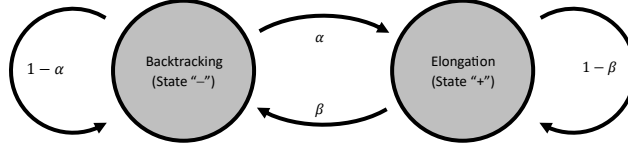


# Two-State Model of Backtracking Rescue by GreB

Xiang Gao

September 2024

In the presence of GreB, GreB may rescue a backtracked RNA polymerase (RNAP) and bring it into active elongation. Thus the stalling process consists of transitions from two states: backtracking and active elongation. The stochastic transitions between the two states may be described as reversible first-order reactions with reaction rates  $k_-$  and  $k_+$ , corresponding to the transition rate from elongation state to backtracking state and the transition rate from backtracking state to elongation state, respectively.



During backtracking, RNAP moves at a velocity  $v_-$ , during active elongation, it moves at a velocity  $v_+$ . We sought to determine  $k_-$  and  $k_+$  in terms of measured RNAP positions and measured velocities  $v_-$  and  $v_+$ .

## 1 Markov model

We first considered these two-state transitions during an infinitely small time interval  $\Delta t$  (figure). During this time, the probabilities of transitioning from ‘-’ to a ‘+’ state and from ‘+’ to ‘-’ state are  $\alpha = k_+ \cdot \Delta t$  and  $\beta = k_- \cdot \Delta t$  respectively.

These transitions follow the Markov process and the probabilities being in the ‘+’ and ‘-’ states evolve as

$$\begin{pmatrix} P_-^{(i+1)} \\ P_+^{(i+1)} \end{pmatrix} = \begin{pmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{pmatrix} \begin{pmatrix} P_-^{(i)} \\ P_+^{(i)} \end{pmatrix} \quad (1)$$

with the transition matrix:

$$T = \begin{pmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{pmatrix} \quad (2)$$

## 2 The eigenvalues and eigenvectors of transition matrix

The eigenvalues and eigenvectors of the transition matrix can be calculated as follows

$$0 = T - \lambda I = \begin{pmatrix} 1 - \alpha - \lambda & \beta \\ \alpha & 1 - \beta - \lambda \end{pmatrix} \quad (3)$$

The two eigenvalues are  $\lambda_1 = 1$  and  $\lambda_2 = 1 - \alpha - \beta$ . It is worth noting that  $1 - \alpha - \beta \in (-1, 1)$ .

The eigenvectors satisfy:

$$T\mathbf{V}_k = \lambda_k \mathbf{V}_k \quad (4)$$

where  $V_k$  is the  $k$ -th eigenvector ( $k = 1, 2$ ) that corresponds to  $\lambda_k$ . We find the eigenvectors by solving Eq. 4:

$$\mathbf{V}_1 = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}, \mathbf{V}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (5)$$

We define:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, V = (\mathbf{V}_1 \quad \mathbf{V}_2) \quad (6)$$

$$TV = T(\mathbf{V}_1 \quad \mathbf{V}_2) = (\lambda_1 \mathbf{V}_1 \quad \lambda_2 \mathbf{V}_2) = (\mathbf{V}_1 \quad \mathbf{V}_2) \Lambda = V\Lambda \quad (7)$$

Therefore

$$T = V\Lambda V^{-1} \quad (8)$$

and

$$T^N = V\Lambda^N V^{-1} \quad (9)$$

At the time  $t$ ,  $N = t/\Delta t$  number of  $\Delta t$  steps will have been taken. The  $N$ -step transition matrix can be written as:

$$\begin{aligned}
T^N &= V \Lambda^N V^{-1} \\
&= \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & 1 \\ \alpha & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1 - \alpha - \beta)^N \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \alpha & -\beta \end{pmatrix} \\
&= \frac{1}{\alpha + \beta} \begin{pmatrix} \beta + s\alpha & \beta - s\beta \\ \alpha - s\alpha & \alpha + s\beta \end{pmatrix}
\end{aligned} \tag{10}$$

where  $s = (\alpha + \beta - 1)^N$ . When  $N \rightarrow \infty$  and  $s \rightarrow 0$

$$T^\infty = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix} \tag{11}$$

which gives the probabilities when the two states are under equilibrium and is independent of the initial states.

