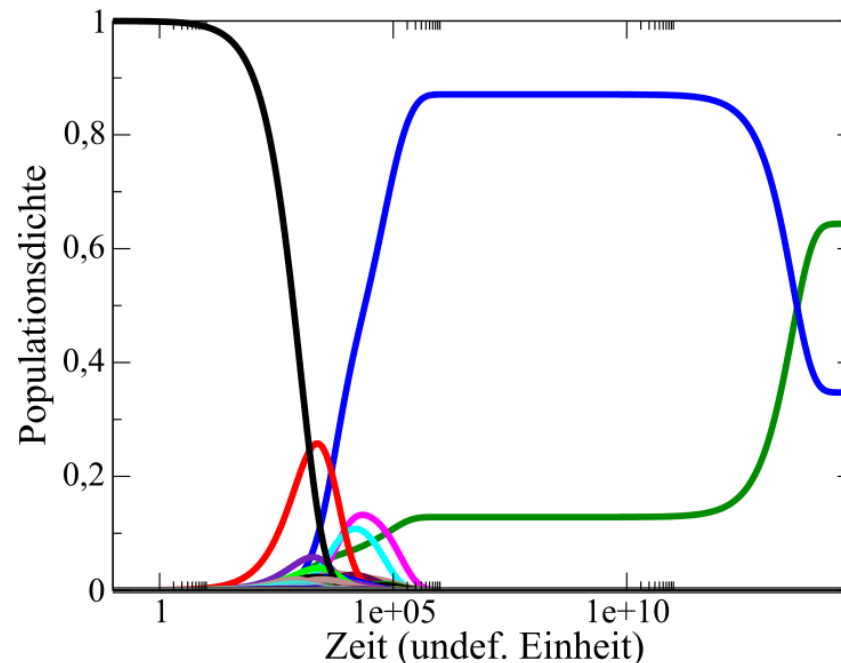


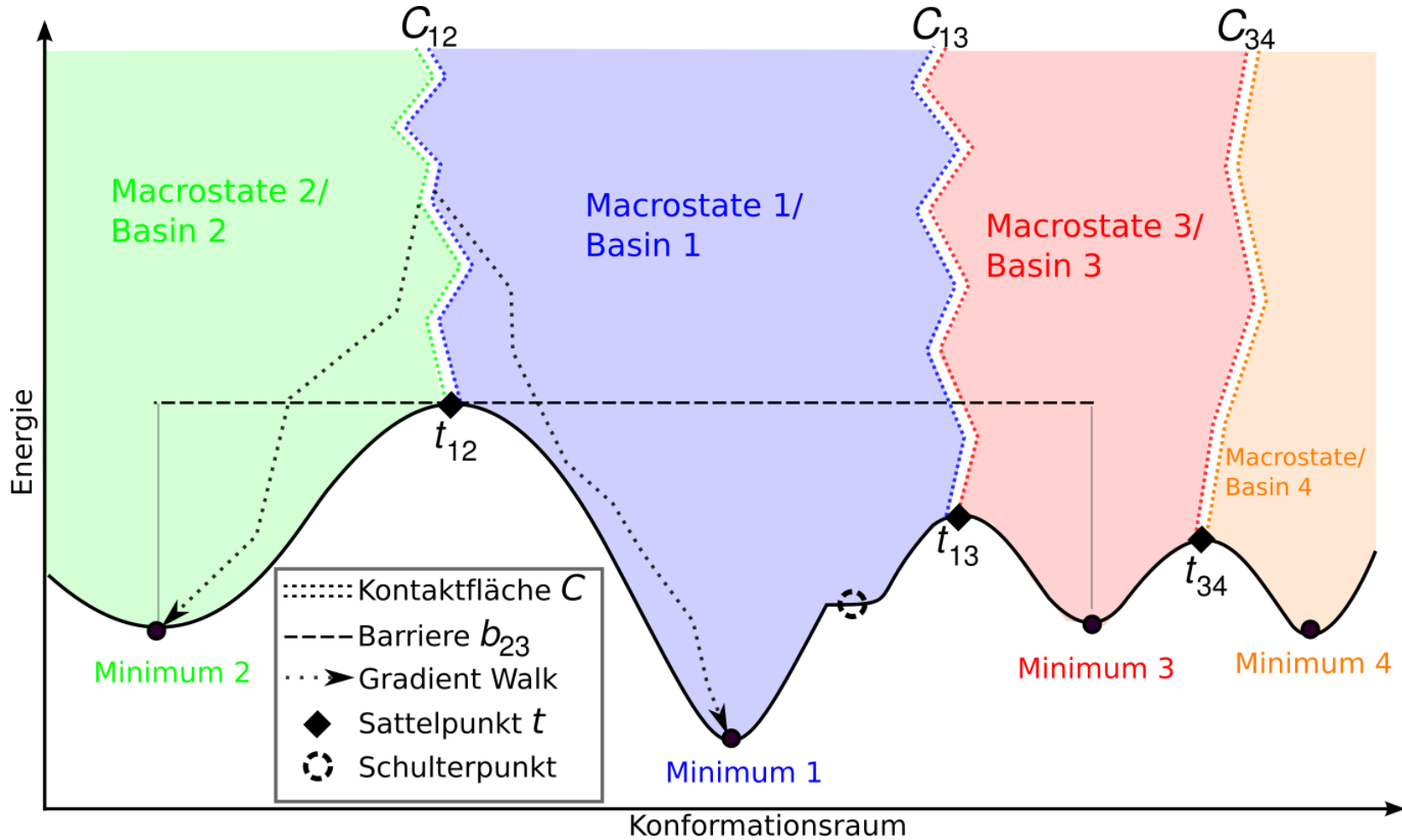
# Approximate RNA kinetics

# RNA Kinetic

- $\frac{d\vec{D}}{dt} = R(t)\vec{D}$  R=rate matrix, D=Populationdensity
- Challenge: Kinetics for long RNA sequences
- Rate matrix calculation

