Reaction Systems Models for the Self-Assembly of Intermediate Filaments

SEPINOUD AZIMI AND CHARMI PANCHAL
EUGEN CZEIZLER AND ION PETRE

Abstract - Reaction systems are a recent addition to the spectrum of computational modeling frameworks. We construct in this paper several reaction systems models for the self-assembly of vimentin tetramers into intermediate filaments. We demonstrate that reaction systems are a versatile modeling framework, able to capture several aspects of the dynamics of the self-assembly of intermediate filaments using only simple, set-theoretical-based concepts.

Key words and phrases : Reaction systems; modeling; intermediate filaments; self-assembly.

1 Introduction

Reaction systems were introduced in [9] as a new modelling framework inspired by the functioning/bio-energetics of the living cell. It differs drastically from the traditional modelling frameworks (such as ordinary differential equations (ODEs), stochastic processes, Boolean networks, state machines) in focusing on reactions and in having the environment as an integral part of the model. Reactions in reaction systems are enabled through promotion and inhibition, see [4]. Reaction systems are based on two main principles. The first one, called the *threshold principle* makes reaction systems a qualitative framework by stating that if available, a resource (reactant) is available abundantly. In other words, reactions can not limit each other through a quantitative competition on resources. The second principle, called the *no permanency principle*, states that a resource or reactant vanishes unless sustained by a reaction. That is to say, the next state of a reaction system is only obtained from the output of the reaction enabled in the previous state plus the contribution of the environment.

Studies on reaction systems have been quite diverse, see for example [1, 3, 5, 10, 11, 15, 19, 22–25]. They can be categorized into two main streams. While the first direction is concerned with more theoretical aspects of reaction systems (e.g., [8,10,11,24,25]), the second direction has taken a more practical spin, mainly regarding reaction systems as a platform to do bio-modelling (e.g., [1, 3, 5]).

In this paper we follow the second line of research and focus on the expressivity of reaction systems as a modelling framework. We consider a case-study

on the self-assembly of intermediate filaments from vimentin tetramers. We start from the molecular models and the ODE-based analysis in [6]. We build several different reaction systems models based on the molecular model of [6] that capture the self-assembly process at two levels of resolution: one where no distinction is made between filaments of different lengths (resolution 0) and one where the model distinguishes between unit-length, short and long filaments (resolution 2).

2 Preliminaries

In this section we recall some of the basic definitions of reaction systems that we need throughout the paper.

Reactions are the building blocks of a reaction system. Intuitively, a reaction is triggered if all resources needed for the reaction are available in the environment and there exist no species that inhibit the reaction. In this case, the reaction transforms the set of resources to the set of products. This intuition is formally captured in the definition of a reaction in reaction systems framework as follows, see [4, 9] for more details.

Definition 2.1 *[9] A* reaction *is a tuple $a = (R_a, I_a, P_a)$, where R_a, I_a and P_a are finite, non-empty sets and $R_a \cap I_a = \emptyset$. The sets R_a, I_a and P_a are called the set of* reactants*,* inhibitors *and* products *of a, respectively. We say a is a reaction over set S, if R_a, I_a, $P_a \subseteq S$. We denote the set of reactions in S by* $\mathsf{rac}(S)$.

We next define the result of applying a reaction and a set of reactions on a given set. In this definition the result of applying a number of reactions to a set is the collective result of applying each reaction to the set independently. Indeed this is true because by *threshold assumption* there is no competition for the resources between different reactions and hence running a reaction does not inhibit enabling any other one.

Definition 2.2 *[9] Let A be a set of reactions, $a \in A$ and T a set.*

(i) The result *of a on T, denoted by $res_a(T)$, is*

$$res_a(T) = \begin{cases} P_a, & \text{if } R_a \subseteq T \text{ and } I_a \cap T = \emptyset \\ \emptyset, & \text{otherwise.} \end{cases}$$

(ii) The result of A on T*, denoted by $res_A(T)$, is*

$$res_A(T) = \bigcup_{a \in A} res_a(T).$$

A reaction system *(RS in short) is defined as an ordered pair $\mathcal{A} = (S, A)$, where S is a finite set and $A \subseteq \mathsf{rac}(S)$. Set S is called the* background set *of A.*

To capture the dynamics of a given reaction system the notion of *interactive process* has been introduced. In what follows we present a formal definition of such a process.

