

Many-body effects on gene expression

Hossein Hosseini

December 4, 2016

Table of contents

- 1 Background
- 2 Single RNA polymerase on a DNA strand
- 3 Two RNA polymerases on a DNA strand

Transcription

- Three compartment of transcription process
 - ▶ Initiation
 - ▶ Elongation
 - ▶ Termination
- Pausing
- traffics of RNA polymerases

Transcription

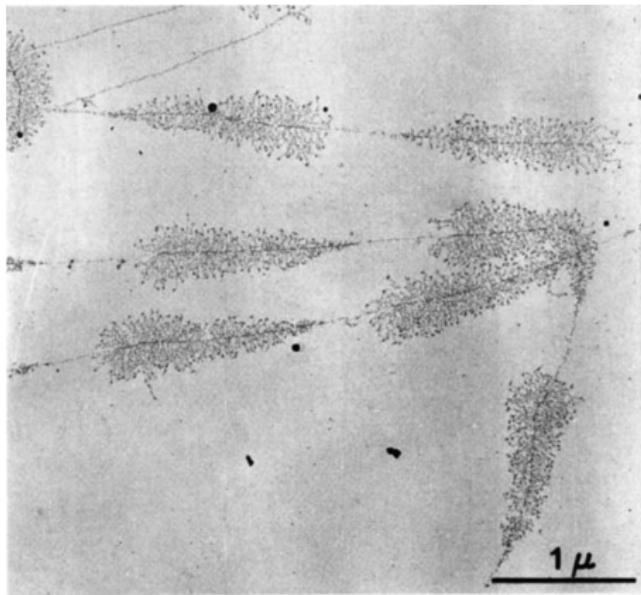
- Three compartment of transcription process
 - ▶ Initiation
 - ▶ Elongation
 - ▶ Termination
- Pausing
- traffics of RNA polymerases

Transcription

- Three compartment of transcription process
 - ▶ Initiation
 - ▶ Elongation
 - ▶ Termination
- Pausing
- traffics of RNA polymerases

Transcription

- Three compartment of transcription process
 - ▶ Initiation
 - ▶ Elongation
 - ▶ Termination
- Pausing
- traffics of RNA polymerases



An example of gene expression modeling

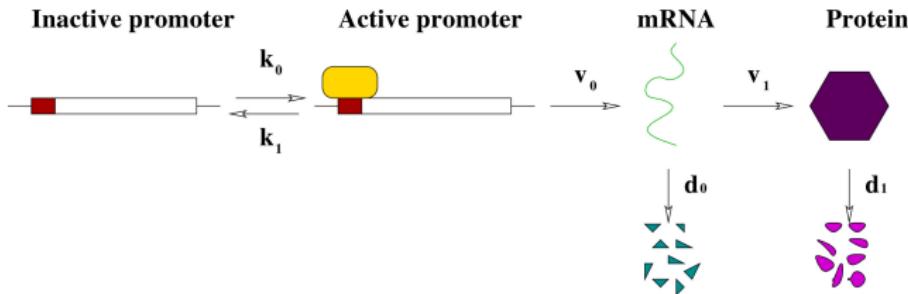


Figure: Three-step model of gene expression [Shahrezaei and Swain, 2009].

$$\begin{aligned}\frac{\partial P_{m,n}}{\partial t} = & v_0(P_{m-1,n} - P_{m,n}) + v_1 m (P_{m,n-1} - P_{m,n}) \\ & + d_0[(m+1)P_{m+1,n} - mP_{m,n}] \\ & + d_1[(n+1)P_{m,n+1} - nP_{m,n}]\end{aligned}$$

Figure: Master equation for a two-step model of gene expression [Shahrezaei and Swain, 2009].

