Dynamics of water and conformational transitions in proteins

Dmitry Nerukh

Aston University

TACC 2012, 4 Sep 2012

Introduction

We study protein (peptide) dynamics.

All biomolecular processes happen in solution: water is a key player.

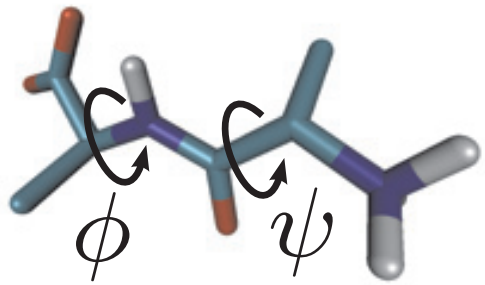
Frauenfelder *et al* [1] have experimentally shown that protein motions are slaved by water.

The protein molecule provides an 'active matrix' necessary for guiding the water's dynamics towards the biologically relevant changes of the 3D structure (the proteins 'function' or folding).

[1] H. Frauenfelder, et al, *PNAS*, **106**, 512 (2009)

Introduction

By ‘dynamics’ we mean time dependent properties of backbone conformation quantified by the values of the dihedral angles:



L-alanyl-L-alanine zwitterion

Introduction

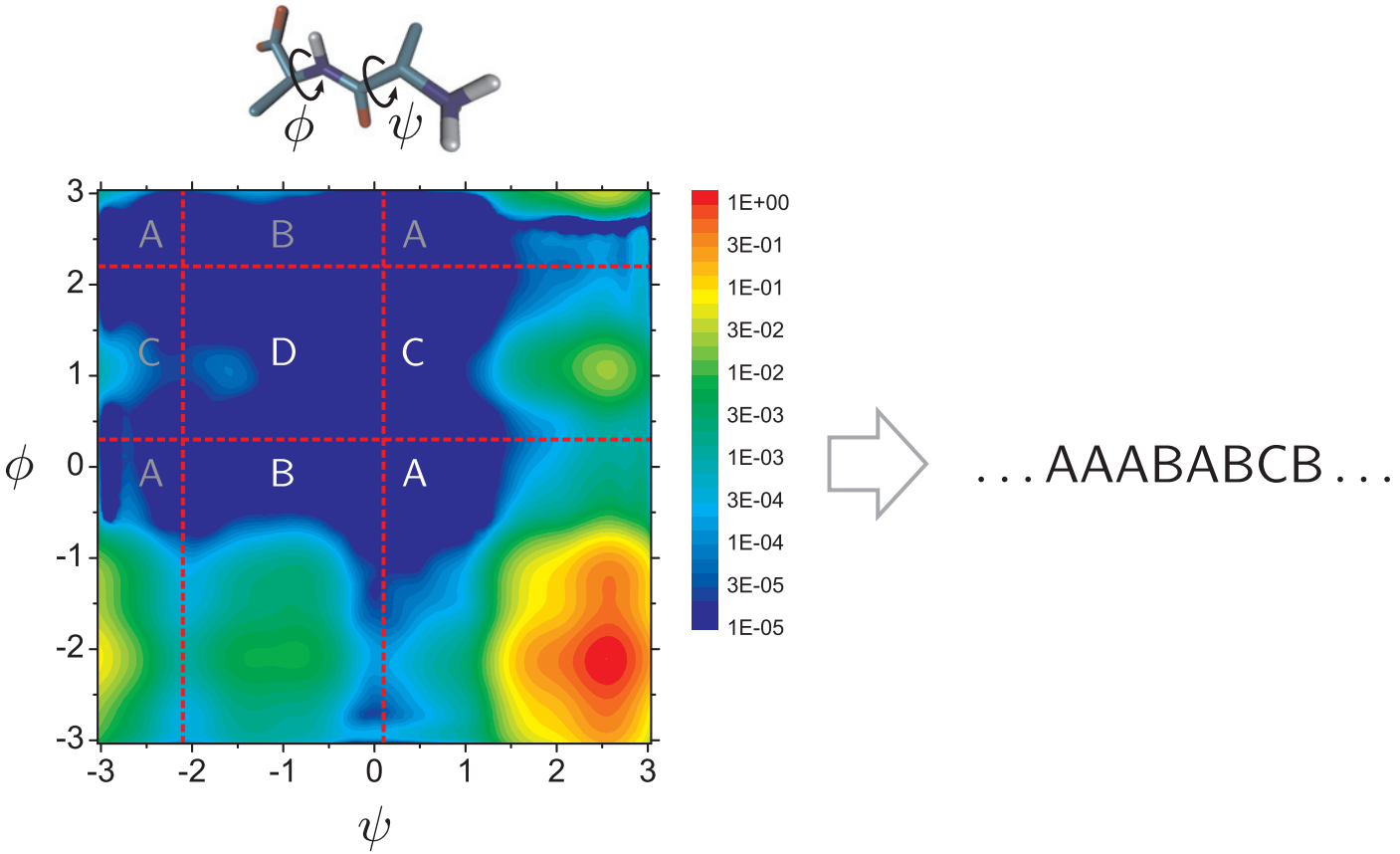
The method: classical MD in *explicit* water.

The protein is assumed to be in 'metastable' conformational states most of the time with infrequent and quick transitions between them.

The main question: how do water molecules influence (if they do) the backbone motion and lead to the conformational transitions?

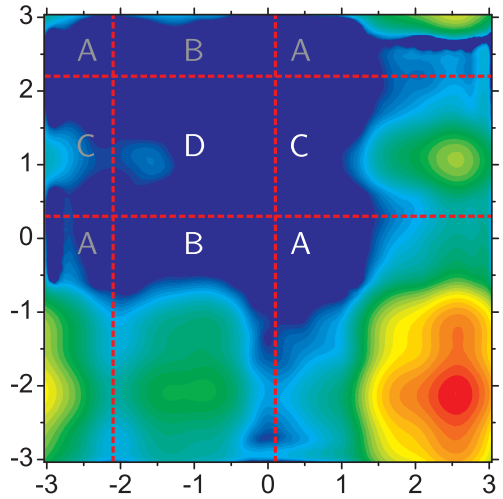
More specifically: how do molecular trajectories behave before the transition and what is the difference compared to the stable periods?

Quantifying conformational dynamics



Probabilities of conformations for a 1μs trajectory and its symbolisation

Markov State Model



s_i \ s_{i+1}	A	B	C
A	0.996	0.003	0.001
B	0.261	0.737	0.002
C	0.097	0.002	0.901

The time step is $6ps$.

