The ViennaRNA Package 2.0.2

The ViennaRNA Package 2.0.2

- Finally published[1]
- Many new features already implemented
- Several new models/algorithms in preparation

[1] Algorithms for Molecular Biology 2011, 6:26 Published: 24 November 2011

The ViennaRNA Package 2.0.2

- Finally published[1]
- Many new features already implemented
- Several new models/algorithms in preparation

Thank You for Your attention!

In silico RNA folding kinetics using 2DKin

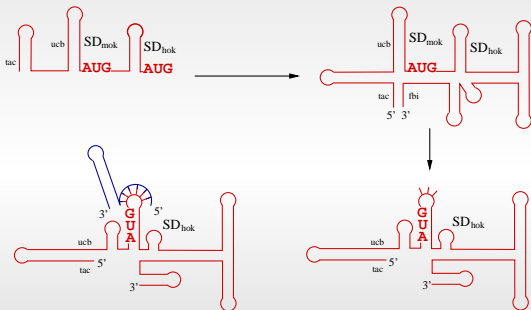Ronny Lorenz
ronny@tbi.univie.ac.at

Institute for Theoretical Chemistry
University of Vienna

Bled, Slovenia, February 17, 2012

Gene regulation by RNA structure changes

Self-induced switches:

- co-transcriptional kinetic traps (e.g. sv11, ms2)
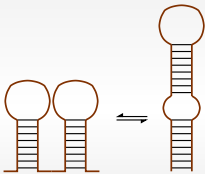- degradation associated rearrangement (e.g. hok/sok)

trans-induced switches:

- transcriptional control (e.g. attenuators)
- metabolite sensors (e.g. add, TPP, SAM)
- temperature sensors (e.g. bacterial virulence genes)
- anti-sense driven rearrangements

RNA switch design (self-induced)

- in silico[2]
- mutational studies
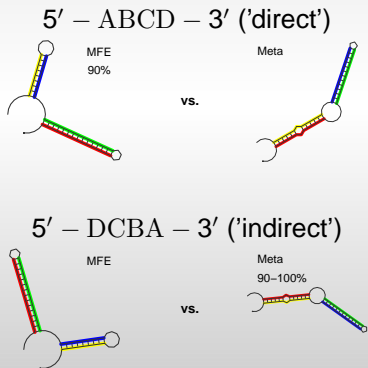- molecular biologists expert knowledge
- ...

Sequential folding effects

- hairpins form first (short range interactions)
- more complex structures later

---

[2]Flamm et al. 2001

RNA switch design[3]

- folding path can be encoded within sequence
- utilization of helix length asymetry
- two switches with almost identical energy landscape
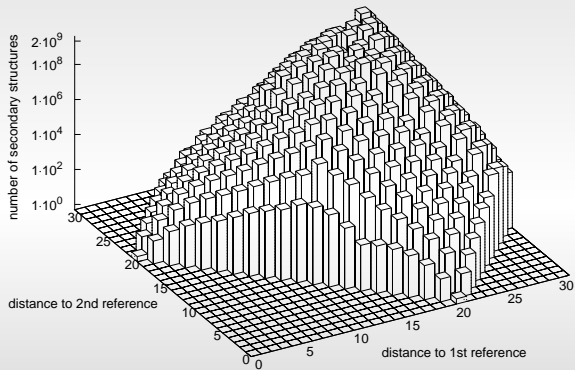- in vitro experiments to see co-transcriptional folding effect

$5' - \mathrm{ABCD} - 3'$ ('direct')

MFE 90%   vs.   Meta

$5' - \mathrm{DCBA} - 3'$ ('indirect')