Protein distributions

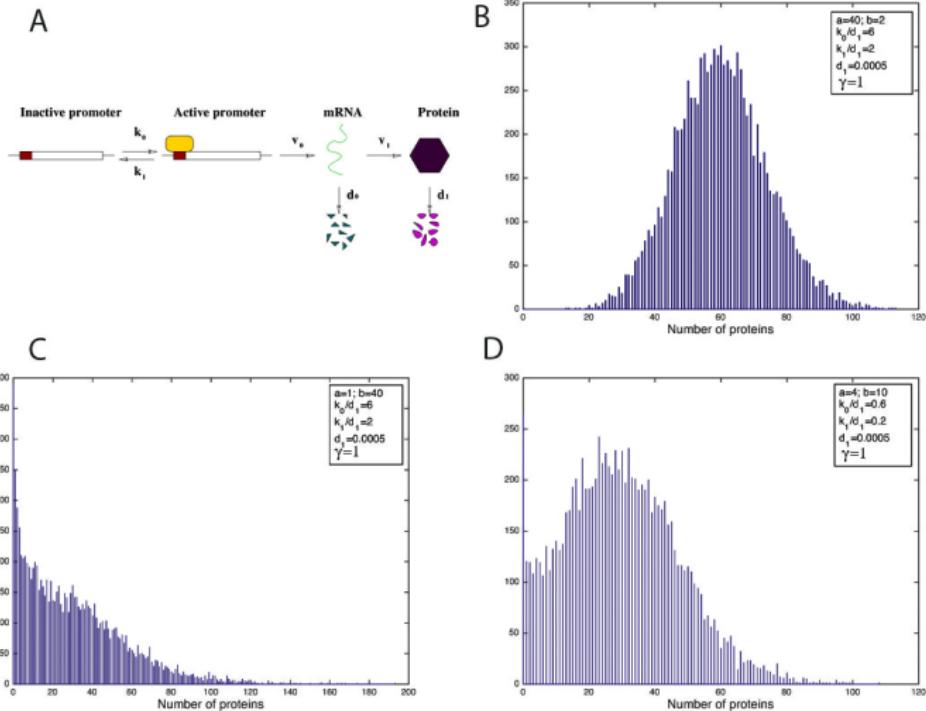


Figure: Stochastic simulation of protein distributions. Reproduced graphs of [Shahrezaei and Swain, 2009].

Modeling of transcription kinetics

- Three stages of transcription are considered mostly separately.
- Translocation mechanism: Brownian ratchet vs. power stroke models
- Sequence sensitivity for the motion of RNA polymerase.

Traffic of RNA polymerases

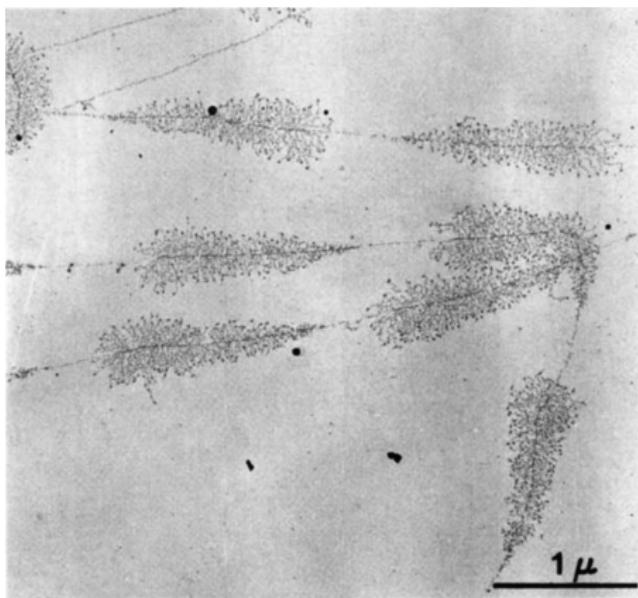


Figure: Electron micrograph image of a segment of unwound nucleolar core from *Triturus viridescens* [Hamkalo and Miller, 1973].

Modeling of RNA ploymerses traffic

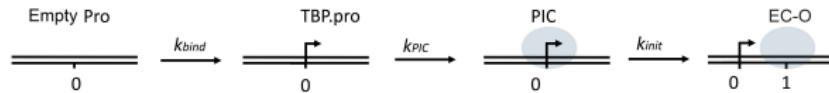
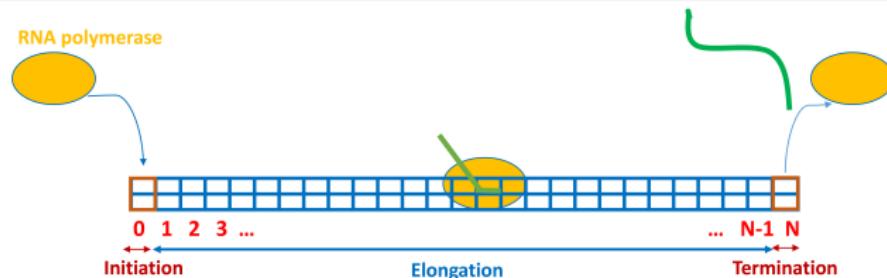
- Co-directional and contra-directional traffic (transcriptional interference)
- TASEP modeling