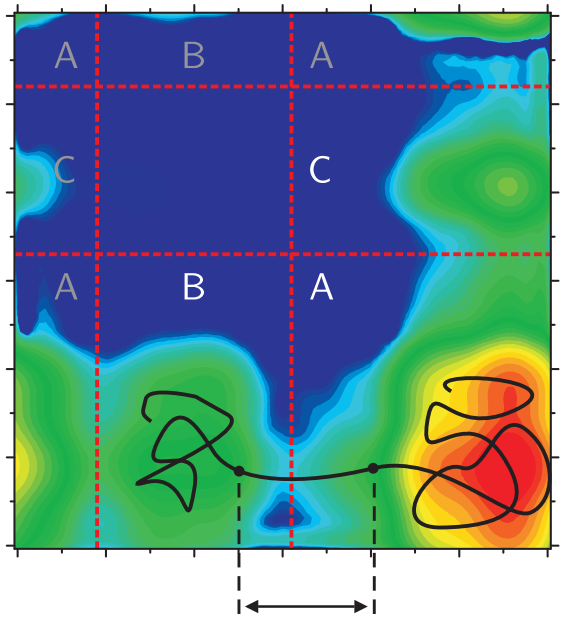
Markov State Model

A transition is when symbol 'A' at time step i changes to symbol 'B' at time step $i+1$.

s_{i+1} / s_i	A	B	C
A	0.996	0.003	0.001
B	0.261	0.737	0.002
C	0.097	0.002	0.901

$A \rightarrow B$ happens once in $3.1ns$ on average.

Markov State Model



Markov State Model is valid only for time steps longer than the length of the transition process.

It does not say anything about *how* and *why* the transitions happen.

Solution: build a model that works for smaller time steps.

Hidden Markov model

Instead of the conformational states, s_i, themselves (A,B,C) we consider the l-long *sequences* of states
$\overleftarrow{s}_i \equiv \{s_{i-(l-1)} \dots s_{i-2} s_{i-1} s_i\}$.

Hidden Markov model

Instead of the conformational states, s_i, themselves (A,B,C) we consider the l-long *sequences* of states
$\overleftarrow{s}_i \equiv \{s_{i-(l-1)} \dots s_{i-2} s_{i-1} s_i\}$.

There are at most 3^l such states.

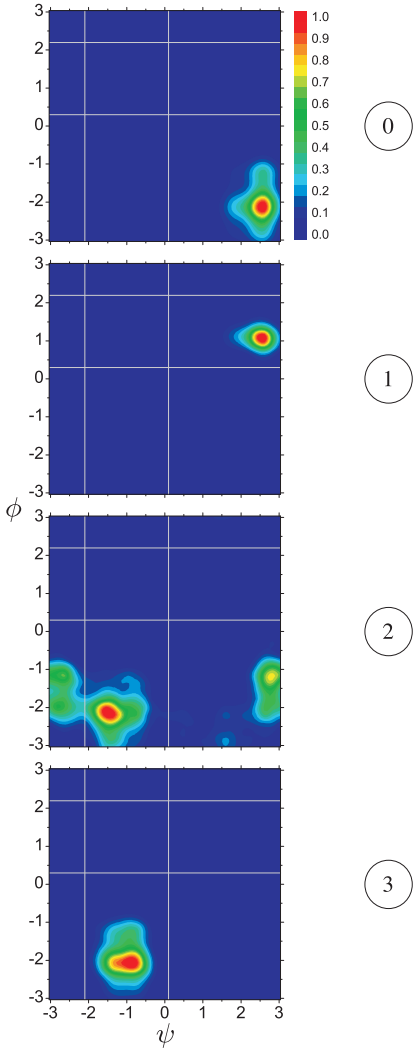
Hidden Markov model

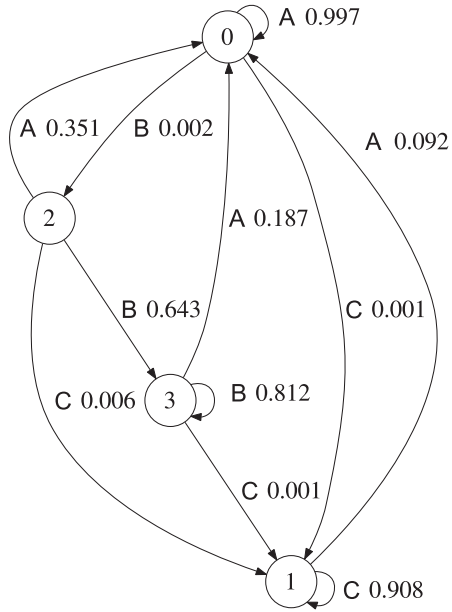
Instead of the conformational states, s_i, themselves (A,B,C) we consider the l-long *sequences* of states
$\overleftarrow{s}_i \equiv \{s_{i-(l-1)} \dots s_{i-2} s_{i-1} s_i\}$.

There are at most 3^l such states.

We build a Markov model on these states.

Hidden Markov model





For $5ps$ time step the model converges at $l = 2$.
State '0' corresponds to the original state 'A': it consists of the sequences 'AA'.
State '3' corresponds to 'B'.
State '2' (sequences 'AB') describes the transitions between 'A' and 'B'.

The tool for studying *how* and *why* transitions happen

The time resolution is limited only by the amount of data (in contrast to the standard MSM where there is a fundamental limit).

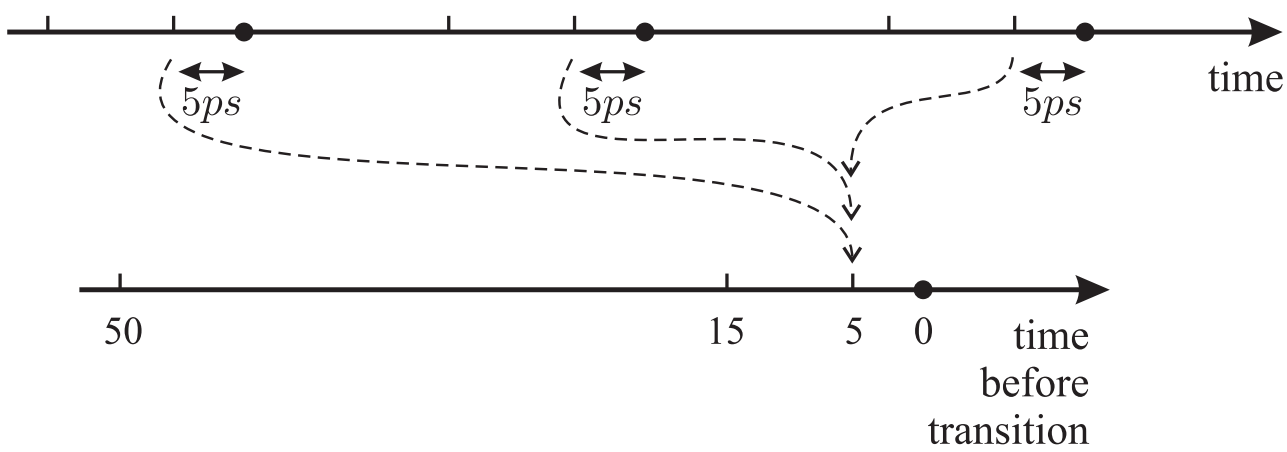
Thus, we can identify the transitions with the precision of $0.3ps$.

