Stochastic Transcription Elongation via Rule Based Modelling

Masahiro HAMANO
hamano@jaist.ac.jp

SASB, Saint Malo 2015
Purpose of This Talk

Rule based investigation for

• Equilibrium Kinetics and Steady State Dynamics of Transcription Elongation such as \{ detailed balance (reversibility) \}
  Wegscheider condition

• Michaelis-Menten Enzyme Kinetics via Probability/Energy of Boltzmann Distribution
mechano-chemical TE as Rule Based Modelling

- TE is a stochastic mechano-chemical interaction consisting of
  - Brownian ratchet mechanism
  - Chemical Reactions: NTP hydrolysis/catalysis and PPI release

![reaction equation]

- This mechano-chemical principle is represented by rule based $\kappa$-calculus syntactically and semantically
Table of Contents

- Transcription Elongation Complex (TEC)
- Rule Based Description in Terms of TEC's Compact Active Sites
- Master Equation Semantics and Abstraction
- Michaelis-Menten Kinetics for TE via Boltzmann distribution (energy/probability)
Table of Contents

- Transcription Elongation Complex (TEC)
  - TE is described by Rule Based $\kappa$ in terms of TEC's compact active sites
- Master Equation Semantics and Abstraction
- Michaelis-Menten Kinetics for TE via Boltzmann distribution (energy/probability)
Transcription Elongation (TE)

- RNAP moves along the template DNA
- (polymerization) NTP unit is added to the (3'-end of) the nascent RNA.

![Structure of Transcription Elongation Complex (TEC)](image-url)
Translocations of TEC (terminology)

- **Back- and forward-track of TEC**
  (The active sites of) TEC moves back (toward 3‘end of DNA) and forth (toward 5’-end) along the DNA template without polymerizing and depolymerizing the transcript.

- **Pre- and post-translocation of TEC**
  - **pre-translocation** = the state immediately after the polymerization
  - **post-translocation** = one forward track step from pre-translocation so that the active site is in position to catalyze the next nucleotide.
Brownian Ratchet Mechanism of Elongation

\[ \text{TEC}_{(n, pre)} \xrightarrow[k_1]{k_1} \text{TEC}_{(n, post)} \xrightarrow[k_2]{NTP} \text{TEC}_{(n, post)} \cdot \text{NTP} \xleftarrow[\text{PPi}]{P} \text{TEC}_{(n+1, pre)} \]

\[ \text{TEC}_{(n, post)} \cdot \text{NTP} \xrightarrow[k_3]{k_3} \text{TEC}_{(n+1, pre)} \cdot \text{PPi} \xleftarrow[k_4]{k_4} \text{TEC}_{(n+1, pre)} \]

Table of Contents

- Transcription Elongation Complex (TEC)
- Rule Based Description in Terms of TEC's Compact Active Sites
- Master Equation Semantics and Abstraction
- Michaelis-Menten Kinetics for TE via Boltzmann distribution (energy/probability)
Stochastic Process Calculi
and their Markov Semantics

Stochastic $\pi$ (Priami, Regev, et al.)
Rule based $\kappa$ (Danos, Feret, Harmer, Krivine, et al.)

Each channel (site) is equipped with an interaction rate of exponential time distribution.

Their Semantical Counterpart of
Chemical Master Equations and Markov processes

U

Stochastic Petri Nets
(e.g., mass action semantics for chemical reactions among species by Danos-Oury)

U

Theory of Waiting Queues
(e.g., Poisson time consuming for customers waiting line)
Stochastic Process Calculi Provide Discrete/Stochastic Modelling in Biology

Multi-step but non-uniform molecular interaction for

\[ \text{stochastic fluctuation} \]
\[ \text{non Poisson time consuming} \]
\[ \text{decay of particular events} \]

In this talk,

Transcription \[ \{ \text{Initiation} \]
\[ \text{Elongation} \]
\[ \text{Termination} \]
Single-nucleotide-level description

1 nucleotide as an agent (process)

E.g., my preceding paper in SASB’12 RNAi was represented by $\kappa$:

agent = single nucleotide N with phosphate and hydrogen bondings

The description yields semantically multitype branching processes determining (non)extinction of certain classes of agents
**Nucleotide and TEC’s Muti-Partite Active Sites as Agents**

<table>
<thead>
<tr>
<th>nucleotide N</th>
<th>window frame W for TEC m-p act sts</th>
<th>W with counters (n,m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N(\ell, r)$</td>
<td>$W$</td>
<td>$W^m_n$</td>
</tr>
</tbody>
</table>

- Post-translocation: $W[N(r), \emptyset]$ schematically $N_{OH}$
- Pre-translocation: $W[N(r^1), N(\ell^1, r)]$ schematically $N - N_{OH}$
- Loading: $W[N(r), NTP]$ schematically $N_{OH}NTP$