Objectives

- How is the kinetics of a single RNAPII during the transcription process?
- How does a strong pause region located on a gene affect the transcription process?
- How does the interaction of two elongating RNA polymerases affect the spatio-temporal organization of RNA polymerases?

Initiation

Transcription process



Initiation

Analytical solution for $P_{PIC}(t)$

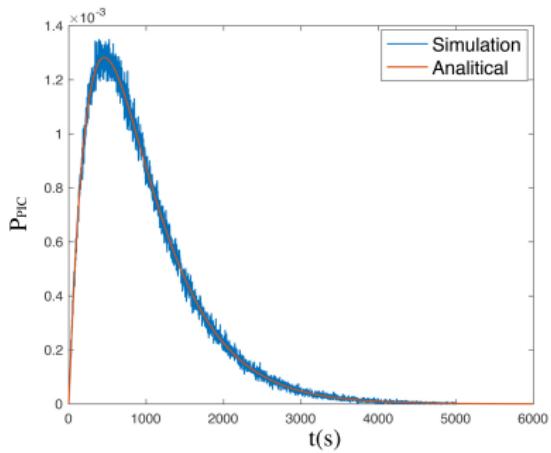
$$P_{PIC}(t) = \frac{k_{PIC} k_{bind} ((k_{PIC} - k_{init}) e^{-k_{bind} t} + (k_{init} - k_{bind}) e^{-k_{PIC} t} - (k_{PIC} - k_{bind}) e^{-k_{init} t})}{(k_{PIC} - k_{bind})(k_{bind} - k_{init})(k_{init} - k_{PIC})}$$

$k_{init} P_{PIC}(t)$ is the rate of probability which transits through $x = 0$, the left-hand edge of the elongation compartment.

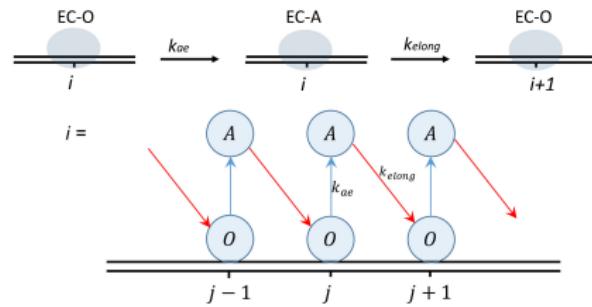
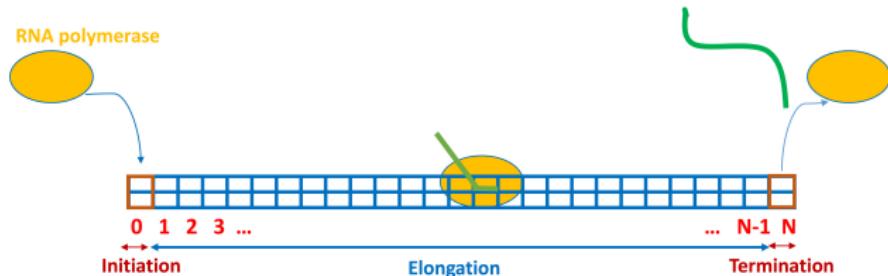
probability density distribution