The tool for studying *how* and *why* transitions happen

The time resolution is limited only by the amount of data (in contrast to the standard MSM where there is a fundamental limit).

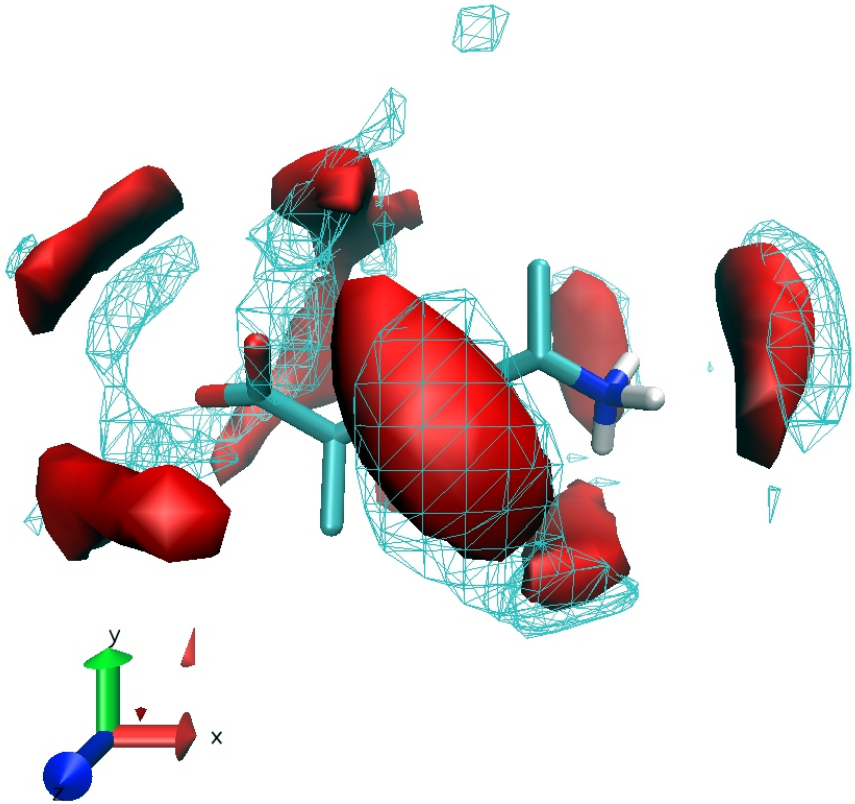
Thus, we can identify the transitions with the precision of $0.3ps$.

We collect the time frames at specific times *before* the transitions.



How water contributes to the transition from A to B

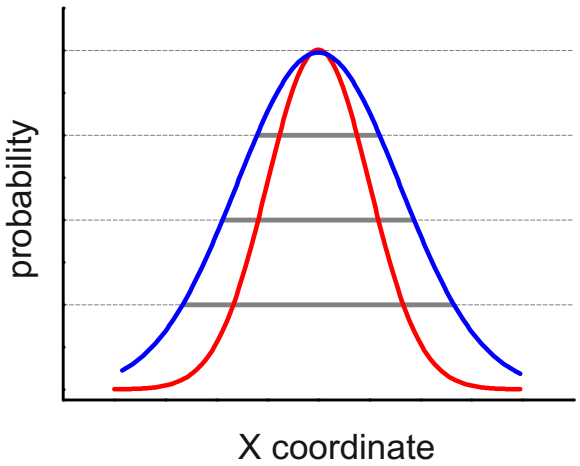
The high probability areas for oxygen (hydrogen) atoms is calculated by averaging over the selected time frames.

