Journal of Theoretical Biology 397 (2016) 68–88

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Estimation from moments measurements for amyloid depolymerisation

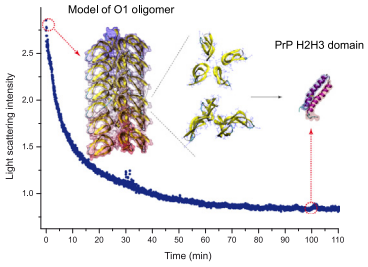
Aurora Armiento [a,b,c,*], Marie Doumic [b], Philippe Moireau [c], H. Rezaei [d]

[a] Univ Paris Diderot, Sorbonne Paris Cité, Lab. J.L. Lions, UMR CNRS 7598, Inria , Paris, France
[b] Sorbonne Universités, Inria, UPMC Univ Paris 06, Lab. J.L. Lions UMR CNRS 7598, Paris, France
[c] Inria and Université Paris-Saclay, Campus de l'Ecole Polytechnique, 91128 Palaiseau, France
[d] Virologie et Immunologie Moléculaires, Institut National de la Recherche Agronomique, F-78352 Jouy-en-Josas, France

HIGHLIGHTS

- A framework methodology for estimating cluster concentration in protein aggregation.
- Theoretical solution with kernel method in a depolymerising system.
- Theoretical solution with 4dVar data assimilation method in a depolymerising system.
- Our moment-based method is tested on experimental data of ovPrP oligomers.
- Smaller aggregates, which are probably the most cytotoxic species, reveal more stable.

GRAPHICAL ABSTRACT



Structural model of PrP O1 oligomer depolymerization into α-enriched monomeric PrP as explored in the present work.

ARTICLE INFO

Article history:
Received 18 December 2015
Received in revised form
23 February 2016
Accepted 26 February 2016
Available online 4 March 2016

Keywords:
Amyloid
Prion
Protein stability
Oligomer
Transport equation
State estimation
Inverse problem
Data assimilation

ABSTRACT

Estimating reaction rates and size distributions of protein polymers is an important step for understanding the mechanisms of protein misfolding and aggregation, a key feature for amyloid diseases. This study aims at setting this framework problem when the experimental measurements consist in the time-dynamics of a moment of the population (i.e. for instance the total polymerised mass, as in Thioflavin T measurements, or the second moment measured by Static Light Scattering). We propose a general methodology, and we solve the problem theoretically and numerically in the case of a depolymerising system. We then apply our method to experimental data of depolymerising oligomers, and conclude that smaller aggregates of ovPrP protein should be more stable than larger ones. This has an important biological implication, since it is commonly admitted that small oligomers constitute the most cytotoxic species during prion misfolding process.

1. Introduction

Protein aggregation is a key feature of a large range of diseases, called *amyloid* diseases, among which we can quote Alzheimer's, Parkinson's, Huntington's, transmissible spongiform encephalopaties (or prion diseases – e.g. Creutzfeldt–Jakob's, Kuru, bovine spongiform encephalopathy/madcow), etc. (Ow and Dunstan, 2014; Ramìrez-Alvarado et al., 2000).

This category of diseases takes its name from the protein fibrils, called *amyloids*, which are formed during the disease and accumulate into the tissue. Their formation arise from misfolded

* Corresponding author.
E-mail address: aurora.armiento@inria.fr (A. Armiento).

http://dx.doi.org/10.1016/j.jtbi.2016.02.037
0022-5193/

versions of proteins present naturally in the body, each disease having its specific precursor protein (e.g. APP for Alzheimer's, PrP for Prion, β_2m for haemodialysis-associated amyloidosis). While their accumulation in organs is characteristic for the disease, the reason for their association as well as their role in tissue damages are still unclear. Moreover, their aggregation mechanisms – most probably specific for each protein involved – are at the moment largely unknown.

The main reasons for so many open questions, despite both the longstanding interest raised in the biological, biophysical and biochemical communities, and the major importance of amyloid diseases for public health, are twofold. First, the number of possible chain-reactions involved is huge, possibly infinite – as the size of aggregates is. Hence model design and discrimination is very complex and conclusions made on a specific protein are hardly translatable to another one. Second, the most common experimental devices can measure averaged quantities on the polymerised proteins, such as the total polymerised mass (Thioflavin T measurements (Batzli and Love, 2015)) or the average size of polymers (Static Light Scattering (SLS) (Some et al., 2013)). How such measurements may be used to estimate reaction rates (which may also be an infinity) and size distribution of aggregates, and thus to select the major mechanisms, is an emerging field of inverse problems with few theoretical progress (Banks et al., 2015) and positive results on experimental data (Xue et al., 2008; Xue and Radford, 2013).

To contribute to this new field, this paper focuses on one of the major concerns in pathologies due to protein misassembly and aggregation: the determination of oligomer size distribution. It has been reported that – while amyloid fibrils present low biological activity – oligomers and small assemblies are the cytopathogenic elements (Simoneau et al., 2007; James Gilbert, 2013). Depending on the type of pathology and the protein involved, oligomers could either be involved into the pathway of amyloid fibrils formation or be associated to an independent pathway, which only leads to the formation of oligomers. Oligomer size characterisation can play a key role in distinguishing between these pathways. Therefore, the investigation on size distribution remains the first step to understand how oligomers are formed, their biological activity and their biophysical characterisation to finally design therapeutic strategies.

