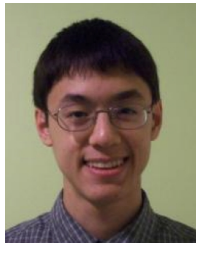
THE COMMIT-AND-DELAY MODEL

... FOR (ALMOST) EXACT, COARSE GRAINED ...

... NUCLEIC ACID FOLDING KINETICS

Erik Winfree (Caltech)



Joseph Berleant (MIT)

Ivo Hofacker

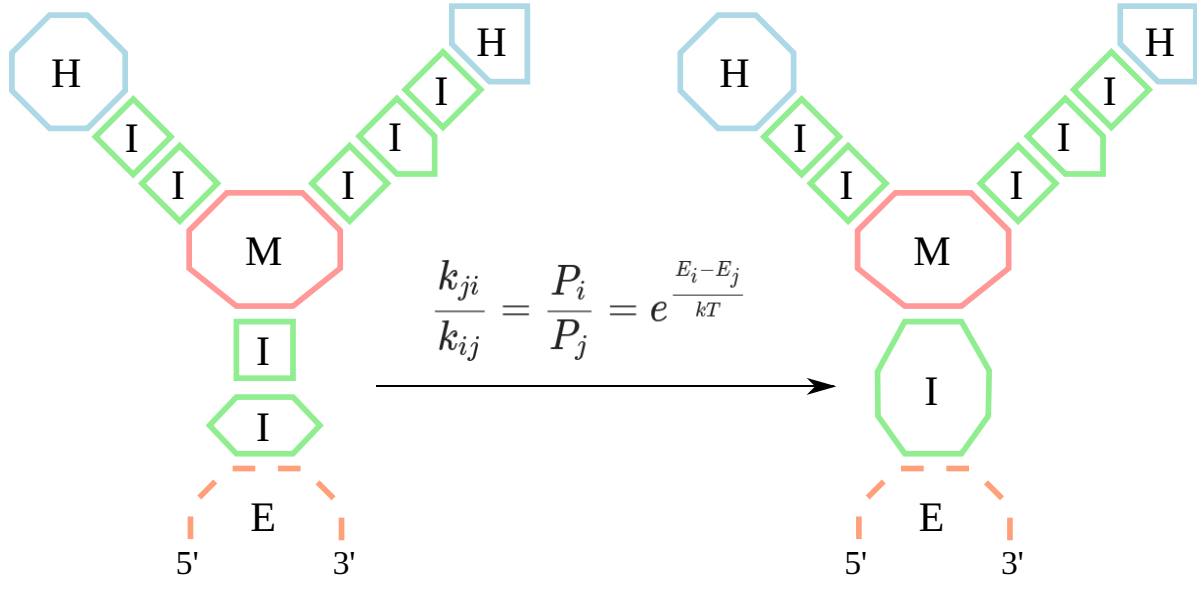
tbi

BLED 2024

Stefan Badelt
Theoretical Biochemistry Group (tbi),
University of Vienna

39th TBI Winterseminar

THE Kinfold MODEL



- Choose a rate model that satisfies detailed balance
- Simulate single stochastic folding trajectories as given by the master equation

$$\frac{dP_i(t)}{dt} = \sum_{i \neq j} (P_j(t)k_{ji} - P_i(t)k_{ij})$$

$\rightarrow$ yields the correct equilibrium distribution

$\rightarrow$ whether kinetics is correct depends on the rate model!

DETERMINISTIC MODELS

- are useful, e.g. because:
  - sometimes simulations take a long time
  - interesting transitions visible on the log scale
  - kinetic evaluation during sequence design
  - parameter inference
- existing solutions:
  - have unreliable dynamics
  - correct equilibrium distribution

EXAMPLES FOR EXISTING MODELS

- barriers/treekin
  - limited sequence length
  - non-Markovian macrostates
  - exact equilibrium distribution
- DrTransformer
  - heuristic
- ... more

WHAT IS THE RATE?

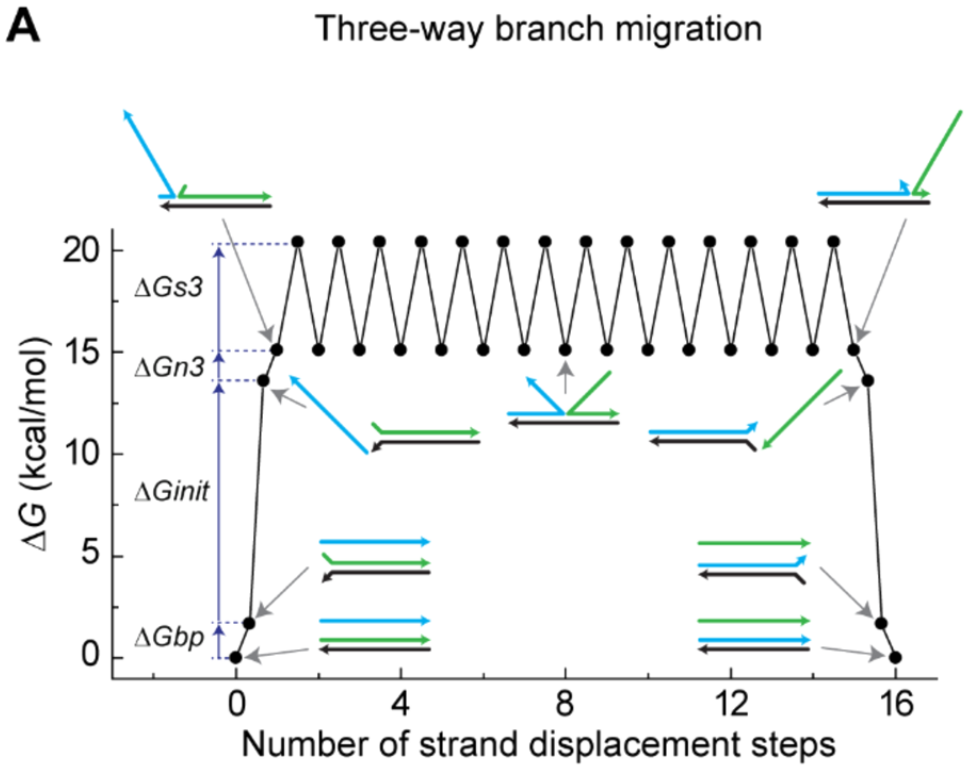


Figure from Kotani & Hughes (2017)

WHAT IS THE RATE?

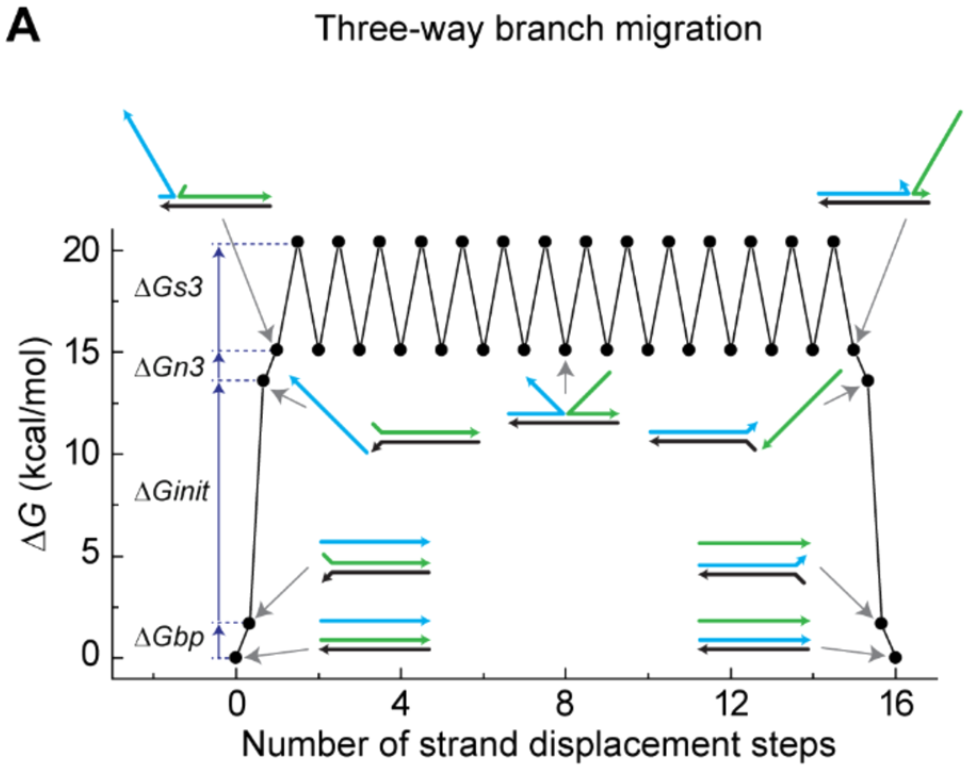


Figure from Kotani & Hughes (2017)

$$P^{success} = 1/l \quad \rightarrow k \propto 1/l$$
$$T^{expected} = l^2 \quad \rightarrow k \propto 1/l^2$$

WHAT IS THE RATE?