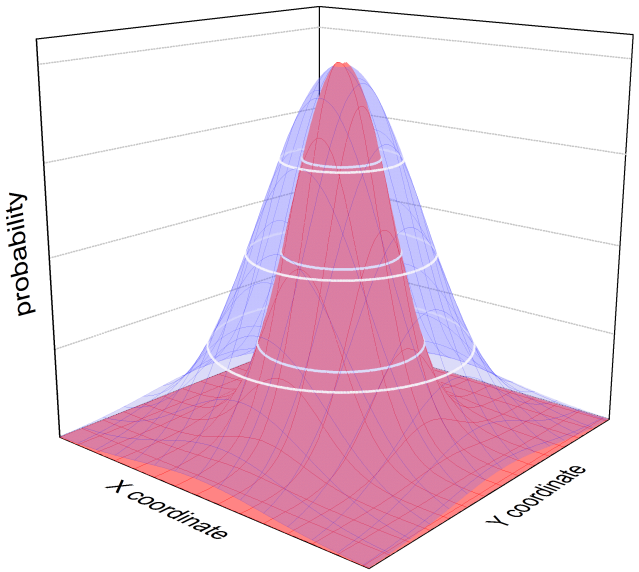


High probability areas of oxygen (red) and hydrogen (blue) in state A

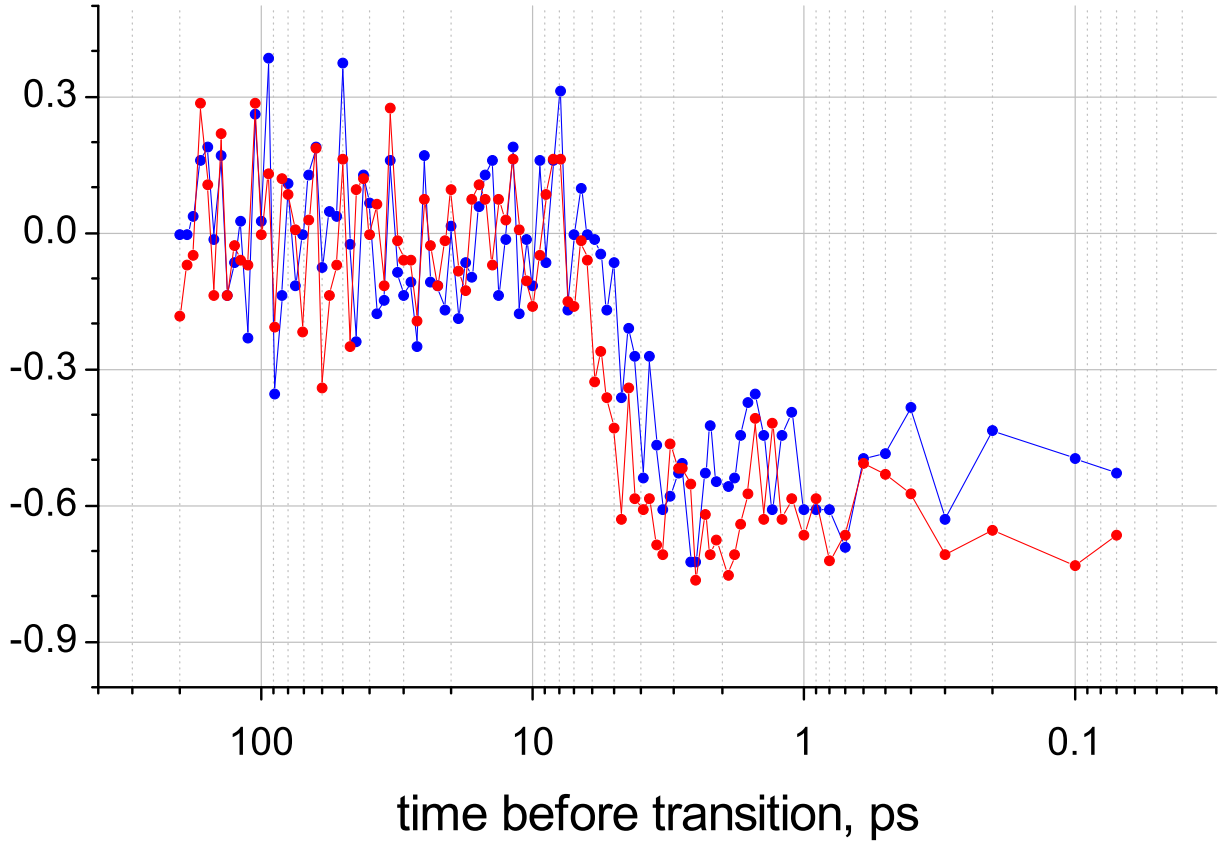
The flow of trajectories

More concerted trajectories form narrower probabilities





The flow of trajectories



The volume of high probability areas for oxygen (red) and hydrogen (blue) during the transition

The flow of trajectories

Water concentrates at more specific locations just before the transition.

This means that the phase space trajectories pass through a 'bottleneck'.

Water distribution correlation with the peptide motion

We quantify the degree of dependence between the high probability water areas and the dihedral angles of the peptide.

Method: Linear Stochastic Estimation.

The probability field $f(\mathbf{x}, t)$ is first converted to time fluctuations:

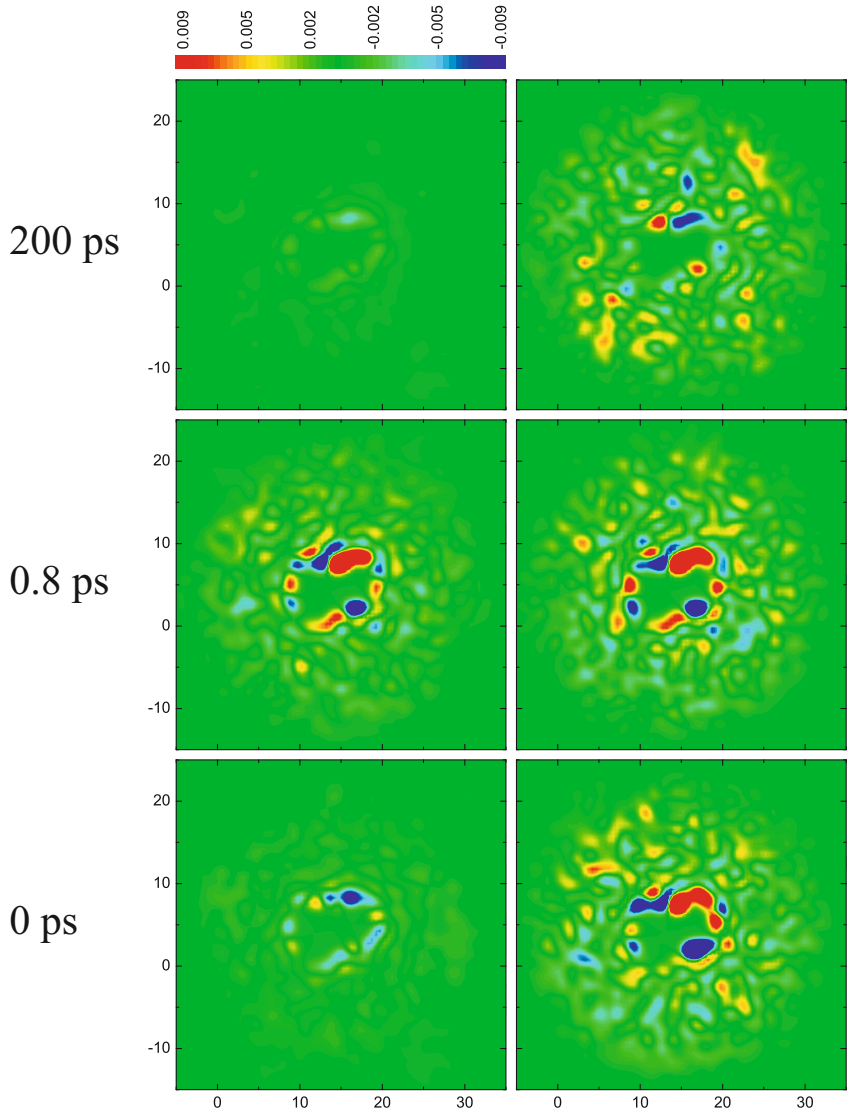
$$f'(\mathbf{x}, t) = f(\mathbf{x}, t) - \langle f(\mathbf{x}, t) \rangle_t.$$

Then a linear stochastic fit

$$f'_c(\mathbf{x}, t) = \alpha(\mathbf{x})\phi(t) + \beta(\mathbf{x})\psi(t)$$

to $f'(\mathbf{x}, t)$ is computed by minimising the statistical error $\langle f' - f'_c \rangle_t$.