Value of rate constants for initiation

$$k_{PIC} = 0.0029 s^{-1}, k_{bind} = 0.0016 s^{-1}, k_{init} = 0.6 s^{-1}$$



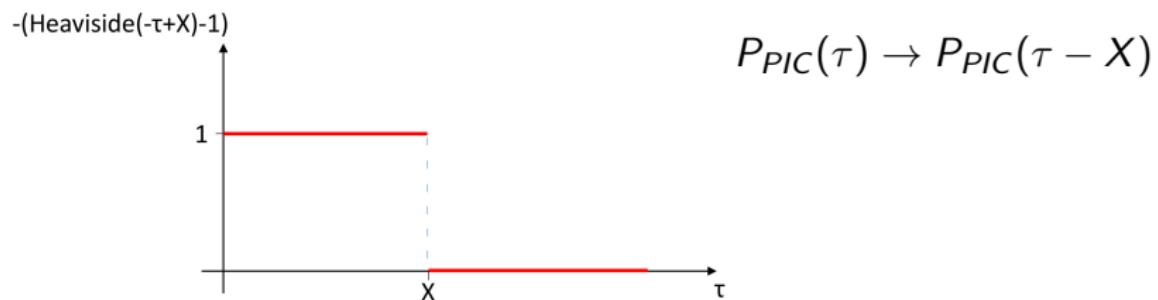
Elongation



Approximate solution

The zero-order approximation

$$q(X, \tau) = \begin{cases} -\frac{k_{init}}{a} (\text{Heaviside}(-\tau + X) - 1) P_{PIC}\left(\frac{N}{a}\tau - X\right) & 0 < X < 1 - \epsilon \\ -\frac{k_{init}}{a} (\text{Heaviside}(-\tau + 1) - 1) P_{PIC}\left(\frac{N}{a}\tau - 1\right) & 1 - \epsilon < X < 1 \end{cases}$$



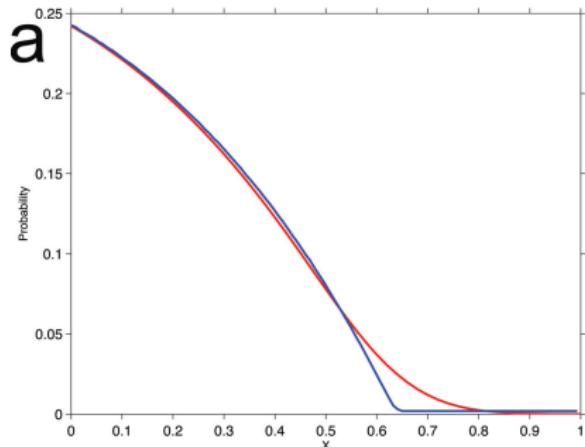
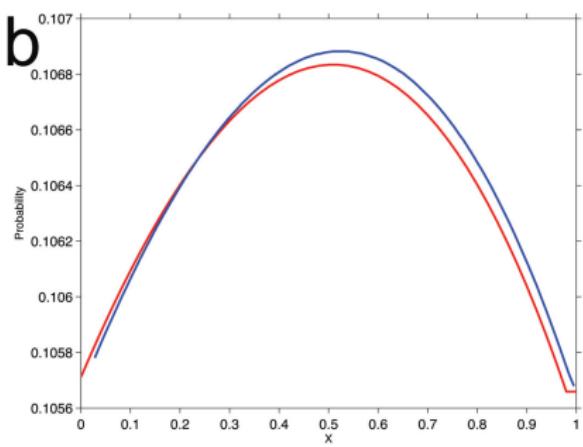
Comparison with numerical solution

The parameter values:

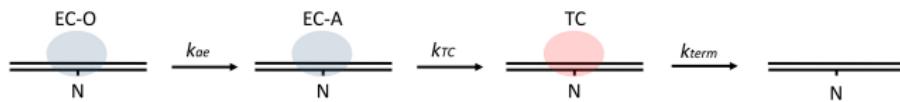
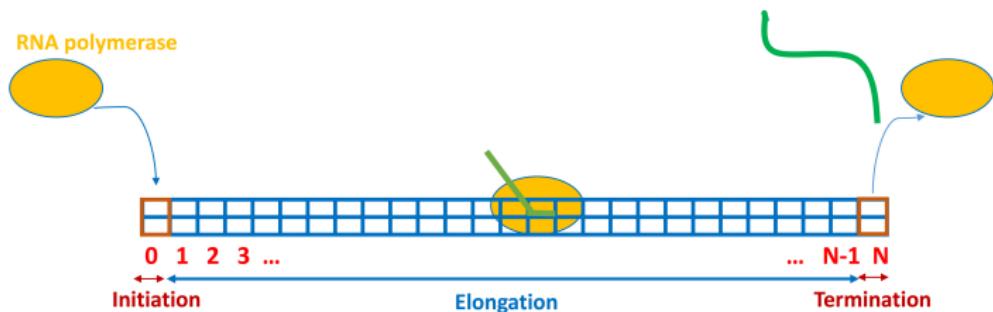
$$\begin{aligned}k_{ae} &= 144\text{s}^{-1}, k_{elong} = 144\text{s}^{-1} \\k_{PIC} &= 0.0029\text{s}^{-1}, k_{bind} = 0.0016\text{s}^{-1} \\k_{init} &= 0.6\text{s}^{-1}, \tau = 3.83, N = 1000.\end{aligned}$$

