Long-Scale Molecular-Dynamics Simulations of Cyclic Peptide Hormones Vasopressin, Urotensin and Analogues

Elke Haensele

The thesis is submitted in partial fulfilment of the requirements

for the award of the degree of Doctor of Philosophy

of the University of Portsmouth

June 2017

FRONT PAGES

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

This is a cumulative thesis comprising three peer-reviewed published full papers showing the main results of this research project. I was the first author of these papers and the contribution of my co-authors is made clear in the text.

To keep the journals´ copyrights, post prints of the peer-reviewed, accepted papers are given.

Word count: 39,044 (without Front Pages, Appendices and References)

Elke Haensele

University of Portsmouth

June 2017

Abstract

This thesis describes unrestrained microsecond-scale molecular-dynamics (MD) simulations of the structurally related peptide hormones Arg^{8}-vasopressin (AVP), urotensin II (UII), urotensin-related peptide (URP), Leu^{8}-oxytocin (OT) and analogous. All are agonistic ligands of G-protein coupled receptors that regulate a multitude of physiological functions. They are thus, connected with many pathophysiological processes, making them a major target for drug design.

The common structural feature of these intrinsically flexible peptides is a cyclic 6-residue moiety closed by a disulphide bridge. The conformational space was explored and systematically clustered with the analysis method DASH. The main conformations were classified: They all show two main classes of ring conformations independently of their primary sequence. One comprises unfolded ring conformations (denoted as *open*) with no significant transannular hydrogen bonds and the other *folded*, ring conformations with multiple turns stabilised by highly populated hydrogen bonds. The conformations of the latter type are often considered as the bioactive structure within the binding pocket of the receptor. C- or N-terminal tails either adopt *extended* or *folded* conformations that generally interconvert more frequently than the ring. An interdependence of ring and tail conformations is possible; however, it is most appropriate to base the conformational classification primarily on the ring conformation. Structure coordinates of the main conformations may serve as input for 3D drug design, receptor/ligand modelling or further simulations.

Fast conformational equilibria in solution are difficult to access with experimental methods. A new technique is introduced that is able to decipher nuclear magnetic resonance (NMR) data of these equilibria with a combination of MD simulations and NMR calculations without classical analysis of Nuclear-Overhauser effect (NOE) distances and coupling constants. The technique was tested successfully for AVP, a “known system”, and subsequently applied to UII/URP, a “less well known system”. Based on these results, current single-conformation descriptions of AVP and UII/URP need to be replaced by a description as fast equilibria of *open* and *folded* conformations with characteristic *open:folded* ratios (AVP 30:70, UII 72:28, URP 86:14). Insights into the pre-allosteric dynamics may contribute to the understanding of factors that influence bioactivity.

The NMR data from experiments performed within this research supplement experimental data from the literature (*e.g.* assignment of *cis*-Pro^{3}-UII; ^{15}N chemical shifts for AVP and UII/URP).

The main results of this thesis have been published in peer reviewed journals.

Table of Contents

List of Tables

Chapters 1-7

Appendix 1 (Online Supporting Information – Paper 1)

Appendix 2 (Online Supporting Information – Paper 2)

Appendix 3 (Online Supporting Information – Paper 3)

Appendix 4-9

List of Figures and Schemes

Chapters 1-7

Appendix 1 (Online Supporting Information – Paper 1)

Appendix 2 (Online Supporting Information – Paper 2)

Appendix 3 (Online Supporting Information – Paper 3)

Appendix 4-8

Abbreviations

Abbreviation	Long Form
1JK4	RCSB Protein Data Bank structure code Lys^{8}-vasopressin
1NPO	RCSB Protein Data Bank structure code oxytocin
1YF4	RCSB Protein Data Bank structure code Arg^{8}-vasopressin
abs	Absolute
ACTH	Adrenocorticotropic Hormone
AMBER	Assisted Model Building with Energy Refinement
av	Average
AVP	8-Arg-vasopressin, Arg^{8}-vasopressin
B3LYP	Becke 3-Parameter (Exchange), Lee, Yang and Parr
BUA	Butanoic Acid
CD	Circular Dichroism
CERMN	Centre d'Etudes et de Recherche sur le Médicament de Normandie
CFWKYC	Cys-Phe-Trp-Lys-Tyr-Cys
CHARMM	Chemistry at HARvard using Molecular Mechanics
circsim	Circular Similarity
clop, cl.open	*clinched open*
CPU	Central Processing Unit
CT	Carbetocin
CUDA	Computer Unified Device Architecture (Language)
DASH	Dynamics Analysis by Salt and Hudson
dAVP	Deamino-Arg^{8}-vasopressin
DFT	Density-Functional Theory
DMS	Dimethyl Sulphate
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic acid
dOT	Deamino-oxytocin
DP4	Probability Measure by Goodman and Smith
DPC	Dodecylphophocholine
DSS	(3-trimethylsilyl)propane sulfonic acid (NMR standard)
EDMC	Electronically Driven Monte Carlo
FAU	Friedrich-Alexander Universität
ff99SB	Force Field 1999 Stony Brooks
g/g'	gauche/gauche'
gHSQC	Gradient Heteronuclear Single Quantum Coherence
GIAO	Gauge-Independent Atomic Orbital, Gauge-Invariant Atomic Orbital
GNU	(a general public license)
GPCR	G-Protein Coupled Receptor
GPU	Graphics Processing Unit
GROMOS	Groningen Molecular Simulation Computer Program Package
Hbond	Hydrogen Bond
h-UII	Human Urotensin
IEFPCM	Integral Equation Formalism variant of PCM
IGLO	Individual Gauges for Localised Orbitals
Interreg EU	Interreg IVA France (Channel) - England 2007-2013 progamme
inv-folded	*inverse folded*
LEaP	Link Edit and Parm
LKH	Lock-and-Key Hypothesis
LVP	8-Lys-Vasopressin, Lys^{8}-Vasopressin
MD	Molecular Dynamics
MM	Molecular Mechanics
MSE	Mean Square Error
MUE	Mean Unsigned Error

