Applied Mathematics and Computation 393 (2021) 125778

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A network thermodynamic analysis of amyloid aggregation along competing pathways

P. Ghosh [a], J. Pateras [b], V. Rangachari [c,d], A. Vaidya [b,*]

[a] Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23220, USA
[b] Department of Mathematics, Montclair State University, Montclair, NJ 07043, USA
[c] Department of Chemistry & Biochemistry, 118 College Dr # 5043 University of Southern Mississippi, Hattiesburg, MS 39406, USA
[d] Center for Cellular and Molecular Biosciences, University of Southern Mississippi, Hattiesburg MS 39406, USA

ARTICLE INFO

Article history:
Received 13 June 2020
Revised 13 September 2020
Accepted 28 October 2020
Available online 18 November 2020

Keywords:
Network thermodynamics
Gibbs energy
Amyloid aggregation
Game-Theory
Categories
Neurodegenerative diseases
Oligomers

ABSTRACT

Self-assembly of proteins towards amyloid aggregates is a significant event in many neurodegenerative diseases. Aggregates of low-molecular weight called oligomers are largely the primary toxic agents in many of these maladies. Therefore, there is an increasing interest in understanding their formation and behavior. In this paper, we build on our previously established theoretical investigations on the interactions between Aβ and lipids (L) that induces off-pathway aggregates under the control of L concentrations. Here, our previously developed competing game theoretic framework between the on- and off-pathway dynamics has been expanded to understand the underlying network topological structures in the reaction kinetics of amyloid formation. The mass-action based dynamical systems are solved to identify dominant pathways in the system with fixed initial conditions, and variations in the occurrence of these dominant pathways are identified as a function of various seeding conditions. The mechanistic approach is supported by thermodynamic free energy computations which helps identify stable reactions. The resulting analysis provides possible intervention strategies that can draw the dynamics away from the off-pathways and potential toxic intermediates. We also draw upon the classic literature on network thermodynamics to suggest new approaches to better understand such complex systems.

1. Introduction

Aggregation of misfolded or disordered proteins and concomitant amyloid formation is a key pathogenic event in many neurodegenerative and other maladies [3]. Amyloid aggregation is a nucleation-dependent process in which a rate-limiting formation of an aggregate with a specific mass and conformation is a critical step for the kinetically-favored elongation process that follows [3]. Molecular details involved during the pre-nucleation events remain unclear, but are believed to be stochastic, especially for intrinsically disordered proteins such as amyloid-β (Aβ) and α-synuclein (αS), which are involved in Alzheimer disease (AD) and Parkinson disease (PD), respectively. In several neurodegenerative disorders, low molecular weight oligomers formed along the aggregation pathway are known to be the most neurotoxic species [3,23], making the rate-limiting step of nucleus formation to be one of the key events in the etiology of these diseases. Highly stochastic reactions during the prenucleation phase renders proetins to adopt multiple aggregation pathways (on- and off-) and different

* Corresponding author.
E-mail address: vaidyaa@mail.montclair.edu (A. Vaidya).

https://doi.org/10.1016/j.amc.2020.125778
0096-3003/

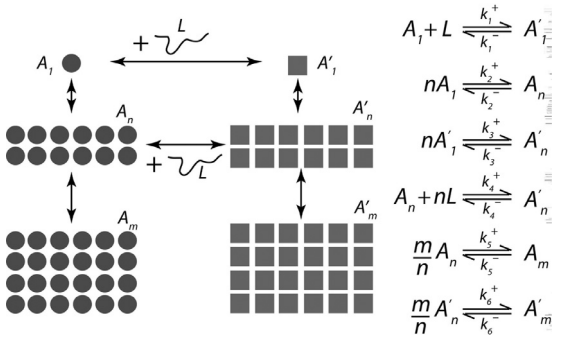

Fig. 1. A schematic of the aggregation pathways and the corresponding reactions.

oligomer conformations depending on the reactions conditions, such as the presence of membrane lipids [7,9,13,17,21]. Moreover, multiple pathways can generate different oligomer and fibrillar structures with varying toxicities. A growing number of studies seem to indicate that polymorphism and heterogeneity in the structure of the aggregates seems to contribute to phenotypes observed in AD patients [19]. Therefore, it is important to understand the dynamics of aggregation in the perspective of factors controlling multiple pathways not just in AD but also in many other neurodegenerative diseases. The phenomenon of seeding among amyloid proteins has been well-established. Addition of sub-stochiometric amounts of preformed amyloid aggregates as seeds accelerates aggregation [8,14,22]. Such seeding reactions further present the possibility of propagating distinct aggregate structures.

Recently, using a six-species model for A β aggregation along competing pathways, we determined which pathway dominates as a function of lipid concentrations using a game theoretic mathematical formulation [5]. In this paper, we extend this approach by examining in detail how the network of pathways could predominate a feasible reaction pathway, and what thermodynamic weight such pathways may endure.

We begin with laying out the modeling approach and the kinetic model system. We also introduce the underlying competitive elements of this model between the on- and off- pathways to aggregation; our modeled network competition reveals itself through the underlying network topological structure in the system, which is discussed in Section 2.3. The identification of the network structure and resulting definitions allow us to lay out the fundamental questions being investigated in a clear manner. Section 3 is devoted to the results of the aggregation kinetics computations and an analysis of the various steady state, dominant phases or winning topological structures. A thermodynamic analysis of these various states in Section 3.2 allows for a determination of stability of these states. The discussion in Section 4, provides a synthesis of the different mechanistic and energetic elements that we have identified - via a categorial analysis. A set-based categorization of the network topological paths helps to identify underlying patterns and unravel the complexity of our chosen system.

2. Mathematical modeling approach

Major obstacles to investigating amyloid systems is that they are complex and stochastic. Complex amyloid systems preclude a meaningful full-scale mathematical analysis. Therefore, analyzing every step of the model and comparing to actual data is not only difficult and time consuming but also unlikely to yield meaningful solutions. The model presented here is simplified, consisting of only six species of $A\beta$ with the objective of understanding the dynamics of the system at a rudimentary level. The six-species model can be considered as a characterization of three bins each for the on and off-pathways; the first bin represents monomeric species, the second bin (denoted by the subscript n, represents the nucleus (small oligmoers) and the final bin, denoted by subscript m, stands for larger oligomers. However, despite the apparent simplicity, the model provides a wealth of biophysical insights, as previously demonstrated [5,20]. Furthermore, the simplified model (shown in Fig. 1) also correlates well with experimental data and larger scale models. In this paper, the problem of amyloid aggregation is analyzed in several phases. In the first stage, the model system and appropriate reaction equations and rates are identified. This results in a system of differential equations based on mass-action principles.

Even with just six species (see Fig. 1) there are infinitely many rate regimes that one could investigate, most of which could be physically insignificant. Thus, an exhaustive sweep of all parameter values was not made. Instead a select few have been chosen to experiment with physically meaningful rate regimes about which we vary our key parameters. In the following model, oligomers can react with lipids L to create off-pathway oligomers, as demonstrated in our previous work [21]. The system of chemical reactions in our model consists of the non-prime species, A_1, A_n, and A_m representing on-pathway Aβ monomers and oligomers; whereas the prime species, A'_1, A'_n, and A'_m, are the off-pathway species which are created through a reaction with the fatty acid micelles, L. In the computations to follow, for each species, $n = 12$ and $m = 24$ unless otherwise specified; which denotes the order of oligomer [6].

2.1. Reaction equations

Using the Law of Mass Action Kinetics, the system of differential equations was formulated as follows [5]

$$\frac{dA_1}{dt} = nk_2^- A_n - nk_2^+ A_1^n + k_1^- A_1^{'} - k_1^+ A_1 L \tag{1}$$

$$\frac{dA_1^{'}}{dt} = nk_3^- A_n^{'} - nk_3^+ A_1^{'n} + k_1^+ A_1 L - k_1^- A_1^{'} \tag{2}$$

$$\frac{dA_n}{dt} = k_2^+ A_1^n + \frac{m}{n} k_5^- A_m + k_4^- A_n^{'} - k_2^- A_n - k_4^+ A_n L - \frac{m}{n} k_5^+ A_n^{\frac{m}{n}} \tag{3}$$

$$\frac{dA_n^{'}}{dt} = k_3^+ A_1^{'n} + k_4^+ A_n L + \frac{m}{n} k_6^- A_m^{'} - k_3^- A_n^{'} - \frac{m}{n} k_6^+ A^{'\frac{m}{n}}_n - k_4^- A^{'}{}_n \tag{4}$$

$$\frac{dA_m}{dt} = k_5^+ A_n^{\frac{m}{n}} - k_5^- A_m \tag{5}$$

$$\frac{dA_m^{'}}{dt} = k_6^+ A^{'\frac{m}{n}}_n - k_6^- A^{'}{}_m \tag{6}$$

This system is then put into non-dimensional form. Using A_0 as the characteristic concentration of monomers and $\frac{1}{k_1^-}$ the characteristic time, the dimensionless species were defined as follows

$$B_1 = \frac{A_1}{A_0}; B_n = \frac{A_n}{A_0}; B_m = \frac{A_m}{A_0}; B_1^{'} = \frac{A_1^{'}}{A_0}; B_n^{'} = \frac{A_n^{'}}{A_0}; B_m^{'} = \frac{A_m^{'}}{A_0} \tag{7}$$

The reaction constants are defined as follows:

$$\alpha_1 = \frac{k_2^-}{k_1^-}; \alpha_2 = \frac{k_2^+ A_0^{n-1}}{k_1^-}; \alpha_3 = \frac{k_1^+ L}{k_1^-}; \alpha_4 = \frac{k_4^+ L}{k_1^-}; \alpha_5 = \frac{k_5^-}{k_1^-}; \alpha_6 = \frac{k_6^+ A_0^{\frac{m}{n}-1}}{k_1^-} \tag{8}$$

$$\beta_1 = \frac{k_3^-}{k_1^-}; \beta_2 = \frac{k_3^+ A_0^{n-1}}{k_1^-}; \beta_3 = \frac{k_5^+ A_0^{\frac{m}{n}-1}}{k_1^-}; \beta_4 = \frac{k_4^-}{k_1^-}; \beta_5 = \frac{k_6^-}{k_1^-} \tag{9}$$