Definition 2.3 *[9] For a given reaction system $\mathcal{A}$, an* interactive process *in $\mathcal{A}$ is a pair $\pi = (\gamma, \delta)$, where $\gamma = C_0, C_1, ..., C_n, \delta = D_1, D_2, ..., D_n \subseteq S$, $n \geq 1$, with $D_1 = \text{res}_{\mathcal{A}}(C_0)$ and, for every $1 < i \leq n$, $D_i = \text{res}_{\mathcal{A}}(C_{i-1} \cup D_{i-1})$.*

The sequences γ and δ are called the context sequence *and the* result sequence *of π, respectively. The sequence $\tau = W_0, W_1, ..., W_n$ is the* state sequence *of π, where $W_0 = C_0$ and $W_i = C_i \cup D_i$, for all $i \in \{0, ..., n\}$. W_0 is called the* initial state *of π.*

Next we recall the definition of a reaction system's *steady state*, introduced in [3].

Definition 2.4 *[3] Let $\mathcal{A} = (S, A)$ be a reaction system and $C \subseteq S$. We say that $D \subseteq S$ is a* steady state *of $\mathcal{A}$ for C if $\text{res}_{\mathcal{A}}(C \cup D) = D$.*

3 A model for self-assembly of intermediate filaments

In this section we describe the *in-vitro* assembly principles of vimentin filaments, as a representative for the class of intermediate filaments proteins. Based on the recent studies in [6] and [21] we present both a basic and a refined molecular model for vimentin assembly, from the level of first stable subunits till the emerging of mature filaments.

Intermediate filaments (IFs) are one of the three types of protein filaments inside the eukaryotic cell [26]. Together with *microtubules* and *actin filaments*, they form the *cytoskeleton*, which is a complex network of filaments with active role in a number of cellular processes, including sustaining the mechanical integrity of the cell, controlling its shape, but also facilitating the intracellular transport [20].

IF subunits are α-helical rods which assemble by both lateral and end-to-end interactions into rope-like filaments [13]. The length of these filaments ranges from hundreds of nm long to micro-meter values, while their width (when in mature state) is preserved at 11 nm. We are particularly interested in the (*in-vitro*) assembly of IF generated from human vimentin proteins; one of the several types of eukaryotic IF proteins [14]. In the case of vimentin-based IFs the *in-vitro* assembly process follows four stages.

The first assembly stage of vimentin intermediate filaments is the fast lateral association of monomers into dimers and subsequently into tetramers (denoted as T). Tetramers are the first stable filament subunits, as both monomers and dimers are not chemically stable. Moreover, for the case of *in-vitro* assembly, the process can be blocked/freezed after tetramer formation, and the system can be initialized starting from this level. Thus, when modelling the *in-vitro* assembly this first stage

is generally omitted, and the process is assumed to start from a mono-populated system of tetrameric subunits.

The second phase of the vimentin IF assembly consists of further lateral associations: two tetramers join to form an octamer (denoted as O), two octamers join to form a hexadecamer (denoted as H), while two hexadecamer join to form a unit length filament (ULF). ULFs are the basic unit of the mature filament structure, as from now on, the formation and elongation of the filaments is performed only through end-to-end association reactions.

The third stage represents the formation and elongation of filaments from individual ULF's. Here, on one hand we have elongation reactions, when filament complexes are enlarged by one ULF at a time, and merger reactions on the other, where two longer filaments join by end-to-end interactions and form a longer complex. Depending on the number k of constituent ULF's within one filament, we can differentiate between the emergent assemblies based on their "size" k.

The final assembly stage represents a radial compaction of the filaments from a ULF diameter of about 15 nm to a filament diameter of about 11 nm, see [13] for details. However, from a modeler point of view, we can consider the assembly process as complete after the first three stages above, as the radial compaction does not modify the ULF per filament ratio.

3.1 Basic and refined molecular models

A common problem in modelling self-assembly systems is dealing with the combinatorial explosion of all different emergent assemblies as possible distinct species. In the case of the IF model above, this translates into the problem of representing and reasoning about all the emergent filaments of size 1, 2, 3, etc. Depending on the level of details the modeler chooses to describe the assembly process, there might be a number of models which could describe such a process. In our study we concentrate over two such detail levels, thus generating two models for IF assembly; we call them the *basic* and the *refined model*.

In the basic model, we consider the ULF's as elementary filaments (generically denoted as F), and we do not differentiate between them and other filaments. The molecular model of this basic representation is presented in Table 1(a). On the refined model, in order to be able to capture the formation and dynamics of short vs long emerging filaments we differentiate between ULFs (filaments of size 1, denoted as F_u), filaments of size 2 (short filaments, denoted as F_s), and filaments of size 3 or more (long filaments, denoted as F_l). The molecular model of this second refined system is presented in Table 1(b).

The above basic model, as well as a refined version differentiating between ULFs and the remaining filaments, were introduced and analyzed in [6] in correlation with experimental results for *in-vitro* vimentin self-assembly taken from [18]. ODE mathematical models based on mass action kinetics formulations were derived, numerically fitted, and validated using data from [18]. A generic method of quantitative model refinement was introduced and discussed in [12] and [21].