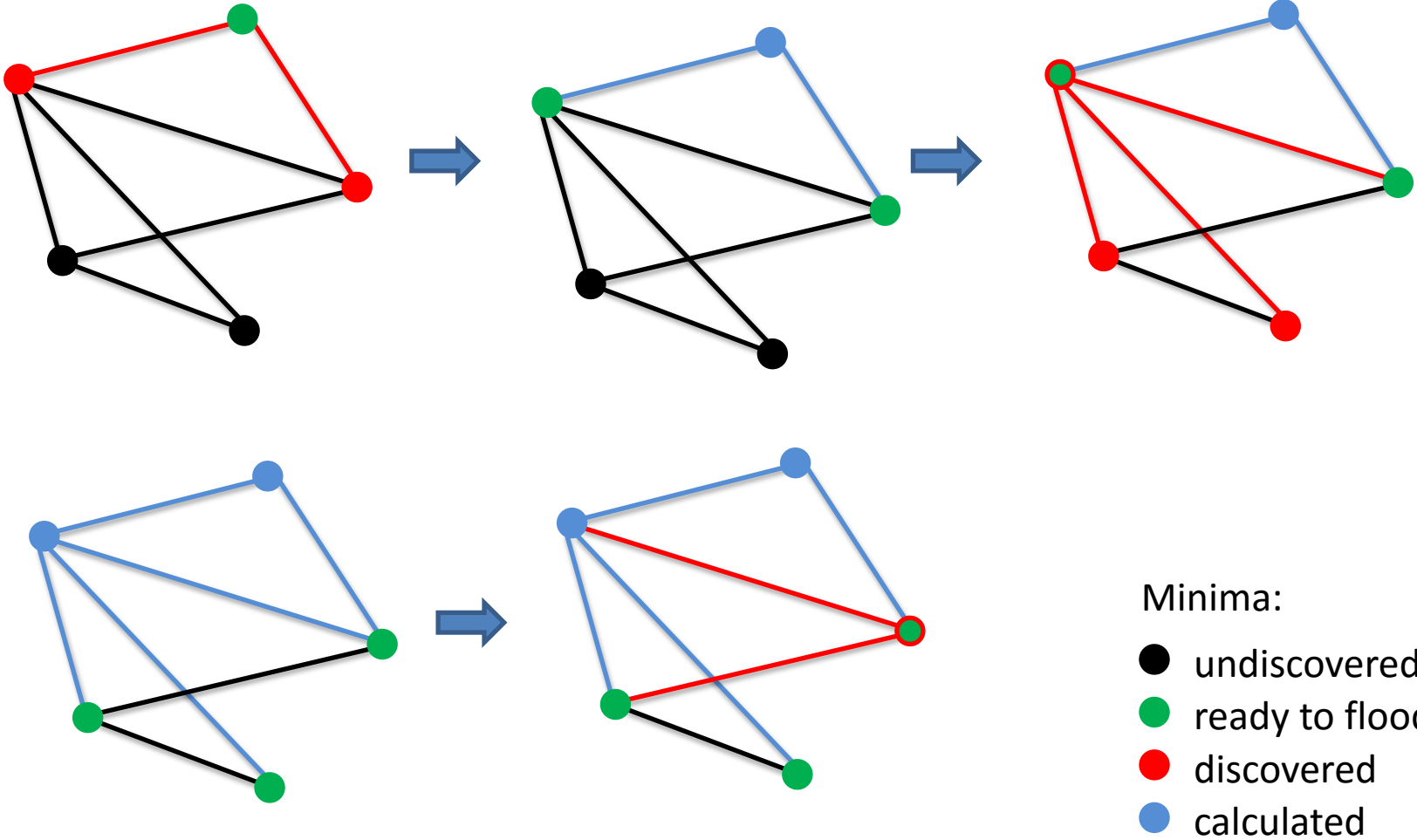


# Energy landscape

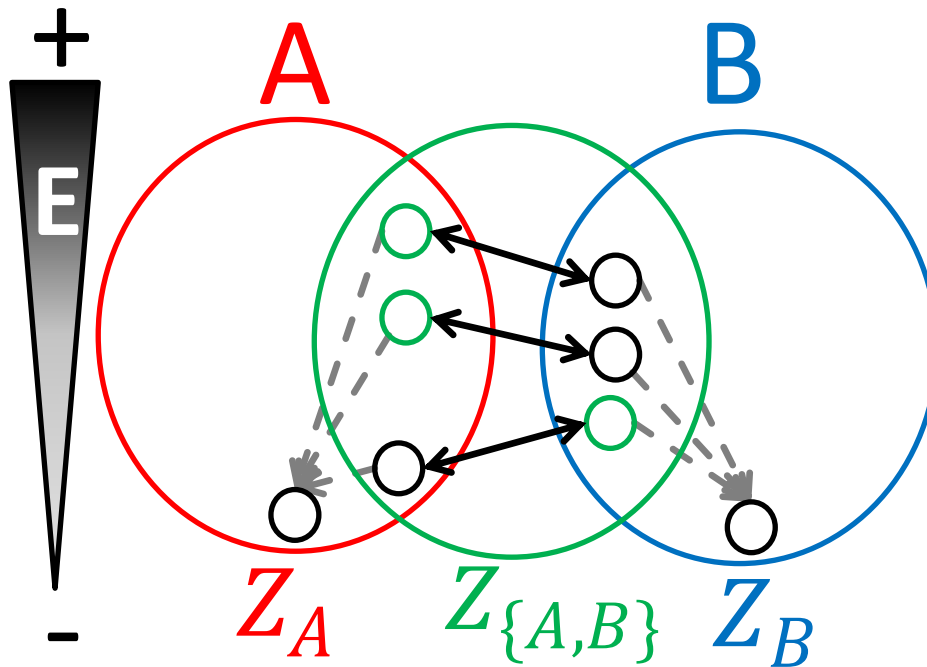


Quelle: D.A. Hannes Kochniß 2008

# Explorative flooding