Note that both α_3 and α_4 have a factor L which is responsible for off-pathway aggregation. These are the bridge variables between on- and off-pathway species in the analysis which follows. The dimensionless system can be written as

$$\frac{dB_1}{ds} = n\alpha_1 B_n - n\alpha_2 B_1^n + B_1^{'} - \alpha_3 B_1 \tag{10}$$

$$\frac{dB_1^{'}}{ds} = n\beta_1 B_n^{'} - n\beta_2 B_1^{'n} + \alpha_3 B_1 - B_1^{'} \tag{11}$$

$$\frac{dB_n}{ds} = \alpha_2 B_1^n - \alpha_1 B_n + \frac{m}{n}\alpha_5 B_m + \beta_4 B_n^{'} - \alpha_4 B_n - \frac{m}{n}\beta_3 B_n^{\frac{m}{n}} \tag{12}$$

$$\frac{dB_n^{'}}{ds} = \beta_2 B_1^{'n} - \beta_1 B^{'}{}_n + \alpha_4 B_n + \frac{m}{n}\beta_5 B_m^{'} - \frac{m}{n}\alpha_6 B^{'\frac{m}{n}}_n - \beta_4 B^{'}{}_n \tag{13}$$

$$\frac{dB_m}{ds} = \beta_3 B_n^{\frac{m}{n}} - \alpha_5 B_m \tag{14}$$

$$\frac{dB_m^{'}}{ds} = \alpha_6 B^{'\frac{m}{n}}_n - \beta_5 B^{'}{}_m \tag{15}$$

The governing Eqs. (10)–(15) contain a plethora of parameters, which include 12 rate constants and the parameters m and n which are indicative of the oligomer/nucleus and fibril size, adding to the complexity of the problem. Based on previous studies [5,6,20], sample non-dimensional rate constants are defined for the study, as indicated in Table 1. The Type 1 model is a symmetric setup whose base case was previously investigated [5], while the Type 2 model is based on more biophysically relevant considerations, studied by Powers and Powers [17]. While the values of the rate constants here are somewhat arbitrary, the ratios for the forward to backward rate constants chosen are based on sound biophysical reasoning and considering the rate constant estimates from more detailed simulation models (considering many more oligomer species) in our prior work [5]. In both models the non-dimensionalization is performed such that one of the bridge parameters is set to unity, ultimately leaving us with two non-dimensional free bridge-parameters, α_3 and α_4 which are varied

Table 1
Table shows two sets of parameters used in the modeling. The first, Model 1, is based on a symmetric distribution of rate constants where all forward and backward rates in the on and off pathways are taken to be the same, respectively. The second model, Model 2, is more biophysically relevant and the distribution is based on the suggestion by Powers and Powers [17].

Model Type	α_1	α_2	α_5	β_1	β_2	β_3	β_4	β_5	α_6
1	0.001	1	0.001	0.001	1	1	1	0.001	1
2	0.001	1	0.0001	0.005	10	1	1	0.005	10

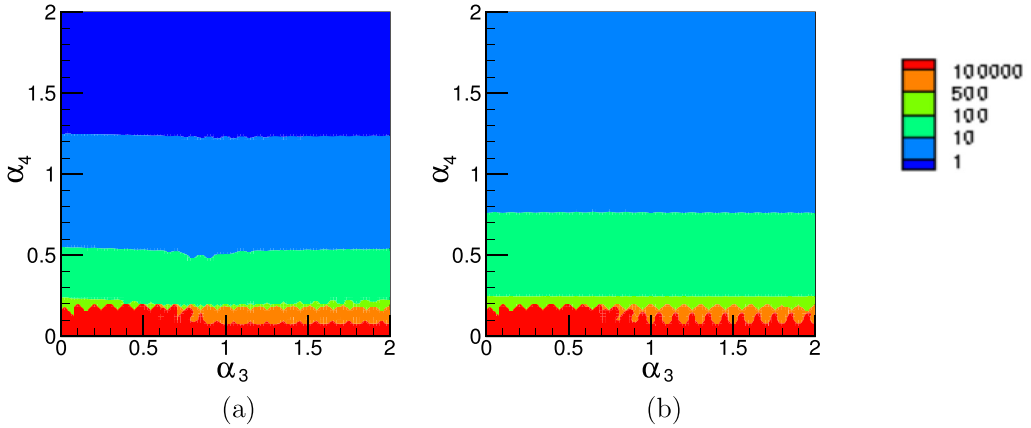

Fig. 2. This figure shows a contour plot of the ratio of B_m/B'_m for (a) the Type 2 base model with $IC = (1, 0, 0, 0, 0, 0)$ and (b) Type 2 seeded model with $IC = (1, 0, 0, 0, 1.5, 0)$.

between 0 and 2. This rate range is indicative of lipid addition to the system, where a maximum of 2 sufficiently displays the possible dynamics in the system.

The system of Eqs. (10)–(15) is also highly sensitive to perturbations of initial conditions. An initial conditions vector will often be referred to, whose values take the following order $IC = (B_1^0, B_1'^0, B_n^0, B_n'^0, B_m^0, B_m'^0)$. Here, B_1^0=1, since our characteristic monomer concentration $A_0 = A_1$ was chosen. The effect of initial conditions on our system by 'seeding' different species was studied. In other words, the seeded species' initial concentration B_i^0 will be chosen such that $B_i^0 > B_1^0 = 1$ and $B_j^0 = 0$ for all $j \neq i$ and $j \neq 1$. (eg., $IC = (1, 0, 0, 0, 1.5, 0)$ is the B_m seeded case).

2.2. Pathway competition

Our previous work [5] has revealed the importance of a game-theoretic approach. One can analyze the problem of on and off pathway persistence through an investigation of the conditions which foster the domination of these pathways. To define this problem from a game-theoretic perspective, the individual players are identified as the species $B_1, B_1', B_n, B_n', B_m, B_m'$. Two major pathways are defined: (1) On-Pathway which is the collection (B_1, B_n, B_m) and (b) Off-Pathway, denoted by (B_1', B_n', B_m'). The game is determined by the Eqs. (10)–(15) and the winner of either pathway is ultimately chosen by which of B_m or B_m' dominates.

An apt way to study the problem is to specifically look at the ratio $S_m = \frac{B_m}{B_m'}$ where $S_m > 1$ suggests domination of the on-pathway while, $S_m < 1$ is indicative of the reverse. The Fig. 2 shows two instances of variations in S_m for the Type 2 model, for changing values of α_3 and α_4, which control the phase change. Fig. 2(a), (b) correspond to two different initial conditions which have particular relevance for this study and will be discussed in greater detail in subsequent sections. In the Fig. 2, the dark blue regions correspond to $S_m < 1$ while in all other regions of the graph, $S_m > 1$. The data suggests that in case (a), with initial condition $IC = (1, 0, 0, 0, 0, 0)$, the outcome of the competition is mostly independent of the first bridge, α_3 but strongly depends on α_4. However, even with a small change in the initial conditions $IC = (1, 0, 0, 0, 1.5, 0)$, the S_m contour changes to reveal a single victor in this domain, namely B_m. Even though S_m reveals the endgame, the process of arriving at this fibril stage along any pathway is more nuanced. Fig. 2 does not reveal anything about the nature of domination of the monomers or other oligomers along the pathway. A more detailed study of the competition between all like-species in the system is needed.

2.3. Network topologies

Next, we identify embedded network structures corresponding to aggregation paths. As mentioned earlier, the biophysical problem of protein aggregation being modeled is a complex, large reaction network. Even if one were to approximate

Table 2
This table depicts the number of possible distinct aggregation pathways (or network topologies) in an $\bar{n} \times \bar{m}$ system. This paper examines a 2×3 system which contains eight distinct pathways.

# Paths → # Species per path ↓	1	2	$\bar{n}$
1	1	2	$\bar{n}$
2	1	4	$\bar{n}^2$
3	1	⑧	$\bar{n}^3$
$\bar{m}$	1	$2^{\bar{m}}$	$\bar{n}^{\bar{m}}$

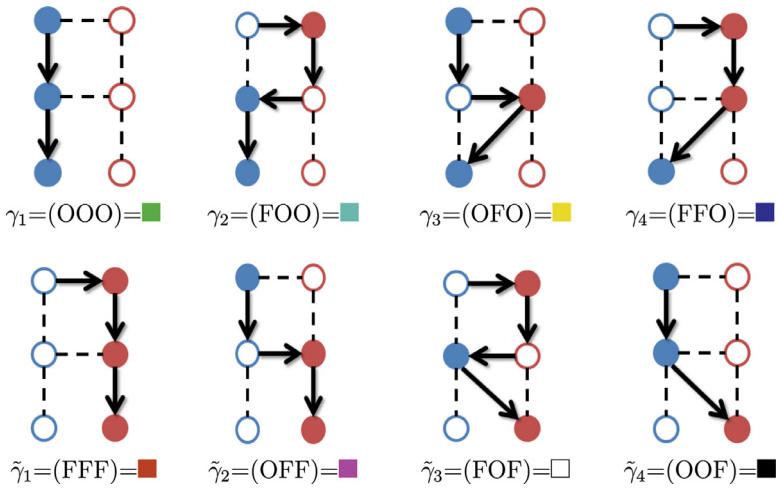

Fig. 3. This figure shows the 8 different network topologies or reaction paths that can be identified for this system. Each path is identified by three different kinds of labels which will be utilized in different context in the rest of the paper. It is important to note that the color scheme adopted, which will be used consistently throughout the paper.

the dynamics using a binary model of on- and off-pathway, the number of species and reactions remain computationally expensive. The Table 2 shows the complexity of paths which exists for any $\bar{n} \times \bar{m}$ system, where $\bar{n}$ is the number of pathways and $\bar{m}$ is the number of species per pathway.

