

Mathematical Modelling of OAS2 activation by
dsRNA and effects of dsRNA lengths

by

Deokro Lee

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Mathematics
University of Manitoba
Winnipeg

Copyright © 2020 by Deokro Lee

Abstract

The activation of 2'-5'-oligoadenylate synthetase (OAS) enzymes by direct interaction with viral double-stranded RNA (dsRNA) is a key part of the innate immune response to viral infection. The downstream effect of the OAS-dsRNA interaction is to degrade the viral single-stranded RNA (ssRNA) to prevent the spread of the virus.

The entire OAS activation mechanism is complex and not yet well understood. Based on experimental data, the process appears to depend on concentrations and lengths of dsRNA; however, it cannot be completely observed in experiments. Hence, mathematical models can help to understand the detailed OAS activation mechanism and the effects of dsRNA lengths. Plausible biochemical scenarios are translated to mathematical models and their responses are compared to *in vitro* experimental data provided by McKenna's lab to test different hypotheses. In total, nine models are derived from enzyme kinetics; and their mathematical analyses and numerical investigations are provided. Model selection methods are used to determine the best model accommodating different dsRNA concentrations and lengths.

Acknowledgment Page

Thanks for all the help me write this thesis.

Thanks to the department of Mathematics at the University of Manitoba for the financial support, which helps me to finish my graduate life successfully. Thanks to the support staffs for help with kindness and smile. Thanks to Dr. Leo Butler, Dr. Craig Cowan, Dr. Richard Slevinsky and Dr. Julien Arino for the qualified course works, which gave me the guidance and enlightenment.

Thanks to Dr. Kang-Ling Liao for the helpful comments to improve this thesis.

I really appreciate Dr. Sean McKenna providing the data for writing this thesis and spending much time to guide me.

I sincerely appreciate everything Dr. Stéphanie Portet has done for me. As a supervisor, she allowed me to be a master student and novice researcher. As a senior, she showed me the perseverance and consideration. Her instruction will be remained in me forever.

Dedication Page

To my father, mother and brother, who have raised me in happiness,
to father-in-law and mother-in-law, who give me warm consideration,
to uncle Ron and aunt Young, who help and love me generously,
to Jae Gu, who is my true friend.

Finally, to my lovely Lovle, Youngeun, who I loved, love and will love.

Contents

Contents	iii
List of Tables	vii
List of Figures	ix
1 Introduction	1
1.1 OAS family in innate immune system	1
1.2 Regulation of activation of OAS2	3
1.3 Organization of thesis	4
2 Mathematical Modelling	7
2.1 Group 1 - Model-S	10
2.2 Group 2	11
2.2.1 Model-E: constant rate	11
2.2.2 Model-ER and Model-EC: length dependent rates	12

2.2.3	Model-ED	21
2.3	Group 3 - Model-N	22
2.4	Group 4	24
2.4.1	Model-M: constant rate	25
2.4.2	Model-MR and Model-MC: length dependent rates	27
2.5	Overview	27
3	Mathematical Analysis	31
3.1	Group 1 - Model-S	32
3.2	Wellposedness - Group 2 to 4	32
3.3	Group 2	34
3.4	Group 3 - Model-N	45
3.4.1	Quasi-steady-state assumption for Model-N	45
3.4.2	Asymptotic behavior of Model-N	56
3.5	Group 4	63
3.6	About the Product	74

4 Numerical Investigations	79
4.1 Models versus experimental data	82
4.1.1 Model-S	82
4.1.2 Fitting: all concentrations at one length	83
4.1.3 Fitting: all concentrations at all lengths	87
4.2 Model Selection	97
5 Discussion and conclusion	103
5.1 Mathematical remarks and discussion	103
5.2 Model evaluations	107
5.3 Limitation and idea for future study	109
Appendices	113
A Model-M: model equations	115
Bibliography	119

List of Tables

4.1	Mean, standard deviation (SD) and best parameter set for Model-E, Model-ED, Model-N and Model-M obtained when fitting all concentrations at one length of dsRNA, 120bp.	86
4.2	Mean, standard deviation (SD) and the best parameter set for Model-E, Model-ED, Model-N and Model-M with all concentrations at all lengths of dsRNA.	88
4.3	Length dependent parts of binding rate; $\sigma_1(L)$ and $\sigma_2(L)$	92
4.4	Mean and standard deviation (SD) of parameters of Model-ER, Model-EC, Model-MR and Model-MC	94
4.5	AIC, AICc and Akaike weights for four models of OAS2 activation by dsRNA (one length).	98
4.6	AIC, AICc and Akaike weights for eight models of OAS2 activation by dsRNA (all lengths). The ranks are obtained by sorting AIC and AICc values from the smallest to the largest.	99

List of Figures

1.1	The simple diagram for OAS2 activation by dsRNA and the immune response. The focus of this thesis is represented in the red frame in Figure 1.1.	2
1.2	Concentrations of product, PPI, for each concentration and length of dsRNA over time. Number codes for the length: (1) 40bp, (2) 50bp, (3) 60bp, (4) 70bp, (5) 80bp, (6) 90bp and (7) 120bp. Letters represent concentrations: (a) $0.5\mu\text{g/mL}$, (b) $1\mu\text{g/mL}$, (c) $2\mu\text{g/mL}$, (d) $4\mu\text{g/mL}$ and (e) $8\mu\text{g/mL}$. The range of y-axis (the concentration of the product) is from 0 to $130\mu\text{g/mL}$ and the range of x-axis (time) is from 0 to 30 minutes.	5
2.1	Compounds considered in models and their corresponding variables. Note that D_i refer to a complex formed of a dsRNA bound to i OAS2.	8
2.2	List of parameters used in models.	9
2.3	Model-S	10

2.4	Model-E	11
2.5	Binding rate depending on cofactor length L . Replace K_1 in Model-E by $k_1\sigma_{1,2}(L)$, where $\sigma_{1,2}(L)$ is $\sigma_1(L)$ for Model-ER or $\sigma_2(L)$ for Model-EC.	13
2.6	Dashed circles delimit the interaction zones for OAS2 and dsRNA. L is length of dsRNA, R_O is the radius of OAS2.	14
2.7	Valid binding domain when dsRNA is represented as a rigid rod.	17
2.8	dsRNA is described as a cylinder of length L and diameter d . R_O : radius of OAS2, k : ratio for valid binding area.	19
2.9	Model-ED with degradation of OAS2 or complex, \emptyset : degradation	21
2.10	Model-N with a two-step binding and non-productive bindings.	23
2.11	Model-M, $n = \left\lfloor \frac{L}{2R_O} \right\rfloor$, n is the maximum number of OAS2 that can be bound to one dsRNA simultaneously.	25
2.12	Tree of models: the relations among the models.	29
3.1	The phase line for (3.6).	38
4.1	The experimental data: concentrations of product as a function of c_R with fixed t	82

4.2	Experimental data (solid line) and prediction for the product P (dashed line) of (a) Model-E, (b) Model-ED, (c) Model-N and (d) Model-M by using all concentrations and one length (120bp) of dsRNA. Color code is given in Figure 4.3.	84
4.3	Legend table for Figure 4.2. High to low concentrations: rainbow order.	84
4.4	Affinities of reactions for (a) Model-E, (b) Model-ED, (c) Model-N and (d) Model-M by using all concentrations at one length (120bp) of dsRNA. Ten outliers are ignored for k_{5m}/k_5 in (c), outliers > 110	85
4.5	Experimental data (solid line) and prediction for the product P (dashed line) of (1) Model-E, (2) Model-ED, (3) Model-N and (4) Model-M. Each column presents one concentration of dsRNA: (a) $0.5\mu\text{g/mL}$, (b) $1\mu\text{g/mL}$, (c) $2\mu\text{g/mL}$, (d) $4\mu\text{g/mL}$ and (e) $8\mu\text{g/mL}$. Color code is given in Figure 4.6.	89
4.6	Legend table for fitting graphs. Long to short dsRNA: rainbow order.	90
4.7	Affinities of reactions for (a) Model-ER, (b) Model-EC, (c) Model-MR and (d) Model-MC by using all concentrations at all lengths of dsRNA.	93
4.8	Affinities of length dependent binding rates in Model-ER, Model-EC, Model-MR and Model-MC for seven lengths.	95

4.9	Experimental data (solid line) and prediction for the product P (dashed line) of (1) Model-ER, (2) Model-EC, (3) Model-MR and (4) Model-MC. Each column presents one concentration of dsRNA: (a) $0.5\mu\text{g/mL}$, (b) $1\mu\text{g/mL}$, (c) $2\mu\text{g/mL}$, (d) $4\mu\text{g/mL}$ and (e) $8\mu\text{g/mL}$. Color code is given in Figure 4.6.	96
5.1	Summary of results of the qualitative analysis carried out on models in Chapter 3. The last column gives time-dependent functions corresponding to predictions for product concentrations when time is large enough ($t > T$) for the different models. Parameters are $b_1 = \frac{K_1}{k_{1m}}$, $b_2 = \frac{k_2}{k_{2m}}$, $b_5 = \frac{k_5}{k_{5m}}$ and $\rho_0 = P(T)$	105
5.2	Predictions of $P(t)$ for longer times: equilibrium values of Model-E, EC, N, ED, M and MC with the best estimated parameters obtained by fitting all lengths.	106
5.3	Summary of conclusions.	107
5.4	The result of fitting for Model-MC in all concentrations and all lengths of dsRNA. Color code is given in Figure 4.6.	110
A.1	Diagram associated to P(2)	115
A.2	ℓ th and $(\ell + 1)$ th bindings	117

Chapter 1

Introduction

In research on enzyme activation, the whole process cannot always be observed. Only some parts of the activation can be examined because of the complexity of mechanisms and technological limitations. However, some possible scenarios can be explored by mathematical modelling to study the process [1]. For this reason, the collaboration between Mathematics and Biochemistry is considered as important [2]. Based on these considerations, this thesis is focused on mathematical and numerical analysis of models for the OAS2 activation by dsRNA to decipher possible scenarios of activation.

1.1 OAS family in innate immune system

The 2'-5'-oligoadenylate synthetase (OAS) is a part of the innate immune response against virus infection [3, 4]. The OAS family has various members, OAS1, OAS2, and OAS3, which consist of one, two, and three OAS domains

respectively [5]. OASs enzymes are ATP polymerases that require double-stranded RNA (dsRNA) as a *cofactor*; ATP is the *substrate*. These enzymes are activated by binding to dsRNA. Among three OAS family members, we only focus on OAS2, which is a dimer consisting of two OAS domains and is activated by dsRNA longer than 35 base pairs (bp) [3, 4].

Once OAS2 binds to dsRNA in the presence of ATPs, OAS2 polymerizes ATPs

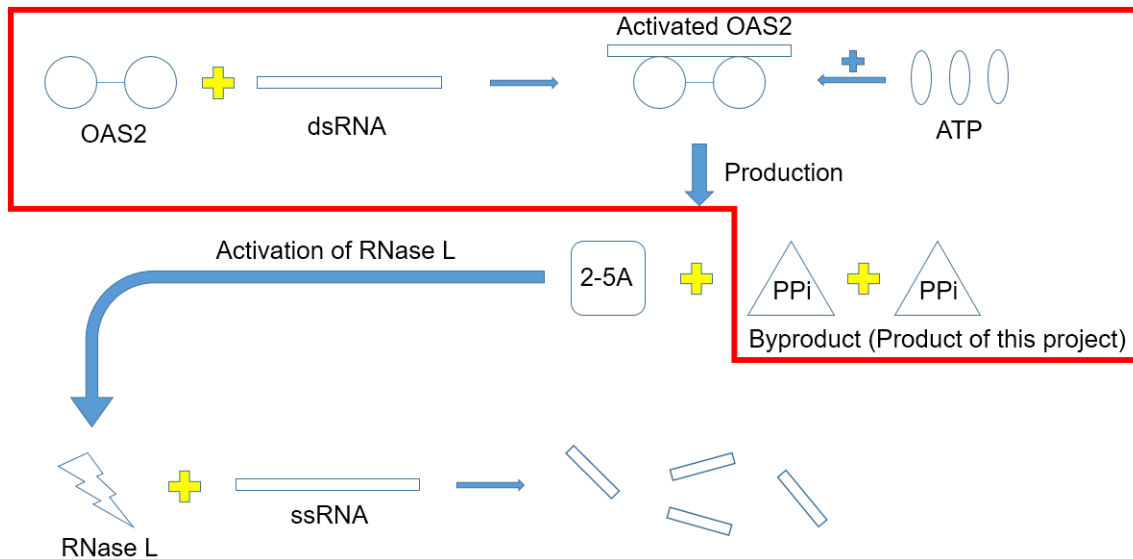


Figure 1.1: The simple diagram for OAS2 activation by dsRNA and the immune response. The focus of this thesis is represented in the red frame in Figure 1.1.

into chains of 2'-5' linked oligoadenylates (2-5A) and pyrophosphate (PPi) is a byproduct of this activation [6]. Then, 2-5A activates another enzyme, RNase L, which degrades viral single-stranded RNA (ssRNA) [7, 8]. The antiviral process driven by OAS2 can be described as in Figure 1.1 [7]. Since OAS enzymes are involved in the regulation of the viral RNA degradation and eventually block

the viral replication in cells, the OAS family is part of the innate immune system.

1.2 Regulation of activation of OAS2

In this thesis, to investigate OAS2 activation mechanism, we design mathematical models based on experimental data measuring the concentration of the byproduct, PPi, over time. Only PPi concentration is experimentally observable instead of the product, 2-5A. PPi is called the product in this work. All experiments are performed and data are provided by McKenna's lab [3].

In contrast to general enzyme kinetics and OAS1 activation, OAS2 activation depends on the length of dsRNA, the cofactor of OAS2 [3]. In general enzyme kinetic, the concentration is the only factor, which affects the production; however, in OAS2 activation, the concentration and length of dsRNA are both factors that regulate the yield of the product [3]. It is experimentally observed that the concentration of the product increases as the concentration or length of dsRNA increases (Figure 1.2) [3]. Therefore, the main goal of this research is to find mathematical models representing the effect of dsRNA concentration and length changes on the activity of OAS2, simultaneously.

Seven sets, which are categorized by lengths of dsRNA (40bp, 50bp, 60bp, 70bp, 80bp, 90bp and 120bp), are considered to investigate the effect of dsRNA length in this work. The concentration of the product in each set is recorded at different time points (0min, 5min, 10min, 15min, 20min and 30min) and

for five concentrations of dsRNA ($0.5\mu\text{g}/\text{mL}$, $1\mu\text{g}/\text{mL}$, $2\mu\text{g}/\text{mL}$, $4\mu\text{g}/\text{mL}$ and $8\mu\text{g}/\text{mL}$). In these *in vitro* experiments, ATP, the substrate for OAS2 enzyme, and dsRNA, the cofactor of OAS2 enzyme, are in abundance. Experimental data is given in Figure 1.2 [3].

1.3 Organization of thesis

In Chapter 2, we assume chemical reactions based on the theory of OAS2 activation by dsRNA [3, 4]. Ordinary differential equation (ODE) systems for nine different models with different assumptions to accommodate changes in concentrations and lengths of dsRNA are proposed. In Chapter 3, all the models introduced in Chapter 2 are mathematically analyzed to conjecture the amount of the product based on the models' ODE system. In Chapter 4, all the models are compared to experimental data from McKenna's lab to estimate the parameter values. The dissociation constants or affinities, $\frac{k_{off}}{k_{on}}$, of reactions considered are also obtained. Furthermore, the best model is selected by model selection methods in the last part of Chapter 4. Finally, Chapter 5 gives an overview of the work and some ideas for future studies.

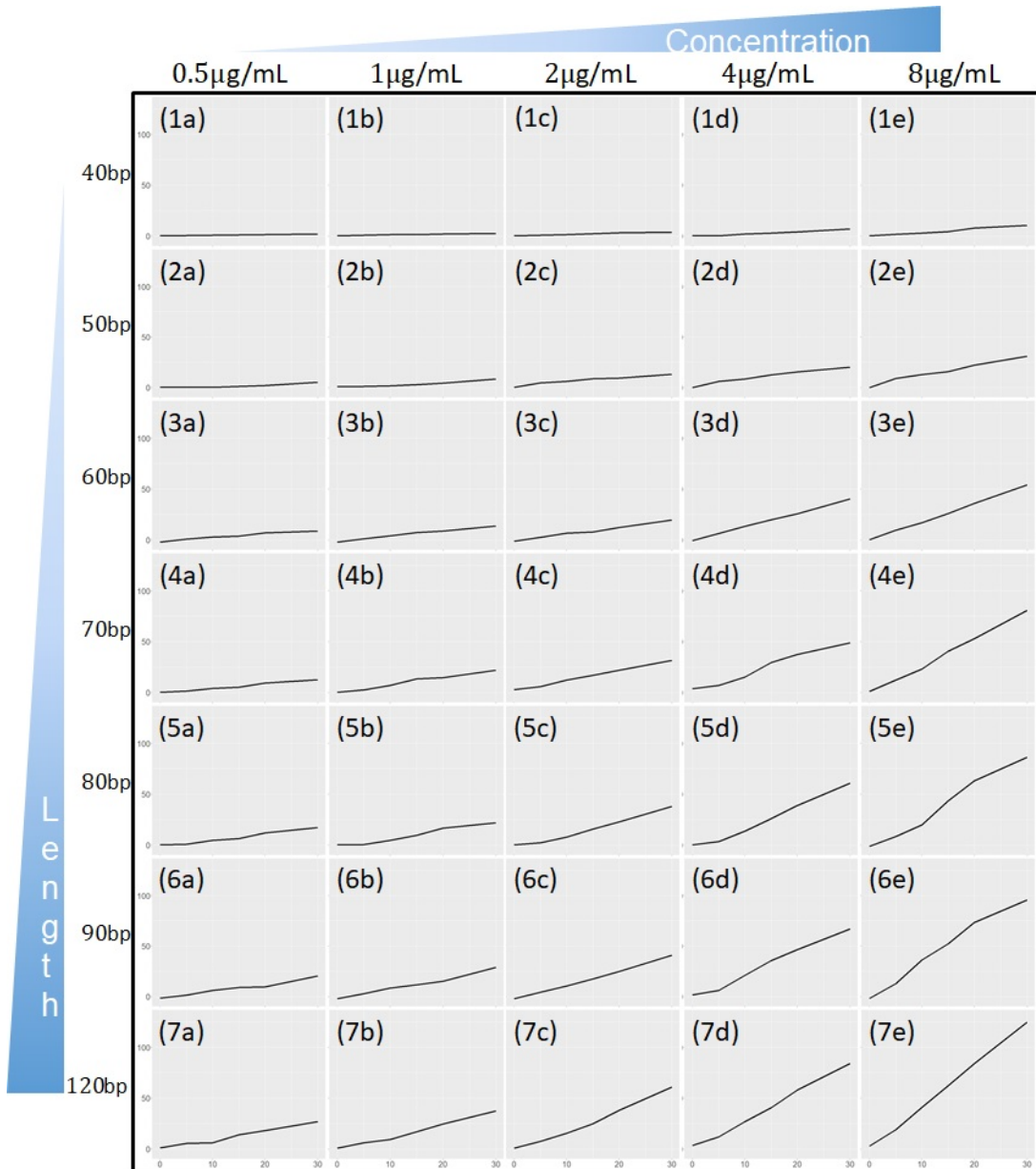


Figure 1.2: Concentrations of product, PPi, for each concentration and length of dsRNA over time. Number codes for the length: (1) 40bp, (2) 50bp, (3) 60bp, (4) 70bp, (5) 80bp, (6) 90bp and (7) 120bp. Letters represent concentrations: (a) 0.5 µg/mL, (b) 1 µg/mL, (c) 2 µg/mL, (d) 4 µg/mL and (e) 8 µg/mL. The range of y-axis (the concentration of the product) is from 0 to 130 µg/mL and the range of x-axis (time) is from 0 to 30 minutes.

Chapter 2

Mathematical Modelling

OAS2 is an enzyme, which plays a role in the innate immune system to prevent from the viral infection; its activation is induced by the binding of viral dsRNA as a cofactor. This chapter discusses nine different mathematical models of OAS2 activation by dsRNA, which are divided into four groups. Each group of models is defined by the dimension of ordinary differential equation systems. In the first group, the model has the simplest structure. This model assumes fixed concentrations for OAS2 and dsRNA. This group, group 1, is composed of only one model, Model-S. In the second group, models are made in consideration of the enzyme kinetics between OAS2 and dsRNA. These models are Model-E, Model-ER and Model-EC. Model-E is the base model of group 2, Model-ER and Model-EC are obtained by considering Model-E with two different kinds of length dependency on the binding rate. Moreover, the degradation of OAS2 or the complex formed by the binding of OAS2 and dsRNA is considered. Models with degradation are called Model-ED; three versions are considered Model-

ED1, Model-ED2 and Model-ED3 depending on which degradation reactions are added to Model-E. In the third group, a model assumes that OAS2 and dsRNA can form a non-productive complex. Two OAS2 domains of OAS2 need to bind to the same dsRNA, there could exist non-productive bindings for which only one domain of OAS2 binds properly to dsRNA. Group 3 has one model, Model-N. The last group, group 4, considers multi-binding models. With enough length of dsRNA, there is a possibility that multiple OAS2 molecules bind to one long dsRNA. This group has three models: Model-M, Model-MR and Model-MC.

Before giving full details of the models in next sections, the diagrams and variables used in equations are given in Figure 2.1 and the parameters of models are explained in Figure 2.2.







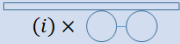
Diagram	Variables	Name
	E	OAS2 Enzyme
	R	dsRNA
	$D (= D_1)$	Binding Stage (Activated OAS2)
	C	Pre-binding Stage
	N	Non-productive binding Stages
	P	Product (PPi)
	D_i	Multi-binding Stages (i activated OAS2)

Figure 2.1: Compounds considered in models and their corresponding variables. Note that D_i refer to a complex formed of a dsRNA bound to i OAS2.

Parameters	Explanation	Models
K_1	Binding rate of OAS2 and dsRNA	Model-E, ED, M
	Binding rate of a pre-binding stage	Model-N
k_1	Constant part of binding rate of OAS2 and dsRNA	Model- ER, EC, MR, MC
k_{1m}	Unbinding rate of a complex of OAS2 and dsRNA	Model-E, ER, EC, ED, M, MR, MC
	Unbinding rate of a pre-binding stage	Model-N
k_2	Binding rate of OAS2 and dsRNA from a pre-binding stage	Model-N
k_{2m}	Unbinding rate of complex of OAS2 and dsRNA to a pre-binding stage	Model-N
k_5	Binding rate of a non-productive binding stages	Model-N
k_{5m}	Unbinding rate of a non-productive binding stages	Model-N
k_{d1}	Degradation rate of OAS2	Model-ED
k_{d2}	Degradation rate of a complex of OAS2 and dsRNA	Model-ED
k_{ib} $i \in \{2,3,\dots,n\}$	Binding rate of OAS2 and D_{i-1} to D_i	Model-M, MR, MC
k_{ibm} $i \in \{2,3,\dots,n\}$	Unbinding rate of D_i to D_{i-1}	Model-M, MR, MC
k_p	Production rate of one OAS2 and dsRNA	All models

Figure 2.2: List of parameters used in models.

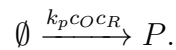
2.1 Group 1 - Model-S

The first group has only one model, Model-S. In general enzyme kinetics, there are binding and catalytic stages when a substrate is changed into a product. However, in OAS2-dsRNA activation, the enzyme (OAS2) attaches to the cofactor (dsRNA) and makes the product (PPi) from the substrate (ATP) while enzyme and cofactor are not consumed. Recall that in the *in vitro* experiments considered, the substrate is in abundance; hence, the dynamics of ATP is not explicitly represented in models. Model-S for the activation reaction of OAS2 by dsRNA is represented in the diagram shown in Figure 2.3; as the concentrations of enzyme and cofactor are considered as not varying over time, a zero order reaction is first assumed for Model-S.



Figure 2.3: Model-S

The model diagram of Model-S can be written as the following chemical equation:



This chemical equation leads to the following ordinary differential equation:

$$\frac{dP}{dt} = k_p c_O c_R, \quad (2.1)$$

where $c_O = E(t)$ and $c_R = R(t)$ for all $t \geq 0$ are the concentrations of OAS2 and dsRNA, respectively. Variables and parameters are listed in Figures 2.1 and 2.2.

2.2 Group 2

2.2.1 Model-E: constant rate

Model-E is the base model of group 2 in which an enzyme kinetics is applied. Once bound to a dsRNA, the OAS2 enzyme is activated and makes the product from the substrate that is in abundance in the *in vitro* experiments considered. Hence, as previously mentioned, the dynamics of the substrate is not of interest in this work; no equation is used to describe ATP in models considered. Furthermore, the binding and unbinding of enzymes and cofactors are reversible reactions; neither enzyme nor cofactor are consumed during these reactions nor the formation of product. Hence, the complex D , which is the activated enzyme or productive stage, is not consumed in the formation of product. The model diagram for Model-E is as in Figure 2.4.

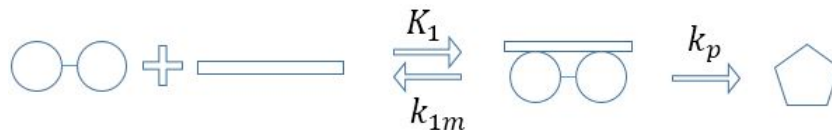
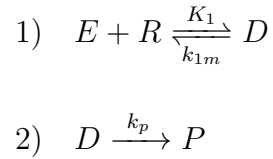


Figure 2.4: Model-E

Based on this diagram, the chemical equations are



and the system of ODE is stated as follows:

$$\begin{aligned} \frac{dE}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D, \\ \frac{dP}{dt} &= k_pD. \end{aligned} \tag{2.2}$$

Variables and parameters are listed in Figures 2.1 and 2.2.

2.2.2 Model-ER and Model-EC: length dependent rates

This section introduces models that are extensions of Model-E. The experimental data of the OAS2 activation by dsRNA shows that the amount of the product increases as the length of the dsRNA increases (Figure 1.2). Models with a rate that depends on the length of dsRNA are developed from Model-E. There are three potential rates, which can be replaced by a length dependent binding rate (Figure 2.4): K_1 , k_{1m} , k_p . Among these three rates, K_1 turned out to be the most reasonably possible candidate for a length dependent binding

rate for several reasons. First, in experiments (Figure 1.2), the longer dsRNA is, the larger amount of the product is made, which means that the longer dsRNA has the greater possibility of interacting with OAS2. Therefore, it is reasonable for the length dependency to be applied to K_1 . Furthermore, for k_p , it is possible to assume that the amount of the product that can be produced by one OAS2 is fixed, so k_p is expected to be a constant rate. Similarly, in the case of k_{1m} , it is also possible to assume that once OAS2 combines with dsRNA, the unbinding rate of complex (binding stage) does not depend on the dsRNA length. For these reasons, the length dependent binding rate is chosen to be K_1 (Figure 2.5).

