Sustainability of RNA-interference in Rule Based Modelling

Masahiro Hamano [1]

PRESTO
Japan Science and Technology Agency (JST)
4-1-8 Honcho Kawaguchi, Saitama 332-0012, JAPAN

Abstract

RNA interference (RNAi) is a mechanism whereby small pieces of RNA directly control gene expression in target messenger RNA by identifying complementary sequences. We model RNAi in terms of rule-based modelling. Interpreting a small interfering RNA (siRNA) as a primitive agent, the model provides a fine-grained interpretation to directly capture the fundamental interactions (e.g., hybridization, denaturation, cleavage, copying, degradation) among double and single strands of RNA and siRNA. We investigate the sustainability of RNAi, which is characterized by the population level of double-stranded RNA (dsRNA) during the interference. Our model aims to capture the individual level of each agent in RNAi in terms of the Galton–Watson multitype branching processes determined by the modelling. Each siRNA has a type that represents its original position inside the dsRNA from which it was cleaved. The probability of extinction of populations of siRNA is investigated and analyzed for both primer-dependent and -independent synthesis of RNAi, which are important topics in experimental biology from the perspective of the difference between animal and plant RNAi. The sustainability is shown to be invariant under some appropriate model refinements for the primer-dependent synthesis. This invariance guarantees the sufficiency of the compact description of our rule based modelling in capturing the sustainability of RNAi.

Keywords: RNA interference, Rule-based Modelling, Multitype Branching Process

1 Introduction

The mechanism of RNA interference (RNAi), also known as RNA silencing, is widely found throughout eukaryotes, where small pieces of RNA (21–26 nts), known as small interfering RNA (siRNA), directly control gene expression [1,23]. RNAi consists of three fundamental biochemical processes: (i) Formation of double-stranded RNA (dsRNA); (ii) The Dicer enzyme's cleavage of dsRNA into siRNAs; and (iii) The incorporation of siRNA into an RNA-induced silencing complex (RISC) by argonaute-protein, targeting and degrading a long single-stranded messenger RNA (mRNA) by Watson–Crick complementary pairing. (See the right half of Figure 1 for (i), (ii), and (iii).)

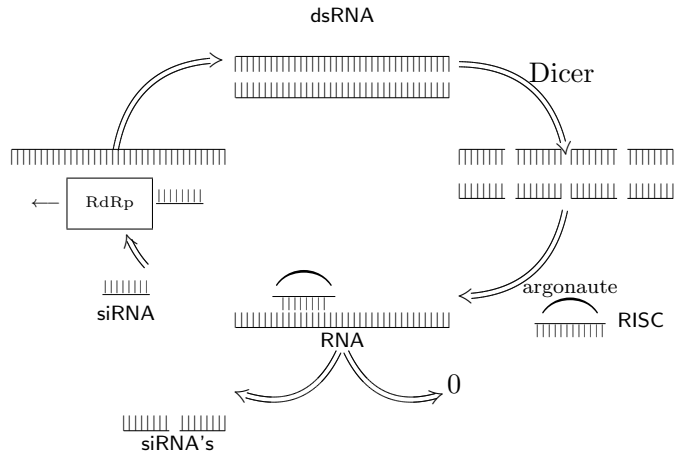[1] Email: hamano@is.s.u-tokyo.ac.jp

Fig. 1. RNA interference

This work concerns a certain circularity among the three processes, which explains the persistence with which RNAi is sustained. We model circular paths from (ii) and (iii) to (i) such that secondary dsRNA is synthesized by RNA-directed RNA polymerase (RdRp) mediation.

The synthesis is performed in two ways [1,2,4,26]: (1) primer-dependent synthesis: An siRNA resulting from (ii) and (iii) triggers polymerization from a single-stranded mRNA template; (2) primer-independent synthesis: RNA containing an aberrant feature resulting from (iii) is duplicated without the trigger.
(See Figure 2 for an illustration of these synthesis mechanisms and Figure 1 for RNAi with primer-dependent synthesis.)
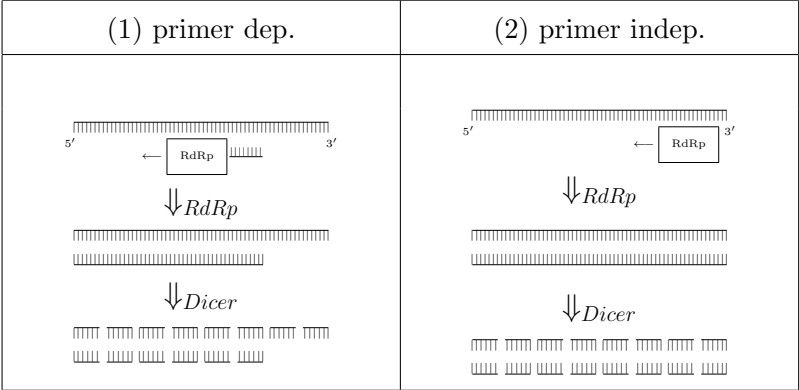
Fig. 2. Two circular paths for the synthesis of dsRNA