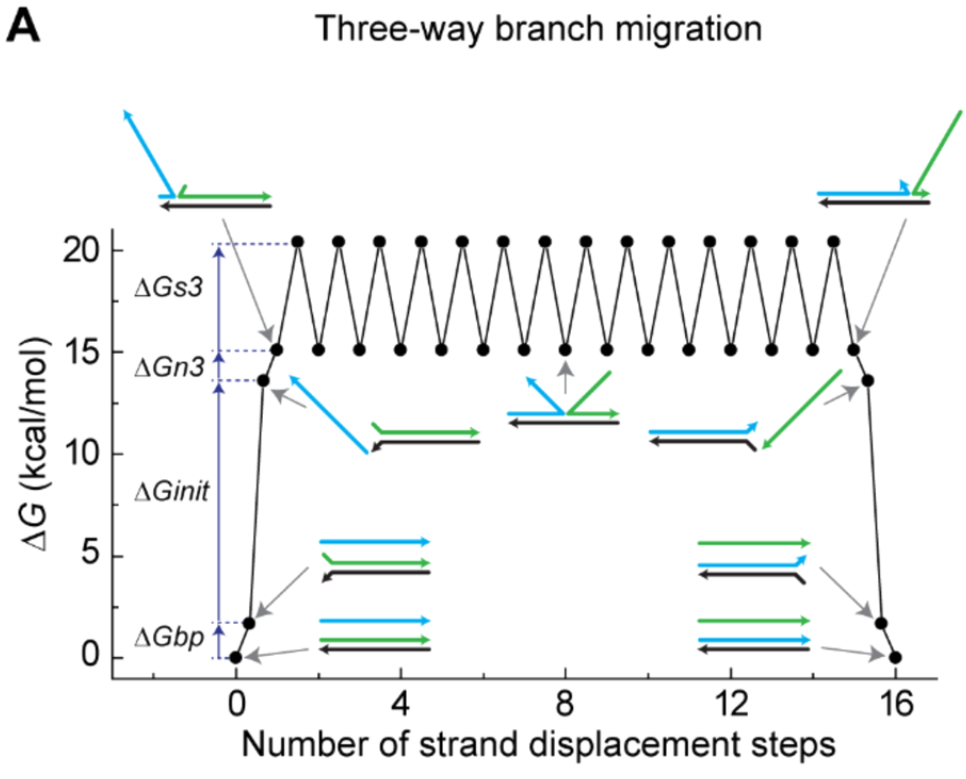
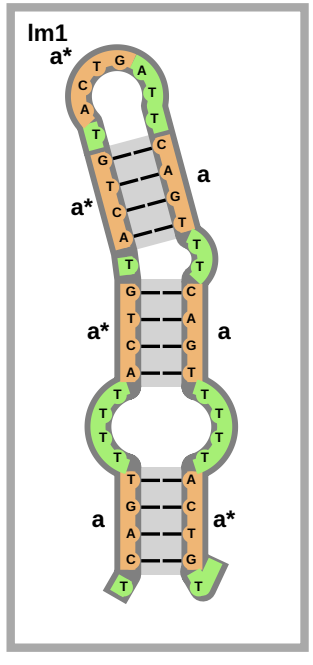
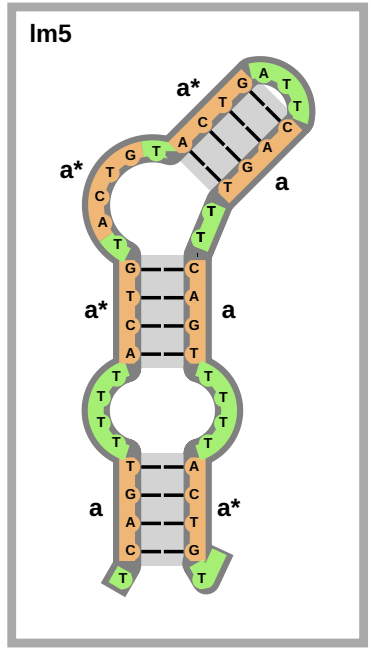


Figure from Kotani & Hughes (2017)

$$P^{success} = 1/l \quad \rightarrow k \propto 1/l$$
$$T^{expected} = l^2 \quad \rightarrow k \propto 1/l^2$$

Berleant et al. (2018): use "first step" model.