MFE   vs.   Meta 90–100%

[3]Xayaphoummine et al. 2007

`RNA2Dfold`[4] **in a nutshell**

- nice way to project high-dimensional energy landscape into 2D
- classified dynamic programming approach
- select two reference structures s_1, s_2 (of interest)
- (base pair) distance classes $\rightarrow \kappa, \lambda-\text{neighborhoods}$
- compute the MFE, partition function for each $\kappa, \lambda-\text{neighborhood}$
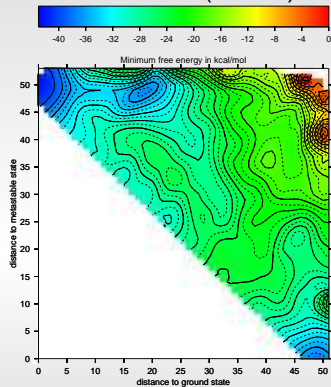
E.g.:

$$\text{MFE}_{\kappa,\lambda} = \min_{\substack{s \in \mathcal{S} \\ \delta(s, s_1) = \kappa \\ \delta(s, s_2) = \lambda}} E(s)$$

---

[4]Lorenz et al. 2009

RNA2Dfold - distance class population

RNA2Dfold - energy landscape analysis

$5' - \mathrm{ABCD} - 3'$ ('direct')

$5' - \mathrm{DCBA} - 3'$ ('indirect')
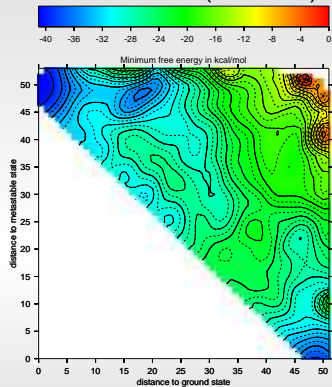
RNA folding as a Markov process

The fitness landscape

- states $\Rightarrow$ secondary structures
- move set $\Rightarrow$ insert/delete base pair
- fitness $\Rightarrow$ free energy of the state

The Markov process

- fitness landscape
- transition rates between states (Metropolis/Kawasaki rule)

The master equation

$$\frac{d}{dt}\vec{p}(t) = \mathbf{R}\vec{p}(t) \quad \text{with formal solution} \quad \vec{p}(t) = e^{t \cdot \mathbf{R}} \cdot \vec{p}(0).$$

The pipeline

- $(\texttt{barriers} + \texttt{RNAsubopt})^5 \mid \texttt{2DKin} \Leftarrow$ fitness landscape $+ \mathbf{R}$
- $\texttt{treekin}^6 \Leftarrow$ solution of master equation

---

[5]Wuchty et al. '99, Flamm et al. 2002
[6]Wolfinger et al. 2004

Folding kinetics on $\kappa, \lambda-$neighborhoods

The rate matrix $\mathbf{R} = (r_{xy})$

- approx. macro rates by Boltzmann sampling S_α from each distance class:

$$r_{\beta\alpha} \approx \frac{1}{|S_\alpha| + |S_\beta|} \left\{ \sum_{x \in S_\alpha} \sum_{y \in \beta \cap \mathcal{N}(x)} k_{yx} + \sum_{y \in S_\beta} \sum_{x \in \alpha \cap \mathcal{N}(y)} \frac{\pi_\beta}{\pi_\alpha} k_{xy} \right\}$$

with:

$$k_{yx} = \begin{cases} e^{-\frac{E(y) - E(x)}{kT}} & \text{if } E(x) < E(y) \\ 1 & \text{otherwise.} \end{cases}$$
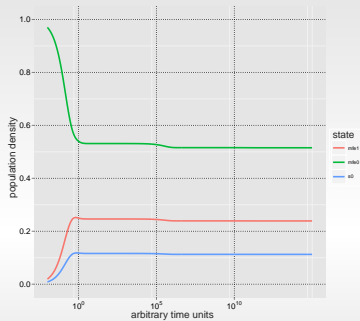
- detailed balance preserved
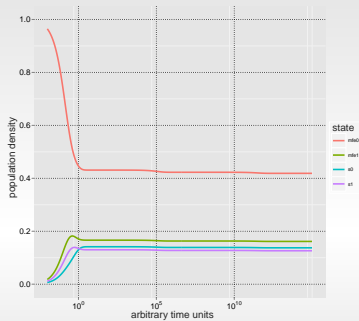
$$r_{\beta\alpha}\pi_\alpha = r_{\alpha\beta}\pi_\beta$$

