Kinwalker

Michael Geis

Bioinformatics Group, Dept. of Computer Science University of Leipzig

21th TBI Winterseminar in Bled



- Aims
- The Master Equation

2 Existing Models

- Kinfold–Stochastic Modelling
- TreeKin
- Helix Models

3 Kinwalker

- Few states, microscopic transitions
- Algorithm
- The Morgan-Higgs Heuristic

< ロ > < 同 > < 回 > < 回 > .

Aims The Master Equation

Questions Folding Kinetics attempts to answer

- What structure is a RNA random coil expected to fold into?
- What states does it pass through in the process?

< 日 > < 同 > < 回 > < 回 > < □ > <



Given the probabilities of conformations and the in- and outflow probabilities, the change in probability can be determined. The **master equation** summarizes this.

Let $P_{\alpha}(t)$ denote the probability of state α at time t, k_{+} the rate of inflow and k_{-} the rate of outflow. Then:

$$\begin{split} \frac{\delta P_{\alpha}}{\delta t}(t) &= \sum_{\beta \neq \alpha} [k_{+} P_{\beta}(t) - k_{-} P_{\alpha}(t)] \\ k_{+/-} &= k_{0} e^{\frac{-\Delta G}{KT}} \end{split}$$

< ロ > < 同 > < 回 > < 回 > .

 Outline

 Introduction
 Aims

 Existing Models
 The Master Equation

 Kinwalker
 The Master Equation

- For t → ∞, transition rates become constant, yielding the eventual probability of each state.
- Only a few of all possible conformations are of interest. Computing all of them becomes soon too expensive. Kinetic models essentially differ in how they group states into macro-states and their modelling of Δ*G* for the transition between these.

< 日 > < 同 > < 回 > < 回 > < □ > <

Aims The Master Equation

The underlying assumption

For $t \to \infty$, the result of the folding process and the state predicted by the thermodynamic ensemble agree.

- 1. Therefore, models based on thermodynamic considerations can be used to predict the final structure.
- 2. This is only true, if the model is *ergodic*: It must be possible to reach every state from every other state.

< □ > < 同 > < 回 > <

Kinfold–Stochastic Modelling TreeKin Helix Models

(日)

Kinetic Folding Algorithm

Sample folding paths using a Monte-Carlo based method to obtain transition rates.

Assign rates to each move, e.g.

$$P_i = \min\left\{1, \exp\left(-\frac{\Delta E}{kT}\right)\right\}$$

Select a move with probability proportional to its rate.

Kinfold–Stochastic Modelling TreeKin Helix Models

・ロット (雪) (日) (日) 日

The Flooding Algorithm

Refer to a set of points in the folding space for which a steepest descent leads to the same minimum as basin. The Flooding Algorithm then divides a landscape into basins.

Start with an energy sorted list of conformations. For each structure *x*, there are three cases:

- 1. If no neighbors of *x* are known, *x* is a local minimum and starts a new basin.
- 2. *x* belongs to a basin B, if all its neighbors belong to it.
- 3. If *x* has neigbors in several basins, it merges these basins. The new basin retains the lowest minimum of the original basins.



Grouping States Into Macro-States

- Use the flooding algorithm to find basins (*macro-states*), i.e. sets of states sharing saddle points.
- Within a macro-state the transition rates are: $k_{\alpha\beta} = \Gamma_{\alpha\beta} \exp \left(-(E^*_{\alpha\beta} G_{\alpha})/kT\right)$ G_{α} : energy of macro-state. $E^*_{\alpha\beta}$: energy of transition state.
- Only transition rates between macro-states require the master equation.





- Helix Models relax state resolution to the level of helices. This prunes the state space by the number of conformations with basepairs outside of helices.
- 2 The difficulty lies in finding a good model for the energy difference between states. Kinwalker uses the Morgan-Higgs heuristic which provides a microscopic resolution of state transitions.