This question – how to estimate size distributions – leads us to setting a framework problem and studying it, both theoretically and numerically, in one of its simplest possible version. We then apply our method to experimental data, using the time-dependent average size of polymers (measured by SLS) to reconstruct the oligomer initial size distribution. We compare our estimation to the experimental estimation obtained by chromatography and discuss the implications of our results. Eventually, we discuss the new problems and possibilities opened-up by these results, and how this methodology could easily be adapted to other models and experiments.

1.1. Mathematical setting

Since protein aggregates can reach extremely large average sizes, we adopt here a continuous framework (Prigent et al., 2012) and denote $x \in (0,\infty)$ the *size* of an aggregate, i.e. x represents the (rescaled) quantity of monomers contained in a given polymer. We thus call $u(x,t)$ the concentration of polymers of size x at time t (see Banks et al. (2014) for a discussion and theoretical justification of the use of a continuous size variable rather than a discrete one).

One of the techniques most widely used is the measurement of Thioflavin T (ThT) fluorescence, (Batzli and Love, 2015), which provides measurements of the total polymerised mass, i.e a linear transformation of the first moment of the concentration function

$$z_{\text{tht}}(t) = c_1 \int_0^\infty x u(x,t)\, dx.$$

The Static Light Scattering (SLS) technique, (Some et al., 2013), could give us an affine transformation of the second moment

$$z_{\text{sls}}(t) = c_1 \int_0^\infty x^2 u(x,t)\, dx + c_2,$$

where $c_1 \geq 0, c_2 \in \mathbb{R}$.

The framework problem we want to contribute stands: Under which assumptions (and limitations) is it possible to estimate the reaction rates and/or the initial size distribution, from a time measurement $z_{\text{tht}}(t)$ or $z_{\text{sls}}(t)$?

As a first simplifying assumption, we model the primary reactions involved in the evolution of polymers with the Lifshitz–Slyozov system, that is one of the most common polymerising/depolymerising model. In this system, polymers (or clusters, in another application context) can only grow by monomer addition, with a size-dependent reaction rate $a(x)$, and depolymerise by monomer loss, with a reaction rate $b(x)$. This results in the following system

$$\begin{cases} \dfrac{\partial}{\partial t}u(x,t) + \dfrac{\partial}{\partial x}((a(x)v(t) - b(x))u(x,t)) = 0, & x \in [0,\ell],\ t \geq 0, \\ u(\ell,t) = 0, \\ u(x,0) = u_0(x), \end{cases} \tag{1}$$

where $\ell \in (0,\infty]$ is the upper bound of polymer sizes. We assume here $\ell < \infty$, in contrast with the initial Lifshitz–Slyozov model (Lifshitz, 1961). The function $v(t)$ is the concentration of monomers in the cuvette and is directly related to polymer concentration from the following mass conservation law

$$v(t) + \int_0^\ell x u(x,t)\, dx = v(0) + \int_0^\ell x u_0(x)\, dx > 0 \quad \forall\ t \geq 0. \tag{2}$$

When applied to amyloid formation, this model may be seen as a qualitative model taking into account what biologists call *primary pathway* and neglecting, as a first approach, *secondary pathways* such as fragmentation or coalescence (Bishop and Ferrone, 1984). Note that there are many other possible applications of this model, such as phase transition, which was the original application for which it had been designed (Lifshitz, 1961).

The problem now stands: measuring $z_{\text{sls}}(t)$ or $z_{\text{tht}}(t)$, or more generally the time-dependence of a n-th moment defined by $\int_0^\ell x^n u(x,t)\, dx$, with $u(x,t)$ solution of Systems (1) and (2), what may be possibly estimated among the unknown quantities, i.e. the initial state $u_0(x)$ and the parameter functions $a(x)$ and $b(x)$?

This problem in its full generality is both nonlinear and highly ill-posed. Hence, we proceed to further simplifications and study the state estimation of a model of pure depolarisation. Assuming to start with no monomers, i.e. $v(0) = 0$, we can neglect the polymerisation term, at least during the beginning of the reaction – see Fig. 12 for measurements of such an experiment. The model then becomes

$$\begin{cases} \dfrac{\partial}{\partial t}u(x,t) - \dfrac{\partial}{\partial x}(b(x)u(x,t)) = 0, & x \in [0,\ell],\ t \geq 0, \\ u(\ell,t) = 0, \\ u(x,0) = u_0(x). \end{cases} \tag{3}$$

The state estimation problem may be formulated as follows:

(IP) How to estimate u_0 – the initial condition of System (3) – from the given *a priori* knowledge of $b(x)$ and measurement $\int_0^\ell x^n u(x,t)dx$?

In order to settle a general framework, easy to adapt to more complex problems in the future, we introduce below the notations