To derive the length dependent binding rate, it is important to take into account the shape of OAS2 and dsRNA [9]. For instance, several studies investigated the effect of length of interacting objects on their association rates [10, 11] or of tethered ligands on enzymatic reactions [12, 13]. In this work two different length dependent binding rates, which are based on the rigid properties of dsRNA, are considered, $k_1\sigma_1(L)$ and $k_1\sigma_2(L)$, where L is the dsRNA length.

To replace the constant binding rate with a length dependent binding rate, it

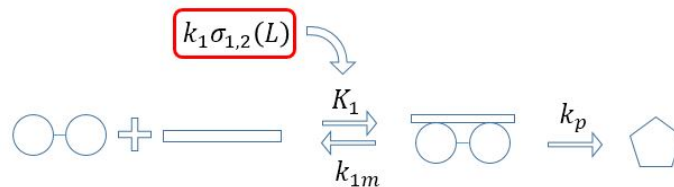


Figure 2.5: Binding rate depending on cofactor length L . Replace K_1 in Model-E by $k_1\sigma_{1,2}(L)$, where $\sigma_{1,2}(L)$ is $\sigma_1(L)$ for Model-ER or $\sigma_2(L)$ for Model-EC.

is important to consider a collision rate and a binding probability [14]. The binding rate is the product of the collision rate and binding probability when a collision occurs. In the next section, the collision rate is calculated based on the shape of OAS2 and dsRNA. After that, two binding probabilities are proposed based on two different assumptions. Thus, the two binding rates share the same collision rate and each of them uses its own binding probability.

Collision rate

To calculate the collision rate of OAS2 and dsRNA, we simplify the shape of OAS2 and dsRNA as in Figure 2.6. OAS2 is described as a sphere. On the other hand, dsRNA is considered as a long cylinder.

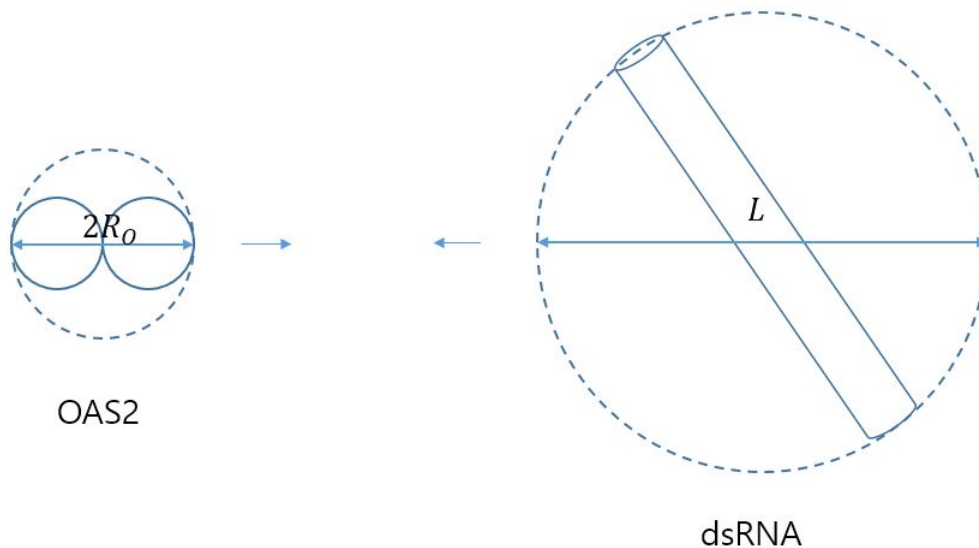


Figure 2.6: Dashed circles delimit the interaction zones for OAS2 and dsRNA. L is length of dsRNA, R_O is the radius of OAS2.

Molecules of OAS2 and dsRNA diffuse and rotate freely in solution, so their effective interaction zone can be represented as the two spheres inferred by dashed circles in Figure 2.6. The radii of interaction spheres are R_R for dsRNA and R_O for OAS2. Furthermore, OAS2 and dsRNA diffuse freely; the diffusion coefficient for dsRNA is D_R and D_O for OAS2. Consequently, the collision rate of dsRNA and OAS2 [9] can be expressed as

$$\text{Collision rate} = 4\pi(D_R + D_O)(R_R + R_O), \quad (2.3)$$

where R_R is the half length of dsRNA, $R_R = \frac{L}{2}$, with L the length of dsRNA, and R_O is the radius of OAS2. The diffusion coefficient of dsRNA is approximated following the work of [15, 16]:

$$D_R = \frac{K_B T}{3\pi\eta L} \left[\log \left(2x - 1 + \sqrt{(2x - 1)^2 + 1} \right) + \frac{1}{x} \left(x + \sqrt{2} - \sqrt{(2x - 1)^2 + 1} \right) + \frac{1}{2x} \log \left(\frac{(\sqrt{2} - 1)^2 (1 + \sqrt{(2x - 1)^2 + 1})}{(2x - 1 + \sqrt{(2x - 1)^2 + 1})} \right) \right], \quad (2.4)$$

where $x = \frac{L}{d}$, d is the diameter of the dsRNA, K_B is Boltzmann constant, T is a absolute temperature in Kelvin, and η is the coefficient of viscosity. The diffusion coefficient for OAS2 is obtained based on Stokes-Einstein equation [9]:

$$D_O = \frac{K_B T}{3\pi\eta} \left(\frac{1}{2R_O} \right).$$

To make (2.4) easier to read, let $A = 2x - 1$, $B = \sqrt{(A^2 + 1)}$ and $C_0 = \frac{K_B T}{3\pi\eta}$, then we get the equation,

$$D_R = \frac{C_0}{L} \left(\log(A + B) + \frac{1}{x}(x + \sqrt{2} - B) + \frac{1}{2x} \log \left(\frac{(\sqrt{2} - 1)^2(1 + B)}{(A + B)} \right) \right).$$

Furthermore, let $\left(\log(A + B) + \frac{1}{x}(x + \sqrt{2} - B) + \frac{1}{2x} \log \left(\frac{(\sqrt{2} - 1)^2(1 + B)}{(A + B)} \right) \right) = Z$, then diffusion coefficients of dsRNA and OAS2 can be rewritten as follows:

$$D_R = \frac{C_0 Z}{L}, \quad D_O = \frac{C_0}{2R_O}, \quad (2.5)$$

where only $\frac{Z}{L}$ depends on dsRNA lengths. Substituting (2.5) into the collision rate equation (2.3), we get the collision rate between dsRNA and OAS2 [17] as below.

$$4\pi C_0 \left(\frac{Z}{L} + \frac{1}{2R_O} \right) \left(\frac{L + 2R_O}{2} \right). \quad (2.6)$$

Binding Probability

The binding probability depends on the assumption how OAS2 binds to dsRNA. Because two OAS domains need to bind to the same dsRNA, if one domain of OAS2 binds too close to the tip of dsRNA, then the other domain cannot bind, there is no formation of product. Based on this idea, it can be induced that there is the valid binding area on dsRNA for OAS2 binding to result in the OAS2 activation. Because the valid binding area is determined by the shape of

dsRNA, the binding probability also depends on the shape of dsRNA. We discuss two assumptions for the shape of dsRNA in this section. The first assumption is that dsRNA is a rigid rod, which can bind to OAS2 on one side, in two dimensional space, and the second one is that dsRNA is a cylinder in three dimensional space. Consequently, with those binding probabilities, we get two binding rates of OAS2 and dsRNA.

Binding rate: dsRNA as a rigid rod

First, a dsRNA is described as a rigid rod (Figure 2.7). Since OAS2 is a dimer

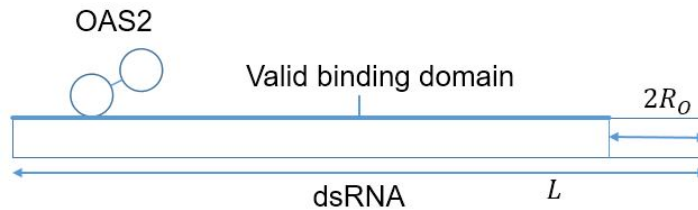


Figure 2.7: Valid binding domain when dsRNA is represented as a rigid rod.

having a non-negligible length, there is a valid domain of binding on dsRNA for OAS2 to bind. When one of the OAS2 domain binds to dsRNA's valid binding domain, it needs enough space for the second OAS2 domain to bind in the valid binding domain. Therefore, with L , the length of dsRNA, $L - 2R_O$ is the valid length and it brings the following probability for both domains of OAS2 to bind on the valid domain (in the sense of uniform distribution):

$$\frac{L - 2R_O}{L}.$$

Then, the probability of binding given a collision occurs is

$$k_0 \left(\frac{L - 2R_O}{L} \right), \quad (2.7)$$

where the factor k_0 is a positive constant accounting for, for instance, electrostatic interactions or other types of forces that can modulate the binding.

Since the binding rate is the multiplication of collision rate (2.6) and binding probability (2.7), the binding rate of OAS2 with a rigid rod-shaped dsRNA is described as follows:

$$\begin{aligned} k_1 \sigma_1(L) &= 4\pi C_0 \left(\frac{Z}{L} + \frac{1}{2R_O} \right) \left(\frac{L + 2R_O}{2} \right) k_0 \left(\frac{L - 2R_O}{L} \right) \\ &= k_1 \underbrace{\left(\frac{Z}{L} + \frac{1}{2R_O} \right) \left(\frac{L + 2R_O}{2} \right) \left(\frac{L - 2R_O}{L} \right)}_{\sigma_1(L)}, \end{aligned} \quad (2.8)$$

where $k_1 = 4\pi C_0 k_0$. Note that d and R_O are constant, $\sigma_1(L)$ only depends on the length of dsRNA.

Binding rate: dsRNA as a cylinder

The second assumption is that dsRNA is a cylinder in three dimensional space (Figure 2.8). The surface of dsRNA consists of a rod part and two cap parts, so it is $2\pi \left(\frac{d}{2}\right)^2 + d\pi L$. The valid binding area is only on the rod part, but we do not know how much area is allowed for binding. Thus, let the percentage of the valid binding area be k and then the valid binding area can be written

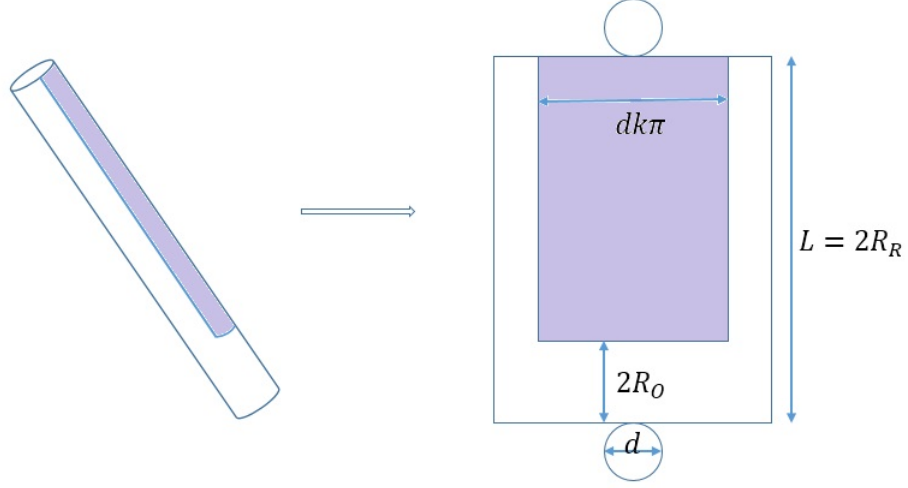


Figure 2.8: dsRNA is described as a cylinder of length L and diameter d . R_O : radius of OAS2, k : ratio for valid binding area.

as $kd\pi(L - 2R_O)$. Consequently, assuming the uniform distribution of binding sites, the probability of binding is

$$k_0 \left(\frac{kd\pi(L - 2R_O)}{2\pi \left(\frac{d}{2}\right)^2 + d\pi L} \right) = k_0 \left(\frac{k(L - 2R_O)}{\frac{d}{2} + L} \right),$$

where the parameter k_0 represents the modulation of the binding from other factors.

Then, the binding rate with the cylindrical-shaped dsRNA is described as

follows:

$$\begin{aligned}
 k_1\sigma_2(L) &= 4\pi C_0 \left(\frac{Z}{L} + \frac{1}{2R_O} \right) \left(\frac{L + 2R_O}{2} \right) k_0 \left(\frac{k(L - 2R_O)}{\frac{d}{2} + L} \right) \\
 &= k_1 \underbrace{\left(\frac{Z}{L} + \frac{1}{2R_O} \right) \left(\frac{L + 2R_O}{2} \right) \left(\frac{L - 2R_O}{\frac{d}{2} + L} \right)}_{\sigma_2(L)}, \tag{2.9}
 \end{aligned}$$

where $k_1 = 4\pi C_0 k_0 k$.

In (2.8) and (2.9), $\sigma_1(L)$ and $\sigma_2(L)$ are functions of L and k_1 is the constant part of the binding rates. In both cases, k_1 is unknown and will be estimated by comparing models' responses to experimental data in Chapter 4.

Model-ER and Model-EC

Model-ER is obtained by considering Model-E with $K_1 = k_1\sigma_1(L)$ (2.8) and Model-EC uses $K_1 = k_1\sigma_2(L)$ (2.9) (Figure 2.5). For simplification, in both length dependent binding rates, the constant part is called k_1 , but they are different values as previously explained.

Chemical reactions considered in Model-ER and Model-EC are as in Model-E with $K_1 = k_1\sigma_1(L)$ and $K_1 = k_1\sigma_2(L)$, respectively (Figure 2.5). Similarly, the ODE system for Model-ER and Model-EC is system (2.2) considered with $K_1 = k_1\sigma_1(L)$ and $K_1 = k_1\sigma_2(L)$, respectively.

2.2.3 Model-ED

The activity of OAS2 can decrease after several reactions because of the limit of durability of enzymes. In any case, it can be interpreted that the active enzyme is degraded at a rate constant k_{d1} (Figure 2.9) and it leads to the reduced rate of product formation. Moreover, after some activation reactions, it can be assumed that the complex of OAS2 and dsRNA is degraded at a rate constant k_{d2} (Figure 2.9). Model-E considered with the degradation reaction(s) is named Model-ED (Figure 2.9).

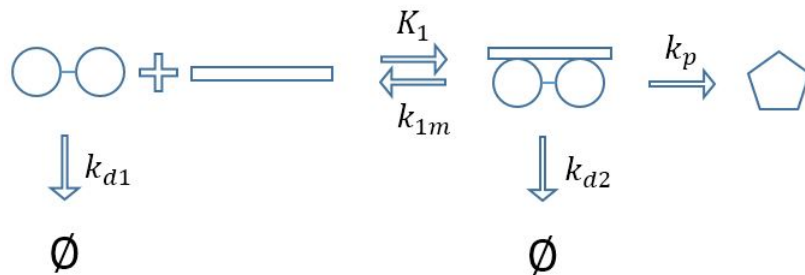
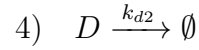
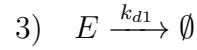
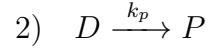
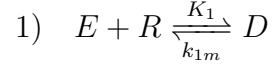


Figure 2.9: Model-ED with degradation of OAS2 or complex, \emptyset : degradation

The case $k_{d1} = k_{d2} = 0$ represents Model-E, and there are three versions of Model-ED: Model-ED1 ($k_{d1} \neq 0, k_{d2} = 0$), Model-ED2 ($k_{d1} = 0, k_{d2} \neq 0$) and Model-ED3 ($k_{d1} \neq 0, k_{d2} \neq 0$). Here, the general model for Model-ED1,

Model-ED2 and Model-ED3 assumes the following chemical reactions.



The system of ODE with degradation of E and D is

$$\begin{aligned} \frac{dE}{dt} &= -K_1ER + k_{1m}D - k_{d1}E, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D - k_{d2}D, \\ \frac{dP}{dt} &= k_pD. \end{aligned} \tag{2.10}$$

2.3 Group 3 - Model-N

Another way to represent the impact of the dsRNA length on OAS2 activation is to consider that some interactions are not complete or invalid [18]. Model-N is established considering that the binding of only one domain of OAS2 to dsRNA cannot cause OAS2 activation. Based on this assumption, we add two stages to Model-E. First, a pre-binding stage (C) of the productive binding stage is added

to represent a stage, which is incomplete with one OAS2 domain binds properly. Secondly, the invalid bindings are added and called non-productive binding stages as it cannot result in the formation of the product. Although there are many possible scenarios for the non-productive binding, we cannot specify all the cases, so we collectively refer to them as stage N . After a non-productive binding, OAS2 can detach from dsRNA and attach to dsRNA again.

A two-step binding process and non-productive binding stages are added to

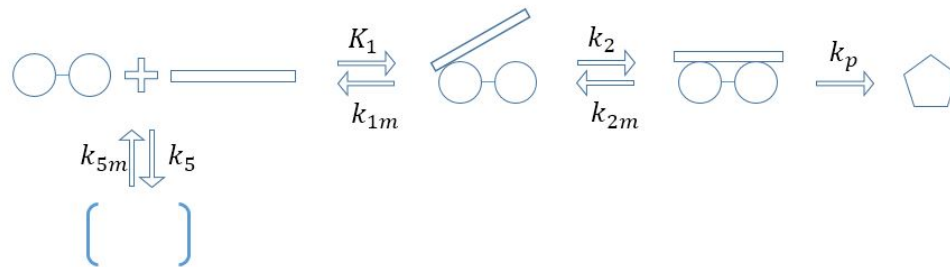
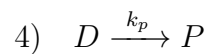
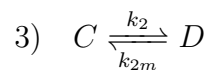
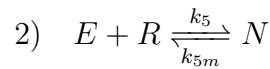
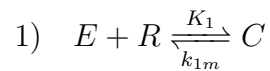


Figure 2.10: Model-N with a two-step binding and non-productive bindings.

Model-E that leads to Model-N (Figure 2.10), the chemical equation set is



and the system of ODE for Model-N is constructed as follows:

$$\begin{aligned}
\frac{dE}{dt} &= -K_1ER + k_{1m}C - k_5ER + k_{5m}N, \\
\frac{dR}{dt} &= -K_1ER + k_{1m}C - k_5ER + k_{5m}N, \\
\frac{dC}{dt} &= K_1ER - k_{1m}C - k_2C + k_{2m}D, \\
\frac{dD}{dt} &= k_2C - k_{2m}D, \\
\frac{dN}{dt} &= k_5ER - k_{5m}N, \\
\frac{dP}{dt} &= k_pD.
\end{aligned} \tag{2.11}$$

2.4 Group 4

From experiments [3], there are two significant observations that drive the derivation of another model. The first observation is that dsRNA is not consumed and changed by the OAS2 activation. OAS2 interacts with dsRNA, facilitates product formation, and then detaches from dsRNA. The second observation is that long dsRNA used in experiments makes more products than short dsRNA with the same amount of OAS2 (Figure 1.2). Hence, another reasonable scenario is that dsRNA having more than double the length of OAS2 can be bound with more than two OAS2 at the same time, leading to increased product formation. We focus on this scenario, the multi-binding assumption,

- 1) $E + R \xrightleftharpoons[k_{1m}]{K_1} D$
- 2) $E + D_{i-1} \xrightleftharpoons[k_{ibm}]{k_{ib}} D_i, \quad i \in \{2, \dots, n\}$
- 3) $D_i \xrightarrow{k_p} P,$

and the system of ODE is

$$\begin{aligned}
\frac{dE}{dt} &= -K_1ER + k_{1m}D_1 + \sum_{i=1}^{n-1} [k_{(i+1)bm}D_{i+1} - k_{(i+1)b}D_iE], \\
\frac{dR}{dt} &= -K_1ER + k_{1m}D_1, \\
\frac{dD_1}{dt} &= K_1ER - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E, \\
&\vdots \\
\frac{dD_j}{dt} &= k_{jb}D_{j-1}E - k_{jbm}D_j + k_{(j+1)bm}D_{j+1} - k_{(j+1)b}D_jE, \tag{2.12} \\
j &\in \{2, 3, \dots, n-1\}, \\
&\vdots \\
\frac{dD_n}{dt} &= k_{nb}D_{n-1}E - k_{nbm}D_n, \\
\frac{dP}{dt} &= k_p \sum_{i=1}^n iD_i.
\end{aligned}$$

The model equation for Model-M allowing up to n ($= \lfloor \frac{L}{2R_0} \rfloor$) OAS2 per dsRNA is validated by mathematical induction; details can be found in Appendix.

2.4.2 Model-MR and Model-MC: length dependent rates

For the binding of one OAS2 molecule and dsRNA, constant or length dependent binding rates can be used. However, since the length dependent binding rates are calculated for one OAS2 binding an “unoccupied” dsRNA, these rates can be only used for the binding of the first OAS2 and cannot be applied to the binding of a second OAS2 to the same dsRNA. In the previous section, we derived two length dependent binding rates (2.8 and 2.9) that can be used for Model-M because the principle of the first binding between OAS2 and dsRNA is not changed. Thus, we discuss two variations of Model-M, Model-MR and Model-MC, which have a length dependent binding rate for the first binding. We define Model-MR as Model-M with $K_1 = k_1\sigma_1(L)$ and Model-MC as Model-M with $K_1 = k_1\sigma_2(L)$.

Chemical reactions considered in Model-MR and Model-MC are as in Model-M with in the first reaction rate $K_1 = k_1\sigma_1(L)$ and $K_1 = k_1\sigma_2(L)$, respectively. Similarly, the ODE system for Model-MR and Model-MC is system (2.12) considered with $K_1 = k_1\sigma_1(L)$ and $K_1 = k_1\sigma_2(L)$, respectively.

2.5 Overview

The main motivation of this work is finding a model which can reflect the effects of changes in concentrations and lengths of dsRNA on OAS2 activation. The change of dsRNA concentration is accommodated in general enzyme kinetics,

but the change of dsRNA length is not reflected in the kinetics. There are two ways to accommodate the change of length of dsRNA.

The first way is to reduce or extend the common enzyme kinetic model, which is Model-E, the base model of all groups. Reducing Model-E gives Model-S, in which the concentrations of OAS2 and dsRNA are assumed to stay constant over time. Expanding the structure of Model-E following two different assumptions leads to Model-N and Model-M. In the Model-N, the existence of reversible non-productive binding stages is considered. In Model-M, it is assumed that long dsRNA can be bound with more than one OAS2 at the same time.

The second method to accommodate the effect of dsRNA's length is to apply the effect of the length directly to the binding parameter without changing the structure of models. Two length dependent binding rates are proposed. Model-ER and Model-EC are the extended versions of Model-E, and similarly, Model-MR and Model-MC are the extensions of Model-M. The R-type (Model-ER and Model-MR) uses the length dependent binding rate based on the rigid rod representation of dsRNA (2.8) and the C-type (Model-EC and Model-MC) is based on the cylindrical representation of dsRNA (2.9). Finally, the degradation of OAS2 or complex of OAS2 and dsRNA are also considered in Model-ED.

The relations among all the models considered in this work are shown in Figure 2.12.

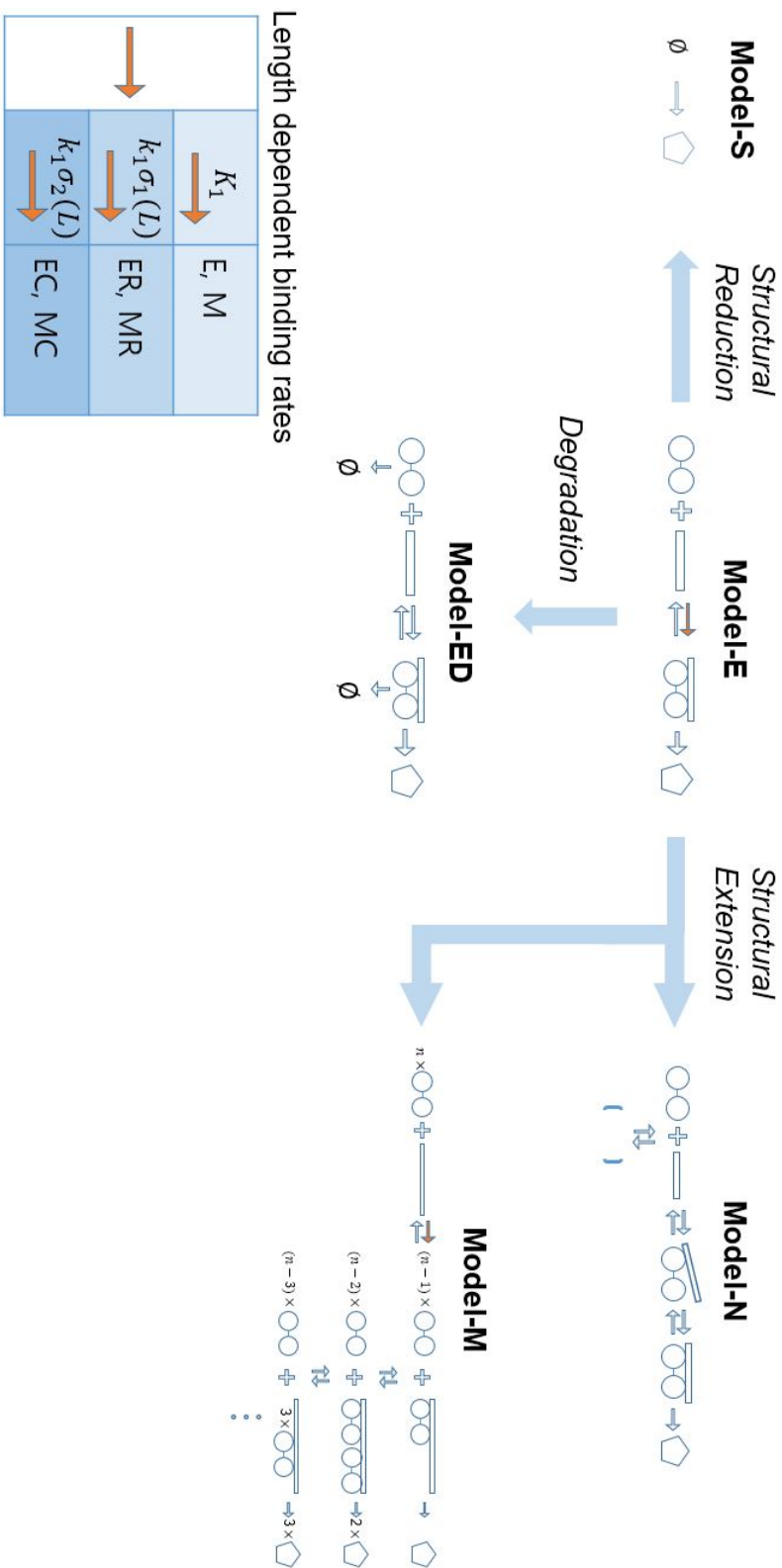


Figure 2.12: Tree of models: the relations among the models.