# Rate calculation



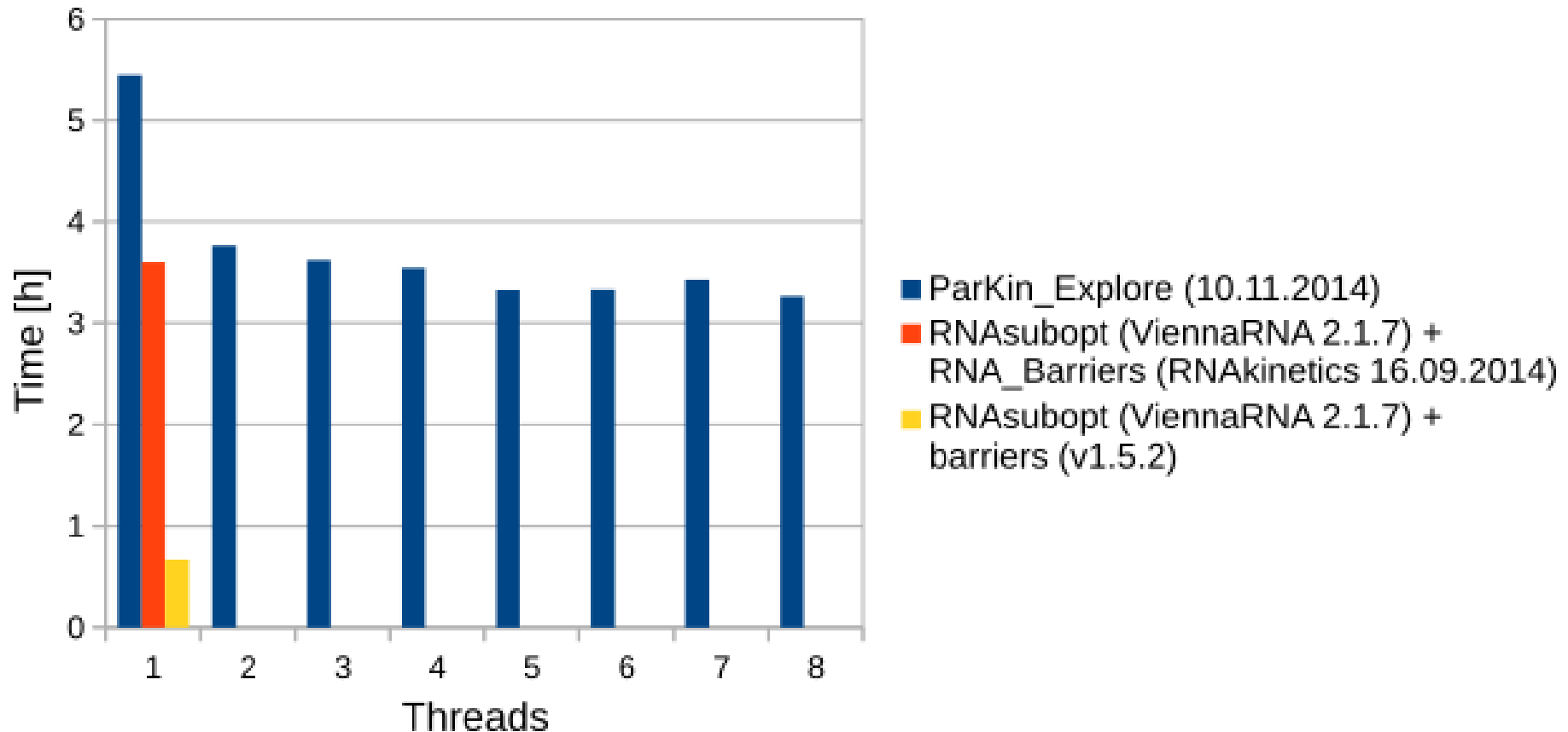
- $Z_A = \sum_{x \in A} e^{\frac{-E(x)}{RT}}$
- $rate_{\{A \rightarrow B\}} = \frac{Z_{\{AB\}}}{Z_A}$

