
Metastable conformational structure and dynamics: Peptides between gas phase and aqueous solution

Eike Meerbach, Christof Schütte, Illia Horenko, and Burkhard Schmidt

Freie Universität, Institute of Mathematics II, Arnimallee 6, D-14195 Berlin, Germany, burkhard@math.fu-berlin.de

1 Introduction

The dynamics of biomolecular systems is characterized by the existence of biomolecular conformations which can be understood as metastable geometrical large scale structures [1]. On the longest time scales, biomolecular dynamics is dominated by flipping processes between these conformations, while on shorter time scales, the dynamical behavior is governed by flexibility within these conformations, resulting in a rich temporal multiscale structure of time-dependent observables. The purpose of this chapter is to elucidate on the possibility of constructing reduced models reflecting the "effective dynamics".

It is a promising idea to describe the effective dynamics of a biomolecular systems by means of a Markov chain with discrete states D_1, \dots, D_m , representing the metastable conformations, and a transition matrix $P = (p_{kj})$, describing the "flipping dynamics" between these states. The problem of efficient algorithmic identification of the metastable conformations is a challenging problem, recently there have been several *set-oriented* approaches to this problem [2–4]. In the context of the present work Hidden Markov Models (HMM) are used to extract the effective dynamics between *hidden* metastable molecule conformations from *observable* time series. e.g. the torsional angles of the backbone of biopolymers obtained by MD simulations [5]. In addition, the flexibility within conformations can be modelled by stochastic differential equations (SDE), thus comprising the HMMSDE model [6–8]. As the description of internal flexibility by SDEs also accounts for relaxation from one metastable conformation to another, this approach narrows the gap between "flipping dynamics" and transition path computation, as described in, e.g., [9].

We will herein first explain the background of the (set-oriented) transfer operator approach and the HMMSDE scheme. In the second part of this contribution we will discuss application to conformation dynamics of a prototypical dipeptide, spanning the range from gas phase to aqueous solution.

2 Metastability and the transfer operator approach

In the following we shortly summarize the algorithmic idea of the transfer operator approach, omitting most of the theoretical background. Instead we concentrate on the question how to set up an effective dynamics from a given time series, e.g., trajectory data. The reader interested in a mathematically more rigorous description is referred to [3, 10–13], readers not familiar with the basic notations of Markov chain theory are referred to [14]. First we explicate the concept of metastability of a Markov chain and the key idea for the identification of metastable states. Note that the transfer operator is an object in continuous state space, while we present the concept on discrete state space. Therefore only the discretized equivalent of the transfer operator, the transition matrix, appears in the following.

Consider a Markov chain $\{X_k\}_{k \in \mathbb{N}}$ on a discrete state space $\mathbf{X} = \{1, 2, \dots, n\}$ specified by a stochastic transition matrix $P = (p_{kj})$, with

$$p_{kj} = \mathbb{P}[X_{l+1} = j | X_l = k],$$

denoting the *conditional probability* to jump from k to j within one timestep. Furthermore, assume that the Markov chain is irreducible, aperiodic and reversible, i.e. a unique and strictly positive stationary distribution $\pi = (\pi_k)$ exists with $\pi_k p_{kj} = \pi_j p_{jk}$ for all $k, j \in \mathbf{X}$. A subset $B \subset \mathbf{X}$ is called *metastable* if

$$\mathbb{P}[X_{l+1} \in B | X_l \in B] \approx 1,$$

i.e., if the process is in subset B it is very likely to stay there within the next time step.

A decomposition $\mathbf{d} = \{D_1, \dots, D_m\}$ of the state space \mathbf{X} is defined as a collection of disjoint subsets $D_k \subset \mathbf{X}$ covering \mathbf{X} , i.e. $\cup_{k=1}^m D_k = \mathbf{X}$. The *metastability of a decomposition* \mathbf{d} is defined as the sum of the metastabilities of its subsets, i.e. for each arbitrary decomposition \mathbf{d}_m of the state space \mathbf{X} into m sets its metastability measure is defined as

$$M(\mathbf{d}_m) = \sum_{j=1}^m \mathbb{P}[X_{l+1} \in D_m | X_l \in D_m].$$

For given m , the *optimal metastable decomposition* into m sets maximizes the functional M . In particular the appropriate number m of metastable subsets must be identified. Both the determination of m and the identification of the metastable subsets can be achieved via spectral analysis of the transition matrix P , as the following holds

Due to reversibility, all eigenvalues of the transition matrix P are real. Metastable subsets can be detected via eigenvalues close to the maximal dominant eigenvalue $\lambda = 1$, i.e., the number of metastable subsets in the metastable decomposition is equal to the number of

eigenvalues close to 1, including $\lambda = 1$ and accounting for multiplicity, while the rest of the spectrum is separated through a *spectral gap* from 1. Among other possibilities, the sign structure of the eigenfunctions allows the identification of the metastable subsets [3, 12, 15].

