13,12], [−9.6418189e+080,13,11],
[−1.0458066e+081,13,10], [1.6647952e+081,13,9],
[−1.1693855e+081,13,8], [4.2077681e+081,13,7],
[7.1950842e+081,13,6], [−1.2000525e+081,13,5],
[−1.2513165e+082,13,4], [−1.9196752e+081,13,3],
[1.8642207e+081,13,2], [−3.5232504e+081,13,1],
[−5.1449328e+081,13,0], [−4.5186753e+074,12,13],
[−4.930812e+075,12,12], [3.2790033e+076,12,11],
[−5.2522596e+076,12,10], [3.6701297e+076,12,9],
[−1.4601403e+076,12,8], [1.3247384e+076,12,7],
[−6.2401347e+076,12,6], [2.7873854e+076,12,5],
[4.5542938e+076,12,4], [−5.5216856e+075,12,3],
[−9.4185061e+075,12,2], [1.0892999e+076,12,1],
[1.4048427e+076,12,0], [−4.7509557e+069,11,14],
[5.5980973e+070,11,13], [−1.5500089e+071,11,12,
[6.3068098e+070,11,11], [2.4366484e+071,11,10],
[−3.5156078e+071,11,9], [1.8727398e+071,11,8],
[−1.1993891e+071,11,7], [2.8244041e+071,11,6],
[−1.5692993e+071,11,5], [−1.1954837e+071,11,4],
[4.1657839e+070,11,3], [2.846992e+070,11,2],
[−2.732297e+070,11,1], [−2.844044e+070,11,0],
[4.8488476e+064,10,15], [−3.1915702e+065,10,14],
[5.8048465e+065,10,13], [3.6474845e+064,10,12],
[−8.7967097e+065,10,11], [3.4833596e+065,10,10],
[5.5126089e+065,10,9], [−3.3361219e+065,10,8],
[4.1054745e+064,10,7], [−4.8146114e+065,10,6],
[2.788692e+065,10,5], [3.4531377e+065,10,4],
[−1.6711513e+065,10,3], [−5.6272472e+064,10,2],
[5.5350546e+064,10,1], [4.2863852e+064,10,0],
[−2.5536597e+059,9,16], [1.3060273e+060,9,15],
[−1.7801094e+060,9,14], [−8.4979954e+059,9,13],
[3.3909958e+060,9,12], [−2.0624336e+060,9,11],
[1.3869602e+060,9,10], [−3.3300779e+060,9,9],
[3.432252e+060,9,8], [−1.9475238e+060,9,7],
[1.455227e+060,9,6], [−3.6089797e+059,9,5],
[−8.9422063e+059,9,4], [4.7407832e+059,9,3],
[6.7449981e+058,9,2], [−8.9075827e+058,9,1],
[−4.7026428e+058,9,0], [9.2275768e+053,8,17],
[−3.9858525e+054,8,16], [3.7003966e+054,8,15],
[6.9218668e+054,8,14], [−1.7588816e+055,8,13],
[1.7394459e+055,8,12], [−1.3603213e+055,8,11],
[6.1410763e+054,8,10], [1.0000524e+055,8,9],
[−2.3052824e+055,8,8], [2.3461508e+055,8,7],
[−1.5811292e+055,8,6], [6.1998177e+054,8,5],
[−1.5238637e+053,8,4], [−5.1602121e+053,8,3],
[−9.8295588e+052,8,2], [1.1823208e+053,8,1],
[3.4805285e+052,8,0], [−4.2715205e+048,7,18],
[2.5002609e+049,7,17], [−6.5082392e+049,7,16],
[1.0854905e+050,7,15], [−1.3859769e+050,7,14],
[1.3623707e+050,7,13], [−1.0596883e+050,7,12],
[1.0045107e+050,7,11], [−1.2784344e+050,7,10],
[1.2054274e+050,7,9], [−5.6259226e+049,7,8],
[−9.6365845e+048,7,7], [3.2881831e+049,7,6],
[−2.1893138e+049,7,5], [6.7771534e+048,7,4],
[−1.1700843e+048,7,3], [3.9439372e+047,7,2],
[−1.5173035e+047,7,1], [−1.2010939e+046,7,0],
[1.8344341e+043,6,19], [−1.1460602e+044,6,18],
[2.9201012e+044,6,17], [−3.4077848e+044,6,16],
[1.2944621e+043,6,15], [5.0547818e+044,6,14],
[−6.1475558e+044,6,13], [1.2277942e+044,6,12],
[3.3459425e+044,6,11], [−1.7895782e+044,6,10],
[−3.0372074e+044,6,9], [5.1862997e+044,6,8],
[−3.6879081e+044,6,7], [1.4159437e+044,6,6],
[−2.4124151e+043,6,5], [−2.3180767e+042,6,4],
[2.3189526e+042,6,3], [−8.0207995e+041,6,2],
[1.8507606e+041,6,1], [−7.4973474e+039,6,0],
[−5.2760208e+037,5,20], [2.9842084e+038,5,19],
[−5.5649015e+038,5,18], [−7.368612e+037,5,17],
[1.9934532e+039,5,16], [−3.5079662e+039,5,15],
[2.7065741e+039,5,14], [−8.5550695e+038,5,13],
[7.9578629e+038,5,12], [−2.3374901e+039,5,11],
[2.7787882e+039,5,10], [−1.3092327e+039,5,9],
[−3.4153881e+038,5,8], [8.6343954e+038,5,7],
[−5.7280719e+038,5,6], [2.1308735e+038,5,5],
[−4.7610813e+037,5,4], [5.5985893e+036,5,3],
[4.7929313e+034,5,2], [−1.2100609e+035,5,1],
[1.2402605e+034,5,0], [8.6146332e+031,4,21],
[−3.6388015e+032,4,20], [3.333436e+031,4,19],
[2.5621365e+033,4,18], [−6.0683368e+033,4,17],
[4.7351588e+033,4,16], [3.062306e+033,4,15],
[−9.2723333e+033,4,14], [7.1330098e+033,4,13],
[−1.2707034e+033,4,12], [−8.043933e+032,4,11],
[−5.7438818e+032,4,10], [1.1366916e+033,4,9],
[−9.3285215e+031,4,8], [−7.587488e+032,4,7],
[7.1419871e+032,4,6], [−3.3877417e+032,4,5],
[9.7448844e+031,4,4], [−1.7204173e+031,4,3],
[1.6692312e+030,4,2], [−4.9553177e+028,4,1],
[−4.6522425e+027,4,0], [−1.9866517e+026,3,22],
[1.2368348e+027,3,21], [−3.1210446e+027,3,20],
[3.9620807e+027,3,19], [−2.5514036e+027,3,18],
[1.0388144e+027,3,17], [−1.8445009e+027,3,16],
[6.7084785e+027,3,15], [−1.9774072e+028,3,14],
[4.0860257e+028,3,13], [−5.7047765e+028,3,12],
[5.4581226e+028,3,11], [−3.6362018e+028,3,10],
[1.7005739e+028,3,9], [−5.699677e+027,3,8],
[1.5498599e+027,3,7], [−4.7674467e+026,3,6],
[1.7809808e+026,3,5], [−5.5799554e+025,3,4],
[1.1669431e+025,3,3], [−1.4566339e+024,3,2],
[9.1222409e+022,3,1], [−9.7865459e+020,3,0],
[−1.2753215e+020,2,23], [2.4384729e+021,2,22],
[−1.4968714e+022,2,21], [4.5545238e+022,2,20],
[−7.2802281e+022,2,19], [3.4896565e+022,2,18],
[1.031598e+023,2,17], [−2.6460921e+023,2,16],
[3.0693425e+023,2,15], [−1.797041e+023,2,14],
[−1.3113554e+022,2,13], [1.3058726e+023,2,12],
[−1.3064788e+023,2,11], [7.0751978e+022,2,10],

[−1.8563761e+022,2,9], [−3.2896387e+021,2,8],
[5.6333968e+021,2,7], [−2.8387217e+021,2,6],
[8.7814651e+020,2,5], [−1.8348848e+020,2,4],
[2.6087065e+019,2,3], [−2.4354957e+018,2,2],
[1.3652652e+017,2,1], [−3.696187e+015,2,0],
[−1.2884138e+015,1,24], [1.4560902e+016,1,23],
[−7.665514e+016,1,22], [2.4598678e+017,1,21],
[−5.270426e+017,1,20], [7.7095874e+017,1,19],
[−7.3933463e+017,1,18], [3.7941886e+017,1,17],
[2.3563215e+016,1,16], [−1.1917975e+017,1,15],
[−1.0929399e+017,1,14], [3.5797727e+017,1,13],
[−3.8199368e+017,1,12], [2.1742659e+017,1,11],
[−4.2005832e+016,1,10], [−4.1111787e+016,1,9],
[4.6276401e+016,1,8], [−2.5761581e+016,1,7],
[9.6541071e+015,1,6], [−2.5984241e+015,1,5],
[5.0706653e+014,1,4], [−7.0331238e+013,1,3],
[6.5888569e+012,1,2], [−3.7424662e+011,1,1],
[9.7428225e+009,1,0], [86577431,0,25], [−1.3643544e+009,0,24],
[1.0582414e+010,0,23], [−5.1364102e+010,0,22],
[1.6985674e+011,0,21], [−3.9765133e+011,0,20],
[6.65372e+011,0,19], [−7.7386874e+011,0,18],
[5.4980386e+011,0,17], [−7.7188377e+010,0,16],
[−3.0731599e+011,0,15], [3.2674373e+011,0,14],
[−3.048412e+010,0,13], [−2.9639712e+011,0,12],
[4.3152714e+011,0,11], [−3.6720244e+011,0,10],
[2.23716e+011,0,9], [−1.0358591e+011,0,8], [3.7237533e+010,0,7],
[−1.0434499e+010,0,6][2.2617331e+009,0,5],
[−3.7208436e+008,0,4], [44908944,0,3], [−3748279.2,0,2],
[193150.36,0,1], [−4625.8546,0,0]

Coefficients for $\log_{10}\{Q_{ROD}(L_i,R_i)\}$ at $\lambda = 400$ nm

[6.2194908e+123,25,0], [1.0246939e+127,24,1],
[−1.5476224e+120,24,0], [5.3029801e+130,23,2],
[−3.4264609e+123,23,1], [1.5176376e+116,23,0],
[−8.2090908e+133,22,3], [−1.5445347e+127,22,2],
[5.126531e+119,22,1], [−5.615606e+111,22,0],
[1.8147228e+137,21,4], [2.0853884e+130,21,3],
[2.081173e+123,21,2], [−4.4721637e+115,21,1],
[−2.2027218e+107,21,0], [5.4083333e+140,20,5],
[−7.4317651e+133,20,4], [−2.0334348e+126,20,3],
[−1.7363514e+119,20,2], [2.4407981e+111,20,1],
[3.6406581e+103,20,0], [−3.5309476e+143,19,6],
[−1.4001085e+137,19,5], [1.2993372e+130,19,4],
[6.4147725e+121,19,3], [1.0187696e+115,19,2],
[−7.781331e+106,19,1], [−1.941784e+099,19,0],
[−8.6897916e+146,18,7], [1.4719091e+140,18,6],
[1.5168641e+133,18,5], [−1.3100479e+126,18,4],
[5.1263256e+117,18,3], [−4.5521141e+110,18,2],
[5.5975122e+101,18,1], [4.9115424e+094,18,0],
[1.1812509e+150,17,8], [1.5346027e+143,17,7],
[−2.3560993e+136,17,6], [−8.441285e+128,17,5],
[8.5634799e+121,17,4], [−7.1592599e+113,17,3],
[1.6371135e+106,17,2], [8.3968929e+097,17,1],
[1.8154879e+089,17,0], [1.0469184e+153,16,9],
[−3.8351324e+146,16,8], [−5.7296111e+138,16,7],
[2.0051735e+132,16,6], [2.0026831e+124,16,5],
[−3.8302566e+117,16,4], [4.2764645e+109,16,3],
[−4.8751088e+101,16,2], [−5.298464e+093,16,1],
[−6.8311316e+085,16,0], [−3.9734793e+156,15,10],
[8.913005e+147,15,9], [4.7761405e+142,15,8],
[−7.509221e+134,15,7], [−9.941633e+127,15,6],
[3.9863332e+119,15,5], [1.1941228e+113,15,4],
[−1.614181e+105,15,3], [1.1935721e+097,15,2],
[1.8802614e+089,15,1], [3.1124474e+081,15,0],
[−2.0894894e+160,14,11], [2.559138e+153,14,10],
[−8.8405286e+145,14,9], [−1.9193232e+138,14,8],
[8.9987814e+130,14,7], [2.7809221e+123,14,6],
[−4.6844802e+115,14,5], [−2.5546162e+108,14,4],
[4.1695821e+100,14,3], [−2.2945852e+092,14,2],
[−4.7645291e+084,14,1], [−8.748587e+076,14,0],
[1.7387835e+162,13,12], [4.0705561e+156,13,11],
[−4.2051738e+149,13,10], [1.5527142e+142,13,9],
[−1.2441068e+134,13,8], [−3.4245359e+126,13,7],
[−3.2117534e+118,13,6], [1.6680818e+111,13,5],
[3.4479931e+103,13,4], [−7.3887582e+095,13,3],
[3.1206821e+087,13,2], [9.2627485e+079,13,1],
[1.7739873e+072,13,0], [4.9159626e+166,12,13],
[−4.6514806e+159,12,12], [−1.8644075e+152,12,11],
[3.0949186e+145,12,10], [−1.2388464e+138,12,9],
[1.8780354e+130,12,8], [−1.8609804e+121,12,7],
[−2.2935392e+113,12,6], [−3.3036218e+106,12,5],
[−1.8610061e+098,12,4], [8.2720667e+090,12,3],
[−2.0602351e+082,12,2], [−1.4280878e+075,12,1],
[−2.7301097e+067,12,0], [1.5636597e+170,11,14],
[−2.3756872e+163,11,13], [1.4322571e+156,11,12],
[−2.7254506e+148,11,11], [−8.3695736e+140,11,10],
[5.3049839e+133,11,9], [−1.0273315e+126,11,8],
[6.9464746e+117,11,7], [−4.0356061e+108,11,6],
[4.0117416e+101,11,5], [−2.9609602e+093,11,4],
[−3.4540414e+085,11,3], [−2.2781684e+077,11,2],
[1.7770569e+070,11,1], [3.2530757e+062,11,0],
[3.8031394e+173,10,15], [−6.4279916e+166,10,14],
[4.8778109e+159,10,13], [−2.1064206e+152,10,12],
[5.0171333e+144,10,11], [−4.1442248e+136,10,10],
[−9.6900307e+128,10,9], [3.041528e+121,10,8],
[−3.0618215e+113,10,7], [1.0220923e+105,10,6],
[−4.3895398e+096,10,5], [8.5692857e+088,10,4],
[−6.2136962e+080,10,3], [8.7989012e+072,10,2],
[−1.7988655e+065,10,1], [−3.0189906e+057,10,0],
[−6.7562032e+176,9,16], [3.9324903e+169,9,15],
[1.4771894e+162,9,14], [−2.3358456e+155,9,13],
[1.108846e+148,9,12], [−2.8152458e+140,9,11],
[3.7200606e+132,9,10], [−8.0990257e+123,9,9],
[−5.0944441e+116,9,8], [7.4357459e+108,9,7],
[−3.9639963e+100,9,6], [1.0500939e+092,9,5],
[−1.2117584e+084,9,4], [1.4734808e+076,9,3],
[−1.3645589e+068,9,2], [1.4800287e+060,9,1],
[2.1758409e+052,9,0], [2.5500949e+180,8,17],
[−2.3456764e+173,8,16], [9.7392545e+165,8,15],
[−2.9673554e+158,8,14], [9.7792177e+150,8,13],
[−3.131118e+143,8,12], [7.1899209e+135,8,11],
[−9.7702791e+127,8,10], [4.703414e+119,8,9],
[6.980814e+111,8,8], [−1.3618509e+104,8,7],
[9.3606657e+095,8,6], [−3.0857853e+087,8,5],
[1.5073841e+079,8,4], [−1.7020721e+071,8,3],
[1.366781e+063,8,2], [−9.8012527e+054,8,1],
[−1.2055186e+047,8,0], [1.005261e+183,7,18],
[−4.1939638e+176,7,17], [3.9506068e+169,7,16],
[−1.8383246e+162,7,15], [5.2274995e+154,7,14],
[−1.0118554e+147,7,13], [1.4406986e+139,7,12],
[−1.5579258e+131,7,11], [1.0572206e+123,7,10],
[3.1480443e+114,7,9], [−2.0173106e+107,7,8],
[2.6759464e+099,7,7], [−1.8382101e+091,7,6],
[7.0086586e+082,7,5], [−2.1726571e+074,7,4],
[1.415455e+066,7,3], [−9.6845028e+057,7,2],
[5.1310043e+049,7,1], [5.0439048e+041,7,0],
[−9.6695418e+186,6,19], [1.2610056e+180,6,18],
[−6.209486e+172,6,17], [1.199626e+165,6,16],
[1.2189251e+157,6,15], [−1.23214e+150,6,14],

[3.2148279e+142,6,13], [−4.6845536e+134,6,12], [3.8192024e+126,6,11], [−4.5057734e+117,6,10], [−3.5130781e+110,6,9], [5.7532589e+102,6,8], [−5.3032323e+094,6,7], [3.1585425e+086,6,6], [−1.2012248e+078,6,5], [3.159559e+069,6,4], [−1.0389051e+061,6,3], [5.0846397e+052,6,2], [−2.0733119e+044,6,1], [−1.5501985e+036,6,0], [5.0344676e+189,5,20], [−2.8268137e+181,5,19], [−4.603527e+175,5,18], [3.5902838e+168,5,17], [−1.3101017e+161,5,16], [2.6959849e+153,5,15], [−2.9530426e+145,5,14], [5.8888967e+136,5,13], [2.7326092e+129,5,12], [−2.7728359e+121,5,11], [−1.9431105e+113,5,10], [7.6573437e+105,5,9], [−9.7905356e+097,5,8], [7.7082031e+089,5,7], [−4.1000603e+081,5,6], [1.4740535e+073,5,5], [−3.5889718e+064,5,4], [7.4249615e+055,5,3], [−2.1036881e+047,5,2], [6.3126176e+038,5,1], [3.3516109e+030,5,0], [−6.6155867e+191,4,21], [−2.9166738e+185,4,20], [3.2210323e+178,4,19], [−1.1013403e+171,4,18], [2.1237178e+161,4,17], [1.1169568e+156,4,16], [−3.9374833e+148,4,15], [7.121679e+140,4,14], [−7.6507746e+132,4,13], [5.0697303e+124,4,12], [−3.3004624e+116,4,11], [5.3848763e+108,4,10], [−8.9274473e+100,4,9], [9.667674e+092,4,8], [−6.9864619e+084,4,7], [3.4697447e+076,4,6], [−1.1834047e+068,4,5], [2.7315428e+059,4,4], [−4.5129382e+050,4,3], [7.3559324e+041,4,2], [−1.4275314e+033,4,1], [−4.729692e+024,4,0], [−7.6780977e+195,3,22], [1.1751181e+189,3,21], [−7.2371234e+181,3,20], [2.2135472e+174,3,19], [−2.5665222e+166,3,18], [−5.5950062e+158,3,17], [3.1187139e+151,3,16], [−7.3550592e+143,3,15], [1.1897284e+136,3,14], [−1.4999833e+128,3,13], [1.5159622e+120,3,12], [−1.1457808e+112,3,11], [4.9478093e+103,3,10], [1.023715e+095,3,9], [−3.6414142e+087,3,8], [3.1340161e+079,3,7], [−1.6092594e+071,3,6], [5.4421035e+062,3,5], [−1.2227802e+054,3,4], [1.8331319e+045,3,3], [−2.0813988e+036,3,2], [2.3857721e+027,3,1], [3.6685219e+018,3,0], [1.5827018e+199,2,23], [−2.1849373e+192,2,22], [1.3159831e+185,2,21], [−4.4261539e+177,2,20], [8.4742339e+169,2,19], [−6.3405648e+161,2,18], [−1.1491367e+154,2,17], [4.1695184e+146,2,16], [−5.9594365e+138,2,15], [3.9072608e+130,2,14], [1.5396884e+122,2,13], [−6.6777361e+114,2,12], [8.0869557e+106,2,11], [−5.7549788e+098,2,10], [2.3485219e+090,2,9], [−1.6963874e+081,2,8], [−4.4690073e+073,2,7], [3.1787624e+065,2,6], [−1.1737864e+057,2,5], [2.6855158e+048,2,4], [−3.8999507e+039,2,3], [3.723306e+030,2,2], [−2.7519062e+021,2,1], [−5.1518788e+011,2,0], [−2.4567022e+202,1,24], [3.5649293e+195,1,23], [−2.3242433e+188,1,22], [8.876351e+180,1,21], [−2.1440845e+173,1,20], [3.1724082e+165,1,19], [−1.9879667e+157,1,18], [−2.5480392e+149,1,17], [8.0313894e+141,1,16], [−9.4990484e+133,1,15], [4.5754335e+125,1,14], [3.4467646e+117,1,13], [−8.9591231e+109,1,12], [9.4715316e+101,1,11], [−6.694984e+093,1,10], [3.516689e+085,1,9], [−1.4368875e+077,1,8], [4.685726e+068,1,7], [−1.2307613e+060,1,6], [2.565548e+051,1,5], [−4.0635848e+042,1,4], [4.5841114e+033,1,3], [−3.4735732e+024,1,2], [1.7838465e+015,1,1], [−1606147.2,1,0], [−3.1603274e+205,0,25], [4.9032112e+198,0,24], [−3.4548419e+191,0,23], [1.4523456e+184,0,22], [−4.0123348e+176,0,21], [7.5370022e+168,0,20], [−9.6506659e+160,0,19], [8.8095292e+152,0,18], [−9.2672009e+144,0,17], [2.010042e+137,0,16], [−4.4638467e+129,0,15], [7.212214e+121,0,14], [−8.5110815e+113,0,13], [7.5856919e+105,0,12], [−5.2205813e+097,0,11], [2.8056421e+089,0,10], [−1.1810805e+081,0,9], [3.8812668e+072,0,8], [−9.8631626e+063,0,7], [1.9077199e+055,0,6], [−2.7426121e+046,0,5], [2.8316055e+037,0,4], [−1.9969537e+028,0,3], [8.9277722e+018,0,2], [−2.254163e+009,0,1], [0.25964777,0,0]

Appendix B. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ab.2015.11.021.

References

[1] M. Stoppini, V. Bellotti, Systemic amyloidosis: lessons from β2-microglobulin, J. Biol. Chem. 290 (2015) 9951–9958.
[2] L.C. Walker, M. Jucker, Neurodegenerative diseases: expanding the prion concept, Annu. Rev. Neurosci. 38 (2015) 87–103.
[3] D. Hall, H. Edskes, Computational modeling of the relationship between amyloid and disease, Biophys. Rev. 4 (2012) 205–222.
[4] R.B. Wickner, H.K. Edskes, F. Shewmaker, T. Nakayashiki, Prions of fungi: inherited structures and biological roles, Nat. Rev. Microbiol. 5 (2007) 611–618.
[5] G. Invernizzi, E. Papaleo, R. Sabate, S. Ventura, Protein aggregation: mechanisms and functional consequences, Int. J. Biochem. Cell Biol. 44 (2012) 1541–1554.
[6] R. Mezzenga, P. Fischer, The self-assembly, aggregation and phase transitions of food protein systems in one, two, and three dimensions, Rep. Prog. Phys. 76 (2013) 046601.
[7] S.K. Singh, N. Afonina, M. Awwad, K. Bechtold-Peters, J.T. Blue, D. Chou, M. Cromwell, H.J. Krause, H.C. Mahler, B.K. Meyer, L. Narhi, D.P. Nesta, T. Spitznagel, An industry perspective on the monitoring of subvisible particles as a quality attribute for protein therapeutics, J. Pharm. Sci. 99 (2010) 3302–3321.
[8] P. Doty, R.F. Steiner, Light scattering and spectrophotometry of colloidal solutions, J. Chem. Phys. 18 (1950) 1211–1220.
[9] C.F. Bohren, D.R. Huffman, Absorption and Scattering of Light by Small Particles, John Wiley, New York, 2008.
[10] J.M. Andreu, S.N. Timasheff, The measurement of cooperative protein self-assembly by turbidity and other techniques, Methods Enzymol. 130 (1986) 47–59.
[11] H.C. Mahler, W. Friess, U. Grauschopf, S. Kiese, Protein aggregation: pathways, induction factors and analysis, J. Pharm. Sci. 98 (2009) 2909–2934.
[12] R.D. Camerini-Otero, L.A. Day, The wavelength dependence of the turbidity of solutions of macromolecules, Biopolymers 17 (1978) 2241–2249.
[13] F.H. Silver, D.E. Birk, Kinetic analysis of collagen fibrillogenesis: I. Use of turbidity–time data, Collagen Relat. Res. 3 (1983) 393–405.
[14] D. Hall, A.P. Minton, Turbidity as a probe of tubulin polymerization kinetics: a theoretical and experimental re-examination, Anal. Biochem. 345 (2005) 198–213.
[15] T.P. Moody, M.A. Donovan, T.M. Laue, Turbidimetric studies of *Limulus* coagulin gel formation, Biophys. J. 71 (1996) 2012–2021.
[16] A.C. Garcia-Lopez, A.D. Snider, L.H. Garcia-Rubio, Rayleigh–Debye–Gans as a model for continuous monitoring of biological particles: I. Assessment of theoretical limits and approximations, Opt. Express 14 (2006) 8849–8865.
[17] A.C. Garcia-Lopez, L.H. Garcia-Rubio, Rayleigh–Debye–Gans as a model for continuous monitoring of biological particles: II. Development of a hybrid model, Opt. Express 16 (2008) 4671–4687.
[18] L.B. Korolevskaya, N.G. Khlebtsov, Spectroturbidimetric determination of the sizes of poly(ethylene glycol)-induced insoluble immune complex particles, Colloid J. 72 (2010) 504–511.
[19] A. Garcia-Lopez, Hybrid Model for Characterization of Submicron Particles Using Multiwavelength Spectroscopy (PhD thesis), University of South Florida, 2005.
[20] M. Elimlech, J. Gregory, X. Jia, R. Williams, Particle Deposition & Aggregation, Chap. 9, Butterworth Heinemann, Oxford, UK, 1995.
[21] R. Khurana, J.R. Gillespie, A. Talapatra, L.J. Minert, C. Ionescu-Zanetti, I. Millett, A.L. Fink, Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates, Biochemistry 40 (2001) 3525–3535.
[22] Y. Yoshimura, Y. Lin, H. Yagi, Y.H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki, Y. Goto, Distinguishing crystal-like amyloid fibrils and glass-like

amorphous aggregates from their kinetics of formation, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 14446–14451.
[23] E.Y. Chi, S. Krishnan, T.W. Randolph, J.F. Carpenter, Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation, Pharm. Res. 20 (2003) 1325–1336.
[24] C.H. Bennett, Serially deposited amorphous aggregates of hard spheres, J. Appl. Phys. 43 (1972) 2727–2734.
[25] S.D. Durbin, G. Feher, Protein crystallization, Annu. Rev. Phys. Chem. 47 (1996) 171–204.
[26] J.P. Sethna, Order parameters, broken symmetry, and topology 1992, arXiv cond-mat/9204009.
[27] D. Hall, Semi-automated methods for simulation and measurement of amyloid fiber distributions obtained from transmission electron microscopy experiments, Anal. Biochem. 421 (2012) 262–277.
[28] D. Hall, A.P. Minton, Effects of inert volume-excluding macromolecules on protein fiber formation: I. Equilibrium models, Biophys. Chem. 98 (2002) 93–104.
[29] D. Hall, A.P. Minton, Effects of inert volume-excluding macromolecules on protein fiber formation: II. Kinetic models for nucleated fiber growth, Biophys. Chem. 107 (2004) 299–316.
[30] O. Kratky, H. Leopold, H. Stabinger, The determination of the partial specific volume of proteins by the mechanical oscillator technique, Methods Enzymol. 27 (1973) 98–110.
[31] D. Hall, J. Kardos, H. Edskes, J.A. Carver, Y. Goto, A multi-pathway perspective on protein aggregation: implications for control of the rate and extent of amyloid formation, FEBS Lett. 589 (2015) 672–679.
[32] A. Shahravan, C. Lucas, T. Matsoukas, Nanowire charging in collisionless plasma, J. Appl. Phys. 108 (2010) 083303.
[33] K.E. van Holde, W. Cutiss Johnson, P. Shing Ho, Principles of Physical Biochemistry, Chap. 7, Prentice Hall, 2015.
[34] L. Rayleigh, On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky, London, Edinb. Dublin Philos. Mag. J. Sci. 47 (1899) 375–384.
[35] G.E. Perlmann, L.G. Longsworth, The specific refractive increment of some purified proteins,, J. Am. Chem. Soc. 70 (1948) 2719–2724.
[36] G. Oster, Light scattering, in: G. Oster, A.W. Pollister (Eds.), Physical Techniques in Biological Research, Vol. 1: Optical TechniquesAcademic Press, San Diego, 2013 chap. 2.
[37] M. Kerker, The Scattering of Light and Other Electromagnetic Radiation, Physical Chemistry: a Series of Monographs, vol. 16, Academic Press, San Diego, 2013 chap. 7.
[38] M.L. Wallach, W. Heller, A.F. Stevenson, Theoretical investigations on the light scattering of colloidal spheres: XII. The determination of size distribution curves from turbidity spectra, J. Chem. Phys. 34 (1961) 1796–1802.
[39] P.J. Wyatt, Measurement of special nanoparticle structures by light scattering, Anal. Chem. 86 (2014) 7171–7183.
[40] W. Hergert, T. Wriedt (Eds.), The Mie Theory: Basics and Applications, vol. 169, Springer, 2012 chap. 2.
[41] E.P. Geiduschek, A. Holtzer, Application of light scattering to biological systems: deoxyribonucleic acid and the muscle proteins, in: C.A. Tobias, J.H. Lawrence (Eds.), Advances in Biological and Medical Physics, Academic Press, 1958, pp. 431–551.
[42] N. Mikati, J. Nordh, B. Norden, Scattering anisotropy of partially oriented samples: turbidity flow linear dichroism (conservative dichroism) of rod-shaped macromolecules, J. Phys. Chem. 91 (1987) 6048–6055.
[43] F. Gaskin, C.R. Cantor, M.L. Shelanski, Turbidimetric studies of the in vitro assembly and disassembly of porcine neurotubules, J. Mol. Biol. 89 (1974) 737–758.
[44] B.J. Berne, Interpretation of the light scattering from long rods, J. Mol. Biol. 89 (1974) 755–758 (Appendix 1).
[45] M. Adachi, M. So, K. Sakurai, J. Kardos, Y. Goto, Supersaturation-limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation, J. Biol. Chem. 290 (2015) 18134–18145.
[46] M.W. Head, J.W. Ironside, Creutzfeldt–Jakob disease: prion protein type, disease phenotype, and agent strain [review], Neuropathol. Appl. Neurobiol. 38 (2012) 296–310.
[47] D. Tilley, I. Levit, J.A. Samis, Development of a microplate coagulation assay for Factor V in human plasma, Thrombosis J. 9 (2011) 1–11.
[48] K. Moffat, Q.H. Gibson, The rates of polymerization and depolymerization of sickle cell haemoglobin, Biochem. Biophys. Res. Commun. 61 (1974) 237–242.
[49] D. Hall, The effects of tubulin denaturation on the characterization of its polymerization behaviour, Biophys. Chem. 104 (2003) 655–682.
[50] W.A. Voter, H.P. Erickson, The kinetics of microtubule assembly: evidence for a two-stage nucleation mechanism, J. Biol. Chem. 259 (1984) 10430–10438.
[51] D.C. Thorn, S. Meehan, M. Sunde, A. Rekas, S.L. Gras, C.E. MacPhee, J.A. Carver, Amyloid fibril formation by bovine milk κ-casein and its inhibition by the molecular chaperones αS- and β-casein, Biochemistry 44 (2005) 17027–17036.
[52] K.N. Pearce, J.E. Kinsella, Emulsifying properties of proteins: evaluation of a turbidimetric technique, J. Agric. Food Chem. 26 (1978) 716–723.
[53] F. Salmannejad, N. Nafissi-Varcheh, A. Shafaati, R. Aboofazeli, Study on the effect of solution conditions on heat-induced aggregation of human alpha interferon, Iran. J. Pharm. Res. 13 (2014) 27–34.
[54] H. Wu, M. Lattuada, M. Morbidelli, Dependence of fractal dimension of DLCA clusters on size of primary particles, Adv. Colloid Interf. Sci. 195 (2013) 41–49.
[55] L. Nicoud, M. Owczarz, P. Arosio, M. Morbidelli, A multiscale view of therapeutic protein aggregation: a colloid science perspective, Biotechnol. J. 10 (2015) 367–378.
[56] A.P. Minton, Static light scattering from concentrated protein solutions: I. General theory for protein mixtures and application to self-associating proteins, Biophys. J. 93 (2007) 1321–1328.
[57] J.C. Thomas, The determination of log normal particle size distributions by dynamic light scattering, J. Colloid Interface Sci. 117 (1987) 187–192.
[58] G. Bryant, T. Mortensen, S. Henderson, S. Williams, W. van Megen, Optical contrast variation study of nonaqueous suspensions of polymer particles, J. Colloid Interface Sci. 216 (1999) 401–408.
[59] J.K.G. Dhont, Multiple Rayleigh–Gans–Debye scattering in colloidal systems-general theory and static light scattering, Phys. A Stat. Mech. Appl. 120 (1983) 238–262.
[60] D. Kashchiev, Kinetics of protein fibrillation controlled by fibril elongation, Proteins Struct. Funct. Bioinform. 82 (2014) 2229–2239.
[61] D. Hall, L. Huang, On the use of size exclusion chromatography for the resolution of mixed amyloid aggregate distributions: I. Equilibrium partition models, Anal. Biochem. 426 (2012) 69–85.

View publication stats

Paper 2: Recognizing and analyzing variability in amyloid formation kinetics: Simulation and statistical methods

The second of the three papers published during my postgraduate studies at the Australian National University Research School of Chemistry (Hall et al. 2016b) was concerned with how to assess differences between assays of chaperone action conducted on protein aggregation reactions. In this paper, together with Dr. Hall, I helped to develop a method for quantifying the variability in the amyloid aggregation assay. We investigated the variability in the amyloid aggregation kinetics, and developed methods for its simulation, identification and analysis. Initially I searched for an already developed method in the literature. However, rather unexpectedly, such an analysis had not been previously developed despite it being the fundamental cornerstone of all differential analyses of drug and condition effects upon the protein aggregation reaction.

Analytical Biochemistry 510 (2016) 56–71

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Recognizing and analyzing variability in amyloid formation kinetics: Simulation and statistical methods

Damien Hall [a, c, *], Ran Zhao [a], Masatomo So [c], Masayuki Adachi [c], Germán Rivas [b], John A. Carver [a], Yuji Goto [c]

[a] Research School of Chemistry, Australian National University, Acton ACT 2601, Australia
[b] Centro de Investigaciones Biológicas, CSIC, 28006 Madrid, Spain
[c] Institute for Protein Research, Osaka University, 3-1- Yamada-oka, Suita, Osaka 565-0871 Japan

ARTICLE INFO

Article history:
Received 4 May 2016
Received in revised form
11 July 2016
Accepted 13 July 2016
Available online 16 July 2016

Keywords:
Amyloid
Aggregation assay
Statistical significance
Anti-amyloid drug screening

ABSTRACT

We examine the phenomenon of variability in the kinetics of amyloid formation and detail methods for its simulation, identification and analysis. Simulated data, reflecting intrinsic variability, were produced using rate constants, randomly sampled from a pre-defined distribution, as parameters in an irreversible nucleation-growth kinetic model. Simulated kinetic traces were reduced in complexity through description in terms of three characteristic parameters. Practical methods for assessing convergence of the reduced parameter distributions were introduced and a bootstrap procedure was applied to determine convergence for different levels of intrinsic variation. Statistical methods for assessing the significance of shifts in parameter distributions, relating to either change in parameter mean or distribution shape, were tested. Robust methods for analyzing and interpreting kinetic data possessing significant intrinsic variance will allow greater scrutiny of the effects of anti-amyloid compounds in drug trials.
Crown Copyright © 2016 Published by Elsevier Inc. All rights reserved.