Abbreviation	Long Form
MWC model	Monod-Wyman-Changeux model
NH	Amide Hydrogen
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
NP	Neurophysin
O	Carbonyl Oxygen
OPLS	Optimised Potentials for Liquid Simulation
OT	Oxytocin
OTR	Oxytocin Receptor
Pauling-KFN	Pauling-Koshland, Nemethy, Filmer
PBC	Periodic Boundary Conditions
PC	Principal Component
PCA	Principal Component Analysis
PCM	Polarizable Continuum Model
PE	Potential Energy
PDB	Protein Data Bank (file format)
PDF	Portable Document Format (Adobe Acrobat)
PeReNE	Peptide Research Network of Excellence
PME	Particle Mesh Ewald
PMEMD	Particle Mesh Ewald Molecular Dynamics
QSAR	Quantitative Structure Activity Relationship
R^2	Coefficients of Determination
RadGyr	Radius of Gyration
RCSB	Research Collaboratory for Structural Bioinformatics
ref	Reference
rel	Relative
REMD	Replica Exchange Molecular Dynamics
RF	Reaction Field
RMSD	Root Mean Square Deviation
SCI	Science Citation Index
SDS	Sodium Dodecyl Sulphate
SI	Supporting Information, Supplementary Information
stddev	Standard Deviation
T6	DASH analysis of 6 torsions, *e.g.* $\Phi\Psi$ 7 to 9 AVP tail states
T10	DASH analysis of 10 torsions, *e.g.* $\Phi\Psi$ 2 to 6 AVP ring states
T16	DASH analysis of 16 torsions, *e.g.* $\Phi\Psi$ 2 to 9 AVP overall states
TIP4P-Ew	Transferable Intermolecular Potential 4 Point - Ewald
TM	Trans Membrane
TMS	Tetramethylsilane, Si(CH3)4 (NMR standard)
TOCSY	Total Correlated Spectroscopy
tws, tw.saddle	*twisted saddle*
UII	Urotensin II (here, also used for human urotensin II)
UNICAEN	Université de Caen Basse−Normandie
URP	Urotensin-Related Peptide
UTR, UTS2R	Urotensin II Receptor
V1aR	Vasopressin-1a Receptor (blood pressure)
V1bR	Vasopressin-1b Receptor (ACTH secretion)
V2R	Vasopressin-2 Receptor (antidiuresis)
WHAM	Weighted Histogram Analysis Method
WKY	Trp-Lys-Tyr
WRMSE	Weighted Root Mean Square Error
YMe	O-methyl-L-tyrosine
ZIP	Compressed File Format
Δ_σ	Coefficients of Distinctiveness

Acknowledgements

First and foremost, I would like to thank **Prof. Tim Clark**, who was the supervisor of my Diplomarbeit in Computational Chemistry in 1987. But as life goes, I did not continue a scientific career. Nevertheless, he held the door open and offered me to come back any time. Almost 30 years later, I did and Tim established the contact to Dr. Lee Banting, who became my first supervisor at the School of Pharmacy and Biomedical Sciences of the University of Portsmouth for my PhD project. Dear Tim: Thank you so much for giving me this opportunity. I am blessed to have you as a mentor and Doktorvater and I am proud to be your "most persistent" PhD student.

Dr. Lee Banting was the best supervisor I could have wished for. He gave me a comprehensive introduction to the field of biochemistry, which was totally new to me, but made clear that I should aim to become the expert in my field. He not only was always open to every scientific concern but also supported, encouraged and accompanied me on the long way of ups and downs through my research project.

Special thanks go to **Dr. David Whitley**. I could always count on his excellent expert knowledge in all mathematical and computational questions, which I abundantly used. His kind and un-frantic manner made teamwork a pleasure.

Also, many thanks to **Dr. Chris Read**, the most competent NMR expert. Thank you very much for your patience in introducing a computational chemist to experimental NMR spectroscopy. Your scepticism towards "*in silico science"* was always challenging; it made me learn to critically question my own work and to deeply respect experimental work.

PD Dr. Harald Lanig, a gifted computational chemist and teacher, has helped me make a good start in the field of molecular dynamics simulations. You have laid the foundations on which I was able to build upon.

Dr. Nico van Eikema-Hommes: I would have despaired without you! Nico is the system administrator of the Computer-Chemie-Centrum (FAU Erlangen-Nürnberg) and a real trouble shooter. He was responsive to all computational interconnecting problems at any time, without which it would have been impossible to perform my MD simulations in Erlangen, Germany, while physically been based in Portsmouth, UK, or Puebla, Mexico. Thank you!

A warm thank you also to my PhD colleagues, their supervisors and all scientists, who contributed to the success of the peptide project with their own work, their expertise, their feedback and their

co-authorship on my papers – namely **Marija Miljak, Nawel Mele** and **Prof. Jon Essex** from University of Southampton, **Dr. Carla Delépée, Prof. Jana Sopkova**, and **Prof. Ronan Bureau** from CERMN Normandie Université Caen, and **Dr. Nour Saleh** from FAU Erlangen-Nürnberg.

I must not forget the academic, technical and administrative staff, which I met during my time at the University of Portsmouth. I, as a stranger to the UK, always felt comfortable and I would like to thank you because this is not a matter of course.

Furthermore, the support from my friends is truly acknowledged: **Sabine Becker**, for moral support and excellent proof reading, **Mireida y Gabriela de Ramírez Macuil**, for culinary well-being and daily company, and **all friends and relatives** that stood in contact despite long-distances during my time in UK and Mexico. *Thanks a million*, *muchas gracias* and *vielen herzlichen Dank* for not letting me walk alone!

Finally, I would like to thank the people that are closest to me. This is my family - my husband **Dr. Klaus Haensele** and my beautiful sons **Christian Haensele** and **Andreas Haensele**. They experienced the emotional side of a research project at first hand. They showed interest in the scientific matter. They shared cheerful and frightened days. They believed in me when I doubted. They gave me (mother and wife) the freedom and space to do the PhD. Thank you so much - I love you <3.

…and it should not be forgotten, the funding from PeReNE (European project "Peptide Research Network of Excellence" as part of the Interreg IVA France (Channel) – England 2007-2014 program), enabling the NMR experiments and the attendance in PeRENE meetings in France and England; and the generous travel budget of the Partner-Support-Program from AUDI AG, which allowed me an appropriate commute between England and Mexico in the final two years.