The parameter values:

$$\begin{aligned}k_{ae} &= 250\text{s}^{-1}, k_{elong} = 30\text{s}^{-1} \\k_{PIC} &= 0.0029\text{s}^{-1}, k_{bind} = 0.0016\text{s}^{-1} \\k_{init} &= 0.6\text{s}^{-1}, \tau = 0.64, N = 1000.\end{aligned}$$



Termination



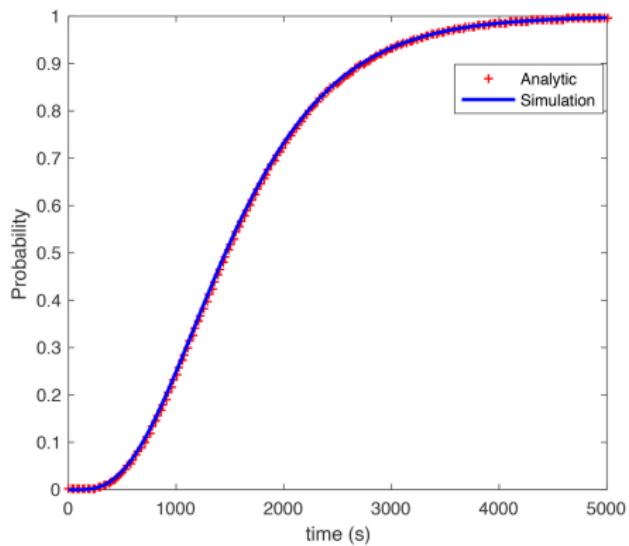
Master equations for termination

$$\frac{dP_A(t)}{dt} = J(x = N, t) - k_{TC} P_A(t)$$

$$\frac{dP_{TC}(t)}{dt} = k_{TC} P_A(t) - k_{term} P_{TC}(t)$$

$$\frac{dP_T(t)}{dt} = k_{term} P_{TC}(t)$$

Analytical vs. stochastic



Velocity profile

$$v(x) = a \left(1 - \frac{A^n}{(x - x_P)^n + A^n}\right) \left(1 - \frac{B^m}{(x - x_P)^m + B^m}\right) \quad (1)$$

Numerical results

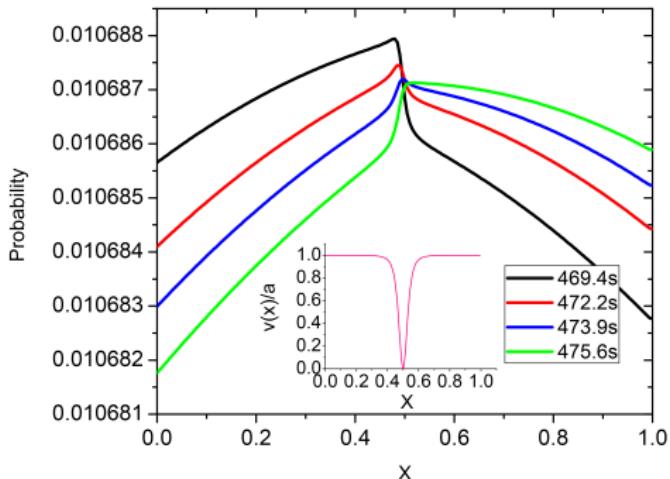


Figure: Effect of a strong pause on the middle of gene on the probability distribution. The parameter

values: $k_{ae} = 144 s^{-1}$, $k_{elong} = 144 s^{-1}$, $k_{PIC} = 0.0029 s^{-1}$, $k_{bind} = 0.0016 s^{-1}$ and $k_{init} = 0.6 s^{-1}$, $n = m = 2$, $A = B = 2$, $x_P = \frac{1}{2}$ and $N = 1000$.