1. Introduction

The proteinaceous fibrous polymer known as amyloid is much studied due to its potentially causal association with a number of fatal amyloidosis [1,2]. Of particular interest to the medical research community has been the question of what chemical and environmental factors cause normally soluble protein to convert to the insoluble-polymeric amyloid form [3,4]. The vast majority of prior *in vitro* based investigations directed towards this topic have taken the form of differential kinetic measurements of control and perturbed sample groups, i.e. the kinetics are simultaneously measured in the absence and presence of the component/condition of interest, and the relative change in kinetics, rather than the absolute change, reported [5–7]. In these studies, the time-dependent change of a suitable experimental measure, such as thioflavin T dye binding [8–10] or turbidity/light scattering [11,12], is followed as a proxy marker of fibril formation.

A great variety of amyloid-forming model systems[1] are known in the literature [13], with some directly related to amyloidosis [14–17], whilst others are studied for their biological significance [18,19] or potential use as bio/nanotechnology agents [20]. Some amyloid-forming systems yield kinetic data that are relatively robust and reproducible [21,22] whilst others seem to lack the property of reproducibility [16,17,23,24], possibly because of a strong sensitivity of the reaction rate to initial conditions [25,26] or the pre-existing potential for diversity of competing reaction pathways [27–29]. Although a comprehensive understanding as to why some systems are more difficult to work with than others has not yet been obtained, such sensitivity may be empirically confronted, and characterized, using a suitable application of statistics. In this paper, we explore how to deal with intrinsic variability in amyloid formation kinetics produced by either inherent stochastic

Abbreviations: SD, Standard Deviation; SE, Standard Error; GLV, Gaussian Low Variance; FHV, Flat High Variance; dof, degrees of freedom; KS2D, Kolmogorov Smirnov Two Sample D Test.

* Corresponding author. Research School of Chemistry, Australian National University. Acton ACT 2601, Australia.
E-mail addresses: damien.hall@anu.edu.au, damien.hall@protein.osaka-u.ac.jp, damienhall30@gmail.com (D. Hall).
URL: http://chemistry.anu.edu.au/research/groups/physical-biochemistry-disease

[1] By amyloid system we are referring to the protein, the solution constituents and the physical parameters.

http://dx.doi.org/10.1016/j.ab.2016.07.013
0003-2697/Crown Copyright © 2016 Published by Elsevier Inc. All rights reserved.

variation [30] or intrinsic non-linear sensitivity to initial conditions [25,26]. The basic methodology of our approach involves the following three steps,

(i) Simulate kinetic data of amyloid formation that exhibit defined extents of intrinsic variation.
(ii) Reduce the primary kinetic data to a set of characteristic parameters.
(iii) Analyze the parameter distributions to examine when/how statistically significant conclusions may/or may not be extracted.

This work represents the first in a two-part series aimed at developing reliable methods for analyzing amyloid kinetics. Part II will focus on the experimental application of the concepts outlined in this paper.

2. Theory and results

The simulation aspect of this work required the generation of synthetic amyloid kinetic traces, concordant with an irreversible nucleated-growth scheme, that featured a defined extent of intrinsic variation (Eqn. (1), Fig. 1A and B). Variation, designated by ξ_1 and ξ_2, was incorporated into the rate constants k_1 and k_2, respectively governing the nucleation (Eqn. (1a)) and growth (Eqn. (1b)) stages.

$$nM \xrightarrow{k_1 \pm \xi_1} N \tag{1a}$$

$$\begin{bmatrix} N+M & \xrightarrow{k_2 \pm \xi_2} & A_{n+1} \\ A_{n+1}+M & \xrightarrow{k_2 \pm \xi_2} & A_{n+2} \\ \cdot & \cdot & \cdot \\ A_{z-1}+M & \xrightarrow{k_2 \pm \xi_2} & A_z \end{bmatrix} \tag{1b}$$

To simulate differing extents of intrinsic variation the rate constants, $k_1 \pm \xi_1$ and $k_2 \pm \xi_2$, were randomly selected from one of two general types of distribution, either a Gaussian distribution with low variance (GLV), or a flat distribution with high variance (FHV) (Fig. 1C). Using this approach, large data sets exhibiting a known level of intrinsic variation, were generated for each of the four possible pairings of rate constant distribution types, through repeated simulation of the differential equation set pertaining to Eqn. (1) (Appendix 1). Once simulated, the primary kinetic data were then reduced using a model-free parameterization strategy based on a characteristic point analysis (Fig. 1D). The decomposition of the primary data was carried out using three empirical parameters, $A_{t\to\infty}$, the maximal fraction of monomer present as amyloid, t_{10}, the tenth time, i.e. the time to reach one tenth of the asymptotic value, and $t_{50}-t_{10}$, the difference between the half-time and the tenth-time.

Four general cases of amyloid kinetics reflecting intrinsic variance were simulated from the two types of distribution for each rate constant.

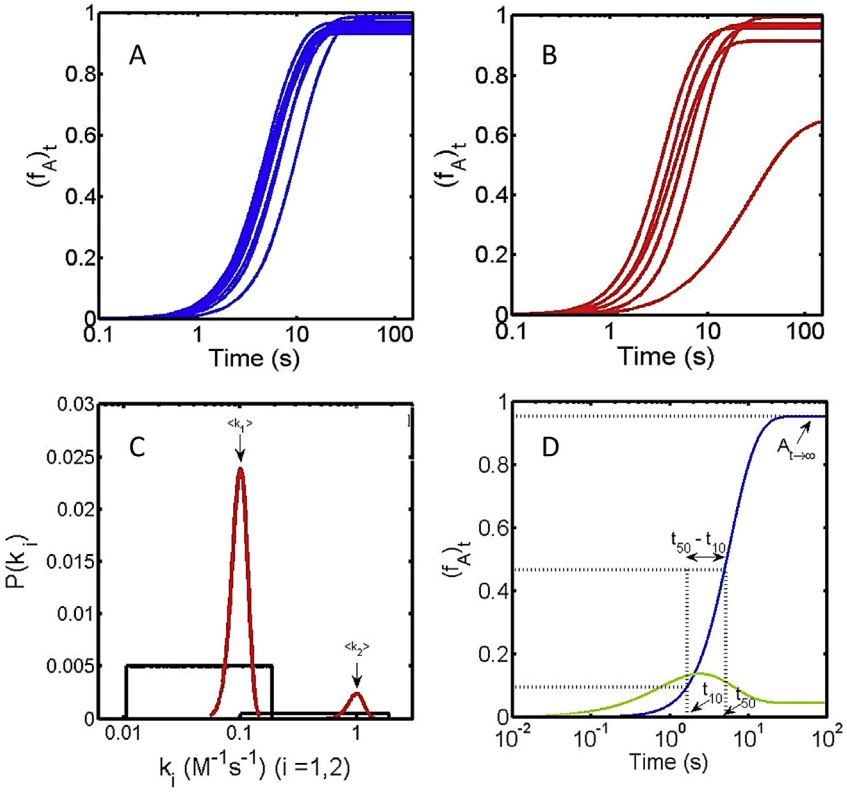


Fig. 1. Simulated replicate kinetic data sets showing the fraction of amyloid formed as a function of time. Simulations were carried out using Eqns. A1 and A2 whereby the input rate constants are sampled from distributions affording (A) Low variation (blue) (B) High variation (red). (C) Probability distribution describing two disparate types of distribution for $k_1 \pm \xi_1$ and $k_2 \pm \xi_2$ along with their mean values (indicated by the arrows) $\langle k_1 \rangle = 0.1\ M^{-1}s^{-1}$, $\langle k_2 \rangle = 1.0\ M^{-1}s^{-1}$. The two distribution types were a Gaussian Low Variance (GLV) $\sigma_k = 0.15\langle k \rangle$, $k \in [\langle k \rangle - 6\sigma_k, \langle k \rangle + 6\sigma_k,]$ or a Flat High Variance (FHV) distribution, $k \in [0.1\langle k \rangle, 1.9\langle k \rangle]$ (D) The three characteristic points measured on the amyloid growth curve used in the data reduction strategy, t_{10}, $t_{50}-t_{10}$ and $A_{t\to\infty}$. The blue line describes the fraction of amyloid formed as a function of time. The green line describes the time course of nucleus formation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(i) Low variance Gaussian k_1 and low variance Gaussian k_2 - [k_1(GLV), k_2(GLV)]
(ii) High variance, flat distribution k_1 and low variance Gaussian k_2 - [k_1(FHV), k_2(GLV)]
(iii) Low variance Gaussian k_1 and high variance flat distribution k_2 - [k_1(GLV), k_2(FHV)]
(iv) High variance flat distribution k_1 and high variance flat distribution k_2 - [k_1(FHV), k_2(FHV)].

For each of the above cases, one-thousand simulations were generated by appropriate random sampling of the rate constant distributions [31] (Fig. 2, column 1). Frequency plots for each of the three characteristic empirical parameters were generated by analysis of all one-thousand simulations (Fig. 2, columns 2–4) for each of the four levels of intrinsic variation (rows 1–4). As might be expected, the well behaved system (Fig. 2 Row 1) yields reduced parameter distributions reflecting the intrinsically low input variance that characterizes its rate constants. Similarly predictable, the poorly defined distribution pairing of k_1 and k_2 produces characteristic parameter distributions with the greatest spread (Fig. 2, Row

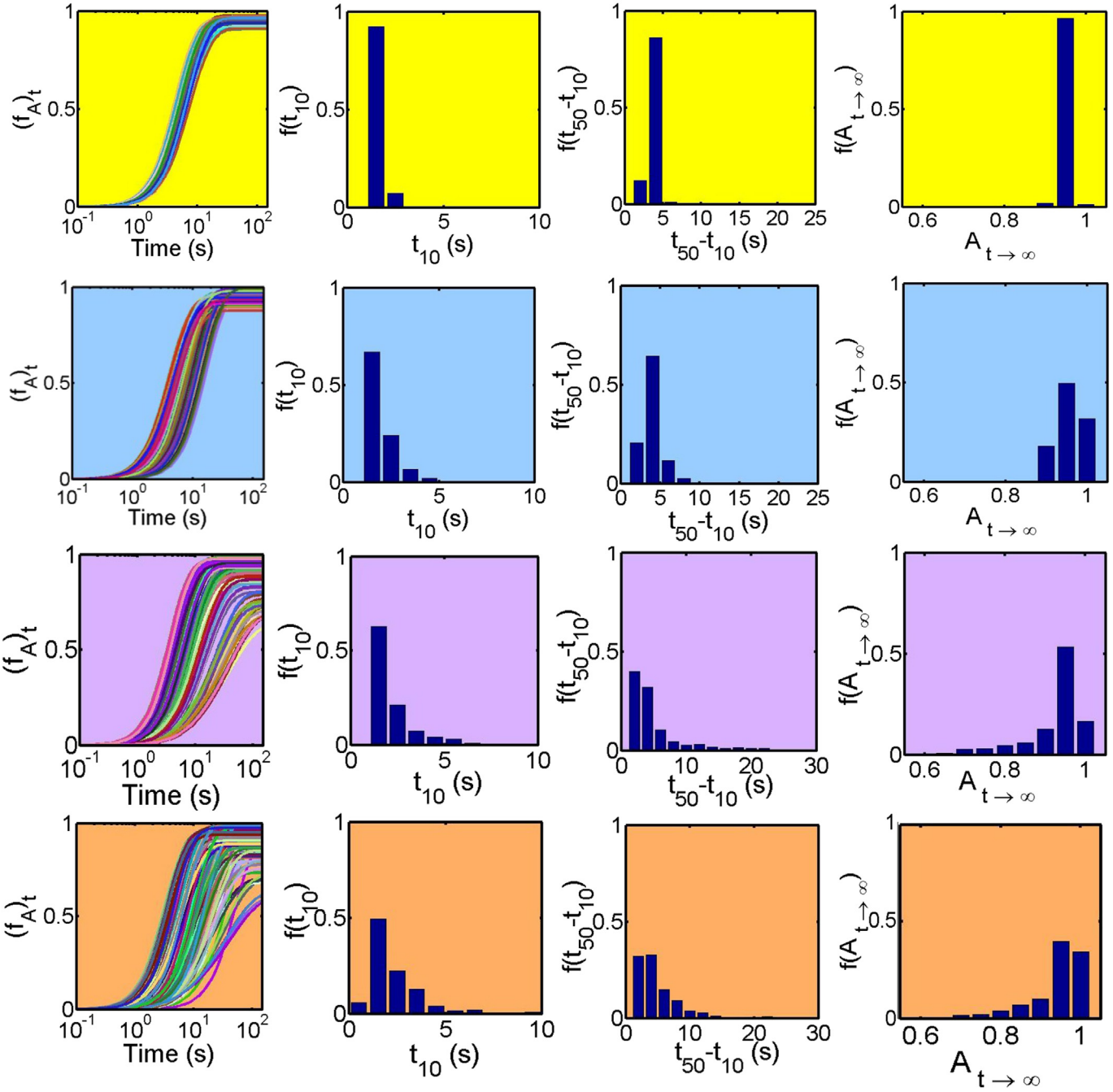


Fig. 2. Raw simulations and histogram analysis of characteristic points, t_{10}, $t_{50}-t_{10}$ and $A_{t\to\infty}$, for all four possible combinations of the two types of distribution for the nucleation (k_1) and growth (k_2) rate constants. **Row 1**: Simulations reflecting [k_1(GLV), k_2(GLV)] combination (yellow background), **Row 2**: [k_1(FHV), k_2(GLV)] (light blue background), **Row 3**: [k_1(GLV), k_2(FHV)] (magenta background), **Row 4**: [k_1(FHV), k_2(FHV)] (orange background). **Column 1**: Limited set of 200 overlaid simulations of fractional amyloid formation vs. time. **Column 2**: Frequency histogram of characteristic t_{10} times. **Column 3**: Frequency histogram of characteristic $t_{50}-t_{10}$ times. **Column 4**: Frequency histogram of characteristic asymptotic levels of amyloid. All histograms were generated from 1000 simulations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4). More interesting, is the comparative failure to see a much tighter spread in the distribution of the t_{10} parameter for the case reflecting a low-variance Gaussian distribution in k_1 and a high-variance flat distribution in k_2 (Fig. 3 Row 3 Column 2). This finding indicates that t_{10} may not necessarily be a good indicator of the magnitude of k_1 and that this experimentally derived parameter may be significantly determined[2] by, or correlated with, the actual value of k_2.

2.1. A first look at determining parameter convergence

The two most common formats for replicating amyloid kinetic experiments involve either the sequential replication of an assay in a single cuvette [32] or the concerted performance of assays using a multi-well microplate procedure [22–24,33]. In either situation, the question arises as to whether the number of replicate experiments is sufficiently large to capture the intrinsic variation in the population. As the only tool available to the experimenter to answer this question is to simply keep repeating the kinetic assay until a suitable test signifies parameter convergence, a suitable numerical indicator of parameter distribution convergence is required [34]. Here we examine three potentially useful measures of parameter convergence (for a generalized parameter β with total replicate number N) (Eqn. (2)). The three indicators include the mean plus or minus a standard error of the mean, $<\beta> \pm SE_\beta$ (Eqn. (2a)), a reduced standard deviation, σ_β (Eqn. (2b)), and a higher moment term, γ_β, reflecting a skew-like description of the distribution (Eqn. (2c)).

$$\langle\beta\rangle \pm \sqrt{\sum_{i=1}^{N} (\beta_i - \langle\beta\rangle)^2 \Big/ (N[N-1])} \tag{2a}$$

$$\sigma_\beta = \sqrt{\sum_{i=1}^{N} (\beta_i - \langle\beta\rangle)^2 \Big/ (N-1)} \tag{2b}$$

$$\gamma_\beta = \frac{\sum_{i=1}^{N} (\beta_i - \langle\beta\rangle)^3}{(N-1)\cdot\sigma_\beta^3} \tag{2c}$$

All three descriptors provide useful information on the convergence, although for limited sampling, the higher moment terms (reflecting the standard deviation and the skew of the distribution) can be seen to be more sensitive indicators. The skew-like term has the added benefit of showing convergence of the distribution's asymmetry due to the fact that its positive or negative sign respectively relates to the existence of a right or left-handed distribution tail [34]. Fig. 3 describes a particular convergence pathway of 96 simulated replicates[3] for the four cases of intrinsic variability, generated from the combinations of k_1 and k_2. From this analysis we note that for systems demonstrating greater intrinsic variance the reduced parameter distributions may not have sufficiently converged even after 96 replicates have been performed.[4]

2.2. A closer look at identifying and analyzing variation

Such a single-shot analysis of convergence, as performed in the previous section, suffers the consequences of random sampling, i.e. each convergence series will follow a unique pathway, thus preventing definite conclusions being drawn from a single trace. To escape this situation, a bootstrap procedure was implemented [35]. Using this method, a fixed number of replicates was randomly sampled 200 times each from the starting pool of 1000 simulations. Adapting this approach, we examined the statistical evolution of each of the three markers of convergence outlined in Eqn. (2), for the four different cases of intrinsic variance (Fig. 4). In this manner the bootstrap procedure affords estimation of the likely mean and standard deviations for the three convergence markers at different replicate numbers.

2.3. Assessing significance for different scenarios

The majority of medical research into amyloid kinetics is based on the concept of drug screening, i.e. differential evaluation of the effect of a ligand or a change in solution conditions on fibril formation [3–8,18,22]. In such cases two groups of data, the test and control sets, will be compared and an assessment made as to the significance of any observed change. In discussing the potential change in a parameter, it is important to be clear as to what the potential outcomes for change are. Fig. 5 describes the classes of change that may be affected within the control distribution, corresponding to changes in the mean, variance and distribution shape. In this section, we examine methods for determination of the presence or absence of significance when comparing data sets featuring possible variation in parameter distributions of the type described in Fig. 5. Two general types of comparative case are examined.

Type 1 Comparison: Test and control sets are drawn from systems defined by the same fundamental parameter distributions (therefore possessing the same levels of intrinsic variation).

Type 2 Comparison: Test and control sets are drawn from different parameter distributions (therefore possessing different extents of intrinsic variation).

Three statistical methods, having varying degrees of (statistical) power [34,36,37], were tested in the current work. For each pairing of test replicates and control replicates, each statistical procedure was independently carried out two-hundred times with the end result reported as fractional significance.[5] In what follows, we first provide a short description of the basis of each of these statistical tests and then present the results of their application to the two general types of comparative case listed above.

(i) Comparison of averages with standard error

A large majority of published studies simply overlay the data or alternatively assert significance based on comparison of the

[2] This is a potentially important result as some have interpreted the presence of a correlation between experimentally derived markers of growth and nucleation rates as indicating that these two processes share a fundamentally similar dependence upon reaction conditions [17]. This conclusion can be seen to be wholly derived from the lack of an orthogonal measure of the nucleation rate, i.e. a strong correlation between experimental markers of nucleation and growth will always be found when basing the nucleation parameterization on a signal related to fiber growth, even in the absence of such a correlation at the molecular rate constant level.

[3] As might be generated using a microplate reader [22,24].

[4] The potential exists for fitting an empirical function to the marker of convergence and subsequently using this empirical function to estimate the total replicates required for distribution convergence.

[5] In the comparison of parameters derived from a set of m group 2 test replicates against a set of n group 1 control replicates fractional significance is defined as the fraction of times (out of a total of 200 trials) that test sets were judged to exhibit a less than 5% critical probability of supporting the null hypothesis, i.e. that they were no different from the control

$$\text{Fractional Significance} = \frac{1}{200}\sum_{i=1}^{200} j$$

$j = 1$ if $P_i \le P_{critical}(0.05)$

$j = 0$ if $P_i > P_{critical}(0.05)$

.

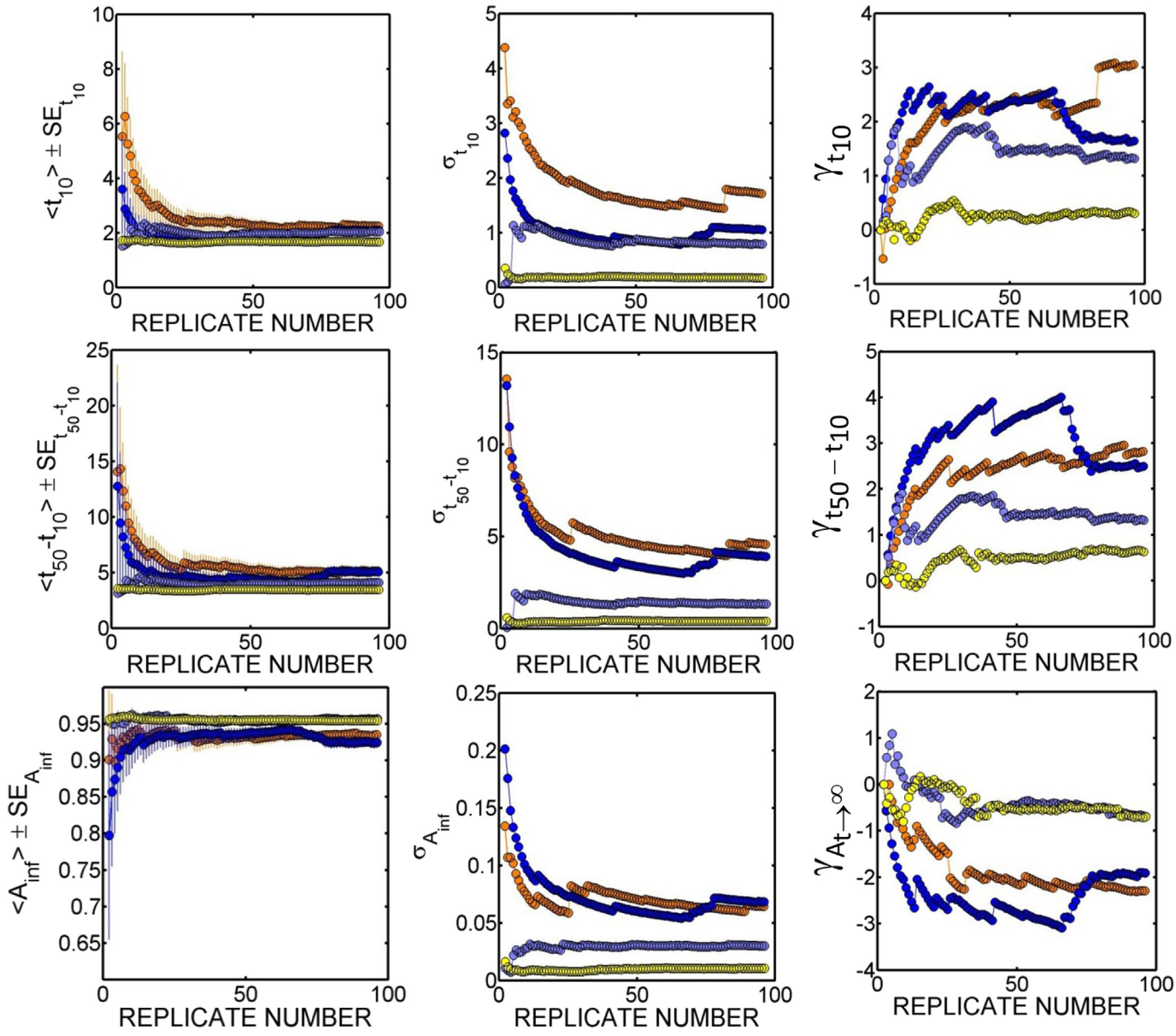


Fig. 3. A single replicate series of convergence indicators corresponding to the distributions of t_{10} (Row 1), $t_{50}-t_{10}$ (Row 2) and $A_{t\to\infty}$ (Row 3). Convergence was monitored using (Column 1) Simple average (< >) plus or minus the standard error (SE), (Column 2) Standard deviation (σ), (Column 3) Skew-like parameter (γ). The four line colours follow the same representation as per Fig. 2 i.e. yellow = [k_1(GLV), k_2(GLV)], light blue = [k_1(FHV), k_2(GLV)], magenta = [k_1(GLV), k_2(FHV)], orange = [k_1(FHV), k_2(FHV)]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parameter mean plus or minus two times its standard error.[6] For the purposes of this article we will call this the <β> ± 2S.E. test (Fig. 6a). With certain restrictions,[7] this approach is valid for assessing how likely it may be that a single value is part of a distribution, but is less correct when comparing two distributions as it does not properly account for distribution overlap [34]. Fig. 7 describes the application of the <β> ± 2S.E. test to a type 1 comparison of data where each row corresponds to one of the four different cases of intrinsic variance tested against itself (as both sample and control). The <β> ± 2S.E. test does a fair job of not falsely asserting significance although some degree of pseudo-significance is seen in the off-diagonal low replicate regions, especially as the variance within the sample increases. Fig. 8 shows the application of the <β> ± 2S.E. test to the type 2 comparison in which the first three cases of amyloid kinetic data, featuring increasing levels of variation, are tested as samples against the control data set possessing the most intrinsic variance. We see that the <β> ± 2S.E. test is capable of identifying statistically meaningful differences in the three characteristic parameters, t_{10}, $t_{50}-t_{10}$ and $A_{t\to\infty}$, between the lowest intrinsic variance data set and the highest intrinsic variance control case (Row 1 of Fig. 8). However relatively large replicate numbers of the low variance sample case are required to identify a significant difference (i.e. the test is asymmetric with regards to replicate number of control and sample groups). The simple <β> ± 2S.E. test

[6] More often the error envelope is simply approximated as ± 1 S.E. (e.g. [21]) but we persist here with ~2 S.E. to maintain a ~95% confidence estimate for the two-tailed test [34].

[7] More properly, the confidence interval should be assessed using the 95% critical t-value for a two-tailed test, $t(0.95,N-1) = (\beta-<\beta>)/(\sigma_\beta/\sqrt{N})$ for a single distribution (Appendix 2). This value will only approach ~2 (actually 1.96) for replicate numbers of sixty or greater.

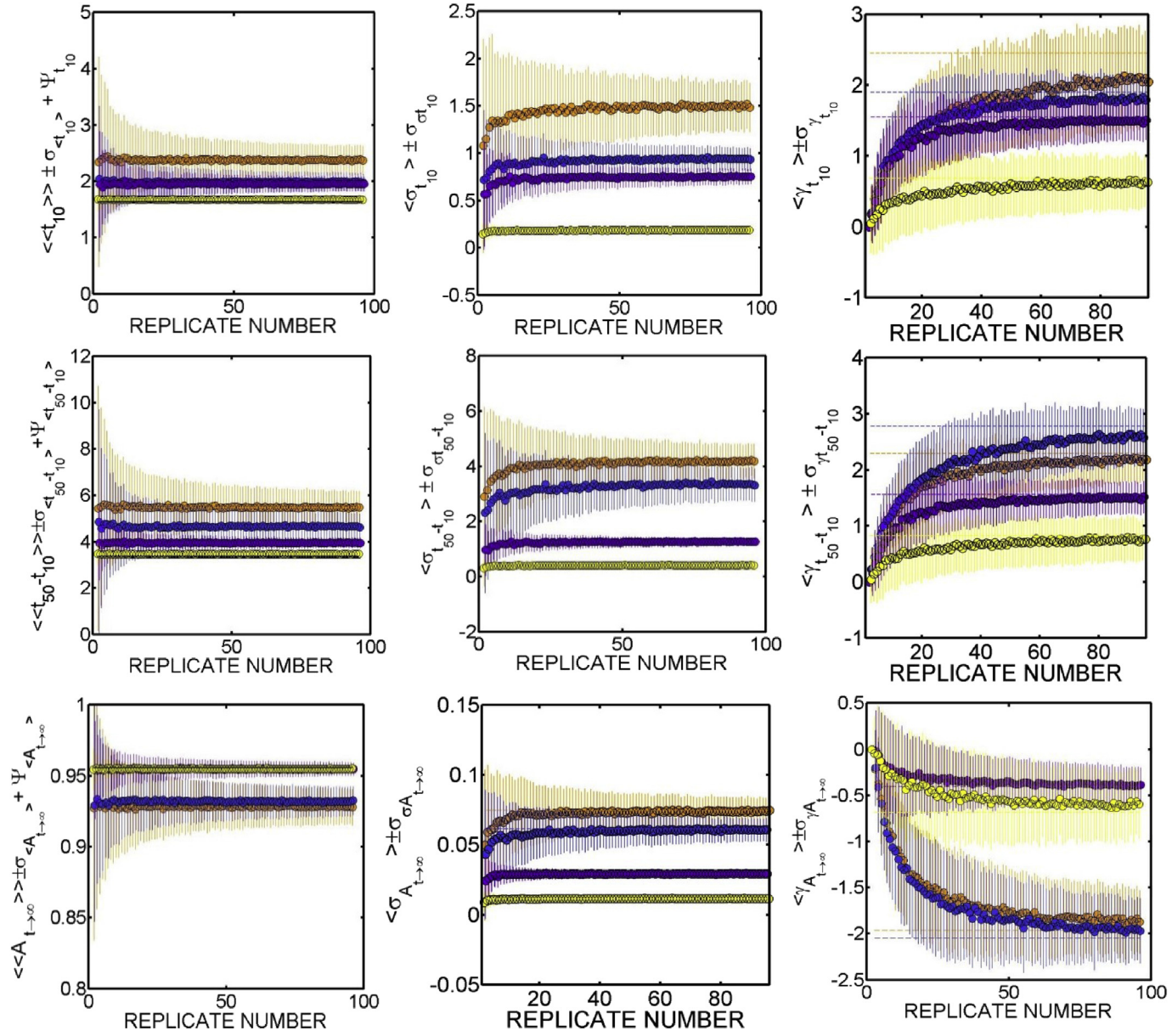


Fig. 4. Bootstrap analysis using 200 selections from a pool of 1000 initial simulations of convergence indicators corresponding to the distributions of t_{10} (Row 1), $t_{50}-t_{10}$ (Row 2) and $A_{t\to\infty}$ (Row 3). Convergence was monitored using (Column 1) Average of the bootstrap parameter average (<<β>>) plus or minus the standard deviation of the bootstrap averages plus the standard deviation of the bootstrap standard error (Ψ), (Column 2) Average of the bootstrap standard deviation (<σ>) plus or minus the standard deviation of the bootstrap standard deviation, (Column 3) Average of the bootstrap skew-like parameter <γ> plus or minus the standard deviation of the bootstrap value of γ. Colour representation is as follows, yellow = [k_1(GLV), k_2(GLV)], magenta = [k_1(FHV), k_2(GLV)], blue = [k_1(GLV), k_2(FHV)], orange = [k_1(FHV), k_2(FHV)]. Error bars reflect plus or minus 1 standard deviation of the 200 estimates used in the bootstrap series. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fairs less well at identifying significance between the intermediate variance sample groups and the high variance control groups (Rows 2, 3 of Fig. 8).

(ii) Comparison using Welch's-t-test

The first alternative test considered is Welch's t-test, a variant of the standard Student t-test, developed for statistical comparison of the independence of the means of two groups of normally distributed data having unequal sizes and unequal variance [34,35]. Welch's t-test is a parametric test as it is based on the assumption that the parameter distributions have a previously known analytical form (in this case the normal distribution). Welch's t-test calculates a t value on the basis of difference in group means relative to grouped standard errors evaluated as a vector sum (Fig. 6b), rather than as a conventional sum as implied for the previous procedure. The t-statistic and number of degrees of freedom (dof) is calculated as per Eqn. (3) [35].

$$t(0.95, dof) = \frac{\langle\beta_1\rangle - \langle\beta_2\rangle}{\sqrt{\frac{\sigma^2_{\beta_1}}{N_1} + \frac{\sigma^2_{\beta_2}}{N_2}}} \tag{3a}$$

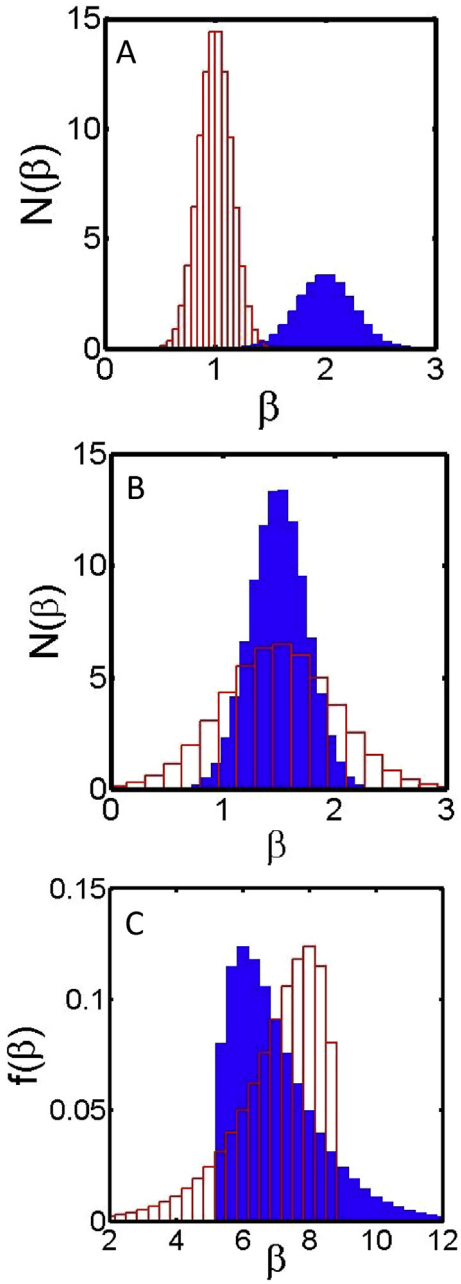


Fig. 5. Some possible changes in distribution of a parameter between control (blue) and sample (red) groups. (A) Raw histogram describing distributions with changed averages and variance (note unequal distribution size). (B) Raw histogram describing distributions with equal averages but with changed variance (for equal distribution sizes) (C) Normalized histogram reflecting distributions with equal means and variance but with different distribution shapes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

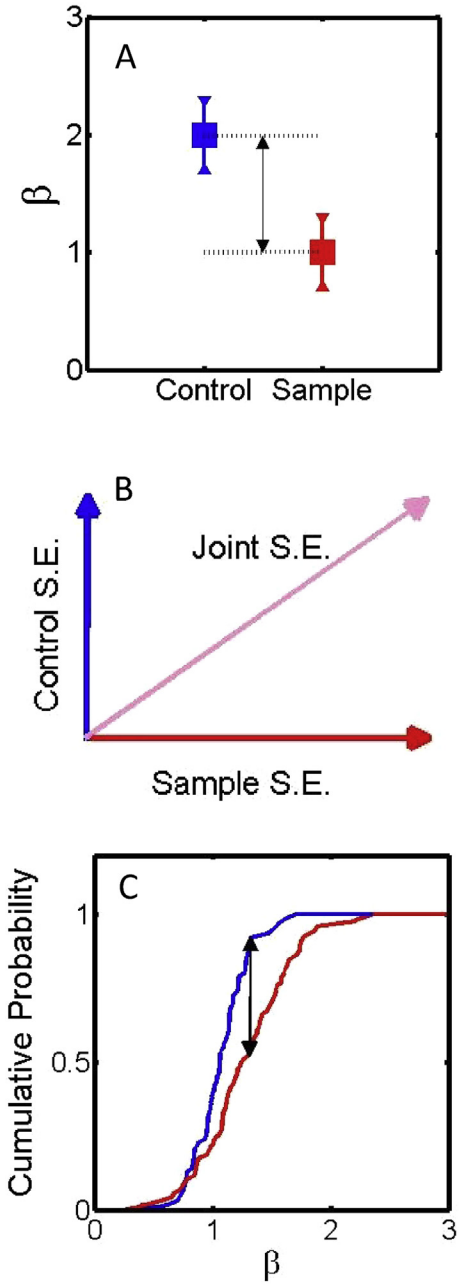


Fig. 6. Critical features of the three statistical tests. (A) <β> ± 2S.E. test (difference in mean vs. plus or minus two times the standard error) operates on the basis of linear combination of standard errors. (B) Welch's t-test: Tests the null hypothesis that the two distributions are the same by comparing the difference in means of the two distributions against the magnitude of the vector sum of standard errors and uses a variance weighted degree of freedom for calculation of significance based on a critical t-distribution. (C) Kolmogorov–Smirnov two-sample D-test: Tests the null hypothesis that the two distributions are the same by comparing the maximum difference in cumulative distribution plots against a table of critical values from the Kolmogorov distribution describing likelihood of that distance being achieved via random chance (for a given number of replicates).

$$dof = \frac{\left(\frac{\sigma_{\beta_1}^2}{N_1} + \frac{\sigma_{\beta_2}^2}{N_2}\right)^2}{\left(\frac{\sigma_{\beta_1}^4}{N_1^2(N_1-1)} + \frac{\sigma_{\beta_2}^4}{N_2^2(N_2-1)}\right)} \tag{3b}$$