"What was the motivation of this study?" There will be no universal answer satisfying the expectation of every reader. Funding for this project was clearly motivated by the prospect that the results would contribute to the development of future drugs. I was already fascinated by 3-dimensional structure and even more intrigued by the dynamics of conformations. Long-scale MD simulations give the unique opportunity to visualise atomistic dynamics. For me personally, it was great motivation to be given the chance to contribute to fundamental research in the area of structure determination of cyclic peptide hormones.

Dissemination

Publications

(1) Steinke T, Hänsele E, Clark T. The solvent effect on the electronic nature of 1,3-dipoles: an *ab initio* SCRF study. J Am Chem Soc. 1989;111:9107-9.

(2) Hofmann H, Hänsele E, Clark T. A cautionary note on the use of the frozen-core approximation for correlation energy calculations involving alkali metals. J Comput Chem. 1990;11(10):1147-50.

(3) Hänsele E, Clark T. *Ab initio* simulation of electron-transfer reactions - the reaction of alkali-metal atoms with ethylene. Z Phys Chem. 1991;171:21-31.

(4) Alex A, Hänsele E, Clark T. The ethylene/metal(0) and ethylene/metal(I) redox system: model *ab initio* calculations. J Mol Model. 2006;12(5):621-9. Epub 2005/12/13.

(5) Haensele E, Banting L, Whitley DC, Clark T. Conformation and dynamics of 8-Arg-vasopressin in solution. J Mol Model. 2014;20(11):2485(17). Epub 2014/11/07.

(6) Saleh N, Saladino G, Gervasio FL, Haensele E, Banting L, Whitley DC, *et al.* A three-site mechanism for agonist/antagonist selective binding to vasopressin receptors. Angew Chem Int Ed Engl. 2016;55(28):8008-12. Epub 2016/05/18.

(7) Haensele E, Saleh N, Read CM, Banting L, Whitley DC, Clark T. Can simulations and modeling decipher NMR data for conformational equilibria? Arginine-vasopressin. J Chem Inf Model. 2016;56(9):1798-807.

(8) Haensele E, Mele N, Miljak M, Read CM, Whitley DC, Banting L, *et al.* Conformation and dynamics of human urotensin II and urotensin-related peptide in aqueous solution. J Chem Inf Model. 2017;57(2):298-310.

Papers (5), (7), and (8) form part of this thesis.

Poster Presentations and Talks

(9) Haensele E, Banting L, Clark T. The necessity of long-term molecular-dynamics simulations: deamino-oxytocin - novel conformational insights. (a) 26th Molecular Modeling Workshop, March 12th, 2012. Erlangen, Germany. (b) IBBS Day, May 11th, 2012. University of Portsmouth, UK.

(10) Haensele E, Banting L, Clark T. Molecular dynamics and umbrella sampling simulations of 8-Arg-vasopressin. (a) 27th Molecular Modeling Workshop, Feb 25th, 2013. Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany. (b) IBBS Day, Jun 7th, 2013. University of Portsmouth, UK.

(11) Haensele E, Banting L, Clark T. Urotensin-related peptide (URP): long-term molecular-dynamics simulation. 28th Molecular Modeling Workshop, Mar 18th, 2014. FAU Erlangen-Nürnberg, Germany.

(12) Haensele E, Whitley D, Banting L, Clark T. DASH: Analysis of microsecond-scale molecular-dynamics trajectories (Talk). 28th Molecular Modeling Workshop, Mar 18th, 2014. FAU Erlangen-Nürnberg, Germany.

(13)Haensele E, Banting L, Whitley D, Read C, Cary P, Clark T, *et al.* Cyclic peptide hormones: conformation, dynamics and pharmacophores of urotensin and vasopressin (Joint Lecture). Final PeReNE Meeting, Jan 15-16th, 2015. University de Le Havre, France.

(14)Haensele E, Mele N, Miljak M, Read CM, Whitley DC, Banting L, *et al.* Urotensin II and urotensin-related peptide: how to decipher NMR-data for conformational equilibria with molecular-dynamics simulation and modelling. 13th German Peptide Symposium (DECHEMA), Mar 20-23, 2017. FAU Erlangen-Nürnberg, Germany.

Abstracts and poster reprints are given in the Appendix.

INTRODUCTION

Chapter 1: Objectives and Outline

The aim of this study was to elucidate the conformational space and dynamics of the cyclic peptide hormone AVP and structurally related peptides (OT, UII, URP, dOT, CT) in order to predict their conformational equilibria in solution. Sufficiently long simulations should expose all possible conformations (convergence) and should enable structural classification. Here, the first microsecond long-scale simulations were performed with these peptides. To predict conformational equilibria, a novel technique was tested and established combining results from NMR spectroscopy, density-functional theory (DFT)/ NMR calculations, long-scale MD simulations and enhanced sampling. The peptides investigated are agonists of their cognate G-protein coupled receptors (GPCRs). These peptides exhibit multiple physiological functions that make them a major target for drug design. As structure and function are interdependent, an atomistic understanding of the conformational dynamics of these peptides will contribute to the understanding of their bioactivity.

The thesis is structured as follows:

Introduction. The Introduction includes the current *Chapter 1 (Objectives and Outline) and Chapter 2 (Peptides – Biological Function and Structure).* Chapter 2 gives an introduction to bioactive peptides and conformation in general, and an overview of known structural data for AVP, OT, UII, URP, dOT and CT, in particular. It supplements the information on AVP, UII and URP given in the Introductions of Paper 1 and 3 (Chaps. 4 and 6).

Methods. The Methods part comprises *Chapter 3*, which explains the principles of the methods used in this work and discusses their advantages and limitations. The chapter supplements methodological details given in the Methods sections and Supporting Information of Papers 1, 2 and 3 (Chaps. 4-6).