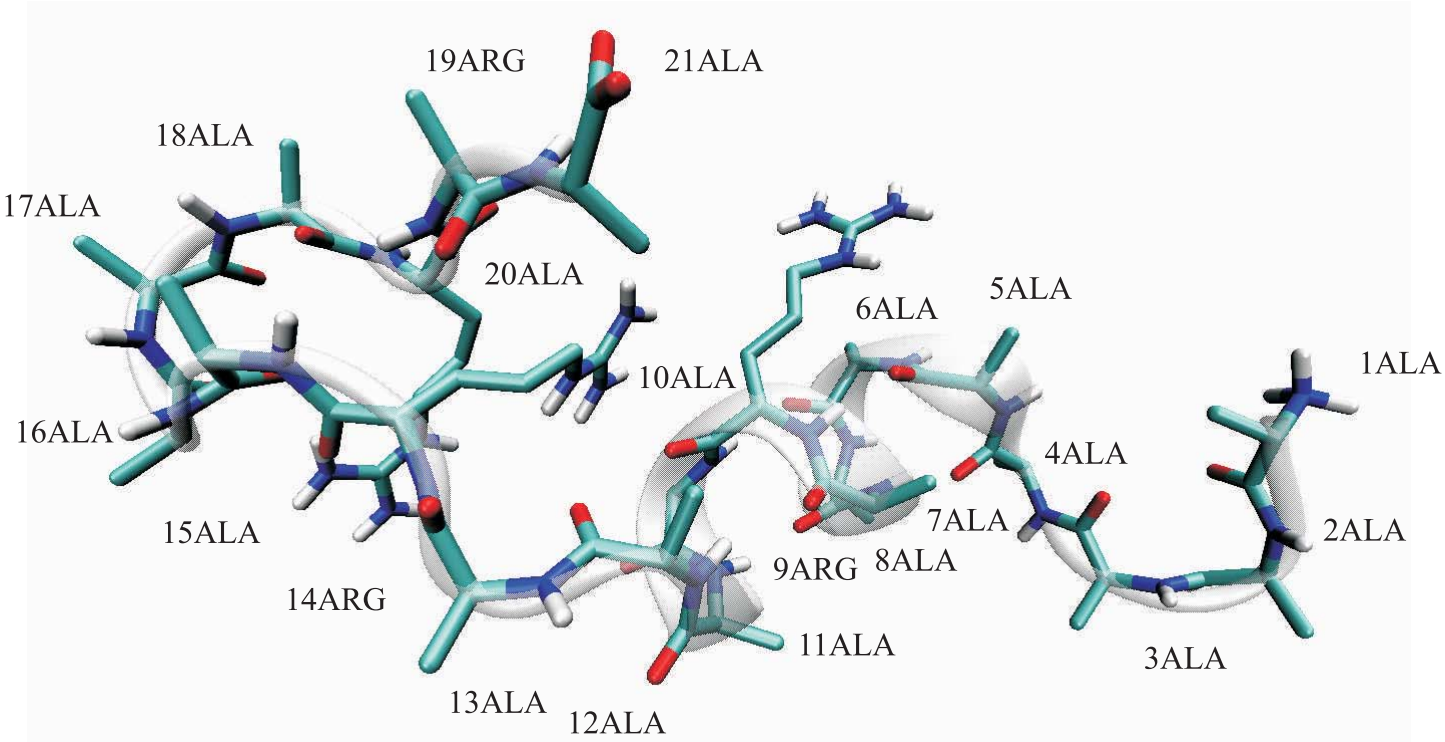
Water density correlation with the peptide motion



Intermediate summary

- From 10 to 1ps before the transition the water molecules tend to be located at more specific positions around the peptide compared to more uniform distribution at other times.
- During the transition the dynamics of water distribution becomes highly correlated with the dynamics of the dihedral angles.
- These correlations are completely absent during the stable conformation periods.