Table 1: The molecular models of (a) the basic and (b) the refined representations of the IF assembly process.

(a) Basic model		(b) Refined model	
$2\,\mathsf{T} \to \mathsf{O}$	(1)	$2\,\mathsf{T} \to \mathsf{O}$	(5)
$2\,\mathsf{O} \to \mathsf{H}$	(2)	$2\,\mathsf{O} \to \mathsf{H}$	(6)
$2\,\mathsf{H} \to \mathsf{F}$	(3)	$2\,\mathsf{H} \to \mathsf{F}_\mathsf{u}$	(7)
$2\,\mathsf{F} \to \mathsf{F}$	(4)	$2\,\mathsf{F}_\mathsf{u} \to \mathsf{F}_\mathsf{s}$	(8)
		$\mathsf{F}_\mathsf{u} + \mathsf{F}_\mathsf{s} \to \mathsf{F}_\mathsf{l}$	(9)
		$\mathsf{F}_\mathsf{u} + \mathsf{F}_\mathsf{l} \to \mathsf{F}_\mathsf{l}$	(10)
		$2\,\mathsf{F}_\mathsf{s} \to \mathsf{F}_\mathsf{l}$	(11)
		$\mathsf{F}_\mathsf{s} + \mathsf{F}_\mathsf{l} \to \mathsf{F}_\mathsf{l}$	(12)
		$2\,\mathsf{F}_\mathsf{l} \to \mathsf{F}_\mathsf{l}$	(13)

Based on those methods we can estimate the kinetic rate constants of the current refined model in Table 1(b), in order to obtain a perfect fit-preserving refinement of the models in [6].

3.2 Variants of the kinetic model

The previously introduced basic and refined models were numerically fitted in [6] in order to corroborate the experimental data of vimentin assembly reported in [18]. However, we show in the following that by modifying the kinetic rate constants of these reactions we can generate several setups with different overall behaviours. These behaviours could be differentiated by measuring the average length of the emerging filament populations. The formula used for generating these measurements was introduced in [18] and updated in [6]. The formula can be easily extended for any level of refinement, including the one considered in our models.

In Figure 1 we present the average filament length over time for the model of [6], as well as for two different kinetic setups of the same model. The plot in the case of the original model is shown in Figure 1 with a solid line. We modified the model of [6] by inhibiting reactions (11) and (12); the result is shown with a dashed line in Figure 1. In another modification of the model in [6], we inhibited reactions (9) and (11); the result is shown with a dotted line in Figure 1. As it can be seen from Figure 1, the three kinetic setups of the model in [6] yield very different results. In the original setup of [6], the model favors the formation of (fewer and fewer) long filaments. The other two setups favour the formation of both short and long filaments, and that of only short filaments, respectively.

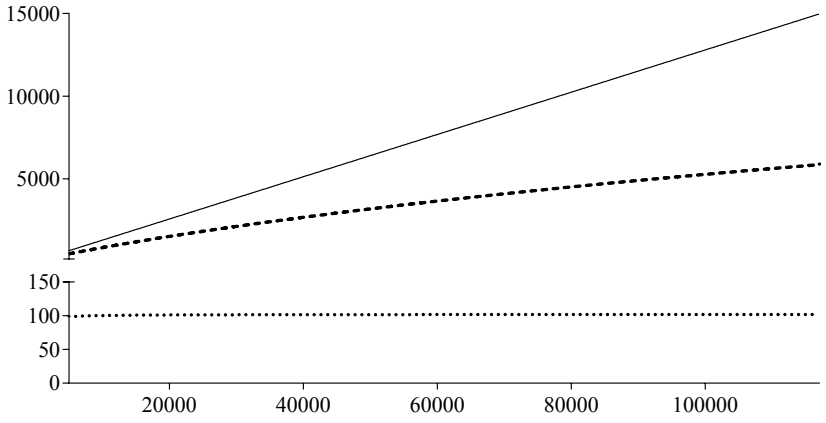

Figure 1: The plot of the time evolution of the average filament length in three different setups. The solid line plot represents the original model of [6], favouring the formation of long filaments. The dashed line plots represents a variant of the model where both short and long filaments are formed. The dotted line plot represents a variant of the model where only short filaments are formed. The three models have the same set of reactions and only differ in their kinetic setups.

4 Reaction systems models

In this section we present the corresponding RS-based models for self-assembly of intermediate filaments introduced in Section 3.