Our previous work based upon four species [20] and six species [5] models have shown to yield fruitful results in comparison with experiments. In this paper, we continue to focus our attention on the six species model with greater emphasis on the reaction network aspect of the problem which offers a lot of room for meaningful exploration of the infinitely possible dimensionalities of our system, demonstrated by Table 2. Other choices of n and m have been explored (see Appendix B and also [5]) and the outcomes are seen to be qualitatively similar to the one shown here. Thus we choose to present a reasonably simple representative case. In the entire phase space ($0 < \alpha_3 < 2, 0 < \alpha_4 < 2$) taken up by each of the paths, with changes to the pair (n, m), small changes regarding a path's domination at any given bridge-rate combination were seen. The path's resiliency to changes in n and m reveal their ubiquity and theoretical significance. Before proceeding further, we first define the topological features of our model and the language used to describe them. These definitions allow us to discuss features of this problem in a more consistent manner.

Definition 2.1. We define the dominant path for an $\bar{n} \times \bar{m}$ system by $(\gamma_i^{(1)}, \gamma_i^{(2)}, ..., \gamma_i^{(\bar{n})})$ where $1 \le i \le \bar{n}^{\bar{m}}$. Here $\gamma_i^{(1)}$ is the collection of network paths that end in the on-pathway fibril while all other $\gamma_i^{(k)}$ $(k \neq 1)$ ends in the off-pathway. In the special case of $\bar{n} = 2$ and $\bar{m} = 3$, we have the following definitions in accordance with Fig. 3, with $\gamma_i^{(2)} = \tilde{\gamma}_i$:

$$\gamma_1 = OOO;\ \gamma_2 = FOO;\ \gamma_3 = OFO;\ \gamma_4 = FFO;$$
$$\tilde{\gamma}_1 = FFF;\ \tilde{\gamma}_2 = OFF;\ \tilde{\gamma}_3 = FOF;\ \tilde{\gamma}_4 = OOF;$$

Definition 2.2. We define an $\bar{n} \times \bar{m}$ pathway-label by $\mathcal{P}(\bar{m}, \bar{n})$ as follows

$$\mathcal{P}(\bar{m}, \bar{n}) = \sum_{i=1}^{\bar{n}^{\bar{m}}} \sum_{k=1}^{\bar{n}} p_i^{(k)} \gamma_i^{(k)} \tag{16}$$

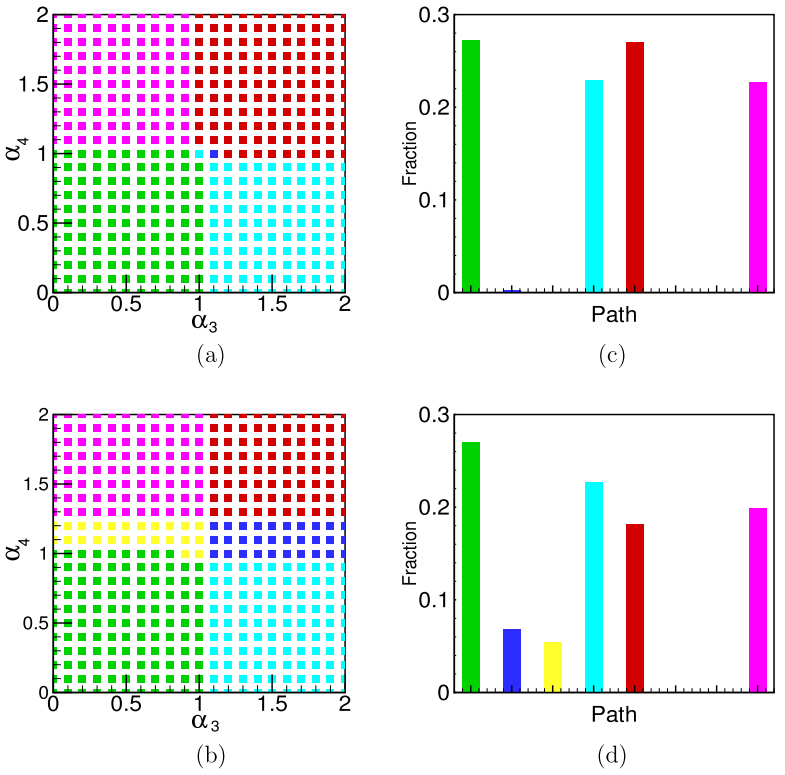

Fig. 4. This figure shows the phase diagram for the base case for the two models (see Table 1). Panel (a) and (b) depict the emergence of different dominant topologies as a function of α_3 and α_4, for type 1 and 2 models respectively. Panels (c) and (d) show the distribution of the various network paths in the phase space α_3-α_4.

$$1 = \sum_{i=1}^{\bar{n}^{\bar{m}}} \left(p_i^{(1)} \gamma_i^{(1)} + p_i^{(2)} \gamma_i^{(2)} + \cdots + p_i^{(\bar{n})} \gamma_i^{(\bar{n})} \right) \tag{17}$$

$$1 = p_1^{(1)} \gamma_1^{(1)} + p_1^{(2)} \gamma_1^{(2)} + \cdots + p_1^{(\bar{n})} \gamma_1^{(\bar{n})} + p_2^{(1)} \gamma_2^{(1)} + p_2^{(2)} \gamma_2^{(2)} + \cdots + p_2^{(\bar{n})} \gamma_2^{(\bar{n})} + p_{2^{\bar{m}}}^{(1)} \gamma_{2^{\bar{m}}}^{(1)} + p_{2^{\bar{m}}}^{(2)} \gamma_{2^{\bar{m}}}^{(2)} + \cdots + p_{2^{\bar{m}}}^{(\bar{n})} \gamma_{2^{\bar{m}}}^{(\bar{n})} \tag{18}$$

where $p_i^{(k)}$, $1 \le k \le \bar{n}$ represent the probability of the dominant path and are computed using the solutions of the Eqs. (10)–(15).

Definition 2.3. We define the set of dominant networks as follows:

$$\mathcal{S}_{IC} = \left\{ \gamma_i^{(k)} \mid 1 \le k \le \bar{n}, 1 \le i \le \bar{n}^{\bar{m}} \right\} \tag{19}$$

with the default baseline case $\mathcal{S}_{(1,0,0,0,0,0)} \equiv \mathcal{S}_0$.

It is worth noting that the rate constants α_k, β_k $(1 \le k \le 6)$ determine the nature of the sets $\mathcal{S}_{(IC)}$ and the probabilities depend strongly upon the initial conditions, i.e. $(B_1^0, B_1'^0, B_n^0, B_n'^0, B_m^0, B_m'^0)$. With these definitions in hand, the central questions that we address in this paper can be restated in the topological language we introduce here. More specifically,

Q1: How do we identify the elements of $\mathcal{S}$?
Q2: Can we define the mapping Γ such that $\Gamma : \mathcal{S}_i \to \mathcal{S}_j$, where $i \neq j$?
Q3: Under what conditions do the various network topologies dominate?
Q4: What are the conditions under which specific paths remain stable?

In our recent paper [5], the idea of analyzing this problem from a game-theoretic point of view was introduced. This entails examining the concentration of the different species under conditions of dynamic equilibrium along the two different pathways and accordingly identifying the *dominant paths* in the system. The players in the game are the paths, $\gamma_i^{(k)}$, whose concentrations are controlled by the various parameters in the problem. The outcome of such a 'game' is best visualized using a phase diagram (see Fig. 4) which shows the dominant topologies as a function of the bridge parameters α_3 and α_4. Each pathway is identified using a different color, shown in Fig. 3 and also identified using a three-letter code comprising of 'O' and 'F' which respectively denote the domination of an on-pathway or off-pathway species. Our computations were

Table 3
Probabilities of the dominant pathways as a function of seeding.

Model	Seeding	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
Type 1	(1,0,0,0,0,0)	0.272	0.229	0.000	0.002	0.270	0.227	0.000	0.000
	(1,0,1.5,0,0,0)	0.272	0.274	0.000	0.000	0.249	0.227	0.000	0.000
	(1,0,0,1.5,0,0)	0.249	0.227	0.000	0.000	0.272	0.274	0.000	0.000
	(1,0,0,0,1.5,0)	0.272	0.274	0.023	0.025	0.224	0.204	0.000	0.000
	(1,0,0,0,0,1.5)	0.247	0.204	0.000	0.000	0.274	0.249	0.023	0.002
Type 2	(1,0,0,0,0,0)	0.270	0.227	0.054	0.068	0.181	0.199	0.000	0.000
	(1,0,1.5,0,0,0)	0.240	0.240	0.245	0.227	0.023	0.025	0.000	0.000
	(1,0,0,1.5,0,0)	0.247	0.229	0.068	0.045	0.204	0.206	0.000	0.000
	(1,0,0,0,1.5,0)	0.274	0.249	0.247	0.229	0.000	0.000	0.000	0.000
	(1,0,0,0,0,1.5)	0.249	0.227	0.050	0.045	0.206	0.222	0.000	0.000

conducted by solving the Eqs. (10)–(15) for α_3 and α_4 between 0 and 2 in increments of 0.1. The phase diagram is uniquely dependent upon the initial conditions chosen, which we also refer to as 'seeding'. Variations in the phase diagrams provide valuable information about bifurcations and emergence of new structures.

3. Results

We examined the evolution of network topologies and their emergence and disappearance under various conditions, which are best visualized by means of a 'phase diagram'. We also examined the possible control mechanisms that move the phase diagrams in desirable ways. First, the 'base-case' for the Type 1 and Type 2 models described in Table 1 was examined, with the initial conditions such that $B_1^0 = 1$ with all others set to zero. The approach adopted in this section specifically addresses questions **Q1**–**Q3**, at least in part.

