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.

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/

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.

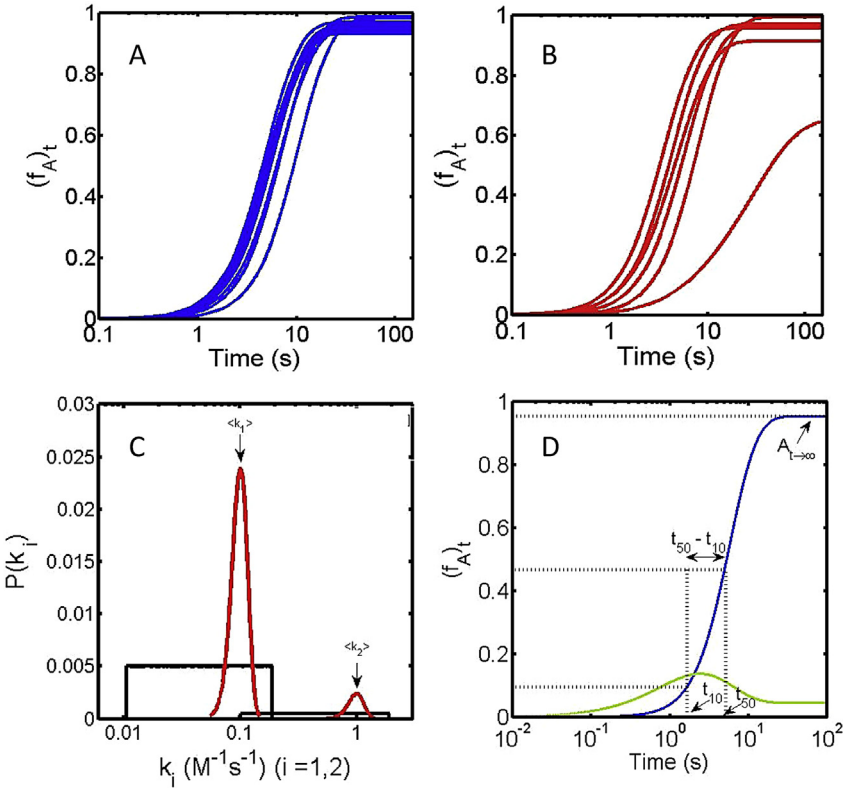

Fig. 1. Simulated replicate kinetic data sets showing the fraction of amyloid formed as a function of time. Simulations were carried out using Eqns. A1 and A2 whereby the input rate constants are sampled from distributions affording (A) Low variation (blue) (B) High variation (red). (C) Probability distribution 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.)