The purpose of this paper is to demonstrate the effectiveness of primer-independent duplication (2) compared to primer-dependent polymerization (1). To show this, we first represent RNAi as a rule-based model using kappa calculus, which is a well-established means of handling the combinatorics of molecular interactions [9,10,12]. The machinery of RNAi is seen as the combinatorial expression of various sized segments of nucleotides in terms of hybridization, denaturation, ligation,

cleavage, copying and so on. Our representation takes siRNA as a primitive agent that has appropriate sites for *phosphate* bonds and *hydrogen* bonds. Each siRNA has an associated type denoting its original position inside the dsRNA from which it was cleaved. Other agents can be defined in terms of these primitive agents and types. (mRNA is represented as a sequence of siRNAs with phosphate bonds to each predecessor and successor, and dsRNA is described by two complementary strands of RNA joined by hydrogen bonds.) Interactions involving RNAi are described by a set of rules according to the binding and unbinding of certain sites on the different types of siRNA. (The cleavage of dsRNA by Dicer is a simultaneous unbinding reaction via phosphate bonds. RISC matching of a complementary sequence is the binding of hydrogen bonds between two complementary sites of mRNA and siRNA.) Importantly, these principal reactions consisting of RNAi are digital in nature, with certain stochastic rates for each rule and site. This fits into the stochastic semantics [25,5] of stochastic processes calculi, which include kappa calculus.

We analyze how RNAi evolves under these rule-based dynamics. In such modelling, the sustainability of RNAi, which is our main concern in this paper, is captured by individual populations of each agent in each generation. As seen in Figure 1, dsRNA is an initiator of RNAi, and is itself produced by an siRNA trigger. Hence, the population level of siRNA is a prime factor in sustaining RNAi. In general, the kappa syntactical modelling yields stochastic Markov processes as its semantical counterpart. Moreover, our RNAi modelling with typed siRNA agents yields Galton–Watson multitype branching processes, the semantics of which allow us to analyze the growth of each siRNA population. The use of a multitype branching process was suggested in [3], though most mathematical modelling of RNAi is phenomenologically and deterministically given by differential equations [3,6,13,18]. Rule-based modelling naturally provides the stochastic process of branching by directly capturing the interactions among agents.

Besides its practical agility for statistical analysis [7,9,21], the main advantages of rule-based modelling in our work are its compactness for describing rules and its model refinement to incorporate less-compact details. Our aim in this paper is to show that a certain property of the evolution of RNAi is sufficiently captured by this compact modelling to allow polymerization to synthesize dsRNA. To ensure this, the model is refined in such a manner that the property obtained in the compact model is shown to be invariant under some appropriate class of model refinements.

We show that, with only primer-dependent synthesis, the population of each type of siRNA becomes extinct. Hence, RNAi ceases to be sustained. The probability of these extinctions is proved to be 1 by the compact description of the rules, and the probability is shown to remain invariant under plausible classes of model refinements, which globalize the compact rules by incorporating contextual traits of the complexes. Then, as soon as primer-independent synthesis is augmented, RNAi is shown to become sustainable because the extinction probabilities become less than 1.

2   Rule Based Modelling of RNAi

In this section, we model RNAi syntactically using kappa calculus [10]. In our rule-based modelling of RNAi, each siRNA is considered a primitive agent, denoted by S_k, where the index k (called the *type*) designates the position inside the dsRNA from which the agent originates. (In this paper, siRNA stands for *single-stranded* one.) Types are natural numbers $1, 2, \ldots, m$ from downstream to upstream ($3'$ to $5'$ of mRNA). T denotes the set of all types. Each primitive agent has three sites

$$\text{siRNA} = S_k(l, h, r),$$

where r and l are for *phosphate* bonds and h is a segment for a series of *hydrogen* bonds with 21–26 nts. Binding of two sites is represented by a common superscript. The mRNA and dsRNA agents are represented as complexes consisting of primitive S_ks with appropriate bonds.