4.1 Basic reaction systems model

First we build an RS-based model for the basic self-assembly of intermediate filaments introduced in Table 1(a). To formulate the reactions of the corresponding RS, every reaction of type $A_1 + A_2 \to B$ in Table 1(a) is translated to a reaction $(\{A_1, A_2\}, \{\mathsf{d_I}\}, \{B\})$ in the corresponding reaction system. The dummy variable, $\mathsf{d_I}$, is only used to respect the constraint that the set of inhibitors of all RS reactions should be non-empty, see [9]. Note that the coefficients of the species in the chemical reaction do not play a role here since we are translating a quantitative framework to a qualitative one and by *threshold assumption* we know that existence of a reactant in the environment implies the abundance of it as well. The RS-based model for the basic self-assembly of intermediate filaments is presented in Table 2.

We are interested in analysing the dynamics of the RS-based model and compare it with the corresponding properties of the quantitative ODE-based model for IF assembly, discussed in Section 3. To construct an environment in which tetramers are always present, we consider $\{\mathsf{T}\}$ as the given context in every step of the interactive process. This corresponds to the ODE models having an initial

Table 2: The direct translation of the biochemical reactions of the basic model to a reaction system $\mathcal{A} = (S, A)$ where $S = \{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F}, \mathsf{d_I}\}$.

Reaction in the chemical network	Reaction in the reaction system	
$2\,\mathsf{T} \to \mathsf{O}$	$(\{\mathsf{T}\}, \{\mathsf{d_I}\}, \{\mathsf{O}\})$	(14)
$2\,\mathsf{O} \to \mathsf{H}$	$(\{\mathsf{O}\}, \{\mathsf{d_I}\}, \{\mathsf{H}\})$	(15)
$2\,\mathsf{H} \to \mathsf{F}$	$(\{\mathsf{H}\}, \{\mathsf{d_I}\}, \{\mathsf{F}\})$	(16)
$2\,\mathsf{F} \to \mathsf{F}$	$(\{\mathsf{F}\}, \{\mathsf{d_I}\}, \{\mathsf{F}\})$	(17)

large pool of tetramers that are assembling into IFs. The interactive process is illustrated in Table 3. The result state sequences of the examples of this paper were obtained by using the reaction system simulator proposed in [2]. The simulator can be reached at [27]. The interactive process thus obtained shows that the model enters into a steady state, see [3], where tetramers, octamers, hexadecamers and filaments are present.

Table 3: An interactive process for the basic RS model. The interactive process enters a loop after the second state from which every state contains all species of the system.

State	C_i	D_i	W_i
0	$\{\mathsf{T}\}$	$\emptyset$	$\{\mathsf{T}\}$
1	$\{\mathsf{T}\}$	$\{\mathsf{O}\}$	$\{\mathsf{T}, \mathsf{O}\}$
2	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}\}$
3	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F}\}$
4	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F}\}$

4.2 Refined reaction systems model

In this section we are taking a step further to increase the level of resolution of our model. To do so, we apply the proposed refinement approach of Table 1(b) and modify our basic model to fit this new information. The methodology for translating the chemical reaction network to the RS-based model does not change with the new setup. The obtained RS-based model is presented in Table 4. The species $\mathsf{F_u}$, $\mathsf{F_s}$ and $\mathsf{F_l}$ correspond to *unit length filament*, *short filament* and *long filament* respectively.

Similarly as in the case of the basic model, we analyze the dynamics of the refined model by running an interactive process with a constant $\{\mathsf{T}\}$ as the given context in every step. The interactive process is presented in Table 5. We conclude

Table 4: The direct translation of the biochemical reactions of the refined model to a reaction system $\mathcal{A}' = (S', A')$ where $S' = \{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}, \mathsf{d_l}\}$.

Reaction in the chemical network	Reaction in the reaction system	
$2\,\mathsf{T} \to \mathsf{O}$	$(\{\mathsf{T}\}, \{\mathsf{d_l}\}, \{\mathsf{O}\})$	(18)
$2\,\mathsf{O} \to \mathsf{H}$	$(\{\mathsf{O}\}, \{\mathsf{d_l}\}, \{\mathsf{H}\})$	(19)
$2\,\mathsf{H} \to \mathsf{F_u}$	$(\{\mathsf{H}\}, \{\mathsf{d_l}\}, \{\mathsf{F_u}\})$	(20)
$2\,\mathsf{F_u} \to \mathsf{F_s}$	$(\{\mathsf{F_u}\}, \{\mathsf{d_l}\}, \{\mathsf{F_s}\})$	(21)
$\mathsf{F_u} + \mathsf{F_s} \to \mathsf{F_l}$	$(\{\mathsf{F_u}, \mathsf{F_s}\}, \{\mathsf{d_l}\}, \{\mathsf{F_l}\})$	(22)
$2\,\mathsf{F_s} \to \mathsf{F_l}$	$(\{\mathsf{F_s}\}, \{\mathsf{d_l}\}, \{\mathsf{F_l}\})$	(23)
$\mathsf{F_u} + \mathsf{F_l} \to \mathsf{F_l}$	$(\{\mathsf{F_u}, \mathsf{F_l}\}, \{\mathsf{d_l}\}, \{\mathsf{F_l}\})$	(24)
$\mathsf{F_s} + \mathsf{F_l} \to \mathsf{F_l}$	$(\{\mathsf{F_s}, \mathsf{F_l}\}, \{\mathsf{d_l}\}, \{\mathsf{F_l}\})$	(25)
$2\,\mathsf{F_l} \to \mathsf{F_l}$	$(\{\mathsf{F_l}\}, \{\mathsf{d_l}\}, \{\mathsf{F_l}\})$	(26)