3.1. Analysis of topological phase diagrams

Understanding network topologies is significant in the context of strain generation among amyloid aggregates. Not only do aggregates generated via different networks and pathways differ in their temporal resolutions, they may also do so in their thermodynamic stabilities. Therefore, understanding network topologies in conjunction with delineating energetic-mechanistic landscape is critical to identify thermodynamically stable off-pathway aggregates. Such a contour will enable predictions of strain generation as a function of temporal parameters.

As in our previous work [5], it is instructive to examine the phase diagram which depicts the dominant network paths in the system as a function of the bridge parameters. Fig. 4 shows the dominant phases in the $0 \leq \alpha_3, \alpha_4 \leq 2$ space, for the base case of the Type 1 and 2 models along with the distributions of the respective components.

The Fig. 4(c) and (d) show stark contrast; while the Type 1 base model only permits essentially four network paths, i.e. $\mathcal{S}_0^{Type1} = \{\gamma_1, \gamma_2, \tilde{\gamma}_1, \tilde{\gamma}_2\}$, for the Type 2 model, $\mathcal{S}_0^{Type2} = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma}_1, \tilde{\gamma}_2\}$. The emergence of the new paths in a substantial manner is primarily due to the choice of rate constants. The path labels allow us to represent the system more compactly with the respective probabilities,

$$\mathcal{P}_0^{Type1} = 0.27\gamma_1 + 0.23\gamma_2 + 0.27\tilde{\gamma}_1 + 0.23\tilde{\gamma}_2$$
$$\mathcal{P}_0^{Type2} = 0.27\gamma_1 + 0.23\gamma_2 + 0.05\gamma_3 + 0.07\gamma_4 + 0.18\tilde{\gamma}_1 + 0.2\tilde{\gamma}_2$$

Next are seeding studies, which were computed with various initial conditions. In each of the studies, B_1^0 is at unity, however one of the other species deviates from the base case and contains a non-zero initial concentration ranging from 0 to 1.5 times B_1^0. The Fig. 5 shows the phase diagrams corresponding to various seedings: (a) $B_n^0 = 1.5$, (b) $B_n'^0 = 1.5$, (c) $B_m^0 = 1.5$ and (d) $B_m'^0 = 1.5$. The related Fig. 6 shows the respective fractions of the dominant pathways in each of the phase diagrams. Of particular importance is the change in the percent of the $\tilde{\gamma}_i$ pathways and the emergence of new pathways which are not observed in the base cases in either of the Type 1 or 2 models. The Table 3 provides the path probabilities for the Type 1 and 2 models with various seeding. The table clearly shows the impact of the rate constants upon the probabilities but also how seeding controls the emergence of the network paths. Of particular importance is the cell highlighted in green, which correspond to seeding via B_n and B_m, i.e. along the on-pathway. Clearly these seedings are significant and result in the depletion or elimination of all network paths which terminate in the off-pathway.

3.2. Thermodynamics of aggregation pathways

It seems that a correlation exists between thermodynamic stabilities of aggregates and their cellular effects and toxicities. It is well known that aggregates such as fibrils, profibrils and low-molecular weight oligomers differ in their stabilities and show differential effect on cells. Therefore, it is prudent to understand aggregate stabilities formed along the networks. In our previous work [5], we discussed the outcome of a linearized stability analysis for the same six-species system. Here, an

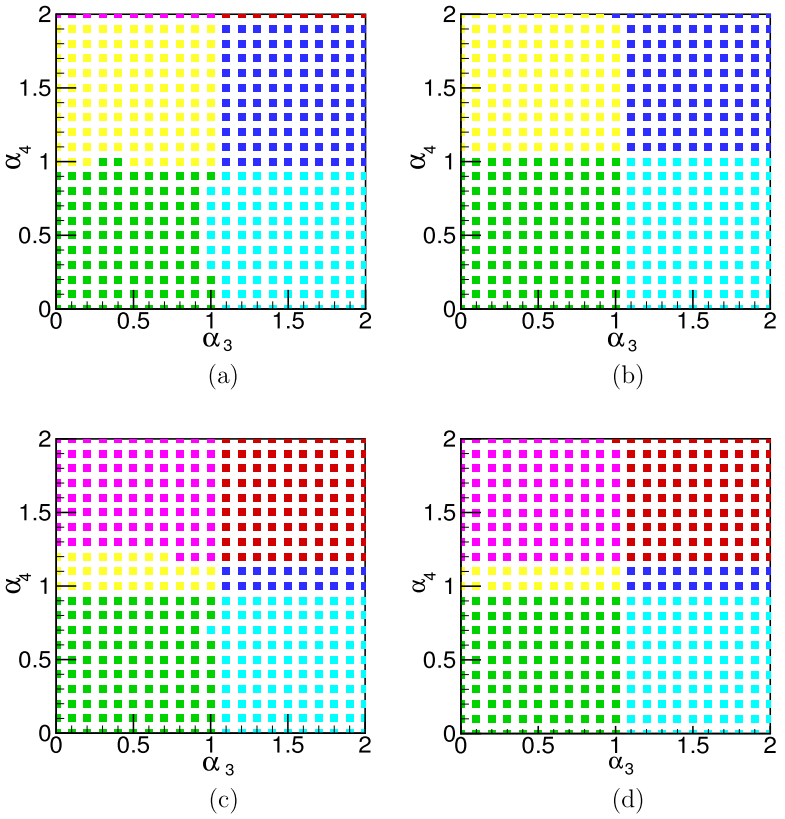

Fig. 5. This figure shows the phase diagram for Type 2 model under different seedings, beyond the base case : (a) $B_n^0 = 1.5$; (b) $B_m^0 = 1.5$; (c) $B_n'^0 = 1.5$; (d) $B_m'^0 = 1.5$.

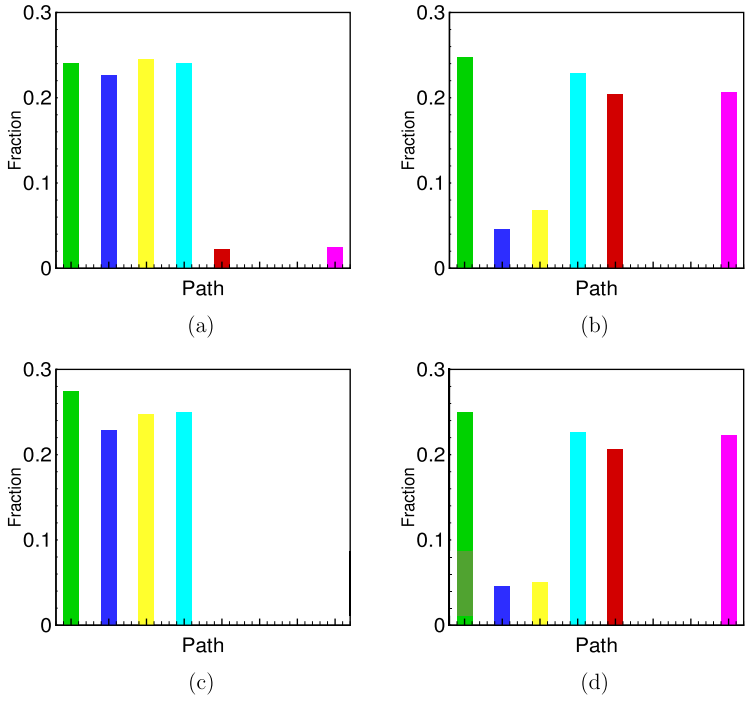

Fig. 6. Histogram of path probabilities as a function of various seedings for the Type 2 model: (a) $B_1^0 = 1$, $B_n^0 = 1.5$; (b) $B_1^0 = 1$, $B_n'^0 = 1.5$; (c) $B_1^0 = 1$, $B_m^0 = 1.5$; (d) $B_1^0 = 1$, $B_m'^0 = 1.5$.

Table 4
Construction of the different topologies as a union of sub-paths.

Path	γ_{1n}	γ_{nm}	$\gamma_{1'n'}$	$\gamma_{n'm'}$	$\gamma_{11'}$	$\gamma_{nn'}$	$\gamma_{nm'}$	$\gamma_{n'm}$
γ_1	+	+						
γ_2		+	+		+	−		
γ_3	+					+		+
γ_4			+		+			+
$\tilde{\gamma}_1$			+	+	+			
$\tilde{\gamma}_2$	+			+		+		
$\tilde{\gamma}_3$			+		+	−	+	
$\tilde{\gamma}_4$	+						+	

alternative approach to understanding the stability of this complex system was taken by investigating the thermodynamics of the system under variations of the bridge parameters, through the analysis of quantities such as fluxes, forces and entropy production (see [10]). For ease of computation, the network paths are further sub-divided into sub-paths as indicated in Table 4.

Each pathway in Fig. 3 is the union of single paths which can be constructed in accordance with Table 4. For instance, $\gamma_1 := \gamma_{1n} \cup \gamma_{nm}$ and $\gamma_2 := \gamma_{11'} \cup \gamma_{1'n'} \cup \gamma_{n'n} \cup \gamma_{nm}$, where $\gamma_{n'n}$ is the reverse path of $\gamma_{nn'}$. The *path-matrix* can be defined as follows:

$$\underline{\underline{C}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & -i \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & -i \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & i & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & i & 0 \end{pmatrix}$$

The matrix can be further decomposed into $\underline{\underline{C}}_{ab} = \underline{\underline{C}}^{(re)}_{ab} + i\underline{\underline{C}}^{(im)}_{ab}$, where

$$\underline{\underline{C}}^{(im)}_{ab} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}; \underline{\underline{C}}^{(re)}_{ab} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where $\underline{\underline{C}}^{(im)}_{ab}$ is indicative of *emergent pathways*, namely $\gamma_{nm'}$ and $\gamma_{n'm}$ which are not physical reaction paths in the original system (Fig. 1). The emergence of these 'diagonal' paths can be accounted for using certain isomorphisms, which are identified in Fig. 7 which essentially state the following identities, in terms of sub-paths:

$$\gamma_3 \equiv \gamma_{1i} \cup \gamma_{nn'} \cup \gamma_{nm}$$
$$\gamma_4 \equiv \gamma_{11'} \cup \gamma_{1'n'} \cup \gamma_{n'm}$$
$$\tilde{\gamma}_3 \equiv \gamma_{11'} \cup \gamma_{1'n'} \cup \gamma_{n'n} \cup \gamma_{n'm'}$$
$$\tilde{\gamma}_4 \equiv \gamma_{1n} \cup \gamma_{nm'}.$$

This equivalency permits computation of several thermodynamic quantities since rate constants for these isomorphic paths are well defined. These isomorphic paths are chosen such that: (a) there is no re-tracing of steps and (b) the terminal diagonal path is replaced by an appropriate direct one.