mRNA =
$$\ldots, S_{n+1}(l^{n+2}, r^{n+1}), S_n(l^{n+1}, r^n), \ldots, S_2(l^3, r^2), S_1(l^2, r)$$

dsRNA =
$$\ldots, S_{n+1}(l^{n+2}, h^{1_{n+1}}, r^{n+1}), S_n(l^{n+1}, h^{1_n}, r^n), \ldots, S_2(l^3, h^{1_2}, r^2), S_1(l^2, h^{1_1}, r)$$
$$\ldots, S_{n+1}(l^{\overline{n+2}}, h^{1_{n+1}}, r^{\overline{n+1}}), S_n(l^{\overline{n+1}}, h^{1_n}, r^{\overline{n}}), \ldots, S_2(l^{\overline{3}}, h^{1_2}, r^{\overline{2}}), S_1(l^{\overline{2}}, h^{1_1}, r)$$

Note that in the definition of mRNA, site h for each S_k is not written, whereas it is present in the definition of dsRNA. See Figure 3 for a visual representation of each agent with their respective sites, where bound (res. unbound) sites are represented by black (res. white) circles.
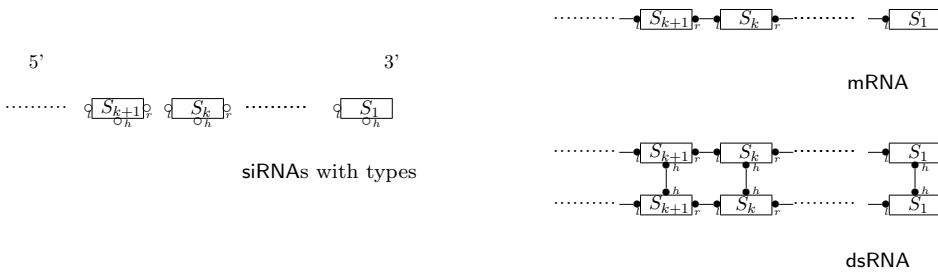

Fig. 3. Agents defined using S_k

Three fundamental reactions of RNAi are described by the following rules:

Definition 2.1

(i) polymerization

$$S_k(l, h^{1_k}, r^k) \longrightarrow S_k(l^{k+1}, h^{1_k}, r^k), S_{k+1}(l, h^{1_{k+1}}, r^{k+1})$$

This rule describes how siRNA of type $k + 1$ is produced from its predecessor type. The bounded site h (res. r) on the left-hand side (LHS) gives the

condition that the hydrogen (res. ligation) bond must connect to the template mRNA (res. to the predecessor of S_k). This rule is compact in that only the local part of the domains and ranges of the rule are specified. The contextual siRNA so far produced (i.e., S_i with $i < k$) and the template mRNA (i.e., the upper strand mRNA having the complementary site for the hydrogen bond h of S_k) are not relevant to this rule.

(ii) cleavage

$$\prod_{i \in T} (S_i(l^{i+1}, h^{1_i}, r^i) \mid S_i(\overline{l^{i+1}}, h^{1_i}, r^{\bar{i}})) \longrightarrow \prod_{i \in T} (S_i(l, h^{1_i}, r) \mid S_i(l, h^{1_i}, r))$$

where the LHS of the rule describes dsRNA, written with $\mid$ in place of , and $\prod$ to denote compositions of $\mid$. All the phosphate bonds on the RHS are released by the rule.

(iii) degradation

$$\mathsf{RISC}(h^{1_k}), \quad S_k(l^{k+1}, h^{1_k}, r^k) \mid \prod_{i \in T \setminus \{k\}} S_i(l^{i+1}, r^i) \longrightarrow \mathsf{RISC}(h), \ 0$$

where RISC is an agent with a site for hydrogen bonding to a complementary sequence of mRNA. The second agent on the LHS is mRNA with one bound site for hydrogen bonds. RISC is recycled, that is, it occurs on both the LHS and RHS.

See Figure 4 for an illustration of each rule.

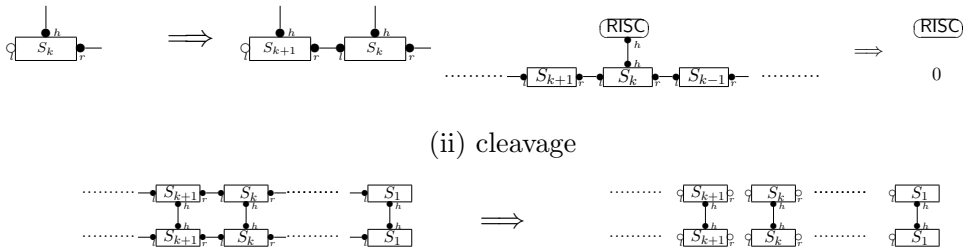(i) polymerization   (iii) degradation

(ii) cleavage

Fig. 4. Rules of RNAi

Remark 2.2 In the above cleavage rule (ii), we choose a modelling process that destroys several bonds simultaneously. However, for a smaller time-scale, intermediate steps may be needed in terms of successive Dicer interactions against the dsRNA. Given that Dicer is a molecular ruler [19] for recognizing dsRNA and cleaving 21–26 nts for the length of siRNA, we define it as an agent $D(1, \ldots, k, \ldots, m)$ with m sites for measuring/cleaving dsRNA. The cleaving of dsRNA is then represented by the successive interactions of Dicer, for which siRNA needs to be augmented with another site d. The following is an alternative, fine-grained representation of the cleavage rule that produces a series of (ii-k) rules, $k = 1, \ldots, m$, corresponding to (ii).