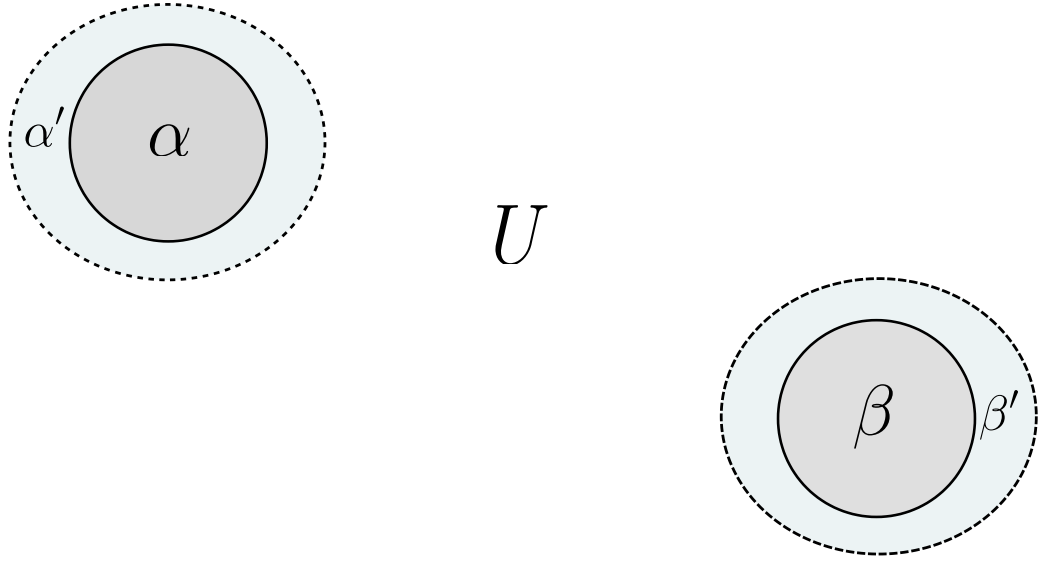

lm1
a*
a
a*
a*
a
a
a*


lm5
a*
a*
a
a*
a
a
a*

$\alpha\ \beta$ macostates

$\alpha'\ \beta'$ macostate neighborhoods

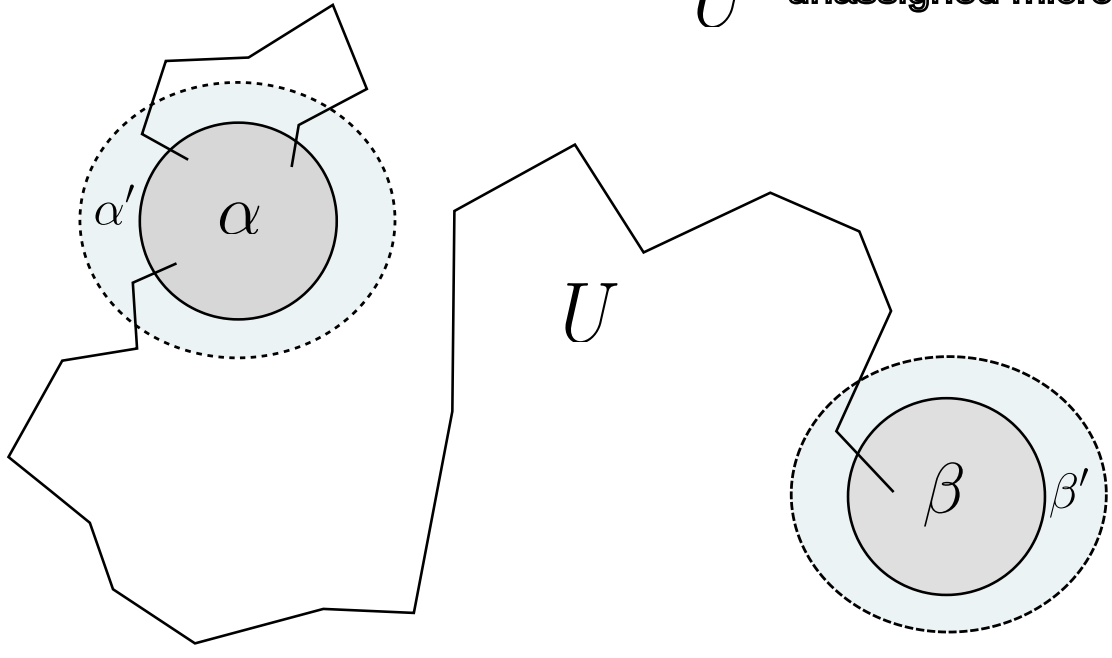
U unassigned microstates



$\alpha\ \beta$ macostates

$\alpha'\ \beta'$ macostate neighborhoods

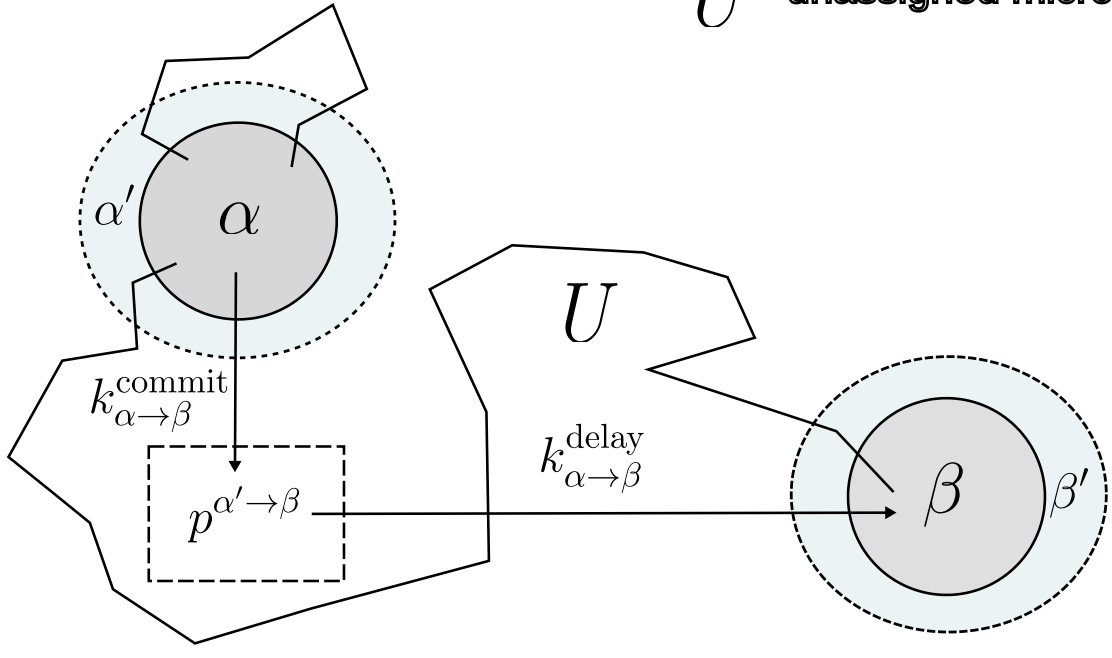
U unassigned microstates

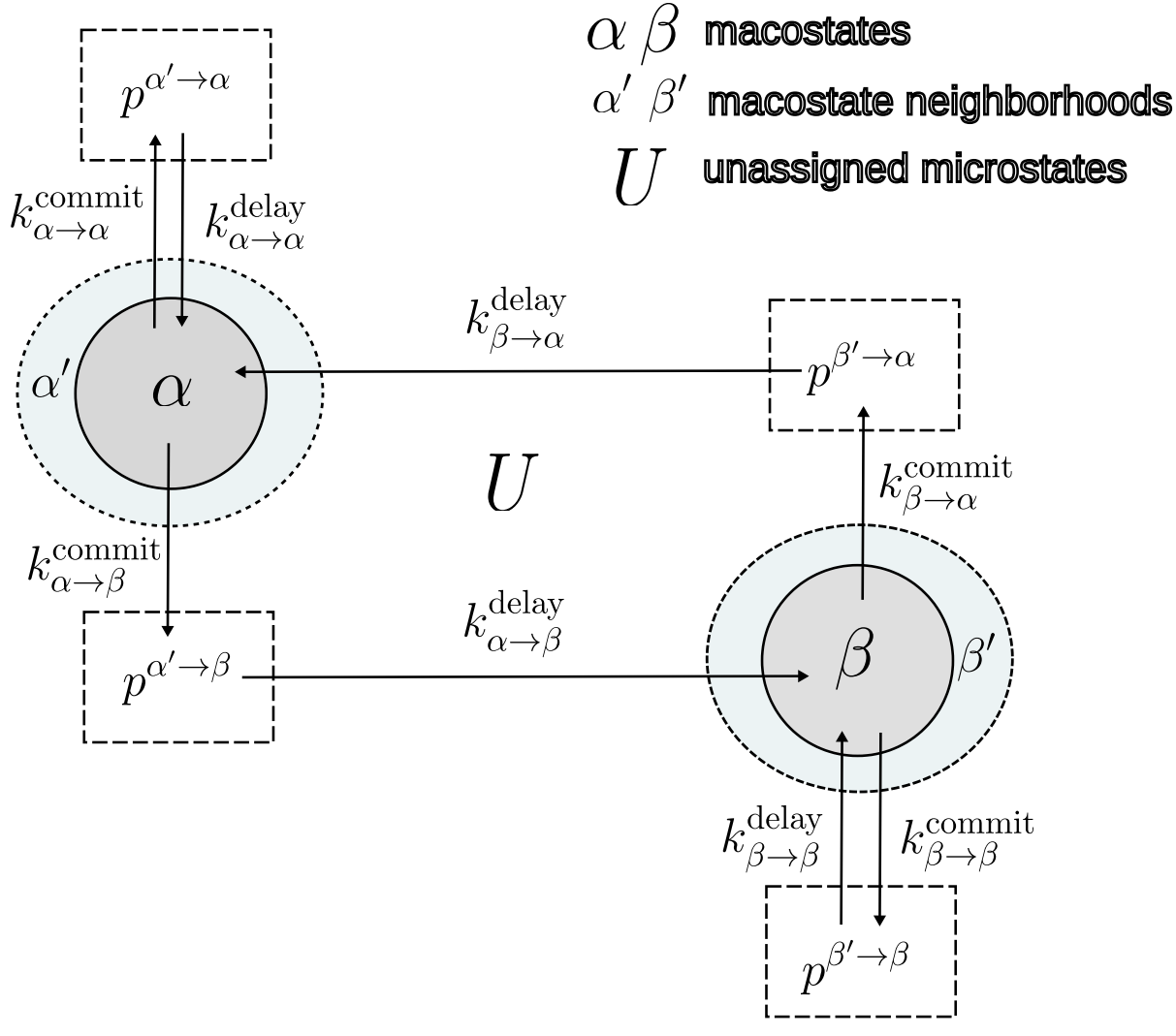


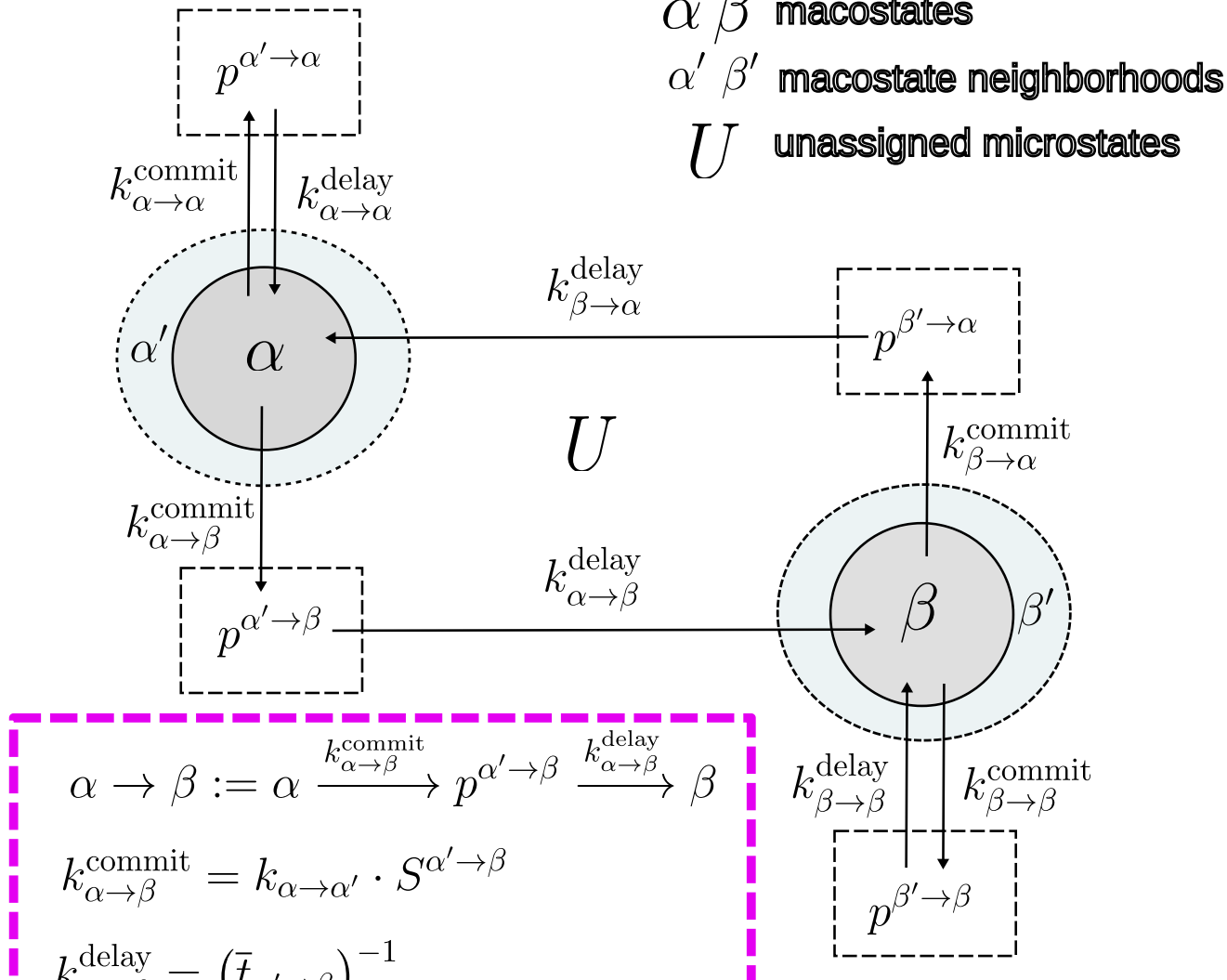
$\alpha\ \beta$ macostates

$\alpha'\ \beta'$ macostate neighborhoods

U unassigned microstates




$\alpha\ \beta$ macostates
$\alpha'\ \beta'$ macostate neighborhoods
U unassigned microstates
$p^{\alpha'\to\alpha}$
$k^{\text{commit}}_{\alpha\to\alpha}$
$k^{\text{delay}}_{\alpha\to\alpha}$
$k^{\text{delay}}_{\beta\to\alpha}$
$p^{\beta'\to\alpha}$
α'
α
U
$k^{\text{commit}}_{\beta\to\alpha}$
$k^{\text{commit}}_{\alpha\to\beta}$
$k^{\text{delay}}_{\alpha\to\beta}$
$p^{\alpha'\to\beta}$
β
β'
$k^{\text{delay}}_{\beta\to\beta}$
$k^{\text{commit}}_{\beta\to\beta}$
$p^{\beta'\to\beta}$