Importantly, the number of degrees of freedom in Welch's t-test is weighted such that it penalizes high variance sample or control data by returning a lower numerical value for the joint degrees of freedom. A table of critical t-values (Appendix 2) is used for evaluation of the 95% confidence interval in rejecting the null hypothesis that the control and sample groups belong to the same distribution [34]. Fig. 9 describes the use of Welch's t-test in a type 1 comparison of the data (with each of the four rows describing a case of intrinsic variance tested against itself as sample and control). Interestingly Welch's t-test performs about as well as the simple <β> ± 2S.E. test with some slight tendency for pseudo-

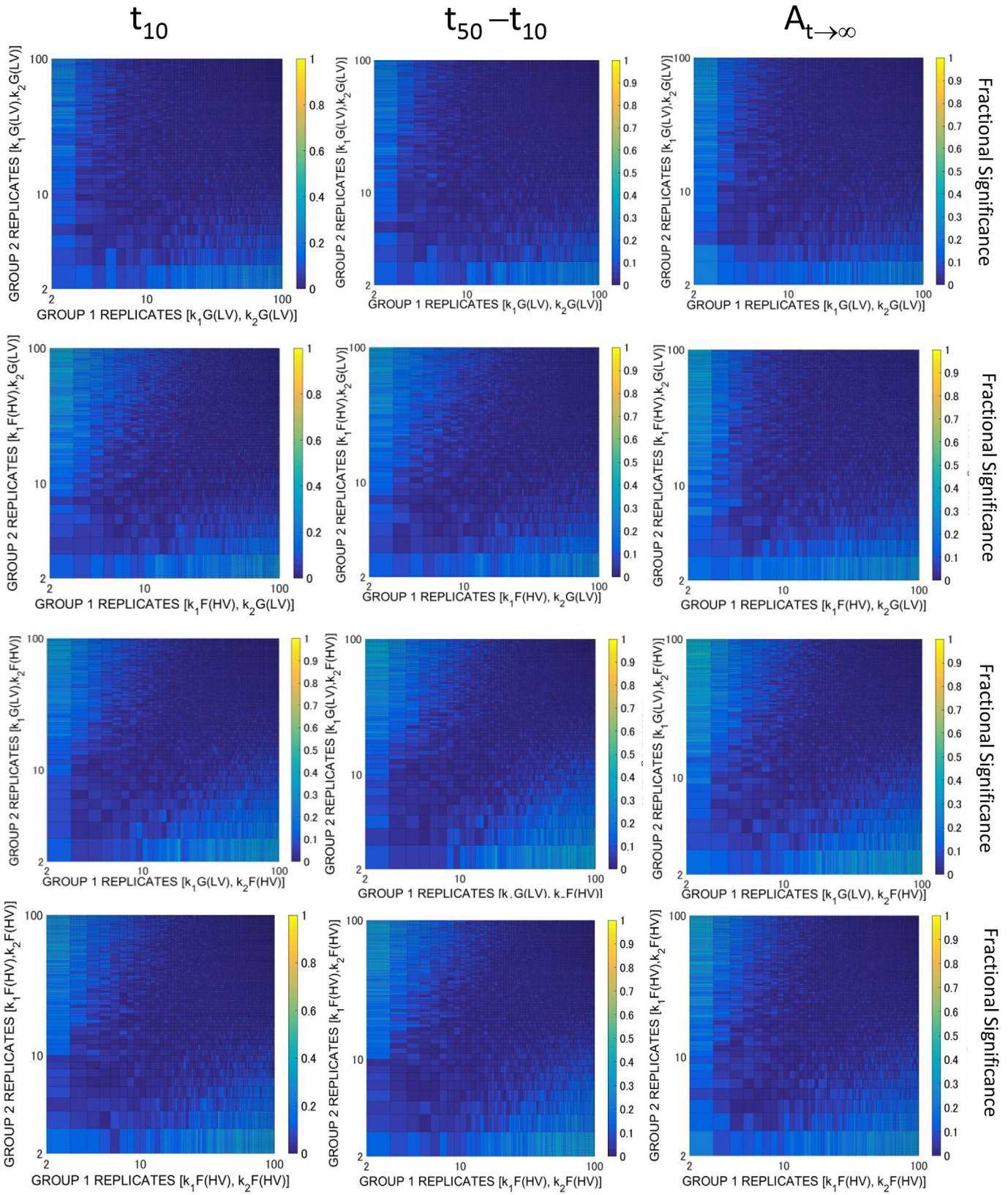


Fig. 7. Application of the <β> ± 2S.E. test to a type 1 comparison of sample and control data as a function of sample and control replicate number. Type 1 comparison - the null hypothesis that the sample and control distributions are the same is true. **Row 1**: Control group = Sample group = [k_1(GLV), k_2(GLV)] **Row 2**: Control group = Sample group = [k_1(FHV), k_2(GLV)] **Row 3**: Control group = Sample group = [k_1(GLV), k_2(FHV)], **Row 4**: Control group = Sample group = [k_1(FHV), k_2(FHV)]. **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

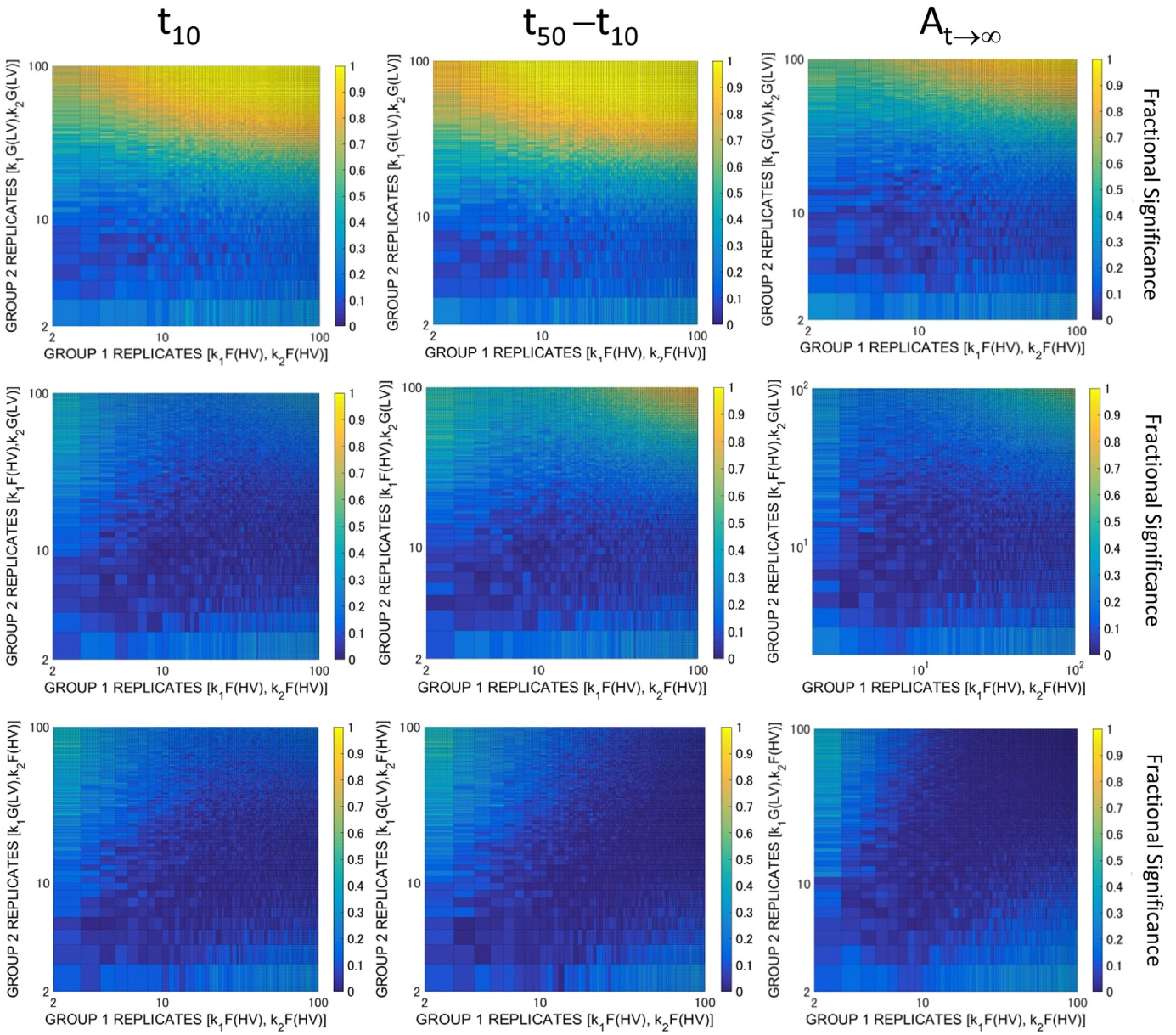


Fig. 8. Application of the <β> ± 2S.E. test to a type 2 comparison of sample and control data as a function of sample and control replicate number. Type 2 comparison - the null hypothesis that the sample and control distributions are the same is not true. **Row 1**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(GLV)]. **Row 2**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(FHV), k_2(GLV)]. **Row 3**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(FHV)] **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

significance being spotted in the off-diagonal regions. In the type 2 comparative test (Fig. 10), Welch's test performs noticeably better than the <β> ± 2S.E. test, as indicated by the clear identification of significance between the intermediate cases of variance when compared with the high variance control group (Row 2 of Fig. 10). This finding suggests a preference for using Welch's t-test over the <β> ± 2S.E. test. It is worth noting however that Welch's t-test cannot identify a significant difference when contrasting the penultimate high variance sample group against the high variance control group (Row 3 of Fig. 10).

(iii) Kolmogorov–Smirnov two-sample D-test

The third test considered is the Kolmogorov–Smirnov two-sample D-test (KS2D) (Fig. 6c). The KS2D-test is a non-parametric statistical test which is capable of identifying significant differences in both the mean and shape of the distribution and therefore is applicable to detecting the type of changes shown in Fig. 5A, B and C [35,36]. This makes the KS2D test different to the previously considered simple <β> ± 2S.E. tests and Welch's t-test, which only test for the differences in means between sample and control groups. The fundamental measure in the KS2D test is the parameter D, which is the maximum difference between cumulative frequency distributions of the sample and control groups over the domain of the measured variable β [35,36]. The process for calculation of D is shown below [35,37].

(a) Cumulative distributions of frequency plots of the sample, $C_{sample}(\beta)$, and control, $C_{control}(\beta)$, groups are calculated according to Eqn. (4).

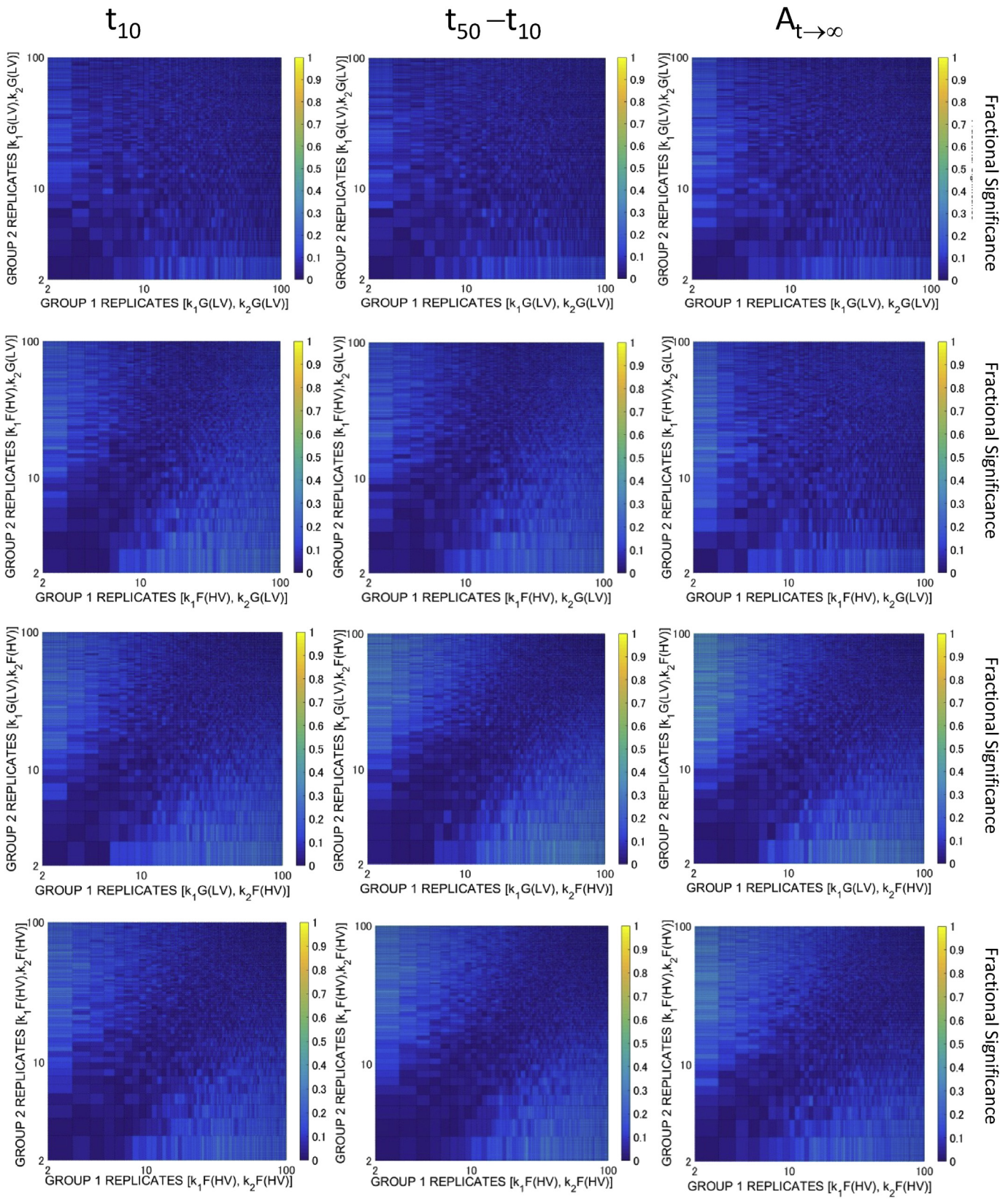


Fig. 9. Application of Welch's t-test to a type 1 comparison of sample and control data as a function of sample and control replicate number. Type 1 comparison - the null hypothesis that the sample and control distributions are the same is true. **Row 1**: Control group = Sample group = [k_1(GLV), k_2(GLV)] **Row 2**: Control group = Sample group = [k_1(FHV), k_2(GLV)] **Row 3**: Control group = Sample group = [k_1(GLV), k_2(FHV)], **Row 4**: Control group = Sample group = [k_1(FHV), k_2(FHV)]. **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

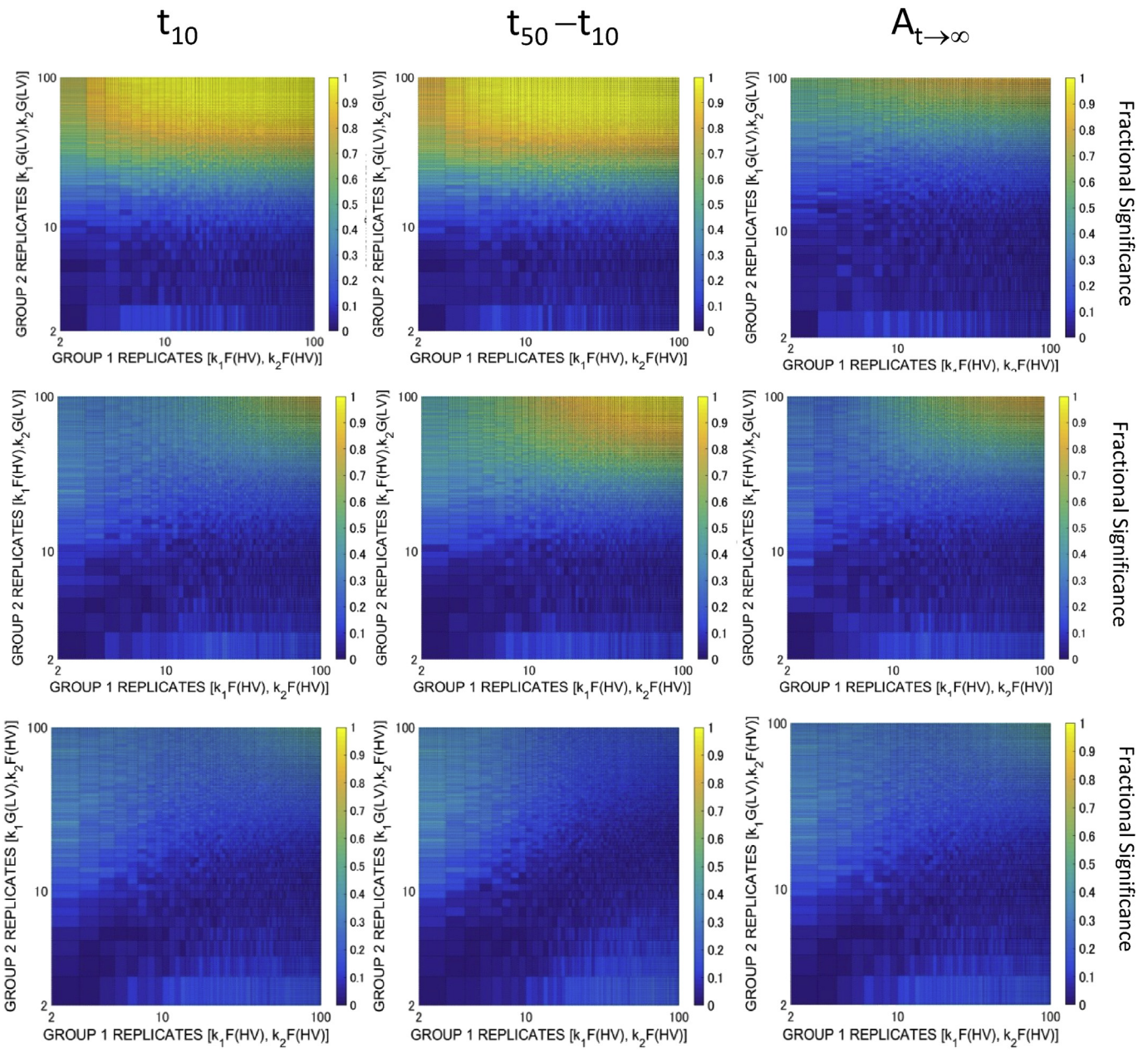


Fig. 10. Application of Welch's t-test to a type 2 comparison of sample and control data as a function of sample and control replicate number. Type 2 comparison - the null hypothesis that the sample and control distributions are the same is not true. **Row 1**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(GLV)]. **Row 2**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(FHV), k_2(GLV)]. **Row 3**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(FHV)] **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

$$C_{sample}(\beta_i) = \sum_{i}^{i_{\max}} \left(\frac{N(\beta_i)}{N_1} \right) \tag{4a}$$

$$C_{control}\left(\beta_j\right) = \sum_{j=1}^{j_{\max}} \left(\frac{N\left(\beta_j\right)}{N_2} \right) \tag{4b}$$

(b) The domain is scanned and the ordinate difference at common abscissa values between β_1 and β_N is determined. The maximum value of this difference is termed the Kolmogorov–Smirnov D value (Eqn. (5)).

$$D = \max \left| C_{sample}(\beta) - C_{control}(\beta) \right| \tag{5}$$

(c) The D value is compared against the appropriate critical value. If D is greater than the critical value, the null hypothesis that sample and control distributions are from the same distribution is rejected, and the alternative hypothesis that the samples are statistically different, is accepted. Critical values for the two-sample D statistic (for rejecting the null hypothesis that the two distributions are identical) are provided in Appendix 2.

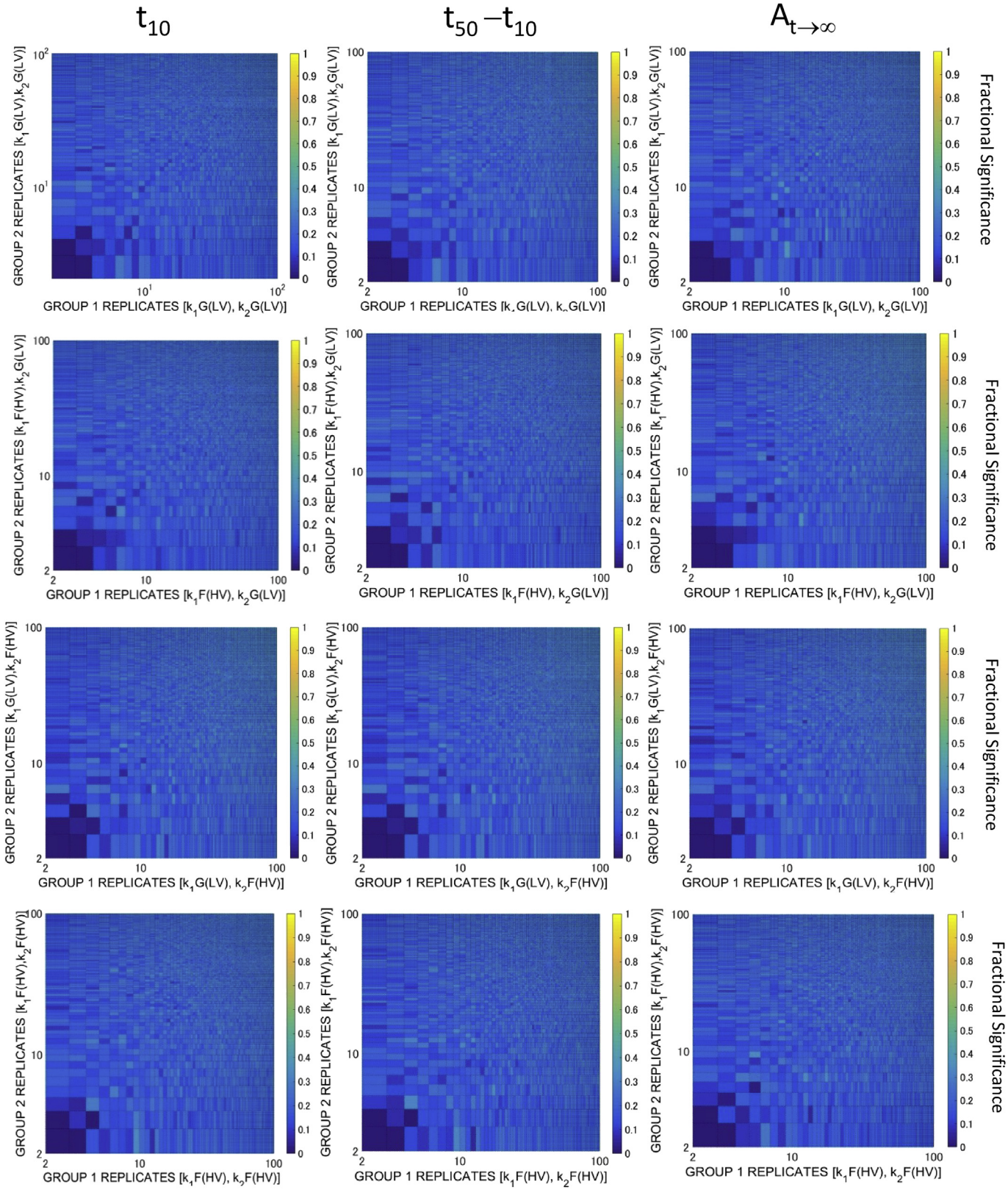


Fig. 11. Application of the Kolmogorov–Smirnov two-sample D-test to a type 1 comparison of sample and control data as a function of sample and control replicate number. Type 1 comparison - the null hypothesis that the sample and control distributions are the same is true. **Row 1**: Control group = Sample group = [k_1(GLV), k_2(GLV)] **Row 2**: Control group = Sample group = [k_1(FHV), k_2(GLV)] **Row 3**: Control group = Sample group = [k_1(GLV), k_2(FHV)], **Row 4**: Control group = Sample group = [k_1(FHV), k_2(FHV)]. **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

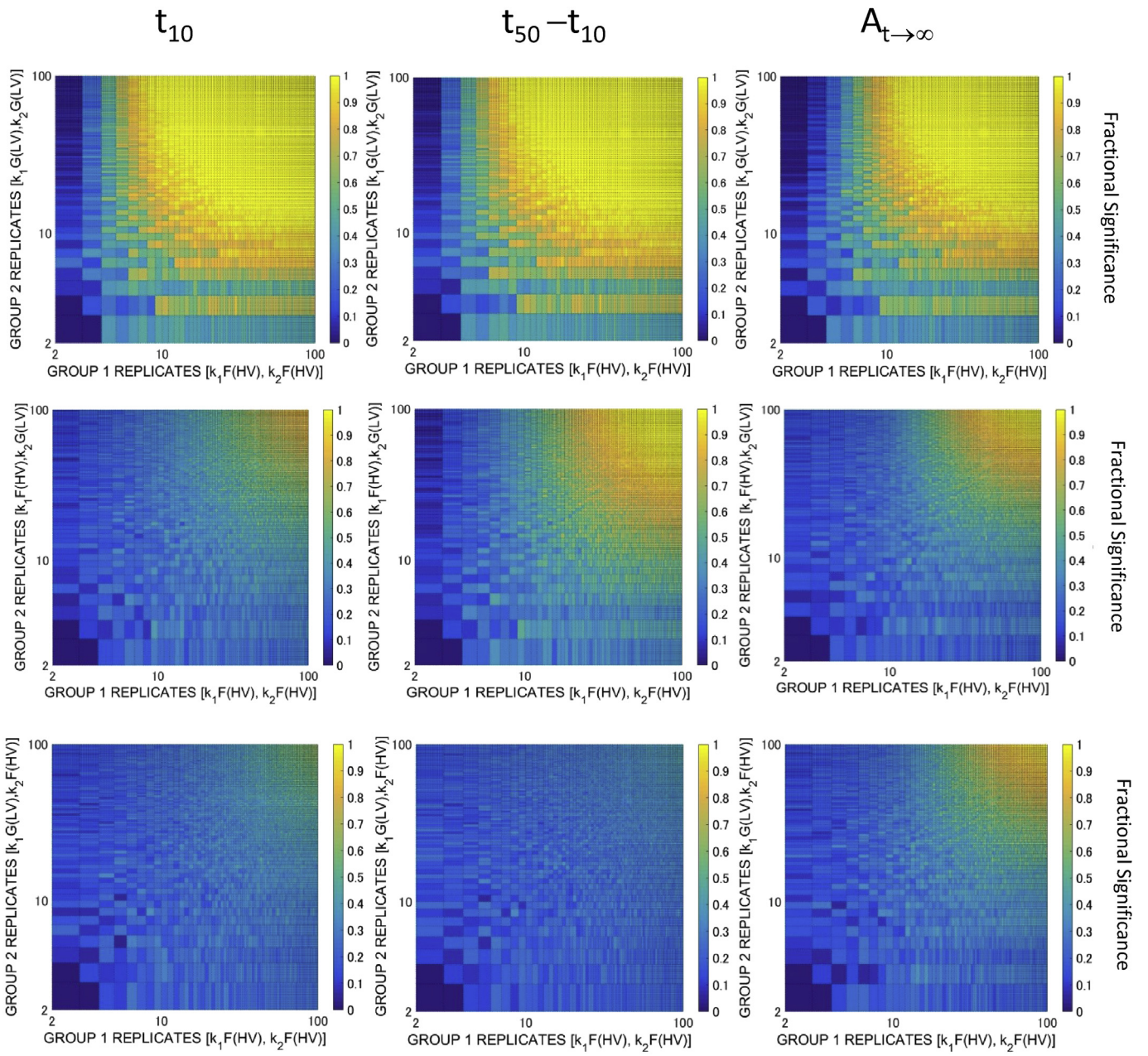


Fig. 12. Application of the Kolmogorov–Smirnov two-sample D-test to a type 2 comparison of sample and control data as a function of sample and control replicate number. Type 2 comparison - the null hypothesis that the sample and control distributions are the same is not true. **Row 1**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(GLV)]. **Row 2**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(FHV), k_2(GLV)]. **Row 3**: Control group = [k_1(FHV), k_2(FHV)], Sample group = [k_1(GLV), k_2(FHV)] **Column 1**: Test applied to t_{10} values. **Column 2**: Test applied to $t_{50}-t_{10}$ values. **Column 3**: Test applied to $A_{t\to\infty}$ values.

Application of the KS2D test to the type 1 and type 2 comparative cases are shown as Figs. 11 and 12 respectively. We note that the KS2D test does a good job in not falsely identifying significance in the type 1 comparative case, although the general level of fractional pseudo-significance is slightly greater than for the <β> ± 2S.E. and Welch's t-test (although there is no apparent off-diagonal dependence). From Fig. 12, we see that the KS2D-test excels in the type 2 comparative test, identifying significant differences across the range of low to high variance sample cases when tested against the highest variance control group (Rows 1–3 of Fig. 12).

3. Discussion

The current paper has considered the topic of inherent variability in the analysis of amyloid kinetics. Before embarking upon a discussion of our results, we stress that the type of variability that we are contemplating is that which would be observed even with exemplary experimental technique. Although we used an empirical process for introducing variability into our data, we first speculate upon the origin of such inherent variability. One potential cause may be a high-order concentration dependence of reaction rate coupled with limits on the accuracy of concentration formulation such as has been observed in sickle cell hemoglobin polymerization kinetics [38]. Such a statement may be formalized as Rate $\propto (C_1 \pm \xi_{C1})^g$ where ξ_{C1} is the error in the monomer concentration and $g >> 1$. Another potential source of inherent variation may stem from uncontrolled/unknown system effects such as the ratio of air/solution interface [39] or

potential volumetric effects related to a stochastically rare process[8] [17,22,30,40]. Other variation-inducing effects may relate to unique properties of the particular amyloid forming protein system, such as the potential existence of a condition sensitive secondary nucleation pathway (as might be realized by fiber fragmentation or surface nucleation [23–26,41–43]), the potential existence of a condition-sensitive alternative aggregation pathway (such as a competitive amorphous pathway [28,29]) or the sensitivity of the 'initial conditions' to protein age and preparation route [discussed in Ref. [44]]. However irrespective of the cause, the end result is variation and this is the topic considered in this paper.

To carry out our study into inherent variability we have had to develop computational and analytical tools that, if adopted, will sharpen the field's collective ability to extract substantial information from *in vitro* kinetic assays of amyloid growth. From a purely practical perspective, we have developed a novel procedure for simulating inherently noisy amyloid kinetic data possessing a known level of variation. On a similarly practical note, we have also demonstrated the systematic application of a data reduction strategy based on three characteristic points taken from a sigmoidal amyloid aggregation curve.

The combined ability to both produce (statistically) large numbers of well-defined synthetic data sets and reduce these data sets to tabular form, allowed us to ask questions related to sufficient sampling and evaluation of statistically significant differences between control and sample groups. With regards to the sampling question, we examined three markers of parameter distribution convergence and found a skew-like descriptor to be the most sensitive of the three markers tested. With regards to the question of assessing statistical significance, we examined three different procedures based on parametric and non-parametric analyses. The three procedures included a test based on comparison of means ± standard error, Welch's t-test and the Kolmogorov–Smirnov two-sample D test. Of the three, the Kolmogorov–Smirnov test was found to be the superior method. This may be largely attributed to the greater flexibility of the Kolmogorov–Smirnov test in assessing differences in distribution shape in addition to differences in mean values. A further major finding of this aspect of the study was the advantage gained from performing statistical testing across the range of characteristic parameters defined by t_{10}, ($t_{50}-t_{10}$) and $A_{t\to\infty}$. Such a concerted multi-parameter testing strategy affords maximum opportunity for identification of significant differences (as can be appreciated from examination of Row 3 of Fig. 12).

Based on analysis of systems with differing levels of pre-defined intrinsic variation, we suggest a useful classification of potential-amyloid forming model systems into four types, those exhibiting low intrinsic variance in nucleation and growth phases, those displaying significant levels of variation in either nucleation or growth phases and those afflicted with high levels of intrinsic variance in both nucleation and growth phases. An interesting caveat to this suggestion, raised in footnote 2, was based on the observation of an artificially large apparent variance induced in the nucleation marker t_{10} caused by a high variance in k_2 growth rate (even when the k_1 nucleation rate had an intrinsically low variance). As such, the long-held strategy of using t_{10} as a marker of nucleation [45–47] may need modification.

The common structural properties of amyloid [48,49] along with the potentially general ability of all proteins to form amyloid [as discussed in Ref. [50]] has led many researchers to substitute general amyloid-forming model systems in the place of particular disease-relevant protein model systems when conducting anti-amyloid inhibitor drug-screening tests [3,5–10,21,22]. Although often having a practical basis, relating to cost or ease of working, such decisions may limit appreciation of the potential effects of anti-amyloid drug treatment due to the fact that the change in a distribution, as outlined in Fig. 5, may often be more subtle than simple shift in mean value. Changes in amyloid distribution shape and variance have potential for modulating disease progression [51], as outlined in the Two-Hit model of amyloidosis and prion infection [41,42,51]. Physical demonstration of such changes is yet to be observed (although methods exist that can potentially measure these changes [52,53]).

With some notable exceptions [15,22,24], standard experimental approaches for measuring amyloid kinetics typically involve performing duplicate or triplicate measurements with subsequent averaging for comparison [21,32]. In this study, we have stressed that the kinetics displayed by some amyloid systems may appear inherently non-reproducible when tackled with such 'small data' type approaches. Such a situation invokes the specter of good versus bad model systems with the natural tendency for experimenters to gravitate towards the more well-behaved experimental models [21]. However, as the present work has shown, knowledge of variation in the system potentially represents important information on the likelihood of amyloid formation and hence also disease progression [51]. In this paper we have developed an alternative means for incorporating apparent stochasticity into the simulation of the aggregation reaction rates. If an analytical expression for the variability were to arise from either developments in theory, or directly from empirical measurement, such variability could be decomposed into distributions of k_1^* and k_2^* via direct fitting of the individual kinetic traces to Eqn. (A1). For particular aggregation reactions, use of such tailored parameter distributions in pre-experimental generation of test data sets, would enable tuning of the experimental design, in terms of required numbers of trials, to achieve more statistically robust outcomes. The methods outlined in this paper make a proper reckoning of these factors possible and therefore should help to improve analytical efforts aimed at quantifying noisy/seemingly recalcitrant amyloid kinetic assays. We hope that this work will supplement the literature directed at the development of anti-amyloid drugs [51,54].

Acknowledgements

DH would like to thank Dr. Nami Hirota, Assoc. Prof. Akira Kinjo and Dr. Daniel Clayton for helpful discussions and comments on an early version of this paper. The work of DH was jointly funded by an Australian National University Senior Research Fellowship and an Osaka University Cross-Appointment as an Associate Professor. RZ thanks the Chinese government and the ANU for the award of the CSC-ANU Scholarship. The research of J.A.C. is supported by a Project Grant from the National Health and Medical Research Council of Australia. The research of M.S. and Y.G. is supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology.

Appendix 1. Amyloid kinetic model

The amyloid kinetic model is an Oosawa type model (Eqn. (A1)) [ref. A1-a] recently used in the exploration of the turbidity generated by amyloid [ref. A1-b].

$$\frac{dC_N}{dt} = [k_1^*](C_M)^n - [k_2^*]C_N C_M \qquad \text{(A1-1)}$$

[8] i.e. the rare event becomes more common with greater system size.

$$\frac{d\sum C_{A_i}}{dt} = [k_2{*}]C_N C_M \quad \text{(A1-2)}$$

$$\frac{d\sum iC_{A_i}}{dt} = (n+1)[k_2{*}]C_N C_M + [k_2{*}]\sum C_{A_i} C_M \quad \text{(A1-3)}$$

where,

$$k_1{*} = k_1 \pm \xi_1 \quad \text{(A1-4a)}$$

$$k_2{*} = k_2 \pm \xi_2 \quad \text{(A1-4b)}$$

For each simulation, the relevant random variables were sampled from the appropriate distributions just once (at the zero time point) and then used throughout that particular simulation. A new sampling procedure was then begun for the next simulation until a total pool of one-thousand replicates was constructed. The above equation set was solved by numerical integration using a modified mid-point adaption of Euler's procedure [ref. A1-c]. Monomer concentration was calculated via mass conservation at each time point in the numerical integration.

