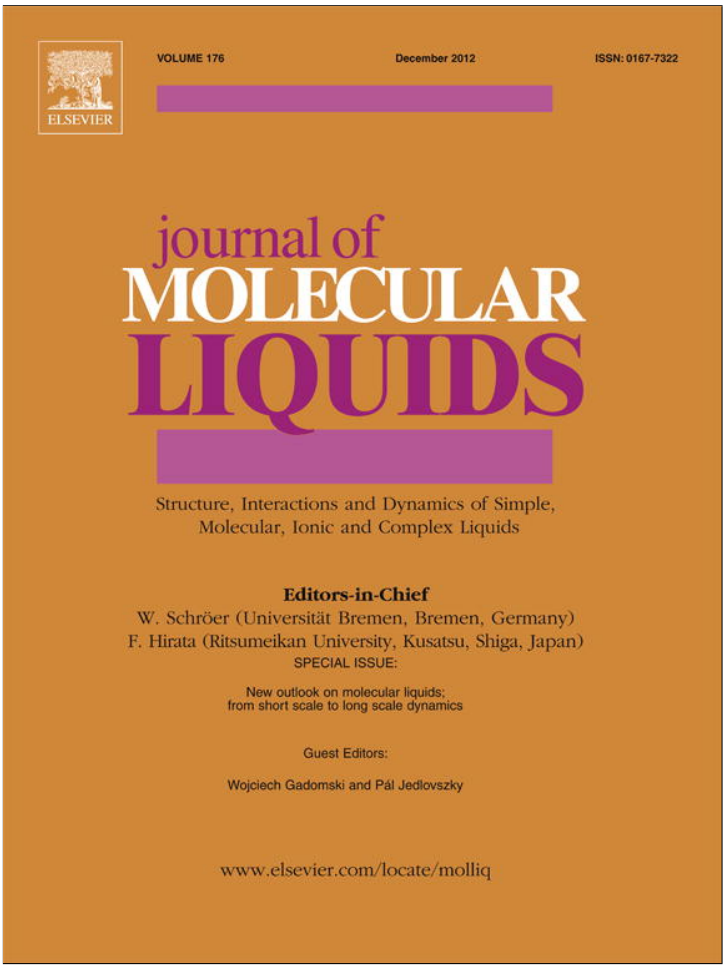

Journal of Molecular Liquids 176 (2012) 65–70

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Non-Markov state model of peptide dynamics

Dmitry Nerukh

Non-linearity and Complexity Research Group, Aston University, Birmingham, B4 7ET, UK

ARTICLE INFO

Available online 7 July 2012

Keywords:
Protein dynamics
Markov state model
Non-Markov
Conformational transitions

ABSTRACT

A hidden Markov state model has been applied to classical molecular dynamics simulated small peptide in explicit water. The methodology allows increasing the time resolution of the model and describe the dynamics with the precision of 0.3 ps (comparing to 6 ps for the standard methodology). It also permits the investigation of the mechanisms of transitions between the conformational states of the peptide. The detailed description of one of such transitions for the studied molecule is presented.

1. Introduction

Protein motion and function are defined by elementary conformational transitions described by the changes of the backbone dihedral angles of the biopolymer. Typically, the angles fluctuate around an averaged value attributed to 'metastable' conformational states and occasionally change the value significantly that correspond to conformational transitions.

The dynamics of such metastable states is often described using the increasingly popular Markov State Model (MSM) [1–7]. The essence of the model is in defining discrete protein conformations such that the transitions between them at discrete time moments are represented by a Markov process. As the mathematical apparatus of Markov processes is well developed, the model provides useful information about the time evolution of the protein system.

One of the active themes in the MSM research is the validation of the states definition. There is no rigorous description of how to divide the continuous space of the protein's dihedral angles into areas designating the Markov states. The important requirement of MSM is the Markov property of the states, that is the independence of the transitions from previous steps. This puts a conceptual restriction on the time scale of the model: the time step should be no less than the memory of the system.

In practice the states are either assigned to the areas roughly corresponding to the basins of attraction on the potential or free energy surface of the protein or constructed from the 'microstates'. The latter are defined using a fine grid covering the whole available dihedral space of the system. The microstates are clustered into 'macrostates' based on the time spent by the system in them [8]. The macrostates are the collections of microstates for which the transitions within a macrostate are more probable than between the macrostates. As the microstates should be 'short lived', that is the transitions within a microstate should be avoided, many of them are needed for correct reconstruction of the macrostates.

The shortcomings of MSM can be avoided by allowing the states to be non-Markov. A natural development of Markov processes theory is the hidden Markov processes framework. Here, the transitions can depend on the previous steps and these time dependent states are grouped into sequences of consecutive states such that the sequences themselves become Markov.

