Conformational Studies of UDP-GlcNAc in Environments of Increasing Complexity

Martin Held[1], **Eike Meerbach**[1,3], **Stephan Hinderlich**[2], **Werner Reutter**[2], **and Christof Schütte**[1]

[1] Fachbereich Mathematik und Informatik,
Freie Universität Berlin
Arnimallee 2-6, 14195 Berlin
E-mail: {*held, meerbach, schuette*}*@mi.fu-berlin.de*

[2] Institut für Biochemie und Molekularbiologie
Campus Benjamin Franklin
Charité Universitätsmedizin Berlin
Arnimallee 22, 14195 Berlin
E-mail: {*Stephan.Hinderlich, Werner.Reutter*}*@Charite.de*

[3] Supported by the DFG research center "Mathematics for key technologies".

The "effective" dynamics of a biomolecular system can often be described by means of a Markov chain describing "flipping dynamics" between metastable (geometrical large scale) molecule conformations on "long" time scales. In this work we present our methods for the identification of metastable conformations by a study of UDP-GlcNAc - a key substrate in sialic acid synthesis. We investigate the system in different environments of increasing complexity (vacuum, water, peptide) by performing force field based molecular dynamic simulations. By applying our analysis techniques to the obtained data we get insights into the conformational dynamics of the system. A comparison of the results reveals interesting environment-dependent properties of UDP-GlcNAc.

1 Introduction

Sialic acids play an important role in various biological processes, e.g., immune response, tumor metastasis or inflammatory reactions. In mammalian cells the key reaction of the sialic acid synthesis pathway is carried out by the bifunctiononal enzyme UDP-GlcNAc-2-Epimerase/ManNAc-Kinase[1]. In cooperation with researchers at the Charité we investigate UDP-N-Acetylglucosamine (UDP-GlcNAc) (see Fig. 1), the ligand of this key reaction by means of molecular dynamic simulations. Thereby, we expect to gain a better understanding of the actual reaction and, in the long term, insights that assist in the design of potential inhibitors.

2 Simulation

By a systematic conformational search in dihedral space we generated 20 initial conformers of UDP-GlcNAc. For each conformer we run molecular dynamic simulations *in vacuo* and in explicit water using AMBER 8. Additionally, we simulated a set of possible binding conformations of UDP-GlcNAc. To derive these, crystallographic data of UDP-GlcNAc-2-Epimerase in complex with UDP was utilized. To compensate for the missing information

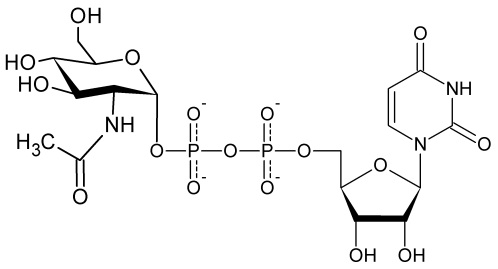

Figure 1. UDP-N-acetylglucosamine

of the GlcNAc orientation, we generated conformers of UDP-GlcNAc by systematically turning the three dihedrals that qualify the orientation of the GlcNAc moiety . The force field parameters needed for this non-standard residue were obtained from the literature[12, 14].

	Atoms	Sim. Time	Step Size	Temperature	Pressure
In Vacuo	64	100 ns (2 μs)	1 fs	300 K	-
Water	3964	10 ns (200 ns)	2 fs	300 K	1 atm
Protein	46646	100 ps	2 fs	300 K	1 atm

Table 1. Simulation Parameters

3 Analysis Methods

Our aim is to identify so-called *metastable* conformations of UDP-GlcNAc, i.e., geometries which are persistent for long periods of time, from the generated trajectory data. Such identification corresponds to the set up of a coarse-grained model for the dynamics, as on long time scales the typical dynamics of biomolecular systems can often be described as a (Markovian) flipping process between such conformations[6, 8, 13]. The challenge of identification lies in the rich temporal multiscale structure of the data, induced by flexibility within the metastable conformations.

The groups of P. Deuflhard and C. Schütte developed the so-called *Perron Cluster Cluster Analysis* (PCCA) for the identification of metastable conformations[2, 3, 5]. The approach is based on assuming a Markov dynamics of the observed system on a certain time scale. Discretization of the state space, e.g., the dihedral angle space of some molecule, allows to set up a stochastic matrix, w.r.t. the chosen time scale, by counting transitions between discrete states in the trajectory. The information contained in the spectrum of this matrix allows to aggregate parts of the state space to metastable states, as the number of metastable states corresponds to the number of eigenvalues close to unity, while information about an appropriate clustering is encoded in the eigenvectors.

In addition a variety of methods based on Hidden Markov Models (HMMs) were recently developed in the group of C. Schütte[7, 9–11]. An intriguing feature of HMM analysis is that the chosen set of observables need not to enable a geometric separation of the

metastable conformations. Instead, the *observed* time series is used to fit a model for an *unobserved* Markov chain, by extracting geometric as well as dynamical properties out of the given observation sequences, and thus allowing for separation of overlapping conformations.