Therefore the road map to determine metastable states on basis of a time series reads as follows

1. Discretize the state space of the time series and extract a transition matrix by counting transitions between the discrete states.
2. Use the spectral properties of the transition matrix to obtain metastable sets, yielding a coarse-grained description.

There are two remarks to be made on this road map.

First: Discretizing the state space is a non-trivial task, as typical biomolecular systems contain hundreds or thousands of degrees of freedom. Fortunately, chemical observations reveal that—even for larger biomolecules—the curse of dimensionality can be circumvented by exploiting the hierarchical structure of the dynamical and statistical properties of biomolecular systems: only relatively few *essential degrees of freedom* may be needed to describe the conformational transitions.

Second: After discretizing the state space there is a choice in the *lag time* τ used to obtain the transition matrix. If τ^* is the timestep between subsequent data points then the lag time τ can be set to $r\tau^*$ by evaluating transitions from every k th sampled step to every $(k+r)$ th, $r \geq 1$, sampled step. Taking $r > 1$ corresponds to a coarser discretization of the time domain of the originally continuous dynamics. Different values of r give rise to different transition matrices. *Therefore subsets of the state space are metastable with respect to a certain timescale.* By choosing r sufficiently large one can decrease correlations between subsequent timesteps and therefore ensure that the Markov description is a proper description.

2.1 Illustrative Example

We give a short and simplistic example to highlight the procedure outlined above. Consider the one dimensional time series $(Y_t)_{t=t_1, \dots, t_N}$, with constant sampling time $\tau^* = t_{j+1} - t_j$, shown in Fig. 1, which clearly exhibits metastable behavior. We discretize the state space $[-180 \ 180]$ into 9 equidistant boxes, the numbering of the boxes randomly chosen. If $N_r(j, k)$ denotes the number of transitions from box j to box k in r steps and $N_r(k)$ the number of data points in box k , we obtain a reversible transition matrix $P = (p_{kj})$, with respect to the timelag $\tau = r * \tau^*$, by setting

$$p_{kj} = \frac{N_r(j, k) + N_r(k, j)}{N_r(j) + N_r(k)}. \quad (1)$$

The obtained matrix seems to exhibit no special structure, but computing the spectrum, for $r = 1$, yields

$$\sigma(P) = \{1, 0.98, 0.55, 0.34, \dots\},$$

indicating two metastable states. The information contained in the eigenvector belonging to the second eigenvalue is used to identify the metastable subsets, i.e. boxes with the same eigenvector sign are assigned to the same metastable state. Permuting the matrix such that boxes belonging to the same metastable set are neighbors, results in a dominantly blockdiagonal structure. Aggregating the states in each metastable set results in a two state "effective dynamics" with transition matrix

$$\begin{pmatrix} 0.989 & 0.011 \\ 0.013 & 0.987 \end{pmatrix},$$

and a stationary distribution $\pi = (0.56 \ 0.44)^T$.