Results. The research project went through three stages and the results of each stage were published consecutively. The thesis presents a cumulative form of these scientific papers (Chapters 4, 5 and 6) complemented with unpublished results. The papers are given as postprints (unmodified content embedded in the formatting of the thesis). Each paper is preceded by a short foreword and a clarification of co-author contributions. The Online Supporting Information from the original papers is included as reprints in Appendices A1 to A3. The chapters content in particular:

Chapter 4 (Paper 1: Conformation and Dynamics of Arg^8-Vasopressin). In the first stage of the research project, the structurally well-known system AVP was simulated to gain experience with long-scale simulations and to assess the reliability of this method. The applicability of the analysis method DASH for long trajectories was tested and optimised. It is shown that a separate consideration of ring and tail conformations is best suited to characterise the conformations of AVP, further supported by the finding that ring and tail conformations are not correlated. AVP comprises four main conformations that are described in detail. The results were published in 2014.[1]

Chapter 5 (Paper 2 Deciphering NMR-Data for Conformational Equilibria). The second stage of the project focused on the determination and evaluation of the conformational equilibrium of AVP. A protocol was established and developed to decipher the experimental NMR-spectra of AVP with respect to its conformational equilibrium. The technique was validated and showed promise for generic application to assess the conformation (or conformational equilibria) of flexible peptides in solution. AVP exhibits approximately 70 % folded (saddle) and 30 % open (clinched open) conformations. The results and technique were published in 2016.[2]

Chapter 6 (Paper 3: Conformation and Dynamics of Urotensin II and Urotensin-Related Peptide). In the third stage of the research project, the novel technique was applied to UII and URP. In this case, it was shown that conformational equilibria of open and folded conformations are better suited to describe the solution structures of UII and URP than single conformations. UII and URP favour open conformations in contrast to AVP. UII exhibits approximately 28 % folded and 72 % open conformations, URP 14 % folded and 86 % open. These findings were preceded by an in-depth exploration of the conformational space of these intrinsically flexible peptides and a systematic classification of their conformations. The results were published in 2017.[3]

Chapter 7 (unpublished results: Related Peptides and General Conformational Classification). Ongoing projects are the investigation of structurally related peptides and analogues (OT, dOT, CT). The current results are summarised and common features of all peptides investigated are discussed. Finally, a general conformational classification of cyclic peptides with 6-residue ring moiety is given.

Final Conclusions and Outlook. The last section summarises and reflects the work as a whole. The conclusions given in the papers (Chaps. 4 to 6) are combined with the unpublished results (Chap. 7) and placed in a general context. The relevance of the results is discussed and outlooks are given.

Appendices. The Appendices include the Online Supporting Information of the published papers and further supplementary material to chapters of this thesis.

Chapter 2: Peptides - Biological Function and Structure

Biological Function of Peptides

Bioactivity. Peptides consist of amino-acid sequences of variable length. Depending on the chain length, one can distinguish between oligopeptides (<= 10 residues) and polypeptides (>10 residues). The latter are called proteins when the chain length exceeds 50-100 residues.[4,5] Natural peptides are synthesised *via* both ribosomal and non-ribosomal processes.[6] When a peptide shows an effect on body function, it is deemed bioactive. The functionality ranges from toxic, antioxidant, antimicrobial, antihypertensive to neurotransmittant.[7,8] Malfunction of peptide signalling may lead to diabetes, cardiovascular diseases, arthritis, allergies, digestive dysfunctions, infections and inflammation, growth perturbation, obesity, cancer, diseases of the central nervous system, and many more.[9-11] Some examples of bioactive functions of peptides are given in Table 2.1.

Table 2.1 The diversity of bioactive peptide functions

Peptide Function	Effect (example)	Peptide (example)	Ref.[a]
Neurotransmitter, Ion channel gating ligand	Neuronal signal transduction	Vasopressin	12
Hormone, Growth factor	Cellular signalling	Glucagon	13
Neurotoxin	Paralysis	Cobratoxin	14
Antifungal	Immunosuppression	Cyclosporine A	
Antioxidant	Inhibition of cellular oxidation processes	Glutathione	15,16
Antimicrobial	Killing or inhibition of microorganisms	α-Defensins	17
Antihypertensive peptides	Enzyme inhibition	Angiotensin	16

[a] References for further reading.

Arg[8]-vasopressin, oxytocin, urotensin II and urotensin-related peptide are examples of natural cyclic peptide hormones found in humans. They mainly perform their function by activating GPCRs.[18-23] The proposed mechanism for this agonism includes interaction with the cell surface,[24,25] the extracellular loops and intrusion of the peptide ligand into the transmembrane binding pockets of its cognate receptors, where a signal is triggered (for references, see Table 2.2). However, more complicated reaction paths have been discussed, including multiple conformations,[26,27] upstream complexes[21,28] and biased agonism.[29]

Table 2.2 lists the corresponding receptors and main physiological functions of the peptides in the focus of this investigation, and gives references for reviews and further reading. Deamino-oxytocin (dOT, 1-(beta-mercaptopropionic acid)-oxytocin) and carbetocin (CT, (1-butanoic acid-2-(O-methyl-L-tyrosine)-1-carbaoxytocin) are synthetic analogues of OT. CT is an approved pharmaceutical substitute of OT with a considerably longer half-life.[30] It is used for the treatment of excessive postpartum bleeding after Caesarean sections.[31] Deamino-OT demonstrates superagonistic activity

toward the OT receptor but is not used pharmaceutically. It was the first crystal structure determined for the aforementioned peptides.[32]

Table 2.2 Physiological function and receptors of AVP, human UII, human URP, OT and the synthetic analogue CT

Peptide	Sequence [a] Physiological function (examples)	Receptor	References [b]
AVP [c]	[CYFQNC]PRG$_{NH2}$	V1aR, V1bR, V2R	20,22,33-38
	Antidiuretic, antipyretic, regulation of blood pressure, regulation of social and sexual behaviour		39-42
OT [d]	[CYIQNC]PLG$_{NH2}$	OTR	19,37,38,43
	Milk ejection, uterotonic activity, regulation of social and sexual behaviour		41,42,44,45
CT [e]	[(BuA)(YMe)IQNC]PLG$_{NH2}$	OTR	46
	Prevents postpartum bleeding [f]		30,31
UII, URP	ETPD[CFWKYC]V, A[CFWKYC]V	UTR (= UTS2R)	21,23,47,48
	Vasoconstrictive (cardiovascular homeostasis)		49

[a] Cyclic motif in square brackets. [b] Reviews and further reading. [c] AVP is also a partial agonist to OTR. [d] Trivial names: "trust, cuddle, love hormone". [e] Synthetic analogue of OT. [f] Intravenous application. Abbreviations: see p. xii.