that, similarly as in the case of the basic model, we reach a steady state where tetramers, octamers, hexadecamers, and the three types of filaments are present. This is of course consistent with this model being a refinement of the basic model.

Table 5: An interactive process for the refined RS model. The interactive process enters a loop after the fourth state from which every state contains all species of the system.

State	C_i	D_i	W_i
0	$\{\mathsf{T}\}$	$\emptyset$	$\{\mathsf{T}\}$
1	$\{\mathsf{T}\}$	$\{\mathsf{O}\}$	$\{\mathsf{T}, \mathsf{O}\}$
2	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}\}$
3	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$
4	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$
5	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}\}$
6	$\{\mathsf{T}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}\}$

4.3 Variants of the refined reaction systems model

We consider in this section the corresponding reaction systems of the setup considered in Section 3.2 for the ODE-based model: we focus on modifying our RS model to control the length of the filaments produced within the system. Three different variants of the RS-based model proposed in Section 3 are presented. In the first variant, we are interested in having only long filaments in the result state,

whereas in the second one we expect to produce only short filaments. The last variation is responsive to what is asked by the modeler through the context of the system, i.e. reaction system produces only short filaments if short is part of the context and long filaments if long is included in the context. If the context contains neither short nor long, the filaments of all lengths are present in the result state. In what follows we discuss each of these variations separately.

Variant one - only long filaments. We have modified the reaction system presented in Table 4 to produce only short filaments by adding $\{F_l\}$ to the set of inhibitors of reactions (21) and (22). In this way, we effectively favour reaction (23) over reactions (21) and (22) and the unit length filaments immediately get to elongate the existing long filaments rather than short filaments. The corresponding RS model is presented in Table 6. The efficiency of the reaction system structure can be observed through the steps of the interactive process illustrated in Table 7: in the steady state there are no short filaments.

Variant two - only short filaments. In this case we modified the reaction system of Table 4 by removing reactions (22) and (23). The corresponding reaction system is presented in Table 6. On one hand, if no long filaments are introduced by the context, the system produces only short filaments. On the other hand, if long filaments are added by the context, they would get extended recurrently. Both cases can be observed in the interactive process shown in Table 8.

Table 6: The list of reactions of the reaction systems corresponding to the first two variants of the refined model. For both reaction systems the background set is $\{T, O, H, F_u, F_s, F_l, d_l\}$.

Reactions in variant one		Reactions in variant two	
$(\{T\}, \{d_l\}, \{O\})$	(27)	$(\{T\}, \{d_l\}, \{O\})$	(36)
$(\{O\}, \{d_l\}, \{H\})$	(28)	$(\{O\}, \{d_l\}, \{H\})$	(37)
$(\{H\}, \{d_l\}, \{F_u\})$	(29)	$(\{H\}, \{d_l\}, \{F_u\})$	(38)
$(\{F_u\}, \{F_l\}, \{F_s\})$	(30)	$(\{F_u\}, \{d_l\}, \{F_s\})$	(39)
$(\{F_u, F_s\}, \{F_l\}, \{F_l\})$	(31)	$(\{F_u, F_l\}, \{d_l\}, \{F_l\})$	(40)
$(\{F_s\}, \{d_l\}, \{F_l\})$	(32)	$(\{F_s, F_l\}, \{d_l\}, \{F_l\})$	(41)
$(\{F_u, F_l\}, \{d_l\}, \{F_l\})$	(33)	$(\{F_l\}, \{d_l\}, \{F_l\})$	(42)
$(\{F_s, F_l\}, \{d_l\}, \{F_l\})$	(34)		
$(\{F_l\}, \{d_l\}, \{F_l\})$	(35)		

Table 7: An interactive process for the first variant of the refined RS model. The interactive process enters a loop after the fifth state from which no short filament is produced.