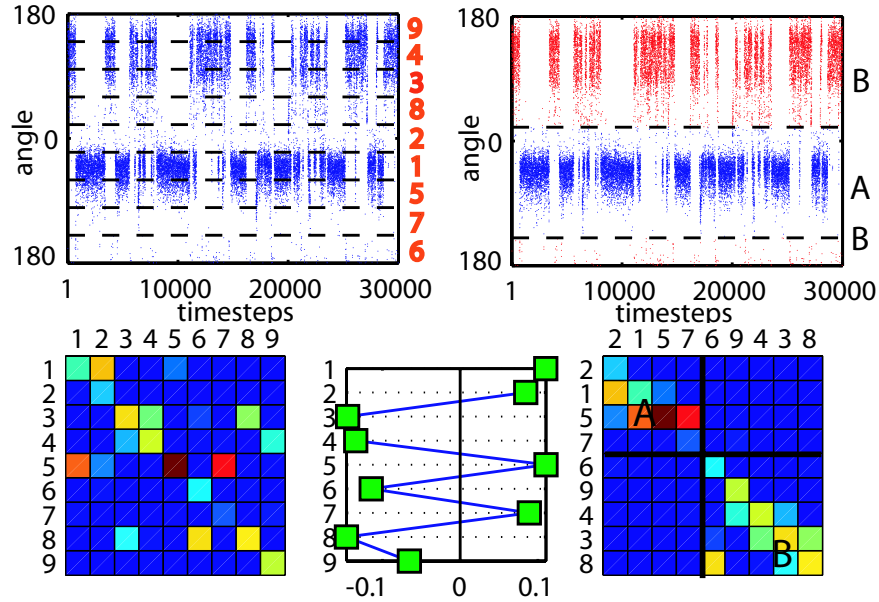


Fig. 1. *Top left:* A time series of circular data. The equidistant space discretization in 9 randomly numbered boxes is indicated with dashed lines. *Bottom left:* The obtained stochastic transition matrix, blue (dark) color represents entries near zero, while red (light) entries are corresponding to entries close to one. *Bottom middle:* The sign structure of the second eigenvector allows assignment to metastable states. *Bottom right:* The matrix permuted according to the eigenvector structure exhibits a block structure. *Top right:* Aggregating the discretization boxes belonging to the same metastable state yields a two state model.

3 The Hidden Markov Model Approach

Assume that we extracted a time series Y_t from, e.g., MD-simulation, which do not necessarily completely specify the state of the molecule at time t , but rather some low-dimensional observable, for example, some or all torsion angles or a set of essential degrees of freedom. As the Markov property does not hold for projections of Markov processes in general, we have to be aware that the process on the (torsion angles) subspace might no longer be Markovian. Nevertheless, we assume that there is an unknown metastable decomposition into m sets D_1, \dots, D_m , in the full dimensional system. We then can premise that, at any time t , the system is in one of the metastable states D_{j_t} to which we simply refer by j_t in the following. However, the time series (j_t) is *hidden*, i.e., neither known in advance nor observed, while the series (Y_t) is called the output series or the *observed* sequence.

This scenario can be represented by a Hidden Markov Model (HMM). A HMM abstractly consists of two related stochastic processes: a hidden process j , that fulfills the Markov property, and an observed process Y_t , that depends on the state of the hidden process j_t at time t . A HMM is fully specified by the initial distribution μ , the transition matrix P of the hidden Markov process j , a rate matrix in continuous time, and the probability distributions that govern the observable Y_t depending on the respective hidden state j_t .

In the standard versions of HMMs the observables are assumed to be identically and independent distributed (i.i.d.) random variables with stationary distributions that depend on the respective hidden states [5]. Within the scope of molecular dynamics this setting corresponds to the simple case where the timelag τ is comparable with the relaxation times within the metastable states, while the relaxation times are sufficiently smaller than the mean exit times of the metastable states. In other words one expects the process to sample the restricted invariant density before exiting from a metastable state, and the sampling time of the time series is long enough to assume statistical independence between steps. Nevertheless, if this is not the case, only a slight modification of the model structure is required to include the relaxation behavior: Instead of i.i.d. random variables, an Ornstein-Uhlenbeck (OU) process serves as a model for the output behavior in each hidden state. The HMM then takes the form [6]:

$$\dot{Y}_t = -\nabla V^{(j_t)}(Y_t) + \sigma^{(j_t)} \dot{W}_t, \quad (2)$$

$$j_t : \mathbf{R}^1 \rightarrow \{1, 2, \dots, m\}, \quad (3)$$

where j_t are the realizations of the hidden Markov process with discrete state space, W_t is standard "white noise", and $\{V^{(j)}, \sigma^{(j)}\}$ is a set of the state-specific model parameters with harmonic potentials $V^{(j)}$ of the form

$$V^{(j)}(Y) = \frac{1}{2}(Y - \mu^{(j)})^T D^{(j)}(Y - \mu^{(j)}) + V_0^{(j)}, \quad (4)$$

$\mu^{(j)}$ and $D^{(j)}$ denoting *equilibrium position* and Hesse-matrix of the OU process within conformation j . This process is therefore specified by the parameters $\Theta^{(j)} = (\mu^{(j)}, D^{(j)}, \sigma^{(j)})$. Since the output process is given by a stochastic differential equation we will refer to this model modification as HMMSDE. Its entire parameter set is $\Theta = (\Theta^{(1)}, \dots, \Theta^{(m)}, P)$, where P denotes the transition matrix of the Markov chain in (3).