$\circ \longleftrightarrow \circ$  = neighbored structures; +/- 1 basepair

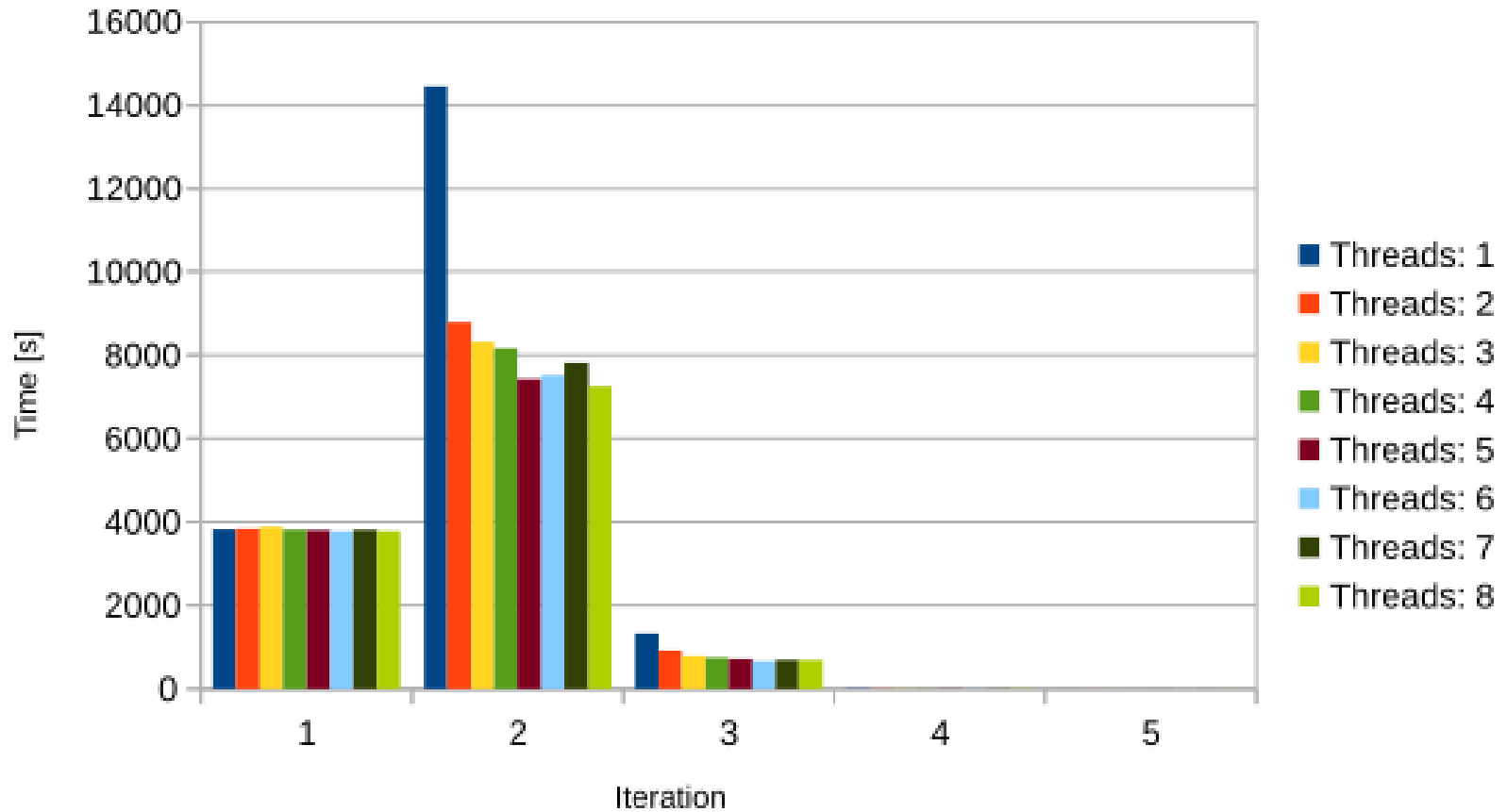
# Solution

- Global explorative flooding
- Disadvantage:
  - more computing time
- Advantage:
  - less memory consumption
  - allows more filters
  - easy to parallelize

# Comparison



# Parallelization

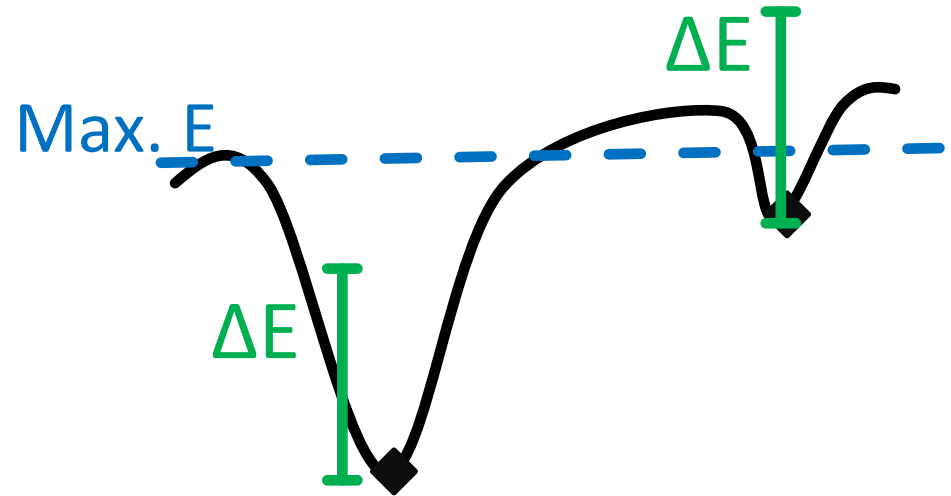




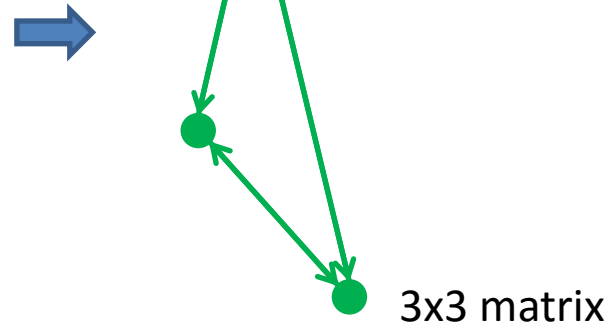
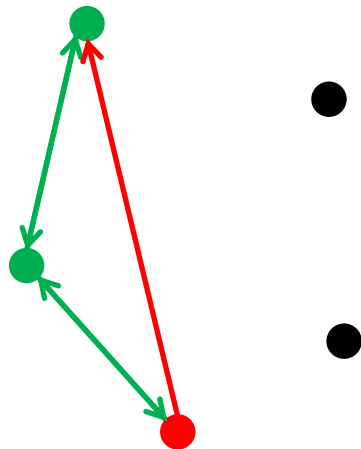
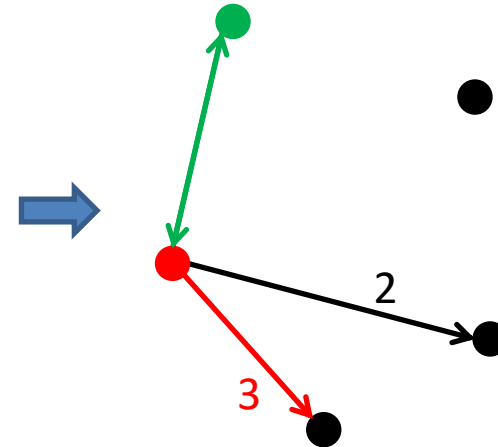
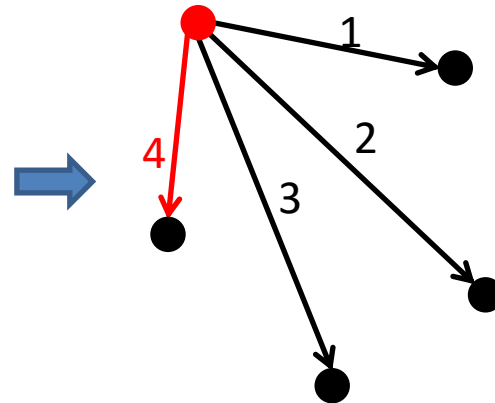
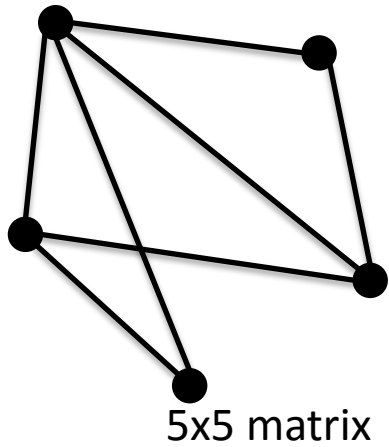
# Filter