(ii-k) k-th measuring/cleaving of Dicer over dsRNA

$$D(1, \ldots, k, \ldots, m) \mid (S_k(l^{k+1}, h^{1_k}, r, d) \mid S_k(l^{\overline{k+1}}, h^{1_k}, r)) \mid C_{k+1,k}$$

$$\longrightarrow \qquad\qquad\qquad\qquad\qquad \text{(binding of k-th site of Dicer to dsRNA)}$$

$$D(1, \ldots, k^0, \ldots, m) \mid (S_k(l^{k+1}, h^{1_k}, r, d^0) \mid S_k(l^{\overline{k+1}}, h^{1_k}, r)) \mid C_{k+1,k}$$

$$\longrightarrow \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(cleaving S_ks)}$$

$$D(1, \ldots, k^0, \ldots, m) \mid (S_{k+1}(l^{k+2}, h^{1_{k+1}}, r) \mid S_{k+1}(l^{\overline{k+2}}, h^{1_{k+1}}, r))$$

$$\qquad\qquad\qquad\qquad \mid (S_k(l, h^{1_k}, r, d^0) \mid S_k(l, h^{1_k}, r)) \mid C_{k+2,k}$$

$$\longrightarrow \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(unbinding of D)}$$

$$D(1, \ldots, k, \ldots, m) \mid (S_{k+1}(l^{k+2}, h^{1_{k+1}}, r) \mid S_{k+1}(l^{\overline{k+2}}, h^{1_{k+1}}, r)) \mid C_{k+2,k+1}$$

where $C_{n_1,n_2} = \displaystyle\prod_{i \geq n_1} U_i \mid \prod_{i < n_2} T_i$ with

$U_i = S_i(l^{i+1}, h^{1_i}, r^i) \mid S_i(l^{\overline{i+1}}, h^{1_i}, r^{\bar{i}})$ and $T_i = S_i(l, h^{1_i}, r) \mid S_i(l, h^{1_i}, r)$.
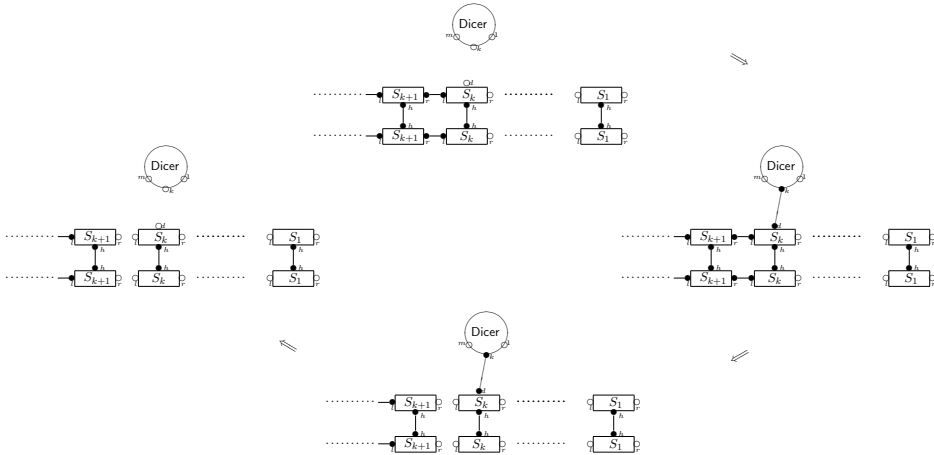
See Figure 5 for a visual description of this rule.

Fig. 5. Rule (ii-k) for cleavage

An important aspect of the fine-grained rules (ii-k) is that they make the description of cleavage compact, involving only local structures without mentioning the whole structure.

3   RNAi as a Multitype Branching Process for siRNA

This section describes a semantical study of the rule-based modelling presented in Section 2. We show that the two syntheses of dsRNA described in Section 1

are captured and discriminated by Galton–Watson multitype branching processes [20,14] for the different siRNA types introduced in Section 2.

A sequence $\{\mathbf{Z}(n)\}$ of vector random variables represents the number of individuals of the various types of siRNA in the n-th generation

$$\mathbf{Z}(n) = (Z_1(n), \dots, Z_m(n))$$

so that Z_i is a random variable for S_i (i.e., for siRNA of type i).