Two elongating RNA polymerases on a DNA strand

Initiation: The leading polymerase

- Chemical reactions, master equations and the resultant $P_{PIC}^{(1)}(t)$ are all the same as the single RNA polymerase initiation.

Initiation: The leading polymerase

- Chemical reactions, master equations and the resultant $P_{PIC}^{(1)}(t)$ are all the same as the single RNA polymerase initiation.
- The flux departing the initiation part and entering the elongation part is $k_{init}P_{PIC}^{(1)}(t)$

$$a p(x, t) - \frac{D}{2} \frac{\partial p(x, t)}{\partial x} = k_{init} P_{PIC}^{(1)}(t)$$

Initiation: The trailing polymerase

- Master equations

$$\frac{dP_{pro}^{(2)}(t)}{dt} = -k_{bind} Q(t) P_{pro}^{(2)}(t)$$

$$\frac{dP_{TBP.pro}^{(2)}(t)}{dt} = k_{bind} Q(t) P_{pro}^{(2)}(t) - k_{PIC} P_{TBP.pro}^{(2)}(t)$$

$$\frac{dP_{PIC}^{(2)}(t)}{dt} = k_{PIC} P_{TBP.pro}^{(2)}(t) - k_{init} P_{PIC}^{(2)}(t)$$

$$\frac{dP_O^{(2)}(t)}{dt} = k_{init} P_{PIC}^{(2)}(t)$$

$$\text{where } Q(t) = 1 - \left(\int_{x=0}^{\Delta} p(x, t) dx + P_{pro}^{(1)}(t) + P_{TBP.PRO}^{(1)}(t) + P_{PIC}^{(1)}(t) \right)$$

Initiation: The trailing polymerase

- Master equations

$$\frac{dP_{pro}^{(2)}(t)}{dt} = -k_{bind} Q(t) P_{pro}^{(2)}(t)$$

$$\frac{dP_{TBP.pro}^{(2)}(t)}{dt} = k_{bind} Q(t) P_{pro}^{(2)}(t) - k_{PIC} P_{TBP.pro}^{(2)}(t)$$

$$\frac{dP_{PIC}^{(2)}(t)}{dt} = k_{PIC} P_{TBP.pro}^{(2)}(t) - k_{init} P_{PIC}^{(2)}(t)$$

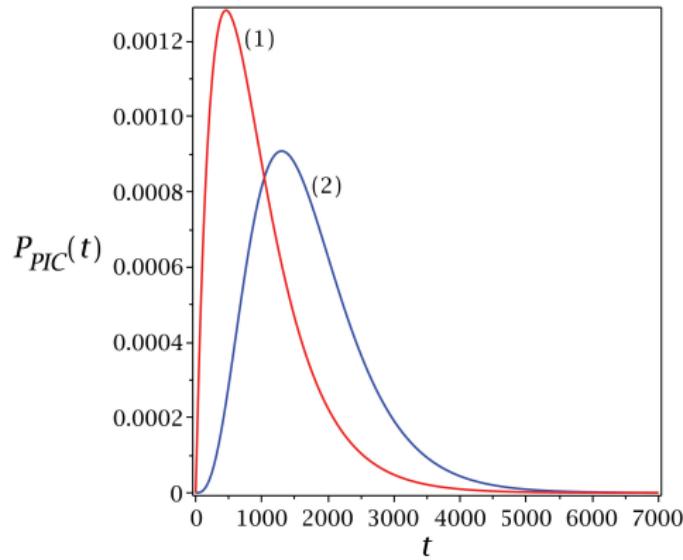
$$\frac{dP_O^{(2)}(t)}{dt} = k_{init} P_{PIC}^{(2)}(t)$$

where $Q(t) = 1 - \left(\int_{x=0}^{\Delta} p(x, t) dx + P_{pro}^{(1)}(t) + P_{TBP.PRO}^{(1)}(t) + P_{PIC}^{(1)}(t) \right)$