State	C_i	D_i	W_i
0	$\{\mathsf{T}\}$	$\emptyset$	$\{\mathsf{T}\}$
1	$\{\mathsf{T}\}$	$\{\mathsf{O}\}$	$\{\mathsf{T},\mathsf{O}\}$
2	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H}\}$
3	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u}\}$
4	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s}\}$
5	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$
6	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_l}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_l}\}$
7	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_l}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_l}\}$

Table 8: An interactive process for the second variant of the refined RS model. The interactive process enters a loop after the third state from which no long filament is produced until $\mathsf{F_l}$ is introduced to the fifth state which triggers the production of long filaments in the consecutive states.

State	C_i	D_i	W_i
0	$\{\mathsf{T}\}$	$\emptyset$	$\{\mathsf{T}\}$
1	$\{\mathsf{T}\}$	$\{\mathsf{O}\}$	$\{\mathsf{T},\mathsf{O}\}$
2	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H}\}$
3	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u}\}$
4	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s}\}$
5	$\{\mathsf{T},\mathsf{F_l}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$
6	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$
7	$\{\mathsf{T}\}$	$\{\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$	$\{\mathsf{T},\mathsf{O},\mathsf{H},\mathsf{F_u},\mathsf{F_s},\mathsf{F_l}\}$

Variant three - context dependent. To fit the requirements of the final variation, the reaction system of Table 4 is modified as follows:

- The background set is extended by adding three species, i.e. $\mathsf{short}, \mathsf{long}$ and $\mathsf{F_{long}}$, to set S;
- $\mathsf{F_{long}}$ is added to the inhibitor set of reaction (21) to prevent producing short filaments whenever the context asks for the long ones;
- short is added to the inhibitor set of reactions (22), (23), (24), (25) and (26) to prevent producing long filaments whenever the context asks for the short ones and
- the set of reactions is extended by adding $(\{\mathsf{F_l}, \mathsf{long}\}, \{\mathsf{d_l}\}, \{\mathsf{F_{long}}\})$ to signal the start of long filament production to the system.

The corresponding reaction system is presented in Table 9, while its interactive process and the steady state with constant context $\{\mathsf{T}, \mathsf{short}\}$ and with constant context $\{\mathsf{T}, \mathsf{long}\}$ is described in Table 10.

Table 9: The list of reactions of the last variant of the refined model for reaction system $\mathcal{A}'' = (S'', A'')$ where $S'' = \{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}, \mathsf{long}, \mathsf{short}, \mathsf{F_{long}}, \mathsf{d_l}\}$.

Reaction		Reaction	
$(\{\mathsf{T}\}, \{\mathsf{d_l}\}, \{\mathsf{O}\})$	(43)	$(\{\mathsf{F_s}\}, \{\mathsf{short}\}, \{\mathsf{F_l}\})$	(48)
$(\{\mathsf{O}\}, \{\mathsf{d_l}\}, \{\mathsf{H}\})$	(44)	$(\{\mathsf{F_u}, \mathsf{F_l}\}, \{\mathsf{short}\}, \{\mathsf{F_l}\})$	(49)
$(\{\mathsf{H}\}, \{\mathsf{d_l}\}, \{\mathsf{F_u}\})$	(45)	$(\{\mathsf{F_s}, \mathsf{F_l}\}, \{\mathsf{short}\}, \{\mathsf{F_l}\})$	(50)
$(\{\mathsf{F_u}\}, \{\mathsf{F_{long}}\}, \{\mathsf{F_s}\})$	(46)	$(\{\mathsf{F_l}\}, \{\mathsf{short}\}, \{\mathsf{F_l}\})$	(51)
$(\{\mathsf{F_u}, \mathsf{F_s}\}, \{\mathsf{short}\}, \{\mathsf{F_l}\})$	(47)	$(\{\mathsf{F_l}, \mathsf{long}\}, \{\mathsf{d_l}\}, \{\mathsf{F_{long}}\})$	(52)

5 Discussion

We investigated in this paper the expressivity of the reaction systems framework as a modelling formalism for biology. Our case study, on protein self-assembly, involves intermediate filaments of arbitrary size. This made the case study a good choice to test the expressivity of modelling with reaction systems. We showed that our reaction systems model is a good qualitative counterpart to quantitative modelling with ODE. Both types of models could demonstrate, albeit with a different language and tools, the formation of intermediate filaments from vimentin tetramers. In the case of ODE models we showed that by disabling various reactions, we could observe different outputs of the self assembly model. Thus, while

Table 10: An interactive process for the third variant of the refined RS model. In this example short is added to the context and the interactive process enters a loop after the third state from which no long filaments are produced. Next, long is added to the context and the interactive process produces only long filaments up to state 9. Once short is added again to the context in state 10, the system only produces short filaments.