The $m \times m$-matrix $M = (m_{ij})$, called the *mean matrix*, is defined by

$$m_{ij} = E[Z_j(1) \mid \mathbf{Z}(0) = \mathbf{e}_i]$$

for all $i, j = 1, 2, \dots, m$, where $\mathbf{e}_i$ denotes the vector whose i-th component is 1 and whose other components are 0. That is, each element m_{ij} gives a type i individual's expected number of children of type j. Let us write

$$u(n) = E[\mathbf{Z}(n)] = (E[Z_1(n)], \dots, E[Z_m(n)])$$

so that

$$u(n) = u(0)M^n.$$

Corresponding to the sequence of vector-valued random variables is the sequence $\mathbf{f}(s) = (f_1(s), \dots, f_m(s))$ of generating functions for $\mathbf{s} = (s_1, \dots, s_m) \in [0,1]^T$, defined by

$$f_i(\mathbf{s}) = \sum_r P[\mathbf{Z}(n) = r \mid \mathbf{Z}(0) = \mathbf{e}_i] s_1^{r_1} s_2^{r_2} \cdots s_M^{r_m}.$$

The generating functions characterize trivial processes without any branching, such that each individual has exactly one offspring of any type with probability 1. The characterization of this *singular process* is $\mathbf{f}(\mathbf{s}) = A\mathbf{s}^T$ for some matrix A.

In this paper, we are interested in the probabilities q_i of the eventual extinction of the process initiated with a single particle of type i. These are given by

$$q_i = \lim_{n \to \infty} q_i(n) \text{ where } q_i(n) = P[\mathbf{Z}(n) = 0 \mid \mathbf{Z}(0) = \mathbf{e}_i].$$

The generating functions $f_i(s)$ yield a recursive definition of the probability $q_i(n)$ such that

$$q_i(1) = f_i(0) \text{ and } q_i(n+1) = f_i(\mathbf{q}(n))$$

for $\mathbf{q}(n) = (q_1(n), \dots, q_m(n))$. Taking the limit of this recursion, we have ($i \in \{1, \dots, m\}$)

$$q_i = f(\mathbf{q}) \text{ for } \mathbf{q} = (q_1, \dots, q_m).$$

In an *irreducible branching process*, each type of individual may eventually have progeny of any other type. For every pair (i, j) of types, there exists an integer $n \geq 1$ such that

$$P[Z_j(n) \geq 1 \mid \mathbf{Z}(0) = \mathbf{e}_i] > 0.$$

That is, the (i, j)-th element of M^n is strictly positive.

The irreducibility is a criterion to discriminate primer-independent from primer-dependent synthesis:

Proposition 3.1

(1) RNAi *with primer-dependent synthesis yields a reducible branching process.*

(2) RNAi *with primer-independent synthesis yields an irreducible branching process.*

Proof. In primer-dependent synthesis, it is directly observed that no children of type S_k with $k < n$ are produced by a parent of type S_n. In contrast, in primer-independent synthesis, every type of offspring is produced immediately, thus this process becomes irreducible. □

The well-known Perron–Frobenius theorem for the classical theory of matrices (cf. Theorem 6.1 of [20]) says that, for every irreducible process, the mean matrix M has a unique positive eigenvalue ρ (called its *Perron–Frobenius root*) that is greater in absolute value than any other eigenvalue, and the powers of M have the property that

$$M^n = \rho^n M_1 + o(\rho^n),$$

where M_1 is the matrix whose (i, j)-th element is given by $u_i \cdot v_j$ for the normalized right and left eigenvectors ${}^t u$ and v such that $\rho v = vM$ and $Mu = \rho u$.

Irreducibility is the property that any initial configuration can lead to any other composition. Hence, in irreducible populations, all types grow at the same rate according to the single parameter ρ of the Perron–Frobenius root. Therefore, this parameter completely characterizes the extinction and growth:

Lemma 3.2 (Theorem 7.1 of pg. 16 [20] for irreducible processes)   For a non-singular and irreducible process, the probability of extinction is the solution of

$$\mathbf{f}(\mathbf{s}) = \mathbf{s}$$

that is closest to the origin in the unit cube $[0, 1]^T$. Moreover,

(i) If $\rho \leq 1$, then $q_i = 1$ for all $i = 1, \ldots, m$.
(ii) If $\rho > 1$, then $q_i < 1$ for all $i = 1, \ldots, m$.

In reducible processes, distinct groups of types that do not produce those of other groups can be distinguished. Reducible systems can, in general, display great heterogeneity. One type may become extinct, whereas another thrives, and different types may grow at different rates. Nevertheless, uniform extinctions occur in the following case, with λ given by the maximal eigenvalue of the mean matrix:

Lemma 3.3 (Theorem 3.1 of pg. 65 [20] for reducible processes) If a reducible multitype branching process is non-singular and $\lambda \leq 1$, then the population becomes extinct with probability 1 given that $\mathbf{Z}(0) = \mathbf{e}_i$ for all i.