$\alpha\ \beta$ macostates
$\alpha'\ \beta'$ macostate neighborhoods
U unassigned microstates
$p^{\alpha'\to\alpha}$
$k^{\text{commit}}_{\alpha\to\alpha}$
$k^{\text{delay}}_{\alpha\to\alpha}$
$k^{\text{delay}}_{\beta\to\alpha}$
$p^{\beta'\to\alpha}$
α'
α
U
$k^{\text{commit}}_{\beta\to\alpha}$
$k^{\text{commit}}_{\alpha\to\beta}$
$k^{\text{delay}}_{\alpha\to\beta}$
$p^{\alpha'\to\beta}$
β
β'
$k^{\text{delay}}_{\beta\to\beta}$
$k^{\text{commit}}_{\beta\to\beta}$
$p^{\beta'\to\beta}$
$\alpha \to \beta := \alpha \xrightarrow{k^{\text{commit}}_{\alpha\to\beta}} p^{\alpha'\to\beta} \xrightarrow{k^{\text{delay}}_{\alpha\to\beta}} \beta$
$k^{\text{commit}}_{\alpha\to\beta} = k_{\alpha\to\alpha'} \cdot S^{\alpha'\to\beta}$
$k^{\text{delay}}_{\alpha\to\beta} = \left(\bar{t}_{\alpha'\to\beta}\right)^{-1}$

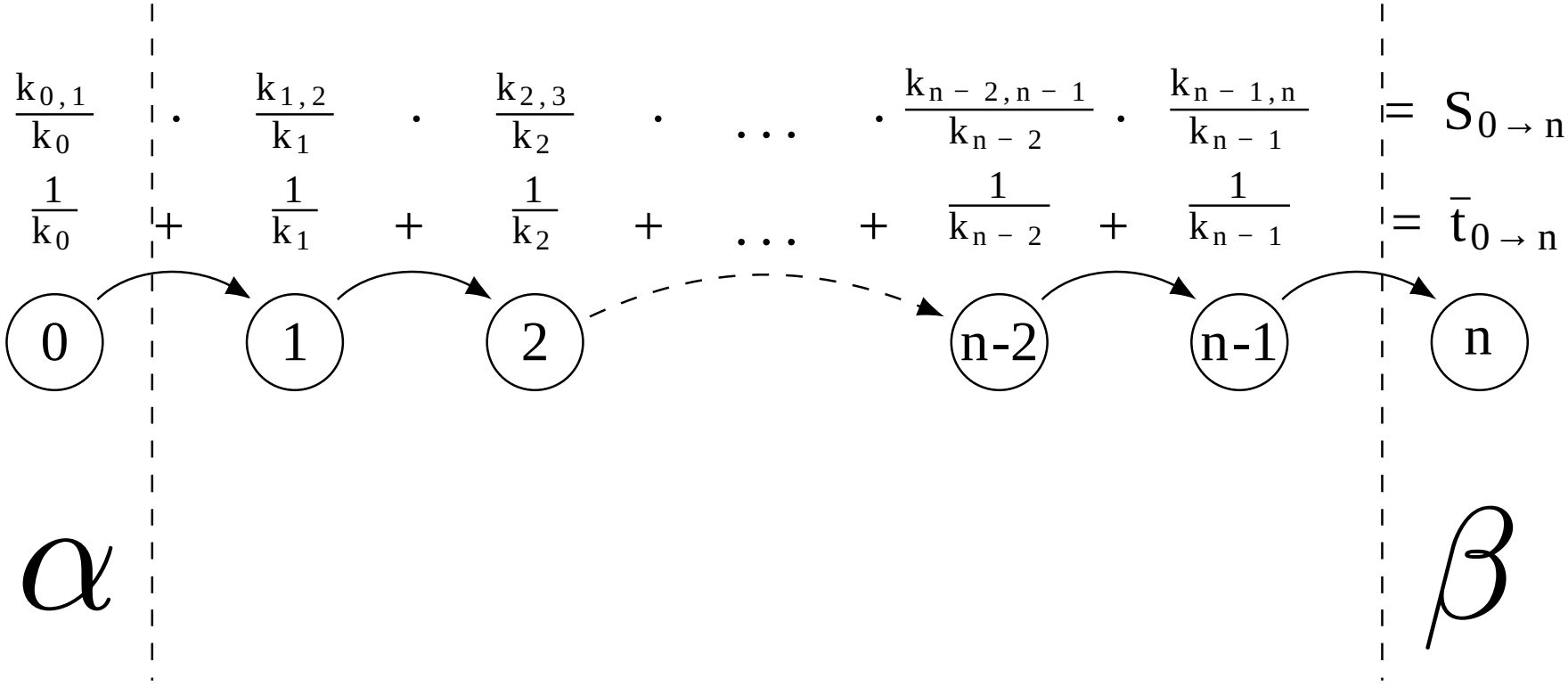
THE "ROOT" MODEL

$$\alpha \to \beta := \alpha \xrightarrow{k^{\text{commit}}_{\alpha\to\beta}} p^{\alpha'\to\beta} \xrightarrow{k^{\text{delay}}_{\alpha\to\beta}}$$

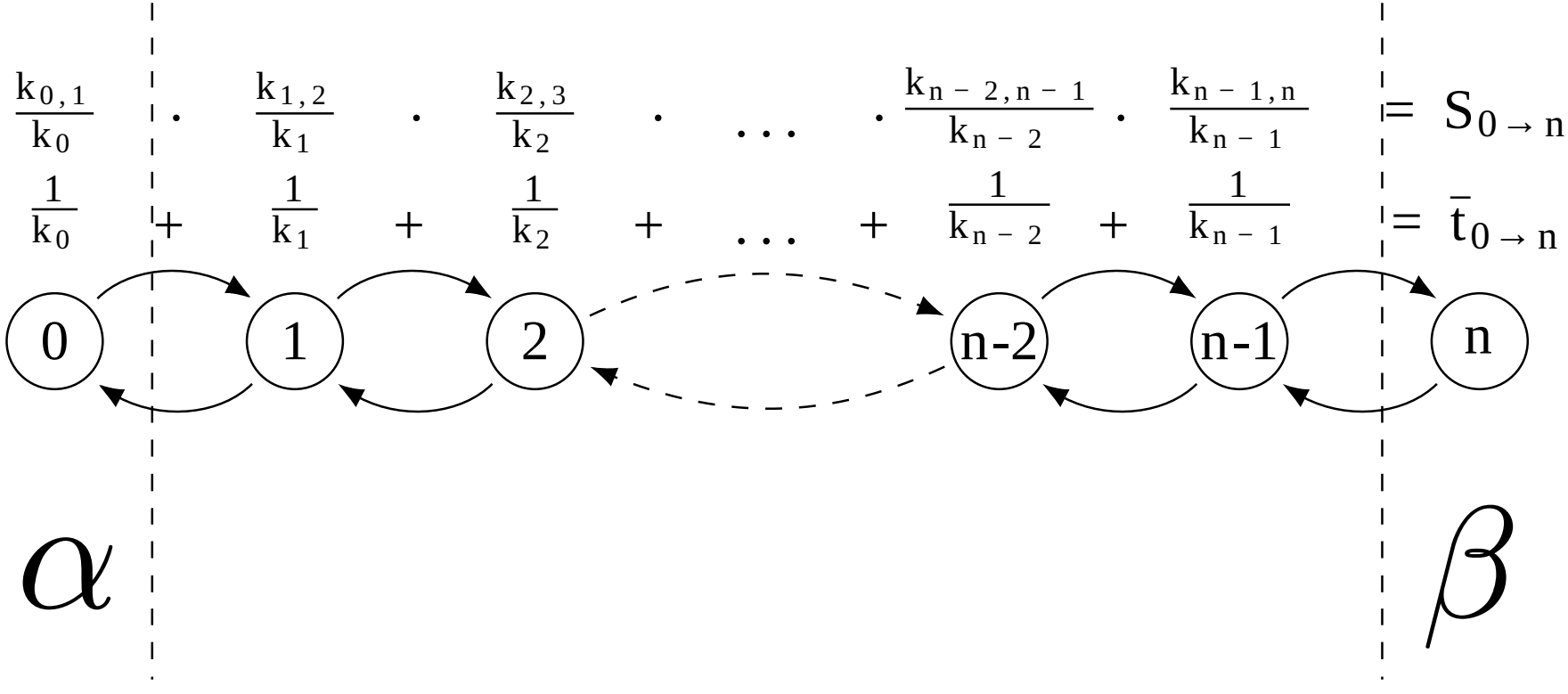
$$k^{\text{commit}}_{\alpha\to\beta} = k_\alpha \cdot \hat{S}^{\alpha'\to\beta} = \sum_{\substack{i\in\alpha \\ j\in\alpha'}} P(i|\alpha) k_{i\to j} \cdot \frac{\sum_{j\in\alpha'} |t^j}{\sum_{j\in\alpha'} |t}$$

$$k^{\text{delay}}_{\alpha\to\beta} = \frac{1}{\bar{t}_{\alpha'\to\beta}} = \left(\frac{1}{n} \sum_{l=1}^{n} t_l^{j\to\beta} \right)^{-1}$$