Appendix 1. References

[A1-a] Oosawa, F., & Asakura, S. (1975). Thermodynamics of the Polymerization of Protein. Academic Press.
[A1-b] Hall, D., Zhao, R., Dehlsen, I., Bloomfield, N., Williams, S.R., Arisaka, F., Goto, Y. and Carver, J.A. (2016). Protein aggregate turbidity: Simulation of turbidity profiles for mixed aggregation reactions. Analytical biochemistry. 498, 78–94.
[A1-c] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996). Numerical recipes in C (Vol. 2). Cambridge: Cambridge University Press. Chapter 16 – Integration of Ordinary Differential Equations.

Appendix 2. Critical values for listed statistical tests.

Critical t-values: Critical values for the Student's t-test and Welch's t-test for group sizes up to dof = 100 are given in Supplementary Table 1 (ref. A2-a) (note interpolated values were used for sample sizes greater than 30).

Critical D-values: Critical values for the Kolmogorov Smirnov two sample D-test group for $N_1 < N_2 \leq 3$ and $N_2 < N_1 \leq 3$ were set as greater 1 i.e. no significant difference could be inferred. Beyond this limit a suitable approximation for assessing confidence at the 95% level was based on the following approximate empirical Eqs. (A2-1) [ref. A2-b 37 discussed to a greater extent in ref. A2-c].

$$D_{crit} = 1.36\sqrt{\left(\frac{N_1 + N_2}{N_1 \cdot N_2}\right)} \quad \text{(A2-1)}$$

Appendix 2. References

[A2-a] C. Dougherty (2002) Introduction to Econometrics (Second edition 2002, Oxford University Press, Oxford). Supplementary Tables.
[A2-b] Van Belle, G., Fisher, L. D., Heagerty, P. J., & Lumley, T. (2004). Biostatistics: a methodology for the health sciences (Vol. 519). John Wiley & Sons. Ch 8.7 Kolmogorov–Smirnov two sample test.
[A2-c] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996). Numerical recipes in C (Vol. 2). Cambridge: Cambridge University Press. Chapter 15.6 Confidence Limits on Estimated Model Parameters

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ab.2016.07.013.

References

[1] J.D. Sipe, A.S. Cohen, Review: history of the amyloid fibril, J. Struct. Biol. 130 (2000) 88–98.
[2] R.H. Sayed, P.N. Hawkins, H.J. Lachmann, Emerging treatments for amyloidosis, Kidney Int. 87 (2015) 516–526.
[3] J. McLaurin, D.S. Yang, C.M. Yip, P.E. Fraser, Review: modulating factors in amyloid-β fibril formation, J. Struct. Biol. 130 (2000) 259–270.
[4] M. So, D. Hall, Y. Goto, Revisiting supersaturation as a factor determining amyloid fibrillation, Curr. Opin. Struct. Biol. 36 (2016) 32–39.
[5] Q. Wang, X. Yu, L. Li, J. Zheng, Inhibition of amyloid-β aggregation in Alzheimer's disease, Curr. Pharm. Des. 20 (2014) 1223–1243.
[6] A. Porzoor, B. Alford, H.M. Hügel, D. Grando, J. Caine, I. Macreadie, Anti-amyloidogenic properties of some phenolic compounds, Biomolecules 5 (2015) 505–527.
[7] K. Ono, Y. Yoshiike, A. Takashima, K. Hasegawa, H. Naiki, M. Yamada, Potent anti-amyloidogenic and fibril-destabilizing effects of polyphenols in vitro: implications for the prevention and therapeutics of Alzheimer's disease, J. Neurochem. 87 (2003) 172–181.
[8] H. Naiki, N. Hashimoto, S. Suzuki, H. Kimura, K. Nakakuki, F. Gejyo, Establishment of a kinetic model of dialysis-related amyloid fibril extension in vitro, Amyloid 4 (1997) 223–232.
[9] H. Levine, Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution, Protein Sci. 2 (1993) 404–410.
[10] S. Ghosh, N.K. Pandey, S. Dasgupta, (−)-Epicatechin gallate prevents alkali-salt mediated fibrillogenesis of hen egg white lysozyme, Int. J. Biol. Macromol. 54 (2013) 90–98.
[11] D. Hall, R. Zhao, I. Dehlsen, N. Bloomfield, S.R. Williams, F. Arisaka, Y. Goto, J.A. Carver, Protein aggregate turbidity: Simulation of turbidity profiles for mixed aggregation reactions, Anal. Biochem. 498 (2016) 78–94.
[12] Y. Wang, S. Petty, A. Trojanowski, K. Knee, D. Goulet, I. Mukerji, J. King, Formation of amyloid fibrils in vitro from partially unfolded intermediates of human γC-crystallin, Investig. Ophthalmol. Vis. Sci. 51 (2010) 672–678.
[13] A.M. Morris, M.A. Watzky, R.G. Finke, Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature, Biochim. Biophys. Acta (BBA)-Proteins Proteomics 1794 (2009) 375–397.
[14] M. So, Y. Yoshimura, Y. Goto, Ultrasonication-forced Amyloid Fibrillation of Proteins. In Advances in Organic Crystal Chemistry, Springer Japan, 2015, pp. 15–29.
[15] W.F. Xue, S.W. Homans, S.E. Radford, Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly, Proc. Natl. Acad. Sci. 105 (2008) 8926–8931.
[16] J.R. Brender, J. Krishnamoorthy, M.F. Sciacca, S. Vivekanandan, L. D'Urso, J. Chen, C. La Rosa, A. Ramamoorthy, Probing the sources of the apparent irreproducibility of amyloid formation: drastic changes in kinetics and a switch in mechanism due to micellelike oligomer formation at critical concentrations of IAPP, J. Phys. Chem. B 119 (2015) 2886–2896.
[17] P. Hortschansky, V. Schroeckh, T. Christopeit, G. Zandomeneghi, M. Fändrich, The aggregation kinetics of Alzheimer's β-amyloid peptide is controlled by stochastic nucleation, Protein Sci. 14 (2005) 1753–1759.
[18] C.M. Pfefferkorn, R.P. McGlinchey, J.C. Lee, Effects of pH on aggregation kinetics of the repeat domain of a functional amyloid, Pmel17, Proc. Natl. Acad. Sci. 107 (2010) 21447–21452.
[19] N.D. Hammer, B.A. McGuffie, Y. Zhou, M.P. Badtke, A.A. Reinke, K. Brännström, M.R. Chapman, The C-terminal repeating units of CsgB direct bacterial functional amyloid nucleation, J. Mol. Biol. 422 (2012) 376–389.
[20] M. Garvey, S.L. Gras, S. Meehan, S.J. Meade, J.A. Carver, J.A. Gerrard, Protein nanofibres of defined morphology prepared from mixtures of crude crystallins, Int. J. Nanotechnol. 6 (2009) 258–273.
[21] S.A. Hudson, H. Ecroyd, T.W. Kee, J.A. Carver, The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds, FEBS J. 276 (2009) 5960–5972.
[22] L. Giehm, D.E. Otzen, Strategies to increase the reproducibility of protein fibrillization in plate reader assays, Anal. Biochem. 400 (2010) 270–281.
[23] H. Yagi, A. Mizuno, M. So, M. Hirano, M. Adachi, Y. Akazawa-Ogawa, Y. Hagihara, T. Ikenoue, Y.H. Lee, Y. Kawata, Y. Goto, Ultrasonication-dependent formation and degradation of α-synuclein amyloid fibrils, Biochim. Biophys. Acta (BBA)-Proteins Proteomics 1854 (2015) 209–217.
[24] A. Umemoto, H. Yagi, M. So, Y. Goto, High-throughput analysis of ultrasonication-forced amyloid fibrillation reveals the mechanism underlying the large fluctuation in the lag time, J. Biol. Chem. 289 (2014) 27290–27299.

[25] S.A. McBride, S.P. Sanford, J.M. Lopez, A.H. Hirsa, Shear-induced amyloid fibrillization: the role of inertia, Soft Matter 12 (2016) 3461–3467.
[26] S.Y. Ow, D.E. Dunstan, The effect of concentration, temperature and stirring on hen egg white lysozyme amyloid formation, Soft Matter 9 (2013) 9692–9701.
[27] K. Eden, R. Morris, J. Gillam, C.E. MacPhee, R.J. Allen, Competition between primary nucleation and autocatalysis in amyloid fibril self-assembly, Biophys. J. 108 (2015) 632–643.
[28] D. Hall, J. Kardos, H. Edskes, J.A. Carver, Y. Goto, A multi-pathway perspective on protein aggregation: Implications for control of the rate and extent of amyloid formation, FEBS Lett. 589 (2015) 672–679.
[29] Y. Yoshimura, Y. Lin, H. Yagi, Y.H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki, Y. Goto, Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation, Proc. Natl. Acad. Sci. 109 (2012) 14446–14451.
[30] J. Szavits-Nossan, K. Eden, R.J. Morris, C.E. MacPhee, M.R. Evans, R.J. Allen, Inherent variability in the kinetics of autocatalytic protein self-assembly, Phys Rev Lett 113 (2014) 098101.
[31] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Vol. 2, Cambridge University Press, Cambridge, 1996 (Chapter 7). Random Numbers.
[32] Y. Ohhashi, M. Kihara, H. Naiki, Y. Goto, Ultrasonication-induced amyloid fibril formation of β2-microglobulin, J. Biol. Chem. 280 (2005) 32843–32848.
[33] L. Giehm, N. Lorenzen, D.E. Otzen, Assays for α-synuclein aggregation, Methods 53 (2011) 295–305.
[34] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Vol. 2, Cambridge University Press, Cambridge, 1996 (Chapter 14). Statistical Description of Data.
[35] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Vol. 2, Cambridge University Press, Cambridge, 1996 (Chapter 15).6 Confidence Limits on Estimated Model Parameters.
[36] G.D. Ruxton, The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test, Behav. Ecol. 17 (2006) 688–690.
[37] G. Van Belle, L.D. Fisher, P.J. Heagerty, T. Lumley, Biostatistics: a Methodology for the Health Sciences, Vol. 519, John Wiley & Sons, 2004. Ch 8.7 Kolmogorov-Smirnov two sample test.
[38] Z. Cao, F.A. Ferrone, A 50th order reaction predicted and observed for sickle hemoglobin nucleation, J. Mol. Biol. 256 (1996) 219–222.
[39] D. Jiang, K.L. Dinh, T.C. Ruthenburg, Y. Zhang, L. Su, D.P. Land, F. Zhou, A kinetic model for β-amyloid adsorption at the air/solution interface and its implication to the β-amyloid aggregation process, J. Phys. Chem. B 113 (2009) 3160–3168.
[40] A. Pawar, G. Favrin, Finite Size Effects in Simulations of Protein Aggregation, PLOS One 3 (2008) e2641.
[41] D. Hall, H. Edskes, Silent prions lying in wait: a two-hit model of prion/amyloid formation and infection, J. Mol. Biol. 336 (2004) 775–786.
[42] D. Hall, H. Edskes, A model of amyloid's role in disease based on fibril fracture, Biophys. Chem. 145 (2009) 17–28.
[43] K. Sasahara, D. Hall, D. Hamada, Effect of lipid type on the binding of lipid vesicles to islet amyloid polypeptide amyloid fibrils, Biochemistry 49 (2010) 3040–3048.
[44] M.M. Pallitto, R.M. Murphy, A mathematical model of the kinetics of β-amyloid fibril growth from the denatured state, Biophys. J. 81 (2001) 1805–1822.
[45] D. Hall, The effects of tubulin denaturation on the characterization of its polymerization behavior, Biophys. Chem. 104 (2003) 655–682.
[46] F.A. Ferrone, J. Hofrichter, H.R. Sunshine, W.A. Eaton, Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism, Biophys. J. 32 (1980) 361.
[47] W.A. Voter, H.P. Erickson, The kinetics of microtubule assembly. Evidence for a two-stage nucleation mechanism, J. Biol. Chem. 259 (1984) 10430–10438.
[48] R. Tycko, Solid state NMR studies of amyloid fibril structure, Annu. Rev. Phys. Chem. 62 (2011) 279.
[49] C. Goldsbury, U. Baxa, M.N. Simon, A.C. Steven, A. Engel, J.S. Wall, S.A. Müller, Amyloid structure and assembly: insights from scanning transmission electron microscopy, J. Struct. Biol. 173 (2011) 1–13.
[50] R.N. Rambaran, L.C. Serpell, Amyloid fibrils. Prion 2–3 (2008) 112–117.
[51] D. Hall, H. Edskes, Computational modeling of the relationship between amyloid and disease, Biophys. Rev. 4 (2012) 205–222.
[52] D. Hall, L. Huang, On the use of size exclusion chromatography for the resolution of mixed amyloid aggregate distributions: I. Equilibrium partition models, Anal. Biochem. 426 (2012) 69–85.
[53] D. Hall, Semi-automated methods for simulation and measurement of amyloid fiber distributions obtained from transmission electron microscopy experiments, Anal. Biochem. 421 (2012) 262–277.
[54] J. Masel, V.A. Jansen, Designing drugs to stop the formation of prion aggregates and other amyloids, Biophys. Chem. 88 (2000) 47–59.

Paper 3: Measurement of Amyloid Formation by Turbidity Assay – Seeing Through the Cloud

The third of the three papers published during my studies at the Australian National University Research School of Chemistry (Zhao et al. 2016) was a detailed examination of the relevant literature relating to the basic light scattering behaviour of my chaperone aggregation functional assay. In this paper I reviewed the physical chemistry of the turbidimetric assay methodology, investigating the reviewed information with a series of pedagogical kinetic simulations. We particularly focused upon recent literature relating to ultra-microscope image analysis (Hall, 2012; Usov & Messenga, 2015), light scattering and turbidity development by protein aggregates (Garcia-Lopez *et al.,* 2006, 2008; Hall *et al.,* 2016a), and computer simulation of the kinetics of amyloid and other aggregate types (Adachi *et al.,* 2015; Hall *et al.,* 2015; Hall *et al.,* 2016a).

Biophys Rev (2016) 8:445–471
DOI 10.1007/s12551-016-0233-7

REVIEW

Measurement of amyloid formation by turbidity assay—seeing through the cloud

Ran Zhao[1] · Masatomo So[2] · Hendrik Maat[1] · Nicholas J. Ray[1] · Fumio Arisaka[3] · Yuji Goto[2] · John A. Carver[1] · Damien Hall[1,2]

Received: 9 October 2016 / Accepted: 11 October 2016 / Published online: 23 November 2016
© International Union for Pure and Applied Biophysics (IUPAB) and Springer-Verlag Berlin Heidelberg 2016. This article is published with open access at Springerlink.com

Abstract Detection of amyloid growth is commonly carried out by measurement of solution turbidity, a low-cost assay procedure based on the intrinsic light scattering properties of the protein aggregate. Here, we review the biophysical chemistry associated with the turbidimetric assay methodology, exploring the reviewed literature using a series of pedagogical kinetic simulations. In turn, these simulations are used to interrogate the literature concerned with in vitro drug screening and the assessment of amyloid aggregation mechanisms.

Keywords Amyloid biophysics · Turbidimetric method · Amyloid aggregation kinetics · Data reduction · Nonlinear signal response

Introduction

The word 'amyloid' was first coined over 160 years ago to describe white densities of protein aggregate in autopsied livers, in the mistaken belief that they represented deposits of starch[1] (Virchow 1854). In modern day scientific practice, the meaning of the term amyloid has extended beyond its original histopathological association with disease, to describe a class of nanofiber able to be formed by most proteins upon their adoption of an unfolded structure and subsequent polymerization via intermolecular β-sheet formation (Toyama and Weissman 2011; Tycko 2011) (Fig. 1). From this current perspective, amyloid is interpreted as a particular tertiary fold, whose structural maintenance is conditional upon its stabilization as part of a higher-order quaternary assembly.

Regardless of the motivation for its study, the most basic practical requirement for experimenting with amyloid is an assay procedure for monitoring its formation. Although there are numerous techniques that are capable of achieving this goal (Li et al. 2009; Nilsson 2004), by far the three most common in vitro assay formats are those based on turbidity (Dolado et al. 2005; Sant'Anna et al. 2016), induced fluorescence associated with Thioflavin T dye binding (Dalpadado et al. 2016; Levine 1993; Naiki et al. 1997) and induced absorbance spectral shift exhibited upon Congo Red dye binding (Klunk et al. 1989). In this review, we examine the literature concerned with the underlying theory and experimental interpretation of the turbidity assay (Andreu and Timasheff 1986; Moody et al. 1996). As such, our review differs from many others on the topic of amyloid biophysics (Hall and Edskes 2012; Kashchiev 2015; Ma and Nussinov 2006; Mezzenga and Fischer 2013; Sasahara and Goto 2013; So et al. 2016; Tycko and Wickner 2013) by its restriction to matters directly related to achieving an understanding of the turbidimetric method. Towards this goal our examination will pay particular attention to recent articles concerned with ultra-microscope image analysis (Hall 2012; Usov and Mezzenga 2015), light scattering and turbidity development by protein

[1] Latin: amylum; Greek: amylon *meaning* starch or ground meal

This article is part of a Special Issue on 'Analytical Quantitative Relations in Biochemistry' edited by Damien Hall and Stephen Harding

✉ Damien Hall
damien.hall@anu.edu.au; damien.hall@protein.osaka-u.ac.jp

[1] Research School of Chemistry, Australian National University, Acton ACT 2601, Australia

[2] Institute for Protein Research, Osaka University, 3-1- Yamada-oka, Suita, Osaka 565-0871, Japan

[3] College of Bio-resource Sciences, Nihon University, Chiyoda-ku, Tokyo 102-8275, Japan

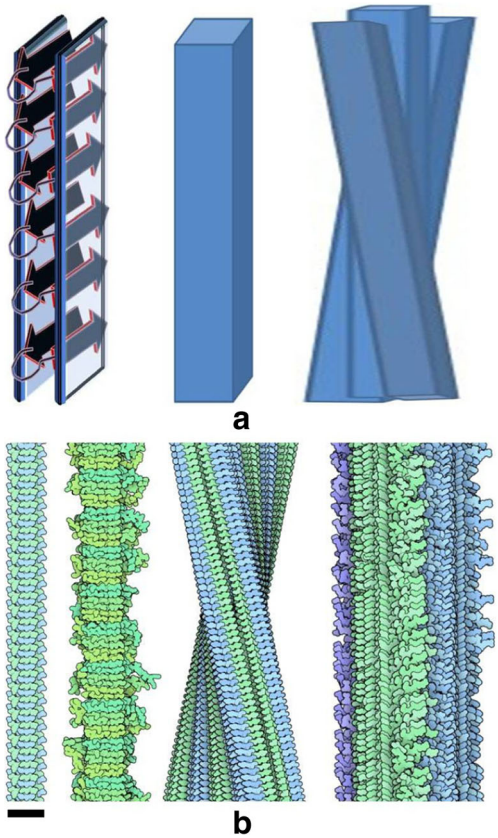


Fig. 1 Amyloid structure. **a** Consensus structural features of the amyloid fibre. *Left* Intermolecular β-sheet stacks are formed between polypeptides along the direction of the fibre. One or more sections of a polypeptide may contribute to the longitudinal β-sheet formation. Hydrophobic-driven lateral packing may occur between the orthogonal faces of the β-sheet elements within the amyloid fibre. *Centre* The simplest possible fibre arrangement is termed a protofibril which can be characterized by a length, width, persistence length and helical pitch. *Right* Hydrophobic packing forces can cause multiple protofibrils to align to yield higher-order quaternary arrangements of amyloid fibres termed 'mature fibres' (figures adapted, with permission, from Hall and Edskes 2012). **b** Artistic renderings of the structures of four different amyloids solved by a combination of solid-state nuclear magnetic resonance and various types of electron microscopy. *From left to right* amyloid fibres derived from the human prion protein (Apostol et al. 2010), yeast prion amyloid fibres formed from the full-length yeast protein HET-s (Van Melckebeke et al. 2010), amyloid formed from a peptide segment of transthyretin (Fitzpatrick et al. 2013) and a mature amyloid fibre, composed of multiple protofibrils, derived from the brain of an Alzheimer's Disease patient (Paravastu et al. 2008). *Bar in lower left hand corner* 5 nm [figures adapted from painted illustrations by D.G. Goodsell (Goodsell et al. 2015)]

aggregates (Garcia-Lopez and Garcia-Rubio 2008; Garcia-Lopez et al. 2006; Hall et al. 2016a) and simulation of the kinetics of amyloid (Ghosh et al. 2010; Hall et al. 2016b; Kaschiev 2015) and other aggregate types (Adachi et al. 2015; Hall et al. 2015). Although placing the focus of the review on a single type of assay procedure may seem like a retreat from the bigger questions, such as the relation between amyloid and disease (Hall and Edskes 2012; Walker and Jucker 2015), we contend that a thorough understanding of principles associated with the turbidimetric monitoring of amyloid growth will sharpen our collective ability to make informed judgements about the biological implications of results gained from in vitro protein aggregation assays.

In the following sections we outline (1) consensus physical models of amyloid aggregates to better understand how they might interact with visible wavelength light, (2) the general physics of the interaction of light with matter, concentrating on the description of utilitarian mathematical transforms able to estimate the value of the turbidity on the basis of attainable experimental quantities and (3) consensus kinetic models of aggregate development capable of predicting broad features of the time course of aggregation for various limiting-case regimes of amyloid growth. As a means for summarizing relevant literature into compact review form, the geometric and turbidimetric transforms discussed in (1) and (2) are applied to the output of the consensus kinetic models presented in (3). These transformed data sets are then used as aids for the interpretation of literature related to amyloid aggregation kinetics.

(i) Consensus physical models of protein aggregates

Solid-state nuclear magnetic resonance experiments and various electron microscopy techniques have been used, in combination, to determine atomic-level structural models for several amyloids (Tycko and Wickner 2013). Figure 1a is a schematic highlighting three consensus features displayed by nearly all amyloid structures observed to date (Tycko and Wickner 2013; Tycko 2014), namely:

(1) intermolecular β-sheet formation directed parallel to the long axis of the fibre
(2) hydrophobic stacking of β-sheet segments perpendicular to the long axis of the fibre (if more than one β-sheet motif is present per polypeptide)
(3) lateral association of protofibrils to form multi-fibre assemblies

Figure 1b shows an artistic rendering by Goodsell (Goodsell et al. 2015) of four different amyloid structures formed from four different proteins, with all structures determined using hybrid-combination approaches (Apostol et al. 2010; Fitzpatrick et al. 2013; Lu et al. 2013; Van Melckebeke et al. 2010). Regular arrangements of stacked β-sheets can be seen in all rendered images. Packing restraints associated with these bonding patterns induce a differential diffraction of incident X-rays from fibres aligned perpendicularly to the incident radiation, with this image providing the

basis of the 'cross-β' structural nomenclature[2] often used to describe internal amyloid bonding patterns (Liu et al. 2016; Makin and Serpell 2005).

Although atomic models provide maximum structural information, they are often not representative of the population of amyloid fibres in typical in vitro, or in vivo, experiments, for which fibre heterogeneity tends to be the norm, rather than the exception (Guo and Akhremitchev 2006; Liu et al. 2016; Meinhardt et al. 2009). With regard to fibre structure, two general classes of variation exist. The first type of variation is related to heterogeneity in polymer length/width[3] (Hall and Minton 2002, 2004; Hall et al. 2016b; Szavits-Nossan et al. 2014). Such variation in fibre length and width is a natural feature of any polymerization reaction capable of longitudinal growth and lateral association[4] (Hall 2012; Ridgley and Barone 2013; Umemoto et al. 2014). The second type of variation is generated by intrinsic structural differences in the core amyloid/aggregate stemming from competing nucleation pathways available to a single polypeptide sequence (King and Diaz-Avalos 2004; Paravastu et al. 2008; Petkova et al. 2005; Tanaka et al. 2004; Toyama and Weissman 2011; Tycko and Wickner 2013). The availability of multiple aggregation pathways is thought to represent the basis of amyloid polymorphism (also known as strain formation) (King and Diaz-Avalos 2004; Tanaka et al. 2004; Tycko and Wickner 2013). Regardless of the cause of the variation, the end result is a heterogeneous mixture of fibres and other aggregate products (Hall 2012; Umemoto et al. 2014). Indeed, in order to be used in high-resolution structural studies, such heterogeneous fibre distributions must first be carefully treated by either selective degradation, purification or re-cultivation (via selection and re-seeding) in a manner analogous to crystal farming (Qiang et al. 2011; Scherpelz et al. 2016).

Atomic-level differences in amyloid structure are often invisible, or muted, when coarser assay methods are used (Li et al. 2009; Nilsson 2004). This is the case for amyloid scattering/turbidity experiments conducted using visible wavelength light (~400–700 nm) for which the large wavelength—relative to the aggregate size—makes anything more than a mesoscopic[5] description of amyloid structure superfluous. The most common and direct means for such estimations of mesoscopic structures involve the use of ultramicroscopy techniques, atomic force microscopy (Adamcik et al. 2010; Harper et al. 1997), transmission electron microscopy (TEM) (Goldsbury et al. 2011; (Hall 2012) or total internal reflection fluorescence microscopy (Ban et al. 2003, Ban and Goto, 2006)[6]. A small number of researchers (Hall 2012; Usov and Mezzenga 2015) have quantitatively reduced amyloid ultramicroscope images into equivalent hard particle models based on a description of the aggregates as spheres (for small amyloid/oligomers and large amorphous aggregates) or cylinders (for amyloid fibers) with the asymmetric bodies assigned a characteristic rigidity value (Adamcik and Mezzenga 2011; Hall 2012) or a defined chiral twist (Usov and Mezzenga 2015). In the study carried out by Hall (2012), semi-automated analysis software was developed and applied to TEM images of amyloid formed from bovine insulin under high temperature and low pH conditions (Fig. 2). In that work, two algorithms were used to reduce the fibre images to a table of characteristic values. The first algorithm (Eq. 1a, b, c) enabled deconvolution of the measured perimeter, P, and area, A, of an individual fibre (calculated from pixel analysis of the TEM image bitmap) in terms of an equivalent sphero-cylinder, with the result that each fibre was reduced to an internal length L and a fibre width W (Fig. 2c).

$$P = 2\left(L + \frac{\pi W}{2}\right) \tag{1a}$$

$$A = L \times W + \frac{\pi W^2}{4} \tag{1b}$$

$$\left(\frac{-\pi}{4}\right) W^2 + \left(\frac{P}{2}\right) W - A = 0 \tag{1c}$$

The second algorithm (Eq. 2; Hall 2012)– enabled the estimation of a quantity, θ_{AV}, reflective of the linear persistence of amyloid fibres (Adamcik and Mezzenga 2011; Cantor and Schimmel 1980), defined as the average absolute deviation of the piecewise projection vector I, measured over a

[2] Differential diffraction is dependent upon the relative rotation of the fibre along the cylindrical coordinate.

[3] For a reversible process considered in the thermodynamic limit, this can be a consequence of the statistical/entropic factors associated with a multi-step reaction pathway (Hall and Minton 2002, 2004). For both reversible and irreversible processes considered in the kinetic limit, heterogeneity in reaction products may also be caused by spatial fluctuations in concentration, molecular fluctuations in configuration and/or energetic fluctuations brought about through collisions with the solvent.

[4] Lateral association of fibres is an example of a hierarchical mode of aggregation, in which multiple thinner fibres (sometimes termed protofibrils) laterally self-associate to form thicker fibres designated as mature fibres (Makin and Serpell 2002). Other types of hierarchical aggregation products are also observed (Krebs et al. 2004; Ridgley and Barone 2013).

[5] Mesoscopic physics is concerned with materials of an intermediate length, i.e. from the nanometer to the micrometer range, which roughly spans the experimental space between the atomic and the macroscopic. Mesoscopic models do not display atomic/molecular features but tend to be based on approximate regular structures, such as spheres and hard convex/cylindrical bodies.

[6] Indirect estimates of aggregate shape can also be made from model-based deconvolution of hydrodynamic (Lomakin et al. 1996; (Rogers et al. 2005) or molecular weight measurements (Pallitto and Murphy 2001; Nichols et al. 2002).

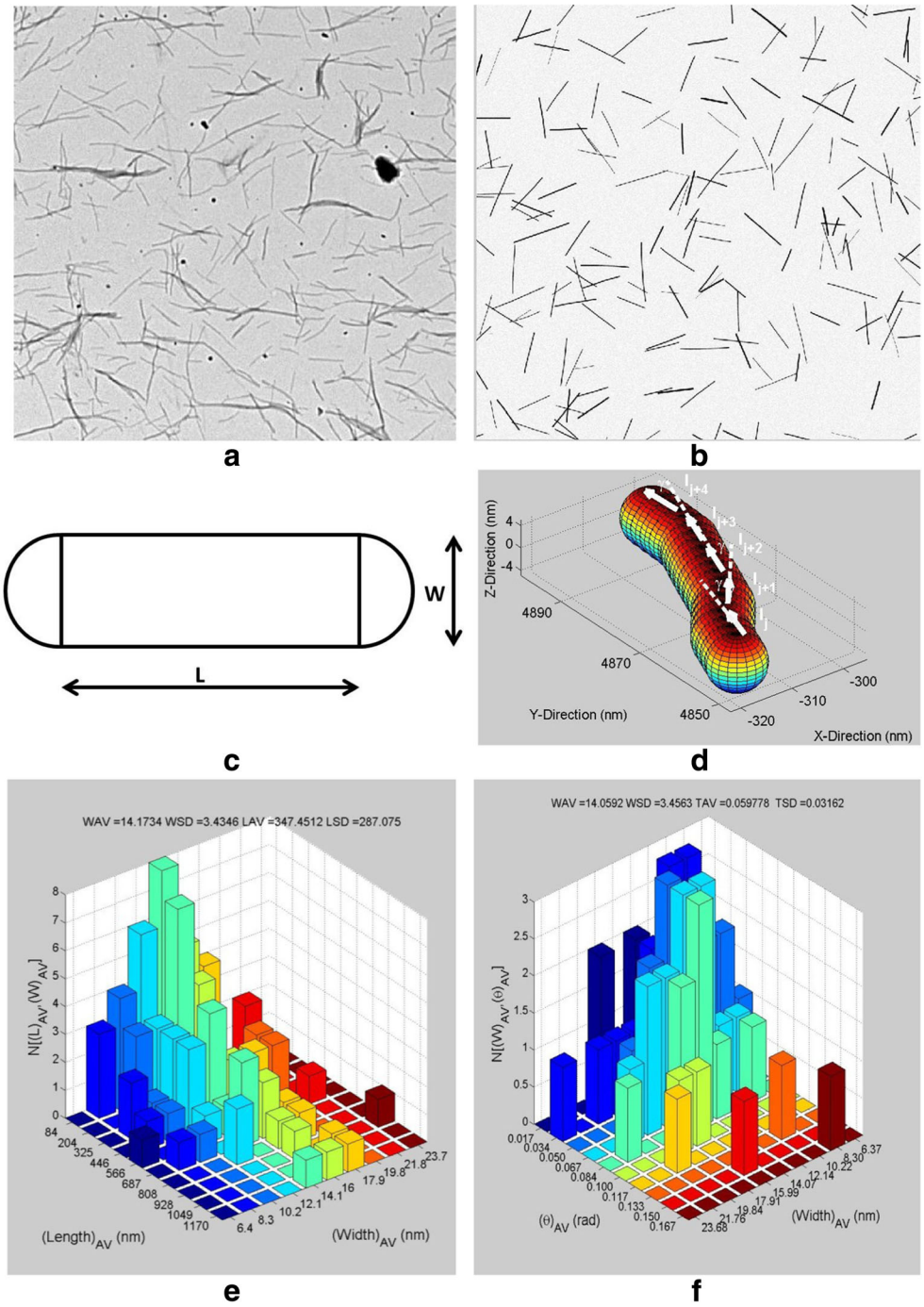


Fig. 2 Ultramicroscopy-based analysis of protein aggregates can provide the necessary mesoscopic-level structural information for estimating turbidity via the methods outlined in the text of this review. **a** Typical experimental transmission electron microscopy (TEM) image of amyloid fibres (made from pig insulin at pH 3.0 and 60 °C, recorded at 6000× magnification (adapted, with permission, from Fig. 9 of Hall 2012). **b** Example of a pseudo-TEM image generated using the Amyloid Distribution Measurement (ADM) software useful for calibrating and testing image analysis routines and designing better ultramicroscope experiments (adapted, with permission, from Fig. 2 of Hall 2012). **c** Mesoscopic representation of fibre by a sphero-cylinder of variable internal length (L) and width (W) (adapted, with permission, from Fig. 8 of Hall 2012). **d** Average angle of deviation (θ_{av}) for an individual fibre as determined by Hall et al. (2016a) using successive calculation of the dot product between projection vectors that trace along the backbone of the amyloid fibre (adapted, with permission, from Fig. 3a of Hall 2012). **e** Analysis of simulated TEM data yielding two-dimensional histograms of length and width (adapted, with permission, from Fig. 12 of Hall 2012). **f** Analysis of simulated TEM data yielding two-dimensional histogram of width and average deviation (adapted, with permission, from Fig. 12 of Hall 2012)

characteristic distance, *l*, selected as the fibre width[7] (Fig. 2d).

$$\theta_{AV} = \frac{1}{(N-1)} \sum_{j=2}^{N} \left| \cos^{-1} \left(\frac{\mathbf{I}_j . \mathbf{I}_{j-1}}{l_j . l_{j-1}} \right) \right| \quad (2)$$

As shown in Fig. 2e, f, the analytical software developed by Hall (2012) is able to decompose fibre images into two-dimensional histograms of fibre length versus width or width versus average deviation. This approach was later used to analyze TEM images of size exclusion chromatography-purified amyloid fibres, thereby facilitating development of a quantitative theory of fibre elution by that technique (Hall and Huang 2012).

An extension to this geometrical description, useful for modelling aggregates in solution (Hall et al. 2016a), involves representing aggregate geometry in terms of similarly limited shape possibilities, along with an extra variable relating to the internal volume packing fraction. Using this approach, amorphous, crystalline and fibrous protein aggregates can all be represented (Fig. 3). In this model, a protein aggregate composed of i monomers is defined by three properties, the molecular weight, M_i, the shape, S_i and the volume trace $(V_i)_{TRACE}$ (Eq. 3a, b, c). With regards to the shape, S_i, protein aggregates are treated as either arbitrarily diffuse rods, defined by a trace length L_i and a trace radius R_i, or arbitrarily diffuse spheres, characterized solely by R_i.

$$M_i = iM_1 \quad (3a)$$

$$(V_i)_{TRACE} = \frac{iM_1 \upsilon}{\alpha_i N_A} \quad (3b)$$

$$S_i = \begin{cases} \text{rod;} & L_i = (V_i)_{TRACE} \Big/ (\pi R_i^2) \\ \text{sphere;} & R_i = \sqrt[3]{3 \ (V_i)_{TRACE} \Big/ (4\pi)} \end{cases} \quad (3c)$$

In Eq. 3, M_1 describes the monomer molecular weight, N_A is Avogadro's number, α_i is the fraction of the trace volume occupied by protein (Fig. 3b) and υ refers to protein partial specific volume[8] (Lee et al. 2009). As noted (Hall et al. 2016a), defining α_I in the manner outlined by Eq. Eq. 3a, b, c allows it to be used to parameterize the transition between compact and diffuse aggregate states ($\alpha_{DIFFUSE} < \alpha_{COMPACT} \leq 1$) such that a higher value of α would be more appropriate for crystal-like aggregates whereas a lower value would describe a less dense amorphous[9] aggregate (Bennett 1972; Zurdo et al. 2001).

In principle, the geometrical information contained within an ultramicroscope image can be used to model the distribution of fibres within the solution from which it was generated. Although some research groups have made great strides forward (Arosio et al. 2012; Hall and Huang 2012; Lomakin et al. 1996; Rogers et al. 2005), this process is as yet a not fully realized proposition[10]. Here we take the liberty of pointing out how fibre shape parameters, derived from analysis of the ultramicroscope images, can be used to define a fibre trace volume. In conjunction with an assumed fractional volume packing, α_i, the fibre molecular weight and degree of polymerization can be estimated from Eq. 4a, b, c.