The ratchet pathway as Rules I

Finest grained rules $\text{Elong}_N^Z$

\[
\begin{align*}
&W_n^3 [\emptyset, \emptyset] \\
&W_n^2 [\emptyset, \emptyset] \\
&W_n [N(r^1), N(\ell^1, r)] \leftrightarrow W_n [N(r), \emptyset] \leftrightarrow W_{n+1} [N(r), \text{NTP}] \leftrightarrow W_{n+1}^{\text{PPi}} [N(r^1), N(\ell^1, r)] \\
&W_{n+1} [N(r^1), N(\ell^1, r)] \leftrightarrow W_{n+1} [N(r^1), N(\ell^1, r)] \\
&W_{n-1} [N(r^1), N(\ell^1, r^1)] \\
&W_{n-2} [N(r^1), N(\ell^1, r^1)] \\
&\vdots
\end{align*}
\]
The ratchet pathway as Rules II:
To coarser/coarsest grained rules

Accordingly to forgetting the sites \( m \) and \( n \) of \( W \)

\[
W^m_n \rightarrow W_n \rightarrow W,
\]

rules are abstracted (collapsed) into:

\[
\text{Elong}_{Z_N} \rightarrow \text{Elong}_{N} \rightarrow \text{Elong}
\]
The ratchet pathway as Rules III

\[
\text{Elong}_{\mathbb{Z}^N} \longrightarrow \text{Elong}_{\mathbb{N}}
\]

\[
W_n \left[ N(r), \emptyset \right] \leftrightarrow W_{n+1} \left[ N(r), N(r^1) \right] \leftrightarrow W_{n+1} \left[ N(r), N(r^1) \right] \leftrightarrow W_{n+1} \left[ N(r), N(r^1) \right] \leftrightarrow W_{n+1} \left[ N(r), N(r^1) \right]
\]
The ratchet pathway as Rules IV:

Finally, coarsest grained rules Elong

\[
\begin{align*}
&W[\emptyset, \emptyset] \\
&W[N(r), \emptyset] \\
&W[N(r^1), N(l^1, r)] & W[N(r^1), N(l^1, r^1)] & W^\text{PPI}[N(r^1), N(l^1, r)] & W[N(r), \text{NTP}]
\end{align*}
\]
Conversely, more complicated pathway by augmenting sites on $W$

$$\text{TEC}_{(n,\text{pre})} \cdot \text{NTP} \xrightleftharpoons[k_{-i_1}]{\kappa_{i_1}} \text{TEC}_{(n,\text{post})} \cdot \text{NTP}$$

$$\text{TEC}_{(n,\text{pre})} \xrightleftharpoons[k_{-i_2}]{\kappa_{i_2}} \text{TEC}_{(n,\text{post})}$$

Bai et al, Kinetic modeling of transcription elongation, In “RNA polymerases as molecular motors” in the Royal Society of Chemistry (2009)

In our rule based framework

$$W = \frac{P_{site}A_{site}E_{site}}{}$$

$$W_n [N(r^1), N(\ell^1, r), \text{NTP}] \leftrightarrow W_n [N(r), \text{NTP}, \emptyset]$$

$$W_n [N(r^1), N(\ell^1, r), \emptyset] \leftrightarrow W_n [N(r), \emptyset, \emptyset]$$
• Rule based $\kappa$ syntax has uniformly derived variation of pathway of TEC.

• How about semantical counterpart?
Table of Contents

- Transcription Elongation Complex (TEC)
- Rule Based Description in Terms of TEC's Compact Active Sites
- Master Equation Semantics and Abstraction
- Michaelis-Menten Kinetics for TE via Boltzmann distribution (energy/probability)
Markov Process and Stationary Distribution

Transition probability $p_t(i,j)$ of Markov process is governed by

(Kolmogorov fwd equation) $p'_t = p_tQ$ with an infinitesimal matrix $Q$.

- **stationary distribution**
  
  $\pi$ is a solution of
  (invariance)
  
  $\pi p_t = \pi$ for all $t$

  $\pi Q = 0$

  i.e.,

  $\pi(j) \sum_k q(j, k) = \sum_k \pi(k) q(k, j)$

  rate of leaving $j = \text{rate of arriving } j$

- **detailed balance, i.e., reversibility** (invariance under time-reversing)

  $\pi(j) q(j, k) = \pi(k) q(k, j)$
Wegscheider Condition Is Chemist's Algorithm to Check Reversibility

For every cycle path $\phi$

$$\prod_{(i,j)\in\phi} \frac{q(i,j)}{q(j,i)} = 1$$

rewritten by

$$\sum_{(i,j)\in\phi} \log \frac{q(i,j)}{q(j,i)} = 0$$

cf. Danos-Oury's undecidability of the W condition
Chemical Master Equation (CME): Kolmogorov Eq Arisen from Chemical Reactions