In this paper we develop a variant of the hidden Markov description of conformational transitions. We use the methodology called 'Computational Mechanics' (CM) developed by Crutchfield et al. [9–11]. We show that applying this approach to protein dynamics (which we term as non-Markov state model, nMSM) the detection of short lived 'transition' states is possible by reducing the time step by an order of magnitude compared to the usual MSM. By describing the transitions at such fine time scale, the detailed mechanisms of the transitions can be elucidated. In particular, the analysis of the elementary conformational transitions, the building blocks of all protein conformational motions becomes possible. We have applied Computational Mechanics to molecular systems in several contexts [12–14]. Here we will concentrate on direct description of conformational motions of a peptide.

As the mechanisms of the transitions can now be uncovered, the duration of the transitions themselves (not the times between the transitions, but the actual process of conformational rearrangements) can be quantified. This is usually impossible to obtain in MSM as the time of transitions is normally smaller than the time step of MSM. Also, the 'recrossings' can be identified and studied, when the system exhibits several quick crossings of the boundary separating the conformational states, the important problem recognised by the community for many years.

2. The molecular system

We study a zwitterion L-alanyl-L-alanine, Fig. 1, a very convenient model because i) the conformation of the molecule is completely

E-mail address: D.Nerukh@aston.ac.uk.

0167-7322/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.molliq.2012.06.011

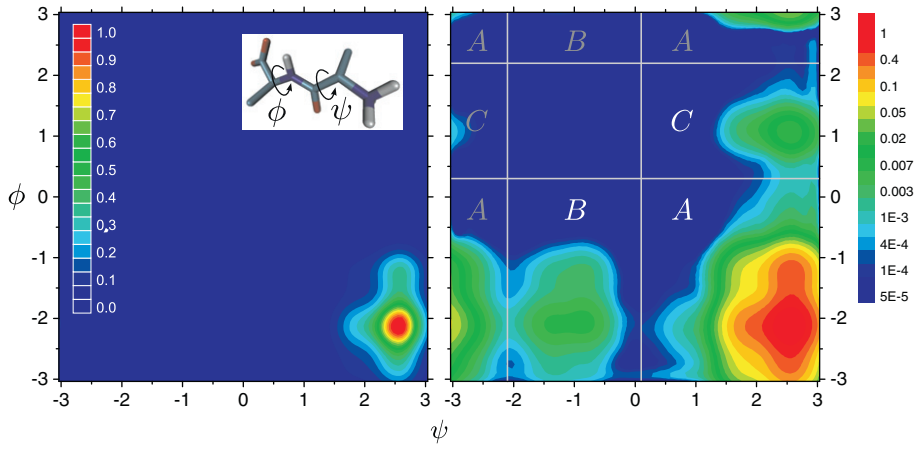

Fig. 1. Left: L-alanyl-L-alanine zwitterion and the normalised density of its conformations (Ramachandran plot) formed by a 1 μs trajectory; right: the same probabilities emphasising the presence of two minor conformations and the partitioning for symbolisation.

defined by the two dihedral angles ψ and ϕ, ii) in water the conformation $\psi \approx 2.5$, $\phi \approx -2.2$ radians is prevalent, however very rare transitions to two other metastable conformations take place, Fig. 1 right, and iii) the transitions only happen in water because the molecule's charged ends lock it in a 'loop' like conformation in vacuum.

The molecule is, thus, a good model of protein conformational changes at the same time being technically advantageous: the transitions are clearly defined and well separated in time.

The simulation was performed using the software package GROMACS [15]. The peptide was solvated by 396 water molecules. The simulation box was $25 \times 22 \times 22$Å. The force field was Gromos 53a6 [16–18]. This is optimized for bimolecular systems interacting with water. Electrostatic interactions were treated by Particle Mesh Ewald (PME) summation technique. Periodic boundary conditions were used. The temperature was kept at 300 K using the Berendsen thermostat [19]. The integration algorithm was a Verlet type and the integration step was 0.002 ps. The system was carefully equilibrated before it was sampled.

3. Methodology

The three metastable states, clearly visible on the density of states, Fig. 1, allow to introduce a simple natural discretisation of the conformational states. By also discretising time with a step Δt the continuous MD trajectory can be converted into a string of symbols $\{s_i\}, i = 0 \ldots N$, where s_i equals to '*A*', '*B*', or '*C*' depending on where the trajectory point falls at the time moment t_i, N is the number of time steps.

The analysis of $\{s_i\}$ can be done by building an MSM on them, or, as we do here, by applying a hidden Markov model.

Computational Mechanics builds a statistic on infinitely long histories ('pasts') of symbols s_i representing the state of the system at times t_i, $\overleftarrow{s_i} \equiv \{\ldots s_{i-2} s_{i-1} s_i\}$, by analysing the 'futures' $\overrightarrow{s_i} \equiv \{s_{i+1} s_{i+2} \ldots\}$ following each past. The focus is shifted from the physical conformational states s_i to their sequences $\overleftarrow{s_i}$ and $\overrightarrow{s_i}$ In this manner, the natural interdependencies of the conformational states are included in the picture, at the same time preserving the well developed mathematical apparatus of Markov processes.