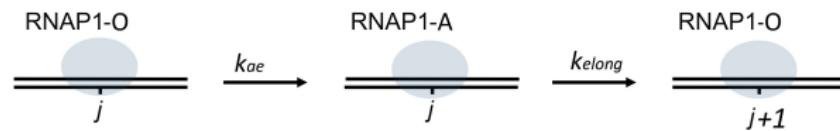
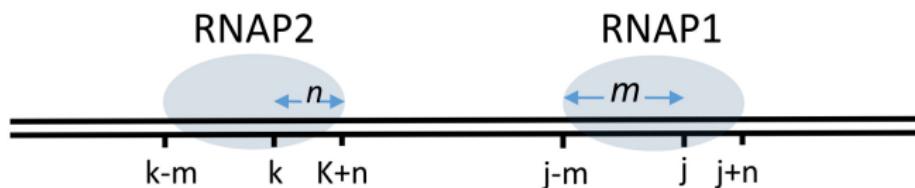
- The flux departing the initiation part and entering the elongation part is $k_{init} P_{PIC}^{(2)}(t)$

$$a p(y, t) - \frac{D}{2} \frac{\partial p(y, t)}{\partial y} = k_{init} P_{PIC}^{(2)}(t)$$

Comparing the initiation rates of the two RNAPs



Two elongating RNA polymerases on a track



$$j = 1, 2, \dots, N-1$$

$$j - k \geq \Delta \text{ where } \Delta = m + n$$

Master equations

$P(j, t; k, t)$: The joint probability distribution for the two RNA polymerases being at site j and k at time t .

Four different possible combination of states:

$$P(j, A, t; k, O, t)$$

$$P(j, A, t; k, A, t)$$

$$P(j, O, t; k, O, t)$$

$$P(j, O, t; k, A, t)$$

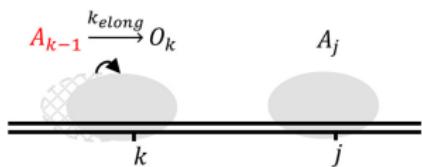
and

$$P(j, t; k, t) = \sum_1^4 \text{ all possible combination of states}$$

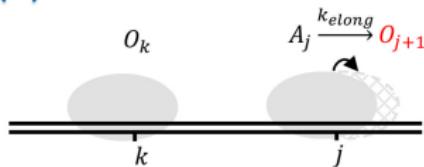
Master equation for $P(j, A, t; k, O, t)$

$$\begin{aligned}\frac{dP(j, A, t; k, O, t)}{dt} = & k_{\text{elong}} P(j, A, t; k-1, A, t) H(j - k - \Delta) \\ & - k_{ae} P(j, A, t; k, O, t) H(j - k - \Delta) \\ & + k_{ae} P(j, O, t; k, O, t) H(j - k - \Delta) \\ & - k_{\text{elong}} P(j, A, t; k, O, t) H(j - k - \Delta)\end{aligned}$$

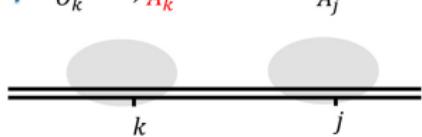
(i)



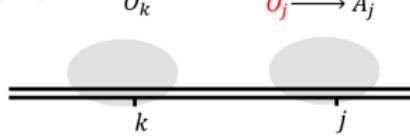
(ii)



(iii)



(iv)



Elongation: Governing equations for two RNAPs

Dimensionless PDEs

By defining dimensionless variables, $X = \frac{x}{N}$, $Y = \frac{y}{N}$, $\tau = \frac{t \cdot a}{N}$, $q(X, Y, \tau) = N p(x, y, t)$, $\alpha = \frac{\Delta}{N}$, $\beta = \frac{1}{N}$ and the Peclet number $Pe = 2N$.

1- $X - Y \geq \alpha + \beta$

$$\begin{aligned}\frac{\partial q(X, Y, \tau)}{\partial \tau} &= -\frac{\partial q(X, Y, \tau)}{\partial Y} + \frac{1}{Pe} \frac{\partial^2 q(X, Y, \tau)}{\partial Y^2} \\ &\quad -\frac{\partial q(X, Y, \tau)}{\partial X} + \frac{1}{Pe} \frac{\partial^2 q(X, Y, \tau)}{\partial X^2}\end{aligned}$$