The parameter set of this model can be estimated from a time series via a modified EM (expectation-maximization) algorithm [16], as described in [6, 7, 17]. Once the model parameters are estimated one can use the Viterbi algorithm [18] to compute the most probable path of hidden states, the *Viterbi path*, given observation sequence. So both can be obtained, a dynamical model and the assignment of data points to the hidden, not observed, states. In contrast to the transfer operator approach, where the number of metastable states is extracted from the spectral properties of the transition matrix, we have to specify the number of metastable states as an input parameter for the EM algorithm. Since this number is in general unknown, a combination of both algorithms is used: First, guess a sufficiently large number of metastable sets, compute a Viterbi path and then reduce the number of states by set up a transition matrix from the (discrete) Viterbi path and cluster with the transfer operator approach.

4 Conformation analysis of a Glycine Di-Peptide Analogue (GLDA)

As an example we investigate the dynamics of glycine dipeptide analogue ($CH_3-CO-NH-CH_2-CO-NH-CH_3$), which is one of the smallest (artificial) peptide containing two peptide bonds ($CO-NH$). Thus the essential degrees of freedom are the torsional rotations of the individual peptide units ($-CO-NH-$) about the backbone of the chain, where Φ and Ψ describe the torsion of the N-terminus ($CH_3-CO-NH-$) and the C-terminus ($-CO-NH-CH_3$), respectively, with regard to the central CH_2 group, see Fig. 2. The plane spanned by the two angles Φ, Ψ is referred to as Ramachandran plane [19], with values of $(\pm 180^\circ, \pm 180^\circ)$ corresponding to a fully extended conformation of the chain. For longer polypeptide chains these angles serve to characterize typical secondary structural motifs such as helices and sheets.

4.1 GLDA in the gas phase

We used an empirical force field (Gromos 53a6 [20, 21]) to obtain a potential energy surface in the two essential degrees, i.e., varying the Ramachandran angles and minimizing the potential energy wrt. the other degrees of freedom. The potential energy surface shown in Fig. 2, reflects the symmetry of the peptide. Local minima correspond to energetically favorable formations of ringlike structures, including seven or five atoms (C5 and C7), closed by

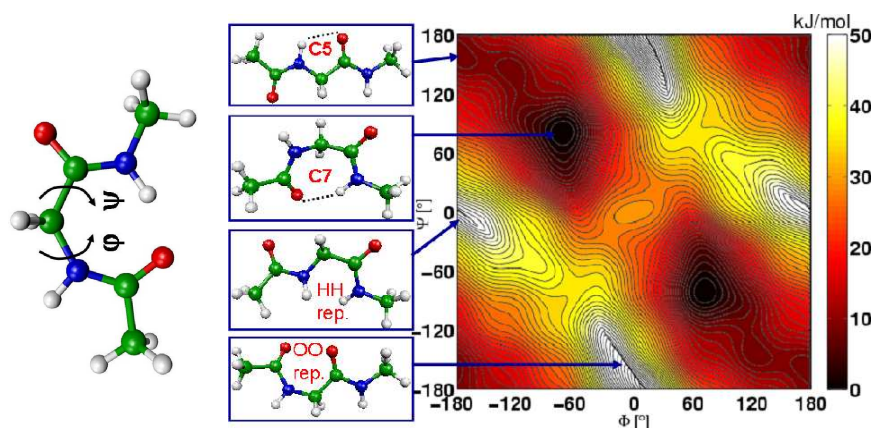


Fig. 2. *Left:* Glycine dipeptide analogue. The two marked torsion angles are the essential degrees of freedom as the peptide units are essentially planar. *Right:* Potential energy surface with respect to the two torsion angles. The local minima are due to the formation of hydrogen bonds between the peptide groups, the local maxima are due to repulsion.

(strongly) frustrated intramolecular hydrogen bonds. The maximal regions are corresponding to intramolecular repulsion of $-O O-$ and $-H H-$. The accuracy of the potential energy surface is of course limited by the quality of the empirical force field used, but comparison with the potential energy surface computed by quantum chemical calculations yields a qualitatively similar picture [22–26].