The algorithm groups the pasts into classes called 'causal states' ϵ_j. The criterion of grouping is the equivalence of the probabilities of the futures, that is two pasts $\overleftarrow{s_i}$ and $\overleftarrow{s_j}$ are assigned to the same causal state if the distributions of their futures are the same:

$$P\left(\overrightarrow{s} \middle| \overleftarrow{s_i}\right) = P\left(\overrightarrow{s} \middle| \overleftarrow{s_j}\right), \tag{1}$$

where $P(X|Y)$ is the probability of X given Y. Thus, instead of the transitions between the states s_i the dynamics of the system is described by the probabilistic transitions between the causal states ϵ_i. Importantly, the sequence of ϵ_i is *Markovian by definition* regardless of the properties of the original process s_i. The collection of the causal states together with the transition probabilities between them is called an 'ϵ-machine'. The detailed definition of an ϵ-machine is provided in Appendix A. ϵ-machines can be reconstructed from observed data $\{s_i\}$ using the CSSR algorithm described and implemented in [20]. Computational Mechanics has been successfully applied to many real world physics systems. One of the most important example is the use of the ϵ-machine to analyse the structure of complex disordered materials [21,22].

Theoretically, CM is formulated using the assumption of infinitely long pasts and futures. In practice a finite history length l has to be chosen and this is one of the adjustable parameters of the CSSR algorithm (the length of the future is always chosen to be 1). The number of possible histories grows exponentially with the history length. Therefore, for long histories an exponential increase in the number of data points is also needed.

The second parameter of the CSSR algorithm is the significance level σ used in comparing the distributions $P\left(\overrightarrow{s} \middle| \overleftarrow{s_i}\right)$ for grouping the histories into causal states, Eq. (1) (the Kolmogorov–Smirnov test is used). Too large σ values (too strict threshold for two distributions to be considered equivalent) lead to artificially too many causal states. This is equivalent to under-sampling the histories. The same situation takes place for too long history length since the number of possible histories is too large and, for moderately long experimental time series, the distributions $P\left(\overrightarrow{s} \middle| \overleftarrow{s_i}\right)$ become not statistically significant.

Therefore, for obtaining the robust results, it appears necessary to perform the analysis of the ϵ-machine as a function of these two parameters. Too long a history or too large a σ value leads to statistically incorrect results. As the authors of CSSR recommend, the value of σ should be chosen such that there is a 'plateau' in the number of causal states as a function of l. If there are several such values of l then the lowest one has to be chosen (according to the minimality principle of CM). This constant value of l is the 'true' value of the history length for a stable ϵ-machine architecture.

The states of the ϵ-machine consist of the sequences of the conformational states. Therefore, strictly speaking it is incorrect to identify the states of the machine ϵ_i with the conformational states s_i. However, if the system spends long time in one of the conformational states, the ϵ-machine states at those times would prevalently consist of a sequence of repeating conformational states (for

example, for an '*A*' state, that would be $\overleftarrow{s_i} \equiv \ldots AAAAA$). Thus, an approximate correspondence between the metastable MSM and nMSM states is possible.

The most interesting moments are when the ϵ-machine generates short lived transition states. These describe the mechanisms of transition between the metastable conformational states.

4. Results

For the studied peptide, the minimal time step for which the conformational states are Markov is $\Delta t = 6$ ps. The dynamics is represented by a usual transition matrix, and a transition is when symbol '*A*' (Fig. 1) at time step i changes to symbol '*B*' at time step $i+1$. The matrix gives the probabilities of the transitions:

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

In this representation the $A \to B$ transition happens once in 3.1 ns on average. We will concentrate on this particular transition for the rest of the paper.

For the time step of $\Delta t = 5$ ps the dynamics becomes non-Markov, as the ϵ-machine demonstrates. The following matrix shows the number of reconstructed causal states ϵ_i as a function of the history length l and the significance level σ (see Section 3):

σ \ l	1	2	3	4	5	6	7	8	9
0.1000	3	6	6	6	6	6	6	6	25
0.0656	3	7	7	8	8	8	8	8	8
0.0430	3	4	11	13	14	14	18	24	57
0.0282	3	4	4	10	11	13	20	24	41
0.0185	3	4	4	4	4	4	4	4	4
0.0122	3	4	4	4	4	4	4	4	4
0.0080	3	4	4	4	4	4	4	4	4
0.0052	3	4	4	4	4	4	4	4	4
0.0034	3	4	4	4	4	4	4	4	4
0.0023	3	4	4	4	4	4	4	4	4
0.0015	3	4	4	4	4	4	4	4	4
0.0010	3	4	4	4	4	4	4	4	4

A clear 'plateau' is present with the minimum number of two steps necessary to reconstruct the ϵ-machine.