CORRECTNESS OF THE "ROOT" MODEL

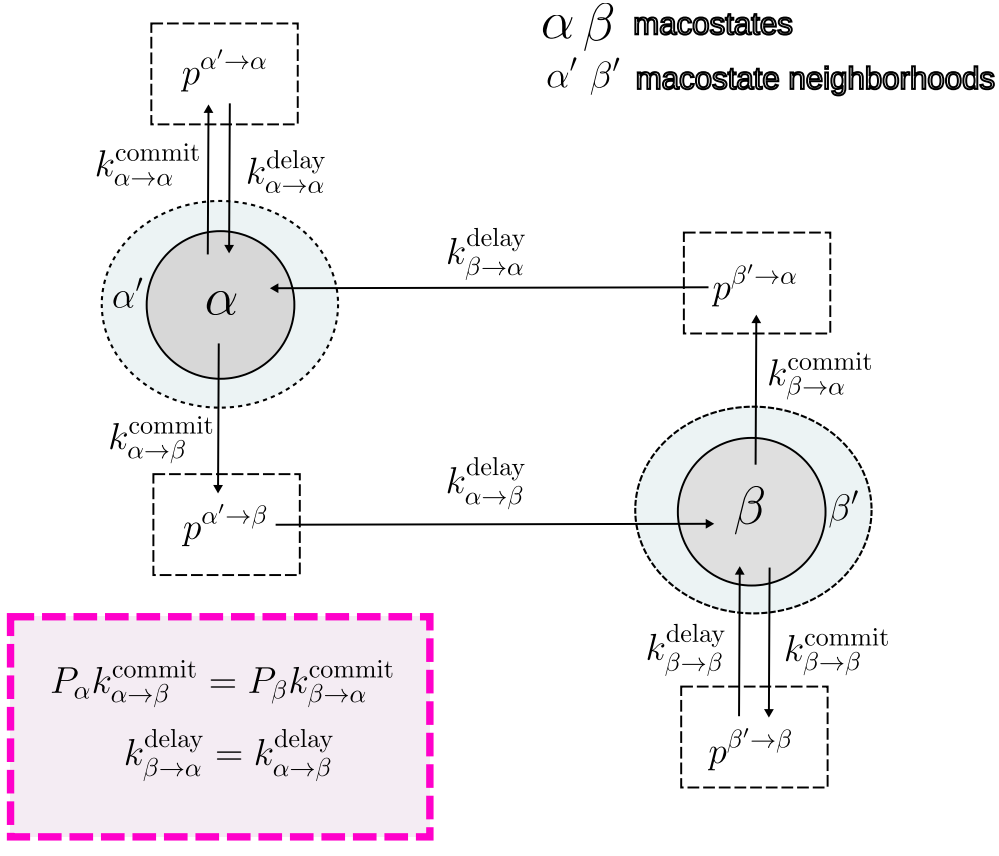


CORRECTNESS OF THE "ROOT" MODEL



So how is $\bar{t}_{\alpha' \to \beta}$ related to $\bar{t}_{\beta' \to \alpha}$?

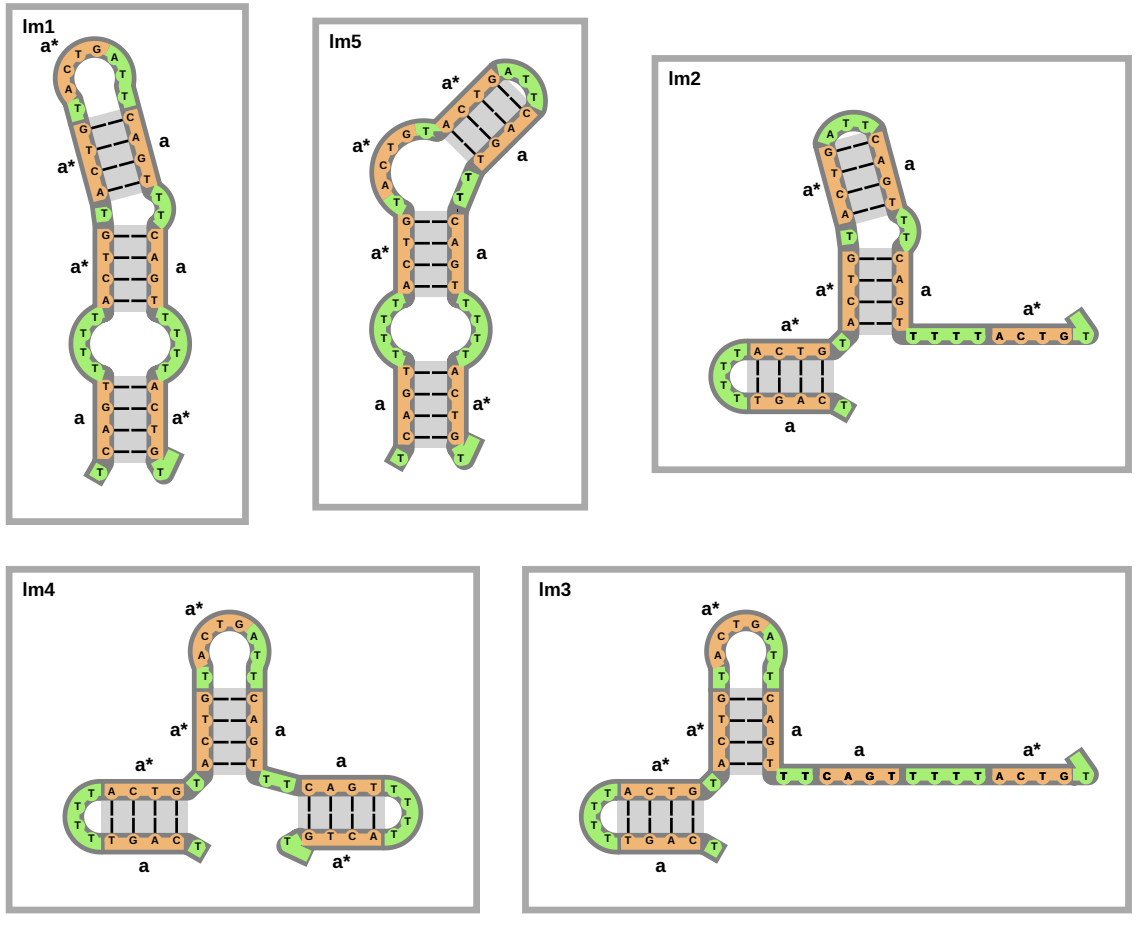
CORRECTNESS OF THE "ROOT" MODEL



LIMITATIONS

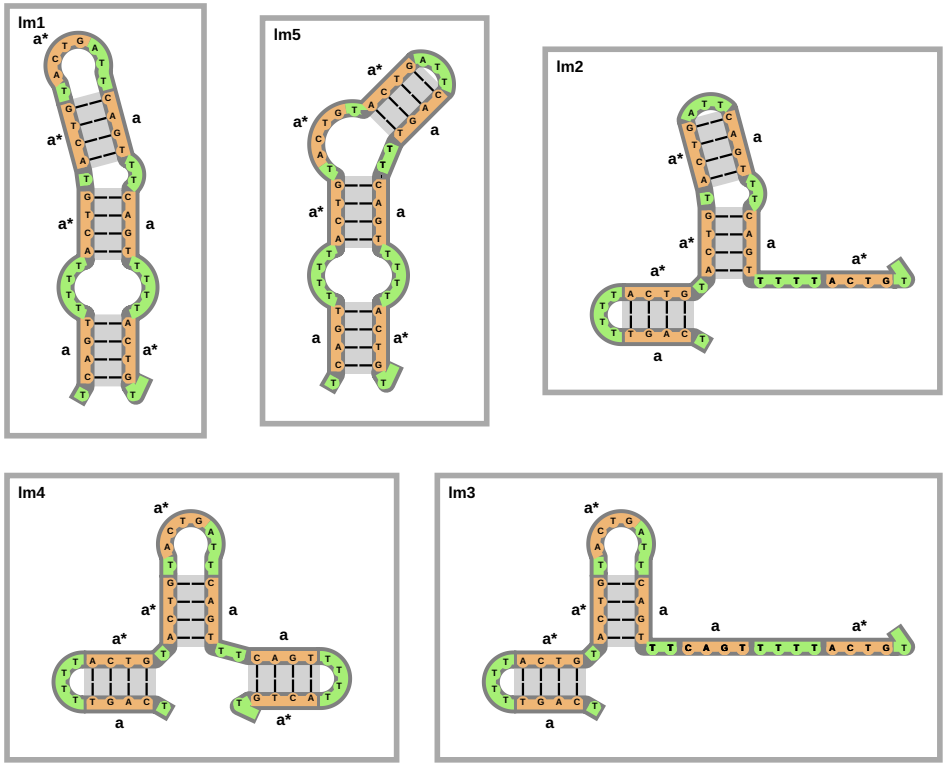
- Correct equilibrium, ... so what?
- What are the limitations of the "root" model?
- What the heck is missing for the final model?