Chapter 3

Mathematical Analysis

In this chapter, the mathematical analysis of models introduced in Chapter 2 is carried out. In Chapter 2, models are categorized by groups based on the dimension of ODE systems. For each group, a general model can be defined; each model of the group is a particular case of the general model. Hence, the analysis is carried out on the general system, the representative of the group. First, the simplest model of group 1 is solved. Then, because the ODE systems of all the models in group 2 to 4 are composed of non-linear and autonomous equations of same type, their well-posedness is commonly dealt. The equilibria of general models and their nature for each group are investigated and the amount of product is predicted. In all the analyses, c_O and c_R are referred as the initial concentrations of OAS2 and dsRNA, respectively.

3.1 Group 1 - Model-S

In Model-S, it is assumed that there are no changes in the concentrations of OAS2 and dsRNA, so $E(t) = c_O$ and $R(t) = c_R$ for all t . Recall the ODE equation of Model-S, (2.1):

$$\frac{dP}{dt} = k_p c_O c_R, \quad (3.1)$$

with k_p , c_O , c_R positive constants.

Theorem 3.1.1. *In Model-S (3.1), the product increases linearly as t increases, $P(t) = k_p c_O c_R t$, for all $t \geq 0$.*

Proof. Integrate (3.1) with respect to t with $P(0) = 0$, then we get the equation,

$$P(t) = k_p c_O c_R t. \quad (3.2)$$

Therefore, in Model-S, P is only proportional to t and the amount of the product increases linearly in time. \square

3.2 Wellposedness - Group 2 to 4

All models considered in group 2 to 4 take the following generic form:

$$\frac{dX}{dt} = f(X), \quad (3.3)$$

where the state variables $X = (E, R, D, D_j, C, N) \in \mathbb{R}^m$ with $j = \{1, 2, 3, \dots, n\}$ and $m = \{3, 4, \dots, n + 4\}$ represent the concentrations of chemical compounds.

For all models of group 2 to 4, note that P is not involved in any equation of systems and the equation $\frac{dP}{dt}$ only depends on D or D_j ; hence, $\frac{dP}{dt}$ is decoupled from all the systems.

The vector field is a vector function $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $f_i(X)$ are polynomials of several variables with $i \in \{1, 2, \dots, m\}$. Hence, all models are nonlinear and autonomous and considered with the initial condition, $X(0) = X_0 = (c_O, c_R, 0, 0, 0, 0) \in \mathbb{R}_{\geq 0}^m$.

Theorem 3.2.1. *System (3.3) considered with $X_0 \in \mathbb{R}_{\geq 0}^m$ is well-posed.*

Proof. An initial value problem is defined by considering system (3.3) with the initial condition, $X_0 \in \mathbb{R}_{\geq 0}^m$. If $X_i = 0$, then $\frac{dX_i}{dt}|_{X_i=0} \geq 0$; solutions $X_i(t)$ stay non-negative, for $i \in \{1, 2, \dots, m\}$.

Since $f_i(X)$ are polynomials on \mathbb{R}^m , they are on C^∞ on \mathbb{R}^m . Therefore, the solution of an initial value problem exists and is unique for $t \geq 0$ by Existence and Uniqueness Theorem for ODE [19]. Thus, system (3.3) is well-posed. \square

Since the equations of $\frac{dP}{dt}$ only depend on D or D_j and the solution of (3.3) exists and is unique, $P(t)$ is also determined uniquely for all models.

3.3 Group 2

Model-ED, (2.10), is the general model of group 2. It is reduced to:

$$\begin{aligned}\frac{dE}{dt} &= -K_1ER + k_{1m}D - k_{d1}E, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D - k_{d2}D.\end{aligned}\tag{3.4}$$

Model-ED1 is obtained with $k_{d1} \neq 0, k_{d2} = 0$, Model-ED2 with $k_{d1} = 0, k_{d2} \neq 0$ and Model-ED3 with $k_{d1} \neq 0, k_{d2} \neq 0$. Setting $k_{d1} = k_{d2} = 0$ gives the three versions of Model-E with no degradation: Model-E, Model-ER and Model-EC. Note that the analysis of the three models can be carried out all at once because these models only differ from their dependency of binding rate K_1 on dsRNA lengths, which are only parameters for the systems.

1) Model-E

Model-E is obtained by setting $k_{d1} = 0$ and $k_{d2} = 0$ in (3.4).

$$\begin{aligned}\frac{dE}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D.\end{aligned}\tag{3.5}$$

From (3.5), $\frac{dE}{dt} + \frac{dD}{dt} = 0$ and $\frac{dR}{dt} + \frac{dD}{dt} = 0$; so by integrating with respect to t we obtain two equations, $E(t) + D(t) = c_O$ and $R(t) + D(t) = c_R$, for all $t \geq 0$. These two equations are referred as conservation equations of Model-E. Then, using conservation equations to describe $E(t) = c_O - D(t)$ and $R(t) = c_R - D(t)$, system (3.5) can be reduced to one equation for D , as follows,

$$\frac{dD}{dt} = K_1 D^2 - (K_1(c_O + c_R) + k_{1m})D + K_1 c_O c_R. \quad (3.6)$$

Furthermore, the conservation equations lead to $D(t) \leq \min(c_O, c_R)$, for all $t \geq 0$.

Theorem 3.3.1. *The unique positive equilibrium of system (3.5) is*

$$\begin{aligned} & (E^*, R^*, D^*) \\ &= \left(c_O - \frac{K_1(c_O + c_R) + k_{1m} - \sqrt{(K_1(c_O + c_R) + k_{1m})^2 - 4K_1^2 c_O c_R}}{2K_1}, \right. \\ & \quad c_R - \frac{K_1(c_O + c_R) + k_{1m} - \sqrt{(K_1(c_O + c_R) + k_{1m})^2 - 4K_1^2 c_O c_R}}{2K_1}, \\ & \quad \left. \frac{K_1(c_O + c_R) + k_{1m} - \sqrt{(K_1(c_O + c_R) + k_{1m})^2 - 4K_1^2 c_O c_R}}{2K_1} \right). \end{aligned} \quad (3.7)$$

and the equilibrium is globally asymptotically stable.

Proof. Since (3.6) is the reduced system of (3.5), we find an equilibrium of (3.6) first, and applying the conservation equations of Model-E to find an equilibrium

of (3.5).

To find the steady state of the model (3.6), let $\frac{dD}{dt} = 0$, then

$$0 = K_1 D^{*2} - (K_1(c_O + c_R) + k_{1m})D^* + K_1 c_O c_R.$$

Let $g = (K_1(c_O + c_R) + k_{1m}) > 0$ and $h = K_1 c_O c_R > 0$ then we get

$$0 = K_1 D^{*2} - gD^* + h. \quad (3.8)$$

Thus, the two solutions are

$$D_1 = \frac{g - \sqrt{g^2 - 4K_1 h}}{2K_1}, \quad D_2 = \frac{g + \sqrt{g^2 - 4K_1 h}}{2K_1}.$$

Note that the solutions have to satisfy the following properties: real, positive and based on conservation laws:

$$D^* \leq \min(c_O, c_R). \quad (3.9)$$

To determine that D_1 and D_2 are real solutions, we consider the discriminant:

$$\begin{aligned} g^2 - 4K_1 h &= (K_1(c_O + c_R) + k_{1m})^2 - 4K_1^2 c_O c_R \\ &= (K_1(c_O + c_R))^2 + 2(c_O + c_R)K_1 k_{1m} + (k_{1m})^2 - 4K_1^2 c_O c_R \\ &= (K_1(c_O - c_R))^2 + 2(c_O + c_R)K_1 k_{1m} + (k_{1m})^2 > 0. \end{aligned}$$

Since c_O , c_R , K_1 and k_{1m} are positive, both D_1 and D_2 are real solutions.

Consider $D^* = D_2$.

$$D^* = D_2 = \frac{g + \sqrt{g^2 - 4K_1 h}}{2K_1} > \frac{g}{2K_1} > \frac{c_O + c_R}{2} = \frac{E^* + D^* + R^* + D^*}{2},$$

re-arranging yields

$$0 > E^* + R^*.$$

Since E^* and R^* are positive, there is a contradiction. Therefore, we get $D_2 > \min(c_O, c_R)$ and then $D^* \neq D_2$.

Now, we consider that $D^* = D_1$ and prove that $0 \leq D_1 \leq \min(c_O, c_R)$. Let $\min(c_O, c_R) = m$, then $K_1 m(c_O + c_R) = K_1 m^2 + K_1 c_O c_R$.

$$k_{1m} m \geq 0,$$

$$m(K_1(c_O + c_R) + k_{1m}) \geq K_1 m^2 + K_1 c_O c_R,$$

$$mg - K_1 m^2 \geq h,$$

$$g^2 - 4K_1(mg - K_1 m^2) \leq g^2 - 4K_1 h,$$

$$(g - 2K_1 m)^2 \leq \left(\sqrt{g^2 - 4K_1 h}\right)^2,$$

$$g - \sqrt{g^2 - 4K_1 h} \leq 2K_1 m,$$

$$D_1 = \frac{g - \sqrt{g^2 - 4K_1 h}}{2K_1} \leq m = \min(c_O, c_R).$$

Therefore, $D_1 \leq \min(c_O, c_R)$. Moreover, D_1 is positive as follows: $g > \sqrt{g^2 - 4K_1 h}$.

Therefore, there is a unique equilibrium for (3.6), $D^* = D_1$.

To determine the stability of the equilibrium, we use the phase line analysis of (3.6). The right hand side of (3.6) is a quadratic equation with a positive leading coefficient, K_1 , that has two distinct real positive solutions, $D_1 < D_2$. Phase line is shown in Figure 3.1. The unique equilibrium $D^*(= D_1)$ is locally

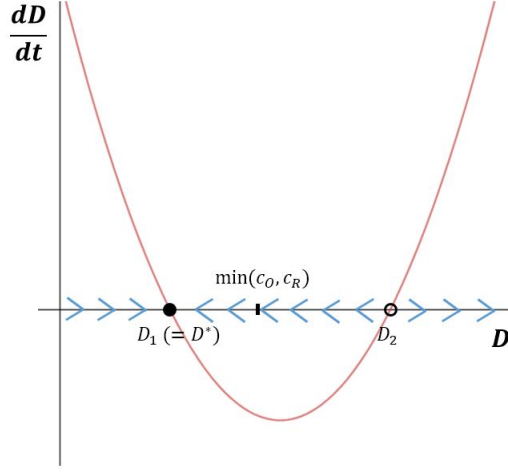


Figure 3.1: The phase line for (3.6).

asymptotically stable. Note that $0 \leq D(0) \leq \min(c_O, c_R)$, $\lim_{t \rightarrow \infty} D(t) = D^*$; therefore, we obtain that D^* is globally asymptotically stable.

Since $E(t) = c_O - D(t)$ and $R(t) = c_R - D(t)$ for all $t \geq 0$, we have $E^* = c_O - D^*$ and $R^* = c_R - D^*$. Hence, (3.5) has a unique equilibrium (3.7), which is globally asymptotically stable. \square

2) Model-ED1 ($k_{d1} \neq 0, k_{d2} = 0$)

The reduced system for Model-ED1 is:

$$\begin{aligned} \frac{dE}{dt} &= -K_1ER + k_{1m}D - k_{d1}E, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D. \end{aligned} \tag{3.10}$$

From $\frac{dR}{dt} + \frac{dD}{dt} = 0$, we get $R(t) + D(t) = c_R$ for all $t \geq 0$, which is the conservation law for Model-ED1. Then we can reduce further the system to two equations:

$$\begin{aligned}\frac{dE}{dt} &= -K_1E(c_R - D) + k_{1m}D - k_{d1}E, \\ \frac{dD}{dt} &= K_1E(c_R - D) - k_{1m}D.\end{aligned}\tag{3.11}$$

Theorem 3.3.2. *The unique equilibrium of system (3.10) is*

$$(E^*, R^*, D^*) = (0, c_R, 0)\tag{3.12}$$

and the equilibrium is globally asymptotically stable.

Proof. Note that (3.11) is the reduced system of (3.10), we obtain an equilibrium of (3.11) first, and then determine an equilibrium of (3.10) with the conservation law for Model-ED1.

Let $\frac{dE}{dt} = \frac{dD}{dt} = 0$ in (3.11), then

$$0 = -K_1E^*(c_R - D^*) + k_{1m}D^* - k_{d1}E^*,\tag{3.13a}$$

$$0 = K_1E^*(c_R - D^*) - k_{1m}D^*.\tag{3.13b}$$

By adding (3.13a) and (3.13b), we get $E^* = 0$, $D^* = 0$.

From (3.11), we get the Jacobian,

$$J_{ED1} = \begin{pmatrix} -K_1(c_R - D) - k_{d1} & K_1E + k_{1m} \\ K_1(c_R - D) & -K_1E - k_{1m} \end{pmatrix},$$

with the following properties at the equilibrium:

$$\det(J_{ED1}(E^*, D^*)) = \begin{vmatrix} -K_1 c_R - k_{d1} & k_{1m} \\ K_1 c_R & -k_{1m} \end{vmatrix} = \begin{vmatrix} -k_{d1} & 0 \\ K_1 c_R & -k_{1m} \end{vmatrix} = k_{d1} k_{1m} > 0,$$

$$\text{Tr}(J_{ED1}(E^*, D^*)) = -K_1 c_R - k_{d1} - k_{1m} < 0.$$

Therefore, the unique equilibrium point $(0, 0)$ of system (3.11) is locally asymptotically stable.

To determine the global stability, we use Poincaré-Bendixson's trichotomy [20].

First, we show that solutions of (3.11) are bounded.

If $D = 0$ then $\frac{dD}{dt} = K_1 E c_R \geq 0$ and $\frac{dE}{dt} = k_{1m} D \geq 0$ at $E = 0$ in (3.11), so $E(t) \geq 0$ and $D(t) \geq 0$ for all $t \geq 0$. Moreover, $\frac{d(E+D)}{dt} = -k_{d1} E \leq 0$. Thus, solutions are bounded.

Secondly, we use Bendixson's criterion [20] to check the existence of a periodic solution as follows:

$$\begin{aligned} & \frac{\partial \left(\frac{dE}{dt} \right)}{\partial E} + \frac{\partial \left(\frac{dD}{dt} \right)}{\partial D} \\ &= -K_1(c_R - D(t)) - k_{d1} - K_1 E(t) - k_{1m} < 0, \quad \forall t \geq 0. \end{aligned}$$

Thus, there is no periodic solution. Therefore, the unique equilibrium $(0, 0)$ of system (3.11) is globally asymptotically stable by the Poincaré-Bendixson trichotomy.

Since $R(t) = c_R - D(t)$ for all $t \geq 0$, we get $R^* = c_R - D^* = c_R$. Hence, (3.10) has a unique equilibrium $(E^*, R^*, D^*) = (0, c_R, 0)$, which is globally asymptotically stable. \square

3) Model-ED2 ($k_{d1} = 0, k_{d2} \neq 0$)

The reduced system for Model-ED2 is

$$\begin{aligned}\frac{dE}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D - k_{d2}D.\end{aligned}\tag{3.14}$$

In this model, there is no conservation. However, we have $\frac{dE}{dt} = \frac{dR}{dt}$, so $R(t) - R(0) = E(t) - E(0)$ and $R(t) = E(t) + c_R - c_O = E(t) + a_0$ with $a_0 = c_R - c_O$.

Then, we get the reduced system:

$$\begin{aligned}\frac{dE}{dt} &= -K_1E(E + a_0) + k_{1m}D, \\ \frac{dD}{dt} &= K_1E(E + a_0) - k_{1m}D - k_{d2}D.\end{aligned}\tag{3.15}$$

Theorem 3.3.3. *The unique equilibrium of system (3.14) is*

$$(E^*, R^*, D^*) = \begin{cases} (c_O - c_R, 0, 0) & c_O > c_R \\ (0, c_R - c_O, 0) & c_R > c_O \end{cases}\tag{3.16}$$

and the equilibrium is globally asymptotically stable.

Proof. Let $\frac{dE}{dt} = \frac{dD}{dt} = 0$ in (3.15), then

$$0 = -K_1E^*(E^* + a_0) + k_{1m}D^*,\tag{3.17a}$$

$$0 = K_1E^*(E^* + a_0) - k_{1m}D^* - k_{d2}D^*.\tag{3.17b}$$

By adding (3.17a) and (3.17b), we get $D^* = 0$ and two possible values for E^* . The first one is $E^* = 0$ (and $R^* = c_R - c_O$). The second one is $E^* = c_O - c_R$ (and $R^* = 0$). However, we know that E^* and R^* must be positive, so only one value for E^* is possible at the time, which depends on the sign of $a_0 = c_R - c_O$. Therefore, system (3.15) has a unique equilibrium whose the value depends on the initial condition.

To find the stability of the equilibrium, we get the Jacobian for (3.15),

$$J_{ED2} = \begin{pmatrix} -K_1 E - K_1(E + c_R - c_O) & k_{1m} \\ K_1 E + K_1(E + c_R - c_O) & -k_{1m} - k_{d2} \end{pmatrix}.$$

If $c_O - c_R < 0$, at the equilibrium $(E^*, D^*) = (0, 0)$, the determinant and trace of the Jacobian are

$$\det(J_{ED2}(0, 0)) =$$

$$\begin{vmatrix} -K_1(c_R - c_O) & k_{1m} \\ K_1(c_R - c_O) & -k_{1m} - k_{d2} \end{vmatrix} = \begin{vmatrix} -K_1(c_R - c_O) & k_{1m} \\ 0 & -k_{d2} \end{vmatrix} = K_1 k_{d2} (c_R - c_O) > 0$$

and

$$\text{Tr}(J_{ED2}(0, 0)) = -K_1(c_R - c_O) - k_{1m} - k_{d2} < 0.$$

If $c_O - c_R > 0$, at the equilibrium $(E^*, D^*) = (c_O - c_R, 0)$ the determinant and trace of the Jacobian are

$$\det(J_{ED2}(c_O - c_R, 0)) =$$

$$\begin{vmatrix} -K_1(c_O - c_R) & k_{1m} \\ K_1(c_O - c_R) & -k_{1m} - k_{d2} \end{vmatrix} = \begin{vmatrix} -K_1(c_O - c_R) & k_{1m} \\ 0 & -k_{d2} \end{vmatrix} = K_1 k_{d2} (c_O - c_R) > 0$$

and

$$\text{Tr}(J_{ED2}(c_O - c_R, 0)) = -K_1(c_O - c_R) - k_{1m} - k_{d2} < 0.$$

Therefore, the unique equilibrium of (3.15) is locally asymptotically stable.

To use the Poincaré-Bendixson trichotomy, the boundedness of solutions is proved first. Note that $\frac{dE}{dt} = k_{1m}D \geq 0$ at $E = 0$ and $\frac{dD}{dt} = K_1E(E + c_R - c_O) \geq 0$ at $D = 0$ because $\frac{dR}{dt} = k_{1m}D \geq 0$ at $R = 0$ and $R(t) = E(t) + c_R - c_O \geq 0$. Thus, we get $E(t) \geq 0$ and $D(t) \geq 0$ for all $t \geq 0$. Moreover $\frac{d(E+D)}{dt} = -k_{d2}D \leq 0$. Therefore, solutions are bounded.

Moreover, there is no periodic solution by Bendixson's criterion by the following argument:

$$\begin{aligned} & \frac{\partial \left(\frac{dE}{dt} \right)}{\partial E} + \frac{\partial \left(\frac{dD}{dt} \right)}{\partial D} \\ &= -K_1(E(t) + E(t) + c_R - c_O) - k_{1m} - k_{d2} < 0, \quad \forall t \geq 0, \end{aligned}$$

since $E(t) \geq 0$ and $E(t) + c_R - c_O = R(t) \geq 0$ for all $t \geq 0$.

The equilibrium value is determined by the initial condition c_O and c_R and there is a unique equilibrium in any case. Therefore, by the Poincaré-Bendixson trichotomy, the unique equilibrium defined in (3.15) is globally asymptotically stable.

Since, $R(t) = E(t) + a_0$ for all $t \geq 0$, we have $R^* = E^* + a_0$ in (3.14) and the unique equilibrium

$$(E^*, R^*, D^*) = \begin{cases} (c_O - c_R, 0, 0) & c_O > c_R, \\ (0, c_R - c_O, 0) & c_R > c_O, \end{cases}$$

which is globally asymptotically stable. \square

4) Model-ED3 ($k_{d1} \neq 0, k_{d2} \neq 0$)

The reduced system for Model-ED3 is

$$\begin{aligned}\frac{dE}{dt} &= -K_1ER + k_{1m}D - k_{d1}E, \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D, \\ \frac{dD}{dt} &= K_1ER - k_{1m}D - k_{d2}D.\end{aligned}\tag{3.18}$$

Theorem 3.3.4. *The equilibrium of system (3.18) is*

$$(E^*, R^*, D^*) = (0, r, 0)\tag{3.19}$$

where $0 \leq r \leq c_R$.

Proof. To find the steady state of the model, let $\frac{dE}{dt} = \frac{dR}{dt} = \frac{dD}{dt} = 0$ in (3.18).

$$\begin{aligned}0 &= -K_1E^*R^* + k_{1m}D^* - k_{d1}E^*, \\ 0 &= -K_1E^*R^* + k_{1m}D^*, \\ 0 &= K_1E^*R^* - k_{1m}D^* - k_{d2}D^*.\end{aligned}\tag{3.20}$$

From (3.20), we get $E^* = 0$, $D^* = 0$ and $R^* = r$ where r is non-negative constant. \square

We get the line of equilibria, which is $(0, r, 0)$ with $0 \leq r \leq c_R$. Previous methods cannot be used on non-isolated equilibrium.

3.4 Group 3 - Model-N

In this section, Model-N is mathematically studied. To derive an approximation of OAS2 activity, quasi-steady-state is first assumed. Then, the asymptotic behavior of the model is characterized by determining its equilibria and their stability. Recall ODE system of Model-N (2.11).

$$\frac{dE}{dt} = -K_1ER + k_{1m}C - k_5ER + k_{5m}N, \quad (3.21a)$$

$$\frac{dR}{dt} = -K_1ER + k_{1m}C - k_5ER + k_{5m}N,$$

$$\frac{dC}{dt} = K_1ER - k_{1m}C - k_2C + k_{2m}D, \quad (3.21b)$$

$$\frac{dD}{dt} = k_2C - k_{2m}D, \quad (3.21c)$$

$$\frac{dN}{dt} = k_5ER - k_{5m}N, \quad (3.21d)$$

$$\frac{dP}{dt} = k_pD.$$

3.4.1 Quasi-steady-state assumption for Model-N

As the cofactor, dsRNA, is in abundance with respect to the enzyme, it is assumed that enzymes are bound to dsRNA so quickly that there is no free enzyme left. If they detach from dsRNA, they instantly bind to dsRNA without

delay. In this circumstance, quasi-steady-state can be assumed:

$$\frac{dE}{dt} = -K_1ER + k_{1m}C - k_5ER + k_{5m}N \approx 0. \quad (3.22)$$

Moreover, $\frac{dE}{dt} = \frac{dR}{dt}$, we get $\frac{dR}{dt} \approx 0$.

From (3.22), we get

$$ER = \frac{k_{1m}C + k_{5m}N}{K_1 + k_5}. \quad (3.23)$$

By replacing ER in (3.21b), (3.21c) and (3.21d) with (3.23), we get the linear system in C , D , N .

$$\frac{dC}{dt} = K_1 \left(\frac{k_{1m}C + k_{5m}N}{K_1 + k_5} \right) - k_{1m}C - k_2C + k_{2m}D,$$

$$\frac{dD}{dt} = k_2C - k_{2m}D,$$

$$\frac{dN}{dt} = k_5 \left(\frac{k_{1m}C + k_{5m}N}{K_1 + k_5} \right) - k_{5m}N.$$

It can be expressed in the following matrix form

$$\begin{pmatrix} \frac{dC}{dt} \\ \frac{dD}{dt} \\ \frac{dN}{dt} \end{pmatrix} = \begin{pmatrix} \frac{K_1k_{1m}}{K_1+k_5} - k_{1m} - k_2 & k_{2m} & \frac{K_1k_{5m}}{K_1+k_5} \\ k_2 & -k_{2m} & 0 \\ \frac{k_5k_{1m}}{K_1+k_5} & 0 & \frac{k_5k_{5m}}{K_1+k_5} - k_{5m} \end{pmatrix} \begin{pmatrix} C \\ D \\ N \end{pmatrix} \quad (3.24)$$

with initial condition $(C(0), D(0), N(0)) = (\epsilon_C, 0, \epsilon_N)$, where $\epsilon_C > 0$ and $\epsilon_N > 0$.

Solutions are non-negative since Model-N is well-posed. It also is easy to show that when starting with non-negative initial condition in the linear system (3.24), solutions stay non-negative.