Example:

- 1 k_+ is comprised only of the entropic component of bending the structure into the shape required by the helix: $\Delta G^F_+ = T * \Delta S$
- 2 k_{-} is made up of an enthalphic component only, i.e. the energy that is released from opening base pairs: $\Delta G^{F}_{-} = \Delta H$

・ロト ・雪 ・ ・ ヨ ・ ・ ヨ

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

< □ > < 同 > < 回 > <

Lessons

- I. Keep the set of states small.
- II. Model state transitions at the level of base pairs.

Implementation

- I. Only consider global minima on substrings of the RNA chain. There are less than $\frac{n^2}{2}$ of these.
- II. Use the Morgan-Higgs approach for barrier calculations, which determines transition rates at base pair level.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

・ロット (雪) (日) (日) 日

Overview

- Find the minimum free energy of all substructures and divide it by the bases in the substructure.
- Populate a matrix with these values and record its local minima.
- Group non-intersecting structures of local minima into an energy front.
- Continously extend the front subject to energy constraints
- Stop when the global mfE structure is reached.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

・ロット (雪) (日) (日) 日

Find optimal substructures

- Calculate the minimum free energies for substructures [i, j], where $1 \le i \le j \le n$.
- Populate an upper triangular matrix C(i, j) with these $\frac{n*(n+1)}{2}$ entries.
- Divide each entry by j i + 1.

This yields the minimum free energy contribution per base for each substructure.

Outline Kinwalker

Algorithm

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

The relative free energy matrix

A 1	C 2	A 3	G 4	G 5	U 6	U 7	C 8	G 9	с 10	с 11	U 12	G 13	U 14	G 15	U 16	U 17	G 18	С 19	G 20	A 21	A 22	C 23	с 24	U 25	G 26	С 27	G 28	G 29	G 30	U 31	U 32	С 33	G 34	
0	0	0	0	0	0	0	0	0	0	6	8	11	11	13	13	13	13	13	15	15	17	20	20	21	24	24	24	24	24	24	26	26	30	Α
	0	0	0	0	0	0	0	0	0	6	8	8	11	11	11	11	11	12	15	15	15	18	19	21	21	24	24	24	24	24	25	25	30	С
		0	0	0	0	0	0	0	0	6	6	8	8	8	8	8	8	12	12	12	13	16	19	19	21	21	21	21	21	21	22	25	27	A
			0	0	0	0	0	0	0	0	6	6	6	6	6	6	8	10	11	11	13	16	16	19	19	19	19	19	20	20	21	23	25	G
				0	0	0	0	0	0	0	0	0	0	4	4	4	4	6	9	11	13	13	16	16	16	16	16	19	20	20	20	21	24	G
					0	0	0	0	0	0	0	0	0	4	4	4	4	6	9	11	11	13	13	13	13	13	16	19	19	19	19	19	24	U
						0	0	0	0	0	0	0	0	4	4	4	4	6	9	9	11	11	11	11	11	13	16	17	18	19	19	19	24	U
							0	0	0	0	0	0	0	0	4	4	4	6	6	9	9	9	9	9	9	13	13	17	18	19	19	19	20	С
								0	0	0	0	0	0	0	0	4	4	4	6	6	6	7	7	7	9	11	13	17	18	18	19	19	20	G
									0	0	0	0	0	0	0	0	2	4	5	5	5	7	7	7	9	11	12	15	17	18	18	18	18	с
										0	0	0	0	0	0	0	2	4	5	5	5	7	7	7	9	11	12	12	15	15	15	17	18	С
											0	0	0	0	0	0	0	4	4	5	5	7	7	7	7	11	11	12	12	13	15	17	18	U
												0	0	0	0	0	0	0	4	4	4	7	7	7	7	7	11	11	11	12	14	17	17	G
													0	0	0	0	0	0	0	4	4	7	7	7	7	7	8	10	11	12	14	14	16	U
														0	0	0	0	0	0	0	4	4	7	7	7	7	8	10	11	11	11	13	16	G
															0	0	0	0	0	0	0	4	4	4	4	7	8	10	10	10	10	13	16	U
																0	0	0	0	0	0	0	0	0	4	7	7	8	8	10	10	13	16	U
																	0	0	0	0	0	0	0	0	4	4	7	7	7	8	10	13	16	G
																		0	0	0	0	0	0	0	0	4	4	6	6	8	10	13	13	C
																			0	0	0	0	0	0	0	0	0	6	6	8	10	10	13	G
																				0	0	0	0	0	0	0	0	6	6	8	8	10	10	A
																					0	0	0	0	0	0	0	6	6	6	8	8	8	A
																						0	0	0	0	0	0	0	6	6	6	6	6	C
																							0	0	0	0	0	0	0	4	4	4	6	C
																								0	0	0	0	0	0	0	0	0	6	U
																									0	0	0	0	0	0	0	0	6	G
																										0	0	0	0	0	0	0	0	С
																											0	0	0	0	0	0	0	G
																												0	0	0	0	0	0	G
																													0	0	0	0	0	G
																														0	0	0	0	U
																															0	0	0	U
																																0	0	С
	-																																0	G