As each time moment is assigned a causal state (a collection of one or more sequences $\overleftarrow{s_i} \equiv \{\ldots s_{i-2}s_{i-1}s_i\}$), the conformations of the peptide corresponding to each causal state can be collected and analysed. The distribution of the conformations for each state together with the ϵ-machine is shown in Fig. 2.

State '0' in Fig. 2 corresponds to the conformational state '*A*' as it consists of the sequences '*AA*'. The distribution of the peptide conformations (Fig. 2, left) for state '0' coincides with the distribution located in the '*A*' area of Fig. 1. Similarly, states '3' and '1' correspond to the conformational states '*B*' and '*C*' respectively. The additional fourth state labelled '2' describes the transitions between '*A*' and '*B*' (sequences '*AB*' and '*BA*'). At this time resolution (5 ps) the time precision is not enough to distinguish between the forward and the reverse transitions, therefore, state '2' includes both $A \to B$ and $B \to A$ processes. The length $l = 2$ time steps also define the time scale at which the internal mechanism of the transitions becomes important: 5 ps.

By further reducing the time step more details are revealed. At $\Delta t = 2.5$ ps three steps become correlated and the minimal ϵ-machine consists of 7 causal states:

σ \ l	1	2	3	4	5	6	7	8	9
0.1000	4	11	11	12	12	12	12	12	54
0.0656	4	11	11	12	12	12	12	12	40
0.0430	4	11	11	12	12	12	12	12	40
0.0282	4	9	9	9	9	9	9	9	9
0.0185	3	9	9	9	9	9	9	9	9
0.0122	3	9	9	9	9	9	9	9	9
0.0080	3	8	7	7	7	7	7	7	7
0.0052	3	8	7	7	7	7	7	7	7
0.0034	3	8	7	7	7	7	7	7	7
0.0023	3	8	7	7	7	7	7	7	7
0.0015	3	6	8	7	7	16	23	34	29
0.0010	3	6	9	8	8	17	24	35	30

Now all transitions between the pairs of conformational states *A*–*B* and *A*–*C* are resolved (direct transitions between *B* and *C* are almost never realised). The same 5 ps time scale is present, however one additional step in the middle of the process is needed to describe the details.

The most detailed information is obtained for $\Delta t = 0.3$ ps. This is the minimal time step achievable for the amount of data, that is the number of transitions realised during 1 μs. For this at least 4 time steps are needed and the number of causal states reconstructed is 14:

σ \ l	1	2	3	4	5	6	7	8	9
0.1000	4	9	17	21	26	33	41	80	151
0.0656	4	9	16	17	22	27	28	76	134
0.0430	4	9	15	19	20	31	34	84	157
0.0282	4	9	16	19	19	24	31	87	172
0.0185	4	9	15	17	19	20	20	30	139
0.0122	4	9	15	14	14	14	14	31	143
0.0080	4	9	15	14	14	14	14	31	143
0.0052	4	9	15	14	14	14	14	31	143
0.0034	4	9	16	14	14	14	14	31	143
0.0023	4	9	17	18	19	21	29	53	139
0.0015	4	9	17	17	17	17	17	17	140
0.0010	4	9	13	13	13	13	13	13	69

Despite the many states found, the causal states corresponding to the metastable conformations can be similarly identified. State '0' corresponds to '*A*', while state '9' describes the metastable conformation '*B*'. The ϵ-machine very conveniently describes the details of the $A \to B$ transition: these are the causal states '2', '7', '4', and '8', Fig. 3.

The system enters into the $A \to B$ transition when it goes from the causal state '0' to the causal state '2' (going to state '3' from '0' corresponds to the $A \to C$ transition). From state '2' there are two possibilities: one leads to the transition to '9' corresponding to the actual $A \to B$ transition, the other eventually returns to state '0' and represents a failed transition. By assuming that going from '0' to either '9' or back to '0' has the probability of 1, the probabilities of different routes of the system are summarised in Table 1.

The number of returns is very significant: only 63% of the attempts are successful. Also, even thought the prevalent sequences are direct (for example 0279), the repetitions of the states are also non-negligible (for example 024279). These show that the recrossings are very important in the dynamics.

For this model, the transition $A \to B$ (that is the left column of Table 1) happens once in 2.4 ns on average, which is substantially smaller than the number provided by MSM (3.1 ns).