As a result, the path-matrix $\underline{\underline{C}}$ can be redefined as

$$\underline{\underline{C}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{pmatrix}$$

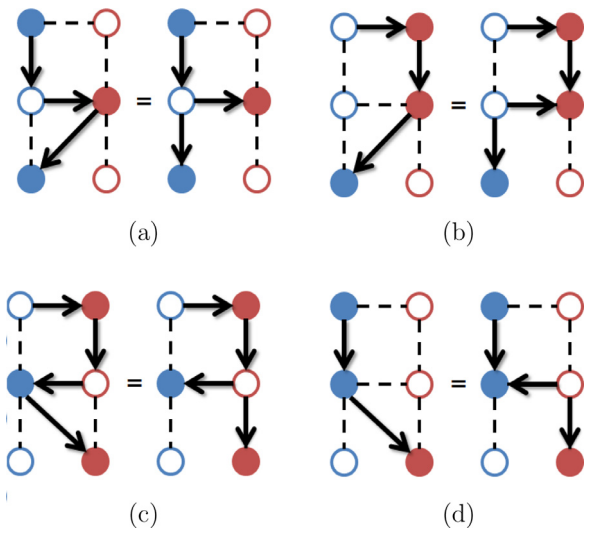

Fig. 7. This figure shows isomorphic versions of the emergent network paths γ_3, γ_4, $\tilde{\gamma}_3$ and $\tilde{\gamma}_4$, respectively.

Similarly, the fluxes corresponding to these pathways can be constructed following the equation

$$\underline{J} = \underline{\underline{C}} \cdot \underline{J}_{sub} \tag{20}$$

or in expanded form

$$\begin{pmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \\ \tilde{J}_1 \\ \tilde{J}_2 \\ \tilde{J}_3 \\ \tilde{J}_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} J_{1n} \\ J_{nm} \\ J_{1'n'} \\ J_{n'm'} \\ J_{11'} \\ J_{nn'} \\ J_{nm'} \\ J_{n'm} \end{pmatrix}$$

Here the fluxes along the individual paths are given by $J_{sub} = R^b - R^a$ where R^b refers to the forward reaction flux and R^a is the backward reaction flux. As a result, the individual fluxes can be written in terms of sub-path fluxes as

$$\begin{pmatrix} J_{1n} \\ J_{nm} \\ J_{1'n'} \\ J_{n'm'} \\ J_{11'} \\ J_{nn'} \\ J_{nm'} \\ J_{n'm} \end{pmatrix} = \begin{pmatrix} \alpha_2 B_1^n - \alpha_1 B_n \\ \beta_3 B_n^{m/n} - \alpha_5 B_m \\ \beta_2 B_1'^n - \beta_1 B_n' \\ \alpha_6 B_n'^{m/n} - \beta_5 B_m' \\ \alpha_3 B_1 - B_1' \\ \alpha_4 B_n - \beta_4 B_n' \\ 0 \\ 0 \end{pmatrix}$$

Similarly, the forces corresponding to the different pathways can be expressed in terms of the reaction affinities or forces which is given by the expression

$$\underline{F} = \underline{\underline{C}} \cdot \underline{F}_{sub} \tag{21}$$

where $F_{sub} = RT \ln\left(\frac{R^b}{R^a}\right)$ and the force along each sub-path is given by

$$\begin{pmatrix} F_{1n} \\ F_{nm} \\ F_{1'n'} \\ F_{n'm'} \\ F_{11'} \\ F_{nn'} \\ F_{nm'} \\ F_{n'm} \end{pmatrix} = \begin{pmatrix} \frac{\alpha_2 B_1^n}{\alpha_1 B_n} \\ \frac{\beta_3 B_n^{m/n}}{\alpha_5 B_m} \\ \frac{\beta_2 B_1'^n}{\beta_1 B_n'} \\ \frac{\alpha_6 B_n'^{m/n}}{\beta_5 B_m'} \\ \frac{\alpha_3 B_1}{B_1'} \\ \frac{\alpha_4 B_n}{\beta_4 B_n'} \\ 0 \\ 0 \end{pmatrix}$$

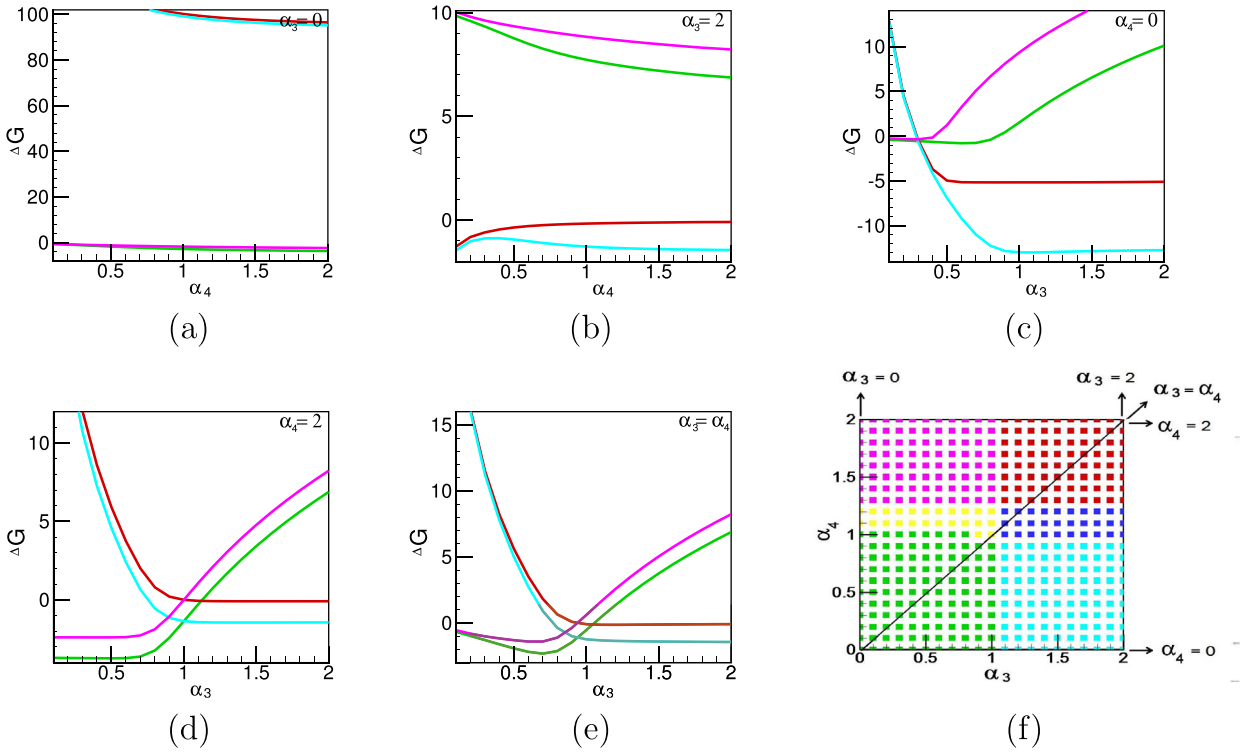

Fig. 8. This figure shows the values of ΔG for changing α_3 and α_4, corresponding to the baseline case (Type 2). The panels (a)-(e) indicate the values of ΔG for different cuts of the phase plane as indicated in panel (f), namely (a) $\alpha_3 = 0$; (b) $\alpha_3 = 2$, (b) $\alpha_0 = 0$. (d) (b) $\alpha_4 = 2$ and (e) $\alpha_3 = \alpha_4$.

These reaction affinities are related to the Gibbs free energy of the system by

$$-\Delta G_i = F_i = RT \ln K_{eq} \tag{22}$$

where K_{eq} is the equilibrium constant for the reaction. If the change in ΔG is negative (i.e. change in entropy is positive), then the chemical reaction occurs spontaneously at constant pressure and temperature. Given that the essence of this study is to understand the biophysics of competition between the on and off pathways, the ΔG profiles for the different pathways, under different conditions in the phase space were examined. In particular, ΔG for five slices or cuts of the α_3-α_4 plane: (a) $\alpha_3 = 0$; (b) $\alpha_3 = 2$, (b) $\alpha_0 = 0$. (d) (b) $\alpha_4 = 2$ and (e) $\alpha_3 = \alpha_4$ were studied as a means for examining the thermodynamic stability fo the system and in particular the stability and spontaneity of the different network paths. The ΔG values (colored according to the definition in Fig. 3) indicate the specific conditions under which specific pathways become spontaneous. Equally important to note is that, for specific values of α_3 and α_4, ΔG become positive for some pathways indicating that they are no longer stable (i.e. they terminate). Often, the coexistence of multiple spontaneous paths was observed. From a game-theoretic perspective this signifies a kind of cooperative game where multiple agents (in this case pathways) co-exist.

The results of two different sets of computations are shown in Figs. 8 and 9 and correspond respectively to the baseline Type 2 model and the case of seeding with $B_m^0 = 1.5$. Both figures show ΔG to be sensitive to changes in α_3 and α_4 and the dominant network paths in the corresponding phase diagram, which is indicated in panel (f) of both figures. Since the different networks share one or more sub-paths, it is understandable that more than one network path prevails for any given condition. It should also be made clear that while computation is a good tool to gauge which path is spontaneous, one cannot necessarily determine the strength of the reaction based on the magnitude of the respective ΔGs.