Performing a finite temperature MD-simulation at 300K using a Berendsen thermostat of the dipeptide in vacuum, samples the low energy regions of C5 and C7 in the Ramachandran plane, see Fig 3(E). As these regions are separated by a barrier of approx. 9 kJ/mol one would expect a metastable behavior at a reasonable timescale. To confirm this assumption, we use HMMSDE to extract a Viterbi path, assuming 4 (hidden) states, for each of the Ramachandran angles Φ, Ψ . Superposition of these two Viterbi paths yields a Viterbi path with 11 states. As the number of hidden states is only determined by our initial guess, we use the transfer operator approach to further reduce the number of states. Setting up the stochastic transition matrix P with time lag $\tau = 0.1ps$, and computing the first five eigenvalues:

$$\sigma(P) = \{1, 0.9948, 0.9296, 0.8152, 0.6540 \dots\},$$

indicates 3 metastable sets. Using the information coded in the three dominant eigenvectors we aggregate the 11 states to 3 states. In Fig. 3 (E) the assignment of the data points to these metastable (hidden) states is shown, the plot reveals, that we identified the (symmetric) C7 conformations and the C5 conformation. Transition probabilities between these hidden states can be obtained by using Eq. 1, Fig. 3 (D). Thus we have obtained a detailed

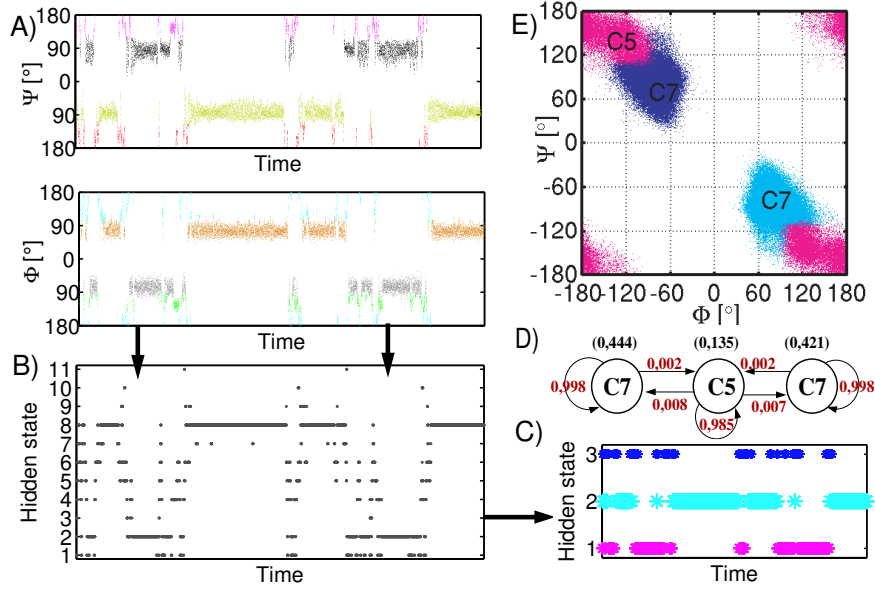


Fig. 3. *A)* A sample of the two dimensional torsion angle time series extracted from a 300K MD-simulation. The data points are colored according to the allocation to hidden states by HMM-SDE. *B)* Superposition of the two Viterbi paths yields into a joint Viterbi path with 11 hidden states. *C)* Using the transfer operator approach lumps the 11 hidden states to 3 metastable hidden states. *D)* A transition network for the 3 hidden states with a time lag of $\tau = 0.1ps$, red numbers denote the conditional transition probabilities, numbers in brackets the weight of each state. *E)* The data points of the torsion angle time series in the Ramachandran plane colored according to their allocation to metastable states.

dynamical picture of the effective finite temperature dynamics of GLDA in vacuum.