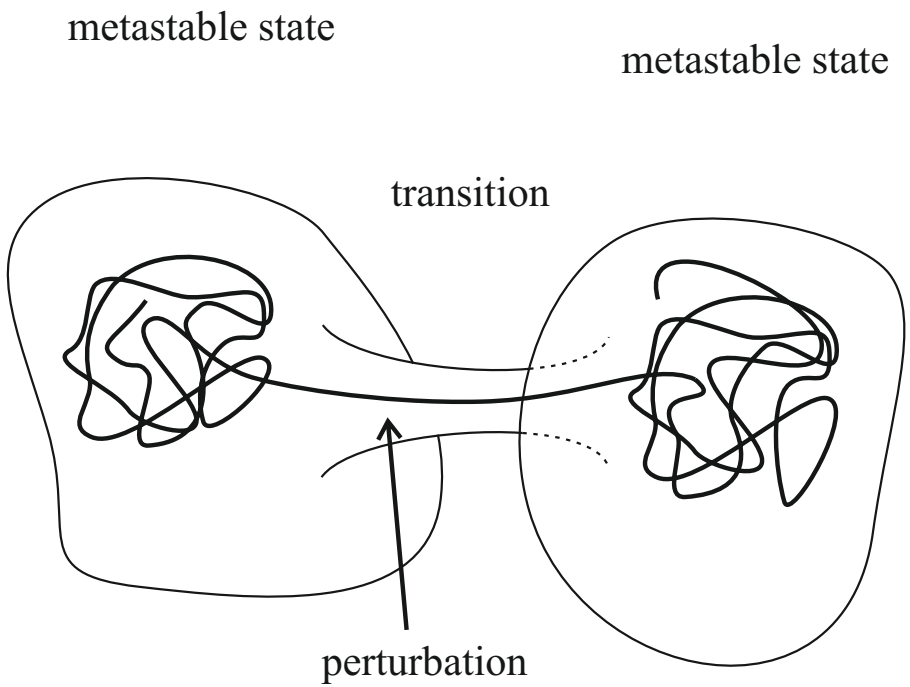
Trajectory perturbation



A5(A3RA)3A molecule

How close are the trajectories during transition

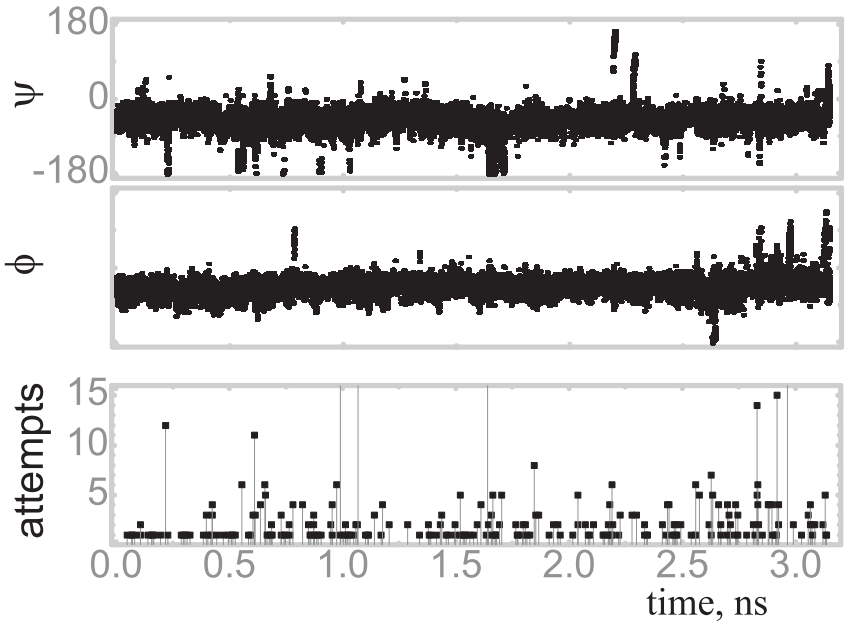
What happens if a trajectory is slightly perturbed just before the transition?



The perturbation is minute: a few degrees rotation of the velocity vector of one of the atoms.

How close are the trajectories during transition

The transition does not take place.



19ARG dihedral angles with perturbations; bottom: the number of 'attempts' needed to prevent the transitions

How close are the trajectories during transition

Perturbations at random times (even very frequent) do not prevent transitions.

Trajectory flow at the transition is very sensitive to perturbation.

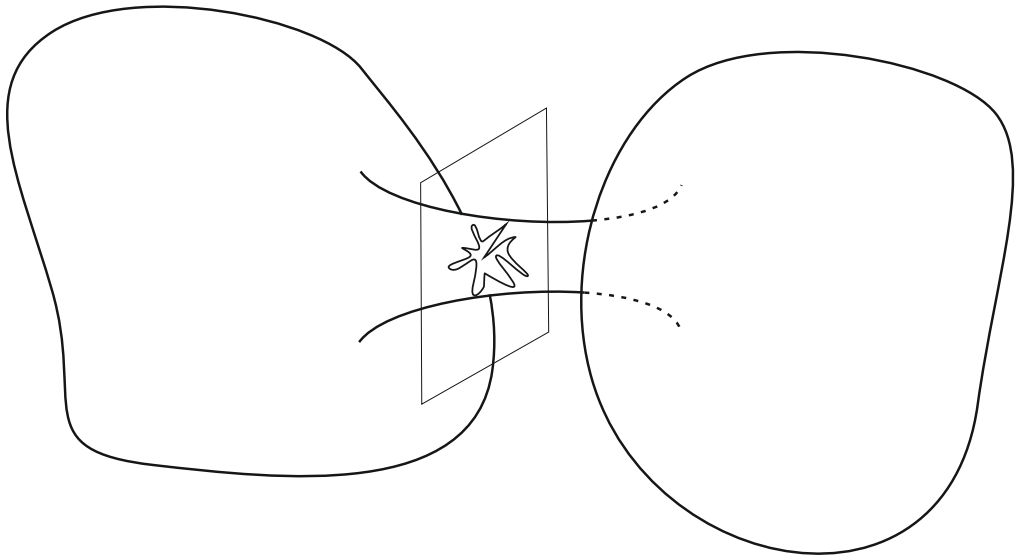
What degrees of freedom are most sensitive

Very surprisingly, it does not matter what atom is perturbed!

Even distant water atoms prevent transitions.

Why this is interesting

The shape of the 'cross-section' of the reactive bottleneck is extremely intricate.



Molecularly this means that during the conformational transition all molecules (including water) have to be carefully aligned. This is not a simple 'downhill' motion!

Conclusions

Conclusions

- Peptides and surrounding water molecules make an integral system for the process of conformational changes; water degrees of freedom must be explicitly included in the reaction coordinates.

Conclusions

- Peptides and surrounding water molecules make an integral system for the process of conformational changes; water degrees of freedom must be explicitly included in the reaction coordinates.
- The flow of the MD trajectories pass through a 'bottleneck' during the transitions, the dimensionality of the bottleneck is high and includes the surrounding water degrees of freedom.

Conclusions

- Peptides and surrounding water molecules make an integral system for the process of conformational changes; water degrees of freedom must be explicitly included in the reaction coordinates.
- The flow of the MD trajectories pass through a 'bottleneck' during the transitions, the dimensionality of the bottleneck is high and includes the surrounding water degrees of freedom.
- The bottleneck 'cross-section' is very complicated: a minute displacement is enough to 'step off' the reactive channel; the direction of displacement can be a distant water degree of freedom.

Collaborators

Funding: G8 Research Councils Initiative on Multilateral Research Funding - Exascale Computing