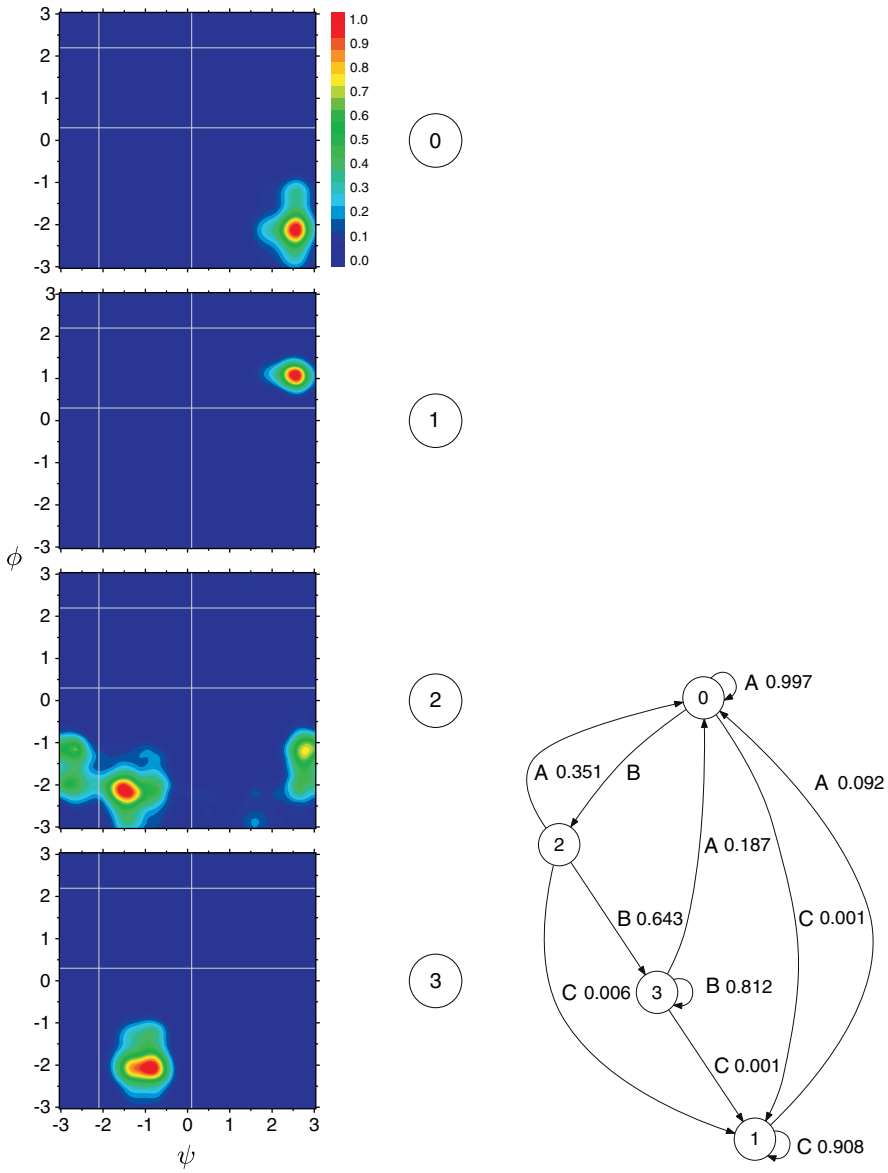

Fig. 2. Right: the ϵ-machine for the time step 5 ps, the length of the sequences is $l=2$, state '0' corresponds to the original state '*A*' (mostly consists of the sequences *AA*), similarly, state '3' corresponds to '*B*', state '2' describes the transitions between *A* and *B* (consists of the sequences *AB* and *BA*); left: the distribution of conformations corresponding to each causal state; the time moments at the last symbol of the histories are chosen for visualisation.

As the ϵ-machine provides a very detailed picture of transitions, the flow of the dihedral angles $\psi-\phi$ can be visualised. The two routes, corresponding to the left and right columns of Table 1 are shown in Fig. 4. The small time step allows a rather precise identification of the transition moment. Thus, different moments in advance of the transition can be shown. Interestingly, the dynamics seems to be very similar for both routes up to ≈ 3 ps before the transition. It then becomes different resulting in either going to conformation *B* or returning to *A*.

Overall, the mechanism shows that the direct transition is very fast, it takes only ≈ 4 time steps, that is 1.2 ps. The significance of this fact is that this time is substantially smaller than the minimally allowed time scale for the standard MSM. Therefore, important details can be overlooked, at least for small peptides like the studied one. The reason for the relatively large required time steps in MSM is in multiple recrossings that are masked by the larger time steps and thus, made the dynamics Markovian. This provides another evidence of the usefulness of nMSM in studying peptide dynamics.

5. Conclusions

We present a development of Markov state method that incorporates non-Markov behaviour of the conformational states of the protein. The method avoids several shortcomings inherent to MSM. In particular, there is no need for many microstates for identifying correct metastable states. The dynamics of the system finds the states automatically.