4.2 GLDA in aqueous solution

To compare these results with the dynamics in solution phase we consider GLDA in a $(3.5nm)^3$ box filled with 1405 rigid water molecules. Using a cutoff for electrostatic interactions of 1.1 nm and a Berendsen-temperature coupling to the solvent of 300K, we performed an MD-simulation over 2.5 ns with an integration timestep of 2 fs using again the Gromos 53a6 force field and recording the atom positions every 20 fs. After discretizing the Ramachandran plane in $5^\circ \times 5^\circ$ boxes, the free energy for each box B_i can be calculated by [27, 28]

$$\Delta G(B_i) = -k_B T (\log(\mathbb{P}[B_i]) - \log(\max_i \mathbb{P}[B_i])).$$

This free energy surface, see Fig. 4, has, due to intermolecular interactions (water-GLDA), a considerably richer structure than the potential energy surface in gas phase, Fig. 2. Analyzing the torsion angle time series with the HMMSDE approach, assuming 24 metastable sets, perfectly distinguishes regions belonging to different local minima in the free energy surface, Fig. 4. As these local minima are separated by low energy barriers, compared with thermally available energies, it is not a priori clear that they correspond to metastable states on timescales of e.g. 1 ps. An instructive picture is obtained by setting up the transition matrix, based on the 24 states of the obtained Viterbi path, and plotting the eigenvalues against the time lag used, Fig. 5. It can be clearly seen that 4 metastable states are persistent even for larger time lags, as there is an obvious gap after the first 4 eigenvalues.

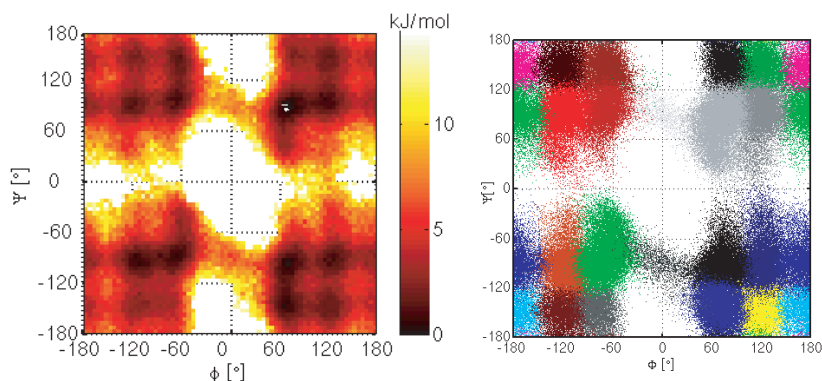


Fig. 4. *Left:* The free energy surface obtained from MD-Simulation with water has a considerably richer structure than the potential energy surface in vacuum. *Right:* Coloring the data points according to state allocation obtained by HMMSDE perfectly distinguishes the local minima in the free energy surface (for 24 metastable sets assumed).

The cause of the metastable states can be revealed by taking the intermolecular interactions into account. These interactions are mainly due to H-bond bridges between the peptide groups ($-CO-NH-$) and neighboring water molecules. Each peptide group provides a donor pair (NH) and an acceptor (O) for H-bond bridges attracting solvent molecules. If we restrict to microsolvation structures, i.e. GLDA with 1 or 2 waters, it is clear that ring like structures, as shown in Fig. 6, are energetically favorable, as each water molecule can participate in two H-bond bridges [29]. Besides the possible extension of the C7 and C5 structure to C7+2 (atoms), C7+2+2, C5+2 and C5+2+2 structures, H-bond bridges can stabilize structures that do not occur in vacuum, namely the C6 structures shown, where the $-O O-$ and $-H H-$ repulsion is overcome by inserting water molecules to form a ring struc-

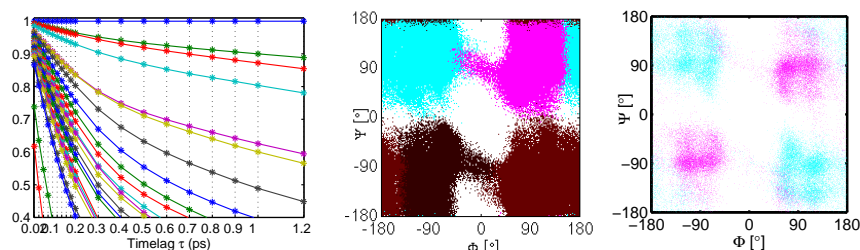


Fig. 5. *Left:* Dependence of the eigenvalues of the transition matrix obtained from the Viterbi path on different time lags. *Middle:* Data points are colored according to a clustering of the Viterbi path in 4 metastable sets. *Right:* Data points exhibiting a C6 microsolvation structure (magenta) and data points exhibiting a C7 or C5 microsolvation structure (cyan).

ture. In the following we denote by C7+X the C7+2+2 and C7+2 structures collectively (with analogous meaning of C6+X and C5+X).