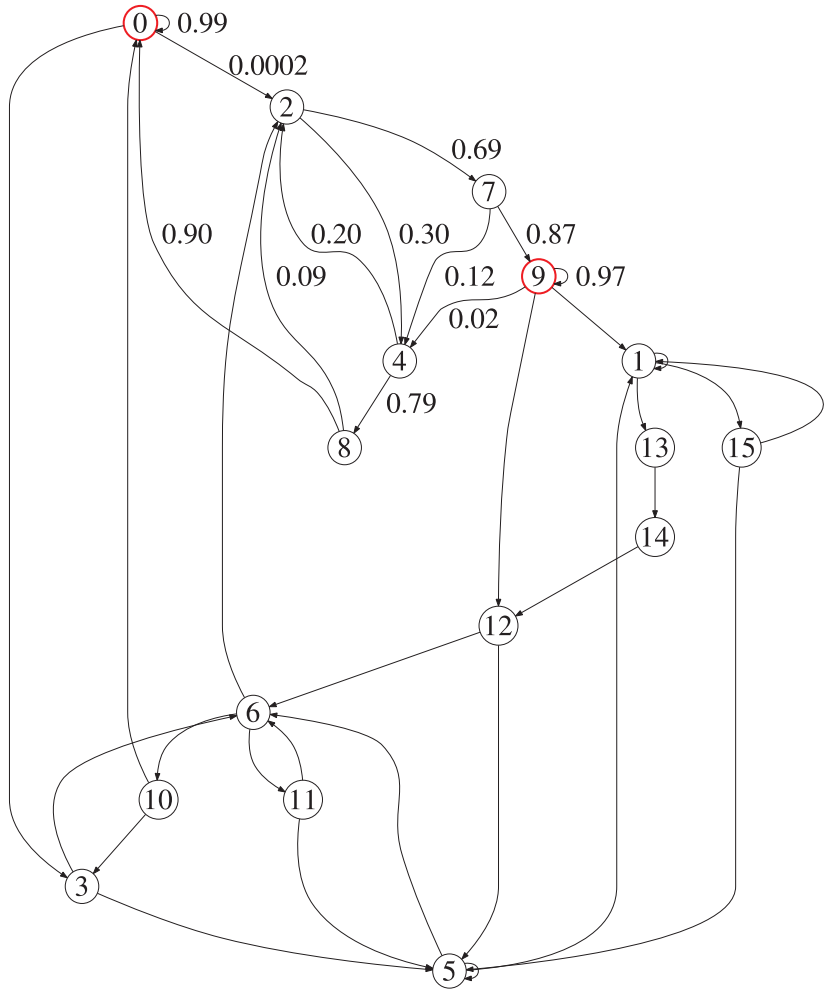
Sergey Karabasov
Cambridge University

Makoto Taiji
RIKEN Institute

Vladimir Ryabov
Hakodate University

Hidden Markov model

For the time step $0.3ps$ the states are Markov for $l \geq 4$.

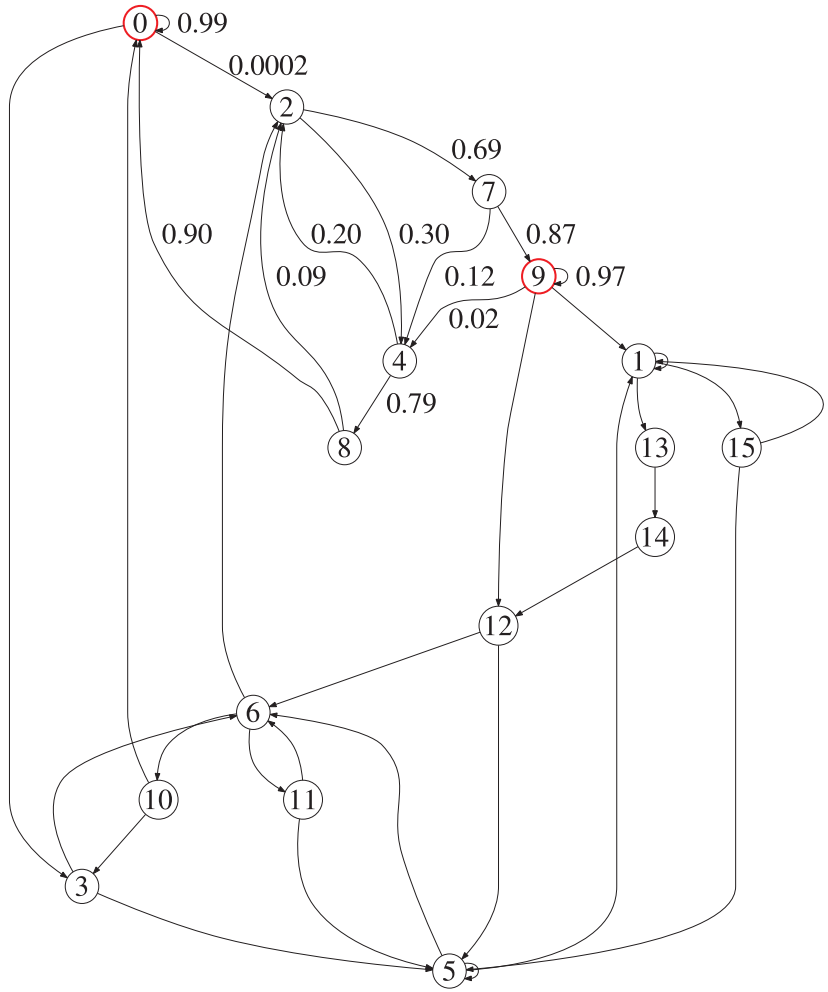


Hidden Markov model

For the time step $0.3ps$ the states are Markov for $l \geq 4$.

State '0' corresponds to the original state 'A': it consists of the sequences $AAAA$.

State '9' corresponds to 'B'.



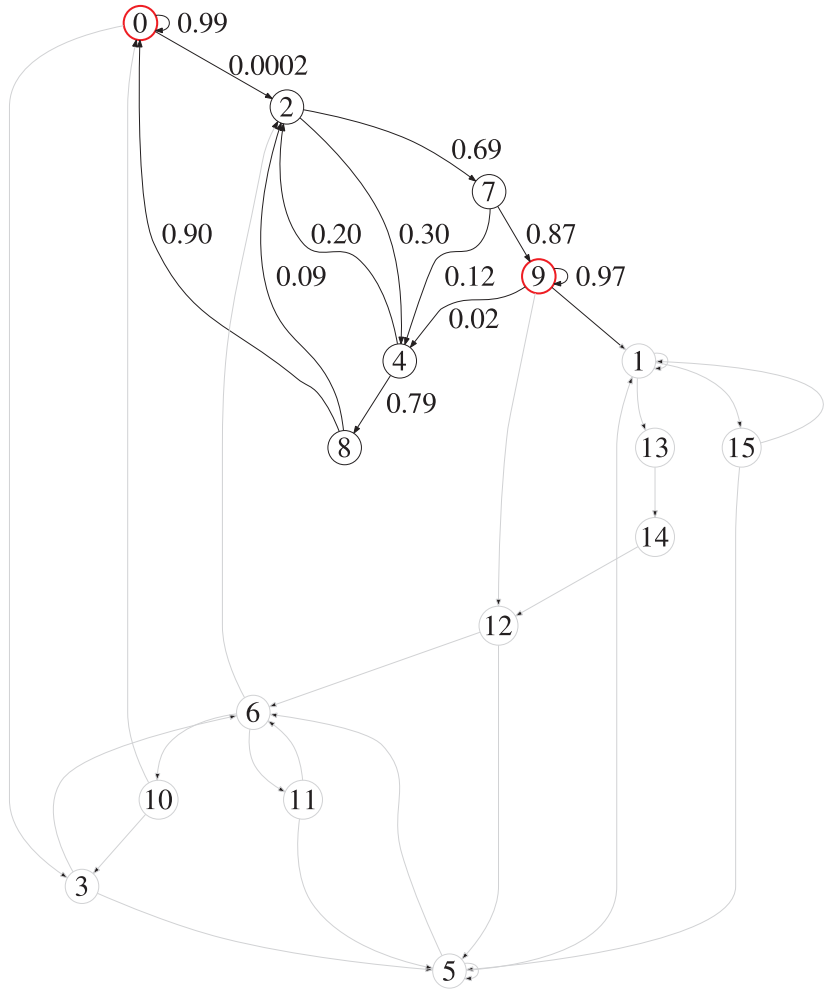
Hidden Markov model

For the time step $0.3ps$ the states are Markov for $l \geq 4$.