Pharmacology. The diversity of their bioactive function makes peptides attractive for their pharmaceutical potential.[8,11,50] GPCR targeting drugs share approximately 25-40 % of the global market.[26] However, traditional drugs are small molecules and orally bioavailable, stable against digestion and able to cross membranes. These are all properties that are commonly not present in peptides. Bio-drugs often need to be delivered by injection (*e.g.* CT), they are usually metabolically unstable and show poor membrane permeability. Nevertheless, they outclass traditional small-molecule drugs, demonstrating high specificity for their targets, high potency and low side effects.[50,51] In 2010, 100 peptide-based drugs were registered holding approximately 10% of the therapeutics market with an increasing share.[50] Most of these therapeutics are peptide hormones[i] with chain lengths of 8-10 residues and cyclic peptides are especially interesting because of higher resistance against proteolytic degradation and increased bioavailability.[11,50] Furthermore, it is assumed for cyclic peptides that their conformational flexibility, in combination with the ability to build intramolecular hydrogen bonds (reduction of hydrophilic surface) may facilitate membrane crossing.[52] Synthetic therapeutic peptides derived from AVP include argipressin (I), desmopressin acetate (II), lypressin (III), and phenypressin (IV), indicated for the treatment of diabetes insipidus (I, II, III), enuresis[ii] (II), Cushing´s syndrome (III), stomatitis and pharyngitis[iii] (IV).[11] Those derived from OT include carbetocin, mentioned above, and atosiban acetate. The latter is used as antagonist (tocolytics[iv]) for the treatment of premature contractions.[11] CT and the mentioned AVP

[i] or derivatives with hormone function
[ii] uncontrolled wetting
[iii] inflammation of mouth/lips and back of the throat, respectively
[iv] labour suppressants

derivatives act agonistically and an activation of 5-20 % of the receptors is sufficient to be effective.[i] UII and URP or their derivatives are not yet used as drugs.

The Structure of Peptides

Conformation. In 1874, the postgraduate student J. H. van´t Hoff (who 26 years later became the first winner of the Nobel prize in Chemistry)[53] made a "*suggestion looking to the extension into space of the structural formulas at present used in chemistry*"[54] and proposed that the four covalent CH bonds of methane (CH_4) adopted a tetrahedral spacial orientation. Though strongly criticised as "childish fantasy" by his contemporary H. Kolbe,[55] van´t Hoff´s idea prevailed. It introduced the third dimension to chemistry and started the fields of stereochemistry and conformational analysis. Conformation is generally understood as the "*arrangement of atoms in a molecule obtained by rotation about one or more single bonds*".[56] However, how can conformation best be described? 2D structural formulae (*e.g.* Fischer projections) give a limited indication of 3D structure but are inadequate to represent the 3D arrangement of macromolecules. A precise definition is given by the Cartesian coordinates of the atoms, but this is too detailed a view for many purposes and needs computational visualisation to be imaginable and figurative names are often used to communicate conformational shapes (*e.g. boat* or *chair* for the 3D structure of cyclohexane).

The prime mover in facilitating the *lingua franca* of protein structure has been decades of X-ray, followed later by neutron, diffraction crystallography and a notation using four structural levels is established. The *primary structure* refers to the defined sequence of amide-bond connected amino acids also called residues. Names of the amino acids are abbreviated by 1- or 3-letter codes. The *secondary structure* describes conformational segments denoted as *secondary structure motifs*. Typical *secondary structure* motifs are *turns* and *helices*, which are illustrative descriptions of the local conformation. They are defined by distinct sequences of the dihedral angles (Φ and Ψ) of the backbone C^{α} atoms and energetically favoured regions for ΦΨ combinations can be visualised with Ramachandran plots.[57] A *helix* describes a periodically repeatable motif, other periodic structure motifs are *e.g.* β-*sheets.* These structures are generally stabilised by repetitive hydrogen bonds and are further characterised by the number of C^{α} atoms involved in the periodic motif. A *turn* characterises a single conformational motif and a motif with consecutive turns is denoted as a *multiple turn*. The overall folding of all the secondary structure elements in a polypeptide chain defines the *tertiary structure*. A fourth structural level, *quaternary structure*, is applied if

[i] antagonists need to occupy at least 50% of all receptors to ensure therapeutic effectiveness

polypeptide chains build complexes of multiple subunits. For further basic details, the reader is referred to biochemical textbooks (*e.g.* Stryer´s "Biochemie"[4]). For a general taxonomy of protein structure, reference is made to Richardson.[58]

In the context of this work, only the secondary structure elements that are relevant for the investigated peptides will be described in detail. Arg[8]-vasopressin and the structurally related peptides exhibit rather short chain lengths of 8 to 11 amino acids and the major motifs of these cyclic peptides are *turns*. Descriptions of the structure of these peptides in solution in terms of *turns*, *turn centres* and *turn types* are very common in the literature (*cf.* Tables 2.3, 2.4, 2.6). The classical turns are known as γ- and β-*turns*. A γ-*turn* comprises three C^{α} atoms and a β-*turn* four C^{α} atoms with characteristic *turn centres* (C^{α}_{i+1} for a γ-*turn*, C^{α}_{i+1} and C^{α}_{i+2} for a β-*turn*), where the "course" of the chain is reversed. The classical β-*turn* includes a hydrogen bond between the carbonyl oxygen of residue *i* and the amide proton of residue *i+3,* and the ΦΨ torsions at the centre C^{α} atoms define the *turn type*. A table of classical *turn types* with further explanations is given (see Appendix A5). The classical terminology, however, falls a little short when considering the dynamical nature of small peptides in solution or in complexes with their receptor, which now can be observed by MD simulation. In this work, the term "*classical hydrogen-bonded turn"* was extended to "*open turn"* to take into account the conformational fluctuation of the peptides in solution and a detailed definition is given in the Supporting Information of Paper 3[3] (see Appendix A3, p. A46). The periodic motifs *3_{10}-helix*, *parallel sheet* and *anti-parallel sheet* were also found in conformations of the cyclic oligopeptides in this study. However, periodic motifs emerge only to some extent and without long-range repetition due to the short peptide-chain lengths. A *3_{10}-helix* motif comprises three residues and a hydrogen bond from the carbonyl oxygen of residue *i* to residue *i+*3 (in contrast to *i+4* for *α-helices*) resulting in a ring motif of 10 atoms;[58] the ΨΦ torsions of this motif fluctuate around -49° and -26° (*cf.* Appendix 5, Table A5.1). *Parallel* and *anti-parallel sheets* are subtypes of β-*sheets*, where adjacent amino-acid sequences are connected *via "evenly spaced"*[58] hydrogen bonds with either ascending (*parallel*) or descending (*antiparallel*) consecutive order.[58] While it is debatable whether it makes sense to use periodic structure elements for short sequence lengths, each specified secondary structure element matches the required torsion angles and/or hydrogen bonds and a notation as *helix* or *sheet* indicates a tendency to periodic motif repetition.