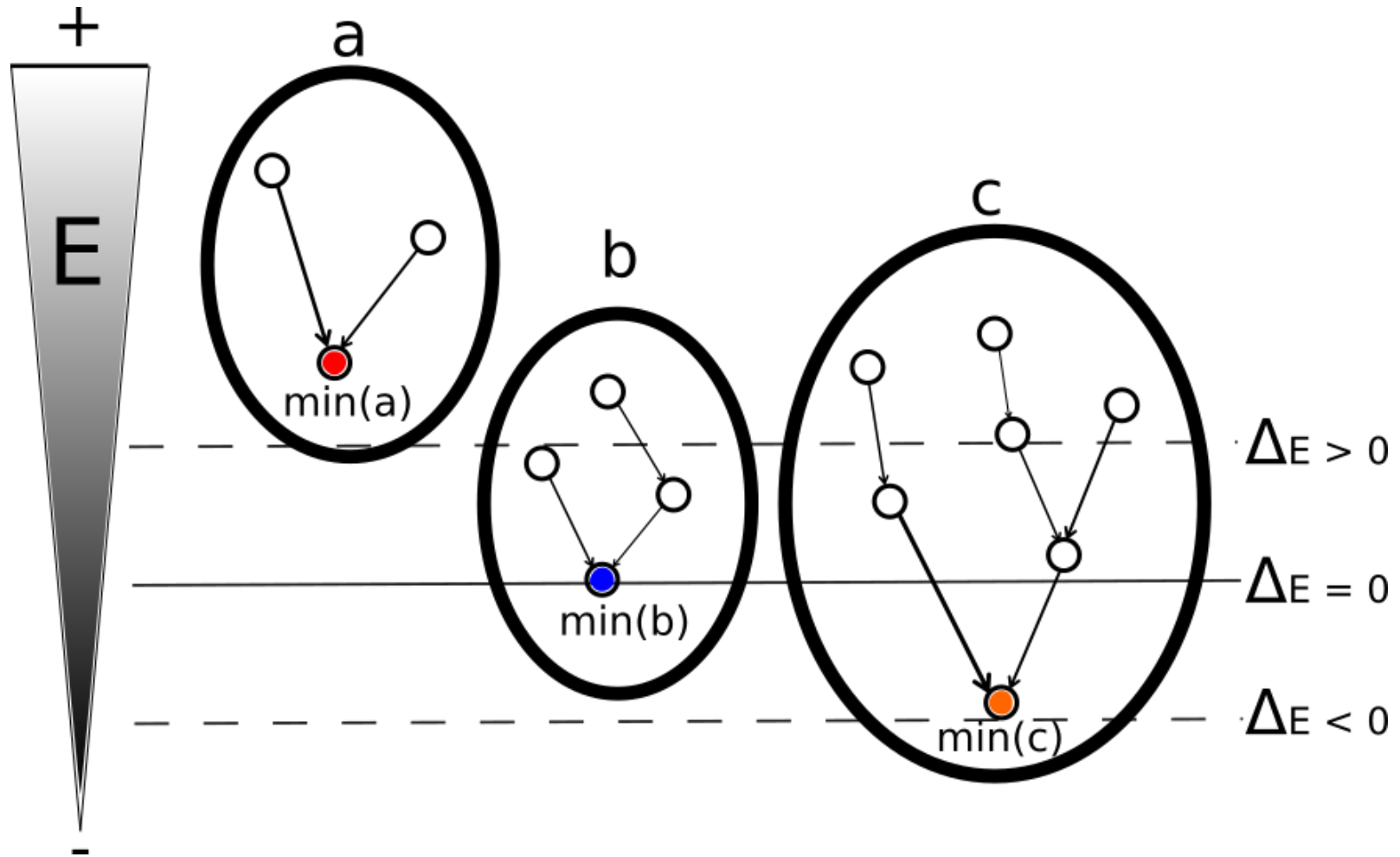
- Micro-state filter
  - Maximum energy
  - Delta energy
- Macro-state filter
  - K-Best filter
  - DeltaMinE filter