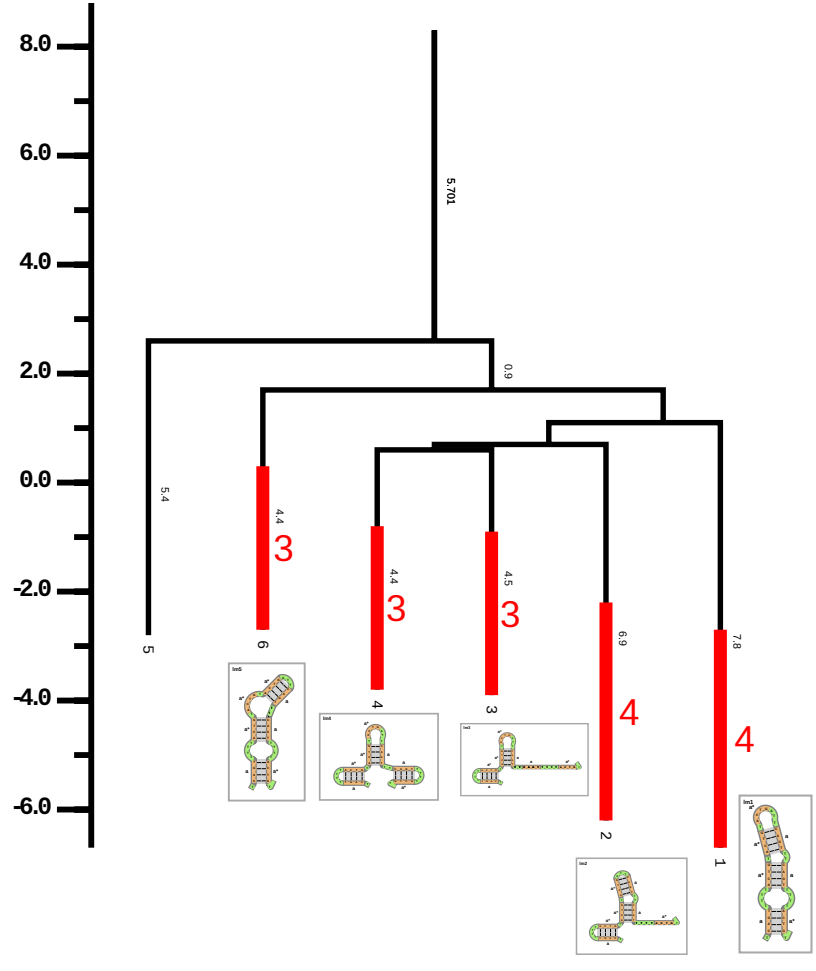
AN EXAMPLE



UCAGUCUUCGCUGCGCUGUAUCGAUUCGGUUUCAGUUUUUAUUGC

AN EXAMPLE

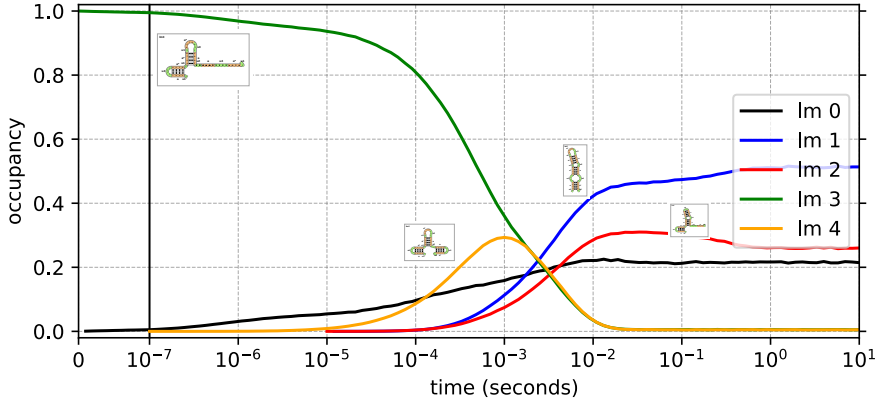




UCAGUCUUCGCUGCGCUGUAUCGAUUCGGUUUCAGUUUUUAUUGC

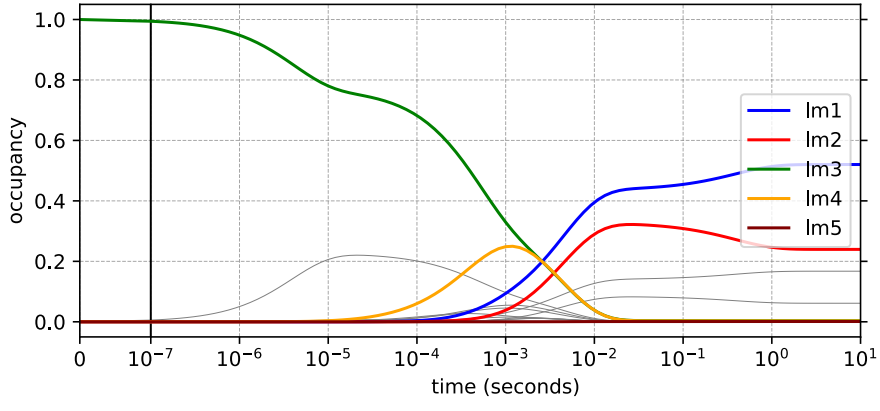
A SIMPLE COMPARISON

kinfold



10^5 simulations;
metropolis model:
$k_0 = 10^6$

commit 'n' delay

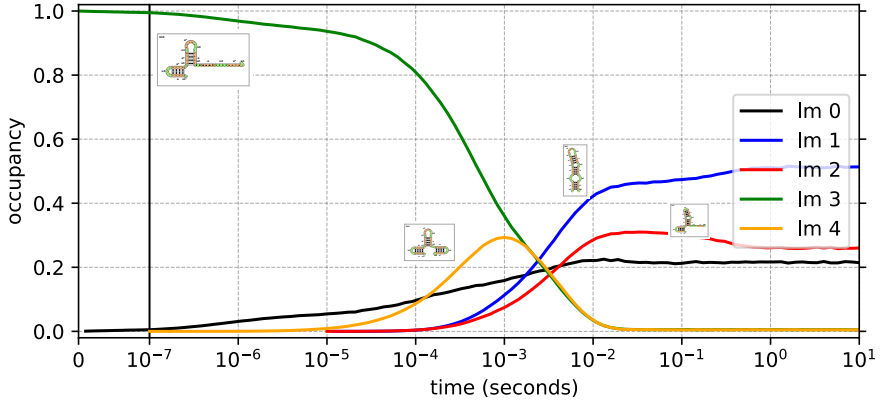


$5 \cdot 10^5$ simulations;
metropolis model:
$k_0 = 10^6$

... so how do they compare?

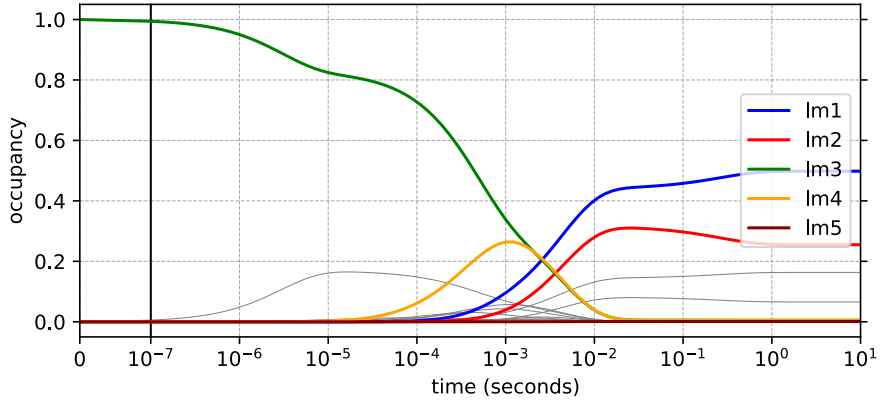
A SIMPLE COMPARISON

kinfold



10^5 simulations;
metropolis model:
$k_0 = 10^6$

commit 'n' delay

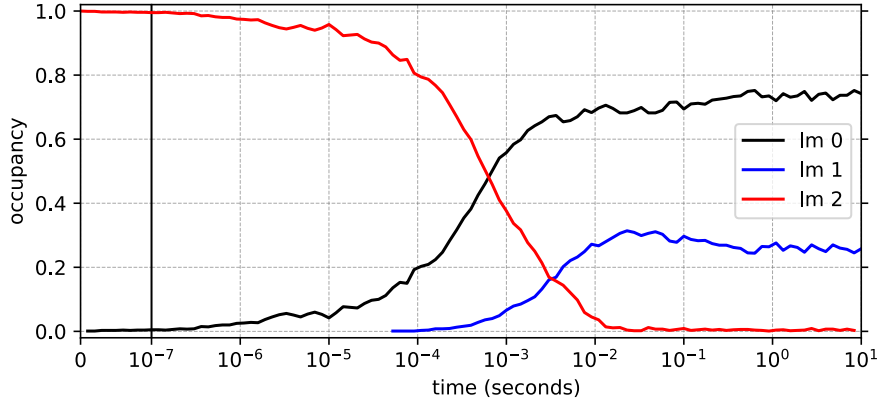


10^7 simulations;
metropolis model:
$k_0 = 10^6$

... so how do they compare?