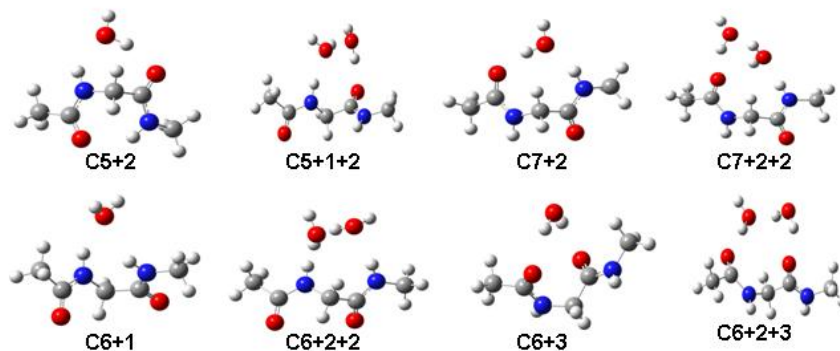


Fig. 6. Energetically favorable microsolvation structures with one or two water molecules.

These microsolvation ring structures can also be identified in the fully solvated system. Comparison of the plots in Fig. 5 reveals the nature of the four metastable states. They correspond to regions where C7+X/C5+X or C6+X ring structures occur.

The assumption that microsolvation structures causes metastability can be further supported by redoing the analysis based on six metastable states, see Fig. 7. Even though there are regions in the Ramachandran plane allowing different microsolvation structures, e.g. regions allowing C7+X or C5+X structures have an overlap, the plots of the data points belonging to a metastable state and the plots of data points with certain microsolvation structures show an obvious similarity. Again, this indicates that the origin of the metastable

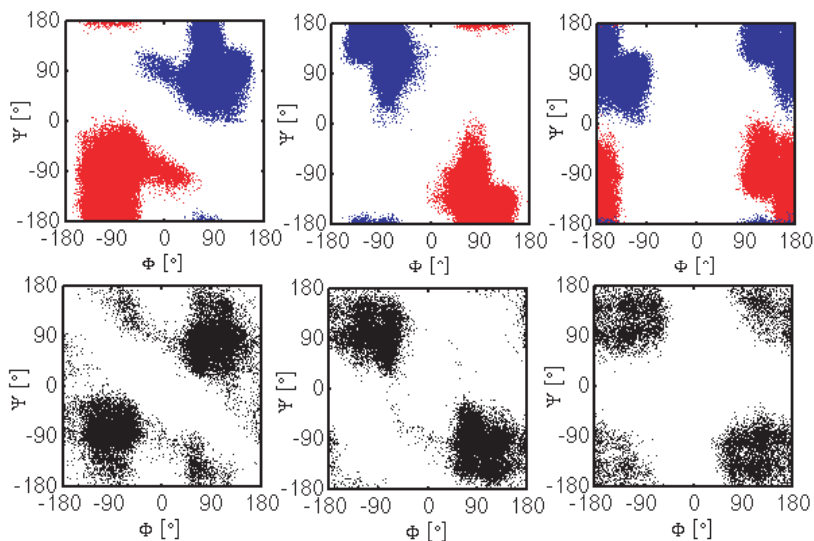


Fig. 7. *Top:* Data points belonging to different metastable states after clustering in 6 states, two symmetric equivalent states are shown in one plot. *Bottom:* Data points exhibiting C6+X (left), C7+X (middle) or C5+X (right) ring structures.

conformational structures is related to the formation of different microsolvation environments of the solute molecule.

5 Conclusion

We demonstrated the ability of our HMMSDE approach to reflect structural properties of the complete simulated system by analysis of only two essential degrees of freedom. The effective reduction of dimensionality achieved for the GLDA example is due to the capability of HMMSDE to distinguish different dynamical behavior in time series. Although the system investigated here is of moderate size, HMMSDE appears to be a promising approach to beat the curse of dimensionality in more complex systems. Currently conformational analysis of DNA fragments containing 15 base pairs has been pursued in our laboratory [17]. Hence, it is believed that this approach is much more general and can be used beyond the context of MD-simulations. Possible applications range, e.g., from transient spectroscopy to the analysis of climate or financial data.