In addition and in the spirit of communicative structure notation, most of the main conformations identified in the long-scale MD simulations of this work were given names that describe their three-dimensional shape, *e.g. saddle*, *scoop*, *twisted saddle* and *omega* (shape of the Greek letter).

Relevance. To understand the functionality of peptides, it is necessary to understand their structure because of the interdependence of conformation with function. Distinct *single conformations* are the basis of the traditional Lock-and-Key hypothesis (LKH)[59] and allosteric models.[60-65] The LKH is the simple view that a ligand and receptor need to fit together to trigger a response, requiring a specific "bioactive" conformation of the ligand. Allostery describes the possibility of conformational changes of ligand and/or receptor caused by mutual contact (or contact with further components) on the reaction path that triggers a biological function.[61,63,65,66] However, more recent models consider *flexibility* and even *disorder* or, in other terms, conformational dynamics as a functional feature: This implies the possibility that the bioactive conformation is not the lowest-energy conformation but a minor populated state that is selected from an *equilibrium of multiple conformations* and reproduced by population shift.[67-69] All polypeptide chains are intrinsically flexible[67,68,70] and exist as equilibria of a vast range of conformations. The thermodynamic stability determines the relative populations of the conformations and the barrier heights determine the timescale of interconversions (kinetics).

Restrictions. There are several factors that restrict conformational freedom and promote distinct conformations.[71] The amino acids of peptide chains are linked by peptide bonds. The rotation of a peptide bond is limited by its double bond character and a *trans* conformation of the adjacent residues is thermodynamically favoured in native peptides.[4] The degrees of freedom increase with the sequence length. AVP, UII and other peptides studied here have relatively short sequence lengths of 8-11 residues. The chemical or stereochemical nature of the sidechains will also affect the 3D structure of peptides *via* attractive or repulsive interactions along with conformational rigidity of residues themselves (*e.g.* Pro). Examples of decisive sidechain interactions include hydrogen bonding, π-π stacking of aromatic rings (Phe, Tyr, Trp), steric hindrance of bulky residues (*e.g.* Arg, Lys), or covalent binding (*e.g.* Cys, disulphide bridge). Cyclisation is a further consideration constraining conformational diversity. The smaller the cyclic chain, the higher the constraint of conformations.[71] AVP, OT, UII, URP and dOT comprise a 6-residue ring closed by a disulphide bond forming a 20-membered macrocycle. They are examples of flexible peptides that nevertheless may exhibit defined conformations due to their cyclic restriction (see below). Besides the intrinsic properties, the environment will also affect possible conformations. Polar solvents (*e.g.* H_2O) favour chemical exchange (*e.g.* between amide protons and solvent), which may hamper the formation of defined secondary-structure motifs, which is why less polar solvents (*e.g.* dimethyl sulphoxide (DMSO)) facilitate them. A peptide in solution certainly has more conformational freedom than a peptide in contact with the cell-surface, an outer binding site or a peptide "trapped" in a binding

pocket. Consequently, a peptide conformation deduced from a crystal structure cannot be readily transferred to the situation in solution and *vice versa*.

Analysis Methods. Classical experimental methods of structure determination include X-ray crystallography and NMR spectroscopy, accompanied by other spectroscopic methods, *e.g.* circular dichroism spectroscopy (CD). Computational methods are summarised under the term *molecular modelling* and range from quantum chemistry to molecular mechanics (MM), including MD simulations. They are able to provide atomistic insight into the conformational space of molecules, here, the peptides considered. The methods used in this thesis are discussed in Chapter 3.

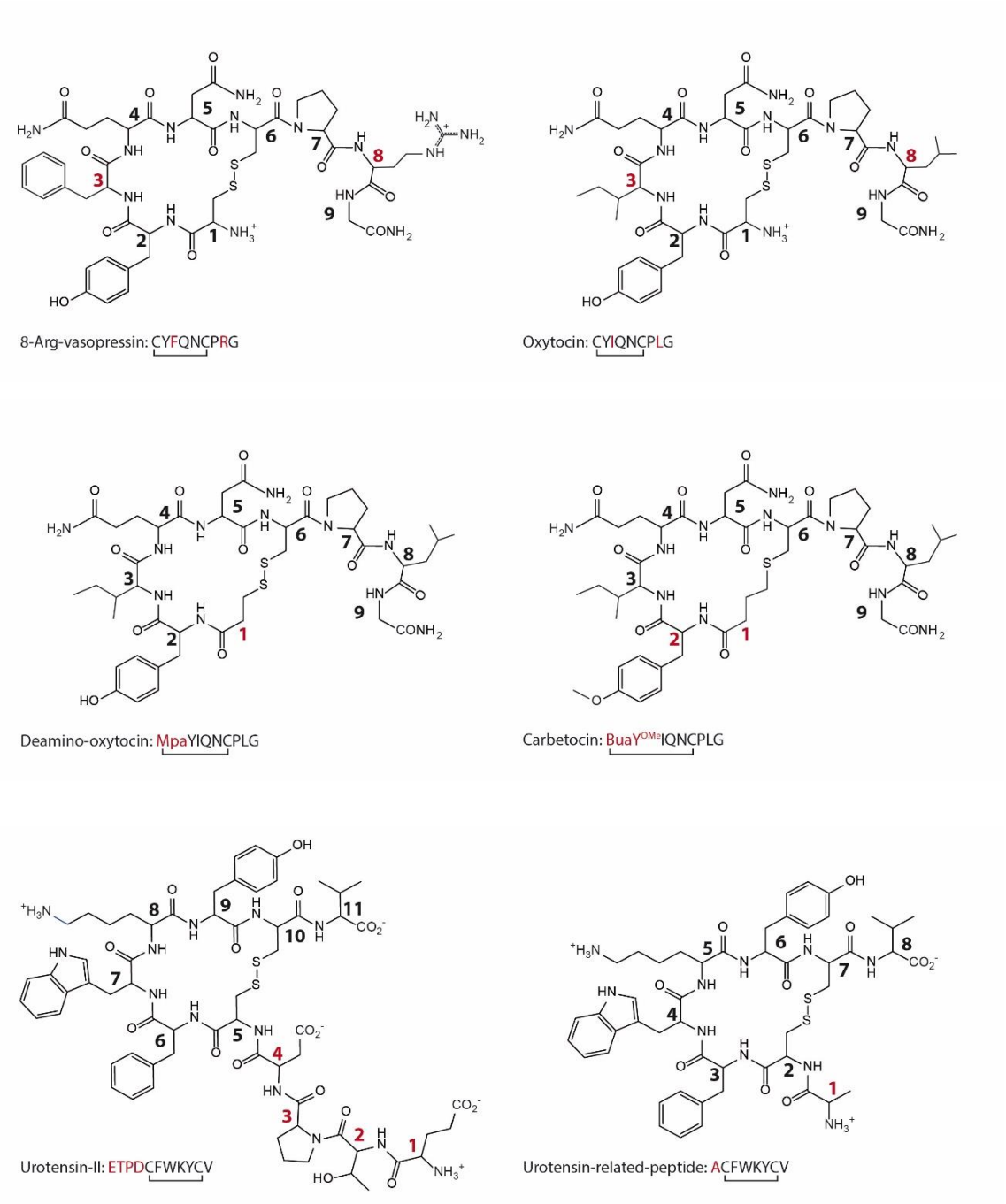
Structural Data for AVP, OT, UII, URP, dOT, and CT