CME

\[
\frac{d}{dt} p_t(x) = \sum_k \{ \lambda_k(x - (\nu'_k - \nu_k)) p_t(x - (\nu'_k - \nu_k)) - \lambda_k(x) p_t(x) \}
\]

(\nu'_k - \nu_k) \text{ stoichiometry (state-change) matrix}

\lambda_k \text{ propensity function}

Though analytically unsolvable CME

Behaviour of average yields a continuous model

\[
< x_i(t) > = \sum_x x_i p_t(x)
\]

\[
\frac{d}{dt} < x_i(t) > = \sum_{i'} (\nu'_{i'k} - \nu_{i'k}) < \lambda_k(x(t)) >
\]

reaction rate equation

\[
= \sum_k (\nu'_{ik} - \nu_{ik}) \lambda_k(x(t))
\]

no fluctuation hypothesis
State Transition for Elongation

\[
\begin{align*}
\frac{d}{dt} w_{n}^{[0]} &= (aw_{n}^{[-1]} + \alpha_{4}w_{n-1}^{[0]} + \beta_{4}w_{n+1}^{[0]} + \alpha_{2}w_{n}^{[\bullet]} + bw_{n}^{[1]}) \\
&\quad - (\alpha_{4} + \beta_{2} + a + b + \beta_{4})w_{n}^{[0]} \\
\frac{d}{dt} w_{n}^{[\bullet]} &= (\alpha_{3}w_{n-1}^{[\bullet]} + \beta_{3}w_{n+1}^{[\bullet]} + \beta_{2}w_{n}^{[0]} + \alpha_{1}w_{n-1}^{[1]}) \\
&\quad - (\alpha_{3} + \beta_{3} + \alpha_{2} + \beta_{1})w_{n}^{[\bullet]} \\
\frac{d}{dt} w_{n}^{[1]} &= (\beta_{1}w_{n+1}^{[1]} + aw_{n}^{[0]}) - (b + \alpha_{1})w_{n}^{[1]}
\end{align*}
\]

Evolution on Back-Track Pathway

\[
\begin{align*}
\frac{d}{dt} w_{n}^{[-(n-1)]} &= bw_{n}^{[-(n-2)]} - aw_{n}^{[-(n-1)]} \\
\frac{d}{dt} w_{n}^{[-j]} &= (aw_{n}^{[-(j+1)]} + bw_{n}^{[-(j-1)]}) - (a + b)w_{n}^{[-j]} \quad \text{for } 1 \leq j < n - 1
\end{align*}
\]
State Transition for $\text{Elong}_N$

Evolution on back-track is reduced by summing over the positions $j$'s

$$w_n^{[-]} = \sum_{1 \leq j \leq n-1} w_n^{[-j]} \quad \text{and} \quad \frac{d}{dt} w_n^{[-]} = \sum_{1 \leq j \leq n-1} \frac{d}{dt} w_n^{[-j]} = bw_n^{[0]} - aw_n^{[-1]}$$
Finally, State Transition for Elong

Evolution on main-track is reduced by summing over the lengths $n$'s

\[
\begin{align*}
\frac{d}{dt}w^{[0]} &= aw^{[-1]} + \alpha_2 w^{[\bullet]} + bw^{[1]} - (\beta_2 + a + b)w^{[0]} \\
\frac{d}{dt}w^{[\bullet]} &= \beta_2 w^{[0]} + \alpha_1 w^{[1]} - (\alpha_2 + \beta_1)w^{[\bullet]} \\
\frac{d}{dt}w^{[1]} &= \beta_1 w^{[\bullet]} + aw^{[0]} - (b + \alpha_1)w^{[1]}
\end{align*}
\]
Equilibrium Distribution of Elong

Wegscheider condition:

\[(a/b) (\alpha_1/\beta_1) (\alpha_2/\beta_2) = 1\]

so that equilibrium is given with a relax time \(\tau\)

\[w_{*}^{[0]} = b\beta_1 \tau \quad w_{*}^{[1]} = a\beta_1 \tau \quad w_{*}^{[\bullet]} = a\alpha_1 \tau \quad \tau = 1/(b\beta_1 + a\beta_1 + a\alpha_1)\]
Quasi-steady state approximation

How sound is the model abstraction arisen from forgetting sites?

The finest grained rules may be approximately retrieved from the coarsest ones under certain biological assumption.