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

< □ > < 同 > < 回 > <

-

tRNA phe

Primary structure: GCGGAUUUAGCUCAGDDGGGAGAGCGCCAGACUGAAYA UCUGGAGGUCCUGUGTPCGAUCCACAGAAUUCGCACCA



Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

Contourplot of tRNA phe for stacksize 1 and 3





◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - シスペ

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

・ロト ・雪 ・ ・ ヨ ・ ・ ヨ

Initial Energy Front

- The state of the structure for a given thermal energy is reflected in the energy front.
- This front is a set of local minima in the relative energy matrix, i.e. disjoint optimal substructures.
- The algorithm progresses through rounds, trying to update the energy front in each turn.
- The order in which minima are tested for admissibility is given by:
 - a Increasing index of the diagonal
 - b Increasing distance from the edge of the matrix
 - c 5' before 3'

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

A D > A P > A D > A D >



Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

◆□▶ ◆□▶ ▲□▶ ▲□▶ ▲□ ◆ ◇ ◇ ◇

Energy Front Progression

- The algorithm extends the energy front by local minima until the mfE structure is reached.
- In each round the maximal energy barrier between the front and admissible extension points is incremented.
- The Morgan-Higgs heuristic determines the lowest saddle point between a minimum and the energy front.
- The minimum is added to the energy front, if
 E(*front*) *E*(*saddle*) <= *energy_barrier*.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

Energy Front Progression





▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● のへで

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

・ロッ ・ 一 ・ ・ ・ ・ ・ ・ ・ ・

3

The Morgan-Higgs heuristic aims at determining the saddle

point between two conformations A and B. It only considers

direct routes, that is routes that only change contacts in the

symmetric difference $A \triangle B$ of A and B.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

◆ロ ▶ ◆ 圖 ▶ ◆ 圖 ▶ ◆ 圖 ■ ● ● ● ●

Algorithm

Let $A_{add} = B \setminus A$ be the contacts to add to get from A to B. Let $A_{remove} = A \setminus B$ be the contacts to remove to get from A to B. Sort A_{add} by ascending number of conflicts with A_{remove} .

For each basepair $p \in A_{add}$:

Remove the conficts of p in A_{remove} from the structure. Add all elements in A_{add} without conficts to the structure. Record the maximum of the old and the new energy.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

< 日 > < 同 > < 回 > < 回 > < □ > <

3

Comments:

- 1 The Morgan-Higgs heuristic returns the energy barrier of the lowest traversed path. There is no guarantee that the choice of routes includes the lowest direct route.
- 2 When there are several base pairs with an equal number of conflicts, paths for each possible ordering must be calculated.

Few states, microscopic transitions Algorithm The Morgan-Higgs Heuristic

◆□ > ◆□ > ◆注 > ◆注 > 「注 」

Acknowledgments

Christoph Flamm, Peter Stadler