State '0' corresponds to the original state 'A': it consists of the sequences $AAAA$.

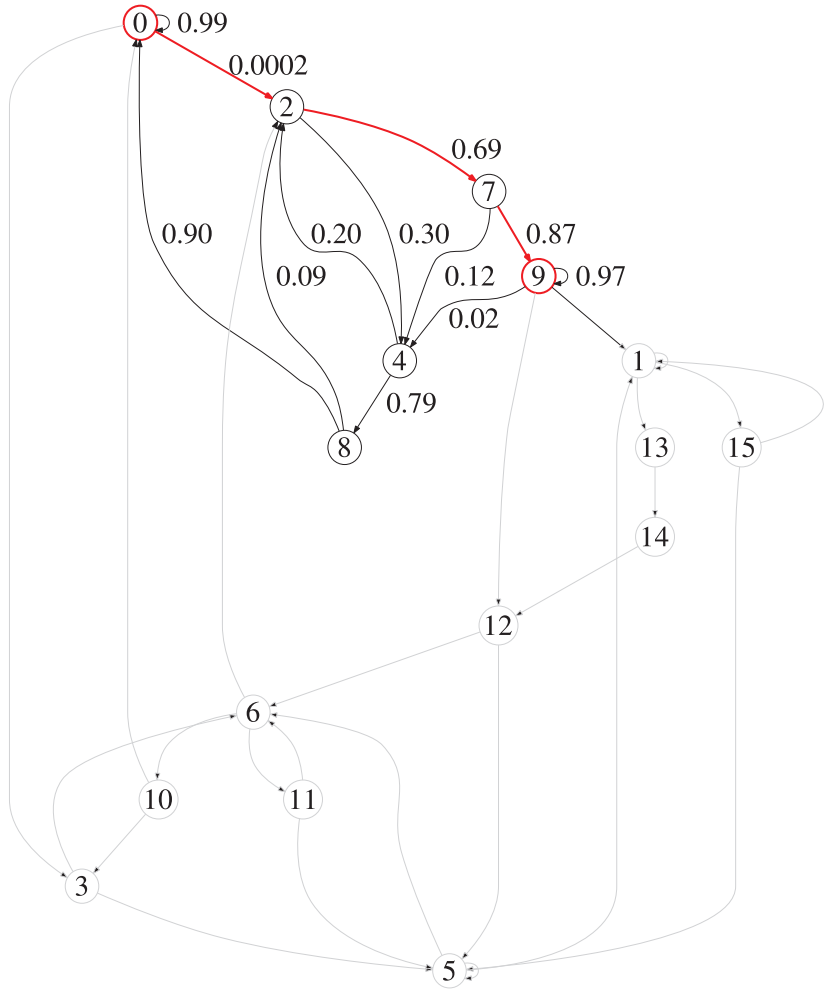
State '9' corresponds to 'B'.

The states '$2, 7, 4, 8$' are the transition states, they describe the mechanism (pathways) of the transition.



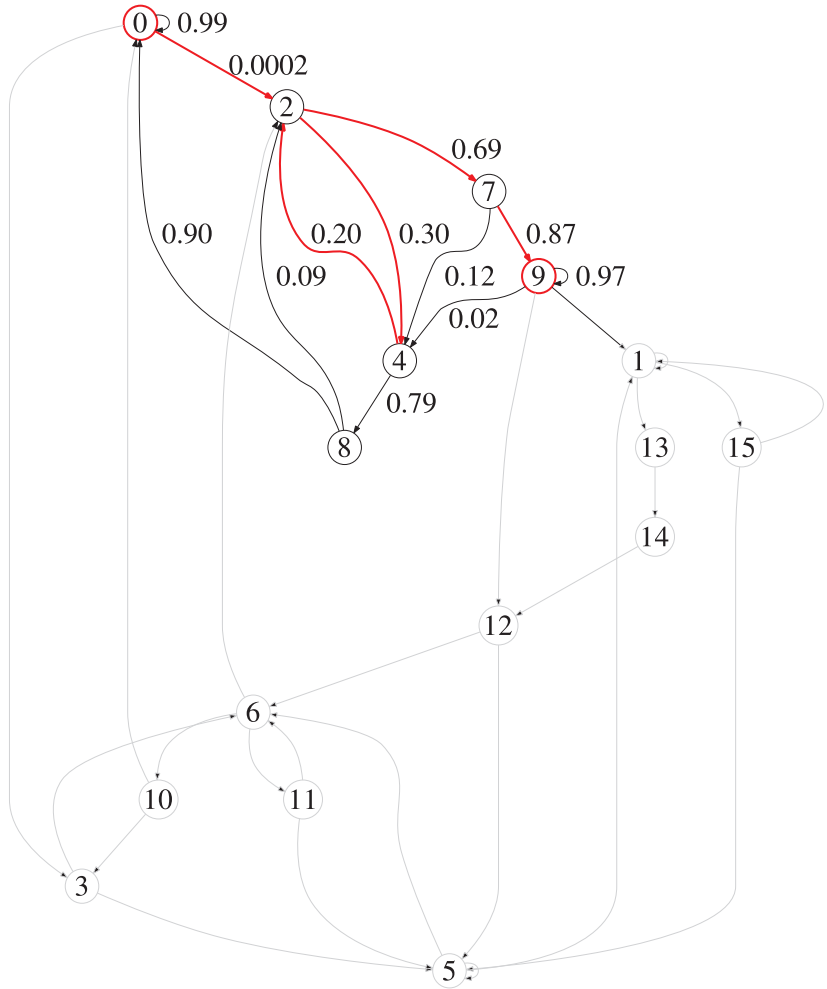
Hidden Markov model

Direct transition
(probability: 0.61)



Hidden Markov model

With several recrossings (probability: 0.04)



Hidden Markov model

Failed attempts of transition (probability: 0.07)