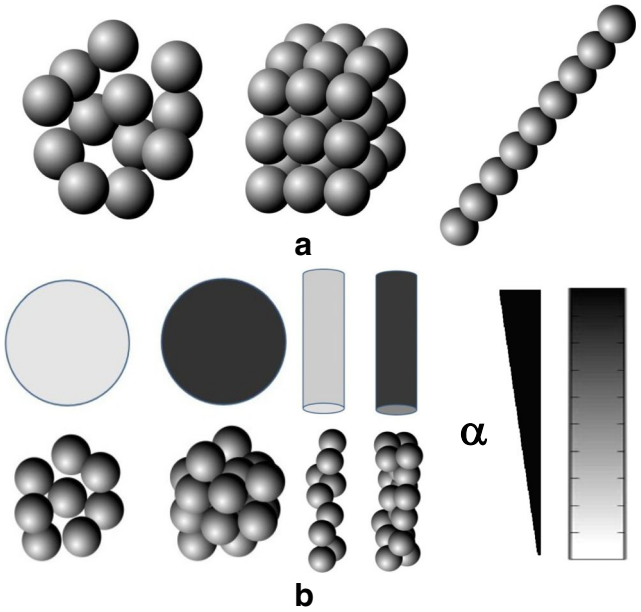


Fig. 3 Coarse structural models of aggregates. **a** Schematic describing coarse-grained conceptualization of bonding arrangements seen in various types of protein aggregate corresponding to amorphous (*left*), crystalline (*middle*) and fibrous (*right*) structures. **b** Schematic describing mesoscopic structural groupings of aggregates as either rod-like or spherical with assignment of a volume packing fraction, defined by the parameter α, such that a *darker colour* represents a greater fractional occupancy of the aggregate trace volume by protein, i.e. a greater internal density (schematic is adapted, with permission, from Fig. 1 of Hall et al. 2016a, b)

$$(V_i)_{TRACE} = \begin{cases} \frac{4}{3}\pi R_i^3 & \text{- for a sphere} \\ L_i \pi R_i^2 & \text{- for a cylinder} \end{cases} \quad (4a)$$

$$M_i = (V_i)_{TRACE} \alpha_i N_A \Big/ \upsilon \quad (4b)$$

$$i = M_i / M_1 \quad (4c)$$

To probe macroscopic-level phenomenon, one must be able to infer the concentration distribution of aggregate in solution based on knowledge of the number of adsorbed fibres, N, possessing properties within the discrete limits set by the element of a histogram[11]. For some techniques, such as

[7] In the preparation of our review we noticed that the corresponding equation of Hall (2012) has a typographical error with regard to bracket placement.

[8] To a first approximation υ can be taken as having a constant value of ~0.73 ml/g although Lee and co-workers (Lee et al. 2009) have shown that this value may decrease in the amyloid state, i.e. amyloid may be more tightly packed than normal globular proteins.

[9] Amorphous means an aggregate in which the monomeric units constituting the aggregates possess no regular internal structure (Bennett 1972).

[10] In a similar fashion, most solid state NMR structures require determination of an average density per unit length by scanning TEM experiment (e.g. Petkova et al. 2005).

[11] As a simple example, we might consider the histogram limits as referring to aggregate degree i, such that we consider the bin elements defined from $i \rightarrow i + \Delta i$.

light scattering (Lomakin et al. 1996; Nichols et al. 2002) or analytical ultracentrifugation (Binger et al. 2008), in which the total signal intensity is defined by the solution distribution, deconvolution can be attempted directly. However, for distributions inferred from ultramicroscope images, either an internal standard (Kirschner et al. 1975) or an independent measure of the total mass concentration of aggregate in solution, $c_{aggregate}$, is required to determine absolute number densities (with $c_{aggregate}$ defined as the total concentration of protein in an oligomeric form having a polymer degree ≥ 2). As shown previously (Borgia et al. 2013; O'Nuallain et al. 2006), a value of $c_{aggregate}$ can be determined by pelleting or filtration assay, with concomitant spectrophotometric analysis of the supernatant. With this information available, the ultramicroscope image data can be converted into a solution distribution on the assumption that the derived distribution is a true representation of the solution state (Eq.5a, b).

$$C(i \rightarrow i+\Delta i) = \frac{N(i \rightarrow i+\Delta i)}{\sum_{j=2}^{Z} N(j+\Delta j).\left(j+\frac{\Delta j}{2}\right)}.c_{aggregate} \quad (5a)$$

$$C\left(i+\frac{\Delta i}{2}\right) = \frac{C(i \rightarrow i+\Delta i)}{\Delta i} \quad (5b)$$

Equation 5a, b provides the means for inferring the solution aggregate distribution from ultramicroscope-derived histograms. Joining a continuous line between discrete concentration estimates (such as that provided by Eq. 5b) allows, in principle, for realization of the form of the fibre distribution in solution. In the next section, we examine the literature concerned with the estimation of the light scattering properties from such a protein aggregate distribution, summarizing germane concepts into a set of equations capable of directly transforming aggregate distributions into turbidity (at a given wavelength and path length).

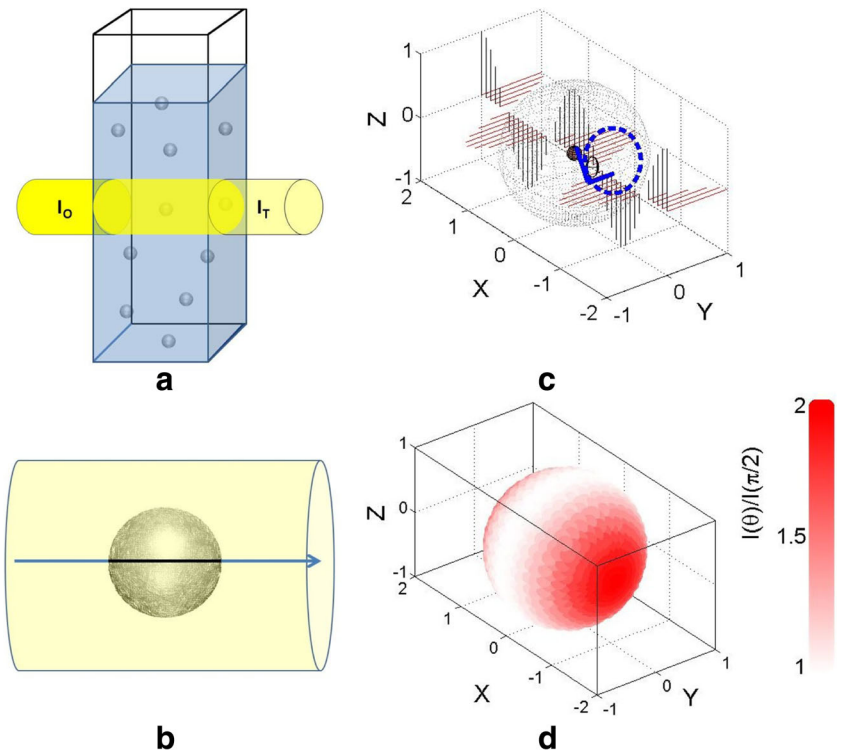


Fig. 4 Principles of light scattering. **a** Schematic describing the transmission-based measurement of excess solution turbidity of protein aggregates in which the transmitted light intensity (I_T) is measured in relation to the incident light intensity (I_0) using a standard spectrophotometer (or plate reader). **b** Ray diagram of the encounter between light and the scattering particle in solution. **c** Simplified schematic of a general goniometric scattering experiment for non-polarized light (although the light wave shown has only one polarization!). Scattering intensity for Rayleigh-type scattering is equivalent when recorded at any point on a sphere (centre located at the scattering particle) defined by the radius (r) and the angle θ, whereby θ is defined as the sub-apex of the spherical solid angle measured from the forward scattering direction (adapted, with permission, from Fig. 2a of Hall et al. 2016a, b). **d** Colour plot indicating the scattering intensity (normalized relative to the scattering recorded at right angles to the incident beam) as a function of the recording angle θ, with the system conforming to limiting Rayleigh scattering conditions described in **c** (adapted, with permission, from Fig. 2b of Hall et al. 2016a, b)

Table 1 Values of F and Q for the three size regimes and two shape types considered

Approximate size range (for which the description is valid)	Idealized turbidity per unit molecular concentration[a] (F_i)	Transmittance form factor[b] (Q_i)
Rayleigh $0 \leq \langle R_i \rangle \leq \lambda/20$	$F_i = \frac{24\left(M_1^2 i^2 v_i^2 / \left[\alpha_i^2 N_A^2\right]\right)}{\lambda^4}\left(\frac{m_i^2-1}{m_i^2+2}\right)^2$	$Q_i = 1$
Rayleigh-Gans-Debye $0 \leq \langle R_i \rangle \leq \lambda/2$	$F_i = \frac{24\left(M_1^2 i^2 v_i^2 / \left[\alpha_i^2 N_A^2\right]\right)}{\lambda^4}\left(\frac{m_i^2-1}{m_i^2+2}\right)^2$	$Q_i = \int_0^{\pi} P(\theta,\lambda).\left(1+\cos^2\theta\right).\sin\theta.d\theta$ Where, $P(\theta,\lambda) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sin\left(h.d_{ij}\right)}{\left(h.d_{ij}\right)}$ $h = \frac{4\pi n \sin(\theta/2)}{\lambda}$ $d_{ij} = \left\lvert \overrightarrow{r}_i - \overrightarrow{r}_j \right\rvert$ For $x_1 = R_i n_o/\lambda$ and $x_2 = L_i n_o/\lambda$ When $x_2 < 1$; $Q_i(sphere) = 1-0.955\left(1-e^{-6.48x_1}\right)^{2.40}$ $Q_i(rod) = Q_i(sphere)\ (x_1)$ $\left[1-0.955\left(1-e^{-1.08x_2}\right)^{1.275}\right]$ When $x_2 \geq 1$; $Q_i(rod) = \frac{Q_{i(sphere)}(x_1)}{2.4x_2^{0.95}}$
Anomalous diffraction approximation[b] $2\lambda \leq \langle R_i \rangle \leq 15\lambda$	$F_i = \frac{24\left(M_1^2 i^2 v_i^2 / \left[\alpha_i^2 N_A^2\right]\right)}{\lambda^4}\left(\frac{m_i^2-1}{m_i^2+2}\right)^2$	$Q_i(rod) = \frac{Q_{i(sphere)}(x_1)}{2.4x_2^{0.95}}$ $Q_i(sphere) = \frac{\left[2-\left(\frac{4}{\rho_i}\right)\sin(\rho_i)+\left(\frac{4}{\rho_i^2}\right)(1-\cos(\rho_i))\right]}{\left((F_i)_{RAYLEIGH}/\pi.\left\langle R_i^2\right\rangle\right)}$ where, $\rho_i = 4\pi R_i(m_i-1)/\lambda$

All expressions are particular for a non-polarized light source and randomly oriented aggregate. All terms are defined in Eq. 3a, b, c; Eq. 4a, b, c; Eq. 6a, b, c, d e; Eq. 7a, b, c; Eq. 9

[a] As defined in Eq. 9 of the text

[b] The anomalous diffraction equation is a good approximation of the Mie scattering description for spheres (Hergert and Wriedt 2012; Kerker 2013; Mie 1908). Arguments have been advanced (Hall et al. 2016a) to suggest that the expression for given for Q_i(rod) would retain validity in the Mie scattering regime (for a discussion see Cassasa, 1955; Bishop 1989; Buitenhuis et al. 1994; Liu et al. 1998)

(ii) Turbidity of aggregates in the visible region

Turbidity describes the attenuation of the incident beam by light scattering (Bohren and Huffman 2008; Elimelech et al. 2013) and thus can be evaluated either directly, via measurement of the loss of intensity by transmission measurement (Fig. 4a, b), or indirectly, by integration of the angle-dependent scattering at a fixed distance (goniometric static light scattering) (Doty and Steiner 1950; Wyatt 2014) (Fig. 4c, d). The relatively straightforward nature of the transmission measurement, requiring only a spectrophotometer or plate reader, has encouraged adoption of the turbidimetric method in the absence of more specialist light scattering equipment (Andreu and Timasheff 1986; Gaskin et al. 1974; Wyatt 2014). Coupled with ease of performance, the generality of light scattering (due to the lack of a requirement for an extrinsic label) has made transmission-based turbidimetric assays the default 'basic' standard for recording protein aggregation kinetics. Historically speaking, turbidity has been used to monitor the growth of a range of protein aggregation reactions, including helical fibre formation by sickle cell haemoglobin (Ferrone et al. 1985; Moody et al. 1996), cytoskeletal fibre formation (Voter and Erickson 1984; Wegner and Engel 1975; Gaskin et al. 1974), virus capsid formation (Tachibana et al. 1977), non-specific amorphous polymerization (Stranks et al. 2009; Ingham et al. 2011) and of course amyloid formation (Anzai et al. 2016; Dolado et al. 2005; Hatters et al. 2001; Necula et al. 2007; Sant'Anna et al. 2016). Although the turbidimetric procedure is a relatively non-demanding experiment to carry out, as with all scattering methods, the downside is that gaining an understanding of the underlying physics generally requires a familiarity with Maxwell's equations not often in simpatico with the background of those performing the work. It is partly towards this divergence that the next section is directed.

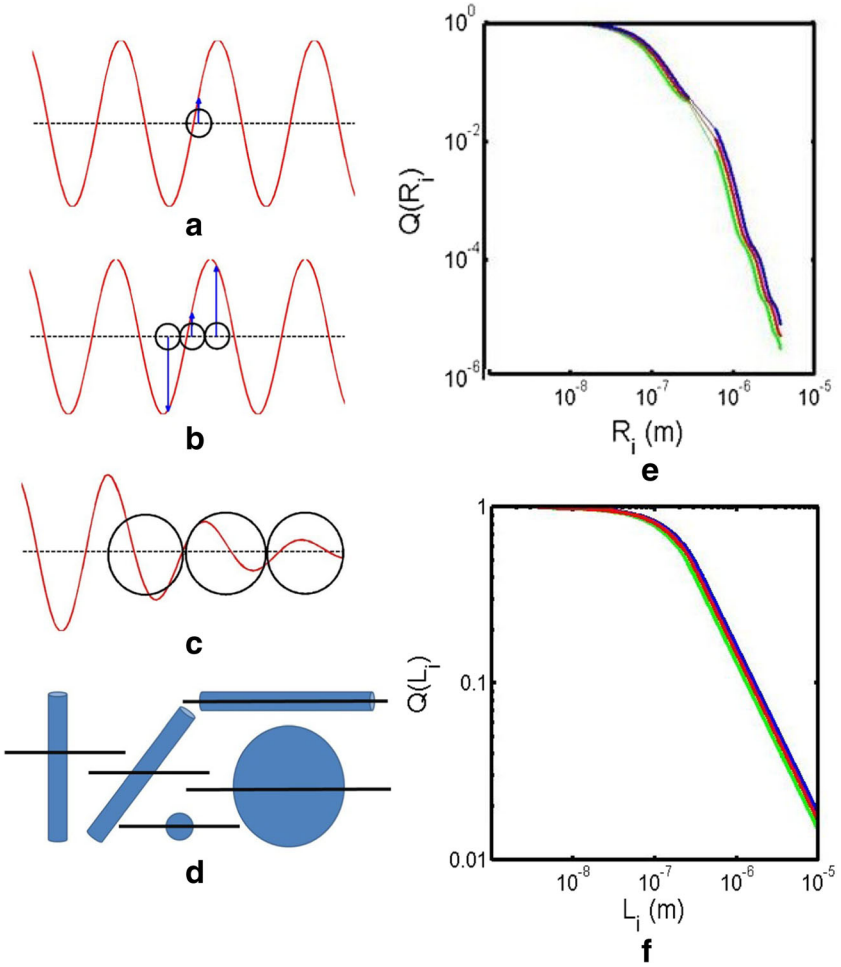


Fig. 5 Theoretical treatments of scattering. **a**–**c** Three general scattering regimes were considered by Hall et al. (2016a), namely **a** Rayleigh limit—where the scattering particle is small in relation to the wavelength of light [$<R_i> < \lambda/20$] (*red line* light wave, *blue arrow* position of the dependent electric field vector). **b** Rayleigh–Gans–Debye limit—where the particle can be reasonably large in relation to the wavelength of light at ~[$0 < <R_i> < \lambda/2$] such that it produces out-of-phase scattering at different centres of the particle but the light suffers no appreciable loss of intensity as it passes through the particle. **c** Mie scattering regime—where the particle is sufficiently large to both generate out-of-phase scattering and to perturb the intensity of the light as it passes through the aggregate. For the anomalous diffraction approximation of the Mie equation used by Hall et al. (2016a) this description is applicable over the size regime of ~[$2\lambda < <R_i> < 15\lambda$]. **d** Schematic highlighting the potential for orientation effects on both the out-of-phase scattering and loss of intensity complications accompanying increasing size and asymmetry of the aggregate. All quantitative descriptions described by Hall et al. (2016a) assume random orientation of the aggregate. **e** Continuous description of the transmittance form factor for a spherical aggregate [$Q(R_{SPHERE})$] at three different wavelengths (*blue line* 400 nm, *red line* 450 nm, *green line* 500 nm). Interpolation based on a polynomial description of spliced simulations from the three characteristic size regimes is shown in Table 1. **f** Continuous description of the transmittance form factor for rods [$Q(L_{ROD})$] over a large size regime for three different wavelengths (**a**–**d** adapted, with permission, from Fig. 3 of Hall et al. 2016a; **e**, **f** adapted, with permission, from Fig. 5 of Hall et al. 2016a)

The truest understanding of light scattering phenomena requires a quantum-level description of both the system and the light source (Chu 1974; Loudon 2000). However from the time of Rayleigh (Rayleigh 1899), theories based on classical electromagnetism (Bohren and Huffman 2008; Doty and Steiner 1950; Oster 1955; Penzkofer et al. 2007), coupled with regular shape approximations of the scattering bodies and continuum approximations of the solvent, have proven effective for extracting shape and molecular weight information from measurement of the angle-dependent intensity of the scattered light (Geiduschek and Holtzer 1958; Wyatt 2014). In the classical approach, light is considered to be a coincident, yet perpendicular pair of travelling electric and magnetic transverse field vectors, oscillating at a frequency f, over a wavelength λ (Bohren and Huffman 2008). The charge distribution associated with any element of matter in the path of the light beam is perturbed by these fields and caused to, itself, oscillate. For the case of elastic scattering, the oscillating (and thus accelerating) charge will produce another light wave of identical wavelength[12] (Bohren and

[12] Inelastic light scattering refers to the case where some energy is absorbed, donated or dissipated by the molecule with subsequent change in wavelength between incident and scattered radiation.

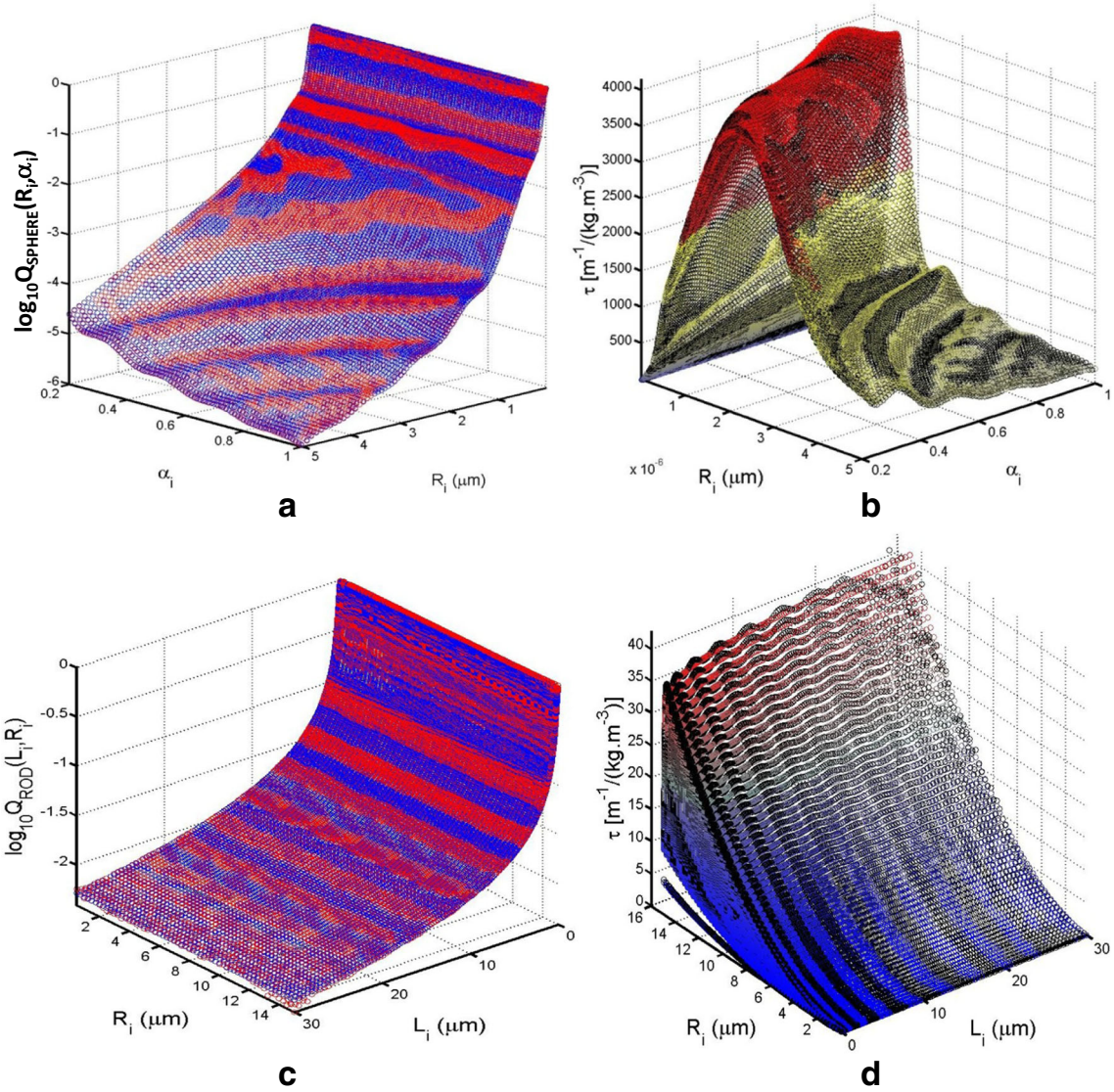


Fig. 6 Utilitarian approach developed by Hall et al. (2016a, b) for estimating turbidity. **a** Two-dimensional polynomial fit of simulated Q values for a sphere: fitted values were overlaid onto large sets of the base ten logarithm of Q calculated for a sphere of arbitrary packing fraction α_I and radius R_i, determined using the interpolation technique describe in Fig. 5e (at $\lambda = 400$ nm). **b** Specific turbidity (turbidity per kg/m^3 of aggregate) for a spherical protein aggregate of arbitrary α_i and R_i, calculated using the corresponding value of Q shown in **a**. Protein concentration and mass were respectively set at 1 mg/ml and 5000 g/mole. **c**, **d** Corresponding plots to **a** and **b**, respectively, but this time describing the case for cylindrical rods of arbitrary length and radius. Specific turbidity in **d** was calculated at the same concentration and mass of the protein monomer with a value of the specific fractional volume occupancy of $\alpha = 1.0$ (adapted, with permission, from Figs. 6, 7, 8 and 9 of Hall et al. 2016a, b)

Huffman 2008; Kerker 2013). In the late 19th century, Strutt (Rayleigh 1899) deduced the quantitative relationship between the characteristics of an illuminated particle and the scattering intensity measured at a set distance and direction (Eq. 6a; Fig. 4). That formulation was derived on the basis of a set of simplifying criteria specifying limiting dilution and small size for the scattering object relative to the wavelength of light (Fig 4a). Under these Rayleigh limiting conditions the total amount of light scattered away from the forward direction, I_S, can be calculated by spherical integration of the angle-specific scattering intensity, i(r,θ), whereby θ represents the forward scattering apical sub-angle of the solid angle and r describes the radial distance from the centre (Kerker 2013; Oster 1955) (Eq. 6b–e).

$$i(r,\theta) = \frac{i_0}{r^2}\left[\frac{9\pi^2 (V_i)^2_{TRACE}}{2\lambda^4}\left(\frac{m_i{}^2-1}{m_i{}^2+2}\right)^2\left(1+\cos^2\theta\right)\right] \quad (6a)$$

$$I_S = 2\pi r^2 \int_0^{\pi} i(r,\theta)\sin(\theta)d\theta \quad (6b)$$

$$I_0 = A_0 i_0 \quad (6c)$$

$$\frac{I_S}{I_0} = \frac{24\pi^3 (V_i)^2_{TRACE}}{\lambda^4} \left(\frac{{m_i}^2 - 1}{{m_i}^2 + 2}\right)^2 \left(\frac{1}{A_o}\right) \quad (6d)$$

$$\frac{I_S}{I_0} = LC_i \frac{24\pi^3 (V_i)^2_{TRACE}}{\lambda^4} \left(\frac{{m_i}^2 - 1}{{m_i}^2 + 2}\right)^2 \quad (6e)$$

In Eq. 6a–e, i_0 refers to the incident light intensity, A_0 to the cross-sectional area of illumination of the incident light, L to the optical path length of the transmission measurement and m_i to the relative refractive index of the aggregate (relative to the solvent). A numerical value of m_i can be calculated (Hall et al. 2016a) on the basis of knowledge of the solvent refractive index, n_0, the aggregate protein refractive increment, dn/dc_i, the fractional volume occupation by protein in the trace volume, α_i, and the partial specific volume, υ (Eq. 7a). The wavelength dependence of the refractive index and refractive increment can be determined using an empirical formula (Perlmann and Longsworth 1948) (Eq. 7b, c).

$$m_i = 1 + \alpha_i \left[\left(dn/dc\right)\left(1/\upsilon\right)\right]/n_0 \quad (7a)$$

$$n(\lambda) = 1.3403 \left[0.9922 + 2.31 \times 10^{-15}/\lambda^2\right] \quad (7b)$$

$$dn/dc_i(\lambda) = 0.19 \times 10^{-3} \left[0.925 + 2.2 \times 10^{-14}/\lambda^2\right] \quad (7c)$$

The Rayleigh scattering relationship, shown in Eq. 6a–e, is able to quantitatively account for the scattering of non-polarized light by a compact solute with average radius of less than one-twentieth of the wavelength of light ($R_i < \lambda/20$). In a standard spectrophotometer arrangement (Fig. 4a), the continual encounter of incident light with particles in its path leads to a length-dependent decrease in transmitted light intensity recorded at the detector (Kerker 2013). The description of how the intensity changes with position due to scattering can be formulated via Eq. 8a, in which turbidity, defined as τ, is the first-order decay constant of light intensity, I, with path length, L (Bohren and Huffman 2008; Oster 1955) Rearrangement and integration yields the central section of Eq. 8a, which in turn can be simplified by a series expansion to produce Eq. 8b. When the ratio of scattered light to incident is <0.1, the first term of the series expansion suffices (Eq. 8c) (Kerker 2013).

$$\tau = -\left(\frac{1}{I}\right)\frac{dI}{dL} = -\frac{1}{L}\log_e\left(1 - \frac{I_S}{I_0}\right) = 2.303(O.D.) \quad (8a)$$

$$\tau = \frac{1}{L}\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{I_S}{I_0}\right)^n \quad (8b)$$

$$\underset{(\lim\tau \to 0)}{\tau} \approx \frac{1}{L}\left(\frac{I_S}{I_0}\right) \quad (8c)$$

Insertion of Eq. 6e into Eq. 8a, b, c yields an expression which accurately describes the path length-corrected turbidity values of small particles at low concentrations, i.e. $R_i < \lambda/20$, $\tau_i \to 0$, $C_i \to 0$).

Although Eq. 8a is capable of predicting the turbidity of small compact particles in the dilute limit, it becomes less suitable as the particles increase either in size, concentration or complexity of their shape[13] (Bohren and Huffman 2008; Garcia-Lopez and Garcia-Rubio 2008; Garcia-Lopez et al. 2006; Hergert and Wriedt 2012; Kerker 2013). With specific regard to the size and shape of a particle, we note that deviation from the ideal Rayleigh case occurs for two reasons (Fig. 5):

1. Different regions of large aggregates will experience different phases of the incident light's electric field, thereby producing a complex superposition of the scattered light with a reduction in overall scattering intensity (Bohren and Huffman 2008; Geiduschek and Holtzer 1958) (Fig. 5b).
2. In the case of scattering from very large aggregates, the incident light will be demonstrably reduced in intensity as it travels through the aggregate, compounding the difficulty of accounting for any phase difference produced upon scattering (Elimelech et al. 2013; Kerker 2013) (Fig. 5c, d).

The Rayleigh–Gans–Debye (RGD) formalism (Debye 1947; Gans 1925; Zimm and Dandliker 1954) is a theoretical approach capable of tackling only the first of these two difficulties and is therefore applicable only to particles of averaged cross-sectional radius, $\langle R_i \rangle$, smaller than $\lambda/\{2n_o(\lambda)\}$ (Bohren and Huffman 2008). In RGD theory, total scattering intensity is calculated as the sum of the scattering from N discretized centres within the aggregate, on the assumption that the incident light intensity is constant throughout the aggregate (Fig. 6b). A quantity known as the particle form factor $P_i(\theta,\lambda)$ reflects the degree to which this type of internal interference, generated by effective phase difference, diminishes the scattering recorded for a real particle, $i(r,\theta)_{real}$, relative to that measured for an idealized scattering particle (same mass, but point-like dimensions), $i(r,\theta)_{ideal}$, such that $P_i(\theta,\lambda) =$

[13] As issues related to high concentration can be effectively solved experimentally by either using a smaller path length or by serial dilution of the solution prior to measurement, we focus in this review on the effects related to particle size and shape complexity. With this point made, we note that the dilute limit refers to the concentration range at which one may neglect significant re-scattering of the scattered light by other aggregate particles in solution back into the collimated detection pathway. This re-scattering effect by other molecules in solution is sometimes called external interference. We refer the reader to work describing the effect of external interference on scattering encountered in the high concentration limit based on virial expansions of the turbidity in terms of solute concentration (Dhont 1983; Minton 2007). Such factors will be important to consider if the turbidimetric technique is be used to study aggregation in near cell-like conditions such as those employed in macromolecular crowding studies (Hall and Minton, 2003; Hall, 2002; Hall and Dobson, 2006; Hall 2006).

$i(r,\theta)_{real}/i(r,\theta)_{ideal}$ (Doty and Steiner 1950; Geiduschek and Holtzer 1958) (Table 1). The equivalent term for the transmittance measurement, known as the transmittance form factor, Q_i, can be directly obtained from $P_i(\theta,\lambda)$ upon integration to account for all possible orientations of the aggregate in relation to all possible polarizations of the light (Table 1). Within the limits of the approximations inherent in their construction, these form factors can be calculated for any arbitrary shape based on knowledge of the centre-to-centre distances of the discretized scattering centres through use of the Debye equation (Table 1) (Bohren and Huffman 2008; Debye 1947).

An alternative approach to the Debye approximation, developed by Gustav Mie for particles of arbitrary size and shape (Hergert and Wriedt 2012; Mie 1908), accounts for both the decrease in light intensity as it passes through the aggregate and the phase difference in scattered light intensity generated by scattering from widely separated regions of the aggregate molecule (Hergert and Wriedt 2012; Kerker 2013) (Fig. 5c, d). The anomalous diffraction (AD) equation (Table 1) developed by Van de Hulst represents a very accurate simplifying approximation to the Mie scattering equations for aggregates having spherical geometry (Elimelech et al. 2013; van de Hulst 1957). The AD approximation for spheres retains validity over the size regime $2\lambda \leq \langle R_i \rangle \leq 15\lambda$ for systems having a relative refractive index, m_i, of <1.3 (van de Hulst 1957). Importantly, this last requirement represents nearly all conceivable cases of proteins aggregating in standard aqueous buffers. Relatively simple approximate forms of Mie-type solutions for other shapes, such as cylindrical rods, have also been developed and compared to 'exact' calculations made using finite element numerical techniques performed over a large range of particle sizes relative to the wavelength of light employed (Bishop 1989; Buitenhuis et al. 1994; Liu et al. 1998) (Table 1).

Based on a recasting of the general turbidity expression into an equation involving three parts, Hall and co-workers (Hall et al. 2016a) laid the foundation for producing an empirical interpolation of the transmittance particle form factor Q_i over a wide range of sizes and shapes suitable for describing amyloid growth (Eq. 9) (Fig. 5e, f).

$$\underset{(\lim \tau \to 0)}{\tau_i} = C_i F_i Q_i \tag{9}$$

As formulated by Eq. 9, τ_i, the turbidity at limiting dilution, is composed of three terms, namely C_i, F_i and Q_i, whereby C_i is the scattering particle concentration (units of molecules m^{-3}), F_i is the idealized turbidity per unit molecular concentration that would be generated if the particle scattered light as a point mass (units: m^2 $molecule^{-1}$) and Q_i is the unitless transmittance form factor discussed above. Hall and co-workers (Hall et al. 2016a) considered the appropriate functionalization of Eq. 9 for two general shapes, a rod and a sphere of arbitrary internal density, over the three particle size regimes of scattering described in Table 1.

Although values of Q_i and F_i exist for other shapes (see Bohren and Huffman 2008; Moody et al. 1996), their evaluation from a turbidity signal is problematic, representing, as it does, a type of inverse problem (Hall and Minton 2005; Mroczka and Szczuczynski 2010; Shmakov 2014). Given that a rod and a sphere respectively demonstrate the least and most scattering potential of any regular body, Hall et al. (2016a) suggested that an experimental signal, presumed to reflect amyloid growth, might be empirically decomposed into amyloid (rod-like) and non-amyloid aggregate (assumed spherical) structures. With this basic premise they went on to provide a continuous description of F and Q over a size range spanning the point scattering ($R < \lambda/20$) to Mie regime ($2\lambda < R < 15\lambda$) in the form of two-dimensional polynomial interpolants for spheres (Eq. 10a) (Fig. 6a) and rods[14] (Eq. 10b) (Fig. 6c).

$$\log_{10} Q_i(R_i, \alpha_i) = \sum_{j=0}^{N} \sum_{k=0}^{N} u_{j,k} (R_i)^j (\alpha_i)^k \tag{10a}$$

$$\log_{10} Q_i(L_i, R_i) = \sum_{j=0}^{N} \sum_{k=0}^{N} w_{j,k} (L_i)^j (R_i)^k \tag{10b}$$

The coefficients for these two polynomials were evaluated at a series of different wavelengths. With the value of F common to all three different size regimes (Table 1), the differences in turbidity due to shape can be seen to be directly defined by the transmittance particle form factor (Bohren and Huffman 2008; Kerker 2013). Previously calculated numerical examples (Hall et al. 2016a) describing the specific turbidity (i.e. turbidity per set mass of scattering component) produced by a spherical aggregate of arbitrary internal fractional occupancy, and a cylindrical rod for which, $\alpha_i = 1$, are reproduced in Fig. 6b, d. In the case of multiple aggregate species, the turbidity for a solution of aggregates of different sizes and shapes can be calculated as the sum of the contributions to turbidity from each particle (Eq. 11) (Bohren and Huffman 2008; Kerker 2013).

$$\underset{(\lim \tau \to 0)}{\tau} = \sum_{i=1}^{N} \tau_i \tag{11}$$

The aim of this review is to consider the kinetics of aggregate formation as monitored by the turbidity assay. In the section Consensus kinetic models of aggregate growth, we review different consensus kinetic behaviours of amyloid/aggregate

[14] Two important points to note here are (1) Q_{SPHERE} in the Mie region was calculated using the AD approximation and (2) Q_{ROD} in the Mie region was calculated by extending the asymptotic form calculated by the RGD approach. Although not exact, this assumption has been calculated to induce an error of <15 % in the estimation of scattering of non-polarized light from randomly oriented fibres extending in length beyond the RGD limit up to an (effectively) infinite length (Bishop 1989; (Buitenhuis et al. 1994).

formation, summarizing them into a set of limited basis models (Table 2). Together with the geometric and turbidimetric transforms reviewed in the previous sections, these consensus kinetic models are used to simulate characteristic turbidity signatures associated with particular mechanistic sub-types (Ghosh et al. 2010; Hall and Edskes 2012; Kashchiev 2015).