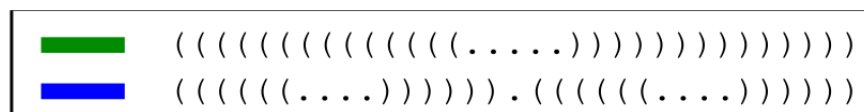
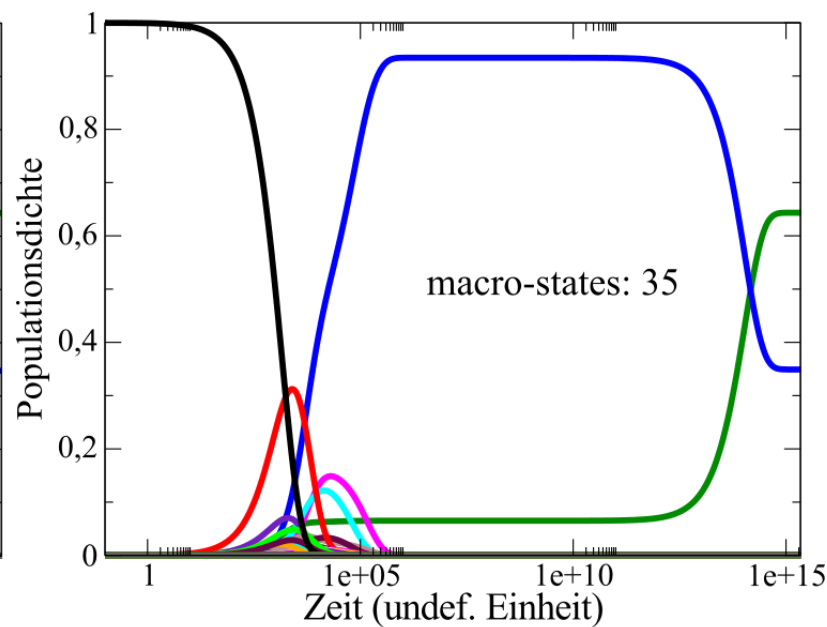
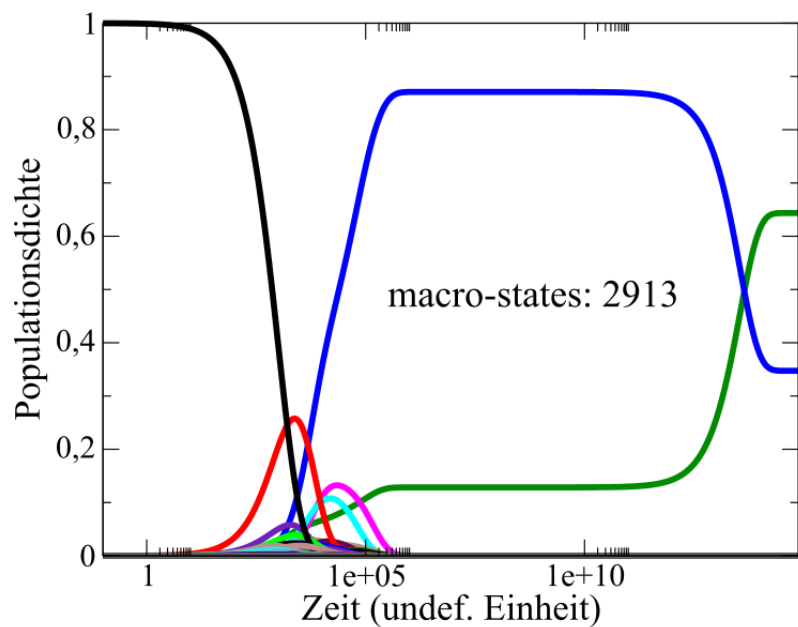
# Best K Pruning



# DeltaMinE Filter



# Filtered data

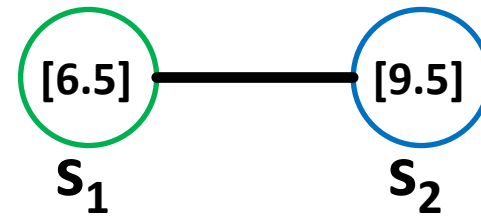
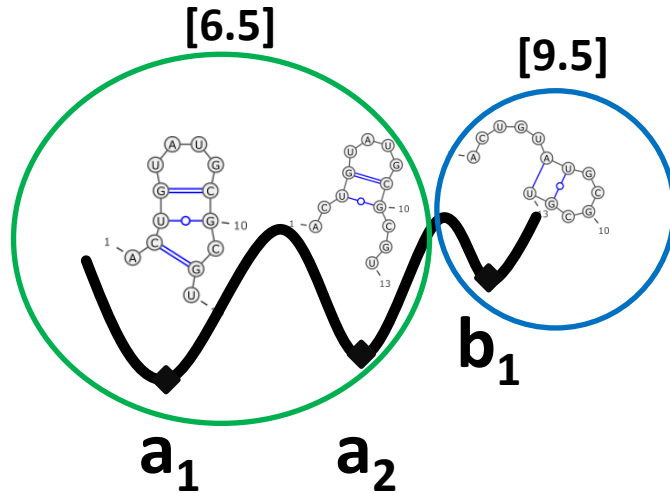


unfiltered  
**Computing time: 0.5h to 3.5h**

Delta-Energie filter with 10 kcal/mol ,  
K-Best filter = 4 ,  
DeltaMinE filter = 1  
**Computing time: 0.5 min**