Most importantly, the methodology provides a possibility of reducing the time step, thus, increasing the time resolution of the model. In principle, the resolution is only restricted by the amount of available data for the analysis, whereas in the standard MSM the time resolution restriction is conceptual.

Our method provides a detailed description of the mechanisms of the transitions. It identifies the importance of multiple recrossing and quantifies their probabilities. As the mechanisms of the transitions can now be elucidated with high time resolution, a detailed molecular picture of conformational motions can be obtained.

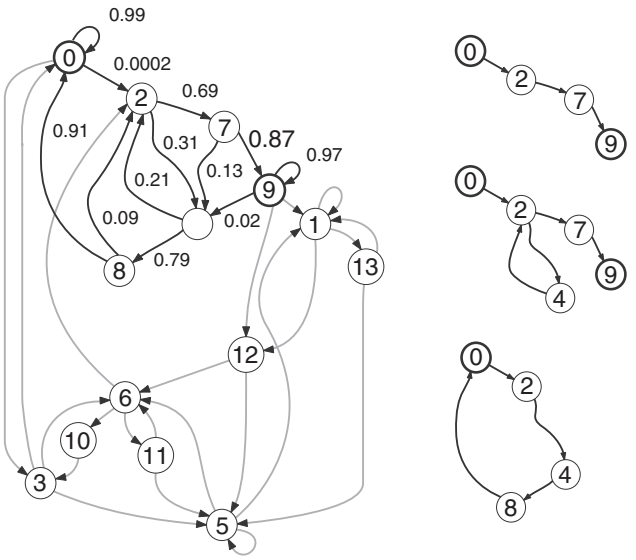

Fig. 3. Left: the ϵ-machine for the time step 0.3 ps, the length of the sequences is $l=4$, state '0' corresponds to the original state '*A*' (mostly consists of the sequences *AAAA*), similarly, state '9' corresponds to '*B*', state '9' can only be reached from state '0' *via* the states '2', '7', '4', '8' that describe the mechanism (pathways) of the $A \rightarrow B$ transition; right: three typical cases during the transition: direct transition (top, probability 0.56), transition with several recrossings (middle, probability 0.03), failed attempt of transition (bottom, probability 0.27), the probabilities of these cases are given assuming that the probability of going from state '0' or to state '9' by any possible route is 1.

The mechanisms of the transitions for this particular peptide shows that there is, probably, other important degrees of freedom that need to be included in the picture. We hypothesise that such degrees of freedom should include water molecules and this is the subject of our subsequent publications.

Appendix A. Computational mechanics

All past $\overleftarrow{s_i}$ and future $\overrightarrow{s_i}$ halves of bi-infinite symbolic sequences centred at times i are considered. Two pasts $\overleftarrow{s_1}$ and $\overleftarrow{s_2}$ are defined equivalent if the conditional distributions over their futures $P\left(\overrightarrow{s}\middle|\overleftarrow{s_1}\right)$ and $P\left(\overrightarrow{s}\middle|\overleftarrow{s_2}\right)$ are equal. A *causal state* $\epsilon\left(\overleftarrow{s_i}\right)$ is a set of all pasts equivalent to $\overleftarrow{s_i}$: $\epsilon_i \equiv \epsilon\left(\overleftarrow{s_i}\right) = \left\{\lambda : P\left(\overrightarrow{s}\middle|\lambda\right) = P\left(\overrightarrow{s}\middle|\overleftarrow{s_i}\right)\right\}$. At a given moment the system is at one of the causal states, and moves to the next one with the probability given by the transition matrix $T_{ij} \equiv P(\epsilon_j|\epsilon_i)$. The transition matrix determines the asymptotic causal state probabilities as its left eigenvector $P(\epsilon_i)T = P(\epsilon_i)$, where $\sum_i P(\epsilon_i) = 1$. The collection of the causal states together with the transition probabilities define an ϵ-machine.

It is proven [23] that the ϵ-machine is

- a *sufficient* statistic, that is it contains the complete statistical information about the data;
- a *minimal sufficient* statistic, therefore the causal states cannot be subdivided into smaller states;
- a *unique minimal sufficient* statistic, any other one simply re-labels the same states.

Table 1
The sequences of causal states describing two routes realised when leaving state '0' and their probabilities, see Fig. **3**.

$A \rightarrow B$		$A \rightarrow A$	
Sequence	Probability	Sequence	Probability
0279	0.56	02480	0.27
024279	0.03	027480	0.08
0248279	0.02	0242480	0.01
0274279	0.009	02482480	0.007
02748279	0.006	02742480	0.004
02482424279	0.001	02427480	0.003
Total:	0.63	0242482480	0.001
		Total:	0.37

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.molliq.2012.06.011.