The primer-dependent polymerization of RNAi gives the following mean matrix M_{dep}, which is triangular with lower left elements of 0.

$$(1) \quad M_{dep} = \begin{pmatrix} s_1 & & & * \\ & s_2 & & \\ & & \ddots & \\ 0 & & & s_m \end{pmatrix}$$

Let u_n denote the n-th row of the matrix (1):

$$u_n = (0, \ldots 0, s_n, m_{n,n+1}, \ldots, m_{nm}).$$

Then, each element of u_n (from left, respectively) corresponds to the birth rate of children S_i (with $i = 1, 2, \ldots, m$, respectively) triggered by the polymerization of agent S_n. The front 0s indicate that no children S_j ($1 \leq j < n$) are produced by the polymerization triggered by S_n. The element $m_{n,i}$, which is the birth rate of the individual S_i, is then determined by its immediate predecessor S_{i-1} according to the polymerization rule of Definition 4. Thus, we have

$$s_n = \mathsf{site}_n(S_n) \quad \text{and} \quad m_{n,i} = \mathsf{site}_{n,i}(S_{i-1}, S_i),$$

where site indicates that s_n and $m_{n,i}$ are determined by the states of all the sites of the agents inside the arguments.

Example 3.4 Each siRNA of type k decays with probability s'_k. Then, with probability q, siRNA binds to a complementary mRNA to trigger RdRp to copy the template mRNA. With probability h, denaturation takes place between S_k and the mRNA, breaking the hydrogen bonds. With probability r, ligation breaks between two siRNAs of types $k - 1$ and k. Under these conditions, u_n for (1) is given by the following, where $\bar{x} = 1 - x$:

$$s_n = 1 - s'_n \quad \text{and} \quad m_{n,i} = \bar{h}(\bar{h}\bar{r})^{i-1} q\, s_n.$$

Proposition 3.5 (Extinction of siRNA in solely primer-dep. synthesis)
The populations of all types of siRNA $S_1, \ldots, S_m$ *eventually become extinct.*

Proof. The eigenvalues of (1) are given by its diagonal elements. Hence, the Perron–Frobenius root is the maximum eigenvalue, and this is less than or equal to 1. Thus, the assertion follows from Lemma 3.3. □

As soon as primer-independent synthesis is enabled, we have the following:

Proposition 3.6 (Sustainablility of RNAi with primer-indep. synthesis)
RNAi *may be sustainable with primer-independent synthesis. That is, The probability of extinction of every type* $S_1, \ldots, S_m$ *becomes less than 1.*

Proof. The mean matrix M_{indep} for the primer-independent synthesis has the following form:

$$M_{indep} = M_{dep} + \sum_{j=1}^{m} \mathbf{u} \otimes {}^t\mathbf{e}_j$$

where

$$\mathbf{u} = (q, \; qc, \; qc^2, \dots, qc^{m-1})$$

is given by a geometric sequence whose initial term q denotes the probability of RdRp mediation. Each element (from left, respectively) represents the birth rate of the corresponding type $S_1, \dots, S_m$, respectively. Following the convention of Example 3.4, the common ratio c is given by $c = \bar{h}\bar{r}$.

The Perron–Frobenius root of $M_{indep} - M_{dep}$ is given by $\rho = \sum_{i=1}^{m} u_i = q\sum_{i=1}^{m-1} c^{i-1}$. With appropriate choices for the parameters q and c, it is possible to ensure $\rho > 1$. By Lemma 3.2 for irreducible processes, the assertion holds.   □

4   Invariance under Model Refinements

The original rule of polymerization is local and compact in that the creation of S_{k+1} of the lower strand is determined only by local knowledge of the immediate predecessor S_k. The knowledge concerns whether sites h and r of S_k are bound to the upper strand and to the predecessor by hydrogen and phosphate bonds, respectively. We can refine this rule to make this local description global, and thus incorporate contextual knowledge from all predecessors of siRNA (i.e., S_j with $j \leq k$) as well as on the template mRNA.