Theorem 3.4.1. *Using quasi-steady-state assumptions in Model-N (3.21) with an initial condition $(C(0), D(0), N(0)) = (\epsilon_C, 0, \epsilon_N)$, where $\epsilon_C > 0$ and $\epsilon_N > 0$, and $P(0) = 0$ leads to*

1. $c_0 > 4d_0$

$$P(t) = k_p k_2 \left[\frac{\alpha_1}{k_{2m}} t + \exp(\lambda_2 t) \left(\frac{\beta_1}{\lambda_2(k_{2m} + \lambda_2)} \right) + \exp(\lambda_3 t) \left(\frac{\gamma_1}{\lambda_3(k_{2m} + \lambda_3)} \right) - \left(\frac{\beta_1}{\lambda_2(k_{2m} + \lambda_2)} \right) - \left(\frac{\gamma_1}{\lambda_3(k_{2m} + \lambda_3)} \right) \right], \forall t \geq 0, \quad (3.25)$$

where $a_0 = \frac{K_1 k_{5m}}{K_1 + k_5}$, $b_0 = \frac{k_5 k_{1m}}{K_1 + k_5}$, $c_0 = k_2 + k_{2m} + a_0 + b_0$, $d_0 = k_2 a_0 + k_{2m} a_0 + k_{2m} b_0$, $\lambda_2 = \frac{-c_0 - \sqrt{c_0^2 - 4d_0}}{2}$, $\lambda_3 = \frac{-c_0 + \sqrt{c_0^2 - 4d_0}}{2}$ and

$$\alpha_1 = \frac{k_{2m} a_0 (\epsilon_N (\lambda_2 + a_0) (\lambda_3 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2 \lambda_3 b_0 (k_{2m} - a_0)},$$

$$\beta_1 = \frac{(\lambda_2 + k_{2m}) (\lambda_2 + a_0) (a_0 \epsilon_N (\lambda_3 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2 b_0 (a_0 - k_{2m}) (\lambda_3 - \lambda_2)},$$

$$\gamma_1 = \frac{(\lambda_3 + k_{2m}) (\lambda_3 + a_0) (a_0 \epsilon_N (\lambda_2 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_3 b_0 (k_{2m} - a_0) (\lambda_3 - \lambda_2)}.$$

2. $c_0 = 4d_0$

$$\begin{aligned}
P(t) = & k_p k_2 \left[\frac{\alpha_2}{k_{2m}} t + \exp(\lambda_2 t) \left(\frac{\beta_2}{\lambda_2(\lambda_2 + k_{2m})} \right) \right. \\
& - \gamma_2 \exp(\lambda_2 t) \left(\frac{(\lambda_2 t - 1)}{\lambda_2^2(\lambda_2 + k_{2m})} - \frac{1}{\lambda_2(\lambda_2 + k_{2m})^2} \right) \\
& \left. - \frac{\beta_2}{\lambda_2(\lambda_2 + k_{2m})} + \gamma_2 \left(\frac{(\lambda_2 t - 1)}{\lambda_2^2(\lambda_2 + k_{2m})} - \frac{1}{\lambda_2(\lambda_2 + k_{2m})^2} \right) \right], \forall t \geq 0,
\end{aligned} \tag{3.26}$$

where $\lambda_2 = \frac{-c_0}{2}$ and

$$\begin{aligned}
\alpha_2 &= \frac{k_{2m} a_0 (\epsilon_N (\lambda_2 + a_0)^2 + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2^2 b_0 (k_{2m} - a_0)}, \\
\beta_2 &= \frac{b_0 \epsilon_C (k_{2m} - a_0) (\lambda_2^2 - k_{2m} a_0) - k_{2m} a_0 \epsilon_N (\lambda_2 + a_0)^2}{\lambda_2^2 b_0 (k_{2m} - a_0)}, \\
\gamma_2 &= \frac{b_0 \epsilon_C (k_{2m} - a_0) (\lambda_2 + k_{2m}) (\lambda_2 + a_0) + a_0 \epsilon_N (\lambda_2 + k_{2m}) (\lambda_2 + a_0)^2}{\lambda_2 b_0 (k_{2m} - a_0)}.
\end{aligned}$$

3. $c_0 < 4d_0$

$$\begin{aligned}
P(t) = & k_p k_2 \left[\frac{\alpha_3}{k_{2m}} t + \frac{4\beta_3 \exp\left(-\frac{c_0}{2}t\right)}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \left((e_0 - c_0) f_0 \sin\left(\frac{f_0}{2}t\right) - (c_0 e_0 + f_0^2) \cos\left(\frac{f_0}{2}t\right) \right) \right. \\
& + \frac{4\gamma_3 \exp\left(-\frac{c_0}{2}t\right)}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \left((c_0 e_0 + f_0^2) \sin\left(\frac{f_0}{2}t\right) + (e_0 - c_0) f_0 \cos\left(\frac{f_0}{2}t\right) \right) \\
& \left. + \frac{4\beta_3(c_0 e_0 + f_0^2) + 4\gamma_3(c_0 - e_0)f_0}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \right], \forall t \geq 0,
\end{aligned} \tag{3.27}$$

where $e_0 = 2k_{2m} - c_0$, $f_0 = 4d_0 - c_0^2$, $g_0 = 2a_0 - c_0$ and

$$\alpha_3 = \frac{k_{2m}a_0(\epsilon_N(g_0^2 + f_0^2) + 2b_0\epsilon_C(e_0 - g_0))}{b_0(k_{2m}(g_0^2 + f_0^2) - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))},$$

$$\beta_3 = \frac{k_{2m}(g_0^2 + f_0^2)(b_0\epsilon_C - a_0\epsilon_N) - a_0b_0\epsilon_C(e_0^2 + f_0^2)}{b_0(k_{2m}(g_0^2 + f_0^2) - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))},$$

$$\gamma_3 = \frac{2k_{2m}e_0(g_0^2 + f_0^2)(b_0\epsilon_C - a_0\epsilon_N) + a_0(e_0^2 + f_0^2)(\epsilon_N(g_0^2 + f_0^2) - 2b_0\epsilon_Cg_0)}{2b_0f_0((g_0^2 + f_0^2)k_{2m} - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))}.$$

Proof. System (3.24) or $\frac{d\mathbb{X}}{dt} = \mathcal{A}\mathbb{X}$, where $\mathbb{X} = \begin{pmatrix} C \\ D \\ N \end{pmatrix}$, is homogeneous so there

is a trivial solution. Non-trivial solutions of (3.24) can be calculated with eigenvalues and eigenvectors of \mathcal{A} .

The eigenvalues of \mathcal{A} are the solutions of the characteristic polynomial of \mathcal{A} :

$$\lambda(\lambda^2 + (k_2 + k_{2m} + a_0 + b_0)\lambda + k_2a_0 + k_{2m}a_0 + k_{2m}b_0) = 0,$$

where $a_0 = \frac{K_1k_{5m}}{K_1+k_5}$ and $b_0 = \frac{k_5k_{1m}}{K_1+k_5}$.

Thus, the eigenvalues of \mathcal{A} are

$$\lambda_1 = 0, \quad \lambda_2 = \frac{-c_0 - \sqrt{c_0^2 - 4d_0}}{2}, \quad \lambda_3 = \frac{-c_0 + \sqrt{c_0^2 - 4d_0}}{2},$$

where $c_0 = k_2 + k_{2m} + a_0 + b_0$ and $d_0 = k_2a_0 + k_{2m}a_0 + k_{2m}b_0$.

For λ_2 and λ_3 , we consider three cases: $c_0^2 - 4d_0 > 0$, $c_0^2 - 4d_0 = 0$, $c_0^2 - 4d_0 < 0$.

First, we find solutions for $c_0^2 - 4d_0 > 0$. Then, there are three distinct real eigenvalues and we get the general solution,

$$\begin{pmatrix} C(t) \\ D(t) \\ N(t) \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m}} \\ \frac{b_0}{a_0} \end{pmatrix} + \beta_1 \exp(\lambda_2 t) \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m} + \lambda_2} \\ \frac{b_0}{a_0 + \lambda_2} \end{pmatrix} + \gamma_1 \exp(\lambda_3 t) \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m} + \lambda_3} \\ \frac{b_0}{a_0 + \lambda_3} \end{pmatrix}, \quad \forall t \geq 0, \quad (3.28)$$

with arbitrary constants α_1 , β_1 , and γ_1 .

To satisfy the initial condition, $(C(0), D(0), N(0)) = (\epsilon_C, 0, \epsilon_N)$, let $t = 0$ in (3.28), then

$$\begin{aligned} \alpha_1 + \beta_1 + \gamma_1 &= \epsilon_C, \\ \alpha_1 \frac{k_2}{k_{2m}} + \beta_1 \frac{k_2}{k_{2m} + \lambda_2} + \gamma_1 \frac{k_2}{k_{2m} + \lambda_3} &= 0, \\ \alpha_1 \frac{b_0}{a_0} + \beta_1 \frac{b_0}{a_0 + \lambda_2} + \gamma_1 \frac{b_0}{a_0 + \lambda_3} &= \epsilon_N. \end{aligned}$$

Solving the linear algebraic system for α_1 , β_1 and γ_1 , we obtain

$$\begin{aligned} \alpha_1 &= \frac{k_{2m} a_0 (\epsilon_N (\lambda_2 + a_0) (\lambda_3 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2 \lambda_3 b_0 (k_{2m} - a_0)}, \\ \beta_1 &= \frac{(\lambda_2 + k_{2m}) (\lambda_2 + a_0) (a_0 \epsilon_N (\lambda_3 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2 b_0 (a_0 - k_{2m}) (\lambda_3 - \lambda_2)}, \\ \gamma_1 &= \frac{(\lambda_3 + k_{2m}) (\lambda_3 + a_0) (a_0 \epsilon_N (\lambda_2 + a_0) + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_3 b_0 (k_{2m} - a_0) (\lambda_3 - \lambda_2)}. \end{aligned} \quad (3.29)$$

The solution starting with the initial condition $(C(0), D(0), N(0)) = (\epsilon_C, 0, \epsilon_N)$ is (3.28) with α_1 , β_1 and γ_1 as defined in (3.29).

Secondly, we assume that $c_0^2 - 4d_0 = 0$. Then, there are two eigenvalues, $\lambda_1 = 0$ and $\lambda_2 = -\frac{c_0}{2}$ with multiplicity 2. For the eigenvalue 0, we obtain the same eigenvector as in the previous case, $\mathbf{v}_1 = \left(1, \frac{k_2}{k_{2m}}, \frac{b_0}{a_0}\right)^T$. For the eigenvalue $-\frac{c_0}{2}$, we get a first eigenvector, $\mathbf{v}_2 = \left(1, \frac{k_2}{k_{2m} + \lambda_2}, \frac{b_0}{a_0 + \lambda_2}\right)^T$. A second linearly independent eigenvector $\boldsymbol{\rho}$ is obtained using the generalized eigenvector method:

$$(\mathcal{A} - \lambda_2)\boldsymbol{\rho} = \mathbf{v}_2.$$

We obtain $\boldsymbol{\rho} = \begin{pmatrix} 0 \\ -\frac{k_2}{(k_{2m} + \lambda_2)^2} \\ -\frac{b_0}{(a_0 + \lambda_2)^2} \end{pmatrix}$. Then the general solution is

$$\begin{pmatrix} C(t) \\ D(t) \\ N(t) \end{pmatrix} = \alpha_2 \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m}} \\ \frac{b_0}{a_0} \end{pmatrix} + \beta_2 \exp(\lambda_2 t) \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m} + \lambda_2} \\ \frac{b_0}{a_0 + \lambda_2} \end{pmatrix} + \gamma_2 \left(t \exp(\lambda_2 t) \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m} + \lambda_2} \\ \frac{b_0}{a_0 + \lambda_2} \end{pmatrix} + \exp(\lambda_2 t) \begin{pmatrix} 0 \\ -\frac{k_2}{(k_{2m} + \lambda_2)^2} \\ -\frac{b_0}{(a_0 + \lambda_2)^2} \end{pmatrix} \right), \forall t \geq 0 \quad (3.30)$$

with arbitrary constants α_2 , β_2 , and γ_2 .

To satisfy the initial condition, the arbitrary constants are chosen to be solutions

to the following linear algebraic system:

$$\begin{aligned}\alpha_2 + \beta_2 &= \epsilon_C, \\ \alpha_2 \frac{k_2}{k_{2m}} + \beta_2 \frac{k_2}{k_{2m} + \lambda_2} - \gamma_2 \frac{k_2}{(k_{2m} + \lambda_2)^2} &= 0, \\ \alpha_2 \frac{b_0}{a_0} + \beta_2 \frac{b_0}{a_0 + \lambda_2} - \gamma_2 \frac{b_0}{(a_0 + \lambda_2)^2} &= \epsilon_N.\end{aligned}$$

Hence, we obtain:

$$\begin{aligned}\alpha_2 &= \frac{k_{2m} a_0 (\epsilon_N (\lambda_2 + a_0)^2 + b_0 \epsilon_C (k_{2m} - a_0))}{\lambda_2^2 b_0 (k_{2m} - a_0)}, \\ \beta_2 &= \frac{b_0 \epsilon_C (k_{2m} - a_0) (\lambda_2^2 - k_{2m} a_0) - k_{2m} a_0 \epsilon_N (\lambda_2 + a_0)^2}{\lambda_2^2 b_0 (k_{2m} - a_0)}, \\ \gamma_2 &= \frac{b_0 \epsilon_C (k_{2m} - a_0) (\lambda_2 + k_{2m}) (\lambda_2 + a_0) + a_0 \epsilon_N (\lambda_2 + k_{2m}) (\lambda_2 + a_0)^2}{\lambda_2 b_0 (k_{2m} - a_0)}.\end{aligned}\tag{3.31}$$

Consequently, the unique solution satisfying the initial condition, when $c_0^2 - 4d_0 = 0$, is (3.30) with α_2 , β_2 and γ_2 as defined in (3.31).

Finally, we consider that $c_0^2 - 4d_0 < 0$ and \mathcal{A} has one zero eigenvalue and two complex conjugate eigenvalues, $\frac{-c_0 \pm i\sqrt{4d_0 - c_0^2}}{2}$, with the negative real part. The

eigenvector corresponding to eigenvalue $\lambda_2 = \frac{-c_0 - i\sqrt{4d_0 - c_0^2}}{2}$ is

$$\begin{aligned} \mathbf{v} &= \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m} + \lambda_2} \\ \frac{b_0}{a_0 + \lambda_2} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{2k_2}{2k_{2m} - c_0 - i\sqrt{4d_0 - c_0^2}} \\ \frac{2b_0}{2a_0 - c_0 - i\sqrt{4d_0 - c_0^2}} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{2k_2(e_0 + if_0)}{e_0^2 + f_0^2} \\ \frac{2b_0(g_0 + if_0)}{g_0^2 + f_0^2} \end{pmatrix} = \\ &= \begin{pmatrix} 1 \\ \frac{2k_2e_0}{e_0^2 + f_0^2} \\ \frac{2b_0g_0}{g_0^2 + f_0^2} \end{pmatrix} + i \begin{pmatrix} 0 \\ \frac{2k_2f_0}{e_0^2 + f_0^2} \\ \frac{2b_0f_0}{g_0^2 + f_0^2} \end{pmatrix} := \mathbf{v}_1 + i\mathbf{v}_2 \end{aligned}$$

where $e_0 = 2k_{2m} - c_0$, $f_0 = \sqrt{4d_0 - c_0^2}$ and $g_0 = 2a_0 - c_0$.

Hence, the solution corresponding to λ_2 is

$$\mathbf{x}(t) = \exp(\lambda_2 t) \mathbf{v}. \quad (3.32)$$

Then,

$$\begin{aligned} \mathbf{x}(t) &= \exp(\lambda_2 t) \mathbf{v} = \exp\left(\frac{-c_0}{2}t\right) \left(\cos\left(\frac{f_0}{2}t\right) - i \sin\left(\frac{f_0}{2}t\right) \right) (\mathbf{v}_1 + i\mathbf{v}_2) \\ &= \exp\left(\frac{-c_0}{2}t\right) (\mathbf{x}_1(t) + i\mathbf{x}_2(t)), \end{aligned}$$

where

$$\mathbf{x}_1(t) = \begin{pmatrix} \cos\left(\frac{f_0}{2}t\right) \\ \cos\left(\frac{f_0}{2}t\right) \left(\frac{2k_2e_0}{e_0^2+f_0^2}\right) + \sin\left(\frac{f_0}{2}t\right) \left(\frac{2k_2f_0}{e_0^2+f_0^2}\right) \\ \cos\left(\frac{f_0}{2}t\right) \left(\frac{2b_0g_0}{g_0^2+f_0^2}\right) + \sin\left(\frac{f_0}{2}t\right) \left(\frac{2b_0f_0}{g_0^2+f_0^2}\right) \end{pmatrix}$$

$$\mathbf{x}_2(t) = \begin{pmatrix} -\sin\left(\frac{f_0}{2}t\right) \\ -\sin\left(\frac{f_0}{2}t\right) \left(\frac{2k_2e_0}{e_0^2+f_0^2}\right) + \cos\left(\frac{f_0}{2}t\right) \left(\frac{2k_2f_0}{e_0^2+f_0^2}\right) \\ -\sin\left(\frac{f_0}{2}t\right) \left(\frac{2b_0g_0}{g_0^2+f_0^2}\right) + \cos\left(\frac{f_0}{2}t\right) \left(\frac{2b_0f_0}{g_0^2+f_0^2}\right) \end{pmatrix}$$

Note that if $\mathbf{y} = \mathbf{y}_1 + i\mathbf{y}_2$ is the solution of $\mathbf{y}' = \mathcal{A}\mathbf{y}$ and \mathcal{A} is a real matrix, then both \mathbf{y}_1 and \mathbf{y}_2 are real solutions of $\mathbf{y}' = \mathcal{A}\mathbf{y}$. Thus, the general solution is

$$\begin{pmatrix} C(t) \\ D(t) \\ N(t) \end{pmatrix} = \alpha_3 \begin{pmatrix} 1 \\ \frac{k_2}{k_{2m}} \\ \frac{b_0}{a_0} \end{pmatrix} + \beta_3 \exp\left(\frac{-c_0}{2}t\right) \mathbf{x}_1(t) + \gamma_3 \exp\left(\frac{-c_0}{2}t\right) \mathbf{x}_2(t), \forall t \geq 0$$

(3.33)

with arbitrary constants α_3 , β_3 and γ_3 .

To determine the values of the arbitrary constants α_3 , β_3 and γ_3 that satisfy the initial condition, the following linear system must be solved:

$$\alpha_3 + \beta_3 = \epsilon_C,$$

$$\alpha_3 \frac{k_2}{k_{2m}} + \beta_3 \frac{2k_2e_0}{e_0^2+f_0^2} + \gamma_3 \frac{2k_2f_0}{e_0^2+f_0^2} = 0,$$

$$\alpha_3 \frac{b_0}{a_0} + \beta_3 \frac{2b_0g_0}{g_0^2+f_0^2} + \gamma_3 \frac{2b_0f_0}{g_0^2+f_0^2} = \epsilon_N.$$

Its solution is:

$$\begin{aligned}
\alpha_3 &= \frac{k_{2m}a_0(\epsilon_N(g_0^2 + f_0^2) + 2b_0\epsilon_C(e_0 - g_0))}{b_0(k_{2m}(g_0^2 + f_0^2) - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))}, \\
\beta_3 &= \frac{k_{2m}(g_0^2 + f_0^2)(b_0\epsilon_C - a_0\epsilon_N) - a_0b_0\epsilon_C(e_0^2 + f_0^2)}{b_0(k_{2m}(g_0^2 + f_0^2) - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))}, \\
\gamma_3 &= \frac{2k_{2m}e_0(g_0^2 + f_0^2)(b_0\epsilon_C - a_0\epsilon_N) + a_0(e_0^2 + f_0^2)(\epsilon_N(g_0^2 + f_0^2) - 2b_0\epsilon_Cg_0)}{2b_0f_0((g_0^2 + f_0^2)k_{2m} - a_0(e_0^2 + f_0^2) + 2k_{2m}a_0(e_0 - g_0))}.
\end{aligned} \tag{3.34}$$

Consequently, the unique solution satisfying the initial condition, when $c_0^2 - 4d_0 < 0$, is (3.33) with α_3 , β_3 and γ_3 as defined in (3.34).

From $\frac{dP}{dt} = k_p D$, we get the following three expressions for $P(t)$ by integrating with respect to t with $P(0) = 0$.

1. $c_0 > 4d_0$

$$\begin{aligned}
P(t) &= k_p k_2 \left[\frac{\alpha_1}{k_{2m}} t + \exp(\lambda_2 t) \left(\frac{\beta_1}{\lambda_2(k_{2m} + \lambda_2)} \right) + \exp(\lambda_3 t) \left(\frac{\gamma_1}{\lambda_3(k_{2m} + \lambda_3)} \right) \right. \\
&\quad \left. - \left(\frac{\beta_1}{\lambda_2(k_{2m} + \lambda_2)} \right) - \left(\frac{\gamma_1}{\lambda_3(k_{2m} + \lambda_3)} \right) \right], \quad \forall t \geq 0
\end{aligned}$$

2. $c_0 = 4d_0$

$$\begin{aligned}
P(t) = & k_p k_2 \left[\frac{\alpha_2}{k_{2m}} t + \frac{\beta_2}{\lambda_2(\lambda_2 + k_{2m})} \exp(\lambda_2 t) \right. \\
& - \gamma_2 \exp(\lambda_2 t) \left(\frac{(\lambda_2 t - 1)}{\lambda_2^2(\lambda_2 + k_{2m})} - \frac{1}{\lambda_2(\lambda_2 + k_{2m})^2} \right) \\
& \left. - \frac{\beta_2}{\lambda_2(\lambda_2 + k_{2m})} + \gamma_2 \left(\frac{(\lambda_2 t - 1)}{\lambda_2^2(\lambda_2 + k_{2m})} - \frac{1}{\lambda_2(\lambda_2 + k_{2m})^2} \right) \right], \quad \forall t \geq 0
\end{aligned}$$

3. $c_0 < 4d_0$

$$\begin{aligned}
P(t) = & k_p k_2 \left[\frac{\alpha_3}{k_{2m}} t + \frac{4\beta_3 \exp\left(-\frac{c_0}{2}t\right)}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \left((e_0 - c_0)f_0 \sin\left(\frac{f_0}{2}t\right) - (c_0 e_0 + f_0^2) \cos\left(\frac{f_0}{2}t\right) \right) \right. \\
& + \frac{4\gamma_3 \exp\left(-\frac{c_0}{2}t\right)}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \left((c_0 e_0 + f_0^2) \sin\left(\frac{f_0}{2}t\right) + (e_0 - c_0)f_0 \cos\left(\frac{f_0}{2}t\right) \right) \\
& \left. + \frac{4\beta_3(c_0 e_0 + f_0^2) + 4\gamma_3(c_0 - e_0)f_0}{(c_0^2 + f_0^2)(e_0^2 + f_0^2)} \right], \quad \forall t \geq 0.
\end{aligned}$$

□

3.4.2 Asymptotic behavior of Model-N

Before discussing equilibria of Model-N, the equation $\frac{dP}{dt}$ is decoupled from system (3.21) similarly to other models of group 2 and it is called reduced Model-N.

Note that $\frac{dE}{dt} + \frac{dC}{dt} + \frac{dD}{dt} + \frac{dN}{dt} = 0$ and $\frac{dR}{dt} + \frac{dC}{dt} + \frac{dD}{dt} + \frac{dN}{dt} = 0$, so by integrating

with respect to time t we obtain:

$$E(t) + C(t) + D(t) + N(t) = c_O, \quad R(t) + C(t) + D(t) + N(t) = c_R.$$

for all $t \geq 0$. These two equations are referred as the conservation laws of Model-N. These laws are used to further reduce the system of equations:

$$\begin{aligned} \frac{dC}{dt} &= k_1(c_O - (C + D + N))(c_R - (C + D + N)) - k_{1m}C - k_2C + k_{2m}D, \\ \frac{dD}{dt} &= k_2C - k_{2m}D, \\ \frac{dN}{dt} &= k_5(c_O - (C + D + N))(c_R - (C + D + N)) - k_{5m}N, \end{aligned} \tag{3.35}$$

and we get $C(t) + D(t) + N(t) \leq \min(c_O, c_R)$ for all $t \geq 0$.

Theorem 3.4.2. *The unique positive equilibrium of reduced Model-N is*

$$\begin{aligned} &(E^*, R^*, C^*, D^*, N^*) \\ &= (c_O - C^* - D^* - N^*, \quad c_R - C^* - D^* - N^*, \\ &\frac{a_0(c_O + c_R) + \frac{1}{b_1} - \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1}\right)^2 - 4a_0^2c_Oc_R}}{2a_0^2}, \\ &b_2 \left(\frac{a_0(c_O + c_R) + \frac{1}{b_1} - \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1}\right)^2 - 4a_0^2c_Oc_R}}{2a_0^2} \right), \\ &\frac{b_5}{b_1} \left(\frac{a_0(c_O + c_R) + \frac{1}{b_1} - \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1}\right)^2 - 4a_0^2c_Oc_R}}{2a_0^2} \right) \right), \end{aligned} \tag{3.36}$$

where $a_0 = \left(1 + b_2 + \frac{b_5}{b_1}\right)$, $b_1 = \frac{K_1}{k_{1m}}$, $b_2 = \frac{k_2}{k_{2m}}$, $b_5 = \frac{k_5}{k_{5m}}$ and the equilibrium is locally asymptotically stable.

Proof. Let $\frac{dC}{dt} = \frac{dD}{dt} = \frac{dN}{dt} = 0$ in (3.35). Then we get the system:

$$\begin{aligned} 0 = & K_1(c_O - (C^* + D^* + N^*))(c_R - (C^* + D^* + N^*)) \\ & - k_{1m}C^* - k_2C^* + k_{2m}D^*, \end{aligned} \quad (3.37a)$$

$$0 = k_2C^* - k_{2m}D^*, \quad (3.37b)$$

$$0 = k_5(c_O - (C^* + D^* + N^*))(c_R - (C^* + D^* + N^*)) - k_{5m}N^*. \quad (3.37c)$$

By substituting (3.37b) into (3.37a),

$$(c_O - (C^* + D^* + N^*))(c_R - (C^* + D^* + N^*)) = \frac{k_{1m}}{K_1}C^*.$$

By arranging (3.37b) and (3.37c), we get

$$D^* = \frac{k_2}{k_{2m}}C^* \quad (3.39)$$

and

$$(c_O - (C^* + D^* + N^*))(c_R - (C^* + D^* + N^*)) = \frac{k_{5m}}{k_5}N^*.$$

Then we get the following equations:

$$\frac{k_{1m}}{K_1}C^* = \frac{k_{5m}}{k_5}N^* \Rightarrow N^* = \frac{k_{1m}k_5}{K_1k_{5m}}C^*, \quad (3.40)$$

then,

$$C^* + D^* + N^* = C^* \left(1 + \frac{k_2}{k_{2m}} + \frac{k_{1m}k_5}{K_1k_{5m}} \right) = a_0C^*, \quad (3.41)$$

where $a_0 = \left(1 + \frac{k_2}{k_{2m}} + \frac{k_{1m}k_5}{K_1k_{5m}} \right) > 0$. Substitute (3.41) to (3.37a), then

$$0 = (c_O - a_0C^*)(c_R - a_0C^*) - \frac{k_{1m}}{K_1}C^*.$$

Hence, C^* satisfies the following quadratic equation:

$$a_0^2C^{*2} - \left(a_0(c_O + c_R) + \frac{1}{b_1} \right) C^* + c_Oc_R = 0, \quad \frac{1}{b_1} = \frac{k_{1m}}{K_1} \geq 0, \quad (3.42)$$

having for solutions

$$C_1 = \frac{a_0(c_O + c_R) + \frac{1}{b_1} - \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R}}{2a_0^2}, \quad (3.43)$$

$$C_2 = \frac{a_0(c_O + c_R) + \frac{1}{b_1} + \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R}}{2a_0^2}.$$

Note that the solutions have to satisfy the following properties: real, positive and based on conservation laws:

$$C^* + D^* + N^* = a_0C^* \leq \min(c_O, c_R).$$

Both solutions, C_1 and C_2 , are real as

$$\begin{aligned} \left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R &> (a_0(c_O + c_R))^2 - 4a_0^2c_Oc_R \\ &= (a_0(c_O - c_R))^2 \geq 0, \end{aligned}$$

and positive as

$$a_0(c_O + c_R) + \frac{1}{b_1} > \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1}\right)^2 - 4a_0^2c_Oc_R} \geq 0.$$

Consider $C^* = C_2$, then

$$C^* + D^* + N^* = a_0C^* = a_0C_2 > a_0 \left(\frac{a_0(c_O + c_R)}{2a_0^2} \right) = \frac{c_O + c_R}{2} \geq \min(c_O, c_R).$$

Hence, $C^* + D^* + N^* \geq \min(c_O, c_R)$ and it is a contradiction to the conservation laws. Thus, $C^* \neq C_2$.