- sample size of 1000 per macro state proved sufficient for the examples tested

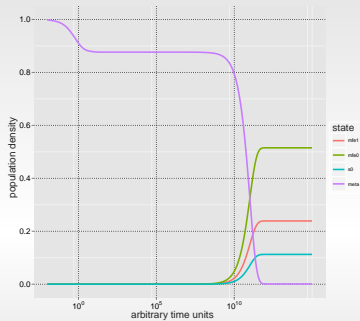# `2DKin` **kinetics - Init: 100% ground state**

$5' - \mathrm{ABCD} - 3'$ ('direct')

$5' - \mathrm{DCBA} - 3'$ ('indirect')

$5' - \text{ABCD} - 3'$ ('direct')

$5' - \text{DCBA} - 3'$ ('indirect')

2DKin kinetics - Init: 100% open chain

$5' - \mathrm{ABCD} - 3'$ ('direct')

$5' - \mathrm{DCBA} - 3'$ ('indirect')
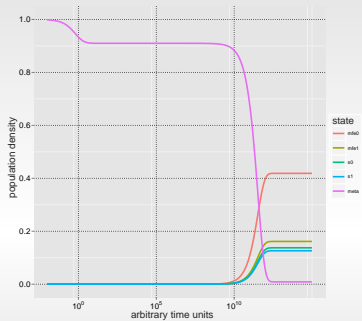
Idea:

- distance classes relative to possible base pairs of references

  ```
  GGAAGCAGAGAGAGACGAAGCAUCUCUCUCUACGAGGUAGAGGGAGAUGGAAAACCGUCUCCCUCUGCCAAGG
  ......(((((((((.......)))))))))....(((((((((((((((....))))))))))))))....
  ....(((((((.(((((((...((((((((((((((...))))))))))))))......))))))).))))))).....
  ```

  ```
  GGAAGCAGAGAGAGACGAAGCAUCUCUCUCUACGAGGUAGAG
  ......(((((((((.......)))))))))...........
  .........................(((((((...)))))))
  ```

- create 2D map for each transcription state
- make a colorful movie

### 2DKin **on varying fitness landscapes**

`barmap`[7] like approach:

- compute rates for each landscape
- select initial population density
- map the states from one landscape onto another:
    - variable temperature/model parameters $\Rightarrow$ mapping=identity
    - variable RNA chain length:
    $\mathrm{seq}[1:i] \to \mathrm{seq}[1:i+1]$:
    $c_{\kappa,\lambda} \to c_{\kappa+\alpha,\lambda+\beta}$ with
    $\alpha = (\mathrm{ref1}[1:i+1] - \mathrm{ref1}[1:i])$ and
    $\beta = (\mathrm{ref2}[1:i+1] - \mathrm{ref2}[1:i])$
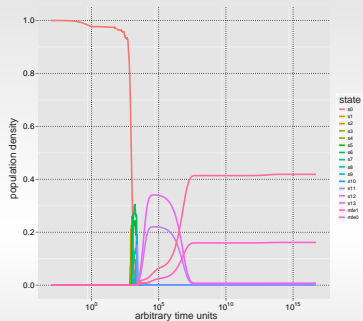- consecutive simulations initialized with population density of previous one

---

[7]Flamm et al. 2009

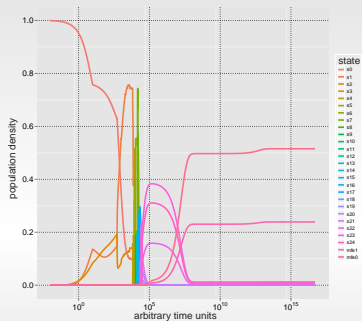# 2DKin **chain growth -** $v_{\text{trans}} = 1\text{nt/atu}$

