

Kinwalker parameter space analysis

Mario Koestl

Institute for Theoretical Chemistry
University of Vienna

February 12, 2016

Overview

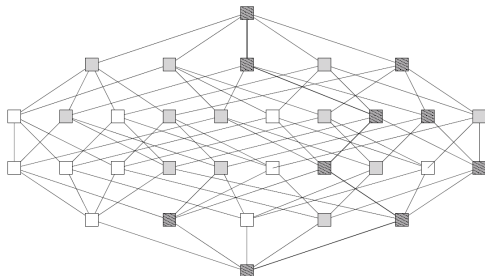
- ▶ co-transcriptional folding = RNA folding at the growing 5' end
- ▶ co-transcriptional pathway prediction difficult
- ▶ Kinwalker algorithm: heuristic approach for RNA folding kinetic prediction

Kinwalker

- ▶ predicts secondary structure trajectories up to 1500 nt
- ▶ uses structures backtracked from $C_{i,j}$ matrix after each base adding step.
- ▶ 2 direct pathfinding algorithm : "Morgan Higgs" and "findpath"

Kinwalker parameters used

- ▶ transcription rate
- ▶ barrier heuristic
- ▶ dangle model
- ▶ max keep



data used

- ▶ 200 different parameter combinations for 3 RNA families
- ▶ SRP, TRP and RNaseP
- ▶ alignment for each family from Rfam database, according to Meyer et al.
- ▶ reference structures taken from literature, reference sequences downloaded from NCBI

Scoring functions

- ▶ comparing Kinwalker output data with different parameters
- ▶ Structure Conservation Index and Sequence Similarity
- ▶ Base-pair diversity, ensemble diversity, ensemble distance

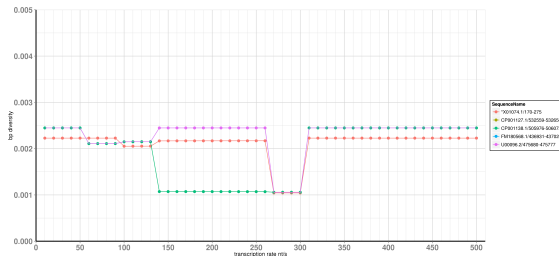
ensemble distance

- ▶ expected distance between folding trajectory and a reference structure

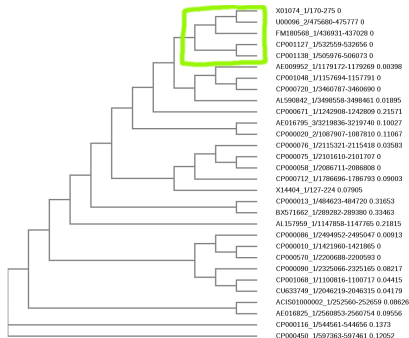
$$\text{ensemble distance}(s, \text{ref}) = \sum_{i,j \in \text{ref}} (1 - p(i,j)) + \sum_{i,j \notin \text{ref}} p(i,j), \quad (1)$$

- ▶ $p(i,j) = \frac{\Delta \text{Time}}{T_{\text{total}}}.$

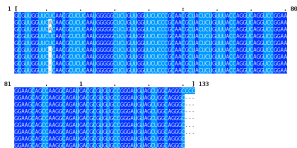
Results



sequence	trap at 260 nt/s	trap at 300 nt/s	trap at 310 nt/s
X01074.1.170-275	yes	no	yes
U00096.2.475680-475777	yes	no	yes
FM180568.1.436931-437028	yes	no	yes
CP001138.1.505976-506073	yes	no	yes
CP001127.1.532559-532656	no	no	yes

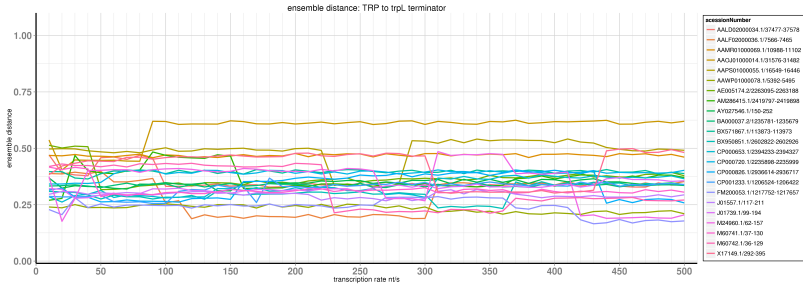


1 X01074_1/170275 100.0%
 2 CP001127_1/5325 99.2%
 3 CP001138_1/5259 99.2%
 4 U00096_2/475680 100.0%
 5 FM180568_1/4369 100.0%
 consensus/100%
 consensus/90%
 consensus/70%
 1 X01074_1/170275 100.0%
 2 CP001127_1/5325 99.2%
 3 CP001138_1/5259 99.2%
 4 U00096_2/475680 100.0%
 5 FM180568_1/4369 100.0%
 consensus/100%
 consensus/90%
 consensus/70%



TRP family

► Ensemble Distance for trp sequences to terminator structure



conclusion

blablabla, auch was das es halt icht geht