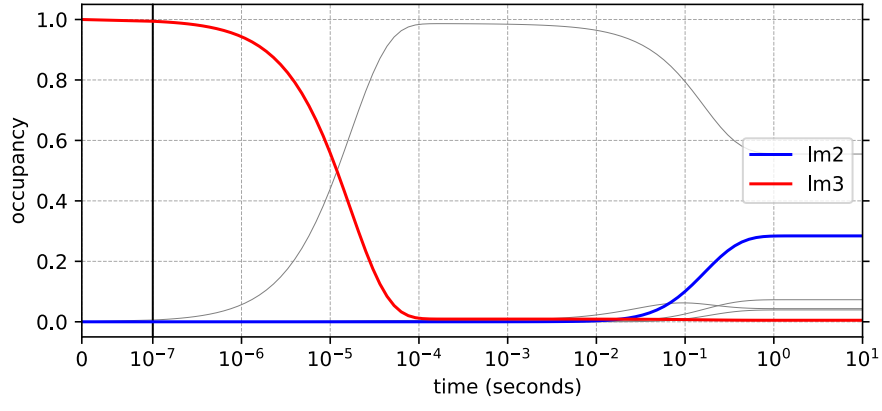
A SIMULATION WITH ONLY TWO SPECIES

kinfold



10^4 simulations;
metropolis model:
$k_0 = 10^6$

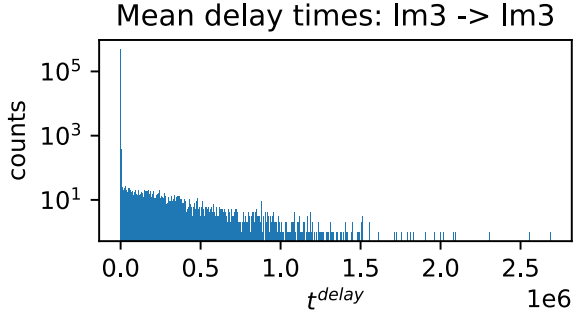
commit 'n' delay

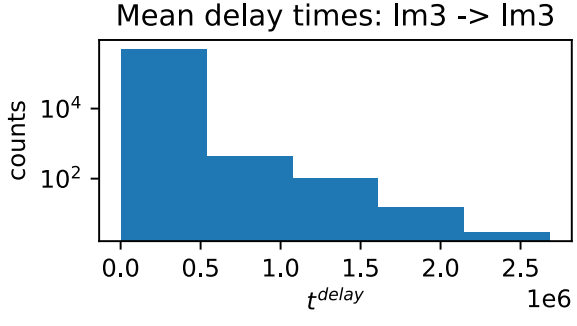


10^5 simulations;
metropolis model:
$k_0 = 10^6$

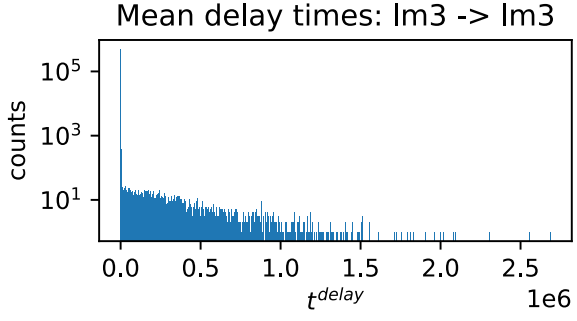
... a pair of reactions is not enough.

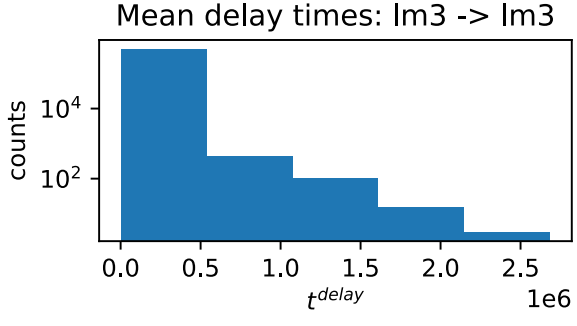
SPLITTING PATHWAYS

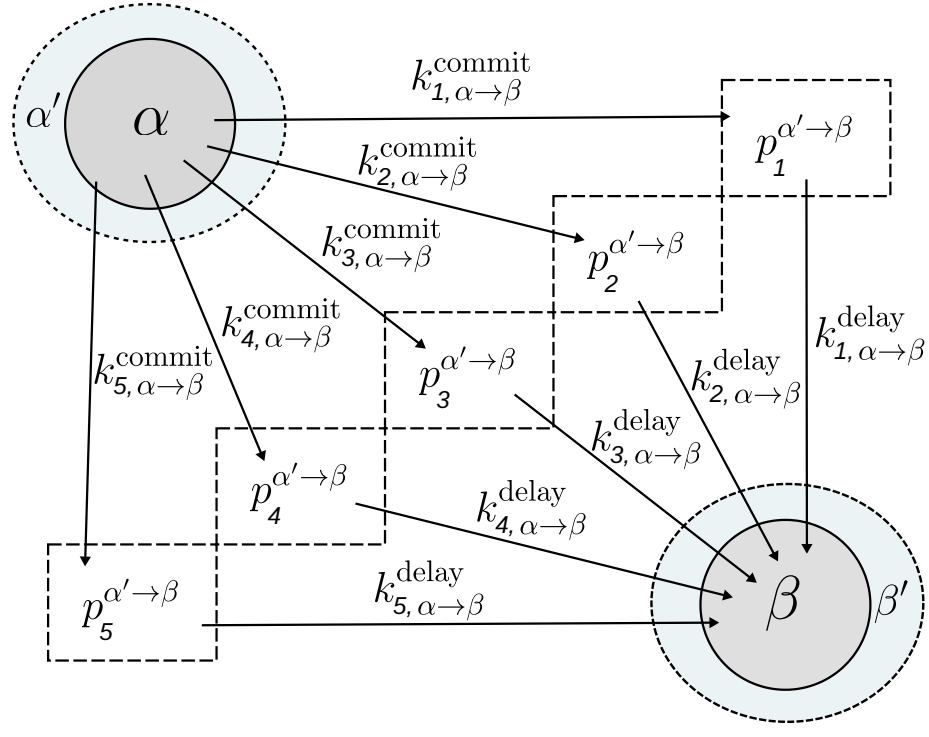




SPLITTING PATHWAYS

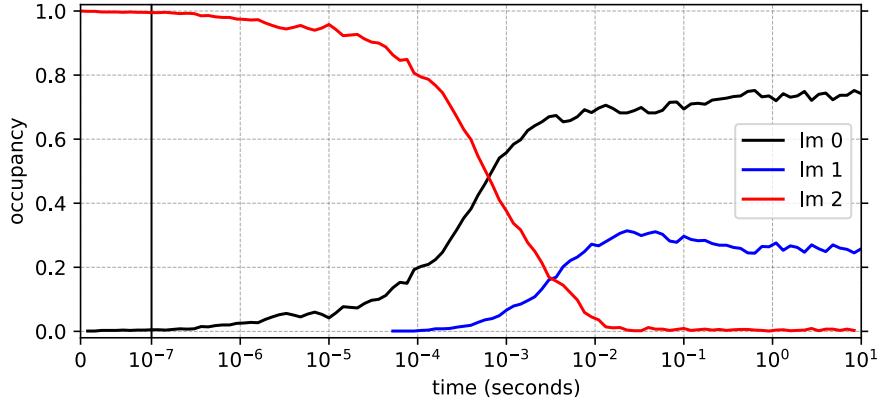






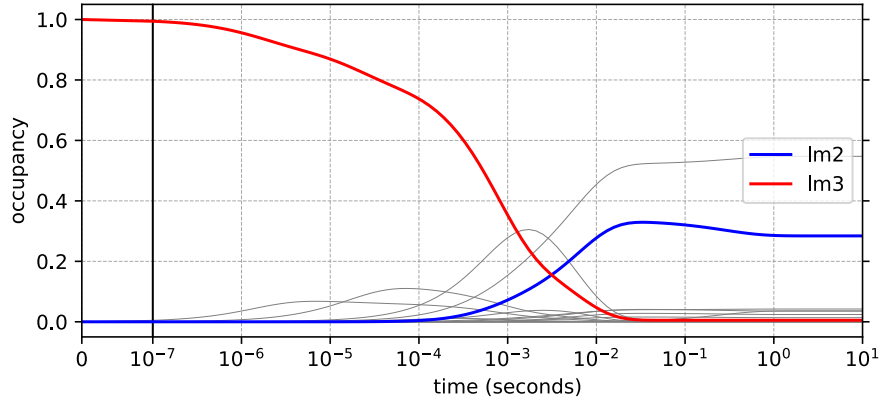
A SIMULATION WITH ONLY TWO SPECIES

kinfold



10^4 simulations;
metropolis model:
$k_0 = 10^6$

commit 'n' delay

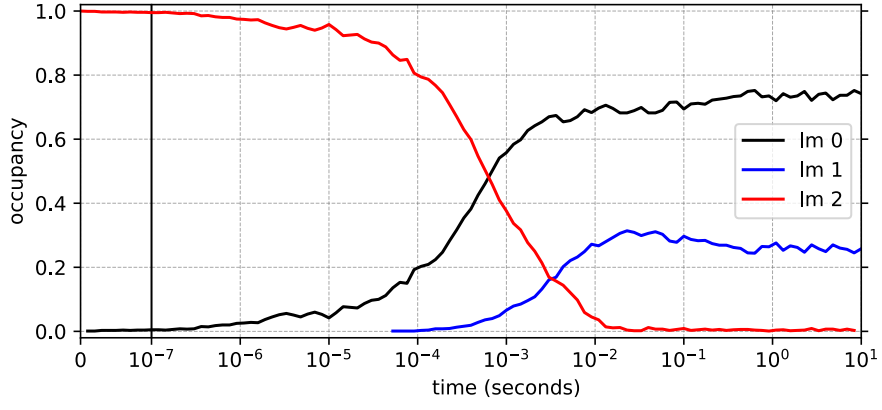


10^5 simulations;
metropolis model:
$k_0 = 10^6$; 5 paths!