$5' - \text{ABCD} - 3'$ ('direct')
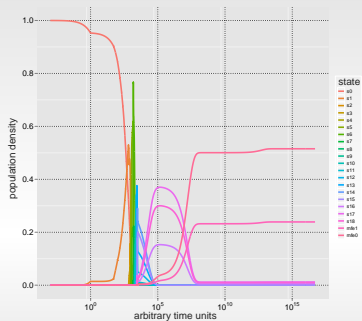
$5' - \text{DCBA} - 3'$ ('indirect')

2DKin **chain growth -** $v_{\text{trans}} = 0.1\,\text{nt/atu}$

$5' - \text{ABCD} - 3'$ ('direct')
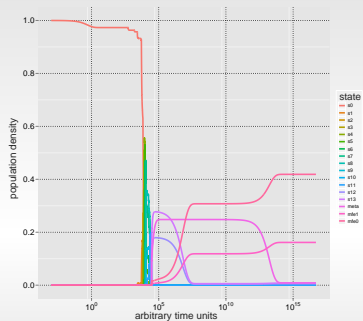
$5' - \text{DCBA} - 3'$ ('indirect')

2DKin **chain growth -** $v_{\mathrm{trans}} = 0.01\,\mathrm{nt/atu}$

$5' - \mathrm{ABCD} - 3'$ ('direct')
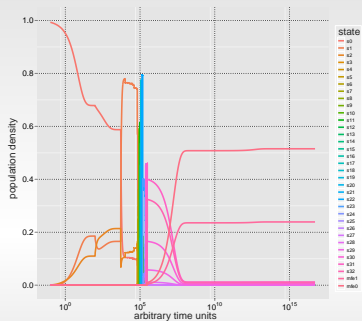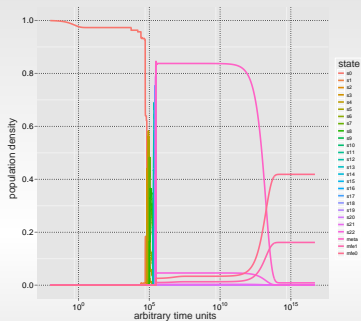
$5' - \mathrm{DCBA} - 3'$ ('indirect')

2DKin **chain growth -** $v_{\mathrm{trans}} = 0.001 \mathrm{nt/atu}$

$5' - \mathrm{ABCD} - 3'$ ('direct')
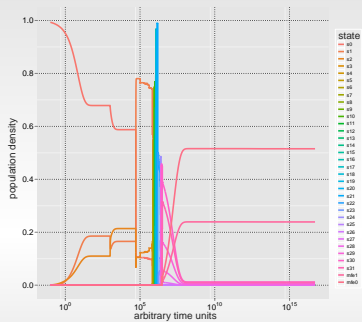
$5' - \mathrm{DCBA} - 3'$ ('indirect')
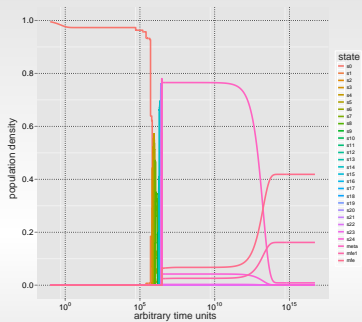
Conclusions, Problems and Outlook

- 2D projection macro states can be used for kinetic simulation
- number of states can be reduced by limiting κ, α
- works well with 'short' examples tested
- for longer sequences, thus higher state numbers, some modifications of treekin and mapper pipeline necessary
- mapping transcription times to arbitrary time units?
- 'very long' example in progress but chain reduction introduces non-trivial mapping

Thanks to

- Michael T. Wolfinger
- Chrisoph Flamm
- Marcel K
- Ivo L Hofacker

Thank You for your attention!