Quasi-Steady State approximation

Under the assumption that poly/depoly merization rates $\ll$ translocation rates, the translocation rates may be run off in our time scale.

\[ w_n^{[j]}(t) \approx w_*^{[j]} \times w_n(t) \quad \text{where} \quad w_n = \sum_{j \in \{1,0,-,\emptyset\}} w_n^{[j]} \]

Summing evolutions on main (res. backtrack) pathways res. of $\text{Elong}_N^Z$ (res.$\text{Elong}_N^N$) yields a simple birth-death master equation with the effective rates $\theta_1/\theta_2$ for poly/depoly merization

\[ \frac{d}{dt} w_n(t) = \theta_+ w_{n-1} - (\theta_+ + \theta_-) w_n + \theta_- w_{n+1} \]

in which $\theta_+ = \alpha_4 b \beta_1 + \alpha_3 a \alpha_1 + \alpha_1 a \beta_1$ and $\theta_- = \beta_4 b \beta_1 + \beta_3 a \alpha_1 + \beta_1 a \alpha_1$
Table of Contents

- Transcription Elongation Complex (TEC)
- Rule Based Description in Terms of TEC's Compact Active Sites
- Master Equation Semantics and Abstraction
- Michaelis-Menten Kinetics for TE via Boltzmann distribution (energy/probability)
Boltzmann distribution

correspondence of probability $\pi$ and energy $\varepsilon$

$$w_n^{[j]} = \pi(W_n^j) = \frac{1}{Z_n} \exp\left(\frac{-\mathcal{E}(W_n^j)}{k_B T}\right)$$

with $$Z_n = \sum_j \exp\left(\frac{-\mathcal{E}(W_n^j)}{k_B T}\right)$$
The main reaction pathway is the following part

In our transition diagram with \( a = k_1, b = k_{-1}, \alpha = k_2, \beta = k_2 \) [NTP]

The generator \( Q \) of Kolmogorov back eqn \( p'(t) = Qp(t) \) is given for a fixed \( n \):

\[
Q = \begin{pmatrix}
-(q_{11} + q_{01}) & q_{11} & q_{10} \\
q_{01} & -q_{10} & 0 \\
0 & 0 & -q_{10}
\end{pmatrix}
\]

\[
p = (p_i)_{i=1,\bullet,0}
\]

\( p_1 \) is \( W_n^i \)
The main reaction pathway as Michaelis–Menten kinetics

M–M kinetics is derivable from the stationary distribution of Kolmogorov (Kolmogorov backward equation)

\[ p'(t) = Qp(t) \]

First, stationary distribution \( p'(t)=0 \) yields via \( q_1 = k_2 \) and \( q_1 = k_2[NTP] \)

\[ p_0 = \frac{1}{q_1 (1+q_{01}/q_{10})+1} = \frac{[NTP]}{k_2 (1 + q_{01}/q_{10}) + [NTP]} \]

Second, the velocity \( v \) of TEC(n+1,pre) is given so that \( v/v_{max} = p_0 \) Michaelis–Menten

\[ v = \frac{v_{max}[NTP]}{k_2 (1 + q_{01}/q_{10}) + [NTP]} = \frac{a/b}{K_{eq}} \] equilibrium const for translocation
More generally combining M–M and Boltzmann

Assuming bck/fwd translocations are in equilibrium:

**Steady–state elongation rate**

\[ v = \frac{v_{max} [\text{NTP}]}{K_n + [\text{NTP}]} \quad \text{with} \quad K_n = \frac{k_{-2}}{k_2} \log \prod_j \frac{w_n^{[j]}}{w_n^{[0]}} \]

\[ \sum_j \exp \left( (\mathcal{E}(W_n^0) - \mathcal{E}(W_n^j))/k_B T \right) \]

energy difference
Future Work

More conformationally faithful details to realize various pawls in TEC

E.g., G-loop configuration to control bending/straightening F-helix

Future Work 2

RNAp as molecular motor

How to augment drift (thermal force) on top of \( \kappa \) description of Brownian motion ??

probability \( p(x,t) \)

probability flux given by Fick’s law wiith drift;

\[
j(x,t) = -D \frac{\partial p}{\partial x} + \frac{F}{\gamma} p(x,t)
\]

where diffusion coef. \( D \) and external force \( F(x,t) \) applied to RNAp with \( \gamma \) drag coef.

Fokker-Planck eqn

\[
\frac{\partial p}{\partial t} = -\frac{\partial j}{\partial x} = D \frac{\partial^2 p}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\partial U}{\partial x} p \right)
\]

1st eqn is flux continuity eqn

(Note: \( F = -\frac{\partial U}{\partial x} \) with energy \( U \))
Future Work 2 (contd.)

Why Fokker-Planck?

At equilibrium of FP-eqn:

\[ 0 = D \frac{dp}{dx} + \frac{1}{\gamma} \frac{dU}{dx}p \]

Boltzmann distribution is a solution

\[ p(x) \propto \exp \left( -\frac{U(x)}{D\gamma} \right) \]

Towards more energy \( U(x) \) efficient description of elongation (e.g., chemist's use of energy-landscape)