State	C_i	D_i	W_i
0	$\{\mathsf{T}, \mathsf{short}\}$	$\emptyset$	$\{\mathsf{T}\}$
1	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}\}$	$\{\mathsf{T}, \mathsf{O}\}$
2	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}\}$
3	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$
4	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$
5	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$
6	$\{\mathsf{T}, \mathsf{long}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}, \mathsf{F_{long}}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}, \mathsf{F_l}, \mathsf{F_{long}}\}$
7	$\{\mathsf{T}, \mathsf{long}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_l}, \mathsf{F_{long}}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_l}, \mathsf{F_{long}}\}$
8	$\{\mathsf{T}, \mathsf{long}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_l}, \mathsf{F_{long}}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_l}, \mathsf{F_{long}}\}$
9	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}\}$
10	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$
11	$\{\mathsf{T}, \mathsf{short}\}$	$\{\mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$	$\{\mathsf{T}, \mathsf{O}, \mathsf{H}, \mathsf{F_u}, \mathsf{F_s}\}$

in the original setup the system favors the formation of long filaments, two altered variants were generated in which we produced either short filaments or a combination of short and long filaments as the output. We showed that reaction systems are equally versatile and we built three versions of the model demonstrating the same behaviour as that of the ODE models: only short, only long and both (short and long filaments). Moreover, our final reaction systems model is able to switch its preferred output target between short and long filaments, depending on the trigger coming from the environment.

Building a reaction systems model with a qualitative behaviour "similar" to the quantitative behaviour of an ODE model is in general a difficult problem. Starting from a chemical reaction network, an ODE model can be built using standard kinetic principles such as the law of mass action, whereas building a reaction systems model can be quite intricate, see, e.g. [3]. In the case of the intermediate filaments, building the reaction systems models was straightforward because the chemical reaction network we started from contained no feedback loops.

We built our reaction systems models in two steps: we first built a basic model that did not distinguish between the filaments and then we refined it to a more detailed one that distinguished between unit-length, short, and long filaments. The technique of building models by adding details step by step in called *model refinement* and it is currently being investigated in connection to several different modelling frameworks, e.g., rule-based models [7], ODE models [16], and Petri net models [17]. Exploring in more details the refinement of reaction systems models seems an interesting research topic and we plan to return to it.

Acknowledgements

The authors acknowledge the support of Academy of Finland through projects 267915 and 272559. Charmi Panchal acknowledges the support from Otto A. Malm foundation and from Oskar Öflunds stiftelse foundation.

References

[1] Sepinoud Azimi, Cristian Gratie, Sergiu Ivanov, Luca Manzoni, Ion Petre, and Antonio E. Porreca. Complexity of model checking for reaction systems. Technical Report 1122, Turku Centre for Computer Science, 2014.

[2] Sepinoud Azimi, Cristian Gratie, Sergiu Ivanov, and Ion Petre. Dependency graphs and mass conservation in reaction systems. Technical Report 1123, 2014.

[3] Sepinoud Azimi, Bogdan Iancu, and Ion Petre. Reaction system models for the heat shock response. *Fundamenta Informaticae*, 131(3):299–312, 2014.

[4] Robert Brijder, Andrzej Ehrenfeucht, Michael Main, and Grzegorz Rozenberg. A tour of reaction systems. *International Journal of Foundations of Computer Science*, 22(07):1499–1517, 2011.

[5] Luca Corolli, Carlo Maj, Fabrizio Marini, Daniela Besozzi, and Giancarlo Mauri. An excursion in reaction systems: From computer science to biology. *Theoretical Computer Science*, 454:95–108, 2012.

[6] Eugen Czeizler, Andrzej Mizera, Elena Czeizler, Ralph-Johan Back, John E Eriksson, and Ion Petre. Quantitative analysis of the self-assembly strategies of intermediate filaments from tetrameric vimentin. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 9(3):885–898, 2012.

[7] Vincent Danos, Jerome Feret, Walter Fontana, Russ Harmer, and Jean Krivine. Rule-based modelling and model perturbation. *Transactions on Computational Systems Biology XI*, pages 116–137, 2009.

[8] Andrzej Ehrenfeucht, Michael Main, and Grzegorz Rozenberg. Functions defined by reaction systems. *International Journal of Foundations of Computer Science*, 22(01):167–178, 2011.

[9] Andrzej Ehrenfeucht and Grzegorz Rozenberg. Reaction systems. *Fundamenta Informaticae*, 75(1):263–280, 2007.