### 3 Velocity correlation

Although we ultimately wished to find the solution of position correlation, it was mathematically more conducive to start by finding the velocity correlation. We assumed that at the starting point of measurement, the system is in equilibrium. The velocity correlation between the initial state and the  $N$ th-step can be written as  $\langle v_0 v_N \rangle$ , which equals to the average of all possible combinations of correlation of stochastic velocity variables, e.g.,

$$\begin{aligned}
\langle v_0 v_N \rangle &= v_- P_-^\infty (v_- \quad v_+) \begin{pmatrix} P_{(-|-)}^N \\ P_{(+|-)}^N \end{pmatrix} + v_+ P_+^\infty (v_- \quad v_+) \begin{pmatrix} P_{(-|+)}^N \\ P_{(+|+)}^N \end{pmatrix} \\
&= (v_- \quad v_+) T^N \begin{pmatrix} v_- P_-^\infty \\ v_+ P_+^\infty \end{pmatrix}
\end{aligned} \tag{12}$$

where  $P_-^\infty$  and  $P_+^\infty$  are the probabilities of state '-' and '+' at 0-th step respectively.  $P_{(i|j)}^N$  is the conditional probability that the state is  $j$  at 0-th step and the state becomes  $i$  after  $N$  steps. Plugging  $T^N$ ,  $P_-^\infty$ , and  $P_+^\infty$  into Eq. 12, we have

$$\langle v_0 v_N \rangle = \frac{1}{\alpha + \beta} (v_- \quad v_+) \begin{pmatrix} \beta + s\alpha & \beta - s\beta \\ \alpha - s\alpha & \alpha + s\beta \end{pmatrix} \begin{pmatrix} v_- \beta / (\alpha + \beta) \\ v_+ \alpha / (\alpha + \beta) \end{pmatrix} \tag{13}$$

To convert this expression to a continuous form, we take the limit that the time step interval  $\Delta t \rightarrow 0$ . We have thus obtained the velocity correlation as a function of time

$$\langle v_0 v_N \rangle = A + B e^{-kt} \tag{14}$$

where  $A = (\frac{k_- v_- + k_+ v_+}{k_- + k_+})^2$ ,  $B = k_- k_+ (\frac{v_- - v_+}{k_- + k_+})^2$ , and  $k = k_- + k_+$ .

## 4 Position correlation

We assume that all position trajectories,  $x(t)$ , are the samples from the same ensemble, therefore:

$$\begin{aligned}
\langle (x(t) - x(0))^2 \rangle &= \langle (\int_0^t v(t_1) dt_1)^2 \rangle = \int_0^t \int_0^t \langle v(t_1)v(t_2) \rangle dt_1 dt_2 \\
&= \int_0^t \int_0^t \langle v(0)v(|t_2 - t_1|) \rangle dt_1 dt_2 \\
&= \int_0^t \int_0^t [A + Be^{-k|t_2 - t_1|}] dt_1 dt_2 \\
&= 2 \int_0^t dt_2 \int_0^{t_2} [A + Be^{-k(t_2 - t_1)}] dt_1 \\
&= At^2 + \frac{2B}{k}t - \frac{2B}{k^2} [1 - e^{-kt}]
\end{aligned} \tag{15}$$

It is worth noting that line 2 of Eq. 15 assumes that the system has already reached equilibrium at the time  $t = 0$ , so  $\langle v(t_1)v(t_2) \rangle = \langle v(0)v(|t_2 - t_1|) \rangle$ . This is true in practice since at  $t = 0$ , we only know the probabilities that RNAP is in either an elongation state or backtracking state.

In Eq. 15, the first term is a linear drift of the mean position:

$$\bar{x}(t) = \frac{k_-v_- + k_+v_+}{k_- + k_+} \tag{16}$$

Therefore, variance of the position is

$$\sigma_x^2(t) = \langle (x(t) - x(0))^2 \rangle - \langle x(t) \rangle^2 = \frac{2B}{k}t - \frac{2B}{k^2}[1 - e^{-kt}] \tag{17}$$

When  $kt \gg 1$ , the variance evolves with time as follows

$$\sigma_x^2(t) = \frac{2B}{k}t = \frac{2k_-k_+}{k_- + k_+}^{\frac{2}{3}}(v_- - v_+)^2ts \tag{18}$$

## Reference

Ma J, Tan C, **Gao X**, Fulbright RM, Roberts JW, Wang MD. Transcription factor regulation of RNA polymerase's torque generation capacity. *Proceedings of the National Academy of Sciences of the United States of America*. 2019 Feb 12;116(7):2583-8. doi:10.1073/pnas.1807031116