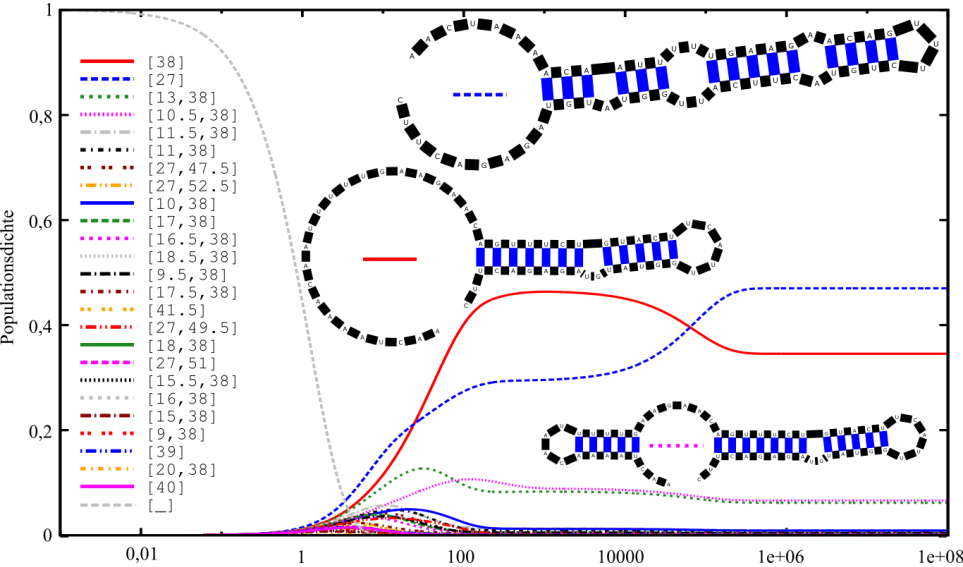
# Hi Shapes

- Despite filter many macro-states
- Large rate matrix  $\rightarrow$  inversion takes much time
- Solution: further clustering with Helix-Index Shapes

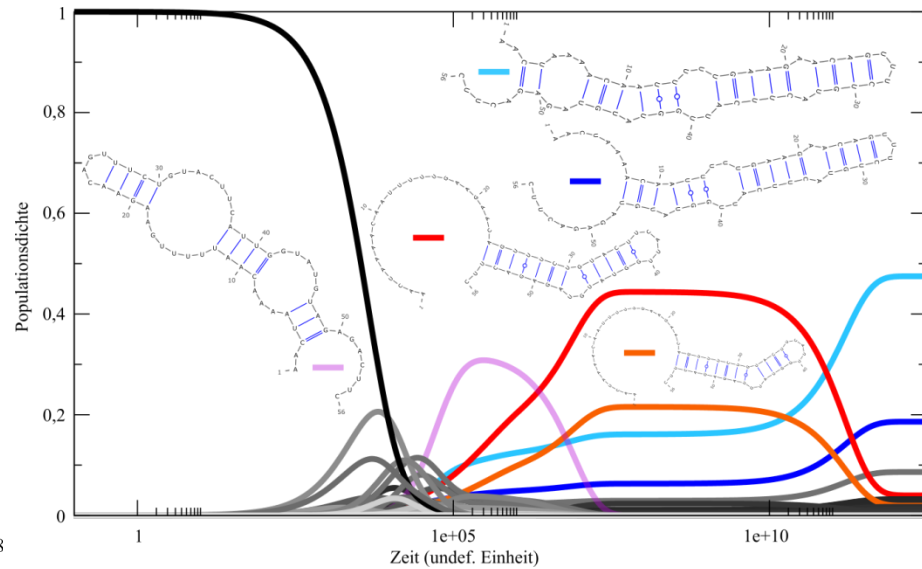


$$rate\_shapes_{\{S_1, S_2\}} = \frac{\sum_{A \in S_1} Z_{\{AB\}}}{\sum_{A \in S_1} Z_A}$$

# HiShape vs Explorative + Filter

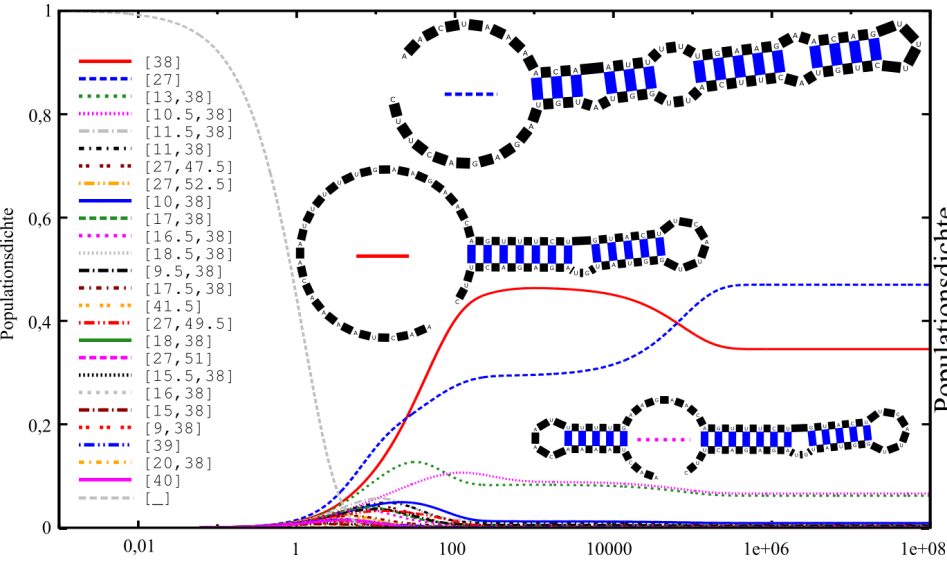


253 min  
 Hardware: 64 cores of a  
 4 x AMD Opteron 6282 SE machine,  
 RAM: 512 GB