(iii) Consensus kinetic models of aggregate growth

An extremely general description of non-specific cluster formation was made over 150 years ago by von Smoluchowski (1916, 1917). In that approach, for a single isomeric state[15], the rate of formation of an aggregate species is given by the total balance of all possible formation and breakage events (Eq. 12a, b),

$$\left(\frac{dC_k}{dt}\right)_{growth} = \sum_{i=1}^{k-1} f_{i,k-i}C_iC_{k-i} + \sum_{j=k+1}^{z} b_{k,j-k}C_j \qquad (12a)$$

$$\left(\frac{dC_k}{dt}\right)_{loss} = -\sum_{i=1}^{z} f_{k,i}C_kC_i - \sum_{j=1}^{k-1} b_{j,k-j}C_k \qquad (12b)$$

Within this cluster growth formalism, the concentration of an aggregate composed of k monomers is denoted by C_k. The chemical rate constant[16] for formation of a species of size k from two smaller species j and k−j is denoted by $f_{j,k-j}$. Similarly, the chemical rate constant describing the breakage of a species of size k into two species, j and k−j, is denoted by $b_{k-j,j}$. With an appropriate choice of rate constants, the Smoluchowski cluster/condensation rate model can be used to describe aggregation processes of great complexity (Aldous 1999). However, despite this potential for diversity, the kinetics of amyloid formation have repeatedly been shown to comport to a subset of the possible model space, defined by Eq. 12a, b, with this subset known as nucleated growth[17] (Jarrett and Lansbury 1992; Lomakin et al. 1996; Masel et al. 1999; Wetzel 2006) (Fig. 7). In its most general form this mode of aggregation involves the introduction, or slow production, of a structural nucleus within a pool of monomeric proteins[18] which are themselves capable of joining to the nucleus and adopting the template structure encoded by it (Jarrett and Lansbury 1992; Petkova et al. 2005; Wetzel 2006). As the amyloid reaction proceeds, fibres can break apart (Hall and Edskes 2009, 2012, 2004; Xue et al. 2008) or join together (Pallitto and Murphy 2001; Binger et al. 2008, Michaels and Knowles 2014). Joining of fibres may occur in either an end-to-end fashion (Binger et al. 2008), a lateral side-to-side manner (Pallitto and Murphy 2001; Nichols et al. 2002; Kanno et al. 2005) or by a random process (Mishra et al. 2011) to produce a low-density matrix. Irrespective of their exact form, idealized nucleation–growth models typically display sigmoidal-type association kinetics if the nucleation step is allowed to occur spontaneously (Jarrett and Lansbury 1992), or exponential-type association kinetics without a lag-phase, if nucleation is bypassed by seeding the system with template (Paravastu et al. 2009; Wetzel 2006). In accordance with Fig. 7, the kinetics can be parameterized with constants describing the nucleation, growth and asymptotic stages of the reaction (Hall et al. 2016b). Alternatively, the kinetic traces may be fitted to equations derived from mechanistic models of amyloid growth, to yield the most parsimonious set of rate constant parameters (Pallitto and Murphy, 2001; Morris et al. 2009; Smith et al. 2006).

In the following subsections we discuss a number of potential variants in the nucleated growth model, summarizing the basic kinetic behaviour with an appropriate chemical rate equation (Table 2). In the formulation and discussion of kinetic data reflecting the various limiting cases of aggregation behavior, we make two assumptions:

(1) Polymer distributions are approximated by their weight average. A more formal description of this approximation for the weight average degree of polymerization, is shown by Eq. 13:

$$\langle i \rangle = \frac{\sum_{i=2}^{z} C_i.i^2}{\sum_{i=2}^{z} C_i.i} \qquad (13)$$

.

(2) In discussing either breakage, competitive growth or fibre joining, a separation of time scales for the monomer/polymer mass and polymer mass/polymer distribution time scales will often be invoked (Bridstrup and Yuan 2016; Hall and Minton 2004). A more formal statement

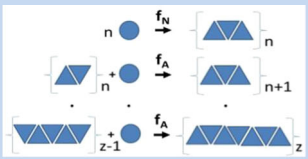
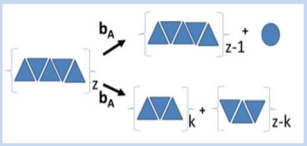
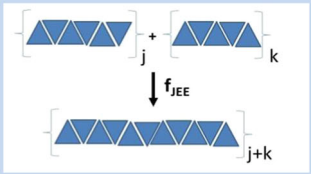
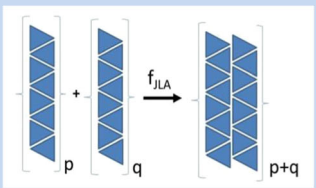
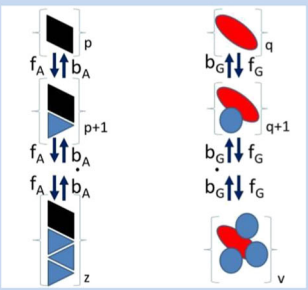
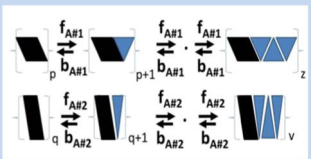
[15] If aggregates are composed of multiple isomeric forms the rate constants become dependent upon the isomeric state and a new dimension must be introduced into the specification of the aggregate and rate constants.

[16] Corrected for statistical and stoichiometric factors.

[17] Also known as templated-growth, nucleated-crystal growth or helical polymerization (Oosawa and Asakura 1975).

[18] As linear polymers of amino acids, proteins potentially have available to them a very great number of possible internal configurations of the polymer chain. Internal bonding preferences for solvent and self often limit this range of possibilities, leading to a single structural state known as the folded state. Under conditions of elevated temperature, the presence of a chemical denaturant or other structure-deforming species, the protein can be induced to unfold.

Table 2 Kinetic equations for six different cases of amyloid aggregation

Kinetic Model	Kinetic Rate Equations[1]	Literature
Irreversible Nucleated Growth		
	$\frac{dC_N}{dt} = f_N(C_M)^n - f_A C_N C_M$ $\frac{dC_A}{dt} = f_A C_N C_M$ $\frac{dC_{M\to A}}{dt} = (n+1) f_A C_N C_M + f_A C_A C_M$	Oosawa and Asakura, 1975. Hall, 2003; Hall et al. 2016b. Powers and Powers, 2004.
Fibre Breakage		
	$\frac{dC_A}{dt} = f_N(C_M)^n + b_A(C_{M\to A} - (n+1).C_A)$ $\frac{dC_{M\to A}}{dt} = n.f_N(C_M)^n + f_A C_M C_A - n b_A C_A$	Masel et al. 1999 Hall and Edskes, 2004, 2009, 2012. Smith et al. 2005. Kashschiev, 2015.
Fibre End to End Association		
	$\frac{dC_N}{dt} = f_N(C_M)^n - f_A C_N C_M$ $\frac{dC_A}{dt} = f_A C_N C_M - f_{JEE}(C_A)^2$ $\frac{dC_{M\to A}}{dt} = (n+1) f_A C_N C_M + f_A C_A C_M$	Hill, 1983 Pallitto and Murphy, 2001 Binger et al. 2008
Fibre Lateral Association		
	$\frac{dC_N}{dt} = f_N(C_M)^n - f_A C_N C_M$ $\frac{dC_A}{dt} = f_A C_N C_M - 2 f_{JLA} C_A{}^2$ $\frac{dC_{A2}}{dt} = f_{JLA} C_A{}^2$ $\frac{dC_{M\to A}}{dt} = (n+1) f_A C_N C_M + f_A [C_A + 2C_{A_2}] C_M$	Pallitto and Murphy, 2001 Ghosh et al. 2010.
Amyloid vs Amorphous Competition[2]		
	$(C_A)_t = (C_A)_{t=0};\ (C_{M\to A})_{t=0} = x$ $(C_G)_t = (C_G)_{t=0};\ (C_{M\to G})_{t=0} = y$ If $C_{M\to A} > x \Rightarrow \frac{dC_{M\to A}}{dt} = f_A C_A C_M - b_A C_A$ If $C_{M\to A} \le x \Rightarrow \frac{dC_{M\to A}}{dt} = f_A C_A C_M$ If $C_{M\to G} > y \Rightarrow \frac{dC_{M\to G}}{dt} = f_G C_G C_M - b_G C_G$ If $C_{M\to G} \le y \Rightarrow \frac{dC_{M\to G}}{dt} = f_G C_G C_M$	Naiki et al. 1997 Stranks et al. 2009 Hall et al. 2015 Hall and Minton, 2004, 2005 Hall, 2003.
Amyloid vs Amyloid Competition[2]		
	As for case above but the subscripts A and G should be replaced with the symbols A#1 and A#2 representing two different kinds of amyloid growth from the same monomeric form.	Naiki et al. 1997 Hall et al. 2015 Hall and Minton, 2004 Hall, 2003.

[1] Monomer was calculated via conservation of mass arguments with terms as appropriate

[2] Seeds are regarded as fixed i.e. non-dissociable

of this conceptual tool is given by the following mechanistic approximation [Eq. 14]:

$$monomer \underset{\leftarrow}{\overset{fast}{\rightarrow}} polymer\ mass \underset{\leftarrow}{\overset{\text{fast or slow}}{\rightarrow}} polymer\ distribution \qquad (14)$$

With regard to the first assumption (described in Eq. 13), we note that a few researchers (Arosio et al. 2012; Ghosh et al. 2010; Hall and Edskes 2004, 2009; Hall et al. 2015) have developed methods for simulating amyloid kinetics that yield full distribution information as a function of time. Although these methods are more informative than the approximation adopted by Eq. 13, they are also necessarily more complex. Due to the focus of this review being on the transformation of the distribution by turbidimetric assay procedures, we have opted to make a trade-off: a level of exactness for ease of discourse[19]. In the cases where no literature-derived chemical rate equation exists (or alternatively no literature derivation possessing a relatively transparent formulation exists), we have cited the relevant literature but put forth an approximate relation.

Of all possible permutations available to Eq. 12a, b, the following six limiting cases of templated growth are regarded as having principal importance in this review:

- Irreversible nucleation–growth model
- Nucleation growth with fibre breakage
- Fibre end-to-end association
- Fibre lateral association
- Amyloid versus amorphous competition
- Amyloid versus amyloid competition

Functional kinetic models for each of these limiting cases have been provided (Table 2). Rate models are presented in differential equation format due to the straightforward manner in which ordinary differential equations (ODE) can be related to mechanism by inspection[20]. Figs. 8, 9, 10, 11, 12, 13, and 14 show the resultant chemical kinetics and turbidity transformation for each case. We discuss both the kinetic behavior and the turbidity transformation (effected through application of Eqs. 9–11 to the simulated chemical data) to each case in turn.

Irreversible nucleation-growth

Oosawa and colleagues developed the first nucleation–growth-type kinetic models to describe the polymerization of helical fibers formed by the cytoskeletal protein actin (Oosawa and Asakura 1975; Oosawa and Kasai 1962). Despite potential structural and mechanistic differences[21], others have adapted these Oosawa class of kinetic models to

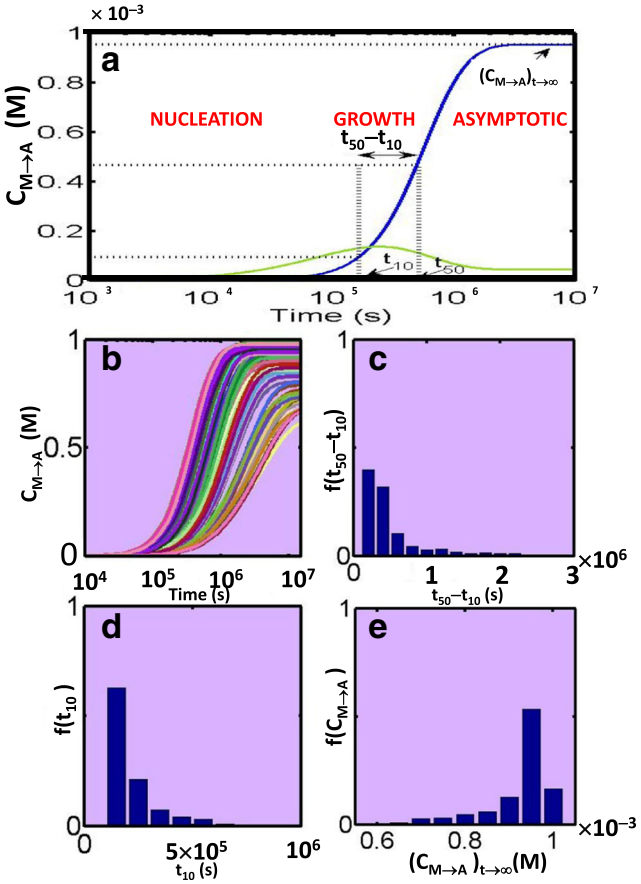


Fig. 7 Schematic of amyloid kinetics. **a** Characteristic features of amyloid nucleation–growth polymerization kinetics include a characteristic lag/nucleation phase, a steep growth phase and an asymptotic endpoint. A simple scheme for reducing the data to parameters reflecting each of these characteristic features is included. These parameters include (1) the kinetic tenth time, t_{10} (time to reach 10 % of reaction), reflecting the nucleation phase, (2) a composite term reflecting the difference between half-time, t_{50}, and kinetic tenth time ($t_{50} - t_{10}$) characteristic of the growth phase and (3) the time-independent value of the extent of the monomer incorporated into amyloid, $(C_{M\rightarrow A})_{t\rightarrow\infty}$, characterizing the asymptotic phase. *Blue line* Value of $C_{M\rightarrow A}$ as a function of time, *green line* the corresponding concentration of monomer as critical nucleus (nC_N) as a function of time (adapted, with permission, from Fig. 1d of Hall et al. 2016a, b). **b** Data reduction and analysis. In the case of drug screening for amyloid inhibitors, replicate measurements of the measured growth kinetics are decomposed into a set of characteristic values (such as the set of parameters described in Fig. 7a), with resultant values represented as a fractional histogram. **c**–**e** Fractional histogram representation of the surrogate markers of the nucleation (**d**), growth (**c**) and asymptotic (**e**) regions derived from the simulations shown in **b** (adapted, with permission from Fig. 2 of Hall et al. 2016b), with simulated results multiplied by a constant value to more closely reflect the time course and concentration profiles shown in subsequent cases)

[19] Approximation of the distribution by the average prior to estimation of the turbidity will introduce another type of error into the process.

[20] In the modern day computer-based numerical integration of ODE sets lessens the requirement for the determination of analytical solutions that may be based on further mathematical approximations (in addition to the already existing chemical approximations).

[21] Interestingly, with regard to this point, it was shown that a nucleation event could be generated for a linear polymerization process based solely on consideration of the conformational entropy of the unfolded protein—that is an entropic nucleus (Hall et al. 2005); Hall and Hirota 2009).

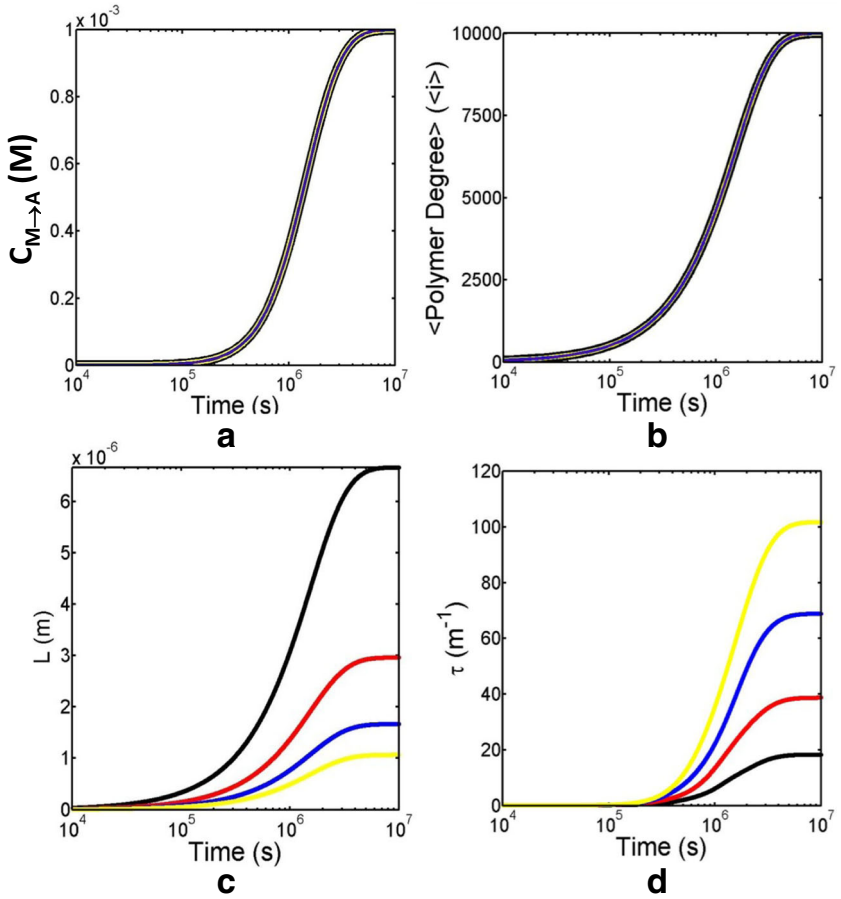


Fig. 8 Irreversible nucleation–growth model—effect of fibre width on the turbidity transform. Simulation of four cases of irreversible amyloid growth which, although exhibiting identical growth kinetics, differ in the radius of the amyloid fibre produced, such that R_A = 4 nm (*black line*), 6 nm (*red line*), 8 nm (*blue line*) or 10 nm (*yellow line*). **a** Concentration of monomer incorporated into amyloid, $C_{M\rightarrow A}$, as a function of time for four different cases of amyloid radius (*single line* for all four cases reflects identical growth kinetics dictated by imposition of identical rate constants. **b** Average polymer degree (<*i*>) of aggregate as a function of time for the four different cases of amyloid fibre radius (*single line* for all four cases is due to identical nucleation and growth kinetics brought about by use of identical rate constants). (**c**) Length (*L*) of amyloid fibres as a function of time for the four different cases of amyloid fibre radius. As per volume conservation requirements, fibres of different width lengthen in a manner proportional to $L_1/L_2 = (R_{ROD2})^2/(R_{ROD1})^2$. **d** Turbidity ($\tau$) of amyloid fibres as a function of time for the four different cases of amyloid fibre radius calculated using the transforms shown in Eqs. 9, 10a, b and 11. For the same average degree of polymerization, wider fibres of shorter length exhibit much greater turbidity than narrow fibres of longer length. Common parameters: f_A =10 M^{-1} s^{-1}, $f_N = 1 \times 10^{-7}$ M^{-1} s^{-1}, $b_A = 0$ s^{-1}, n = 2, $(C_M)_{tot} = 1 \times 10^{-3}$ M, R_1 = 2 nm, M_1 = 27.65 kg/mole, $\upsilon = 0.73 \times 10^{-3}$ m^3 kg^{-1}, $\alpha = 1.0$

empirically describe the time-course of amyloid formation (e.g. Morris et al. 2009). To effect simulation of irreversible nucleation–growth kinetics, all fibre breakage rate constants in Eq. 12a, b are set equal to zero (i.e. all $b_{i,j} = 0$). From the time of the original work by von Smoluchowski (von Smoluchowski 1917, 1916), many have attempted a first principles estimation of association rate constants ($f_{i,j}$) based on the component characteristics (e.g. Hall et al. 2005; Hill 1983; Pallitto and Murphy 2001). On the assumption that amyloid growth occurs primarily via monomer addition, Hall and Hirota (2009) calculated a numerical value for all $f_{1,j}$ association constants (based on hydrodynamic reasoning) and then used these parameter values to perform a full distribution simulation of amyloid growth, exploring effects of peptide position and role of peptide flanking sections. As a further simplification, in the Oosawa-type models all forward association rate constants are assigned one of two different values depending on their positional relation to the polymerization event featuring the nucleus, considered as possessing a size, n (Masel et al. 1999; Oosawa and Asakura 1975). In the Oosawa approximation, association rate constants ($f_{i,j}$) are set equal to either f_N, denoted as the nucleation rate constant for species i+ j ≤ n, or f_A, termed the growth rate constant for association of species i,j where i+j > n. Kinetics comporting to the nucleation–growth scheme are generated by calculating the rate of formation and loss of each species on the condition that $f_A >> f_N$ (Arosio et al. 2012; Hall 2003). A group of three coupled ODEs (shown in Table 2) representing the Oosawa–Kasai–Asakura approximation (Hall 2003; Oosawa and Asakura 1975) is produced upon appropriate summation of the

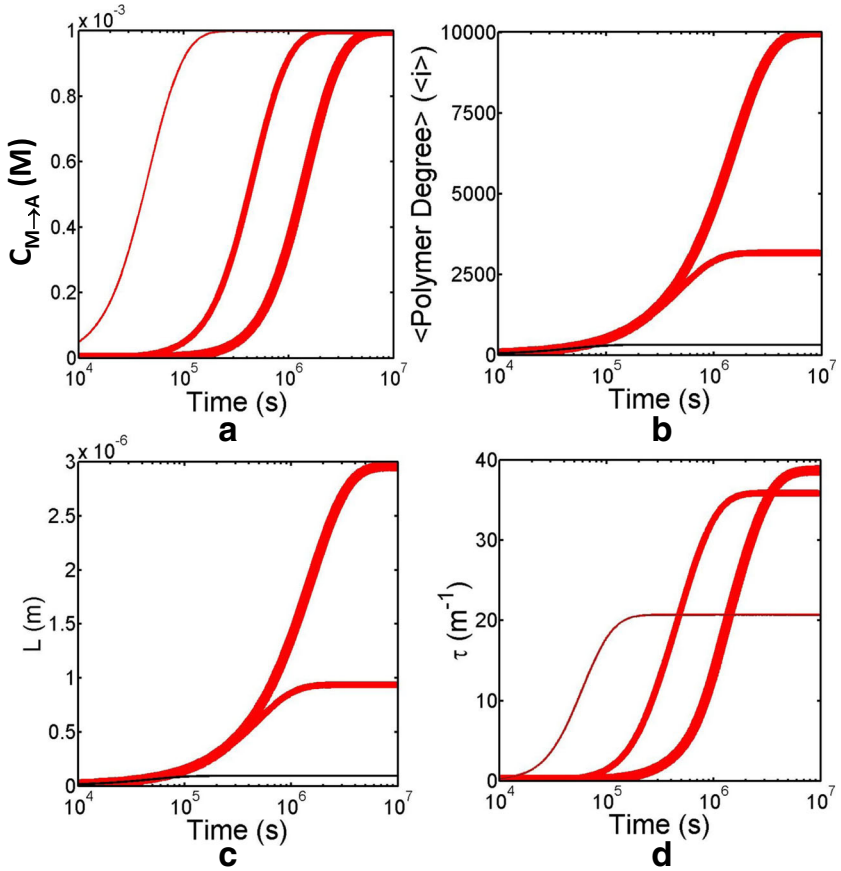


Fig. 9 Irreversible growth model—effect of nucleation rate on the turbidity transform. Simulation of three cases of irreversible amyloid growth which, although rod widths are identical, differ in the rate of nucleation of amyloid fibre produced such that $f_N = 1 \times 10^{-7}$ M^{-1} s^{-1} (*thick red line*), $f_N = 1 \times 10^{-6}$ M^{-1} s^{-1} (*intermediate-thick red line*) and $f_N = 1 \times 10^{-4}$ M^{-1} s^{-1} (*thin line*). **a** Concentration of monomer incorporated into amyloid ($C_{M\rightarrow A}$) as a function of time for the three different cases of amyloid nucleation rate. Faster nucleation rates dictate faster growth kinetics due to a greater number of extendable nuclei being formed. **b** Average polymer degree (<*i*>) of aggregate as a function of time for three different cases of amyloid fibre nucleation. Slower nucleation rates lead to larger average degrees of polymerization. **c** Length (*L*) of amyloid fibres as a function of time for the three different cases of amyloid fibre nucleation rate. As per the average degree of polymerization, for fixed fibre geometry, slower nucleation rates lead to longer fibres. **d** Turbidity (τ) of amyloid fibres as a function of time for the three different cases of amyloid fibre nucleation rate. As can be noted from Fig. 6d, the specific turbidity becomes relatively insensitive to length after the fibres are longer than ~2λ. In practice this finding means that for conditions producing very small fibre distributions, due to rapid nucleation kinetics, the measured turbidity value reflecting the asymptotic limit will be lower than that obtained for a system producing the same mass concentration of amyloid using slower nucleation kinetics. Common parameters: $f_A = 10$ M^{-1} s^{-1}, $f_N = 1 \times 10^{-7}$ M^{-1} s^{-1}, $b_A = 0$ s^{-1}, $n = 2$, $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mole, $\upsilon = 0.73 \times 10^{-3}$ m^3 kg^{-1}. $R_A = 6$ nm, $\alpha = 1.0$

complete set of ODEs specifying the rate of formation and loss of each aggregate species (Hall et al. 2015, 2016a, 2016b). Within this reduced set of equations the nucleus number concentration is given by C_N and the sum of the number concentrations of all amyloid fibre species is described by C_A (whereby $C_A = \Sigma C_i$ from n+1 to the maximum amyloid degree). The number concentration of all monomers within amyloid form is denoted as $C_{M\rightarrow A}$ (whereupon $C_{M\rightarrow A} = \Sigma i.\ C_i$ from n+1 to the maximum amyloid degree). On the basis that the signal measure of amyloid formation reflects $C_{M\rightarrow A}$, methods have been proposed for deducing the nucleus size and the nucleation and growth rate constants from logarithmic transform plots (Hall 2003; O'Nuallain et al. 2006; Oosawa and Asakura 1975; Powers and Powers 2006). A noted feature of the irreversible nucleated growth mechanism is that, dependent upon the relative rates of nucleation versus growth, a demonstrable amount of monomer existing as nucleus species can be present at the reaction end (e.g. see Fig. 7) (Hall et al. 2016b). Another important feature of the irreversible nucleated growth model is that the end state polymer distribution attains a stationary set of values at the same instant as the polymer mass end state, i.e. only the left-hand side equilibrium in Eq. 14 is operative and therefore no slow rearrangement of the distribution takes place (Bridstrup and Yuan 2016; Hall 2003; Hall and Minton 2004).

We considered two different cases of the irreversible nucleated growth model with regard to the turbidimetric transform. The first case (Fig. 8) explores the effects of different fibre geometry upon the turbidity signal. To examine this, four different examples of fibre radii (4, 6, 8 and 10 nm) are considered, with all cases following identical growth kinetics

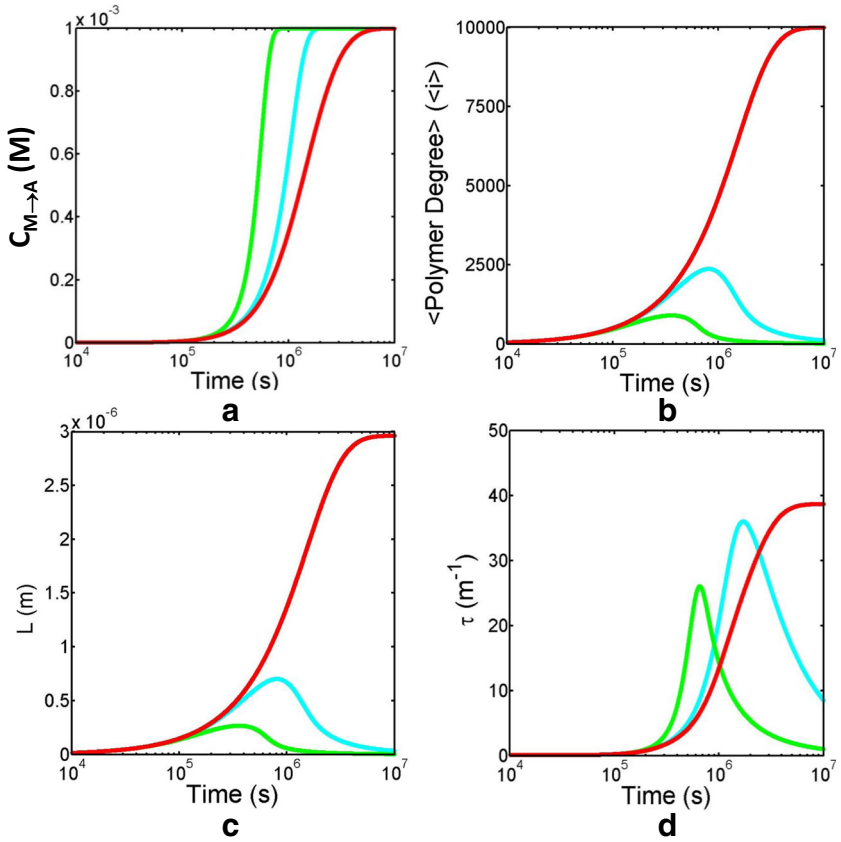


Fig. 10 Reversible growth model—effect of breakage rate on the turbidity transform. Simulation of three cases of reversible growth with breakage, in which the fibre width is the same for all cases, but the fibres differ in their intrinsic tendency towards breakage (or as some have termed 'frangible') such that $b_A = 0\ s^{-1}$ (*red line*), $b_A = 1 \times 10^{-9}\ s^{-1}$ (*cyan line*) and $b_A = 1 \times 10^{-8}\ s^{-1}$ (*green line*). **a** Concentration of monomer incorporated into amyloid ($C_{M\rightarrow A}$) as a function of time for the three different cases of intrinsic breakage rate. Note that faster breakage rates lead to an effective reduction in both the nucleation and growth phases with a subsequent faster attainment of the asymptotic value. **b** The average polymer degree of aggregate (<*i*>) as a function of time for the three different cases of intrinsic breakage rate. Slow breakage rates, relative to the rate of attainment of the polymer mass equilibrium, can lead to a slow reduction in the average polymer degree in a manner effectively temporally decoupled from the time scale of attainment of the monomer/polymer mass equilibrium (see Eq. 14). **c** Length (L) of amyloid fibres as a function of time for the three different cases of breakage rate. As for the just described case of <i> vs. t, slow intrinsic breakage rates can lead to an uncoupling between the times scales of the total mass of protein existing as amyloid and the production of shorter fibre distributions from longer initial distributions. **d** Turbidity (τ) of amyloid fibres as a function of time for three different cases of amyloid breakage rate. As the fibres shorten below the ~2λ length limit the turbidity decreases significantly, even though there is noeffective decrease in $C_{M\rightarrow A}$. Common parameters: $f_A = 10\ M^{-1}\ s^{-1}$, $f_N = 1 \times 10^{-7}\ M^{-1}\ s^{-1}$, $n = 2$, $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mol, $\upsilon = 0.73 \times 10^{-3}\ m^3\ kg^{-1}$. $R_A = 6$ nm, $\alpha = 1.0$

(Fig. 8a,b). Due to volume conservation requirements, the thinner fibres lengthen faster (Fig. 8c), yet it is the shorter, thicker fibres that show the greatest extent of turbidity (Fig. 8d). With respect to this point, we note that relatively short changes in fibre dimension can effect a large change in the recorded turbidity (Fig. 8d—roughly fivefold for the 4 vs. 10 nm case). The second case considered for the irreversible nucleated growth scheme (Fig. 9) involves examination of the effects of slow to fast nucleus production on the chemical kinetics and accompanying turbidity development of a fibre with fixed geometry ($R_A = 6$ nm). Faster nucleation is known to produce a greater number concentration of smaller (Fig. 9b) and shorter (Fig. 9c) amyloid (Lomakin et al. 1996). Interestingly, as the fibre length falls below a limit of ~2λ, the corresponding turbidity value, taken as reflecting asymptotic extent, also falls (Fig. 9d) despite there being the same total amount of monomer in amyloid form for all cases of the nucleation rate. Such a decrease in turbidity for very short fibres was first described and theoretically rationalized for microtubule fibre formation (Berne 1974; Gaskin et al. 1974). This phenomenon was later re-examined (Hall and Minton 2005) specifically for the case of microtubules and recently further developed in relation to amyloid and amorphous growth (Hall et al. 2016a).

Nucleation–growth with fibre breakage

In this mechanism fibres break—both internally, to produce two new fibres, and at their extremities, to release non-amyloid monomers (Hall and Edskes 2004; Masel et al.