References

1. P. Deuffhard, C. Schütte, in *Applied Mathematics Entering the 21st Century. Proc. ICIAM 2003*, ed. by J.M. Hill, R. Moore (2004), pp. 91–119
2. C. Schütte, A. Fischer, W. Huisinga, P. Deuffhard, *J. Comput. Phys.*, Special Issue on Computational Biophysics **151**, 146 (1999)
3. P. Deuffhard, W. Huisinga, A. Fischer, C. Schütte, *Lin. Alg. Appl.* **315**, 39 (2000)
4. M. Dellnitz, O. Junge, *SIAM J. Num. Anal.* **36**(2), 491 (1999)
5. A. Fischer, S. Waldhausen, I. Horenko, E. Meerbach, C. Schütte, *J. Chem. Phys.*, submitted (2004)
6. I. Horenko, E. Dittmer, C. Schütte, *Comp. Vis. Sci.* **9**(2), 89 (2005)
7. I. Horenko, E. Dittmer, A. Fischer, C. Schütte, *Multiscale Modeling and Simulation*, accepted (2005)
8. E. Meerbach, E. Dittmer, I. Horenko, C. Schütte, in *Computer Simulations in Condensed Matter Systems, Lecture Notes in Physics*, vol. 703 (2006), *Lecture Notes in Physics*, vol. 703, pp. 475–497
9. P. Metzner, C. Schütte, E. Vanden-Eijnden, *J. Chem. Phys.*, accepted (2006)
10. F. Cordes, M. Weber, J. Schmidt-Ehrenberg, *Metastable conformations via successive Perron cluster analysis of dihedrals* (2002). ZIB-Report 02-40, Zuse-Institute-Zentrum, Berlin
11. P. Deuffhard, M. Weber, *Robust Perron cluster analysis in conformation dynamics* (2003). ZIB-Report 03-19, Zuse Institute Berlin
12. W. Huisinga, B. Schmidt, in *New Algorithms for Macromolecular Simulation, Lecture Notes in Computational Science and Engineering*, vol. 49, ed. by C. Chipot, R. Elber, A. Laaksonen, B. Leimkuhler, A. Mark, T. Schlick, C. Schütte, R. Skeel (Springer, 2005), *Lecture Notes in Computational Science and Engineering*, vol. 49, pp. 167–182
13. C. Schütte, W. Huisinga, in *Handbook of Numerical Analysis*, vol. X, ed. by P.G. Ciarlet, J.L. Lions (North-Holland, 2003), pp. 699–744
14. P. Bremaud, *Gibbs fields, Monte Carlo simulation, and queues, Texts in applied mathematics*, vol. 31 (Springer, New York, 1999)
15. M. Weber, *Improved Perron cluster analysis* (2004). ZIB-Report 03-04, Zuse Institute Berlin
16. A.P. Dempster, N.M. Laird, D.B. Rubin, *J. Roy. Stat. Soc. B* **39**(1), 1 (1977.)
17. I. Horenko, E. Dittmer, F. Lankas, J. Maddocks, P. Metzner, C. Schütte, *J. Appl. Dyn. Syst.*, submitted (2005)
18. A.J. Viterbi, *IEEE Trans. Informat. Theory* **IT-13**, 260 (1967)
19. G.N. Ramachandran, V. Sasiskharan, *Advan. Prot. Chem.* **23**, 283 (1968)
20. D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. Mark, H.J. Berendsen, *J. Chem. Phys.* **26**, 1701 (2005)
21. C. Oostenbrink, A. Villa, A.E. Mark, W.F.V. Gunsteren, *J. Comp. Chem.* **25**, 1656 (2004)
22. J. Antony, B. Schmidt, C. Schütte, *J. Chem. Phys.* **122**(1), 014309 (2005)
23. T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III, J.A. Pople, *J. Am. Chem. Soc.* **113**(16), 5989 (1991)
24. A. Perczel, Ö. Farkas, I. Jakli, I.A. Topol, I.G. Csizmadia, *J. Comp. Chem.* **24**(9), 1026 (2003)
25. L. Schäfer, C.V. Alsenoy, J.N. Scarsdale, *J. Chem. Phys.* **76**(3), 1439 (1982)

26. H. Hu, M. Elstner, J. Hermans, *Proteins* **50**, 451 (2003)
27. Y. Mu, G. Stock, *J. Phys. Chem. B* **106**(20), 5294 (2002)
28. Y. Mu, P.H. Nguyen, G. Stock, *Proteins* **58**(1), 45 (2004)
29. W.G. Han, K.J. Jalkanen, M. Elstner, S. Suhai, *Journal of Physical Chemistry B* **102**(14), 2587 (1998)