The structural determination of the above peptides started in 1953 when du Vigneaud *et al.*[72] were able to identify the amino-acid sequence of oxytocin. This was the first time that the motif of a 6-residue ring closed by a disulphide bridge was seen in nature, later also identified for vasopressin, urotensin, urotensin-related peptide (Scheme 2.1) and others (*e.g.* insulin).

A literature review of structural information for OT, dOT, AVP, UII and URP from 1960s until today with focus on the ring and tail conformations is given (see Tables 2.3 to 2.6), including results from this work. Table 2.5 lists additional conformational aspects. The conformations from the literature are assigned to the main conformational ring types defined in this work where possible. These ring types are discussed in detail in Chapters 4 (Paper 1), 6 (Paper 3) and 7 (Related Peptides and General Conformational Classification).

The conformational data can be summarised as follows:

Oxytocin and Deamino-Oxytocin. After the determination of OT´s amino-acid sequence in 1953 by du Vigneaud,[72] first crystallographic data of dOT followed in 1964 to 1966.[73-75] Deamino-oxytocin, which only lacks the N-terminal amino group (Scheme 2.1), was considered as model for the 3D structure of OT, although dOT is twice as potent[74] at the OT receptor. The first complete X-ray structure of dOT was published in 1986 by Wood *et al.*,[32] refined by Husain *et al.* in 1995 (PDB ID: 1XY1 and 1XY2, Fig. 2.1).[76] The two crystal forms found for dOT are both characterised by a 3,4 β-II turn and hydrogen bonds Tyr^2O-Asn^5H and Asn^5O-Tyr^2H in the ring and a 7,8 β-III turn of the tail with a hydrogen bond between Cys^6O and Gly^9H.



Scheme 2.1 Primary structure of the natural peptide hormones AVP, OT, UII, URP, and the artificial analogues dOT (1-(beta-mercaptopropionic acid)-oxytocin) and CT (1-(butanoic acid)-2-(O-methyl-Tyr)-1-carbaoxytocin). The main sequence differences (AVP/OT/dOT/CT and UII/URP) are highlighted.

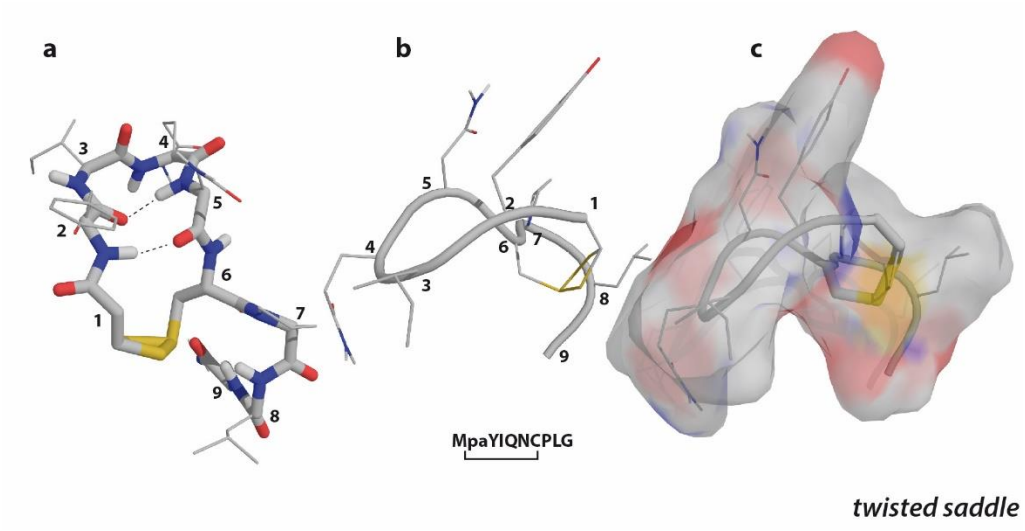


Figure 2.1a-c Deamino-oxytocin, PDB ID: 1XY1. dOT crystallises in two forms, "wet" (space group C2, PDB ID: 1XY1, shown) and "dry" (space group $P2_1$, PDB ID: 1XY2). Both forms show the same ring conformation. The backbone shape resembles the ring-state type *twisted saddle (= folded-IVb2).* Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, transannular hydrogen bonds (dotted lines), residue numbers labelled; (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) all atoms as spheres, perspective like **b**.

In 1996, Rose *et al.*[77] published the X-ray structure of OT bound to its carrier protein neurophysin (NP) (PDB ID: 1NPO, Fig. 2.2). Like dOT, OT_{NP} exhibits a β-turn at residues Tyr^3 and Gln^4 but of different turn type (β-III). The tail of OT_{NP} crystallises in two forms: a *folded* conformation with 7,8 β-turn and an *extended* conformation with a hydrogen bond from Pro^7O to Gly^9H_{NH2}.