...

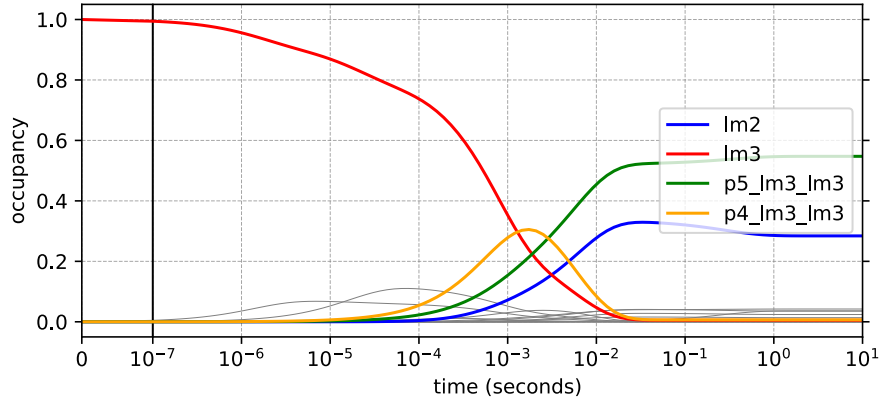
A SIMULATION WITH ONLY TWO SPECIES

kinfold



10^4 simulations;
metropolis model:
$k_0 = 10^6$

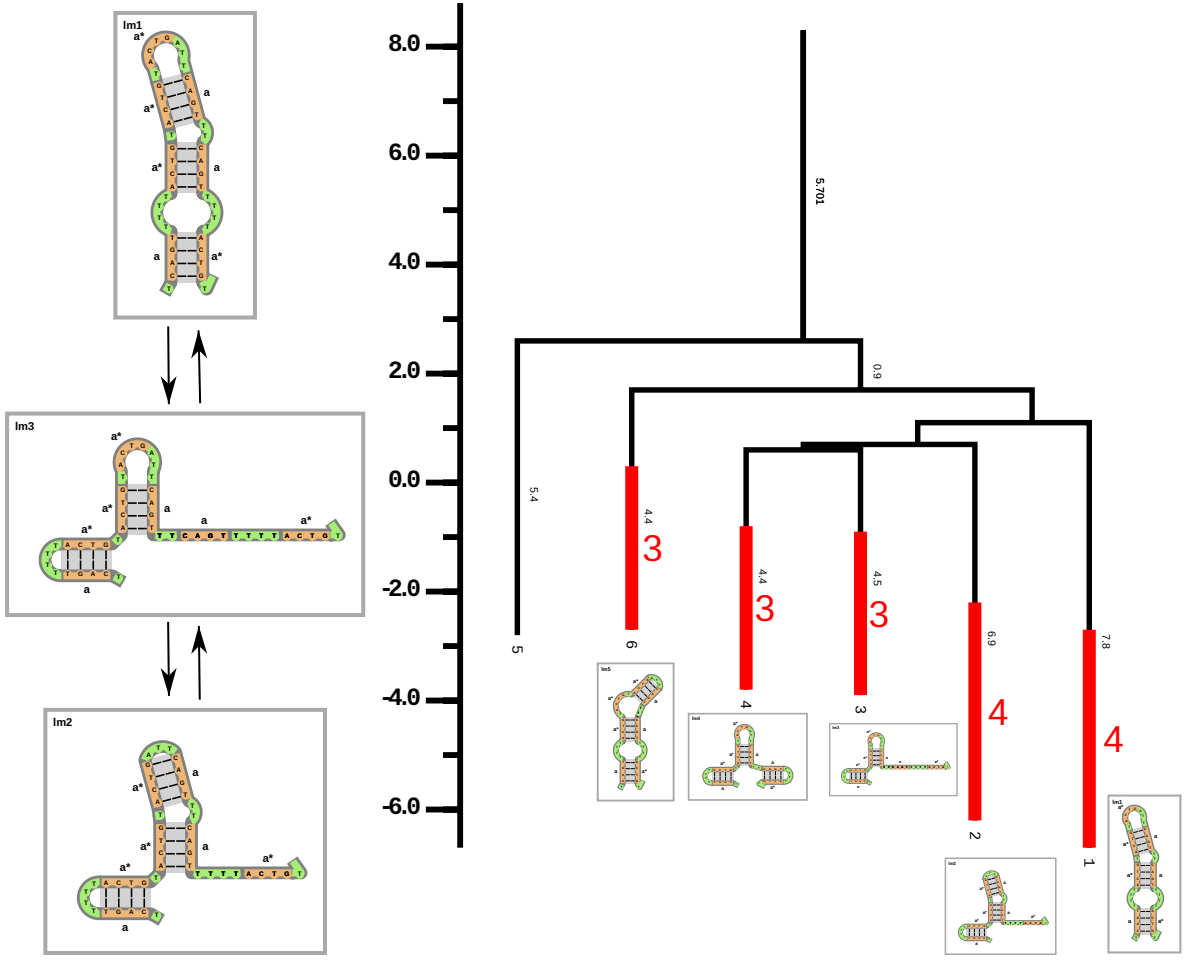
commit 'n' delay



10^5 simulations;
metropolis model:
$k_0 = 10^6$; 5 paths!

...

WHAT ELSE CAN WE LEARN?

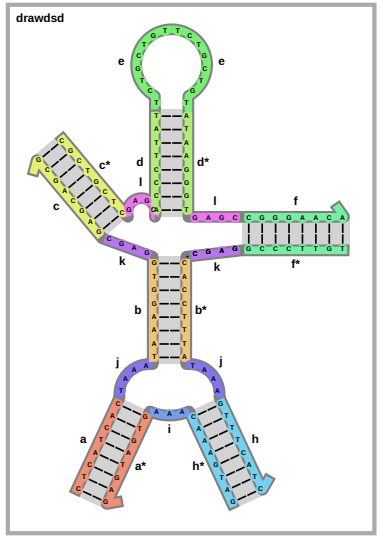


IN PROGRESS / IN PREPARATION

- Estimate the error of simulations?
- Enforce correct equilibrium distribution.
- Learn rate parameters for domain-level folding.

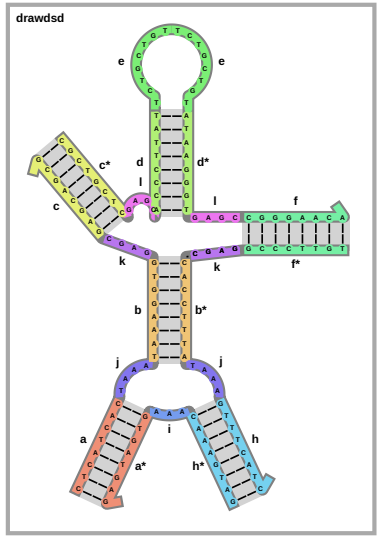
IN PROGRESS / IN PREPARATION

- Estimate the error of simulations?
- Enforce correct equilibrium distribution.
- Learn rate parameters for domain-level folding.



IN PROGRESS / IN PREPARATION

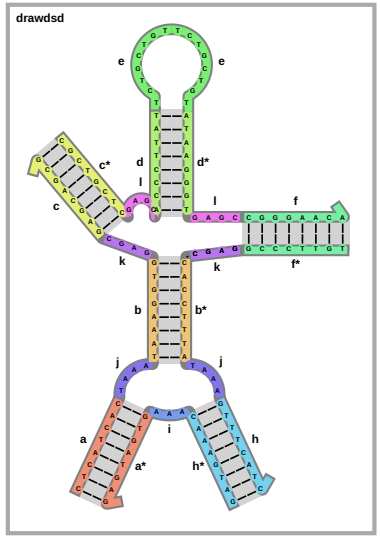
- Estimate the error of simulations?
- Enforce correct equilibrium distribution.
- Learn rate parameters for domain-level folding.



https://github.com/bad-ants-fleet/drawdsd

IN PROGRESS / IN PREPARATION

- Estimate the error of simulations?
- Enforce correct equilibrium distribution.
- Learn rate parameters for domain-level folding.



THANK YOU!

https://github.com/bad-ants-fleet/drawdsd