[10] Enrico Formenti, Luca Manzoni, and Antonio E. Porreca. Cycles and global attractors of reaction systems. In Helmut Jürgensen, Juhani Karhumäki, and Alexander Okhotin, editors, *Descriptional Complexity of Formal Systems*, volume 8614 of *Lecture Notes in Computer Science*, pages 114–125. Springer, 2014.

[11] Enrico Formenti, Luca Manzoni, and Antonio E. Porreca. Fixed points and attractors of reaction systems. In Arnold Beckmann, Erzsébet Csuhaj-Varjú, and Klaus Meer, editors, *Language, Life, Limits, 10th Conference on Computability in Europe, CiE 2014*, volume 8493 of *Lecture Notes in Computer Science*, pages 194–203. Springer, 2014.

[12] Cristian Gratie and Ion Petre. Fit-preserving data refinement of mass-action reaction networks. In *Language, Life, Limits*, pages 204–213. Springer, 2014.

[13] Harald Herrmann and Ueli Aebi. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. *Annual review of biochemistry*, 73(1):749–789, 2004.

[14] Harald Herrmann, Markus Häner, Monika Brettel, Nam-On Ku, and Ueli Aebi. Characterization of distinct early assembly units of different intermediate filament proteins. *Journal of molecular biology*, 286(5):1403–1420, 1999.

[15] Mika Hirvensalo. On probabilistic and quantum reaction systems. *Theoretical Computer Science*, 429:134–143, 2012.

[16] Bogdan Iancu, Elena Czeizler, Eugen Czeizler, and Ion Petre. Quantitative refinement of reaction models. *International Journal of Unconventional Computing*, 8(5-6):529–550, 2012.

[17] Bogdan Iancu, Diana-Elena Gratie, Sepinoud Azimi, and Ion Petre. On the implementation of quantitative model refinement. In *Algorithms for Computational Biology*, pages 95–106. Springer, 2014.

[18] Robert Kirmse, Stephanie Portet, Norbert Mücke, Ueli Aebi, Harald Herrmann, and Jörg Langowski. A quantitative kinetic model for the in vitro assembly of intermediate filaments from tetrameric vimentin. *The Journal of biological chemistry*, 282(25):18563–72, June 2007.

[19] Jetty Kleijn and Maciej Koutny. Membrane systems with qualitative evolution rules. *Fundamenta Informaticae*, 110(1):217–230, 2011.

[20] Elias Lazarides. Intermediate filaments as mechanical integrators of cellular space. *Nature*, 283(5744):249–256, 1980.

[21] Andrzej Mizera, Eugen Czeizler, and Ion Petre. Transactions on computational systems biology xiv. chapter Self-assembly Models of Variable Resolution, pages 181–203. Springer-Verlag, Berlin, Heidelberg, 2012.

[22] Fumiya Okubo, Satoshi Kobayashi, and Takashi Yokomori. Reaction automata. *Theoretical Computer Science*, 429:247–257, 2012.

[23] Gheorghe Păun, Mario J Pérez-Jiménez, and Grzegorz Rozenberg. Bridging membrane and reaction systems–further results and research topics. *Fundamenta Informaticae*, 127(1):99–114, 2013.

[24] Arto Salomaa. Functions and sequences generated by reaction systems. *Theoretical Computer Science*, 466:87–96, 2012.

[25] Arto Salomaa. Functional constructions between reaction systems and propositional logic. *International Journal of Foundations of Computer Science*, 24(1):147–159, 2013.

[26] Manfred Schliwa. The cytoskeleton. cell biology monographs, vol. 13, 1986.

[27] The web interface of the reaction system simulator. http://combio.abo.fi/research/reaction-systems/reaction-system-simulator.

Sepinoud Azimi
Computational Biomodeling Laboratory
Turku Centre for Computer Science
Åbo Akademi University
Joukahaisenkatu 3-5A, 20520 Turku, Finland
E-mail: sazimi@abo.fi

Charmi Panchal
Computational Biomodeling Laboratory
Turku Centre for Computer Science
Åbo Akademi University
Joukahaisenkatu 3-5A, 20520 Turku, Finland
E-mail: cpanchal@abo.fi

Eugen Czeizler
Computational Biomodeling Laboratory
Turku Centre for Computer Science
Åbo Akademi University
Joukahaisenkatu 3-5A, 20520 Turku, Finland
and
Department of Computer Science
Aalto University
PO Box 15400, 00076 Aalto, Finland
E-mail: eczeizle@abo.fi

Ion Petre
Computational Biomodeling Laboratory
Turku Centre for Computer Science
Åbo Akademi University
Joukahaisenkatu 3-5A, 20520 Turku, Finland
E-mail: ipetre@abo.fi