Definition 4.1 The polymerization rule is refined as follows:

$$
\begin{array}{l}
\qquad\qquad\qquad\qquad\qquad S_k(l^{k+1}, h^{1_k}, r^k), \\
S_k(l, h^{1_k}, r^k), \\
\qquad\qquad\qquad \longrightarrow \prod_{j<k} S_j, \; \mathsf{mRNA}(h^{1_k}, h^{1_{k+1}}), \\
\prod_{j<k} S_j, \; \mathsf{mRNA}(h^{1_k}, h) \\
\qquad\qquad\qquad\qquad\qquad S_{k+1}(l, h^{1_{k+1}}, r^{k+1})
\end{array}
$$

Although the states of all sites of S_j ($j < k$) are accounted for, no bonding conditions are required for the hydrogen and phosphate sites. $\mathsf{mRNA}(h, h)$ denotes the two sites of contiguous agents of S_k and S_{k+1} in the template. See Figure 6 for the refined rule.
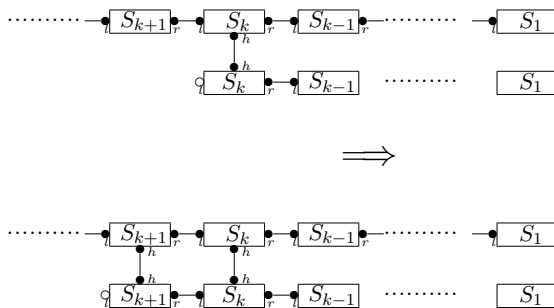

Fig. 6. Rule refinement for polymerization

The branching process for the refined rule is given by the mean matrix $M_{ref(dep)}$. This has the same form as (1), but the n-th row u_n is now given by

$$s_n = \mathsf{site}_n(S_n, \mathsf{mRNA}) \quad \text{and} \quad m_{n,i} = \mathsf{site}_{n,i}(S_n, S_{n+1}, \ldots, S_{i-1}, S_i, \mathsf{mRNA}).$$

Note that the number of arguments of site is enlarged to accommodate all S_j $(j \leq n)$ as well as mRNA.

Proposition 4.2 (Invariance under the rule refinements) *The extinction property of Proposition 3.5 is invariant under the rule refinement of Definition 4.1.*

Proof. The Perron–Frobenius roots of the mean matrices do not increase under the refinement. That is, $\rho_{M_{dep}} \geq \rho_{M_{ref(dep)}}$ for any model refinement $ref(dep)$ defined in Definition 4.1 to the primer-dependent synthesis. This is because the diagonal elements of the matrices do not increase, i.e., $\mathsf{site}_n(S_n) \geq \mathsf{site}_n(S_n, \mathsf{mRNA})$ for $1 \leq n \leq m$. These elements are sufficient to determine the root, given that both $M_{ref(dep)}$ and M_{dep} are triangular. $\qquad \square$

5 Conclusions and Future Works

In this paper, we modelled RNAi using a rule-based approach and investigated Galton–Watson multitype branching processes for several types of siRNA. We demonstrated the extinction of all types of siRNA for primer-dependent synthesis of RNAi in a compact model (Proposition 3.5). Model refinement was used to validate our compact description, so that the extinction property remains invariant throughout the class of the refinement (Definition 4.1), forcing the model to be less compact (Proposition 4.2). We also studied a branching process for primer-independent synthesis, which was shown to make RNAi sustainable (Proposition 3.6).

In future work, we will consider the following three problems: (1) This paper does not discuss any *heterogeneity* peculiar to reducible branching processes for primer-dependent synthesis. We expect that distributions of individuals of each siRNA type may be captured by rule-based modelling. The distribution of spreading concentrations of siRNA has been experimentally observed in [23] for animal RNAi, with the concentration differing according to the origin (5' or 3') of the siRNA. This spreading is being investigated by experimental biologists from the standpoint that the two syntheses discussed in our paper could explain the difference in RNAi [17] between plants [1,4] and animals [23]. Branching in continuous time should be examined, and stochastic semantics and quantitative Monte Carlo simulations [25,5], especially the kappa simulator [11], are naturally applicable in this domain. (2) Although our interest in the interactions of nucleic acids is to capture an autonomous computational mechanism [15] of RNAi, which in this paper results from the stochastic process calculus of the kappa, the interactions via hydrogen bonds also play fundamental roles in Cardelli-Phillips' strand displacement calculus [24] for designing DNA circuits. A challenging angle for future work would be to design analogous information-processing circuits for RNA. (3) The dual notion of model refinement is *model abstraction* [8], which is an important tool for avoiding the

problem of combinatorial explosion. The results of this paper can be seen as the exactness of abstraction employed by local descriptions of rules. A theoretical formulation of the dual notions may characterize the exactness of an abstract model in terms of some invariance under appropriate model refinements.