Elongation: Governing equations for two RNAPs

Dimensionless PDEs

By defining dimensionless variables, $X = \frac{x}{N}$, $Y = \frac{y}{N}$, $\tau = \frac{t \cdot a}{N}$, $q(X, Y, \tau) = N p(x, y, t)$, $\alpha = \frac{\Delta}{N}$, $\beta = \frac{1}{N}$ and the Peclet number $Pe = 2N$.

$$1- X - Y \geq \alpha + \beta$$

$$\begin{aligned}\frac{\partial q(X, Y, \tau)}{\partial \tau} &= -\frac{\partial q(X, Y, \tau)}{\partial Y} + \frac{1}{Pe} \frac{\partial^2 q(X, Y, \tau)}{\partial Y^2} \\ &\quad - \frac{\partial q(X, Y, \tau)}{\partial X} + \frac{1}{Pe} \frac{\partial^2 q(X, Y, \tau)}{\partial X^2}\end{aligned}$$

$$2- X - Y = \alpha$$

$$\frac{\partial q(X, Y, \tau)}{\partial \tau}|_{Y=X-\alpha} = -\frac{\partial q(X, Y, \tau)}{\partial X}|_{Y=X-\alpha} + \frac{1}{Pe} \frac{\partial^2 q(X, Y, \tau)}{\partial X^2}|_{Y=X-\alpha}$$

Elongation: Governing equations for two RNAPs

$X - Y = \alpha$: no flux

$$\hat{n} \cdot \vec{J}|_{X=Y-\alpha} = 0$$

where

$$\hat{n} = \frac{\hat{X} - \hat{Y}}{\sqrt{2}}$$

$$\vec{J} = \left(-q(X, Y, \tau) + \frac{1}{Pe} \frac{\partial q(X, Y, \tau)}{\partial X} \right) \hat{X} + \left(-q(X, Y, \tau) + \frac{1}{Pe} \frac{\partial q(X, Y, \tau)}{\partial Y} \right) \hat{Y}$$

Termination

- The leading RNAP: Chemical reactions, master equations and the resultant $P_T^{(1)}(t)$ are all the same as a single RNA polymerase termination.

Termination

- The leading RNAP: Chemical reactions, master equations and the resultant $P_T^{(1)}(t)$ are all the same as a single RNA polymerase termination.
- The trailing one: The leading polymerase must leave the termination part so that it becomes possible for the the trailing one to occupy the termination part.
The equations are:

$$\frac{dP_{O,N}^{(2)}}{dt} = J^{(2)}(y = N, t) P_T^{(1)}(t) - k_{ae} P_{O,N}^{(2)}(t)$$

$$\frac{dP_{A,N}^{(2)}}{dt} = k_{ae} P_{O,N}^{(2)}(t) - k_{TC} P_{A,N}^{(2)}(t)$$

$$\frac{dP_{TC}^{(2)}}{dt} = k_{TC} P_{A,N}^{(2)}(t) - k_{term} P_{TC}^{(2)}(t)$$

$$\frac{dP_T^{(2)}(t)}{dt} = k_{term} P_{TC}^{(2)}(t)$$

Termination

- The leading RNAP: Chemical reactions, master equations and the resultant $P_T^{(1)}(t)$ are all the same as a single RNA polymerase termination.
- The trailing one: The leading polymerase must leave the termination part so that it becomes possible for the the trailing one to occupy the termination part.
The equations are:

$$\frac{dP_{O,N}^{(2)}}{dt} = J^{(2)}(y = N, t) P_T^{(1)}(t) - k_{ae} P_{O,N}^{(2)}(t)$$

$$\frac{dP_{A,N}^{(2)}}{dt} = k_{ae} P_{O,N}^{(2)}(t) - k_{TC} P_{A,N}^{(2)}(t)$$

$$\frac{dP_{TC}^{(2)}}{dt} = k_{TC} P_{A,N}^{(2)}(t) - k_{term} P_{TC}^{(2)}(t)$$

$$\frac{dP_T^{(2)}(t)}{dt} = k_{term} P_{TC}^{(2)}(t)$$

- The boundary condition for the right-end of the elongation part

$$\frac{\partial p(x, t)}{\partial x} \Big|_{x=N} = 0 \quad \frac{\partial p(y, t)}{\partial y} \Big|_{y=N} = 0$$

What is next?

- Solving the Fokker-Planck equation for two RNAPs.
- Comparing the analytical results with stochastic simulations.
-