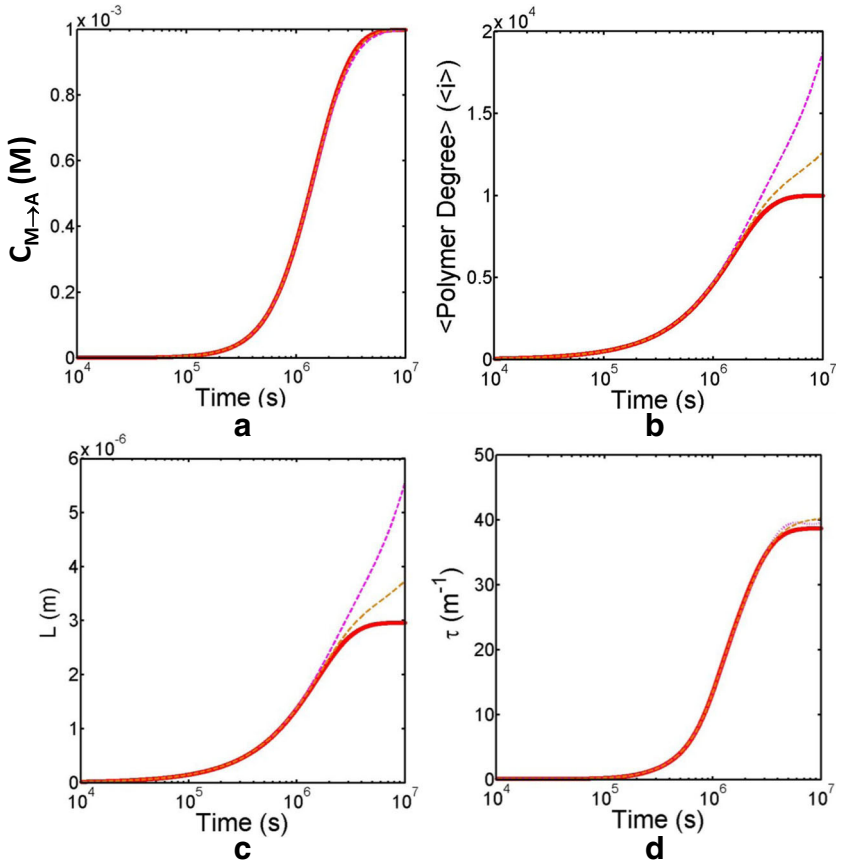


Fig. 11 Fibre end-to-end joining model—effect of association rate on the turbidity transform. Simulation of three cases of the fibre-joining model in which the amyloid fibre width is kept constant but the fibre joining rate (f_{JEE}) is set at $f_{JEE} = 0\ M^{-1}\ s^{-1}$ (*solid red line*), $f_{JEE} = 0.3\ M^{-1}\ s^{-1}$ (*dashed orange line*) and $f_{JEE} = 1.0\ M^{-1}\ s^{-1}$ (*dashed magenta line*). **a** Concentration of monomer incorporated into amyloid ($C_{M\rightarrow A}$) as a function of time for the three different cases of joining rate considered. The relatively low numerical values used for the joining rate constants in these simulations mean that the polymer redistribution kinetics are effectively decoupled from the monomer/polymer mass kinetics (see Eq. 14). As such, no change in the kinetics of monomer incorporation is observed in the three different cases considered. **b** Effect of fibre-joining rate on the average polymer degree ($<i>$) as a function of time. Faster rates of increase in polymer degree are affected by faster joining rates, but this occurs slowly in the present case due to the relatively low values of f_{JEE} specified. **c** Length (L) of amyloid fbres as a function of time for the three different cases of fibre-joining rate considered. Note that the fibres slowly lengthen under the regime of joining rate constants selected. **d** Turbidity (τ) of amyloid fibres as a function of time for the three different cases of fibre-joining rate considered. No change in turbidity is detectable amongst the three cases of fibre-joining rate considered. This result follows from relations summarized in Table 1 (represented pictorially in Fig. 6d) whereby an increase in length, at constant polymer mass concentration, should be largely invisible to detection by turbidity. Common parameters: $f_A = 10\ M^{-1}\ s^{-1}$, $f_N = 1 \times 10^{-7}\ M^{-1}\ s^{-1}$, $n = 2$, $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mol, $\upsilon = 0.73 \times 10^{-3}\ m^3\ kg^{-1}$. $R_A = 6$ nm, $\alpha = 1.0$

1999). The consequences of fibre breakage on the progression of amyloid kinetics have been considered from a number of different perspectives (Hall and Edskes 2004; Masel et al. 1999; Smith et al. 2006; Tanaka et al. 2004), with a detailed model of the potential effects of fibre breakage on amyloidosis-related disease progression being potentially the most important (Hall and Edskes 2009, 2012). With regard to this last point, the importance of both fibre breakage rate and fibre size distributions to aggregate cytotoxicity was demonstrated using a cell culture model (Xue et al. 2009). More recently, Nicoud et al. (2015) have considered further complicating effects upon amyloid growth kinetics associated with potential position dependence of fibre breakage. In the consensus model presented in Table 2, we have reduced Eq. 12a, b to a more tractable form by assuming that all monomer to monomer bonds within the amyloid fibre can break at a rate governed by the first-order rate constant b_A (i.e. $b_{i,j} = b_A$ for all i, j). As per the irreversible nucleated growth model, all second-order association reactions, in which at least one of the species is assumed to be a monomer, are governed by rate constants f_N and f_A, depending upon the size of the reactants. A summation of the set of rate equations describing the growth and loss of all species greater than the monomer produces the set of rate equations described in Table 2 (Hall and Edskes 2009; Smith et al. 2006). Depending upon the rate of internal fibre breakage, the collapse of the polymer size distribution may be either temporally coupled or decoupled from the kinetics of growth of the polymer mass (Hall and Edskes

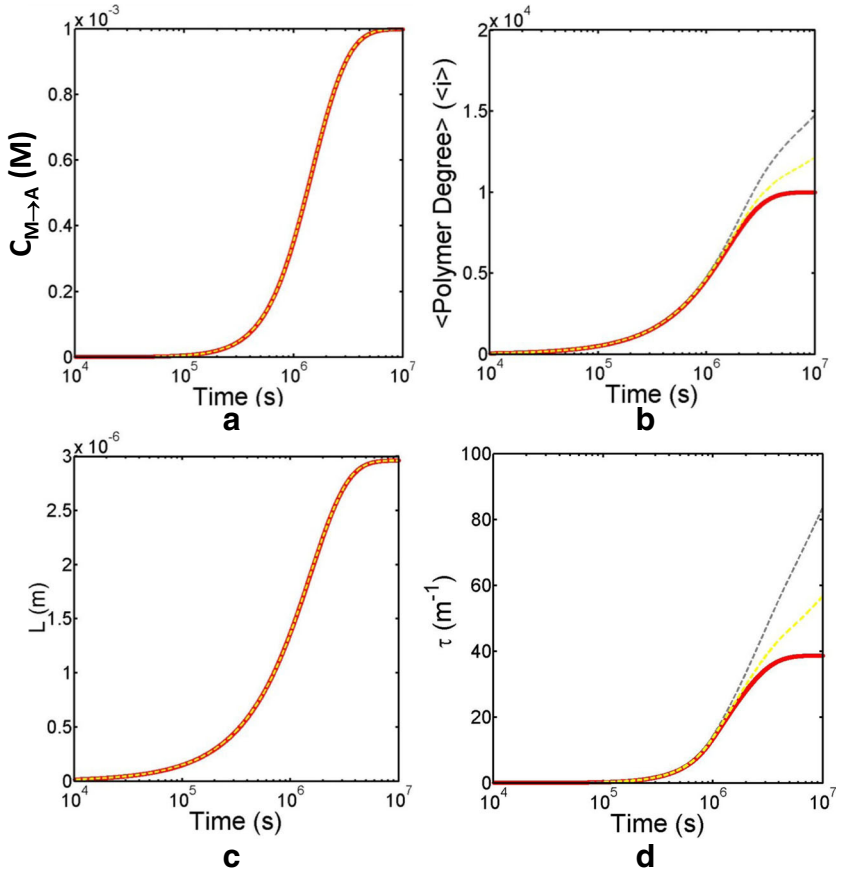


Fig. 12 Fibre lateral association model—effect of lateral association rate on the turbidity transform. Simulation of the fibre lateral association model in which fibres are able to form laterally-associated 'mature' fibres consisting of fibre dimers, for three cases of the joining lateral association rate constant (f_{JLA}) are explored, with $f_{J\,LA} = 0\ M^{-1}\ s^{-1}$ (*solid red line*), $f_{J\,LA} = 0.3\ M^{-1}\ s^{-1}$ (*dashed yellow line*) and $f_{J\,LA} = 10\ M^{-1}\ s^{-1}$ (*dashed grey line*). **a** Concentration of monomer incorporated into amyloid ($C_{M\rightarrow A}$) as a function of time. All three simulated cases of different intrinsic lateral association rate overlap as the fibre-joining rate is assumed not to influence the reactivity of the individual fibre ends. **b** Average polymer degree (<*i*>) as a function of time. The low numerical values selected for the fibre lateral association rate constants mean that the asymptotic limit of the average polymer degree is approached very slowly. **c** Simulated length (L) of amyloid as a function of time for the three examined cases of fibre lateral association rate. The coincident behavior is a consequence of the two simplifying assumptions that fibre size distributions are approximated by their average, <i>, and that fibre lateral association occurs at the fibre midpoint (see text on this point for a discussion). **d** Simulated turbidity (τ) of amyloid solution as a function of time for the three cases of fibre lateral association rate. Attainment of an asymptotic limit in the turbidity profile is delayed (or not apparent) for the cases of faster lateral association rate. Note that based on relations presented in Table 1 and Fig. 6d, a change in fibre width, at constant aggregate mass concentration, will result in an increase in turbidity. Common parameters: $f_A = 10\ M^{-1}\ s^{-1}$, $f_N = 1 \times 10^{-7}\ M^{-1}\ s^{-1}$, n = 2, $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mol, $\upsilon = 0.73 \times 10^{-3}\ m^3\ kg^{-1}$. $R_A = 6$ nm, $\alpha = 1.0$

2012). An important consequence of this mode of amyloid growth is that the end-point size distribution will always approach—albeit often extremely slowly—the critical nucleus size (Hall and Edskes 2009).

Three different rates of intrinsic fibre breakage ($b_A = 0\ s^{-1}$, $b_A = 1 \times 10^{-9}\ s^{-1}$ and $b_A = 1 \times 10^{-8}\ s^{-1}$) were simulated using the consensus reversible fibre growth model (shown in Table 2). Larger values of b_A were found to speed up the incorporation of monomer into the amyloid form (Fig. 10a) (Hall and Edskes 2004, 2009). Due to the relatively low values selected for the breakage rate constants, a very slow relaxation of the fibre distribution is seen (Fig. 10b, c). For constant fibre geometry ($R_A = 6$ nm), we note that fibre breakage produces non-ideal turbidimetric kinetic profiles, exhibiting a decrease in the asymptotic extent of turbidity as the fibre length falls below the ~L > 2λ limit (Fig. 10d).

Fibre end-to-end association

The joining of shorter fibres to form longer ones has been directly observed in some amyloid systems (Binger et al. 2008). Based on theoretical predictions (relating to differences in number concentration[22] and intrinsic orientation effects related to the likelihood of two fibre ends meeting (Hill 1983; Pallitto and Murphy 2001), the rate constant governing longitudinal fibre/fibre

[22] i.e. the monomer number concentration is higher than the fibre number concentration nearly throughout the monomer to polymer mass equilibrium.

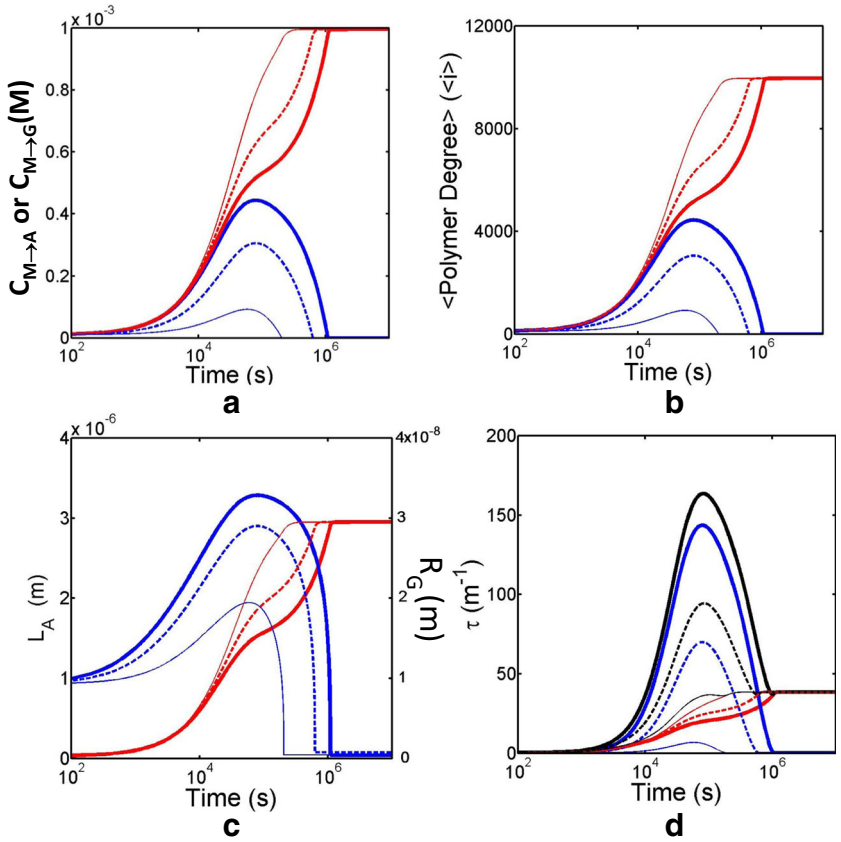


Fig. 13 Amyloid vs. amorphous competition—effect of relative rates of amorphous and amyloid growth on the turbidity transform. Simulation of three cases of competitive reversible-seeded growth in which the rate constants reflecting amyloid growth are kept constant but the amorphous growth kinetics are modified by varying the amorphous aggregate association constant (f_G), such that $f_G = 50\ M^{-1}\ s^{-1}$ (*thin solid lines*), $f_G = 150\ M^{-1}\ s^{-1}$ (*intermediate-thick dashed lines*) and $f_G = 250\ M^{-1}s^{-1}$ (*thick solid lines*) whereby the *red* version of the particular line type refers to the amyloid species and the *blue* version of the line refers to the amorphous species. **a** Concentration of monomer in amyloid ($C_{M\rightarrow A}$) or amorphous aggregate ($C_{M\rightarrow G}$) as a function of time for the three different cases of amorphous relative to amyloid growth. In all cases the choice of rate constants ensures that the amyloid is ultimately more thermodynamically stable than the amorphous aggregate. Relatively fast amorphous association rates lead to a significant extent of monomer being initially converted into the amorphous form, prior to its eventual dissociation and re-incorporation into the amyloid state. **b** Average polymer degree of amyloid ($<i_A>$) and amorphous ($<i_G>$) as a function of time for the different simulated cases of relative rates of amorphous to amyloid growth. Due to the fact that the simulation model specifies seeded growth (in which the number concentration of amyloid and amorphous species are fixed at constant values throughout—see Table 1), $<i_A>$ (*red lines*) attains the same eventual value for all cases of relative growth. Similarly, the average degree of polymerization of the amorphous aggregate, $<i_G>$ (*blue lines*) approaches a value close to the starting value of the amorphous seed, $<i_G>_{t=0}$, in all cases. **c** Average size of aggregate species as a function of time for three simulated cases of relative rates of amorphous vs. amyloid growth, with the *left y-axis* specifying the length (L_A) of the amyloid species and the *right y-axis* describing the radius (R_G) of the amorphous aggregate species. The faster cases of amorphous growth lead to aggregates of larger radius (compare ~32 to 20 nm) whereas L_A never surpasses its maximum value due to a slow approach to equilibrium from below (i.e. no overshoot is seen). **d** Turbidity (τ) as a function of time for the three cases reflecting different relative rates of amorphous to amyloid growth. *Coloured lines* Component turbidity generated by the amyloid (*red line*) and amorphous (*blue line*) species. *Black lines* represent the total resultant turbidity. *Line style* is dictated by the different cases reflecting the rate of amorphous to amyloid growth: *solid thick lines* relatively fast amorphous growth, *dashed intermediate-thick lines* amorphous growth, *thin solid lines* slow amorphous growth. Common parameters: $f_A = 250\ M^{-1}\ s^{-1}$, $b_A = 1 \times 10^{-3}\ s^{-1}$, $b_G = 1 \times 10^{-2}\ s^{-1}$, $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mol, $\upsilon = 0.73 \times 10^{-3}\ m^3\ kg^{-1}$. $R_A = 6$ nm, $\alpha_A = \alpha_G = 1.0$, $(C_A)_{t=0} = 1 \times 10^{-7}$M, $(C_G)_{t=0} = 1 \times 10^{-7}$M, $<i_A>_{t=0} = 100$, $<i_G>_{t=0} = 100$

association ($f_{i,j}$) is assumed to be much smaller than the fibre/monomer association rate constant ($f_{i,1}$). As such, the kinetics of fibre annealing is likely to be relatively slow and decoupled from the (relatively) faster kinetics of the monomer/polymer mass reaction. In terms of the equilibrium described by Eq. 14, fibre joining likely exhibits a slow redistribution phase. In our consensus model, the rate of joining between any two fibres is specified by a single rate constant, f_{JEE}, such that $f_{JEE} << f_A$.

Three cases of fibre joining rate were simulated using the consensus fibre end-to-end joining model shown in Table 2. All simulated cases had the same fibre width, with $R_A = 6$ nm. Following the asymptotic relation predicted for the transmission form factor (Q) for rods (Table 1; Fig. 6), the turbidity is effectively blind to changes in length brought about by joining. In a different kinetic regime (not explored here), fibre joining could significantly influence the turbidimetric profile

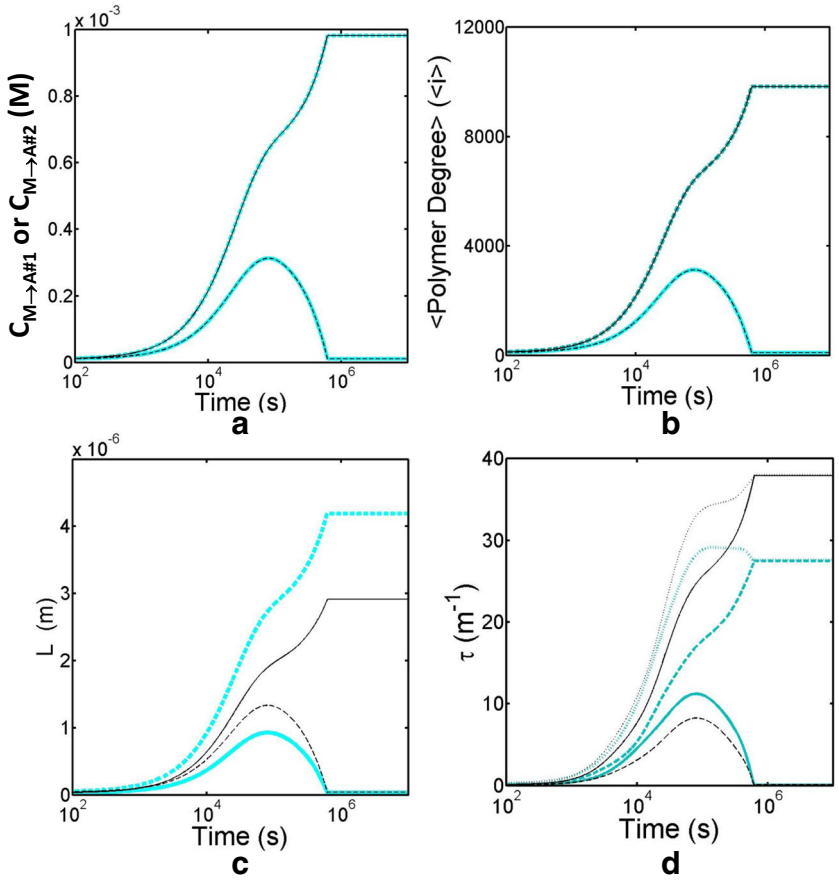


Fig. 14 Amyloid vs. amyloid competition—effect of relative rates of growth between two geometric forms of amyloid on the turbidity transform. Simulations showing two cases of competitive reversible seeded growth between two amyloid types possessing quite subtle differences in geometry such that type #1 fibres have a radius $R_{A\#1}$ of 5 nm (*dashed lines*) and type #2 fibres have a radius of $R_{A\#2}$ = 6nm (*solid lines*). Two cases of reversible growth are produced by swapping the sets of kinetic rate constants. The simulation showing eventual more stable growth of the narrow type #1 fibres in the thermodynamic limit is defined by Case A (*cyan lines*; $f_{A\#1} = 150\ M^{-1}\ s^{-1}$, $b_{A\#1} = 0.001\ s^{-1}$, $f_{A\#2} = 250\ M^{-1}\ s^{-1}$, $b_{A\#2} = 0.01\ s^{-1}$). The simulation ultimately reflecting more stable growth of the thicker type #2 fibres is defined by Case B (*black lines*; $f_{A\#1} = 250\ M^{-1}\ s^{-1}$, $b_{A\#1} = 0.01\ s^{-1}$, $f_{A\#2} = 150\ M^{-1}\ s^{-1}$, $b_{A\#2} = 0.001\ s^{-1}$). **a** Concentration of monomer incorporated into either of the two types of amyloid ($C_{M\rightarrow A\#1}$ *or* $C_{M\rightarrow A\#2}$) as a function of time. As the kinetics are simply reversed between the two different cases, Case A (*cyan lines*) and Case B (*black lines*) are coincident. **b** Average polymer degree (<*i*>) reflecting either type #1 amyloid (<$i_{A\#1}$>) or type #2 amyloid (<$i_{A\#2}$>) as a function of time. As the polymer degree per se is insensitive to the geometry of the amyloid, these two cases are also coincident (being simple reversals of the kinetic rate constants). **c** Simulated length (*L*) of amyloid as a function of time for Case A (*cyan lines*), whereby the thin fibre (*dashed lines*) is eventually dominant, and Case B (*black lines*) for which the thick fibre (*solid line*) is eventually dominant. The differences in width between the two fibre types means that different lengths are produced between the two cases even though the average degree of polymerization is identical. **d** Turbidity (τ) of amyloid solution as a function of time for two cases of competitive amyloid growth. The resultant turbidity for both cases is shown by *dashed lines* (*thin black dashed line* Case B, *thick cyan dashed line* Case A). Note the unusual kinetics (different to the ideal type shown in Fig. 7a) produced by very minor differences in fibre geometry. Common parameters: $(C_M)_{tot} = 1 \times 10^{-3}$ M, $R_1 = 2$ nm, $M_1 = 27.65$ kg/mol, $\upsilon = 0.73 \times 10^{-3}\ m^3\ kg^{-1}$. $\alpha_{A\#1} = 1.0$, $\alpha_{A\#2} = 1.0$

if it occurred between very small fibres (Hall and Minton 2005) having a length smaller than the $L > 2\lambda$ limit, (above this limit, specific turbidity becomes effectively independent of fibre length (Berne 1974; Buitenhuis et al. 1994; Hall et al. 2016a).

Fibre lateral association

Despite widespread descriptions of laterally associated 'mature fibres' (Fitzpatrick et al. 2013; Kanno et al. 2005; Ridgley and Barone 2013; Woolfson and Ryadnov 2006; Yamaguchi et al. 2005) and semi-ordered aggregation of amyloid to form spherulites (Krebs et al. 2004a; Ruth et al. 2000), there is a general dearth[23] of experimental and theoretical studies which have considered the effect of lateral association on amyloid growth kinetics. Three different mechanistic possibilities exist for the production of laterally associated fibres. The first involves heterogeneous nucleation of a new fibre on the surface

[23] With some exceptions (Ghosh et al. 2010; Pallitto and Murphy 2001).

of a pre-existing amyloid fibre (Jeong et al. 2013; Padrick and Miranker 2002); the second involves lateral association taking place while the fibres are small, with subsequent extension of each growing end of the conjoined fibres (Ghosh et al. 2010; Pallitto and Murphy 2001); the third mechanistic option involves self-association of already formed protofibrils with lateral fibre association governed by rate constants much smaller than the rate constants governing the monomer/polymer mass equilibrium (Eq. 14)(Ghosh et al. 2010; Pallitto and Murphy 2001) i.e. $(f_{i,j})_{LAT} << f_{i,j}$. There is a great deal of complexity in the simulation of any of these three cases. To provide a benchmark kinetic description of the effect of fibre lateral growth, we have opted to simulate a kinetic case that is similar in form to the preceding fibre end-to-end joining model. For simplicity we have limited this to lateral protofibril addition governed by a rate constant f_{JLA}, with the level of association fixed to the stage of lateral dimer, denoted as A2 (Table 2). Such a kinetic description dictates a slow association phase in which the polymers become progressively thicker on average (Fig. 12). One potentially misleading assumption in the formulation of the equation set describing lateral association (Table 2) is the subsuming of a multitudinous array of possible lateral associations (involving partial off-centre overlap of fibres) into a single mechanistic path describing centre-to-centre alignment of the fibres. Such off-centre association, or indeed point contact formation, may be responsible for the formation of either extremely long fibres or the birds' nest-type clusters of fibres often seen in amyloid plaques (Merz et al. 1983; Wisniewski et al. 1989) and in ultra-microscope images (Ban et al. 2003; Mishra et al. 2011; Ogi et al. 2014).

Amyloid versus amorphous growth

When there is no orientation or configurational requirement to the association reaction, the internal structure of the aggregate will lack positional order, resulting in the formation of an amorphous product (Bennett 1972; Yoshimura et al. 2012; Zurdo et al. 2001) (Table 2). The production of such amorphous aggregates has been observed in many amyloid-forming systems and often complicates simple interpretation of the reaction. Hall et al. (2015) and Adachi et al. (2015) have treated the case of amyloid growth in competition with amorphous aggregate using a kinetic rate scheme that treated the rate of growth and breakage of all species in an explicit fashion. Here we produce example simulations describing the competition between the amyloid and amorphous aggregate based on a fixed-seeded growth model (Naiki et al. 1997). In this mechanistic format, growth proceeds reversibly, for both amorphous and amyloid aggregate types, from a fixed (i.e. non-dissociable) seed species of degree $\langle i \rangle_{t=0}$. Here we consider growth and shrinkage as occurring via monomer addition and monomer loss only, with no fragmentation or spontaneous nucleation allowed. For amyloid growth this implies the following boundary conditions: $f_{i,1} = f_A$ for $\langle i_A \rangle \geq \langle i_A \rangle_{t=0}$, else all $f_{i,1} = 0$ and $b_{i-1,1} = b_A$ for $\langle i_A \rangle > \langle i_A \rangle_{t=0}$, else $b_{i-1,1} = 0$; for amorphous aggregate growth, the following boundary conditions are implied: $f_{i,1} = f_G$ for $\langle i_G \rangle \geq \langle i_G \rangle_{t=0}$, else all $f_{i,1} = 0$ and $b_{i-1,1} = b_G$ for $\langle i_G \rangle > \langle i_G \rangle_{t=0}$, else $b_{i-1,1} = 0$. Hall and coworkers explored the case where amyloid is more thermodynamically stable than the amorphous aggregate, but slower to initially form (Hall et al. 2015). This study, along with work by Adachi et al. (2015), highlighted a potential dependence of the timescale of amyloid formation on the dissociation rate of the amorphous species.

To explore the effects of competing amorphous growth in extension to that performed previously (Adachi et al. 2015; Hall et al. 2015), we simulated three cases of competition between amorphous aggregate and amyloid using the consensus model presented in Table 2 (Fig. 13). The rate constants were selected to ensure that the amyloid was ultimately more stable, in the thermodynamic limit, than the amorphous aggregate in all cases (Fig. 13a–c). Despite this preponderance for amyloid growth, even relatively small amounts of amorphous aggregate can significantly distort the resultant turbidimetric kinetic profile (Fig. 13d, black lines).

Amyloid versus amyloid growth—two strains in competition

Kinetic competition between two types of amyloid has not yet, to the best of our knowledge, been quantitatively investigated, but it is known to be an important feature of real-case amyloid growth in which different polymorphic strains are observed (Hall and Edskes 2004; Paravastu et al. 2008; Petkova et al. 2005; Tanaka et al. 2004). To explore this behaviour in isolation, we modelled two types of amyloid fibre, both competing for the same monomer pool, using a fixed seeded reversible growth scheme virtually identical to that adopted for the amorphous versus amyloid case (reported in the preceding subsection) with the same requirement that fibres grow or shrink via monomer addition or loss only[24] (Table 2). The two fibre types are characterized by different radii ($R_{A\#1}$ and $R_{A\#2}$). Growth for the type #1 fibre is defined by the following set of boundary conditions: $f_{i,1} = f_{A\#1}$ for $\langle i_{A\#1} \rangle \geq \langle i_{A\#1} \rangle_{t=0}$, else all $f_{i,1} = 0$ and $b_{i-1,1} = b_{A\#1}$ for $\langle i_{A\#1} \rangle > \langle i_{A\#1} \rangle_{t=0}$, else $b_{i-1,1} = 0$; growth for type #2 fibres is defined by $f_{i,1} = f_{A\#2}$ for $\langle i_{A\#2} \rangle \geq \langle i_{A\#2} \rangle_{t=0}$, else all $f_{i,1} = 0$ and $b_{i-1,1} = b_{A\#2}$ for $\langle i_{A\#2} \rangle > \langle i_{A\#2} \rangle_{t=0}$, else $b_{i-1,1} = 0$.

[24] We note that the general behaviour outlined for the two competitive cases ('Amyloid vs. amorphous growth' and 'Amyloid versus amyloid growth—two strains in competition') are also applicable to the description of the growth and dissolution of different crystal forms (Levi and Kotrla 1997).

Two different cases of competitive fibre growth were explored (Fig. 14). The first involved the situation where a relatively thin type #1 fibre ($R_{A\#1}$ = 5 nm) outcompeted a slightly thicker type #2 fibre ($R_{A\#2}$ = 6 nm) for monomer resources (Fig. 14a–c). The second case involved the reverse situation, in which the slightly wider type #2 fibres eventually outcompeted the thinner type #1 fibres for monomer (Fig. 14a–c). As can be noted (Fig. 14d), even relatively subtle differences in dimensions between the dominant and non-dominant fibre types will impart significant non-ideality to the resultant kinetic profile recorded via turbidity (dotted black or cyan lines in Fig. 14d).

Towards the future

Concern over the interpretation of potentially non-linear signal response is a repeating and important theme in science (Araujo 2009). Pursuing this line of investigation, we have focussed on literature capable of informing the reader about the cause and effect relationship between protein aggregates and the turbidity generated by them in solution. Using a slightly non-conventional review format we have combined published transforms with consensus kinetic models to produce 'review-data', cutting out the requirement for worded descriptions otherwise necessary for synthesizing arguments from multiple information streams.

As can be noted from Fig. 6d, straight rod-like fibres possessing a common width should produce a signal that is linearly proportional to the mass of monomer in aggregate form when the fibres are long in relation to the wavelength (Hall et al. 2016a) (or in the words and symbols of this review, $\tau \propto C_{M\rightarrow A}$ when $L_A > 2\lambda$). Outside of this limit a linear relationship will not necessarily hold and should be either (1) investigated experimentally (e.g. Borgia et al. 2013; O'Nuallain et al. 2006), (2) compared against results gained from an orthogonal technique (Li et al. 2009; Nilsson 2004) or (3) examined using some of the theoretical and simulation-based tools highlighted in this review. A fourth option, previously explored by a number of researchers, involves (4) experimental interrogation of the wavelength dependence of the turbidity (Camerini-Otero and Day 1978; Wallach et al. 1961) to gain clues about dominant aggregate sub-types (Andreu and Timasheff 1986; Garcia-Lopez and Garcia-Rubio 2008; Garcia-Lopez et al. 2006; Hall and Minton 2005; Korolevskaya and Khlebtsov 2010; Mahler et al. 2005; Moody et al. 1996; Silver and Birk, 1983)). Advances in computer power make re-visitation of this multi-wavelength approach an attractive area of current and future research (Mroczka and Szczuczynski 2010; Penzkofer et al. 2007; Shmakov 2014).

By far, the major focus of amyloid research remains its association with diseases, collectively termed as amyloidosis, which are all characterized by the deposition of large amounts of amyloid into various organs and tissues of the human body (Symmers 1956; Pepys, 2001; Walker and Jucker 2015). Potential non-linear effects, which complicate the interpretation of the turbidity signal, become very important when turbidity is used as an assay for anti-amyloid drug screening (Anzai et al. 2016; Doig et al. 2004; Dolado et al. 2005; Necula et al. 2007; Sant'Anna et al. 2016). In such cases, the use of an orthogonal technique, such as Congo Red or Thioflavin T dye binding, as a control experimen, should go a long way towards preventing spurious conclusions from being drawn.

From its original negative association with disease, amyloid has since been found to play potentially beneficial roles in non-disease-related areas. Two such positive manifestations of amyloid include (1) the discovery of its role in maintaining the normal biological state as 'functional amyloid' (Greenwald and Riek 2010) and (2) amyloid's potential in biosynthetic applications (Mitraki 2010; Raynes and Gerrard 2013). In this latter role, amyloid's nanometer-scale dimensions (Xu et al. 2016), its inherent capacity for autonomous self-assembly (Lee et al. 2015; Sasahara et al. 2010) and the desirable material properties of the nanofiber product (Paul et al. 2016) all highlight the potential usefulness of amyloid as a 'building block' in nanotechnology applications (Rodina 2012). Due to their simplicity, turbidity assays will continue to be the 'go to' technique for monitoring amyloid formation across these disparate research areas. A greater familiarity with the principles of the turbidimetric technique will undoubtedly facilitate research progress throughout the wider community. Hopefully this review has helped to decrease the foggy nature of turbidity, allowing some metaphorical blue skies to be seen through the cloud.

Acknowledgements All authors would like to wish Professor Donald Winzor the very best for his recently passed 80th birthday and congratulate him on 60 years of outstanding scientific research and mentorship. DH would like to thank Dr. Nami Hirota and Dr. Stephen Williams for helpful discussions and for commenting on an early version of this Review and Miss I Sakura, Miss L Sayuri and Miss M Satoko for kind assistance with the preparation of the manuscript. RZ thanks the Chinese government and the ANU for the award of the CSC-ANU Scholarship. The work of NJR was partly funded by an Australian Postgraduate Research Award. The work of FA is assisted by Funds in Aid provided by Nihon University. The research of YG and MS is supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology. The research of JAC is supported by a Project Grant from the National Health and Medical Research Council of Australia. The work of DH is jointly funded by an Australian National University Senior Research Fellowship and an Osaka University Cross-Appointment as an Associate Professor. DH also receives funding from the Osaka University Institute for Protein Research Program for Multiscale StructuralBiology.

Compliance with ethical standards

Conflict of interests Ran Zhao declares that he has no conflict of interest.
Masatomo So declares that he has no conflict of interest.
Hendrik Maat declares that he has no conflict of interest.
Nicholas J. Ray declares that he has no conflict of interest.
Fumio Arisaka declares that he has no conflict of interest.
Yuji Goto declares that he has no conflict of interest.
John A. Carver declares that he has no conflict of interest.
Damien Hall declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed byany of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Adachi M, So M, Sakurai K, Kardos J, Goto Y (2015) Supersaturation-limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation. J Biol Chem 290:18134–18145

Adamcik J, Mezzenga R (2011) Proteins fibrils from a polymer physics perspective. Macromolecules 45:1137–1150

Adamcik J, Jung JM, Flakowski J, De Los Rios P, Dietler G, Mezzenga R (2010) Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. Nat Nanotechnol 5:423–428

Aldous DJ (1999) Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. Bernoulli 5:3–48

Andreu JM, Timasheff SN (1986) [5] The measurement of cooperative protein self-assembly by turbidity and other techniques. Methods Enzymol 130:47–59

Anzai I, Toichi K, Tokuda E, Mukaiyama A, Akiyama S, Furukawa Y (2016) Screening of drugs inhibiting in vitro oligomerization of Cu/Zn-superoxide dismutase with a mutation causing amyotrophic lateral sclerosis. Front Mol Biosci 3:40

Apostol MI, Sawaya MR, Cascio D, Eisenberg D (2010) Crystallographic studies of prion protein (PrP) segments suggest how structural changes encoded by polymorphism at residue 129 modulate susceptibility to human prion disease. J Biol Chem 285:29671–29675

Araujo P (2009) Key aspects of analytical method validation and linearity evaluation. J Chromatogr B 877:2224–2234

Arosio P, Beeg M, Nicoud L, Morbidelli M (2012) Time evolution of amyloid fibril length distribution described by a population balance model. Chem Eng Sci 78:21–32

Ban T, Goto Y (2006) Direct observation of amyloid growth monitored by total internal reflection fluorescence microscopy. Methods Enzymol 413:91–102

Ban T, Hamada D, Hasegawa K, Naiki H, Goto Y (2003) Direct observation of amyloid fibril growth monitored by thioflavin T fluorescence. J Biol Chem 278:16462–16465

Bennett CH (1972) Serially deposited amorphous aggregates of hard spheres. J Appl Phys 43:2727–2734

Berne BJ (1974) Interpretation of the light scattering from long rods. J Mol Biol 89:755–758

Binger KJ, Pham CL, Wilson LM, Bailey MF, Lawrence LJ, Schuck P, Howlett GJ (2008) Apolipoprotein C-II amyloid fibrils assemble via a reversible pathway that includes fibril breaking and rejoining. J Mol Biol 376:1116–1129

Bishop MF (1989) Calculations of scattered light from rigid polymers by Shifrin and Rayleigh-Debye approximations. Biophys J 56:911

Bohren CF, Huffman DR (2008) Absorption and scattering of light by small particles. John Wiley & Sons, New York

Borgia MB, Nickson AA, Clarke J, Hounslow MJ (2013) A mechanistic model for amorphous protein aggregation of immunoglobulin-like domains. J Am Chem Soc 135:6456–6464

Bridstrup J, Yuan JM (2016) Effects of crowders on the equilibrium and kinetic properties of protein aggregation. Chem Phys Lett 659:252–257

Buitenhuis J, Dhont JK, Lekkerkerker HN (1994) Scattering of light from cylindrical particles: Coupled dipole method calculations and the range of validity of the Rayleigh-Gans-Debye approximation. J Colloid Interface Sci 162:19–24

Cantor CR, Schimmel PR (1980) Configurational statistics of polymer chains. Biophysical chemistry part III: the behaviour of macromolecules. W.H. Freeman, New York

Camerini-Otero RD, Day LA (1978) The wavelength dependence of the turbidity of solutions of macromolecules. Biopolymers 17:2241–2249

Carulla N, Caddy GL, Hall DR, Zurdo J, Gairi M, Feliz M, Giralt E, Robinson CV, Dobson CM (2005) Molecular recycling within amyloid fibrils. Nature 436:554–8

Casassa EF (1955) Light scattering from very long rod-like particles and an application to polymerized fibrinogen. J Chem Phys 23:596–597

Chu B (1974) Laser light scattering. Elsevier, Amsterdam

Dalpadado RC, Maat H, Carver JA, Hall D (2016) Real-time monitoring of amyloid growth in a rigid gel matrix. Anal Biochem 511:13–16

Debye P (1947) Molecular-weight determination by light scattering. J Phys Chem 51:18–32

Dhont JKG (1983) Multiple Rayleigh-Gans-Debye scattering in colloidal systems-general theory and static light scattering. Physica A 120:238–262

Doig AJ, Stott K, Treherne JM (2004) Inhibitors of amyloid aggregation: technologies for the discovery of novel lead compounds. Biotechnol Genet Eng Rev 21:197–214

Dolado I, Nieto J, Saraiva MJM, Arsequell G, Valencia G, Planas A (2005) Kinetic assay for high-throughput screening of in vitro transthyretin amyloid fibrillogenesis inhibitors. J Comb Chem 7:246–252

Doty P, Steiner RF (1950) Light scattering and spectrophotometry of colloidal solutions. J Chem Phys 18:1211–1220

Elimelech M, Gregory J, Jia X (2013) Particle deposition and aggregation: measurement, modelling and simulation. Butterworth-Heinemann, Oxford

Ferrone FA, Hofrichter J, Eaton WA (1985) Kinetics of sickle hemoglobin polymerization: I. Studies using temperature-jump and laser photolysis techniques. J Mol Biol 183:591–610

Fitzpatrick AW, Debelouchina GT, Bayro MJ, Clare DK, Caporini MA, Bajaj VS, Jaroniec CP, Wang L, Ladizhansky V, Müller SA et al (2013) Atomic structure and hierarchical assembly of a cross-ß amyloid fibril. Proc Natl Acad Sci USA 110:5468–5473

Gans R (1925) Strahlungsdiagramme ultramikroskopischer Teilchen. Ann Phys 381:29–38, as cited by Zimm, 1945

Garcia-Lopez AC, Garcia-Rubio LH (2008) Rayleigh-Debye-Gans as a model for continuous monitoring of biological particles: Part II, development of a hybrid model. Opt Express 16:4671–4687