Acknowledgement

References

[1] David Baulcombe, *RNA silencing in plants*, Nature. **431**, 356-63, (2004)

[2] David Baulcombe, *Amplified Silencing*, Science **315**, 199-200 (2007)

[3] C.T. Bergstrom, E. McKittrick and R. Antia, *Mathematical models of RNA silencing: unidirectional amplification limits accidental self-directed reactions*, Proc. Natl. Acad. Sci. USA 100(20), 11511-11516 (2003)

[4] Peter Brodersen and Olivier Voinnet, *The diversity of RNA silencing pathways in plants*, TRENDS in Genetics 22(5), 268-280 (2006)

[5] Luca Cardelli, *On process rate semantics*, Theor. Comput. Sci. **391**(3): 190-215 (2008)

[6] G. Cuccato, A. Polynikis, V. Siciliano, M. Graziano, M. di Bernardo and D. di Bernardo, *Modeling RNA interference in mammalian cells*, BMC Systems Biology 2011, **5**:19

[7] Vincent Danos, Jérôme Feret, Walter Fontana, Russell Harmer and Jean Krivine. *Rule-based modelling, symmetries, refinements*, (FMSB 2008), Lecture Notes in Bioinformatics **5054**. 103-122 (2008), Springer.

[8] Vincent Danos, Jérôme Feret, Walter Fontana, and Jean Krivine, *Abstract Interpretation of Cellular Signalling Networks*, Verification, Model Checking and Abstract Interpretation, VMCAI'08. Lecture Notes in Computer Science **4905**, Springer pp 83-97 (2008).

[9] Vincent Danos, Jérôme Feret, Walter Fontana, Russell Harmer and Jean Krivine, *Rule-based modelling of cellular signalling*, (CONCUR 2007), LNCS **4730**, Springer (2007)

[10] Vincent Danos and Cosimo Laneve, *Core formal molecular biology*, Proc. 12th. ESOP, LNCS **2618**, 302-318, Springer, 2003.

[11] Vincent Danos, Jérôme Feret, Walter Fontana, and Jean Krivine. Scalable simulation of cellular signaling networks. In Proc. APLAS07, LNCS **4807**, 139-157, 2007.

[12] Jérôme Feret, Vincent Danos, Jean Krivine, Russ Harmer and Walter Fontana, *Internal coarse-graining of molecular systems*, Proc. Natl. Acad. Sci. USA, (2009), **106** (16)

[13] M.A.C. Groenenboom, A.F.M. Marée and P. Hogeweg, *The RNA Silencing Pathway: The Bits and Pieces That Matter*, PLoS Comput. Biol. 1(2), 155-165 (2005)

[14] P. Haccou, P. Jagers and V A. Vatutin, "Branching Processes: Variation, Growth, and Extinction of Populations", Cambridge University Press (Cambridge Studies in Adaptive Dynamics) (2007)

[15] Masahiro Hamano, *RNA interference and Register Machines (extended abstract)*, Proceedings of 6th Workshop on Membrane Computing and Biologically Inspired Process Calculi (MecBIC2012), Electronic Proceedings in Theoretical Computer Science 100, (2012), 107112.

[16] Russ Harmer, Vincent Danos, Jérôme Feret, Jean Krivine and Walter Fontana, *Intrinsic Information carriers in combinatorial dynamical systems*, Chaos, 2010, **20** (3), pp.037108

[17] Richard A. Jorgensen, *RNA traffics information systemically in plants*, Proc. Natl. Acad. Sci. USA, **99**(18) 11561-11563 (2002)

[18] E. Levine, Z. Zhang, T. Kuhlman, and T. Hwa, *Quantitative characteristics of gene regulation by small RNA*, PLoS Biol **5**, e229 (2007)

[19] Ian J. MacRae et al, *Structural Basis for Double-Stranded RNA Processing by Dicer*, Science 311(5758), 195-198. (2006)

[20] C. J. Mode, "Multitype Branching Processes- Theory and Applications" , American Elsevier, New York, NY, USA, 1971

[21] Elaine Murphy, Vincent Danos, Jérôme Feret, Russ Harmer and Jean Krivine, *Rule-based modelling and model refinement*, in " Elements of Computational Systems Biology". Ed by H.M.Lodhi and S.H.Muggleton, Wiley (2010).

[22] J. R. Norris, "Markov Chains", Cambridge University Press, Cambridge Series in Statistical and Probabilistic Mathematics (1998)

[23] Julia Pak and Andrew Fire, *Distinct Populations of Primary and Secondary Effectors During RNAi in C. elegans*, Science. **315**, 241-244 (2007)

[24] Andrew Phillips and Luca Cardelli, A programming language for composable DNA circuits, Journal of the Royal Society Interface, 6:S419-S436, (2009)

[25] C. Priami, A. Regev, E. Shapiro, W. Silverman, *Application of a stochastic name-passing calculus to representation and simulation of molecular processes*, Information Processing Letters **80**, 25-31. 2001.

[26] Olivier Voinnet, *Use tolerance and avoidance of amplified RNA silencing by plants*, Trends Plant Sci., **13**(7), 317-328, 2008.