Consider $C^* = C_1$ and prove that $C^* + D^* + N^* \leq \min(c_O, c_R)$. Let $\min(c_O, c_R) = m$, then $(4a_0^2(c_O + c_R)m - 4a_0^2m^2 - 4a_0^2c_Oc_R) = 0$.

$$4a_0 \frac{1}{b_1} m \geq 0,$$

$$4a_0 \frac{1}{b_1} m + (4a_0^2(c_O + c_R)m - 4a_0^2m^2 - 4a_0^2c_Oc_R) \geq 0,$$

$$4a_0m \left(a_0(c_O + c_R) + \frac{1}{b_1} \right) - 4a_0^2m^2 \geq 4a_0^2c_Oc_R,$$

$$\left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0m \left(a_0(c_O + c_R) + \frac{1}{b_1} \right) + 4a_0^2m^2 \leq \left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R,$$

$$\left(a_0(c_O + c_R) + \frac{1}{b_1} \right) - 2a_0m \leq \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R},$$

$$a_0C^* = \frac{\left(a_0(c_O + c_R) + \frac{1}{b_1} \right) - \sqrt{\left(a_0(c_O + c_R) + \frac{1}{b_1} \right)^2 - 4a_0^2c_Oc_R}}{2a_0} \leq m = \min(c_O, c_R).$$

Thus, $C^* + D^* + N^* = a_0 C^* \leq \min(c_O, c_R)$, then $C^* = C_1$.

Hence, system (3.35) has a unique equilibrium (C^*, D^*, N^*) with $C^* = C_1$ defined in (3.43), D^* defined in (3.39) and N^* defined in (3.40).

To determine the stability of the equilibrium, the Jacobian of system (3.35) is evaluated at the equilibrium.

$$J_N(C^*, D^*, N^*) = [A_1 \ A_2 \ A_3], \quad (3.44)$$

where the columns are

$$A_1 = \begin{pmatrix} -K_1(c_O - C^* - D^* - N^*) - K_1(c_R - C^* - D^* - N^*) - k_{1m} - k_2 \\ k_2 \\ -k_5(c_O - C^* - D^* - N^*) - k_5(c_R - C^* - D^* - N^*) \end{pmatrix},$$

$$A_2 = \begin{pmatrix} -K_1(c_O - C^* - D^* - N^*) - K_1(c_R - C^* - D^* - N^*) + k_{2m} \\ -k_{2m} \\ -k_5(c_O - C^* - D^* - N^*) - k_5(c_R - C^* - D^* - N^*) \end{pmatrix},$$

$$A_3 = \begin{pmatrix} -K_1(c_O - C^* - D^* - N^*) - K_1(c_R - C^* - D^* - N^*) \\ 0 \\ -k_5(c_O - C^* - D^* - N^*) - k_5(c_R - C^* - D^* - N^*) - k_{5m} \end{pmatrix}.$$

To determine the sign of the eigenvalues of (3.44), we use the Routh-Hurwitz criterion [21]; but first we simplify the expression of the matrix. From the conservation laws of Model-N, we replace $c_O - C^* - D^* - N^*$ and $c_R - C^* - D^* - N^*$

by E^* and R^* , respectively. Then the Jacobian $J_N(C^*, D^*, N^*)$ is

$$\begin{pmatrix} -K_1(E^* + R^*) - k_{1m} - k_2 & -K_1(E^* + R^*) + k_{2m} & -K_1(E^* + R^*) \\ k_2 & -k_{2m} & 0 \\ -k_5(E^* + R^*) & -k_5(E^* + R^*) & -k_5(E^* + R^*) - k_{5m} \end{pmatrix}.$$

The characteristic polynomial of $J_N(C^*, D^*, N^*)$ is

$$p_{J_N}(C^*, D^*, N^*) = \det(xI - J_N(C^*, D^*, N^*)) = x^3 + c_1x^2 + c_2x + c_3,$$

$$\text{where } c_1 = (E^* + R^*)(K_1 + k_5) + k_{1m} + k_2 + k_{2m} + k_{5m},$$

$$c_2 = (E^* + R^*)((k_2 + k_{2m} + k_{5m})K_1 + (k_{1m} + k_2 + k_{2m})k_5)$$

$$+ (k_{1m} + k_2 + k_{2m})k_{5m} + k_{1m}k_{2m},$$

$$c_3 = (E^* + R^*)((k_2 + k_{2m})K_1k_{5m} + k_{1m}k_{2m}k_5) + k_{1m}k_{2m}k_{5m}.$$

Since all the parameters and $E^* + R^*$ are positive, we get $c_1 > 0$ and $c_3 > 0$.

All the terms of c_1 , c_2 and c_3 are positive. By simple comparison, all the terms of c_3 can be found in c_1c_2 and then c_3 is cancelled in $c_1c_2 - c_3$. For example, the term $(E^* + R^*)(k_2 + k_{2m})K_1k_{5m}$ in c_3 can be made by multiplying $(E^* + R^*)(k_2 + k_{2m})K_1$ in c_2 by k_{5m} in c_1 . And so on, all the other terms of c_3 are cancelled. Hence, we satisfy the last condition of the Routh-Hurwitz criterion $c_1c_2 - c_3 > 0$. Consequently, the equilibrium is locally asymptotically stable by Routh-Hurwitz criterion.

Hence, the reduced Model-N has a unique equilibrium (3.36), which is locally asymptotically stable. \square

3.5 Group 4

Finally, the multi-binding models, Model-M, Model-MR and Model-MC, are mathematically analyzed. We deal with only Model-M in this section because Model-M, Model-MR and Model-MC only differ from the use of a dsRNA length dependent binding rate for K_1 in the version MR and MC. As the dsRNA length is only a parameter, the nature of the system does not change. In the ODE system of Model-M, (2.12), the equation $\frac{dP}{dt}$ only depends on D_i , so it is decoupled.

$$\begin{aligned}
\frac{dE}{dt} &= -K_1ER + k_{1m}D_1 + \sum_{i=1}^{n-1} [k_{(i+1)bm}D_{i+1} - k_{(i+1)b}D_iE], \\
\frac{dR}{dt} &= -K_1ER + k_{1m}D_1, \\
\frac{dD_1}{dt} &= K_1ER - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E, \\
\frac{dD_j}{dt} &= k_{jb}D_{j-1}E - k_{jbm}D_j + k_{(j+1)bm}D_{j+1} - k_{(j+1)b}D_jE, \\
j &\in \{2, 3, \dots, n-1\}, \\
\frac{dD_n}{dt} &= k_{nb}D_{n-1}E - k_{nbm}D_n.
\end{aligned} \tag{3.45}$$

Note that (3.45), $\frac{dE}{dt} + \frac{dD_1}{dt} + 2\frac{dD_2}{dt} + \dots + n\frac{dD_n}{dt} = 0$ and $\frac{dR}{dt} + \frac{dD_1}{dt} + \frac{dD_2}{dt} + \dots + \frac{dD_n}{dt} = 0$, so by integrating with respect to time t , we obtain:

$$E(t) + \sum_{i=1}^n (iD_i(t)) = c_O, \quad R(t) + \sum_{i=1}^n (D_i(t)) = c_R, \tag{3.46}$$

for all $t \geq 0$. These two equations are referred as the conservation laws of Model-M.

Theorem 3.5.1. *ODE system (3.45) has a unique positive equilibrium $(E^*, R^*, D_1^*, \dots, D_n^*)$ with E^* being the unique positive solution of*

$$b_n E^{n+1} + E^n (b_{n-1} + b_n(nc_R - c_O)) + \dots + E(1 + b_1(c_R - c_O)) - c_O = 0,$$

$$\text{where } b_n = \prod_{i=1}^n a_i \text{ with } a_1 = \frac{K_1}{k_{1m}} \text{ and } a_l = \frac{k_{lb}}{k_{lmb}}, l \in \{2, 3, \dots, n\},$$

$$R^* = \frac{c_R}{1 + b_1 E^* + b_2 E^{*2} + \dots + b_n E^{*n}},$$

$$D_i = \frac{c_R b_i E^{*i}}{1 + b_1 E^* + b_2 E^{*2} + \dots + b_n E^{*n}}, \quad i \in \{1, 2, \dots, n\}. \quad (3.47)$$

Proof. Let $\frac{dE}{dt} = \frac{dR}{dt} = \frac{dD_j}{dt} = 0$, $j \in \{1, 2, \dots, n\}$ in (3.45). Then, we obtain:

$$0 = -K_1 E^* R^* + k_{1m} D_1^* + \sum_{i=1}^{n-1} [k_{(i+1)bm} D_{i+1}^* - k_{(i+1)b} D_i^* E^*],$$

$$0 = -K_1 E^* R^* + k_{1m} D_1^*, \quad (3.48a)$$

$$0 = K_1 E^* R^* - k_{1m} D_1^* + k_{2bm} D_2^* - k_{2b} D_1^* E^*, \quad (3.48b)$$

$$0 = k_{jb} D_{j-1}^* E^* - k_{jbm} D_j^* + k_{(j+1)bm} D_{j+1}^* - k_{(j+1)b} D_j^* E^*, \quad j \in \{2, 3, \dots, n-1\},$$

$$0 = k_{nb} D_{n-1}^* E^* - k_{nbm} D_n^*.$$

$\frac{K_1}{k_{1m}} E^* R^* = D_1^*$ is obtained from (3.48a). By substituting (3.48a) into (3.48b),

we get

$$-k_{2b}D_1^*E^* + k_{2bm}D_2^* = 0 \Rightarrow \frac{k_{2b}}{k_{2bm}}E^*D_1^* = D_2^*.$$

Similarly we get the following equations,

$$\frac{K_1}{k_{1m}}E^*R^* = D_1^*, \quad \frac{k_{2b}}{k_{2bm}}E^*D_1^* = D_2^*, \quad \dots, \quad \frac{k_{nb}}{k_{nbm}}E^*D_{n-1}^* = D_n^*.$$

Let $\frac{K_1}{k_{1m}} = a_1$, $\frac{k_{2b}}{k_{2bm}} = a_2$, \dots , $\frac{k_{nb}}{k_{nbm}} = a_n$. Then

$$a_1E^*R^* = D_1^*,$$

$$a_2E^*D_1^* = D_2^* = a_2a_1E^{*2}R^* = D_2^*,$$

$$\vdots$$

$$a_n a_{n-1} \dots a_2 a_1 E^{*n} R^* = D_n^*.$$

Let $\prod_{i=1}^n a_i = b_n$. Then we get equations for D_j^* as follows:

$$D_j^* = \prod_{i=1}^j a_i E^{*j} R^* = b_j E^{*j} R^*. \quad (3.49)$$

where $1 \leq j \leq n$.

The following two equations are valid by conservation laws of Model-M:

$$E^* + \sum_{i=1}^n (iD_i^*) = c_O, \quad R^* + \sum_{i=1}^n (D_i^*) = c_R. \quad (3.50)$$

By substituting (3.49) into (3.50), we get

$$E^* + b_1 E^* R^* + 2b_2 E^{*2} R^* + \dots + nb_n E^{*n} R^* = c_O, \quad (3.51a)$$

$$R^* + b_1 E^* R^* + b_2 E^{*2} R^* + \dots + b_n E^{*n} R^* = c_R. \quad (3.51b)$$

From (3.51b), $R^* = \frac{c_R}{1+b_1E^*+b_2E^{*2}+\dots+b_nE^{*n}}$ and it is plugged into (3.51a). Then we get the following equation for E^* :

$$E^* + R^* (b_1E^* + 2b_2E^{*2} + \dots + nb_nE^{*n}) = c_O,$$

$$E^* + \frac{c_R}{1 + b_1E^* + b_2E^{*2} + \dots + b_nE^{*n}} (b_1E^* + 2b_2E^{*2} + \dots + nb_nE^{*n}) = c_O.$$

Hence, E^* satisfies the following polynomial equation:

$$b_nE^{*n+1} + E^{*n} (b_{n-1} + b_n(nc_R - c_O)) + \dots + E^* (1 + b_1(c_R - c_O)) - c_O = 0. \quad (3.53)$$

The coefficient of the term of degree 0 is $-c_O < 0$; in the experiments $c_R > c_O$ and recall that n is a positive integer. Hence, there is only one change in sign in the coefficients of the polynomial. Therefore, by the Descartes' rule of signs, we can conclude there is only one positive real solution E^* to the polynomial equation (3.53). Thus we get the unique positive equilibrium (3.47). \square

The system of n multi-binding model, (3.45), has $n + 2$ equations; however, by using the conservations of Model-M, the system can be reduced to the

following n -dimensional ODE system.

$$\begin{aligned} \frac{dD_1}{dt} &= K_1 \left(c_O - \sum_{i=1}^n (iD_i) \right) \left(c_R - \sum_{i=1}^n (D_i) \right) - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1 \left(c_O - \sum_{i=1}^n (iD_i) \right), \\ \frac{dD_j}{dt} &= k_{jb}D_{j-1} \left(c_O - \sum_{i=1}^n (iD_i) \right) - k_{jbm}D_j + k_{(j+1)bm}D_{j+1} - k_{(j+1)b}D_j \left(c_O - \sum_{i=1}^n (iD_i) \right), \\ & j \in \{2, 3, \dots, n-1\}, \\ \frac{dD_n}{dt} &= k_{nb}D_{n-1} \left(c_O - \sum_{i=1}^n (iD_i) \right) - k_{nbm}D_n. \end{aligned} \tag{3.54}$$

In this thesis, the cases, $n = 1$, $n = 2$ and $n = 3$, are only investigated because the longest length of dsRNA in the experimental data is 120bp that allows maximum 3 OAS2 to bind to a dsRNA. The case $n = 1$ is Model-E, so the analyses of double ($n = 2$) and triple ($n = 3$) binding of OAS2 to dsRNA are only carried out here.

Based on (3.54) and the conservation laws, the reduced ODE system of Model-M with $n = 2$ is

$$\begin{aligned} \frac{dD_1}{dt} &= K_1(c_O - D_1 - 2D_2)(c_R - D_1 - D_2) - k_{1m}D_1 \\ & \quad + k_{2bm}D_2 - k_{2b}D_1(c_O - D_1 - 2D_2), \\ \frac{dD_2}{dt} &= k_{2b}D_1(c_O - D_1 - 2D_2) - k_{2bm}D_2. \end{aligned} \tag{3.55}$$

Corollary 3.5.1.1. *Model-M with $n = 2$ has a unique positive equilibrium*

(E^*, R^*, D_1^*, D_2^*) with E^* being the unique positive solution of

$$b_2 E^3 + E^2(b_1 + b_2(2c_R - c_O)) + E(1 + b_1(c_R - c_O)) - c_O = 0,$$

$$\text{where } b_2 = a_1 a_2 \text{ with } b_1 = a_1 = \frac{K_1}{k_{1m}} \text{ and } a_2 = \frac{k_{2b}}{k_{2bm}},$$

$$R^* = \frac{c_R}{1 + b_1 E^* + b_2 E^{*2}}, \tag{3.56}$$

$$D_i^* = \frac{c_R b_i E^{*i}}{1 + b_1 E^* + b_2 E^{*2}}, \quad i \in \{1, 2\}.$$

and the equilibrium is globally asymptotically stable.

Proof. The equilibrium of Model-M with $n = 2$ is obtained by substituting $n = 2$ to (3.47). To determine the stability, we get the Jacobian, J_{M_2} , of system (3.55):

$$J_{M_2} = \begin{bmatrix} A_1 & A_2 \end{bmatrix},$$

where

$$A_1 = \begin{bmatrix} -K_1(c_O - D_1 - 2D_2) - K_1(c_R - D_1 - D_2) - k_{1m} - k_{2b}(c_O - D_1 - 2D_2) + k_{2b}D_1 \\ k_{2b}(c_O - D_1 - 2D_2) - k_{2b}D_1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -K_1(c_O - D_1 - 2D_2) - 2K_1(c_R - D_1 - D_2) + k_{2bm} + 2k_{2b}D_1 \\ -2k_{2b}D_1 - k_{2bm} \end{bmatrix}.$$

Note that, from (3.46), $c_O - D_1(t) - 2D_2(t) = E(t)$ and $c_R - D_1(t) - D_2(t) = R(t)$ for all $t \geq 0$ so $c_O - D_1^* - 2D_2^* = E^* > 0$ and $c_R - D_1^* - D_2^* = R^* > 0$. By

using $c_O - D_1^* - 2D_2^* = E^*$ and $c_R - D_1^* - D_2^* = R^*$ to determine the sign of determinant and trace of J_{M_2} more easily, we obtain the following Jacobian evaluated at the equilibrium:

$$J_{M_2}(D_1^*, D_2^*) = \begin{bmatrix} -K_1E^* - K_1R^* - k_{1m} - k_{2b}E^* + k_{2b}D_1^* & -K_1E^* - 2K_1R^* + k_{2bm} + 2k_{2b}D_1^* \\ k_{2b}E^* - k_{2b}D_1^* & -2k_{2b}D_1^* - k_{2bm} \end{bmatrix}.$$

Then we get the determinant of J_{M_2} :

$$\begin{aligned} \det(J_{M_2}(D_1^*, D_2^*)) &= (K_1E^* + K_1R^* + k_{1m})(2k_{2b}D_1^* + k_{2bm}) + (K_1E^* + 2K_1R^*)(k_{2b}E^* - k_{2b}D_1^*) \\ &= k_{2bm}(K_1E^* + K_1R^* + k_{1m}) + k_{2b}E^*(K_1E^* + 2K_1R^*) + k_{2b}D_1^*(K_1E^* + 2k_{1m}) > 0, \end{aligned}$$

and the trace of J_{M_2} :

$$\text{Tr}(J_{M_2}(D_1^*, D_2^*)) = -(K_1E^* + K_1R^* + k_{2b}E^* + k_{2b}D_1^* + k_{1m} + k_{2bm}) < 0,$$

since $E^* > 0$, $R^* > 0$ and $D_1^* > 0$. Thus the equilibrium of (3.55) is locally asymptotically stable.

To determine the global stability, Bendixson's criterion and the Poincaré-Bendixson trichotomy are used. Since $E(t) \geq 0$, $R(t) \geq 0$ and $D_1(t) \geq 0$ for all $t \geq 0$, we have

$$\begin{aligned} \frac{\partial \left(\frac{dD_1}{dt} \right)}{\partial D_1} + \frac{\partial \left(\frac{dD_2}{dt} \right)}{\partial D_2} &= -(K_1E^* + K_1R^* + k_{2b}E^* + k_{2b}D_1^* + k_{1m} + k_{2bm}) < 0, \forall t \geq 0. \end{aligned}$$

Thus, there is no periodic solutions in system (3.55). Moreover, there is only one equilibrium and the solutions are bounded by the conservation laws of Model-M, (3.46), so the equilibrium of (3.55) is globally asymptotically stable by the Poincaré-Bendixson trichotomy.

Since $c_O - D_1^* - 2D_2^* = E^*$ and $c_R - D_1^* - D_2^* = R^*$, the equilibrium (3.56) of Model-M with $n = 2$ is globally asymptotically stable. \square

From (3.54) and conservation laws, the reduced ODE system of Model-M with $n = 3$ is

$$\begin{aligned}
\frac{dD_1}{dt} &= K_1(c_O - D_1 - 2D_2 - 3D_3)(c_R - D_1 - D_2 - D_3) - k_{1m}D_1 \\
&\quad + k_{2bm}D_2 - k_{2b}D_1(c_O - D_1 - 2D_2 - 3D_3), \\
\frac{dD_2}{dt} &= k_{2b}D_1(c_O - D_1 - 2D_2 - 3D_3) - k_{2bm}D_2 + k_{3bm}D_3 \\
&\quad - k_{3b}D_2(c_O - D_1 - 2D_2 - 3D_3), \\
\frac{dD_3}{dt} &= k_{3b}D_2(c_O - D_1 - 2D_2 - 3D_3) - k_{3bm}D_3.
\end{aligned} \tag{3.57}$$

Corollary 3.5.1.2. *ODE system (3.57) has a unique positive equilibrium*

$(E^*, R^*, D_1^*, D_2^*, D_3^*)$ with E^* being the unique positive solution of
 $b_3E^4 + E^3(b_2 + b_3(3c_R - c_O)) + E^2(b_1 + b_2(2c_R - c_O)) + E(1 + b_1(c_R - c_O)) - c_O = 0$,

where $b_n = \prod_{i=1}^n a_i$ with $a_1 = \frac{K_1}{k_{1m}}$ and $a_l = \frac{k_{lb}}{k_{l bm}}$, $l \in \{2, 3\}$,

$$R^* = \frac{c_R}{1 + b_1E^* + b_2E^{*2} + b_3E^{*3}},$$

$$D_i^* = \frac{c_R b_i E^{*i}}{1 + b_1E^* + b_2E^{*2} + b_3E^{*3}}, \quad i \in \{1, 2, 3\}.$$

(3.58)

and the equilibrium is locally asymptotically stable.

Proof. The equilibrium of Model-M with $n = 3$ is obtained by substituting $n = 3$ to (3.47). For the stability, from system (3.57), we get the Jacobian J_{M_3} evaluated at the equilibrium.

$$J_{M_3}(D_1^*, D_2^*, D_3^*) = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix},$$

where

$$A_1 = \begin{bmatrix} -K_1(c_O - D_1^* - 2D_2^* - 3D_3^*) - K_1(c_O - D_1^* - D_2^* - D_3^*) - k_{1m} - k_{2b}(c_O - D_1^* - 2D_2^* - 3D_3^*) + k_{2b}D_1^* \\ k_{2b}(c_O - D_1^* - 2D_2^* - 3D_3^*) - k_{2b}D_1^* + k_{3b}D_2^* \\ -k_{3b}D_2^* \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -K_1(c_O - D_1^* - 2D_2^* - 3D_3^*) - 2K_1(c_O - D_1^* - D_2^* - D_3^*) + k_{2bm} + 2k_{2b}D_1^* \\ -2k_{2b}D_1^* - k_{2bm} - k_{3b}(c_O - D_1^* - 2D_2^* - 3D_3^*) + 2k_{3b}D_2^* \\ k_{3b}(c_O - D_1^* - 2D_2^* - 3D_3^*) - 2k_{3b}D_2^* \end{bmatrix},$$

$$A_3 = \begin{bmatrix} -K_1(c_O - D_1^* - 2D_2^* - 3D_3^*) - 3K_1(c_O - D_1^* - D_2^* - D_3^*) + 3k_{2b}D_1^* \\ -3k_{2b}D_1^* + k_{3bm} + 3k_{3b}D_2^* \\ -3k_{3b}D_2^* - k_{3bm} \end{bmatrix}.$$

Note that, from (3.46), $c_O - D_1(t) - 2D_2(t) - 3D_3(t) = E(t)$ and $c_R - D_1(t) - D_2(t) - D_3(t) = R(t)$ for all $t \geq 0$ so $c_O - D_1^* - 2D_2^* - 3D_3^* = E^*$ and $c_R - D_1^* - D_2^* - D_3^* = R^*$. We replace $c_O - D_1^* - 2D_2^* - 3D_3^* = E^* > 0$ and $c_R - D_1^* - D_2^* - D_3^* = R^* > 0$ to determine the sign of coefficients of characteristic polynomial of J_{M_3} more easily. Hence, the Jacobian is now

$$J_{M_3}(D_1^*, D_2^*, D_3^*) = \begin{bmatrix} -K_1E^* - K_1R^* - k_{1m} - k_{2b}E^* + k_{2b}D_1^* & -K_1E^* - 2K_1R^* + k_{2bm} + 2k_{2b}D_1^* & -K_1E^* - 3K_1R^* + 3k_{2b}D_1^* \\ k_{2b}E^* - k_{2b}D_1^* + k_{3b}D_2^* & -2k_{2b}D_1^* - k_{2bm} - k_{3b}E^* + 2k_{3b}D_2^* & -3k_{2b}D_1^* + k_{3bm} + 3k_{3b}D_2^* \\ -k_{3b}D_2^* & k_{3b}E^* - 2k_{3b}D_2^* & -3k_{3b}D_2^* - k_{3bm} \end{bmatrix},$$

and its characteristic polynomial is

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0,$$

where

$$a_1 = (K_1 + k_{2b} + k_{3b})E^* + K_1R^* + k_{2b}D_1^* + k_{3b}D_2^* + k_{1m} + k_{2bm} + k_{3bm},$$

$$a_2 = (K_1k_{2b} + K_1k_{3b} + k_{2b}k_{3b})E^{*2}$$

$$+ ((2k_{2b} + k_{3b})K_1R^* + (K_1 + 2k_{3b})k_{2b}D_1^* + (K_1 + k_{2b})k_{3b}D_2^*$$

$$+ K_1(k_{2bm} + k_{3bm}) + k_{1m}k_{3b} + k_{2b}k_{3bm})E^*$$

$$+ (k_{2bm} + k_{3bm})K_1R^* + (2k_{1m} + k_{3bm})k_{2b}D_1^* + (k_{1m} + 2k_{2bm})k_{3b}D_2^*$$

$$+ k_{1m}k_{2bm} + k_{1m}k_{3bm} + k_{2bm}k_{3bm},$$

$$a_3 = (K_1k_{2b}k_{3b})E^{*3} + (3K_1k_{2b}k_{3b}R^* + 2K_1k_{2b}k_{3b}D_1^* + K_1k_{2b}k_{3b}D_2^* + K_1k_{2b}k_{3bm})E^{*2}$$

$$+ (2K_1k_{2b}k_{3bm}R^* + (K_1k_{3bm} + 3k_{1m}k_{3b})k_{2b}D_1^* + 2K_1k_{2bm}k_{3b}D_2^* + K_1k_{2bm}k_{3bm})E^*$$

$$+ K_1k_{2bm}k_{3bm}R^* + 2k_{1m}k_{2b}k_{3bm}D_1^* + 3k_{1m}k_{2bm}k_{3b}D_2^* + k_{1m}k_{2bm}k_{3bm}.$$

Since E^* , R^* , D_1^* , D_2^* , K_1 , k_{2b} , k_{3b} , k_{1m} , k_{2bm} , and k_{3bm} are positive, $a_1 > 0$ and $a_3 > 0$. All the terms of a_3 also can be made by multiplication of parts of a_1 and a_2 and only positive terms are remained, since all the terms of a_1 and a_2 are positive. Hence, we satisfy $a_1a_2 > a_3$.