For our analysis we used a combination of these approaches. First each chosen torsion angle is analyzed separately by using an HMM with gaussian output distributions, resulting in a discrete time series corresponding to the hidden states for each angle. Then the obtained information is aggregated by superposition of the discrete one dimensional time series. The (combined) discrete time series is then further analyzed by means of PCCA. For more details refer to, e.g.,[11].

4 Results

Based on the *in vacuo* simulation data we identified 5 metastable conformations via an analysis on the dihedral space. All found conformations are stabilized by intramolecular hydrogen bonds, where the most compact structure corresponds to the metastable set with the highest weight (66.6%). In the water trajectory data we found 6 metastable sets. Here, the related structures appear to be less stiff, as the effect of inner hydrogen bonds is weakened, due to the shielding effect of the present water molecules.

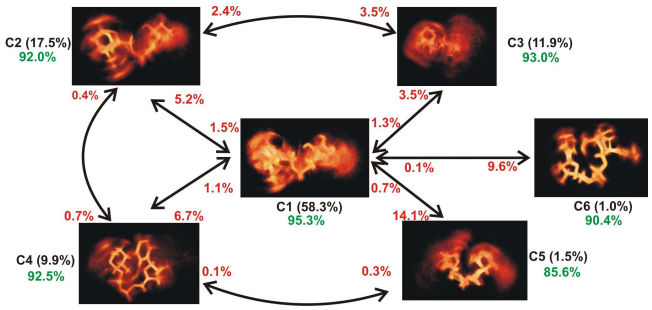

Figure 2. Transition network of UDP-GlcNAc in explicit water with density plots showing the flexibility within each metastable set. (in brackets – relative weighting, below brackets – persistence probability, next to pictures – exit probability) – Visualizations using Amira[15]

Screening the water simulation data for the generated potential protein binding conformations of UDP-GlcNAc did not result in any match. This leads to the conjecture that the binding of the ligand must be accompanied by an induced-fit effect. The protein has to alter the conformation of the ligand in order to bind it. In contrast to a conformational selection scenario, where an available conformation would be ”selected” to bind.

References

1. S. Hinderlich, R. Stäsche, R. Zeitler, and W. Reutter A bifunctional enzyme catalyzes the first two steps in N-acetylneuraminic acid biosynthesis of rat liver. Purification and

characterization of UDP-N-acetylglucosamine2-epimerase/ N-acetylmannosamine kinase. *J Biol Chem 272*, pages 24313-24318 (1997)
2. F. Cordes, M. Weber, and J. Schmidt-Ehrenberg. Metastable conformations via successive Perron cluster analysis of dihedrals. ZIB-Report 02-40, Zuse Institute Berlin, 2002.
3. P. Deuflhard, W. Huisinga, A. Fischer, and C. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Lin. Alg. Appl.*, 315:39–59, 2000.
4. P. Deuflhard and Ch. Schütte. Molecular conformation dynamics and computational drug design. In J. M. Hill and R. Moore, editors, *Applied Mathematics Entering the 21st Century*, pages 91–119. SIAM, 2004.
5. P. Deuflhard and M. Weber. Robust Perron cluster analysis in conformation dynamics. ZIB-Report 03-19, Zuse Institute Berlin, 2003.
6. R. Elber and M. Karplus. Multiple conformational states of proteins: A molecular dynamics analysis of Myoglobin. *Science*, 235:318–321, 1987.
7. A. Fischer, S. Waldhausen, I. Horenko, E. Meerbach, and Ch. Schütte Identification of Biomolecular Conformations from Incomplete Torsion Angle Observations by Hidden Markov Models *J. Comp. Chem.*, accepted (2006).
8. C. Schütte and W. Huisinga. Biomolecular Conformations can be Identified as Metastable Sets of Molecular Dynamics. In P. G. Ciaret and J.-L. Lions, editors, *Handbook of Numerical Analysis X*, Special Volume Computational Chemistry, pages 699–744, 2003.
9. I. Horenko, E. Dittmer, A. Fischer, and Ch. Schütte Automated Model Reduction for Complex Systems exhibiting Metastability, *Mult. Mod. Sim. 5 (3)*, pages 475–491 (2006)
10. I. Horenko, E. Dittmer, F. Lankas, J. Maddocks, Ph. Metzner, and Ch. Schütte Macroscopic Dynamics of Complex Metastable Systems: Theory, Algorithms, and Application to B-DNA Submitted to *J. Appl. Dyn. Syst.* (2005)
11. E. Meerbach, E. Dittmer, I. Horenko, and Ch. Schütte. *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*, volume 1 of *Lecture Notes in Physics*, chapter Multiscale Modelling in Molecular Dynamics: Biomolecular Conformations as Metastable States, pages 475–497. Springer, 2006.
12. P. Petrova, J. Koca, and A. Imberty. Potential energy hypersurfaces of nucleotide sugars: Ab initio calculations, force field parametrization and exploration of the flexibility. *J. Am. Chem. Soc.*, 121:5535–5547, 1999.
13. C. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys.*, 151:146–168, 1999.
14. Woods Group. GLYCAM Web . 15 January 2007. Complex Carbohydrate Research Center , The University of Georgia , Athens , GA. `http://www.glycam.com`
15. Amira—advanced visualization, data analysis and geometry reconstruction, user's guide and reference manual. Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Mercury Computer Systems GmbH and TGS Template Graphics Software Inc., 2000.