As Table 3 indicates, the network paths γ_3, γ_4, $\tilde{\gamma}_3$ and $\tilde{\gamma}_4$ emerge in a dominant manner only in selective cases, depending on the model type. In cases where they are present, the ΔG values corresponding to these paths overlap with their *surrogate paths*, i.e. one that follows closely with one or more regularly dominant paths. In some special cases, such as the Type 2 baseline model, all eight paths are distinguishable (see Fig. 10). Panel (a) of this figure shows the free energy changes for all paths with changing α_3 and $\alpha_4 = 0$. Panel (b) of Fig. 10 depicts a similar graph for Type 2 model with seeding $B_m^0 = 1.5$.

4. Discussion

Results obtained here provide some unique and distinct insights on the network of pathways and thermodynamic stabilities of the aggregates formed.

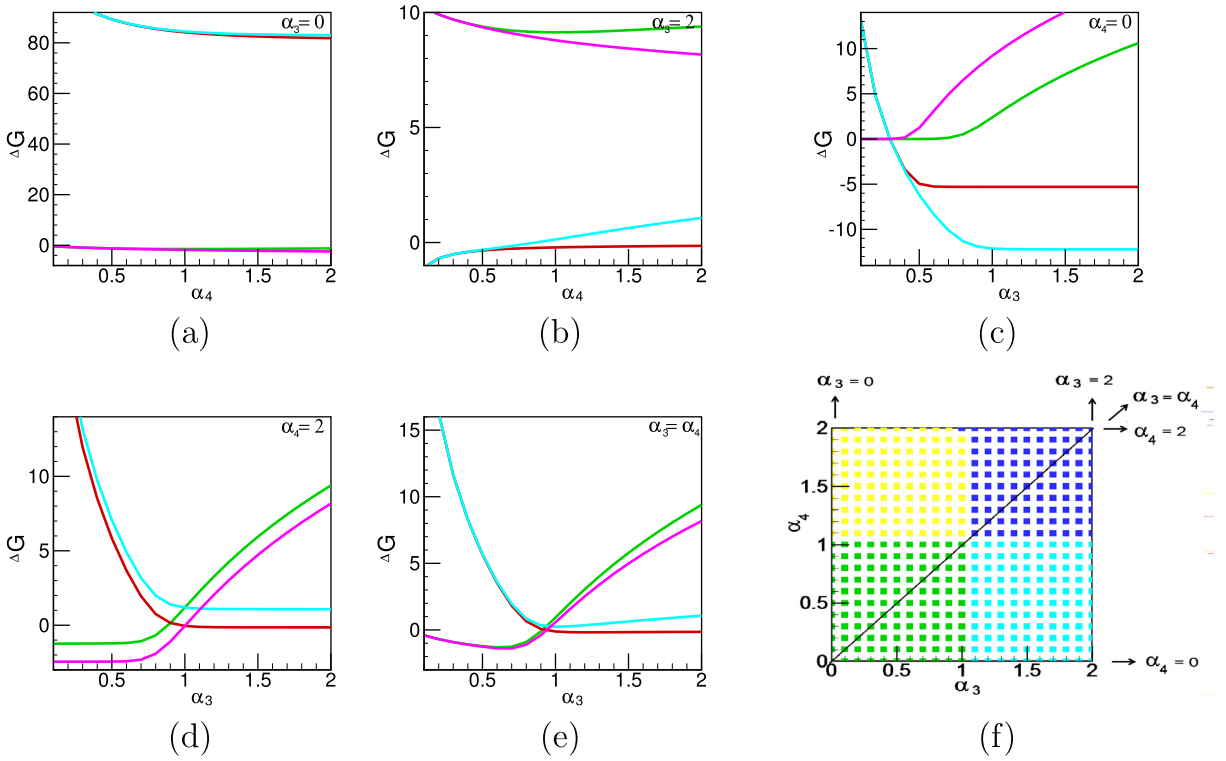

Fig. 9. This figure shows the values of ΔG for changing α_3 and α_4, corresponding to the case of Type 2 with seeding $B_m^0 = 1.5$. The panels (a)-(e) indicate the values of ΔG for different cuts of the phase plane as indicated in panel (f), namely (a) $\alpha_3 = 0$; (b) $\alpha_3 = 2$, (b) $\alpha_0 = 0$. (d) (b) $\alpha_4 = 2$ and (e) $\alpha_3 = \alpha_4$.

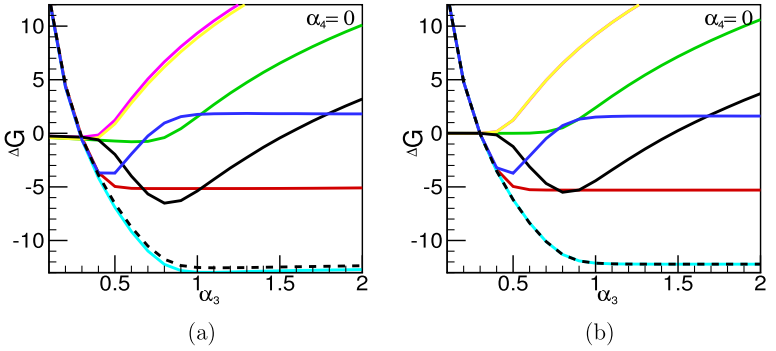

Fig. 10. This figure shows the values of ΔG for changing α_3 and $\alpha_4 = 0$, corresponding to (a) the base case Type 2 model and (b) the seeded case for Type 2 with $B_m^0 = 1.5$. As in earlier graphs Figs. 8 and 9, the colors indicate the respective network path. The solid black curve represents the path $\tilde{\gamma}_4$ and the dashed black curve represents $\tilde{\gamma}_3$.

4.1. Categories and maps

It is important to find a map Γ that moves us from the subset of dominant paths in the base case, to the subset of dominant paths in a seeded case. For example, let Γ_μ be the mapping from the base case to the B_m seeded case defined by $\Gamma_\mu : \mathcal{S}_0 \to \mathcal{S}_M$ where:

$$\mathcal{S}_0 = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}\},\ \mathcal{S}_M = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$$

The morphing caused by each seeding offers unique challenges when defining a mapping. In the B_m seeded case, two-pathways cease to exist as dominant, and the other four grow in domination. It follows that regions of 'bridge-rate space' which were once dominated by the vanished pathways, will make way for the pathways growing in dominance. Thus our mapping must capture $\tilde{\gamma_1}$ and $\tilde{\gamma_2}$ going to $\gamma_1, \gamma_2, \gamma_3$, and γ_4. The competition tradeoffs occurring can be quantififed by examining the base and B_m seeded case probability distributions.

$$\mathcal{P}_m - \mathcal{P}_0 = 0.004\gamma_1 + 0.022\gamma_2 + 0.193\gamma_3 + 0.161\gamma_4 + (-0.181)\tilde{\gamma_1} + (-0.199)\tilde{\gamma_2}$$

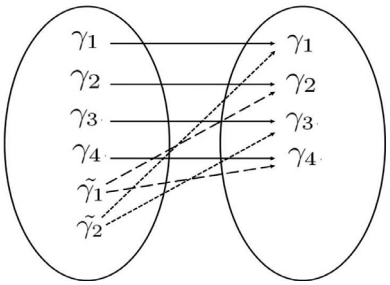

Fig. 11. A schematic of the transformation from set $\mathcal{S}_0 \to \mathcal{S}_m$ showing the infeasibility of the map.

This difference suggests that the map $\Gamma_\mu : \mathcal{S}_0 \to \mathcal{S}_m$ must, beyond the identity map, contain arrows going into paths with positive coefficients which came out of paths with negative coefficients. In other words, path morphisms are determined by pairing donor paths with recipient paths, those with negative and positive coefficients in the above difference, respectively. Thus, in order to define a valid map, the domain must be expanded to contain "hybrid" species which capture the competitive tradeoff happening in pathway domination. In this case, paths like γ_1 are denoted as recipient species and $\tilde{\gamma_2}$ as donor paths. Donor and recipient paths were then paired by coupling species that are synchronous-active according to ΔG values in various rate combinations.

Since, a zero-sum game was played, with the probabilities (see Table 2) determining the dominant persistent network paths, the exchanges during a transformation need to be identified. For instance, in the case of the transformation $\mathcal{S}_0 \to \mathcal{S}_m$, the exchanges illustrated in Fig. 11 can be identified

While such a map is not mathematically possible, a formal map can be identified with the help of the phase-diagrams and the ΔG graphs, Figs. (4), (5 b) and (9). This requires a rearrangement of the co-domain $\mathcal{S}_0$. Before we formulate appropriate mathematical transformations between the sets of network topologies, we will need to define a 'hybrid network path'.

Definition 4.1. A *hybrid network path* is defined as an element union of the union of two or more paths to which a distinct network path contributes. We label such a hybrid path γ_a^b where a represents the donor path (γ_a) and b, the recipient path (γ_b).