By using Routh-Hurwitz criterion, we know that the unique equilibrium of (3.57) is locally asymptotically stable. Then, we get the equilibrium (3.58) of Model-M with $n = 3$ by the conservation laws of Model-M; this equilibrium is locally asymptotically stable. \square

3.6 About the Product

In this chapter, we have obtained the exact solution of Model-S, quasi equilibrium approximation of Model-N and equilibria of Model-E, Model-ED, Model-N and Model-M. Based on these analyses, we summarize the dynamics of $P(t)$ for each model and illustrate the behaviour of the product here.

First, Model-S has the unique solution, $P(t) = k_p c_O c_R t$, for all $t \geq 0$; $P(t)$ is linearly proportional to k_p , c_O , c_R and t .

Assume that t is sufficiently large, $t > T$, then all systems $\frac{dX}{dt}$ for group 2 to 4 approach an equilibrium. Integrating with respect to time, $\frac{dP}{dt} = g(D^*, D_j^*)$ with $j \in \{1, \dots, n\}$ and $P(T) = \rho_0$, where $\rho_0 \geq 0$, an expression of the production of P is obtained as a function of time for each model and assumption.

First, the product under the assumption of Model-E follows:

$$P(t) = \frac{k_p}{2} \left((c_O + c_R) + \frac{1}{b_1} - \sqrt{\left(c_O + c_R + \frac{1}{b_1} \right)^2 - 4c_O c_R} \right) (t - T) + \rho_0, \quad (3.59)$$

for all $t > T$, where $b_1 = \frac{K_1}{k_{1m}}$, so $P(t)$ linearly increases with time t . Replacing K_1 with $k_1 \sigma_{1,2}(L)$ for Model-ER and Model-EC gives

$$P(t) = \frac{k_p}{2} \left((c_O + c_R) + \frac{1}{b_1 \sigma_{1,2}(L)} - \sqrt{\left(c_O + c_R + \frac{1}{b_1 \sigma_{1,2}(L)} \right)^2 - 4c_O c_R} \right) (t - T) + \rho_0, \quad (3.60)$$

where $b_1 = \frac{k_1}{k_{1m}}$, for all $t > T$. The product $P(t)$ increases as $\sigma_{1,2}(L)$ increases when L gets larger since $\frac{dP}{d\sigma_{1,2}(L)} > 0$ for all $t \geq 0$ and $\frac{d\sigma_{1,2}(L)}{dL} > 0$ for all $L > 2R_O$. Hence, $P(t)$ well reflects the relation between the length of dsRNA and the amount of product: if the length of dsRNA increases, the amount of product increases.

When considering non-productive bindings in Model-N, the product follows

$$P(t) = \frac{k_p b_2}{2(1+b_2+\frac{b_5}{b_1})} \left((c_O+c_R) + \frac{1}{b_1(1+b_2+\frac{b_5}{b_1})} - \sqrt{\left((c_O+c_R) + \frac{1}{b_1(1+b_2+\frac{b_5}{b_1})} \right)^2 - 4c_O c_R} \right) (t-T) + \rho_0, \quad (3.61)$$

for all $t > T$, with the association constants $b_1 = \frac{K_1}{k_{1m}}$, $b_2 = \frac{k_2}{k_{2m}}$ and $b_5 = \frac{k_5}{k_{5m}}$. By differentiating with respect to b_i , we obtain $\frac{dP}{db_1} > 0$, $\frac{dP}{db_2} > 0$ and $\frac{dP}{db_5} < 0$. Hence, increasing the association constants of the two-step bindings will promote the production whereas the more non-productive stages the less product.

We also have assumed a quasi-steady-state in Model-N, which hypothesizes that the enzyme is very diligent and $\frac{dE}{dt} \approx 0$ in $t > \epsilon$ with $0 < \epsilon \ll 1$. Under this assumption, we obtained three different forms for the solution, (3.25), (3.26) and (3.27), depending on parameter values. For $t > t_0 > \epsilon$, the contributions to solutions associated to negative eigenvalues or with negative real part vanish. Hence, for $t > t_0$, the three forms of solutions overlap and correspond to the contribution associated with the zero eigenvalue. Hence,

$$P(t) \approx \frac{k_p b_2}{\left(1 + b_2 + \frac{b_5}{b_1}\right)} (\epsilon_C + \epsilon_N)t, \quad \forall t \geq t_0,$$

where $b_1 = \frac{K_1}{k_{1m}}$, $b_2 = \frac{k_2}{k_{2m}}$ and $b_5 = \frac{k_5}{k_{5m}}$ are association constants and ϵ_C and ϵ_N are concentrations of intermediate stages C and N at $t = \epsilon$. Similar to (3.61), we have $\frac{dP}{db_1} > 0$, $\frac{dP}{db_2} > 0$ and $\frac{dP}{db_5} < 0$.

In the long run, the two equations for P in Model-E (3.59) and Model-N (3.61) are similar. However, in (3.61), we have b_5 such that $\frac{dP}{db_5} < 0$ for all $t \geq 0$. Hence, since b_5 is the association constant for the non-productive binding stages, the non-productive bindings play a role of inhibitor in the OAS2 activation by making the process slower.

Model-ED1, which assumes the degradation of enzyme, Model-ED2, which assumes the degradation of complex, and Model-ED3, which assumes the degradations of both enzyme and complex, have a unique equilibrium, $D^* = 0$. Hence, for models with degradations, $P(t) = \rho_0$ as $t > T$. Thus, the concentration of P will be a constant as $t > T$. From Model-ED1, Model-ED2 and Model-ED3, we know that if there is degradation of enzyme, complex or both, then the concentration of the product will reach a plateau after some time.

When multiple bindings are considered as in Model-M, the product follows:

$$\begin{aligned} P(t) &= k_p \sum_{i=1}^n i D_i^*(t - T) + \rho_0 = k_p E^* R^* (b_1 + 2b_2 E^* + \cdots + n b_n E^{*n-1}) (t - T) + \rho_0 \\ &= k_p a_1 E^* R^* \left(1 + 2a_2 E^* + \cdots + n \prod_{i=2}^n a_i E^{*n-1} \right) (t - T) + \rho_0, \end{aligned}$$

for all $t > T$, where, $\prod_{i=1}^n a_i = b_n$, $a_1 = \frac{K_1}{k_{1m}}$ and $a_l = \frac{k_{lb}}{k_{lbm}}$, $l \in \{2, 3, \dots, n\}$.

Constants $b_i = \frac{K_1 k_{2b} \dots k_{ib}}{k_{1m} k_{2bm} \dots k_{ibm}}$, $i \in \{1, 2, \dots, n\}$ are the product of the association constants (a_i) of multiple subsequent OAS2 bindings to the same dsRNA. They can be interpreted as chance of having i OAS2 attached to the same dsRNA. If $\frac{k_{ib}}{k_{lbm}} > 1$, then the total amount of the product increases; if $\frac{k_{ib}}{k_{lbm}} < 1$, the total amount of the product decreases. By the definition of b_i , the association constant of the first binding (a_1) is the most influential of the contributions to the dynamics of the product.

Chapter 4

Numerical Investigations

In this chapter, we fit responses of models introduced in Chapter 2 to experimental data from McKenna's lab to parametrize the models. After fitting every model to data, we use Akaike criterion and weights to compare the models and select the most suitable model to represent the experimental data. In the process, the shortcomings and complements of each model are explained highlighting the development process of the collection of models.

First, the fitting procedures are explained and fitting results are then discussed for each model. Two cost functions are used in this chapter. Both cost functions are based on the square sum of differences between the predicted concentrations of models and the observed concentrations of the experimental data. Let the square sum of the differences be called RSS (the residual sum of squares).

The first cost function is calculated for a given length, l_k , and all five concentrations of dsRNA at all five time points: the parameters are estimated on 25

observation points:

$$RSS_{l_k}(\mathbf{p}) = \sum_{i=1}^5 \left(\sum_{j=1}^5 (P_{c_j, l_k}(\mathbf{p}, t_i) - D_{t_i, c_j, l_k})^2 \right), \quad (4.1)$$

where \mathbf{p} is the parameter set, $\mathbf{p} = (K_1 \text{ or } k_1, k_{1m}, \dots, k_p)$, $P_{c_j, l_k}(\mathbf{p}, t_i)$ is the predicted concentration of the product (solution of model) at time t_i , for a concentration of dsRNA $c_j = c_R$ and the length of dsRNA $l_k = L$, D_{t_i, c_j, l_k} is the experimental concentration of the product at t_i , c_j and l_k . Recall that $t_i \in \{t_1 = 5\text{min}, t_2 = 10\text{min}, t_3 = 15 \text{ min}, t_4 = 20\text{min}, t_5 = 30\text{min}\}$ and $c_j \in \{c_1 = 0.5\mu\text{g/mL}, c_2 = 1\mu\text{g/mL}, c_3 = 2\mu\text{g/mL}, c_4 = 4\mu\text{g/mL}, c_5 = 8\mu\text{g/mL}\}$.

The second cost function quantifies the errors for all five concentrations and seven lengths of dsRNA at all five time points. Parameters are estimated over 175 observation points. The second cost function is:

$$RSS(\mathbf{p}) = \sum_{k=1}^7 RSS_{l_k}(\mathbf{p}) = \sum_{k=1}^7 \left(\sum_{i=1}^5 \left(\sum_{j=1}^5 (P_{c_j, l_k}(\mathbf{p}, t_i) - D_{t_i, c_j, l_k})^2 \right) \right), \quad (4.2)$$

where \mathbf{p} is the parameter set, $\mathbf{p} = (K_1 \text{ or } k_1, k_{1m}, \dots, k_p)$, $P_{c_j, l_k}(\mathbf{p}, t_i)$ is the predicted concentration of the product at t_i , $c_j = c_R$ and $l_k = L$, D_{t_i, c_j, l_k} is the experimental concentration of the product at t_i , c_j and l_k . Recall that $l_k \in \{l_1 = 40\text{bp}, l_2 = 50\text{bp}, l_3 = 60\text{bp}, l_4 = 70\text{bp}, l_5 = 80\text{bp}, l_6 = 90\text{bp}, l_7 = 120\text{bp}\}$.

To find the estimates $\hat{\mathbf{p}}$ of parameter values, we solve $RSS(\hat{\mathbf{p}}) = \min_{\mathbf{p}} RSS(\mathbf{p})$.

This optimization problem is solved using a genetic algorithm developed in

the R package **GA**. The two cost functions return RSS for the parameter values \mathbf{p} . The package **GA** chooses the best parameter set to make RSS as small as possible among candidates of parameter sets by stochastic optimisation.

The package **GA** uses several genetic operators to find better parameter sets than the previous generation's best parameter set. For this genetic algorithm, we choose the following options: `popSize=100`, `pcrossover=0.5`, `pmutation=0.5`, `maxiter=300`, `run=100`, `lower=0` and `upper=50`. In one generation, the program chooses 100 random candidates of parameter set (`popSize=100`) to find the best set of that generation. The probability of crossover is 0.8 and the probability of mutation is 0.1 as a default setting. However, we use the probability of crossover (`pcrossover=0.5`) and mutation (`pmutation=0.5`) to find more suitable parameter sets, which are not located in the neighborhood of the previous best set. We set 300 generations (`maxiter=300`) based on some test simulations, 300 is the best values for the number of generations. If the best set does not change in 100 generations, the simulation is stopped (`run=100`). Moreover, we set upper bounds for parameters, `upper=50`. All the parameters are positive, so the lower limit of parameters is set to 0 (`lower=0`).

One simulation returns one parameter set, which is the best parameters having the smallest RSS through 300 generations. We carry out 50 independent simulations to get 50 parameter sets, so we compute the average, standard deviation (SD) and interval for reaction rates and affinities. We call “the best parameter set” the set of parameter allowing the smallest RSS among the 50 results.

4.1 Models versus experimental data

4.1.1 Model-S

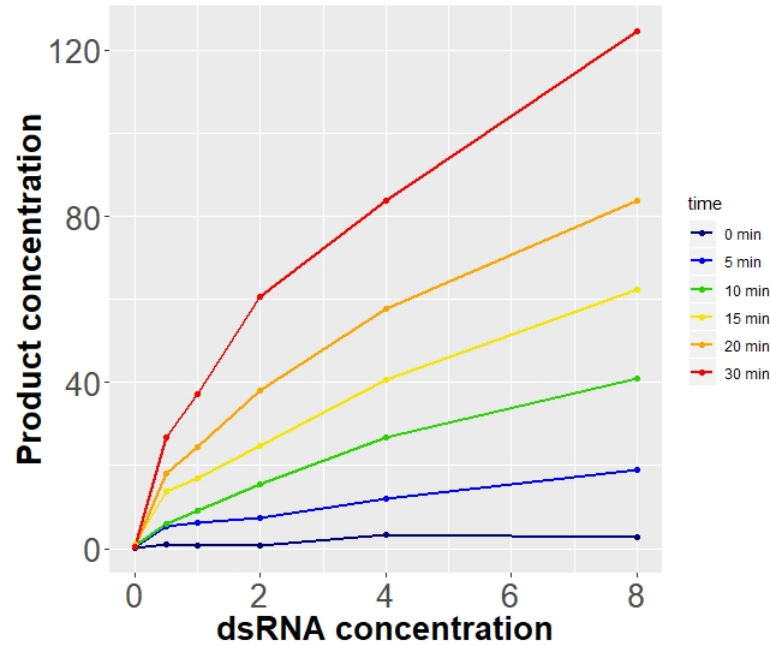


Figure 4.1: The experimental data: concentrations of product as a function of c_R with fixed t .

In the previous chapter for Model-S, we found that the solution for $P(t)$ (3.2) is not only a linear function of t but also a linear function of c_R if t is fixed and c_R is varying. However, for different times, Figure 4.1 shows that, in the experimental data, the change of the concentration of P is not linear in c_R . Thus, Model-S is not a proper model to explain the activation of OAS2 by dsRNA, as it is wrongly accommodated the effects of dsRNA concentrations, c_R .

4.1.2 Fitting: all concentrations at one length

Because the general enzyme kinetics have no length dependency, we firstly fix the length at l_k and carry out fittings for all models (Model-E, Model-ED, Model-N and Model-M) using the cost function (4.1). Since the length binding rates are computed at one length, they are dealt as constants, such that $K_1 = k_1\sigma_1(L) = k_1\sigma_2(L)$ with $L = l_k$. Hence, outputs of Model-E, Model-ER and Model-EC are the same. Same holds for Model-M, Model-MR and Model-MC. For illustration, the length l_k is chosen equal to 120bp, because it gives the most clear fitting results. Graphs of predicted concentrations, shown in Figure 4.2, are obtained with the best parameter set of each model. Table 4.1 displays values and statistics of the best parameters, affinities and RSS for the models; corresponding box-plots are given in Figure 4.4. For the interpretation of Table 4.1, we use [RSS,Model-E,Best] to represent the value at RSS column and Model-E row in Best section.

First of all, all models fit well data for all concentrations at a given length (Figure 4.2); similar results are obtained when other lengths are chosen (results not shown).

Model-E and Model-ED

Model-E, which has the smallest number of parameters, has RSS and parameters with the smallest standard deviations (SD) (see Table 4.1 and Figure 4.4): predictions of Model-E are very sensitive to parameter values. The smallest

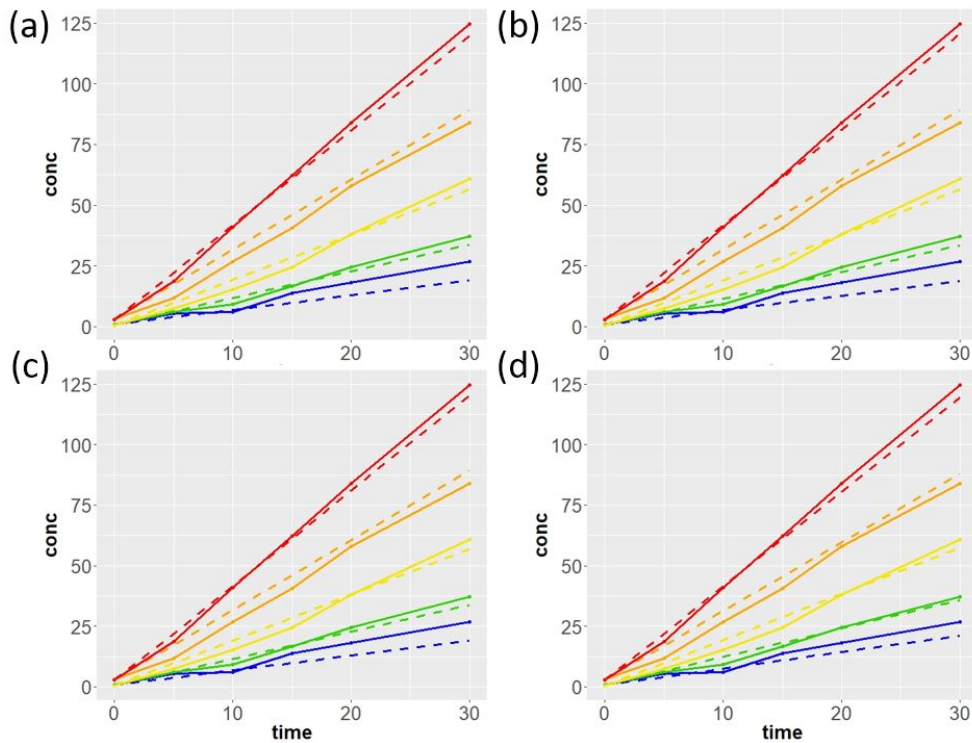


Figure 4.2: Experimental data (solid line) and prediction for the product P (dashed line) of (a) Model-E, (b) Model-ED, (c) Model-N and (d) Model-M by using all concentrations and one length (120bp) of dsRNA. Color code is given in Figure 4.3.

Concentration of dsRNA	Original data	Predicted
0.5 $\mu\text{g/mL}$		
1 $\mu\text{g/mL}$		
2 $\mu\text{g/mL}$		
4 $\mu\text{g/mL}$		
8 $\mu\text{g/mL}$		

Figure 4.3: Legend table for Figure 4.2. High to low concentrations: rainbow order.

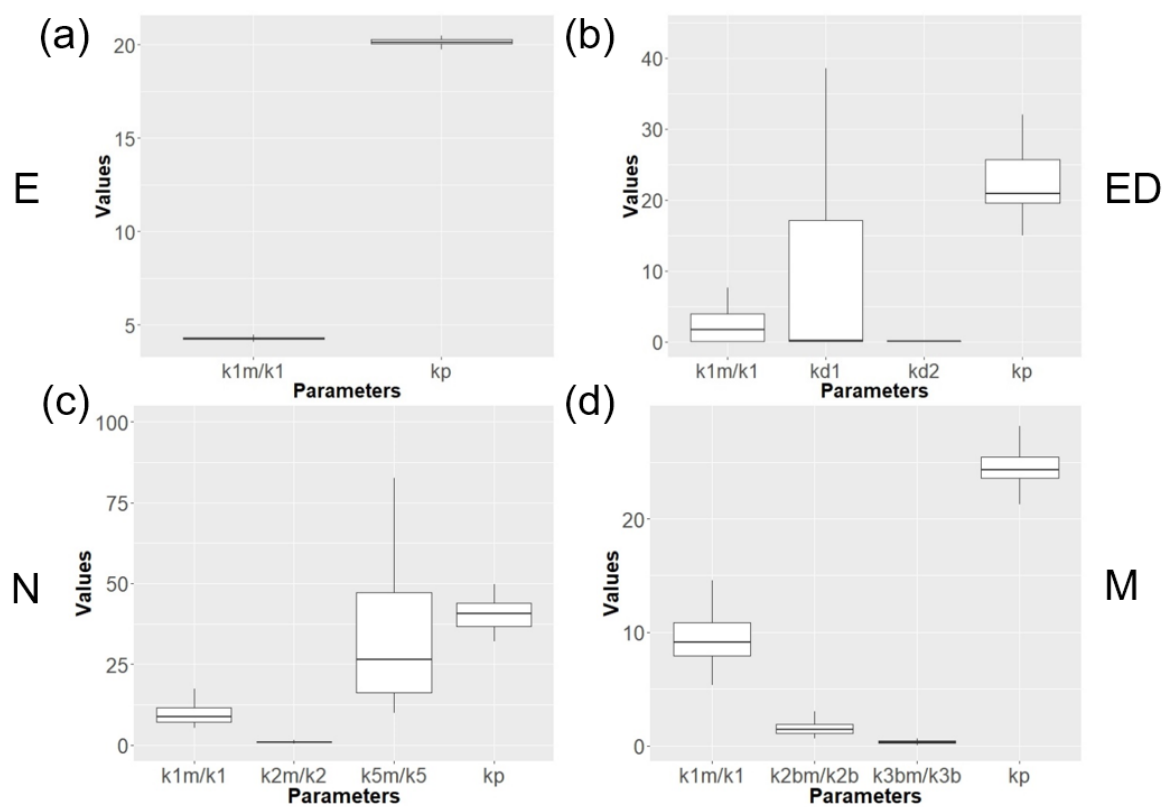


Figure 4.4: Affinities of reactions for (a) Model-E, (b) Model-ED, (c) Model-N and (d) Model-M by using all concentrations at one length (120bp) of dsRNA. Ten outliers are ignored for k_{5m}/k_5 in (c), outliers > 110 .

		k_{1m}/K_1	k_{d1}	k_{d2}	k_{2m}/k_2	k_{5m}/k_5	k_{2bm}/k_{2b}	k_{3bm}/k_{3b}	k_p	RSS
Mean	Model-E	4.278							20.159	352.942
	Model-ED	2.049	7.443	0.026					23.62	2116.951
	Model-N	9.914			0.892	124.805			40.448	394.014
	Model-M	9.81					1.642	0.344	24.574	285.475
SD	Model-E	0.087							0.198	1.219
	Model-ED	1.965	11.620	0.029					6.795	1593.305
	Model-N	4.412			0.287	350.621			4.712	66.908
	Model-M	2.876					0.835	0.156	1.762	11.009
Best	Model-E	4.253							20.108	348.39
	Model-ED	4.37	0	0					20.448	343.203
	Model-N	27.076			0.266	17.094			34.264	336.615
	Model-M	16.711					0.991	0.16	28.146	264.06

Table 4.1: Mean, standard deviation (SD) and best parameter set for Model-E, Model-ED, Model-N and Model-M obtained when fitting all concentrations at one length of dsRNA, 120bp.

values for the rate of production k_p of activated OAS2 are found for Model-E (Figure 4.4). For Model-ED, $[k_{d1}, \text{Model-ED, Best}]$ and $[k_{d2}, \text{Model-ED, Best}]$ are both zero; the best result is obtained with no degradations of OAS2 and complex within 30 minutes (Table 4.1).

Model-N

In Model-N, the affinity k_{5m}/k_5 shows the unbinding power of non-productive complex. Hence, high value of $[k_{5m}/k_5, \text{Model-N, Mean}]$ means that if OAS2 and dsRNA combine in non-productive forms, they detach instantly and then there

is no hindrance for the production. From $[k_{5m}/k_5, \text{Model-N,SD}]$, k_{5m}/k_5 seems to have little effect on model responses.

Model-M

Model-M has the smallest value in $[\text{RSS, Best}]$, a RSS of 264, which is almost 70 less than for other models' $[\text{RSS, Best}]$. Thus, the multi-binding assumption can be considered as a possible scenario for the activation of OAS2 and dsRNA. However, Model-M is the model with the largest number of parameters. Furthermore, from the affinities of the first and subsequent bindings in Figure 4.4, $\frac{k_{1m}}{K_1} > \frac{k_{2bm}}{k_{2b}} > \frac{k_{3bm}}{k_{3b}}$, we can conclude that the system exhibits positive cooperative binding when the multi-bindings are allowed [22]. The interaction of OAS2 to a dsRNA helps the binding of subsequent OAS2 to the same dsRNA.

4.1.3 Fitting: all concentrations at all lengths

By using the second cost function (4.2), we fit all the models to the data that consists of the concentrations of product obtained with all concentrations and all lengths of dsRNA at all time points. For readability, the 175 data points are split into 5 graphs and each graph displays the data for one concentration. Figure 4.5 shows the experimental data and predicted data by Model-E, Model-ED, Model-N and Model-M. The legend of the graphs is given in Figure 4.6. Table

4.2 summarizes values of parameters to obtain the best fits and corresponding RSSs of Model-E, Model-ED, Model-N and Model-M.

		k_{1m}/K_1	k_{d1}	k_{d2}	k_{2m}/k_2	k_{5m}/k_5	k_{2bm}/k_{2b}	k_{3bm}/k_{3b}	k_p	RSS
Mean	Model-E	5.727							13.126	30790.5
	Model-ED	2.155	5.847	0.051					17.056	38013.4
	Model-N	12.145			1.355	182.819			34.996	30896.8
	Model-M	27.691					0.214	0.291	24.247	10058.7
SD	Model-E	0.083							0.11	2.364
	Model-ED	2.998	11.524	0.046					6.132	5092.1
	Model-N	6.714			0.58	1010.015			7.154	139.604
	Model-M	5.413					0.023	0.05	3.344	387.416
Best	Model-E	5.642							13.066	30779.4
	Model-ED	5.357	0	0					12.71	30783.3
	Model-N	44.943			0.317	11.412			33.661	30750.1
	Model-M	35.854					0.241	0.312	29.287	9725.4

Table 4.2: Mean, standard deviation (SD) and the best parameter set for Model-E, Model-ED, Model-N and Model-M with all concentrations at all lengths of dsRNA.

Model-E

In contrast to Figure 4.2(a), the predicted data in Figure 4.5(1) does not represent the experimental data well. Recall, Model-E does not accommodate explicitly the length of dsRNA. Hence, model responses of Model-E do not change with the lengths. Even when fitting Model-E to all concentrations and lengths data, similar responses are obtained for the seven lengths (Figure 4.5(1)). Predicted data does not accommodate neither concentrations nor lengths.