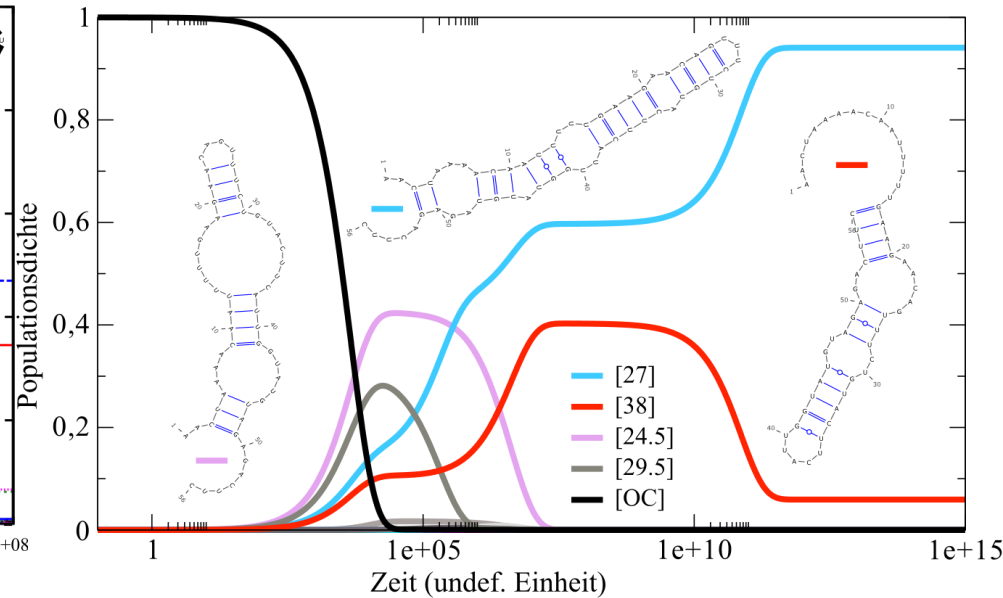


1,6 sec  
 Hardware: Intel i7 Quadcore 2GHz  
 RAM: 14GB

# Leptomonas mapped to Hi-Shapes

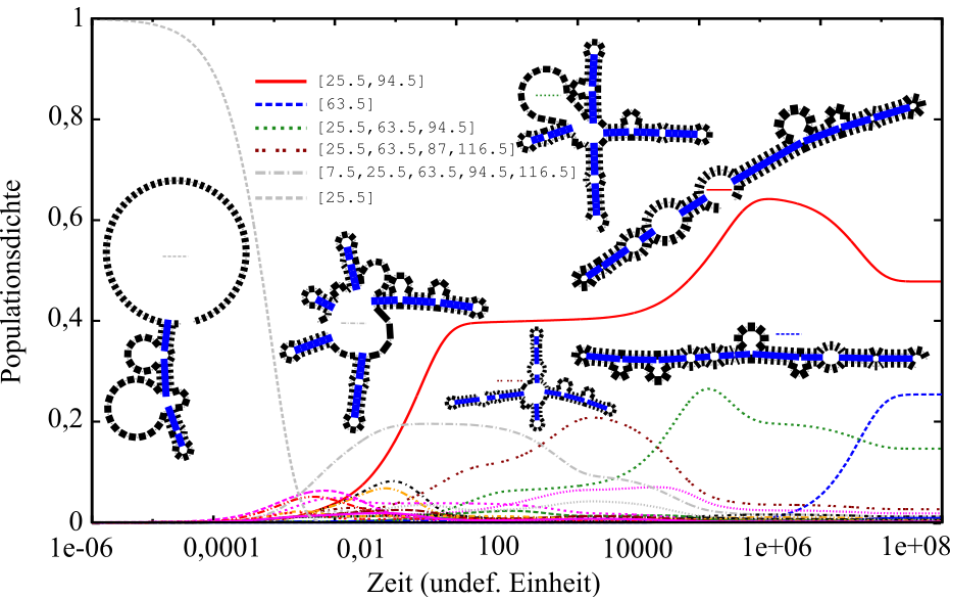


253 min  
 Hardware: 64 cores of a  
 4 x AMD Opteron 6282 SE machine,  
 RAM: 512 GB



1,6 sec  
 Hardware: Intel i7 Quadcore 2GHz  
 RAM: 14GB

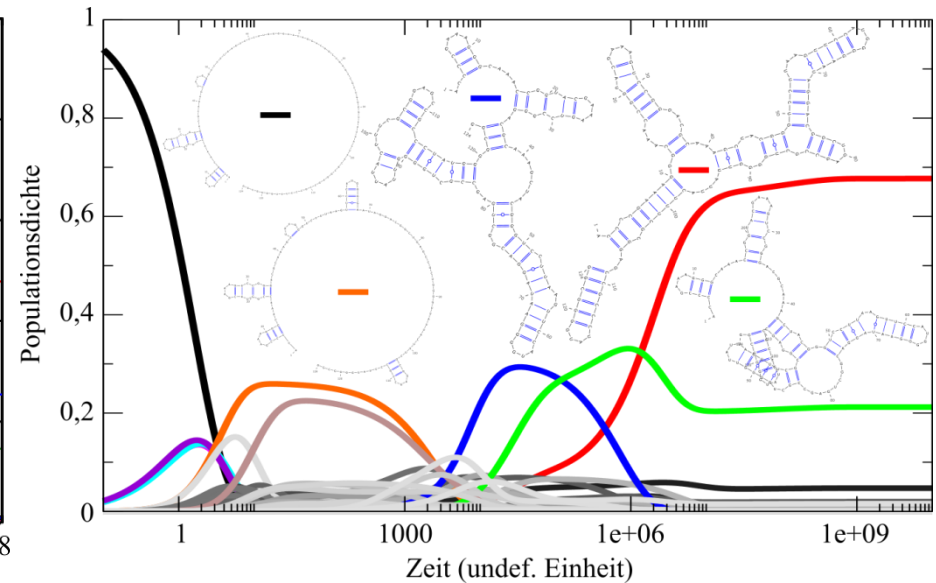
# Explorative + Filter vs HiShape



**24 h**

**26 Hi-Shapes**

Hardware: 64 cores of a  
4 x AMD Opteron 6282 SE machine,  
RAM: 512 GB



**3 h 20 m**

14981 macrostates as **395 Hi-Shapes**

Hardware: Intel i7 Quadcore 2GHz

RAM: 14GB



# Conclusion

- Problem:
  - Long sequences → many states → much time
- Solution:
  - Explorative flooding:
    - Allows more filters
    - Easier parallelization
- Optimization possibilities:
  - Clustering
  - Rate calculation
  - Other filters