References

[1] C. Schuette, A. Fischer, W. Huisinga, P. Deuflhard, Journal of Computational Physics 151 (1) (1999) 146–168, http://dx.doi.org/10.1006/jcph.1999.6231.
[2] W.C. Swope, J.W. Pitera, F. Suits, The Journal of Physical Chemistry. B 108 (21) (2004) 6571–6581, http://dx.doi.org/10.1021/jp037421y arXiv:http://pubs.acs.org/doi/pdf/10.1021/jp037421y, URL http://pubs.acs.org/doi/abs/10.1021/jp037421y.
[3] W.C. Swope, J.W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B.G. Fitch, R.S. Germain, A. Rayshubski, T.J.C. Ward, Y. Zhestkov, R. Zhou, The Journal of Physical Chemistry. B 108 (21) (2004) 6582–6594, http://dx.doi.org/10.1021/jp037422q (arXiv:http://pubs.acs.org/doi/pdf/10.1021/jp037422q, URL http://pubs.acs.org/doi/abs/10.1021/jp037422q).
[4] F. Noe, I. Horenko, C. Schuette, J.C. Smith, The Journal of Chemical Physics 126 (15) (2007) 155102, http://dx.doi.org/10.1063/1.2714539 (URL http://link.aip.org/link/?JCP/126/155102/1).
[5] J.D. Chodera, N. Singhal, V.S. Pande, K.A. Dill, W.C. Swope, The Journal of Chemical Physics 126 (15) (2007) 155101, http://dx.doi.org/10.1063/1.2714538 (URL http://link.aip.org/link/?JCP/126/155101/1).
[6] S. Park, V.S. Pande, The Journal of Chemical Physics 124 (5) (2006) 054118, http://dx.doi.org/10.1063/1.2166393 (URL http://link.aip.org/link/?JCP/124/054118/1).
[7] J.D. Chodera, W.C. Swope, J.W. Pitera, K.A. Dill, Multiscale Modeling and Simulation 5 (4) (2006) 1214–1226, http://dx.doi.org/10.1137/06065146X (URL http://link.aip.org/link/?MMS/5/1214/1).
[8] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J.D. Chodera, C. Schutte, F. Noe, The Journal of Chemical Physics 134 (17) (2011) 174105, http://dx.doi.org/10.1063/1.3565032 (URL http://link.aip.org/link/?JCP/134/174105/1).
[9] J.P. Crutchfield, K. Young, Physical Review Letters 63 (2) (1989) 105–108.
[10] J.P. Crutchfield, K. Young, In: in: W. Zurek (Ed.), Entropy, Complexity, and Physics of Information, SFI Studies in the Sciences of Complexity, VIII, Addison-Wesley, Reading, Massachusetts, 1990.
[11] J.P. Crutchfield, Physica D 75 (1–3) (1994) 11–54.
[12] V. Ryabov, D. Nerukh, Chaos: An Interdisciplinary Journal of Nonlinear Science 21 (3) (2011) 037113, http://dx.doi.org/10.1063/1.3608125 (URL http://link.aip.org/link/?CHA/21/037113/1).
[13] D. Nerukh, C.H. Jensen, R.C. Glen, The Journal of Chemical Physics 132 (8) (2010) 084104, http://dx.doi.org/10.1063/1.3328781 (URL http://link.aip.org/link/?JCP/132/084104/1).
[14] D. Nerukh, V. Ryabov, M. Taiji, Physica A: Statistical Mechanics and its Applications 388 (22) (2009) 4719–4726, http://dx.doi.org/10.1016/j.physa.2009.07.041 (URL http://www.sciencedirect.com/science/article/B6TVG-4WXPXVC-1/2/02ea8773db9afced932b98cdbfc6cb70).
[15] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, Journal of Computational Chemistry 26 (2005) 1701–1718.
[16] B. Hess, N.F.A. van der Vegt, The Journal of Physical Chemistry. B 110 (35) (2006) 17616–17626.
[17] C. Oostenbrink, T.A. Soares, N.F.A. van der Vegt, W.F. van Gunsteren, European Biophysics Journal with Biophysics Letters 34 (4) (2005) 273–284.
[18] C. Oostenbrink, A. Villa, A.E. Mark, W.F. Van Gunsteren, Journal of Computational Chemistry 25 (13) (2004) 1656–1676.
[19] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, The Journal of Chemical Physics 81 (8) (1984) 3684–3690.
[20] C.R. Shalizi, K.L. Shalizi, In: in: M. Chickering, J. Halpern (Eds.), Uncertainty in Artificial Intelligence: Proceedings of the Twentieth Conference, AUAI Press, Arlington, Virginia, 2004, pp. 504–511, URL http://arxiv.org/abs/cs.LG/0406011.
[21] D.P. Varn, G.S. Canright, J.P. Crutchfield, Physical Review B 66 (2002) 174110, http://dx.doi.org/10.1103/PhysRevB.66.174110 (URL http://link.aps.org/doi/10.1103/PhysRevB.66.174110).