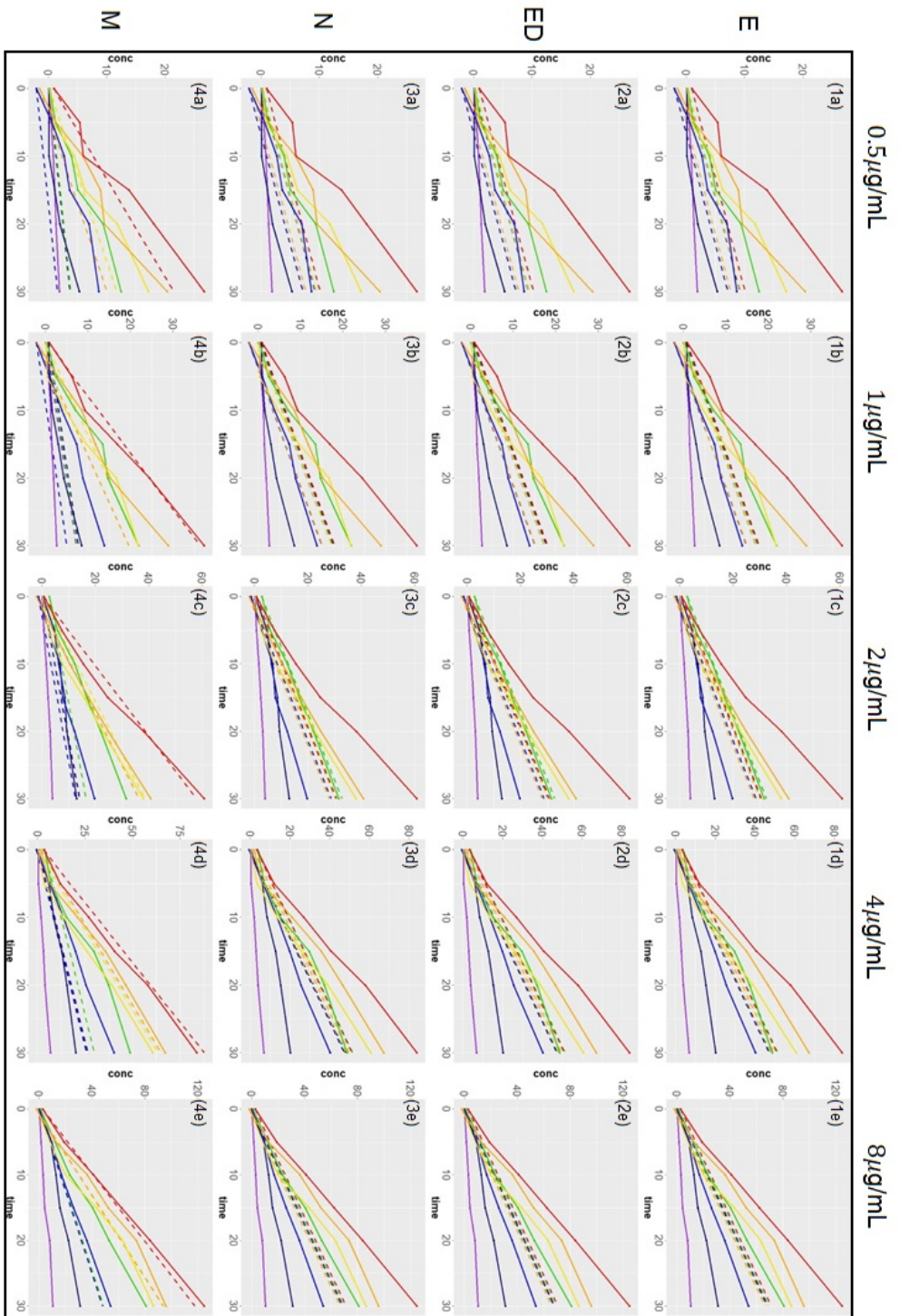


Figure 4.5: Experimental data (solid line) and prediction for the product P (dashed line) of (1) Model-E, (2) Model-ED, (3) Model-N and (4) Model-M. Each column presents one concentration of dsRNA: (a) 0.5µg/mL, (b) 1µg/mL, (c) 2µg/mL, (d) 4µg/mL and (e) 8µg/mL. Color code is given in Figure 4.6.

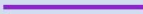
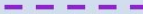




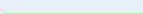

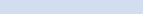
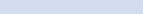
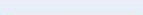

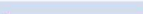
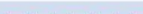
Length of dsRNA	Original data	Predicted
40bp		
50bp		
60bp		
70bp		
80bp		
90bp		
120bp		

Figure 4.6: Legend table for fitting graphs. Long to short dsRNA: rainbow order.

Model-ED

As in Model-E, Model-ED responses do not change with the lengths (Figure 4.5(2)). In Table 4.2, [Model-ED,Best] are almost the same as [Model-E,Best] and they are obtained with $k_{d1} = k_{d2} = 0$, which gives Model-E. As previously observed, we consider that the degradation of the enzyme or complex are not applicable within 30min.

Model-N

The predicted data of Model-N does not properly fit to the experimental data (Figure 4.5(3)). From these results, the two-step binding process with non-productive bindings does not accommodate the impact of dsRNA lengths on the OAS2 activation properly.

Model-M

From Figure 4.5(4), the predicted data from Model-M does not properly fit to the experimental data; however, the smallest RSS is obtained with Model-M (Table 4.2). For fitting all lengths simultaneously, Model-M consists of three systems: for the lengths 40, 50, 60 and 70, we use (2.12) with $n = 1$, for the lengths 80 and 90, (2.12) with $n = 2$, and for the 120 length, (2.12) is used with $n = 3$. However, all systems share the same parameter set that is estimated simultaneously. For example, K_1 , k_{1m} and k_p are used in all systems while k_{3b} and k_{3bm} are used only in (2.12) with $n = 3$. Due to the use of these three systems, Model-M allows three groups of response accommodating slightly the lengths (Figure 4.5(4)). The multi-binding assumption, depicting a sort of length dependency, improves the representation of data.

Accommodation of length by length dependent binding rate

The length dependent parts of binding rates, $\sigma_1(L)$ and $\sigma_2(L)$, are slightly different for a given length of dsRNA as shown in Table 4.3. $\sigma_2(L)$ is smaller than $\sigma_1(L)$ for each length. However, by focusing on the third and fifth rows in Table 4.3, we see that $\sigma_2(L)/\sigma_2(40)$ is greater than $\sigma_1(L)/\sigma_1(40)$ for every length. From these observations, we know that using $\sigma_2(L)$ gives more impact of the length on the binding rates than using $\sigma_1(L)$.

Figure 4.9 shows the experimental data and predicted data by Model-ER,

Length of dsRNA	40bp	50bp	60bp	70bp	80bp	90bp	120bp
$\sigma_1(L)$	0.3455	0.872	1.2738	1.6062	1.8955	2.1563	2.838
$\sigma_1(L)/\sigma_1(40)$	1	2.5232	3.686	4.6476	5.4848	6.2393	8.2121
$\sigma_2(L)$	0.3012	0.7802	1.1601	1.4816	1.7657	2.024	2.7054
$\sigma_2(L)/\sigma_2(40)$	1	2.5896	3.8506	4.9178	5.8604	6.7178	8.9795

Table 4.3: Length dependent parts of binding rate; $\sigma_1(L)$ and $\sigma_2(L)$

Model-EC, Model-MR and Model-MC. Table 4.4 gives the results of the four models' fitting procedures using all data (all concentrations, all lengths and all time points); corresponding box-plots are given in Figure 4.7.

Model-ER, Model-EC, Model-MR and Model-MC fit better to the experimental data their analogue using constant binding rates. Furthermore, M-series (Model-MR and Model-MC) has smaller values of RSS in [RSS,Mean] and [RSS,Best] than E-series (Model-ER and Model-EC). E-series does not represent well data obtained with small concentrations of dsRNA; in contrast, M-series fits well data for all concentrations (Figure 4.9). Note that estimates of parameters are of the same order in both E- and M-series (Table 4.4). Interestingly, $\frac{k_{1m}}{k_1\sigma_1(L)}$ in Model-ER and Model-MR and $\frac{k_{1m}}{k_1\sigma_2(L)}$ in Model-EC and Model-MC similarly decrease as dsRNA length increases (Figure 4.8). The length dependent binding rates allow the increase of the product as length increases. Furthermore, for any length L , the affinity of the first binding $\frac{k_{1m}}{k_1\sigma_{1,2}(L)}$ is always larger than the affinities of subsequent bindings for Model-MR and Model-MC (Figures 4.7, 4.8 and Table 4.4). In other words, the association constant of the second binding

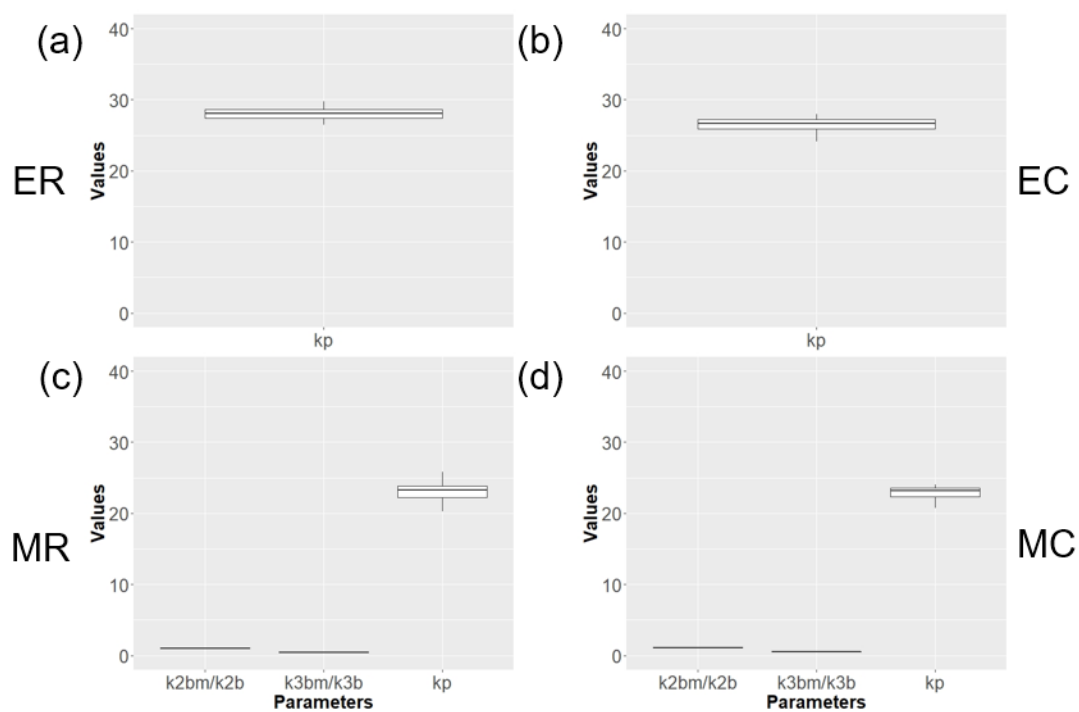


Figure 4.7: Affinities of reactions for (a) Model-ER, (b) Model-EC, (c) Model-MR and (d) Model-MC by using all concentrations at all lengths of dsRNA.

		k_{1m}/k_1	k_{2bm}/k_{2b}	k_{3bm}/k_{3b}	k_p	RSS
Mean	Model-ER	26.2699			27.987	4465.9
	Model-EC	22.6004			26.5732	4037.9
	Model-MR	24.6716	1.0297	0.4925	23.0366	2621.5
	Model-MC	22.2607	1.1071	0.5297	22.84	2498.9
SD	Model-ER	1.1401			0.8431	6.457
	Model-EC	1.2403			0.9726	11.1481
	Model-MR	2.0653	0.0659	0.092	1.3023	68.8882
	Model-MC	1.5716	0.0771	0.0555	1.0517	71.8797
Best	Model-ER	26.599			28.4526	4458.3
	Model-EC	22.9278			26.9336	4026.6
	Model-MR	23.0399	1.0931	0.311	22.2341	2577.6
	Model-MC	23.0398	1.1693	0.4892	23.4766	2464.6

Table 4.4: Mean and standard deviation (SD) of parameters of Model-ER, Model-EC, Model-MR and Model-MC

is higher than that of the first OAS2 binding. Hence, once the first OAS2 is bound to dsRNA, subsequent bindings of OAS2 to the same dsRNA are easier. Therefore, Model-MR and Model-MC exhibit positive cooperative binding of multiple OAS2 enzymes.

The E-series has only one length dependency, the length dependent binding rates. In contrast, the M-series has two length dependencies: the structure of Model-M and length dependent binding rates. As previously mentioned, Model-M structure allows us to specify three types of responses depending on dsRNA lengths; then, the accommodation of length is more refined through the length dependent binding rates.

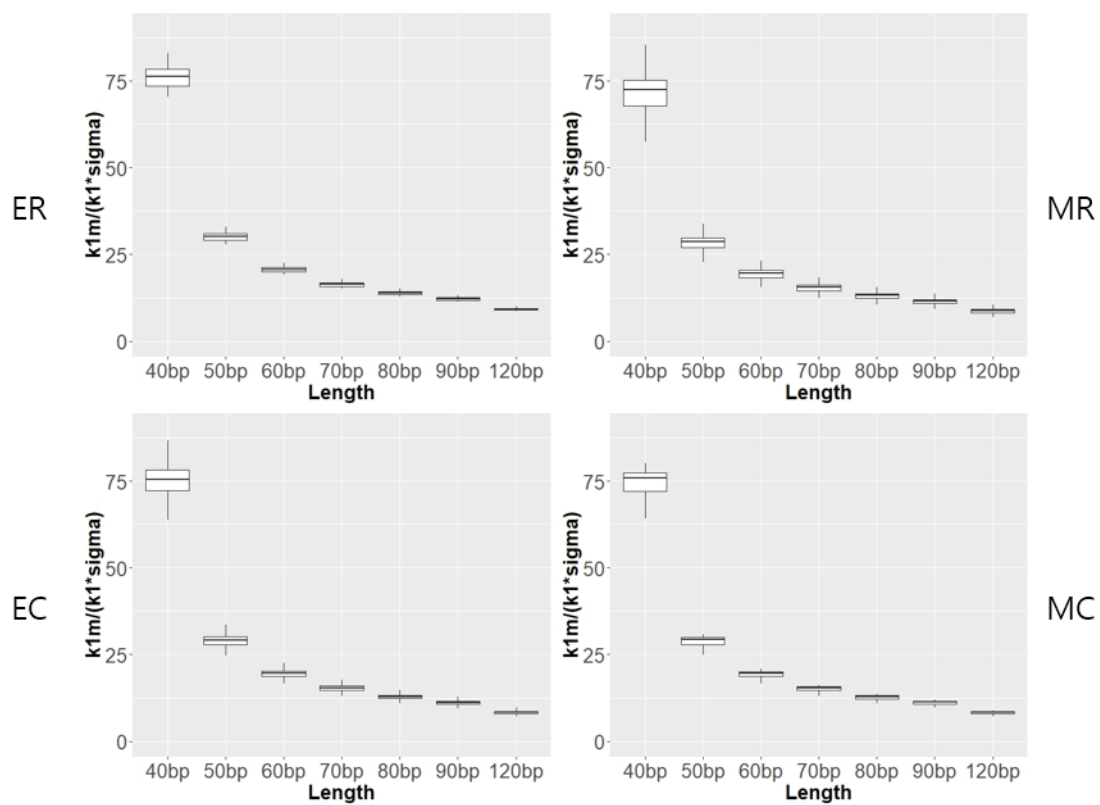


Figure 4.8: Affinities of length dependent binding rates in Model-ER, Model-EC, Model-MR and Model-MC for seven lengths.

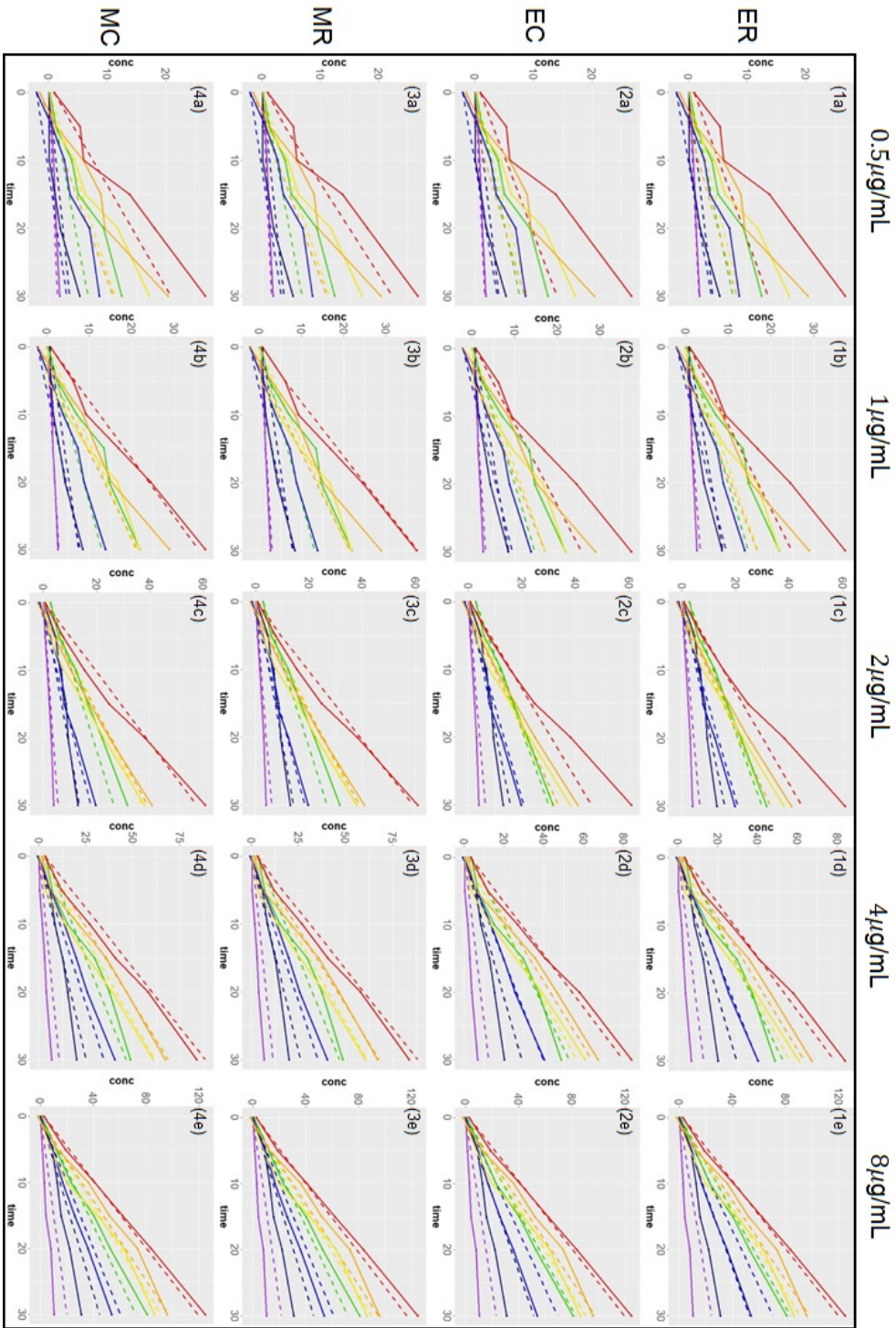


Figure 4.9: Experimental data (solid line) and prediction for the product P (dashed line) of (1) Model-ER, (2) Model-EC, (3) Model-MR and (4) Model-MC. Each column presents one concentration of dsRNA: (a) 0.5 μg/mL, (b) 1 μg/mL, (c) 2 μg/mL, (d) 4 μg/mL and (e) 8 μg/mL. Color code is given in Figure 4.6.

4.2 Model Selection

The best model is selected based on the Akaike Information Criterion (AIC), Akaike Information Criterion corrected (AICc) and Akaike weights. We use the following equations for AIC and AICc [23],

$$1) \text{ AIC} = N \ln \frac{RSS}{N} + 2K, \quad (4.3a)$$

$$2) \text{ AICc} = N \ln \frac{RSS}{N} + \frac{2KN}{N - K - 1}, \quad (4.3b)$$

where K is the number of estimated parameters +1, N is the number of observations. To calculate the AIC and AICc, the best parameter set is used for each model. The best model has the smallest AIC or AICc. However, values of AIC or AICc do not allow us to interpret how a model is credible. Thus, we use Akaike weights to supplement the AIC and AICc.

The Akaike weight is one of the normalization methods that allows to compare models; the Akaike weight of models can be interpreted as the probability of the models to be the most likely given the data and set of models considered. In this section, we use AICc for Akaike weights. The Akaike weight w_i of model i is [23],

$$w_i = \frac{\exp\left(\frac{-\Delta_i}{2}\right)}{\sum_{r=1}^R \exp\left(\frac{-\Delta_r}{2}\right)}, \quad (4.4)$$

where $\Delta_i = AICc_i - \min_i AICc_i$. A model can be considered as the single best model if $w_i > 0.9$.

One length fitting

	Model-E	Model-ED	Model-N	Model-M
N	25	25	25	25
K	4	6	8	8
RSS	348.3	343.2	336.6	264.1
AIC	73.8	77.4	81.0	74.9
$AICc$	75.8	82.1	90.0	83.9
Δ_i	0	6.29	14.14	8.07
$\exp\left(\frac{-\Delta_i}{2}\right)$	1	0.04	0.0008	0.017
w_i	0.94	0.04	0.0008	0.0166

Table 4.5: AIC, AICc and Akaike weights for four models of OAS2 activation by dsRNA (one length).

From Table 4.5, Model-E is the best model among four models when all concentrations are considered with only one length of dsRNA, $w_E = 0.94 > 0.9$. Even though Model-M has the smallest RSS, due to the parameter number, Model-E is considered as the single best model in no length change data.

All lengths fitting

Model-MR and Model-MC have relatively small differences in RSS, AIC and AICc (Table 4.6). However, Model-MC is selected as the best model by an overwhelming margin. Hence, Model-MC is the single best model to accommodate varying lengths.

Effects of length dependencies are observable in Table 4.6. Note that considering both length dependent binding rates and multi-binding improve predictions.

	Model-E	Model-ED	Model-N	Model-M	Model-ER	Model-EC	Model-MR	Model-MC
N	175	175	175	175	175	175	175	175
K	4	6	8	8	4	4	8	8
RSS	30779.4	30783.3	30769.6	9725.4	4458.3	4026.6	2577.6	2464.6
AIC	912.7	916.7	920.6	719.1	574.6	556.7	486.7	478.8
$AICc$	912.9	917.2	921.5	719.9	574.8	557.0	487.5	479.7
Ranks	6	7	8	5	4	3	2	1
Δ_i	433.20	437.49	441.78	240.22	95.09	77.27	7.84	0
$\exp\left(\frac{-\Delta_i}{2}\right)$	8.5E-95	9.9E-96	1.1E-96	6.8E-53	2.2E-21	1.6E-17	0.0198	1
w_i	8.3E-95	9.8E-96	1.1E-96	6.7E-53	2.2E-21	1.6E-17	0.0194	0.9805
Length dependency	0	0	1	1	1	1	2	2

Table 4.6: AIC, AICc and Akaike weights for eight models of OAS2 activation by dsRNA (all lengths). The ranks are obtained by sorting AIC and AICc values from the smallest to the largest.

Evidence ratio

By using evidence ratio of Akaike weights, the effect of length dependencies can be evaluated. The evidence ratio of model i over model j is [24]:

$$\frac{w_i}{w_j}.$$

Note that evidence ratio could be interpreted as an index of relative strength of two models or hypotheses. For instance, if $\frac{w_i}{w_j} = 10$, the hypothesis of model i is 10 times more likely than that of model j [25].

The hypotheses for Model-E and Model-ED are not accommodating changes in the dsRNA lengths but other models are based on at least one length dependent hypothesis. The evidence ratio of Akaike weights for length dependency over

non-length dependency is (Table 4.6)

$$\frac{w_N + w_M + w_{ER} + w_{EC} + w_{MR} + w_{MC}}{w_E + w_{ED}} = 1.07 \times 10^{94}.$$

The models, which have length dependency, are 1.07×10^{94} times more likely than the models having no length dependency.

The effectiveness of each length dependency can be tested. First, the evidence ratio of Model-N based on non-productive binding stages over Model-E, which considers only one-step binding of productive stages, is

$$\frac{w_N}{w_E} = 0.013.$$

The evidence ratio is smaller than 1: non-productive binding stages are unlikely to occur in OAS2 activation for the data used in this work.

Secondly, the evidence ratio for multi-binding assumption over single binding assumption is

$$\frac{w_M + w_{MR} + w_{MC}}{w_E + w_{ER} + w_{EC}} = 6.14 \times 10^{16}.$$

This ratio shows the effectiveness of the multi-binding assumption over the single binding assumption. Since the value is greater than 1 and large, the multiple binding is more likely to occur in OAS2 activation by dsRNA in these data.

We also test the effect of the assumptions for the length dependent binding rates. The relative strengths of 2-dimensional rigid rod assumption versus

3-dimensional cylindrical rigid rod assumption for the dsRNA shape are

$$\frac{w_{EC}}{w_{ER}} = 7402.3, \quad \frac{w_{MC}}{w_{MR}} = 50.48.$$

For both models, the 3-dimensional cylindrical rigid rod assumption is more likely than 2-dimensional assumption for the shape of dsRNA.

Furthermore, we test the effect of the use of length dependent binding rates:

$$\frac{w_{EC} + w_{MC}}{w_E + w_M} = 1.45 \times 10^{52}.$$

Models with length dependent binding rates allow a better representation of data.

Finally, the effect of the number of the length dependency can be evaluated:

$$\frac{w_{MR} + w_{MC}}{w_N + w_M + w_{ER} + w_{EC}} = 6.14 \times 10^{16}.$$

The two length dependencies is significantly more likely than the one length dependency. From evidence ratios, we can conclude that one length dependency, multi-binding assumption or length dependent binding rates, is effective to accommodate the changes in concentrations and lengths; however, both length dependencies are needed to describe the experimental data.

Chapter 5

Discussion and conclusion

In the preceding chapters, we have developed, mathematically analyzed and numerically investigated models. In this chapter, we outline the important results from this work and discuss some future directions.

5.1 Mathematical remarks and discussion

In Chapter 3, we characterized the dynamics of the product P for all the models through mathematical analysis before carrying out calibrations of models.

The different analyses carried out in Chapter 3 and 4 deal with different time periods. Numerical investigations presented in Chapter 4 focus on the time period defined by the experimental data, i.e, from 0 to 30 minutes. The exact solution of Model-S, which is a function of t , holds for all $t \geq 0$. A quasi-equilibrium is assumed in Model-N and we have obtained three different solutions of $P(t)$ based on parameter values. These three solutions differ in

their transient regimes (short run); however, they overlap in their asymptotic regime (long run). The quasi-equilibrium is valid for $t \geq \epsilon > 0$ because $\frac{dE}{dt}$ is assumed to approach to 0 after a very short time.