A mathematically correct domain set can be be defined with the help of the hybrid paths. We therefore write the modified co-domain, $\mathcal{S}_{0M}$ as

$$\mathcal{S}_{0M} = \left\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_{\tilde{1}}^2, \gamma_{\tilde{1}}^4, \gamma_{\tilde{2}}^1, \gamma_{\tilde{2}}^3\right\}$$

with corresponding path label written as

$$\mathcal{P}_{0,m} = 0.270\gamma_1 + 0.227\gamma_2 + 0.054\gamma_3 + 0.068\gamma_4 + 0.21\gamma_{\tilde{1}}^2 + 0.160\gamma_{\tilde{1}}^4 + 0.003\gamma_{\tilde{2}}^1 + 0.198\gamma_{\tilde{2}}^3$$

where the net probability can be seen to be conserved. The mapping, Γ_m: $\mathcal{S}_{0M} \to \mathcal{S}_m$ is therefore well defined. Similarly the mapping for the remaining seedings Γ_n ($\mathcal{S}_{0N} \to \mathcal{S}_n$), $\Gamma_{n'}$ ($\mathcal{S}_{0N'} \to \mathcal{S}_{n'}$), $\Gamma_{m'}$ ($\mathcal{S}_{0M'} \to \mathcal{S}_{m'}$) discussed earlier is given by

$$\Gamma_n : \left\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}, \gamma_1^2, \gamma_1^3, \gamma_{\tilde{2}}^1, \gamma_{\tilde{2}}^3\right\} \to \{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}\}$$

$$\Gamma_{n'} : \left\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}, \gamma_1^2, \gamma_1^3, \gamma_1^{\tilde{4}}, \gamma_4^{\tilde{1}}\right\} \to \{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}\}$$

$$\Gamma_{m'} : \left\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}, \gamma_1^2, \gamma_1^4, \gamma_{\tilde{2}}^1, \gamma_{\tilde{2}}^3\right\} \to \{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}\}$$

which defines the mapping between the set, with respective path labels given by

$$\mathcal{P}_{0,n} = 0.240\gamma_1 + 0.227\gamma_2 + 0.054\gamma_3 + 0.068\gamma_4 + 0.023\tilde{\gamma_1} + 0.025\tilde{\gamma_2} + 0.013\gamma_1^2 + 0.017\gamma_1^3 + 0.158\gamma_{\tilde{1}}^4 + 0.174\gamma_{\tilde{2}}^3$$

$$\mathcal{P}_{0,n'} = 0.247\gamma_1 + 0.227\gamma_2 + 0.054\gamma_3 + 0.045\gamma_4 + 0.181\tilde{\gamma_1} + 0.199\tilde{\gamma_2} + 0.002\gamma_1^2 + 0.014\gamma_1^3 + 0.007\gamma_1^{\tilde{2}} + 0.023\gamma_4^{\tilde{1}}$$

$$\mathcal{P}_{0,m'} = 0.249\gamma_1 + 0.227\gamma_2 + 0.054\gamma_3 + 0.068\gamma_4 + 0.181\tilde{\gamma_1} + 0.199\tilde{\gamma_2} + 0.021\gamma_1^{\tilde{2}} + 0.002\gamma_3^{\tilde{1}} + 0.002\gamma_3^{\tilde{2}} + 0.023\gamma_4^{\tilde{1}}$$

Finally, we define the domain set $\mathcal{S}_0 = \mathcal{S}_{0N} \cup \mathcal{S}_{0M} \cup \mathcal{S}_{0N'} \cup \mathcal{S}_{0M'}$; therefore:

$$\mathcal{S}_0 = \left\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \tilde{\gamma_1}, \tilde{\gamma_2}, \gamma_{\tilde{1}}^2, \gamma_{\tilde{1}}^4, \gamma_{\tilde{2}}^1, \gamma_{\tilde{2}}^3, \gamma_1^2, \gamma_1^3, \gamma_3^{\tilde{1}}, \gamma_3^{\tilde{2}}, \gamma_1^{\tilde{2}}, \gamma_4^{\tilde{1}}\right\}$$

allowing for distinct seeding maps, as depicted in Fig. 12.

Such a categorization addresses the questions **Q1** and **Q2** posed in Section 2.3. Note that the definition of $\mathcal{S}_0$ is not unique and other equivalent or perhaps more convenient definitions are possible, which will be investigated in our future work. This suggests that a recognition of a suitable categorization of the network topologies and their maps is possible (see for instance [1,2]). Of particular significance here are the hybrid network paths which reveal the underlying internal complexity or sub-structure in the system.

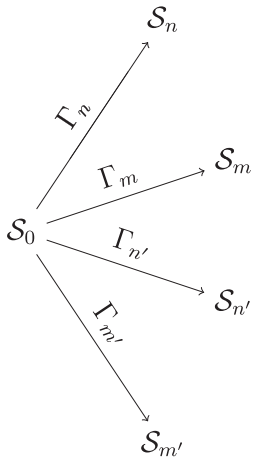

Fig. 12. Unique seeding maps based on the identification of the right domain set $\mathcal{S}_0$.

4.2. Thermodynamics and stability

Next, in the context of the emergence of pathological phenotypes, conformationally distinct and thermodynamically stable aggregates are significant. The work presented here shows theoretical possibilities along the thermodynamic landscape that enable identification of stable 'strain' generation that may lead to phenotypic outcomes.

From the thermodynamic standpoint alone, much can be said about this problem. A significant point to be noted is that, while the system under investigation is in mechanical steady-state, it remains out of thermodynamic equilibrium. There is sufficient evidence that chemical dissipative systems extremize their entropy production [11,18,24]. Therefore, principles of non-equilibrium thermodynamics are most appropriate for the investigation of such systems [15]. Furthermore, the approach adopted in this paper follows the *Network Thermodynamic theory* [12,15,16], which allows for the investigation of irreversible biological, physical and chemical processes through a melding of dynamical systems, network theory and non-equilibrium thermodynamics. Such an approach permits the identification of underlying topologies in the system which helps us understand relations and underlying complexity of the larger system. Non-equilibrium systems rely on the rate of entropy production to characterize its behavior. In this problem, given the Eqs. (20) and (21), the local rate of entropy production, denoted σ, corresponding to the various network pathways can be computed using the equation

$$\sigma = \frac{1}{T}\sum_k F_k J_k \tag{23}$$

where T is the ambient temperature and $k = 1 - 8$, corresponding to the different paths. The entropy production rate σ is the negative of the free energy rate and has been identified as a potential measure of stability and a pattern selection principle in complex systems [4,11] which are out of thermodynamic equilibrium. A particularly useful and somewhat different interpretation of the entropy production principle proposed in [4] appears to be relevant here as well and needs to be further investigated; in the context of a complex fluid-solid mixture, stable patterns observed appear to correspond to the state of maximum entropy production. The maximum here is not in time, but considered over a set of all allowable states that the system can assume; it is a pattern selection principle. A linearized stability argument for the model system, Eqs. (10)–(15) was reported in our previous paper [5]. While such an analysis is useful in determining the nature of the steady states, it speaks less about the physics of the biophysical system. For this reason, we take a free energy approach to determining the stability of the various reactions here. This network thermodynamic approach additionally, allows a categorization of the reaction pathways (Section 4.1) through a holistic examination of the thermo-mechanics of the system. The questions **Q3** and **Q4** are successfully addressed through the hybrid approach. The Figs. 8–10 and Eq. (23) defines the appropriate dominant, stable reaction networks. While we do not pursue the entropy based argument in the current paper, the match between ΔG and the dominant paths in the current study bodes well for entropy production based argument as well which will be pursued further in our future work.

5. Conclusions

In this paper we have modeled the problem of amyloid aggregation with the added complexity of a phase transition induced alternative, off-pathway aggregation. The problem of amyloid aggregation along multiple pathways is well suited to kinetic, mass-action based models, which has been adequately revealed in our previous studies [5] where control mechanisms for transitioning from on- to off-pathway and vice versa were identified. The current work also contains an added layer of thermodynamic analysis, rooted in the Gibbs free energy of the system. Together, they tell us about the complex steady states in the system and their stability.

Specifically, the network thermodynamic approach founded upon the underlying topological structures allows us to identify the dominant reaction pathways for variations of the bridge parameters. The results obtained here, especially under conditions of seeding, brings to light new, emergent network pathways (see Fig. 10) which are strongly dependent on the reaction rates. The biophysical significance of these diverse pathways and their categorization showcases intervention strategies, through appropriate seeding, to control the aggregation process and divert the system towards the less pathological conditions (on-pathway). The particular novelty of this paper lies in bringing a diverse set of theoretical tools to analyze this problem. Our previous theoretical treatments have revealed to be physically insightful and we believe that the outcome of this analysis point us towards futher novel and creative experiments which will be taken up as a natural consequence of this paper.

Acknowledgments

The authors want to thank the National Science Foundation for their financial support; NSF CBET 1802793 (to VR), NSF CBET 1802588 (to PG), and NSF CBET 1802641 (to AV). The authors also thank the National Center for Research Resources [5P20RR01647-11] and the National Institute of General Medical Sciences [8 P20 GM103476-11] from the National Institutes of Health for funding through INBRE (to V.R.). The authors have no competing interests in this work. Authorship is presented in alphabetical order.

Appendix A.

The Fig. 13 shows the phase-diagram for Type 1 model under various conditions of seeding. This should be compared with the Fig. 4 to note the changes between the two models discussed in the paper. Here again, one sees the emergence of new network paths due to seeding.

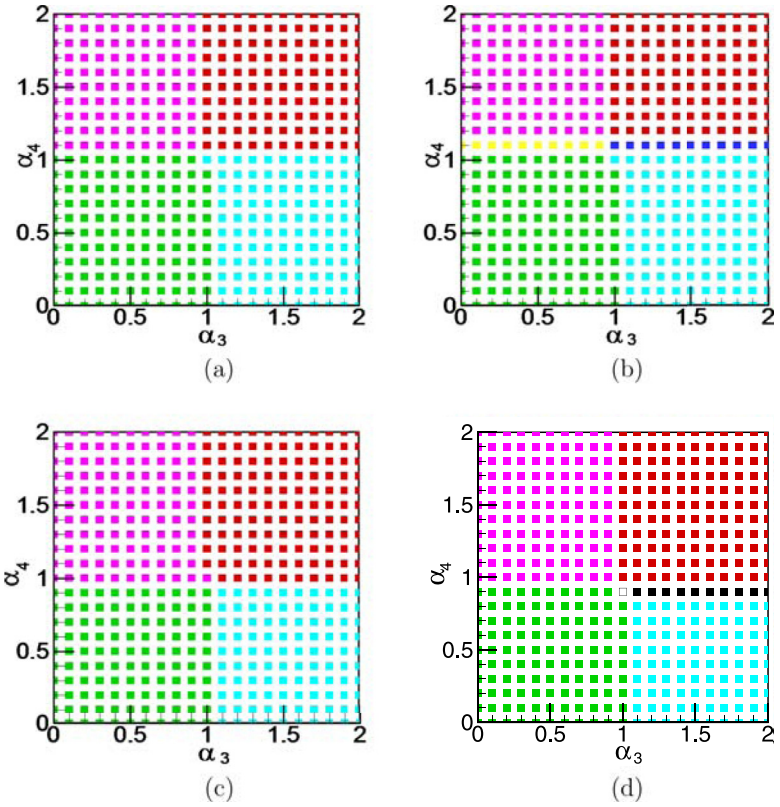

Fig. 13. Changes in phase diagram for Type 1 model under different seedings: (a) $B_1^0 = 1$, $B_n^0 = 1.5$; (b) $B_1^0 = 1$, $B_n'^0 = 1.5$; (c) $B_1^0 = 1$, $B_m^0 = 1.5$; (d) $B_1^0 = 1$, $B_m'^0 = 1.5$.