[22] D.P. Varn, G.S. Canright, J.P. Crutchfield, Acta Crystallographica. Section B 63 (2) (2007) 169–182, http://dx.doi.org/10.1107/S0108768106043084 (URL http://dx.doi.org/10.1107/S0108768106043084).

[23] C. Shalizi, K. Shalizi, R. Haslinger, Physical Review Letters 93 (11) (2004) 118701-1–118701-4.

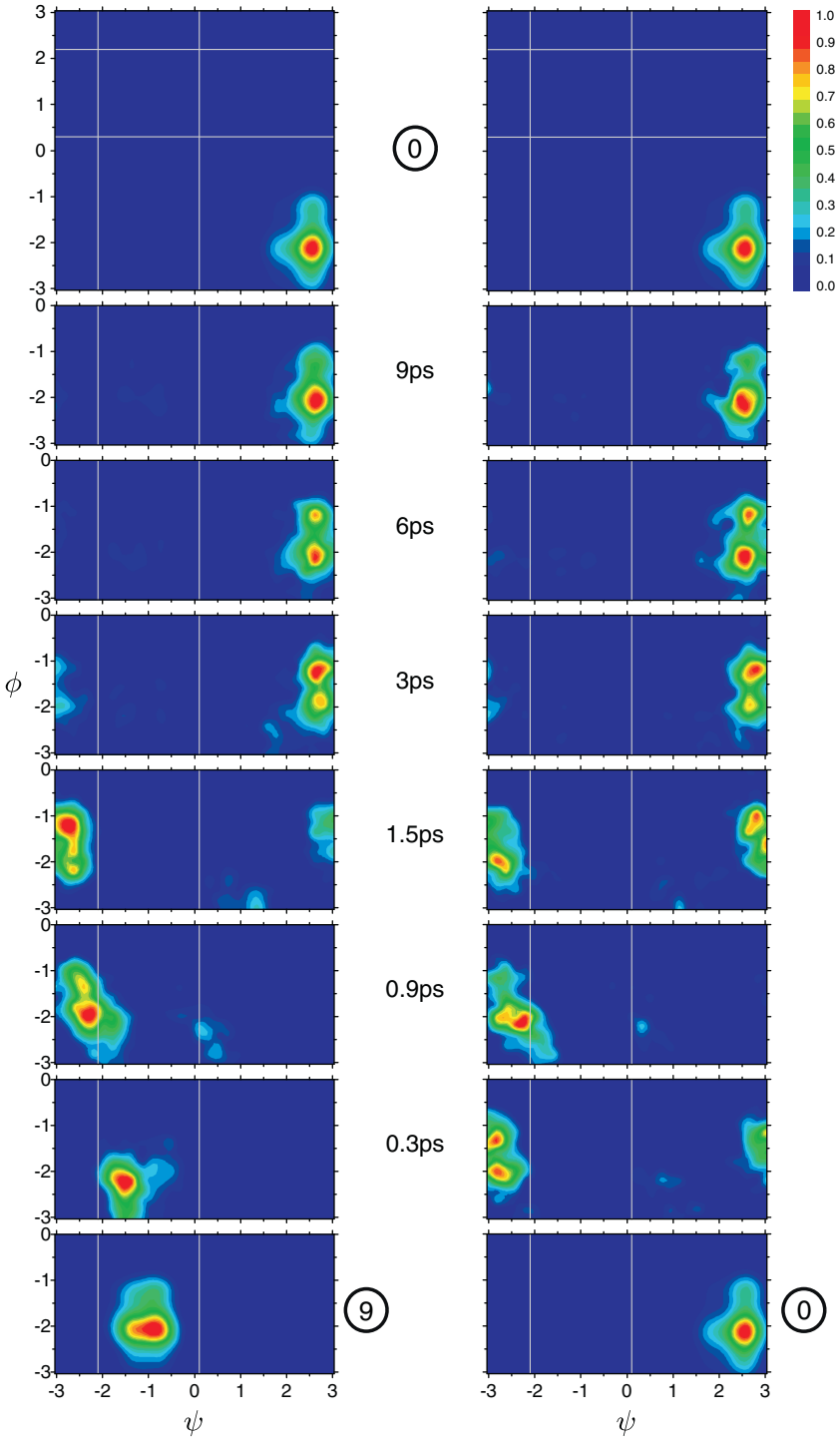

Fig. 4. The flow of the dihedral angles describing two different possibilities of going from state '0': left, the $A \rightarrow B$ transition, right, failed transition and return to state A; the numbers show the time before the transition is complete (that is when the system reaches either state '0' or state '9'); the time moments at the last symbol of the histories are chosen for visualisation.