An asymptotically stable equilibrium is a constant solution that is approached by other solutions when time is large enough ($t > T$). All the reduced models except Model-S and Model-ED3 have a unique non-negative equilibrium. Model-E, Model-ED1, Model-ED2 and Model-M ($n = 2$) have a globally asymptotically stable equilibrium. In other words, regardless of the initial condition, the concentration of P is uniquely determined by parameters and $t > T$. Model-N and Model-M ($n = 3$) have a locally asymptotically stable equilibrium. Even though the global stability is not mathematically proved, it is numerically conjectured. Results of the qualitative analysis carried out in Chapter 3 and the resulting long-time predictions for product concentrations are summarized in Figure 5.1. With the aim of combining results from the mathematical and numerical work, Figure 5.2 presents graphs of $P(t)$ for $t > 30$ minutes using the estimates of parameters from Chapter 4 with functions given in Figure 5.1. These graphs represent the predictions for the product concentrations for $t > 30$ minutes for the different models.

5.1. MATHEMATICAL REMARKS AND DISCUSSION

Models	Equilibria	Stability	Product
E	Unique	GAS	$P(t) = \frac{k_p}{2} \left((c_0 + c_R) + \frac{1}{b_1} - \sqrt{\left((c_0 + c_R + \frac{1}{b_1})^2 - 4c_0c_R \right)} (t - T) + \rho_0 \right)$
ER, EC	Unique	GAS	$P(t) = \frac{k_p}{2} \left((c_0 + c_R) + \frac{1}{b_1 \sigma_{1,2}(L)} - \sqrt{\left((c_0 + c_R + \frac{1}{b_1 \sigma_{1,2}(L)})^2 - 4c_0c_R \right)} (t - T) + \rho_0 \right)$
N	Unique	LAS	$P(t) = \frac{k_p b_2}{2(1+b_2+\frac{b_2}{b_1})} \left((c_0 + c_R) + \frac{1}{b_1(1+b_2+\frac{b_2}{b_1})} - \sqrt{\left((c_0 + c_R) + \frac{1}{b_1(1+b_2+\frac{b_2}{b_1})} \right)^2 - 4c_0c_R} \right) (t - T) + \rho_0$
M, MR, MC	Unique	GAS	$P(t) = k_p \sum_{i=1}^n i D_i^*(t - T) + \rho_0$
		LAS	
ED	$k_{a1} \neq 0 = k_{a2}$	Unique	$P(t) = \rho_0$
	$k_{a1} = 0 \neq k_{a2}$		
	$k_{a1} \neq 0 \neq k_{a2}$	Non-isolated	

Figure 5.1: Summary of results of the qualitative analysis carried out on models in Chapter 3. The last column gives time-dependent functions corresponding to predictions for product concentrations when time is large enough ($t > T$) for the different models. Parameters are $b_1 = \frac{K_1}{k_{1m}}$, $b_2 = \frac{k_2}{k_{2m}}$, $b_5 = \frac{k_5}{k_{5m}}$ and $\rho_0 = P(T)$.

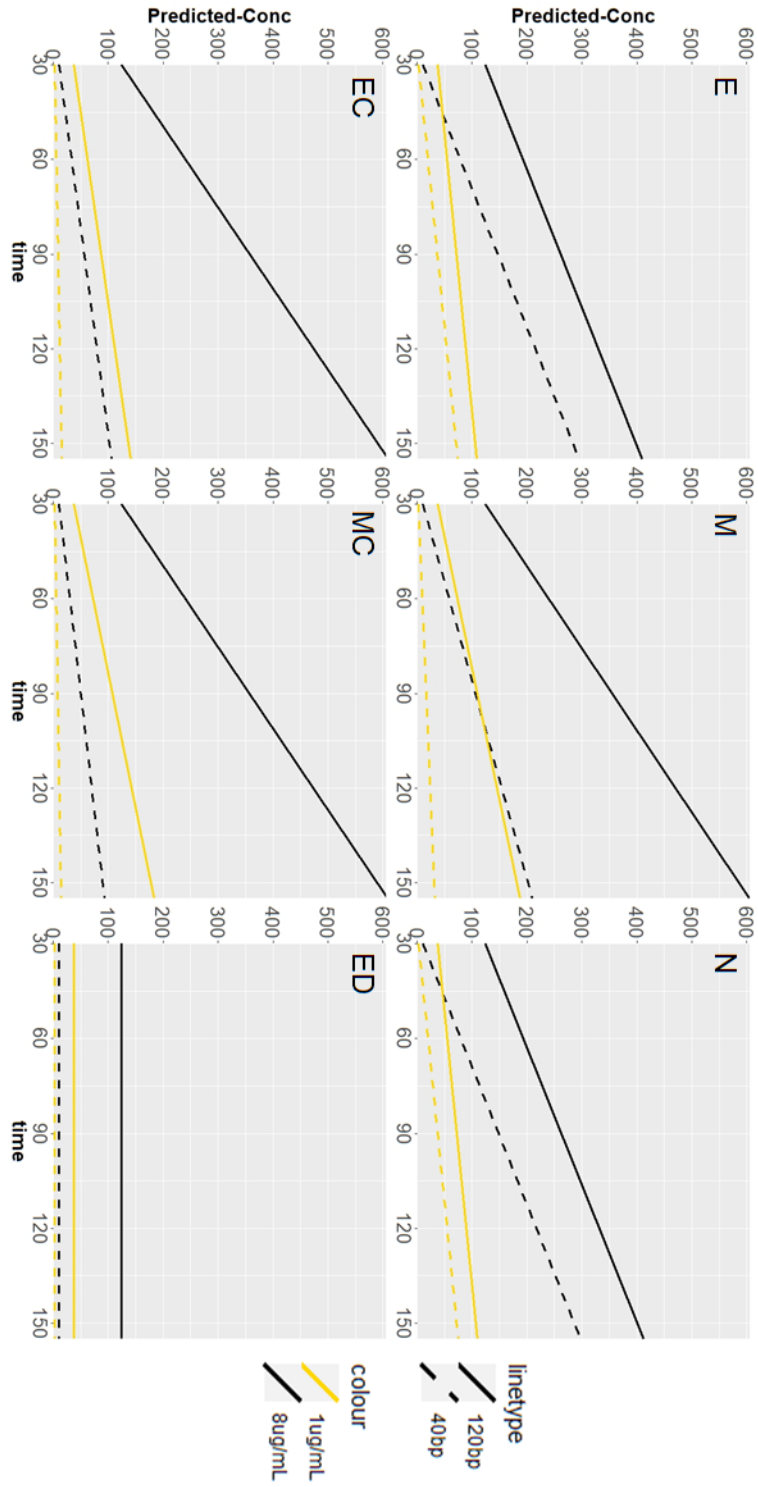


Figure 5.2: Predictions of $P(t)$ for longer times: equilibrium values of Model-E, EC, N, ED, M and MC with the best estimated parameters obtained by fitting all lengths.

5.2 Model evaluations

Since models are based on different assumptions, comparing the dynamics of $P(t)$ from the different models to experimental data (model parametrization and selection) allows us to test different sets of assumptions. A summary of conclusions of the work is given in Figure 5.3.

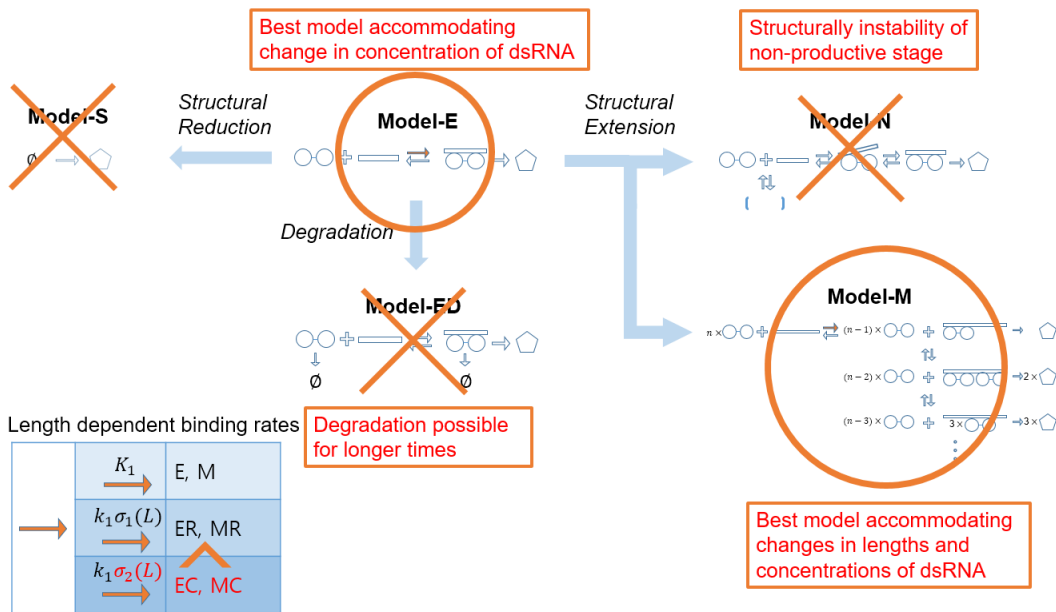


Figure 5.3: Summary of conclusions.

Model-E, which is the base model of the work and based on enzyme kinetics, is the best model accommodating all concentrations for a fixed length. Since Model-S, a structural reduction of Model-E, does not accommodate even the change in dsRNA concentration, Model-E can be considered as the simplest model describing the change in concentrations. However, in Figure 5.2, predictions from Model-E for short (40bp) and long (120bp) dsRNA are parallel

lines; Model-E cannot explain the effect of the cofactor length on the formation of the product.

Model-ER and Model-EC, which are the base model with binding rates depending on the cofactor length, almost fit data; an improved approximation for valid binding area on dsRNA could improve predictions.

In Figure 5.2, Model-E and Model-N have the same predictions. From this observation and affinities of non-productive stages found in Chapter 4, non-productive stages appear to be unfavorable or structurally unstable.

Model-N and Model-M are the models with a length dependent structural expansion of Model-E; Model-N and Model-M are extended by adding to Model-E non-productive binding stages and subsequent multi-binding stages, respectively. The assumption of non-productive stages fails to explain data; however, the multi-binding assumption provides a better explanation of data. Model-MC is the best model to accommodate the changes in concentrations and lengths of dsRNA (Figure 5.4). Each length-dependency type improves fits. They are not enough by themselves; both length dependencies and their interplay are needed to describe well the regulation of the OAS2 enzyme activity by the cofactor.

Even though Model-ED fails to fit data, degradation could be applied to data over longer periods. In this thesis, we use data sets, which measure the concentration of the product within only the first 30 minutes. However, for longer times, we could expect that OAS2, dsRNA or complex would degrade as enzyme activity would decrease over time. In the recent study, Koul *et al.* [3],

the amount of the product reaches a plateau for longer times. A plateau is also observed in the estimated equilibria for Model-ED (Model-ED in Figure 5.2).

5.3 Limitation and idea for future study

In Model-MR and Model-MC, the length dependent binding rates are only applicable to the first binding of OAS2 and dsRNA. A future study could develop the binding probability accommodating the cofactor length for the second and subsequent binding stages in Model-M.

Furthermore, in a recent study, Koul *et al.* [4] shows that OAS2 potentially works as a dimer in solution. From the significant difference in RSS between Model-MR and Model-MC, we can expect that the advanced investigation of the shape of OAS2 will help greatly improve mathematical modelling. For instance, the derivations of the binding probabilities/rates could also consider the dimeric structure of OAS2 to improve and explain better OAS2 activation by dsRNA.

Moreover, since the experimental data has its own inevitable errors [26], the mathematical consideration of how to handle errors in the data could help to improve model responses.

Driven by experimental data, we developed nine models. Since Model-MC is the best model based on the data, length dependent binding rates and multi-binding assumption can be considered as valid hypotheses. The multi-binding of OAS2 to dsRNA should be now investigated experimentally as well as its potential

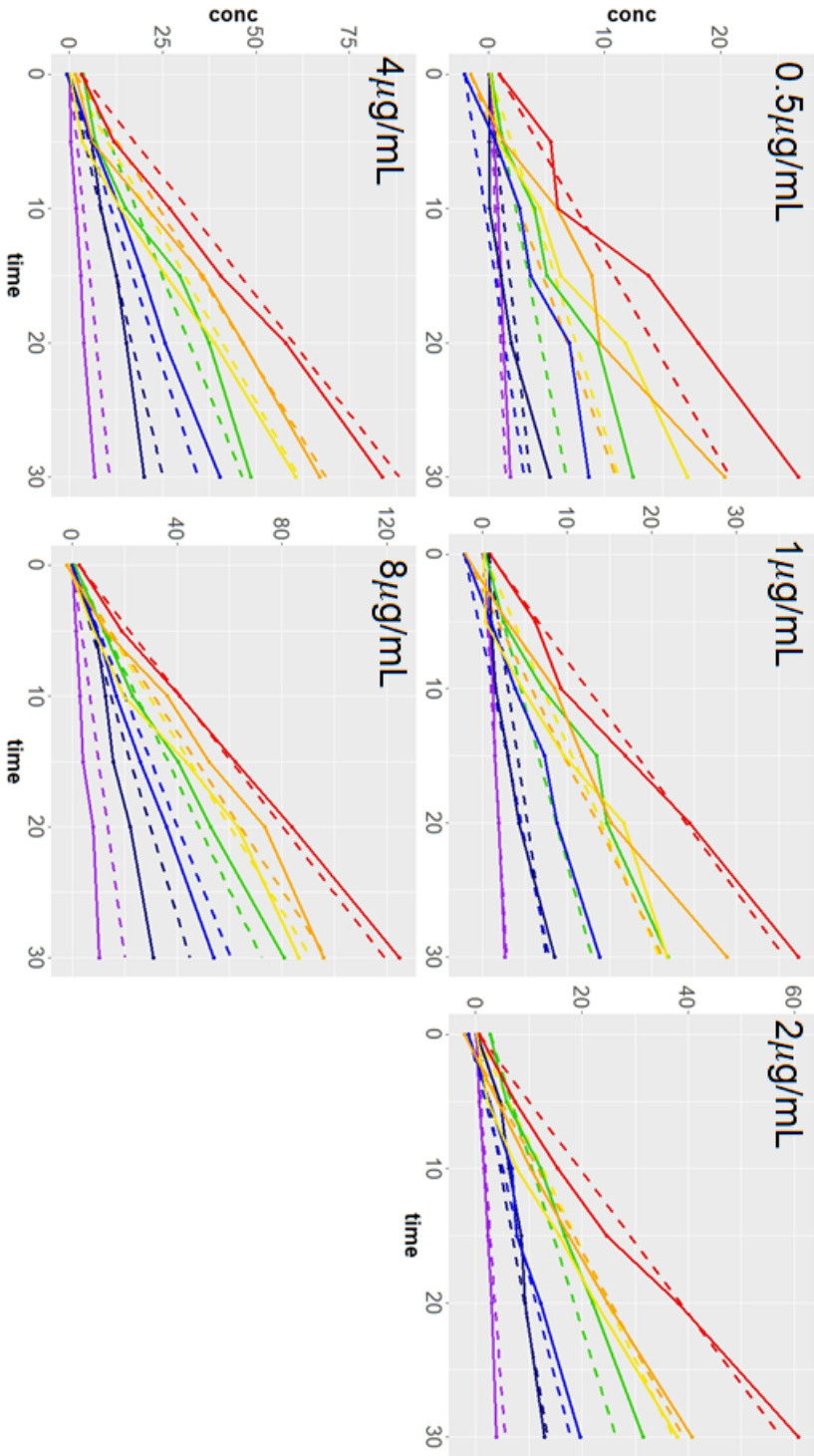


Figure 5.4: The result of fitting for Model-MC in all concentrations and all lengths of dsRNA. Color code is given in Figure 4.6.

positive cooperative nature. For example, the multi-binding assumption could be investigated if there is a noticeable increase between two similar lengths when each of them has the different number of bindings: $n - 1$ and n OAS2 bindings. If multi-binding with positive cooperativity is experimentally validated, we could add one more proof of the effectiveness of mathematical modelling.

Appendices

Appendix A

Model-M: model equations

Theorem: Model-M allowing up to n OAS2 on the same dsRNA has equation (2.12) as the ODE system.

Proof. The theorem is proved by mathematical induction. Let (2.12) be $P(n)$. Check $P(2)$.

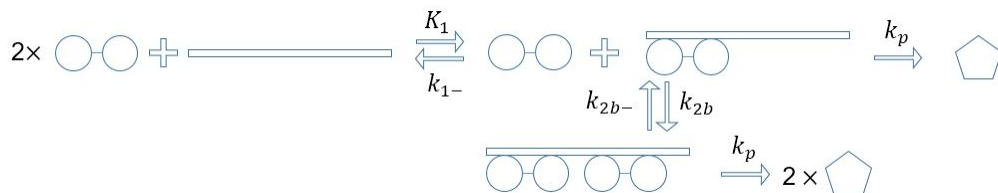


Figure A.1: Diagram associated to $P(2)$

From Figure A.1, the ODE system is:

$$\frac{dE}{dt} = -K_1ER + k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E,$$

$$\frac{dR}{dt} = -K_1ER + k_{1m}D_1,$$

$$\frac{dD_1}{dt} = K_1ER - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E,$$

$$\frac{dD_2}{dt} = k_{2b}D_1E - k_{2bm}D_2,$$

$$\frac{dP}{dt} = k_pD_1 + 2k_pD_2,$$

which is $P(n)$ with $n = 2$, so $P(2)$ is valid.

Fix $\ell \geq 2$, assume that $P(\ell)$ is true. $P(\ell)$ is,

$$\begin{aligned} \frac{dE}{dt} &= -K_1ER + k_{1m}D_1 + \sum_{i=1}^{\ell-1} [k_{(i+1)bm}D_{i+1} - k_{(i+1)b}D_iE], \\ \frac{dR}{dt} &= -K_1ER + k_{1m}D_1, \\ \frac{dD_1}{dt} &= K_1ER - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E, \\ \frac{dD_j}{dt} &= k_{jb}D_{j-1}E - k_{jbm}D_j + k_{(j+1)bm}D_{j+1} - k_{(j+1)b}D_jE, \\ & j \in \{2, 3, \dots, \ell - 1\}, \\ \frac{dD_\ell}{dt} &= k_{\ell b}D_{\ell-1}E - k_{\ell bm}D_\ell, \\ \frac{dP}{dt} &= k_p \sum_{i=1}^{\ell} iD_i. \end{aligned} \tag{A.1}$$

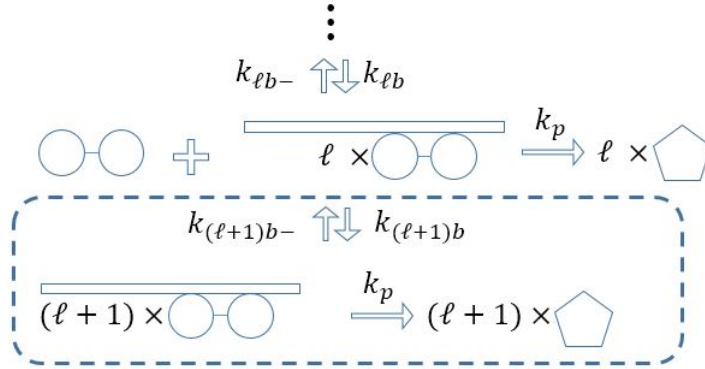


Figure A.2: ℓ th and $(\ell + 1)$ th bindings

By adding the terms of $(\ell + 1)$ th binding and unbinding into (A.1), we get the

following ODE system:

$$\begin{aligned}
\frac{dE}{dt} &= -K_1ER + k_{1m}D_1 + \sum_{i=1}^{\ell-1} [k_{(i+1)bm}D_{i+1} - k_{(i+1)b}D_iE] \\
&\quad + k_{(\ell+1)bm}D_{\ell+1} - k_{(\ell+1)b}D_{\ell}E, \\
\frac{dR}{dt} &= -K_1ER + k_{1m}D_1, \\
\frac{dD_1}{dt} &= K_1ER - k_{1m}D_1 + k_{2bm}D_2 - k_{2b}D_1E, \\
\frac{dD_j}{dt} &= k_{jb}D_{j-1}E - k_{jbm}D_j + k_{(j+1)bm}D_{j+1} - k_{(j+1)b}D_jE, \quad (\text{A.2}) \\
j &\in \{2, 3, \dots, \ell - 1\}, \\
\frac{dD_{\ell}}{dt} &= k_{\ell b}D_{\ell-1}E - k_{\ell bm}D_{\ell} + k_{(\ell+1)bm}D_{\ell+1} - k_{(\ell+1)b}D_{\ell}E, \\
\frac{dD_{\ell+1}}{dt} &= k_{(\ell+1)b}D_{\ell}E - k_{(\ell+1)bm}D_{\ell+1}, \\
\frac{dP}{dt} &= k_p \sum_{i=1}^{\ell} iD_i + (\ell + 1)k_p D_{\ell+1}.
\end{aligned}$$

Then, (A.2) is $P(\ell + 1)$. Consequently, by the Principle of Mathematical Induction, $P(n)$ is valid for all $n \geq 2$. \square

Bibliography

- [1] John J. Tyson, Katherine C. Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221 – 231, 2003. (Cited on page 1.)

- [2] Francisco Dionisio. The importance of mathematics to biology. *Journal of Ecosystem & Ecography*, 2(4):1–1, 2012. (Cited on page 1.)

- [3] Amit Koul, Soumya Deo, Evan P. Booy, George Orriss, Matthew Genung, and Sean A. McKenna. Impact of double-stranded RNA characteristics on the activation of human 2'-5'-oligoadenylate synthetase 2 (OAS2). *Biochemistry and Cell Biology*, 98(1):70–82, 2020. (Cited on pages 1, 2, 3, 4, 24 and 108.)

- [4] Amit Koul, Darren Gemmill, Nikhat Lubna, Markus Meier, Natalie Krahn, Evan P. Booy, Jörg Stetefeld, Trushar R. Patel, and Sean A. McKenna. Structural and hydrodynamic characterization of dimeric human oligoad-

- enylate synthetase 2. *Biophysical Journal*, 118(11):2726 – 2740, 2020. (Cited on pages 1, 2, 4 and 109.)
- [5] Anthony J. Sadler and Bryan R. G. Williams. Interferon-inducible antiviral effectors. *Nature Reviews Immunology*, 8(7):559–568, 2008. (Cited on page 2.)
- [6] Charles E. Samuel. Antiviral actions of interferons. *Clinical microbiology reviews*, 14(4):778–809, 2001. (Cited on page 2.)
- [7] Robert H. Silverman. Viral encounters with 2', 5'-Oligoadenylate synthetase and RNase L during the interferon antiviral response. *Journal of Virology*, 81(23):12720–12729, 2007. (Cited on page 2.)
- [8] Helle Kristiansen, Hans Henrik Gad, Signe Eskildsen-Larsen, Philippe Despres, and Rune Hartmann. The oligoadenylate synthetase family: An ancient protein family with multiple antiviral activities. *Journal of Interferon & Cytokine Research*, 31(1):41–47, 2011. (Cited on page 2.)
- [9] Otto G. Berg and Peter H. von Hippel. Diffusion-controlled macromolecular interactions. *Annual Review of Biophysics and Biophysical Chemistry*, 14(1):131–158, 1985. PMID: 3890878. (Cited on pages 13 and 15.)
- [10] Stéphanie Portet, Norbert Mücke, Robert Kirmse, Jörg Langowski, Michael Beil, and Harald Herrmann. Vimentin intermediate filament formation: in vitro measurement and mathematical modeling of the filament length

- distribution during assembly. *Langmuir*, 25(15):8817–8823, Aug 2009. (Cited on page 13.)
- [11] Stéphanie Portet. Dynamics of in vitro intermediate filament length distributions. *Journal of Theoretical Biology*, 332:20–29, Sep 2013. (Cited on page 13.)
- [12] Jesse Goyette, Citlali Solis Salas, Nicola Coker-Gordon, Marcus Bridge, Samuel A. Isaacson, Jun Allard, and Omer Dushek. Biophysical assay for tethered signaling reactions reveals tether-controlled activity for the phosphatase SHP-1. *Science Advances*, 3(3), 2017. (Cited on page 13.)
- [13] Samuel Bell and Eugene M. Terentjev. Kinetics of tethered ligands binding to a surface receptor. *Macromolecules*, 50(21):8810–8815, 2017. (Cited on page 13.)
- [14] Julio de Paula and Peter Atkins. *Physical Chemistry for the Life Sciences*. W. H. Freeman and Company, 2006. (Cited on page 14.)
- [15] Takenao Yoshizaki and Hiromi Yamakawa. Dynamics of spheroid-cylindrical molecules in dilute solution. *The Journal of Chemical Physics*, 72(1):57–69, 1980. (Cited on page 15.)
- [16] Carlos G. Lopez, Oliva Saldanha, Klaus Huber, and Sarah Köster. Lateral association and elongation of vimentin intermediate filament proteins: A

- time-resolved light-scattering study. *Proceedings of the National Academy of Sciences*, 113(40):11152–11157, 2016. (Cited on page 15.)
- [17] Jonathon Howard. *Mechanics of Motor Proteins & the Cytoskeleton*. Sinauer Associates, 2001. (Cited on page 16.)
- [18] Johan G. Olsen, Kaare Teilum, and Birthe B. Kragelund. Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. *Cellular and molecular life sciences*, 74(17):3175–3183, 2017. (Cited on page 22.)
- [19] Richard K. Miller and Anthony N. Michel. *Ordinary Differential Equations*. Academic Press, 1982. (Cited on page 33.)
- [20] Linda J. S. Allen. *An introduction to mathematical biology*. Pearson, 2006. (Cited on page 40.)
- [21] Leah Edelstein-Keshet. *Mathematical models in biology*. SIAM, 2005. (Cited on page 61.)
- [22] John L. Tymoczko Jeremy M. Berg and Lubert Stryer. *Biochemistry*. W.H. Freeman, 2002. (Cited on page 87.)
- [23] Stéphanie Portet. A primer on model selection using the Akaike information criterion. *Infectious Disease Modelling*, 5:111 – 128, 2020. (Cited on page 97.)

- [24] Eric-Jan Wagenmakers and Simon Farrell. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11:192–196, 2004. (Cited on page 99.)
- [25] Paul M. Lukacs, William L. Thompson, William L. Kendall, William R. Gould, Paul F. Doherty Jr, Kenneth P. Burnham, and David R. Anderson. Concerns regarding a call for pluralism of information theory and hypothesis testing. *Journal of Applied Ecology*, 44:456–460, 2007. (Cited on page 99.)
- [26] John R. Taylor. *An Introduction to Error Analysis*. University Science Books, 1997. (Cited on page 109.)