Appendix B.

The figures below show the variations in probabilities of the various dominant pathways for the Type 2 model for various choices of m and n. The Fig. 14 clearly shows that the outcomes of the study are minimally quantitatively affected by the

choices of m and n and qualitatively remains very similar, showing the emergence of identical paths. The greatest variability occurs for the emergent paths FFO and OFO for higher ratios of m/n. A study on the impact of size of m, n and optimal complexity of the system is currently being conducted and will be reported in our future work.

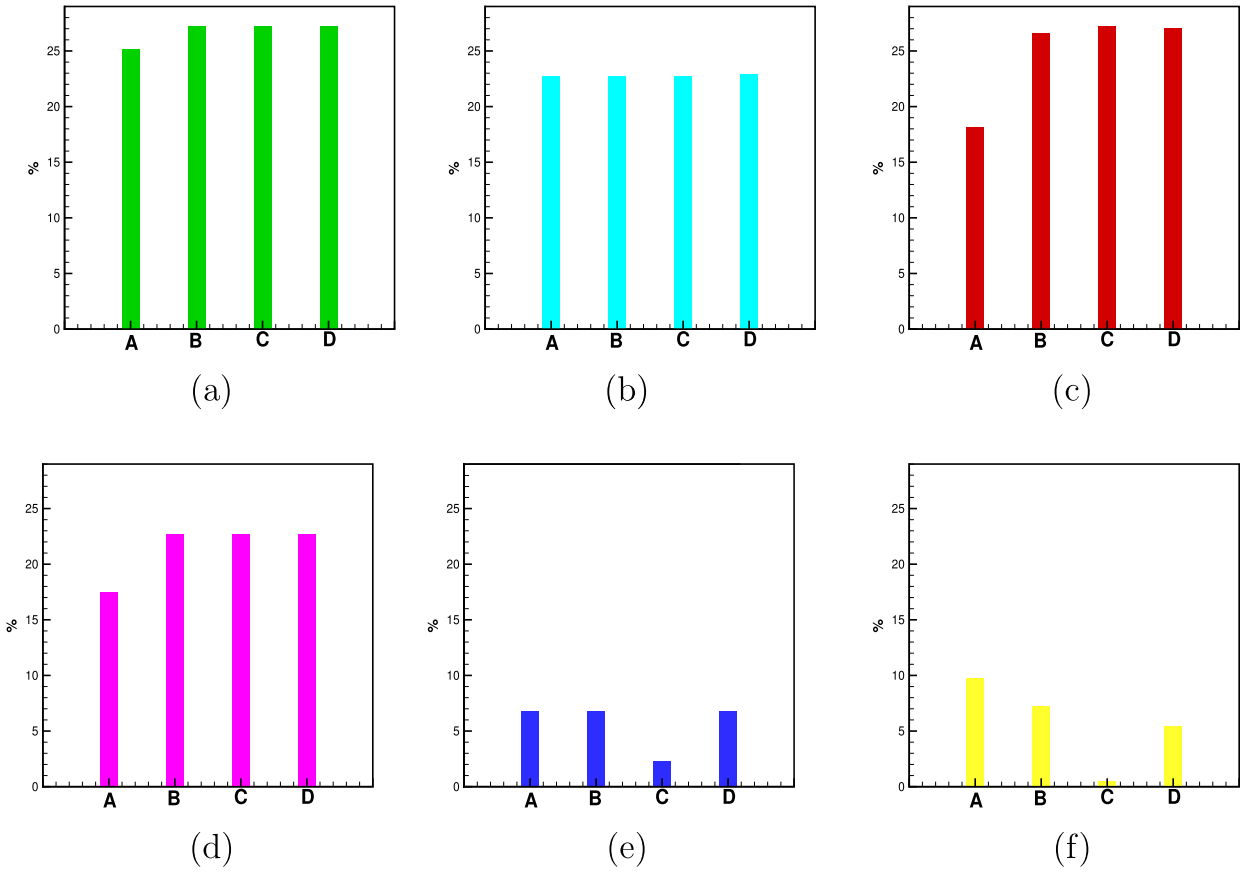

Fig. 14. Percentage of the phase space occupied by different pathways in the six-species system as the choice of (n, m) varies. The panels (a) corresponds to path OOO, (b) to FOO, (c) to FFF, (d) to OFF, (e) to FFO and (f) to OFO. The different cases in the x-axis are to be interpreted as follows: case A: $n = 2, m = 4$; case B: $n = 4, m = 8$; case C: $n = 8, m = 40$ and case D: $n = 12, m = 24$.

References

[1] J.C. Baez, B.S. Pollard, A compositional framework for reaction networks, Rev. Math. Phys. 29 (09) (2017) 1750028.
[2] T. D. Bradley, What is applied category theory?, 2018, arXiv: 1809.05923
[3] F. Chiti, C.M. Dobson, Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade, Annu. Rev. Biochem. 86 (2017) 27–68.
[4] B.J. Chung, B. Ortega, A. Vaidya, Entropy production in a fluid-solid system far from thermodynamic equilibrium, Eur. Phys. J. E 40 (11) (2017) 105.
[5] P. Ghosh, P. Rana, V. Rangachari, J. Saha, E. Steen, A. Vaidya, A game theoretic approach to deciphering the dynamics of amyloid-β aggregation along competing pathways, 2019, arXiv: 581629bioRxiv
[6] P. Ghosh, A. Vaidya, A. Kumar, V. Rangachari, Determination of critical nucleation number for a single nucleation amyloid-β aggregation model, Math. Biosci. 273 (2016) 70–79.
[7] F. Hane, Z. Leonenko, Effect of metals on kinetic pathways of amyloid-β aggregation, Biomolecules 4 (1) (2014) 101–116.
[8] M. Jucker, L.C. Walker, Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders, Ann. Neurol. 70 (4) (2011) 532–540.
[9] S.R. Ji, Y.I. Wu, S.F. Sui, Cholesterol is an important factor affecting the membrane insertion of β-amyloid peptide (aβ1-40), which may potentially inhibit the fibril formation, J. Biol. Chem. 277 (8) (2002) 6273–6279.
[10] D. Kondepudi, I. Prigogine, Modern Thermodynamics: From Heat Engines to Dissipative Structures, John Wiley & Sons, 2014.
[11] L.M. Martyushev, V.D. Seleznev, Maximum entropy production principle in physics, chemistry and biology, Phys. Rep. 426 (1) (2006) 1–45.
[12] D.C. Mikulecky, Network thermodynamics and complexity: a transition to relational systems theory, Comput. Chem. 25 (4) (2001) 369–391.
[13] R.M. Murphy, Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins, Biochimica et Biophysica Acta (BBA)-Biomembranes 1768 (8) (2007) 1923–1934.
[14] B. O'Nuallain, A.D. Williams, P. Westermark, R. Wetzel, Seeding specificity in amyloid growth induced by heterologous fibrils, J. Biol. Chem. 279 (17) (2004) 17490–17499.
[15] G. Oster, A. Perelson, A. Katchalsky, Network thermodynamics, Nature 234 (5329) (1971) 393–399.
[16] A.S. Perelson, Network thermodynamics. an overview, Biophys. J. 15 (7) (1975) 667.
[17] E.T. Powers, D.L. Powers, Mechanisms of protein fibril formation: nucleated polymerization with competing off-pathway aggregation, Biophys. J. 94 (2) (2008) 379–391.
[18] I. Prigogine, Steady states and entropy production, Physica 31 (5) (1965) 719–724.

[19] J. Rasmussen, J. Mahler, N. Beschorner, S.A. Kaeser, L.M. Hasler, F. Baumann, S. Nystrom, E. Portelius, K. Blennow, T. Lashley, N.C, Fox, Amyloid polymorphisms constitute distinct clouds of conformational variants in different etiological subtypes of Alzheimer's disease, Proceedings of the National Academy of Sciences 114 (49) (2017) 13018–13023.
[20] P. Rana, D.N. Dean, E.D. Steen, A. Vaidya, V. Rangachari, P. Ghosh, Fatty acid concentration and phase transitions modulate a-β aggregation pathways, Sci. Rep. 7 (1) (2017) 10370.
[21] V. Rangachari, D.N. Dean, P. Rana, A. Vaidya, P. Ghosh, Cause and consequence of aβ-lipid interactions in alzheimer disease pathogenesis, Biochimica et Biophysica Acta (BBA)-Biomembranes 1860 (9) (2018) 1652–1662.
[22] S. Schilling, T. Lauber, M. Schaupp, S. Manhart, E. Scheel, G. Bohm, H.U. Demuth, On the seeding and oligomerization of pglu-amyloid peptides (in vitro), Biochemistry 45 (41) (2006) 12393–12399.
[23] K.R. Thomas, K.J. Bangen, A.J. Weigand, E.C. Edmonds, C.G. Wong, S. Cooper, ...Alzheimer's disease neuroimaging initiative, Objective subtle cognitive difficulties predict future amyloid accumulation and neurodegeneration, Neurology (2019).
[24] H. Ziegler, C. Wehrli, On a principle of maximal rate of entropy production, J. Non-Equilib. Thermodyn. 12 (3) (1987) 229–244.