Garcia-Lopez AC, Snider AD, Garcia-Rubio LH (2006) Rayleigh-Debye-Gans as a model for continuous monitoring of biological particles: Part I, assessment of theoretical limits and approximations. Opt Express 14:8849–8865

Gaskin F, Cantor CR, Shelanski ML (1974) Turbidimetric studies of the in vitro assembly and disassembly of porcine neurotubules. J Mol Biol 89:737–755

Geiduschek EP, Holtzer AM (1958) Application of light scattering to biological systems. Adv Biol Med Phys 6:431–551

Ghosh P, Kumar A, Datta B, Rangachari V (2010) Dynamics of protofibril elongation and association involved in Aß42 peptide aggregation in Alzheimer's disease. BMC Bioinformatics 11:1

Goldsbury C, Baxa U, Simon MN, Steven AC, Engel A, Wall JS, Aebi U, Müller SA (2011) Amyloid structure and assembly: insights from scanning transmission electron microscopy. J Struct Biol 173:1–13

Goodsell DS, Dutta S, Zardecki C, Voigt M, Berman HM, Burley SK (2015) The RCSB PDB "Molecule of the month": Inspiring a molecular view of biology. PLoS Biol 13:e1002140

Greenwald J, Riek R (2010) Biology of amyloid: structure, function, and regulation. Structure 18:1244–1260

Guo S, Akhremitchev BB (2006) Packing density and structural heterogeneity of insulin amyloid fibrils measured by AFM nanoindentation. Biomacromolecules 7:1630–1636

Hall D (2002) On the role of macromolecular phase transitions in biology in response to changes in solution volume: action as an entropy buffer. Biophys Chem 98:233–48

Hall D (2003) The effects of denaturation of tubulin on the characterization of its polymerization behavior. Anal Biochem 104:655–82

Hall D (2006) Protein self-association in the cell: a mechanism for fine tuning the level of macromolecular crowding? Eur Biophys J 35: 276–280

Hall D (2012) Semi-automated methods for simulation and measurement of amyloid fiber distributions obtained from transmission electron microscopy experiments. Anal Biochem 421:262–277

Hall D, Edskes H (2004) Silent prions lying in wait: a two-hit model of prion/amyloid formation and infection. J Mol Biol 336:775–86

Hall D, Dobson CM (2006) Expanding to fill the gap: a possible role for inert biopolymers in regulating the extent of the 'macromolecular crowding'effect. FEBS Lett 580:2584–2590

Hall D, Edskes H (2009) A model of amyloid's role in disease based on fibril fracture. Biophys Chem 145:17–28

Hall D, Edskes H (2012) Computational modeling of the relationship between amyloid and disease. Biophys Rev 4:205–215

Hall D, Hirota N (2009) Multi-scale modelling of amyloid formation from unfolded proteins using a set of theory derived rate constants. Biophys Chem 140:122–128

Hall D, Huang L (2012) On the use of size-exclusion chromatography for the resolution of mixed amyloid-aggregate distributions (I) Equilibrium partition models. Anal Biochem 426:69–85

Hall D, Minton AP (2002) Effects of inert volume excluding macromolecules on protein fiber formation (I) equilibrium models. Biophys Chem 98:93–104

Hall D, Minton AP (2003) Macromolecular crowding: qualitative and semi-quantitative successes, quantitative challenges. Biochim Biophys Acta 1649:127–139

Hall D, Minton AP (2004) Effects of inert volume-excluding macromolecules on protein fiber formation. (II.) Kinetic models for nucleated fiber growth. Biophys Chem 107:299–316

Hall D, Minton AP (2005) An examination of the use of turbidity as a marker of tubulin polymerization kinetics. Anal Biochem 345:198–213

Hall D, Hirota N, Dobson CM (2005) A toy model for determining the rate of amyloid formation from denatured protein. J Mol Biol 351:195–205

Hall D, Edskes H, Kardos J, Carver JA, Goto Y (2015) A multi-pathway perspective on protein aggregation: implications for control of the rate and extent of amyloid formation. FEBS Lett 589:672–9

Hall D, Zhao R, Dehlsen I, Bloomfield N, Williams SR, Arisaka F, Goto Y, Carver JA (2016a) Protein aggregate turbidity: Simulation of turbidity profiles for mixed aggregation reactions. Anal Biochem 498:78–94

Hall D, Zhao R, So M, Adachi M, Rivas G, Carver JA, Goto Y (2016b) Recognizing and analyzing variability in amyloid formation kinetics: Simulation and statistical methods. Anal Biochem 510:56–71

Harper JD, Lieber CM, Lansbury PT (1997) Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-ß protein. Chem Biol 4:951–959

Hatters DM, Lindner RA, Carver JA, Howlett GJ (2001) The molecular chaperone, a-crystallin, inhibits amyloid formation by apolipoprotein C-II. J Biol Chem 276:33755–33761

Hergert W, Wriedt T (eds) (2012) The Mie Theory: basics and applications. Springer series in optical sciences, vol 169. Springer SBS, Berlin Heidelberg New York

Hill TL (1983) Length dependence of rate constants for end-to-end association and dissociation of equilibrium linear aggregates. Biophys J 44:285

Ingham B, Lim TH, Dotzler CJ, Henning A, Toney MF, Tilley RD (2011) How nanoparticles coalesce: an in situ study of Au nanoparticle aggregation and grain growth. Chem Mater 23:3312–3317

Jarrett JT, Lansbury PT Jr (1992) Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. Biochemistry 31:12345–12352

Jeong JS, Ansaloni A, Mezzenga R, Lashuel HA, Dietler G (2013) Novel mechanistic insight into the molecular basis of amyloid polymorphism and secondary nucleation during amyloid formation. J Mol Biol 425:1765–1781

Kanno T, Yamaguchi K, Naiki H, Goto Y, Kawai T (2005) Association of thin filaments into thick filaments revealing the structural hierarchy of amyloid fibrils. J Struct Biol 149:213–218

Kashchiev D (2015) Protein polymerization into fibrils from the viewpoint of nucleation theory. Biophys J 109:2126–2136

Kerker M, (2013) The scattering of light and other electromagnetic radiation. Physical chemistry: a series of monographs, vol 16. Academic Press, New York

King CY, Diaz-Avalos R (2004) Protein-only transmission of three yeast prion strains. Nature 428:319–323

Kirschner MW, Honig LS, Williams RC (1975) Quantitative electron microscopy of microtubule assembly in vitro. J Mol Biol 99:263–276

Klunk WE, Pettegrew JW, Abraham DJ (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. J Histochem Cytochem 37:1273–1281

Korolevskaya LB, Khlebtsov NG (2010) Spectroturbidimetric determination of the sizes of poly (ethylene glycol)-induced insoluble immune complex particles. Colloid J 72:504–511

Krebs MR, MacPhee CE, Miller AF, Dunlop IE, Dobson CM, Donald AM (2004a) The formation of spherulites by amyloid fibrils of bovine insulin. Proc Natl Acad Sci USA 101:14420–14424

Lee YH, Chatani E, Sasahara K, Naiki H, Goto Y (2009) A comprehensive model for packing and hydration for amyloid fibrils of ß2-microglobulin. J Biol Chem 284:2169–2175

Lee G, Lee W, Lee H, Lee CY, Eom K, Kwon T (2015) Self-assembled amyloid fibrils with controllable conformational heterogeneity. Sci Rep 5:16220

Levi AC, Kotrla M (1997) Theory and simulation of crystal growth. J Phys: Condens Matter 9:299

Levine H 3rd (1993) Thioflavin T interaction with Alzheimer's disease—amyloid peptides: detection of amyloid aggregation in solution. Protein Sci 2:404–10

Li H, Rahimi F, Sinha S, Maiti P, Bitan G. and Murakami K, (2009) Amyloids and protein aggregation —analytical methods. Encyclopedia of Analytical Chemistry. Wiley

Liu Y, Arnott WP, Hallett J (1998) Anomalous diffraction theory for arbitrarily oriented finite circular cylinders and comparison with exact T-matrix results. Appl Opt 37:5019–5030

Liu J, Costantino I, Venugopalan N, Fischetti RF, Hyman BT, Frosch MP, Gomez-Isla T, Makowski L (2016) Amyloid structure exhibits polymorphism on multiple length scales in human brain tissue. Sci Rep 6:33079

Lomakin A, Chung DS, Benedek GB, Kirschner DA, Teplow DB (1996) On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. Proc Natl Acad Sci USA 93:1125–1129

Loudon R (2000) The quantum theory of light. OUP, Oxford

Lu JX, Qiang W, Yau WM, Schwieters CD, Meredith SC, Tycko R (2013) Molecular structure of ß-amyloid fibrils in Alzheimer's disease brain tissue. Cell 154:1257–1268

Ma B, Nussinov R (2006) Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. Curr Opin Chem Biol 10:445–452

Mahler HC, Müller R, Frieß W, Delille A, Matheus S (2005) Induction and analysis of aggregates in a liquid IgG1-antibody formulation. Eur J Pharm Biopharm 59:407–417

Makin OS, Serpell LC (2002) Examining the structure of the mature amyloid fibril. Biochem Soc Trans 30:521–525

Makin OS, Serpell LC (2005) X-ray diffraction studies of amyloid structure. Methods Mol Biol 299: 67–80

Masel J, Jansen VA, Nowak MA (1999) Quantifying the kinetic parameters of prion replication. Biophys Chem 77:139–152

Meinhardt J, Sachse C, Hortschansky P, Grigorieff N, Fändrich M (2009) Aß (1–40) fibril polymorphism implies diverse interaction patterns in amyloid fibrils. J Mol Biol 386:869–877

Merz PA, Wisniewski HM, Somerville RA, Bobin SA, Masters CL, Iqbal K (1983) Ultrastructural morphology of amyloid fibrils from neuritic and amyloid plaques. Acta Neuropathol 60:113–124

Mezzenga R, Fischer P (2013) The self-assembly, aggregation and phase transitions of food protein systems in one, two and three dimensions. Rep Prog Phys 76:046601

Michaels TC, Knowles TP (2014) Role of filament annealing in the kinetics and thermodynamics of nucleated polymerization. J Chem Phys 140:214904

Mie G (1908) Contributions to the optics of turbid media, particularly of colloidal metal solutions. Contributions to the optics of turbid media, particularly of colloidal metal solutions. Ann Phys (Leipzig) 25: 377–445. Transl. into English in 1976 by AA Royal Aircraft Establishment, Farnborough, England)

Minton AP (2007) Static light scattering from concentrated protein solutions, I: General theory for protein mixtures and application to self-associating proteins. Biophys J 93:1321–1328

Mishra R, Sjölander D, Hammarström P (2011) Spectroscopic characterization of diverse amyloid fibrils in vitro by the fluorescent dye Nile red. Mol BioSyst 7:1232–1240

Mitraki A (2010) Protein aggregation: from inclusion bodies to amyloid and biomaterials. Adv Protein Chem Struct Biol 79:89–125

Moody TP, Donovan MA, Laue TM (1996) Turbidimetric studies of Limulus coagulin gel formation. Biophys J 71:2012

Morris AM, Watzky MA, Finke RG (2009) Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. Biochim Biophys Actas 1794:375–397

Mroczka J, Szczuczynski D (2010) Improved regularized solution of the inverse problem in turbidimetric measurements. Appl Opt 49:4591–4603

Naiki H, Hashimoto N, Suzuki S, Kimura H, Nakakuki K, Gejyo F (1997) Establishment of a kinetic model of dialysis-related amyloid fibril extension in vitro. Amyloid 4:223–232

Necula M, Kayed R, Milton S, Glabe CG (2007) Small molecule inhibitors of aggregation indicate that amyloid β oligomerization and fibrillization pathways are independent and distinct. J Biol Chem 282:10311–10324

Nichols MR, Moss MA, Reed DK, Lin WL, Mukhopadhyay R, Hoh JH, Rosenberry TL (2002) Growth of ß-amyloid (1–40) protofibrils by monomer elongation and lateral association. Characterization of distinct products by light scattering and atomic force microscopy. Biochemistry 41:6115–6127

Nicoud L, Lazzari S, Balderas Barragán D, Morbidelli M (2015) Fragmentation of amyloid fibrils occurs in preferential positions depending on the environmental conditions. J Phys Chem B 119:4644–4652

Nilsson MR (2004) Techniques to study amyloid fibril formation in vitro. Methods 34:151–60

Ogi H, Fukukshima M, Hamada H, Noi K, Hirao M, Yagi H, Goto Y (2014) Ultrafast propagation of β-amyloid fibrils in oligomeric cloud. Sci Rep. 4. doi:10.1038/srep06960

O'Nuallain B, Thakur AK, Williams AD, Bhattacharyya AM, Chen S, Thiagarajan G, Wetzel R (2006) Kinetics and Thermodynamics of Amyloid Assembly Using a High-Performance Liquid Chromatography–Based Sedimentation Assay. Methods Enzymol 413:34–74

Oosawa F, Asakura S (1975) Thermodynamics of the polymerization of protein. Academic Press, New York

Oosawa F, Kasai M (1962) A theory of linear and helical aggregations of macromolecules. J Mol Biol 4:10–21

Oster G (ed) (1955) Physical techniques in biological research, vol. 1: optical techniques. Elsevier, Amsterdam

Padrick SB, Miranker AD (2002) Islet amyloid: phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. Biochemistry 41:4694–4703

Pallitto MM, Murphy RM (2001) A mathematical model of the kinetics of β-amyloid fibril growth from the denatured state. Biophys J 81: 1805–1822

Paravastu AK, Leapman RD, Yau WM, Tycko R (2008) Molecular structural basis for polymorphism in Alzheimer's ß-amyloid fibrils. Proc Natl Acad Sci USA 105:18349–18354

Paravastu AK, Qahwash I, Leapman RD, Meredith SC, Tycko R (2009) Seeded growth of ß-amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. Proc Natl Acad Sci USA 106:7443–7448

Paul TJ, Hoffmann Z, Wang C, Shanmugasundaram M, DeJoannis J, Shekhtman A, Lednev IK, Yadavalli VK, Prabhakar R (2016) Structural and mechanical properties of amyloid beta fibrils: a combined experimental and theoretical approach. J Phys Chem Lett 7:2758–2764

Penzkofer A, Shirdel J, Zirak P, Breitkreuz H, Wolf E (2007) Protein aggregation studied by forward light scattering and light transmission analysis. Chem Phys 342:55–63

Pepys M (2001) Pathogenesis, diagnosis and treatment of systemic amyloidosis. Philos Trans R Soc Lond B 356:203–221

Perlmann GE, Longsworth LG (1948) The specific refractive increment of some purified proteins. J Am Chem Soc 70:2719–2724

Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycko R (2005) Self-propagating, molecular-level polymorphism in Alzheimer's ß-amyloid fibrils. Science 307:262–265

Powers ET, Powers DL (2006) The kinetics of nucleated polymerizations at high concentrations: amyloid fibril formation near and above the "supercritical concentration". Biophys J 91:122–132

Qiang W, Yau WM, Tycko R (2011) Structural evolution of Iowa mutant ß-amyloid fibrils from polymorphic to homogeneous states under repeated seeded growth. J Am Chem Soc 133:4018–4029

Rayleigh L (1899) XXXIV. On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky. Philos Mag 47:375–384

Raynes JK, Gerrard JA (2013) Amyloid fibrils as bionanomaterials. In: Rehm BHA (ed) Bionanotechnology: biological self-assembly and its applications. Horizon Scientific Press, Norfolk, pp 85–106

Ridgley DM, Barone JR (2013) Evolution of the amyloid fiber over multiple length scales. ACS Nano 7:1006–1015

Rodina EV (2012) Nanomaterials based on peptides. Polym Sci Ser C 54: 88–95

Rogers SS, Venema P, Sagis LM, van der Linden E, Donald AM (2005) Measuring the length distribution of a fibril system: a flow birefringence technique applied to amyloid fibrils. Macromolecules 38:2948–2958

Ruth L, Eisenberg D, Neufeld EF (2000) α-L-Iduronidase forms semi-crystalline spherulites with amyloid-like properties. Acta Crystallogr, Sect D: Biol Crystallogr 56:524–528

Sant'Anna R, Gallego P, Robinson LZ, Pereira-Henriques A, Ferreira N, Pinheiro F, Esperante S, Pallares I, Huertas O, Almeida MR and Reixach N (2016) Repositioning tolcapone as a potent inhibitor of transthyretin amyloidogenesis and associated cellular toxicity. Nature Commun 7:10787

Sasahara K, Goto Y (2013) Application and use of differential scanning calorimetry in studies of thermal fluctuation associated with amyloid fibril formation. Biophys Rev 5:259–269

Sasahara K, Hall D, Hamada D (2010) Effect of lipid type on the binding of lipid vesicles to islet amyloid polypeptide amyloid fibrils. Biochemistry 49:3040–3048

Scherpelz KP, Lu JX, Tycko R, Meredith SC (2016) Preparation of amyloid fibrils seeded from brain and meninges. Methods Mol Biol 1345:299–312

Shmakov SL (2014) Algorithm to calculate limiting cumulative particle size distribution functions from turbidimetric data. Appl Opt 53:301–305

Silver FH, Birk DE (1983) Kinetic analysis of collagen fibrillogenesis: I. Use of turbidity-time data. Coll Relat Res 3:393–405

Smith JF, Knowles TP, Dobson CM, MacPhee CE, Welland ME (2006) Characterization of the nanoscale properties of individual amyloid fibrils. Proc Natl Acad Sci USA 103:15806–15811

So M, Hall D, Goto Y (2016) Revisiting supersaturation as a factor determining amyloid fibrillation. Curr Opin Struct Biol 36:32–39

Stranks SD, Ecroyd H, Van Sluyter S, Waters EJ, Carver JA, Von Smekal L (2009) Model for amorphous aggregation processes. Phys Rev E 80:051907

Symmers WSC (1956) Primary amyloidosis: a review. J Clin Pathol 9: 187–211

Szavits-Nossan J, Eden K, Morris RJ, MacPhee CE, Evans MR, Allen RJ (2014) Inherent variability in the kinetics of autocatalytic protein self-assembly. Phys Rev Lett 113:098101

Tachibana H, Husimi Y, Wada A (1977) Kinetics of the polymerization reaction of tobacco mosaic virus protein; transient-saturation type polymerization reaction. Biophys Chem 6:173–189

Tanaka M, Chien P, Naber N, Cooke R, Weissman JS (2004) Conformational variations in an infectious protein determine prion strain differences. Nature 428:323–328

Toyama BH, Weissman JS (2011) Amyloid structure: conformational diversity and consequences. Ann Rev Biochem 80:557–585

Tycko R (2011) Solid state NMR studies of amyloid fibril structure. Annu Rev Phys Chem 62:279

Tycko R (2014) Physical and structural basis for polymorphism in amyloid fibrils. Protein Sci 23:1528–1539

Tycko R, Wickner RB (2013) Molecular structures of amyloid and prion fibrils: consensus versus controversy. Acc Chem Res 46:1487–1496

Umemoto A, Yagi H, So M, Goto Y (2014) High-throughput analysis of ultrasonication-forced amyloid fibrillation reveals the mechanism underlying the large fluctuation in the lag time. J Biol Chem 289:27290–27299

Usov I, Mezzenga R (2015) FiberApp: an open-source software for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects. Macromolecules 48:1269–1280

van de Hulst HC (1957) Light scattering by small particles. Wiley, New York

Van Melckebeke H, Wasmer C, Lange A, AB E, Loquet A, Bockmann A, Meier BH (2010) Atomic-resolution three-dimensional structure of HET-s (218–289) amyloid fibrils by solid-state NMR spectroscopy. J Am Chem Soc 132:13765–13775

Virchow R (1854) Virchows Arch Pathol Anat 6:135, as cited by Symmers (1956) Primary Amyloidosis: A Review. J Clin Pathol 9: 187–211

von Smoluchowski M (1916) Drei vortrage uber diffusion. Brownsche bewegung und koagulation von kolloidteilchen. Z Phys 17:557–585

von Smoluchowski M (1917) Grundriß der Koagulationskinetik kolloider Lösungen. Colloid Polym Sci 21:98–104

Voter WA, Erickson HP (1984) The kinetics of microtubule assembly. Evidence for a two-stage nucleation mechanism. J Biol Chem 259: 10430–10438

Walker LC, Jucker M (2015) Neurodegenerative diseases: expanding the prion concept. Annu Rev Neurosci 38:87–103

Wallach ML, Heller W, Stevenson AF (1961) Theoretical investigations on the light scattering of colloidal spheres. XII. The determination of size distribution curves from turbidity spectra. J Chem Phys 34:1796–1802

Wegner A, Engel J (1975) Kinetics of the cooperative association of actin to actin filament. Biophys Chem 3:215–225

Wetzel R (2006) Kinetics and thermodynamics of amyloid fibril assembly. Acc Chem Res 39:671–679

Wisniewski HM, Bancher C, Barcikowska M, Wen GY, Currie J (1989) Spectrum of morphological appearance of amyloid deposits in Alzheimer's disease. Acta Neuropathol 78:337–347

Woolfson DN, Ryadnov MG (2006) Peptide-based fibrous biomaterials: some things old, new and borrowed. Curr Opin Chem Biol 10:559–567

Wyatt PJ (2014) Measurement of special nanoparticle structures by light scattering. Anal Chem 86:7171–7183

Xu Z, Li L, Li H, Gao F (2016) Synthesis of self-assembled noble metal nanoparticle chains using amyloid fibrils of lysozyme as templates. Nanomaterials Nanotechnol. doi: 10.5772/62182

Xue WF, Homans SW, Radford SE (2008) Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. Proc Natl Acad Sci USA 105:8926–8931

Xue WF, Hellewell AL, Gosal WS, Homans SW, Hewitt EW, Radford SE (2009) Fibril fragmentation enhances amyloid cytotoxicity. J Biol Chem 284:34272–34282

Yamaguchi KI, Takahashi S, Kawai T, Naiki H, Goto Y (2005) Seeding-dependent propagation and maturation of amyloid fibril conformation. J Mol Biol 352:952–960

Yoshimura Y, Lin Y, Yagi H, Lee YH, Kitayama H, Sakurai K, So M, Ogi H, Naiki H, Goto Y (2012) Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. Proc Natl Acad Sci USA 109:14446–14451

Zimm BH, Dandliker WB (1954) Theory of light scattering and refractive index of solutions of large colloidal particles. J Phys Chem 58:644–648

Zurdo J, Guijarro JI, Jiménez JL, Saibil HR, Dobson CM (2001) Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. J Mol Biol 311:325–340

Future Studies

In my postgraduate studies I initially set out with the aim of conducting protein aggregation assays to test the efficacy of the molecular chaperones (WT 14-3-3zeta, WT alpha B-crystallin, S59D alpha B-crystallin and S19,45,59D alpha B-crystallin). In each assay, there was one aggregate substrate (made from either ADH, lysozyme, alpha-lactalbumin, RCM alpha-lactalbumin, RCM kappa-casein, insulin) and either one chaperone, or two chaperones used together in combination. Each of these interactions were investigated by use of the turbidimetric assay technique. During the course of my studies I realized that I didn't really understand what factors produced the turbidimetric signal. As this was the basic measuring stick of my project I decided to spend some time digging deeper into the basic physical phenomenon underlying this primary assay technique.

After having looked more closely at the fundamental underlying principles of this assay technique I now have a better understanding of the biological meaning hidden within these assays. In future work I plan to extend the results contained within this thesis by developing an assay deconvolution technique which uses both data from experiment and the developed equations for estimating the turbidity signal to determine the underlying solution composition. While such approaches have been tried before I think that the Ockham's razor type simplification in our simulation approach, basing the signal on just two components (an amorphous and a fibrillar species) will make deconvolution of the signal both experimentally meaningful and achievable. To help this process we plan on factoring in information relating to the wavelength

dependence of the turbidity signal. A future analytical routine will take the form of a constrained non-linear least squares approach in which the total turbidity at a number of different wavelengths is compared to a theoretical estimate of the turbidity generated from combinations of the two different types of protein aggregate (amorphous or rod-like). I have already made some progress in this area by developing an automated diagnostic method for examining turbidimetric wavelength dependence of protein aggregation reactions that represents an improvement of the methods advanced previously (Camerino-Otero and Day, 1978).

One important aspect of my thesis work was the development of statistical measures for discerning meaningful differences in protein aggregation kinetics (Hall et al. 2016b) on the condition that the signal measure was linear with respect to protein aggregate weight concentration. This aspect of my work preceded my attempts to correctly/completely de-convolute the turbidity spectrum and as such, was slightly premature (as my preference was for a linear approach to ironing out the problems inherent in the assay field). With this said however the adoption (by the field) of the rigorous statistical approaches presented in Hall et al. 2016b may turn out to be as, or more, important than the empirical equation set for describing the turbidity of mixed aggregation reactions. Previously, the assessment of differences in aggregation assays has been performed in a very primitive, potentially naïve fashion, manly based on the calculation of simple averages. In my opinion, this reflects a general lack of appreciation of metrological science by the many practicing medical science researchers of the protein aggregation community. The relatively straightforward solutions to basic assay related questions (e.g. what type of things

can provide a measure of difference and how do I measure a statistical difference) was, in my opinion, a much required input to the field and as such my contribution is something that I will remain proud of for a very long time.

A final aspect of my thesis work which deserves mention is the demonstration (and subsequent discussion) of 'non-regular' amyloid kinetics for different kinetic regimes of protein aggregate growth (Zhao et al. 2016). In that Review article I discussed how the turbidity generated by protein aggregation would be affected by changes in fibre width, fibre breakage, fibre end-to-end joining, fibre side-to-side joining, competitive growth between different types of fibre species and cases involving fibre species competing against amorphous aggregate types. Through the use of synthesizing 'Review simulations' I managed to show that turbidity measurements may produce non-asymptotic kinetic forms that defy normal parameterization schemes. In practice, such data is often ignored or heavily truncated and treated as if something went wrong with the experiment. If nothing else was achieved in my thesis the demonstration that such aberrant data is actually informative (rather than problematic) would have made my postgraduate studies worthwhile.

As is often the case in science I did not expect my studies to lead me to this point. My meandering route has certainly meant that I have jumped disciplines (biology, chemistry, physics and mathematics). Ultimately though I feel that I am now at a point where I (or others continuing on from my efforts/achievements) can make meaningful progress into understanding how turbidimetric assays can be used to explore just what effects

chaperones might play on any protein aggregation reaction occurring simultaneously in that chaperone containing solution. Ultimately my interest in biological chemistry has been strengthened by the path followed and I have developed a real appreciation for the non-linear nature of the signals we so often interpret in a facile (and potentially incorrect) fashion.

References

Adachi, M., So, M., Sakurai, K. *et al.,* 2015. Supersaturation-limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation. Journal of Biological Chemistry, 290: 18134-45.

Alberts, B., Johnson, A., Lewis, J., *et al*. (2002). The Shape and Structure of Proteins. Molecular Biology of the Cell; Fourth Edition. New York and London: Garland Science.

Andley, U. (2007). Crystallins in the eye: Function and pathology. Progress in Retinal and Eye Research, 26(1): 78-98.

Bohren, C.F., and Huffman, D.R. (2008). Absorption and scattering of light by small particles. John Wiley & Sons.

Camerini-Otero, R.D. and Day, L.A. (1978). The wavelength dependence of the turbidity of solutions of macromolecules. Biopolymers, 17: 2241-49.

Carver, J.A., Aquilina, J.A., Truscott, R.J.W. *et al*. (1992). Identification by 1H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. FEBS letters, 311(2): 143-9.

Carver, J.A., Rekas, A., Thorn, D.C. *et al*. (2003). Small Heat-shock Proteins and Clusterin Intra- and Extracellular Molecular Chaperones with a Common Mechanism of Action and Function. IUBMB Life, 55(12): 661-8.

Chaudhuri, T.K. and Paul, S. (2006). Protein-misfolding diseases and chaperone-based therapeutic approaches. FEBS J 273(7): 1331-49.

Cohen, F. and Kelly, J. (2003). Therapeutic approaches to protein-misfolding diseases. Nature 426: 905-9.

Dobson, C.M. (2001). The structural basis of protein folding and its links with human disease. Philosophical Transactions of the Royal Society B, 356: 133-45.

Dolado, I., Nieto, J., Saraiva, M.J.M. *et al*. (2005). Kinetic assay for high-throughput screening of in vitro transthyretin amyloid fibrillogenesis inhibitors. Journal of Combinatorial Chemistry, 7: 246-52.

Doty, P. and Steiner, R.F. (1950). Light scattering and spectrophotometry of colloidal solutions. The Journal of Chemical Physics, 18: 1211-20.

Ecroyd, H. and Carver, J. (2008). Unraveling the mysteries of protein folding and misfolding. IUBMB Life, 60(12): 769-74.

Ecroyd, H. and Carver, J. (2009). Crystallin proteins and amyloid fibrils. Cellular and Molecular Life Sciences, 66(1): 62-81.

Ecroyd, H., Meehan, S., Horwitz, J. *et al.* (2007). Mimicking phosphorylation of αB-crystallin affects its chaperone activity. Biochemical Journal, 401: 129-41.

Fu, H., Subramanian, R. and Masters, S. (2000). 14-3-3 Proteins: Structure, Function, and Regulation. Annual Review of Pharmacology and Toxicology, 40: 617-47.

Garcia-Lopez, A. C. and Garcia-Rubio, L. H. (2008). Rayleigh-Debye-Gans as a model for continuous monitoring of biological particles: Part II, development of a hybrid model. Optics Express, 16: 4671-87.

Garcia-Lopez, A. C., Snider, A. D. and Garcia-Rubio, L. H. (2006). Rayleigh-Debye-Gans as a model for continuous monitoring of biological particles: Part I, assessment of theoretical limits and approximations. Optics Express, 14: 8849-65.

Hall, D., (2012). Semi-automated methods for simulation and measurement of amyloid fiber distributions obtained from transmission electron microscopy experiments. Analytical Biochemistry. 421: 262-77.

Hall D., Edskes H., Kardos J. *et al.,* (2015). A Multi-Pathway Perspective on Protein Aggregation: Implications for Control of the Rate and Extent of Amyloid Formation. FEBS Letters. 589: 672-9.

Hall, D. and Minton, A. P. (2005). Turbidity as a probe of tubulin polymerization kinetics: a theoretical and experimental re-examination. Analytical Biochemistry, 345: 198-213.

Hall, D., Zhao. R., Dehlsen, I. *et al*. (2016a). Protein aggregate

turbidity: Simulation of turbidity profiles for mixed aggregation reactions. Analytical Biochemistry, 498: 78-94.

Hall, D., Zhao. R., So, M. *et al*. (2016b). Recognizing and analyzing variability in amyloid formation kinetics: Simulation and statistical methods. Analytical Biochemistry, 510: 56-71.

Hall, D. and Dobson, C.M. (2006). Expanding to fill the gap: a possible role for inert biopolymers in regulating the extent of the 'macromolecular crowding'effect. FEBS letters, 580: 2584-90.

Hall, D. and Edskes, H. (2009). A model of amyloid's role in disease based on fibril fracture. Biophysical chemistry, 145: 17-28.

Hall, D. and Edskes, H. (2012). Computational modeling of the relationship between amyloid and disease. Biophysical reviews, 4: 205-22.

Hall, D. (2002). On the role of the macromolecular phase transitions in biology in response to change in solution volume or macromolecular composition: action as an entropy buffer. Biophysical chemistry. 98: 233-48.

Hlavacek, W.S., Posner, R.G. and Perelson, A.S. (1999). Steric effects on multivalent ligand-receptor binding: exclusion of ligand sites by bound cell surface receptors. Biophysical journal, 76: 3031-43.

Horwitz, J., Huang, Q., Ding, Q. *et al.* (1998). Lens α-crystallin: Chaperone-like properties. Methods in Enzymology, 290: 365-83.

Huang, Y. and Mucke, L. (2012). Alzheimer Mechanisms and Therapeutic Strategies. Cell, 148(6): 1204-22.

Invernizzi, G., Papaleo, E., Sabate, R. *et al.* (2012). Protein aggregation: Mechanisms and functional consequences. The international journal of biochemistry & cell biology, 44: 1541-54.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. Journal of Neurology, Neurosurgery, and Psychiatry, 79: 368-76.

Jarrett, J.T. and Lansbury, P.T. (1993). Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? Cell, 73(6): 1055-8.

Kato, K., Shinohara, H., Kurobe, N. *et al*. (1991). Tissue distribution and developmental profiles of immunoreactive alphaB-crystallin in the rat determined with a sensitive immunoassay system. Biochimica et Biophysica Acta, 1074: 201-8.

Kelly, J. (1998). Alternative conformation of amyloidogenic proteins and their multi-step assembly pathways. Current Opinion in Structural Biology, 8: 101-6.

Korolevskaya, L. B., and Khlebtsov, N. G. (2010). Spectroturbidimetric determination of the sizes of poly (ethylene glycol)-induced insoluble immune complex particles. Colloid Journal, 72: 504-11.

Mahler, H. C., Friess, W., Grauschopf, U. *et al.,* (2009). Protein aggregation: pathways, induction factors and analysis. Journal of Pharmaceutical Sciences, 98: 2909-34.

Martin, H., Rostas, J., Patel, Y. *et al*. (1994). Subcellular localisation of 14-3-3 isoforms in rat brain using specific antibodies. Journal of Neurochemistry, 63: 2259-65.

Moody, T. P., Donovan, M. A. and Laue, T. M. (1996). Turbidimetric studies of Limulus coagulin gel formation. Biophysical Journal, 71: 2012-21.

Moore, BW. and Perez, VJ. (1967). Specific Acid Proteins in the Nervous System. In Physiological and Biochemical Aspects of Nervous Integration, ed. FD Carlson, pp. 343-59.

Muslin, AJ., Tanner, JW., Allen, PM. *et al*. (1996). Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. Cell, 84(6): 889-97.

Sant'Anna, R., Gallego, P., Robinson, L.Z. *et al*. (2016). Repositioning tolcapone as a potent inhibitor of transthyretin amyloidogenesis and associated cellular toxicity. Nature Communications, 7.

Sato, S., Chiba, T., Sakata, E. *et al*. (2006). 14-3-3eta is a novel regulator of parkin ubiquitin ligase. The EMBO Journal, 25: 211-21.

Silver, F. H. and Birk, D. E. (1983). Kinetic analysis of collagen fibrillogenesis: I. Use of turbidity-time data. Collagen and Related Research, 3: 393-405.

Stefani, M. (2004). Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. Biochimica et Biophysica Acta, 1739 (1): 5-25.

Stoppini, M. and Bellotti, V. (2015). Systemic amyloidosis: lessons from β2-microglobulin. Journal of Biological Chemistry, 290: 9951-8.

Sun, S., Wong, E.W., Li, M.W. *et al*. (2009). 14-3-3 and its binding partners are regulators of protein–protein interactions during spermatogenesis. Journal of Endocrinology, 202(3): 327-36.

Treweek, T.M., Morris, A.M. and Carver, J.A. (2003). Intracellular protein unfolding and aggregation: the role of small heat-shock chaperone proteins. Australian Journal of Chemistry, 56 (5): 357-67.

Treweek T.M., Rekas A., Walker M.J., et al. (2010). A quantitative NMR spectroscopic examination of the flexibility of the C-terminal extensions of the molecular chaperones, aA- and aB-crystallin. Experimental Eye Research, 91: 691-99.

Usov, I. and Mezzenga, R. (2015). FiberApp: an open-source software for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects. Macromolecules, 48: 1269-80.

Walker, L. C. and LeVine, III, H. (2000). The cerebral proteopathies: Neurodegenerative disorders of protein conformation and assembly. Molecular Neurobiology, 21: 83-95.

Williams, D., Ecroyd, H., Goodwin, K. *et al*. (2011). NMR spectroscopy of 14-3-3zeta reveals a flexible C-terminal extension: differentiation of the chaperone and phosphoserine-binding activities of 14-3-3zeta. Biochemical Journal, 437(3): 493-503.

Xu, Z., Graham, K., Foote, M. *et al*. (2013).14-3-3 protein targets misfolded chaperone-associated proteins to aggresomes. Journal of Cell Science, 126: 4173-86.

Yano, M., Nakamuta, S., Wu, X., *et al*. (2006). A novel function of 14-3-3 protein: 14-3-3zeta is a heat-shock-related molecular chaperone that dissolves thermal-aggregated proteins. Molecular Biology of the Cell, 17: 4769-79.

Zhao, R., So, M., Maat, H. *et al* (2016). Measurement of Amyloid Formation by Turbidity Assay – Seeing Through the Cloud. Biophysics Review, 8: 445-71.