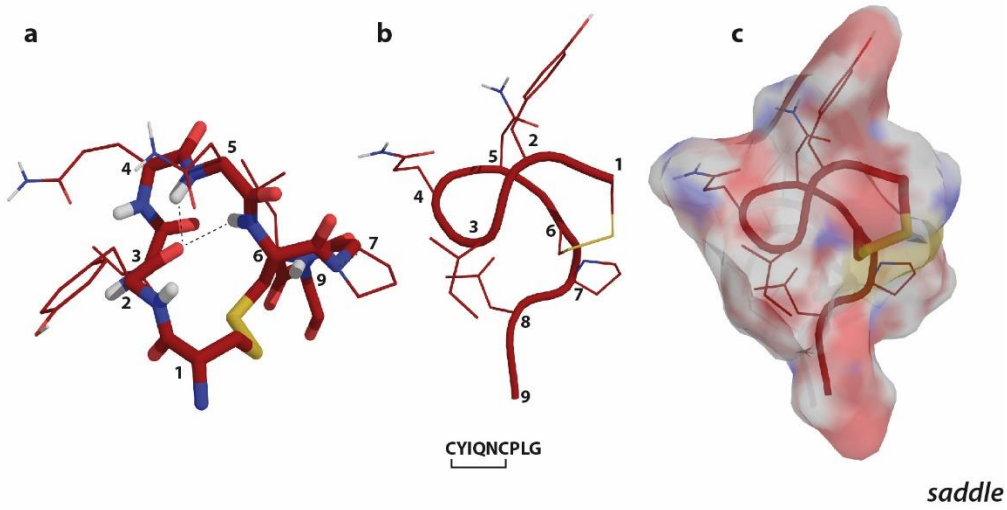


Figure 2.2a-c Oxytocin, PDB ID: 1NPO. Crystal structure of OT bound to its carrier protein neurophysin (not shown). The NP-OT complex crystallises as dimer. Both OT molecules show the same ring conformation but differ in their tail conformation, *extended* (shown) and *folded* (7,8 β-turn), respectively. The backbone shape resembles the ring-state type *saddle (= folded-I)*. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, transannular hydrogen bonds (dotted lines), residue numbers labelled; (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) surface.

However, it was NMR spectroscopy that led to the first plausible 3D structure descriptions of OT as early as 1971.[78] An overview on proposed conformations of OT is given in Table 2.3.

Table 2.3 Ring and tail conformations of OT and dOT (literature review)

Method	Ring conformation	Turns and hydrogen bonds [a]	Ring-state type assignment [b]	References
		OT in H_2O		
MD (50 µs)	*	***folded conformations***	**folded**	Haensele 2017(Chap. 7)
		(1) 3,4,5 multiple turn (3,4 β-I + 2O5H, 2O6H)	**(1)** *saddle*	
		(2) 3,4,5 multiple turn (3,4 β-II + 2O5H)	**(2)** *tws*	
		(3) 1-5 3_{10}-helix (3,4 β)	**(3)** tws_{helix}	
		open conformations	**open**	
		(4) open distorted 4,5 β-VIII/I	**(4)** *clop*	
		(5) open distorted 4,5 β-II	**(5)** $clop_{45pbr}$	
		(6) open, no classical turns	**(6)** *open*	
		Tail: extended and folded		
NMR	*folded (*)*	**(1)**[j] 3,4 β	**(1)** *saddle*	Koehbach 2013[79]
		(*) NMR ensemble, no H-bonds	**(*)** *(saddle/clop)*	
NMR	*folded*	**(1)** 3,4 β + 2O5H, 2H5O	**(1)** *tws*	Ohno 2010[80]
		7,8 β + 6O9H		
EDMC, MD (400 ps)	*	**(1)** 3,4 β	**(1)** *(tws)*	Liwo 1996[81]
		(2) 2,3 β	**(2)** *(open)*	
		(3) 4,5 β [c]	**(3)** *(saddle/clop)*	
NMR	(*)	no transannular H-bonds		Meraldi 1976[82]
				Glickson 1976[83]
				Brewster 1973[84]
		OT in DMSO		
NMR	*folded*	**(1)** 3,4 β-II + 2O5H, 2H5O	**(1)** *tws*	Budesinsky 2005[85]
		7,8 β + 6O9H		Bhaskaran 1992[86]
NMR	(*)	**(1)** no classical turn at 3,4	**(1)** (*tws*)	Kato 1993[87]
		no transannular H-bonds		
NMR	(*)	**(1)** 3,4 β + 2O5H	**(1)** *tws*	Brewster 1973[88,89]
		(2) 1-5 ?-turn + 1O5H	**(2)** (tws_{helix})	
		(3) $4H_{carbamid}5H$	**(3)** -	
		Tail: folded		
NMR	*folded*	**(1)** 3,4 β + 2O5H, no 2H5O	**(1)** *(tws)*	Urry 1970[90]
		7,8 β + 6O9H		Urry 1971[78]
NMR	(*)	(assignment)		Johnson 1969[91]
		OT-NP complex		
X-ray [d]	*folded*	**(1)** 3,4 β-III + 2O(4H,5H,6H), 3O5H	**(1)** *saddle*	Rose 1996[77]
		open 7,8 β and 7-9 γ + 7O9		
NMR	*(folded)*	**(1)** Tyr^2O *exo* toward ring	**(1)** *clop* [e]	Lippens 1993[92]
		OT *in vacuo*		
MM/MD	(*)	**(*)** variants of dOT X-ray conformation	**(*)** tws_{helix} [f]	Ward 1991[93]
PE	(*)	**(*)** 3,4 β (+ 2O5H)	**(*)** tws_{helix} [g]	Nikiforovich 1979[94]
PE	(*)	**(*)** no transannular H-bonds		Kotelchuck 1972[95]
		Tail: *folded* possible		
		dOT in DMSO		
NMR	*folded*	**(1)** 3,4 β + 2O5H, 2H5O	**(1)** *tws*	Urry 1970[90]
		Tail: *folded*		
		dOT crystals		
X-ray [h]	*folded*	**(1)** 3,4 β-II + 2O5H, 2H5O	**(1)** *tws*	Husain 1990[76]
		7,8 β-III + 6O9H		Wood 1986[32]
X-ray	-	(1^{st} X-ray data)[i]		Low 1966[75]

*Conformational flexibility or dynamic equilibrium of multiple conformations. (*)Conformational flexibility suggested. [a] Hydrogen bonds are denoted by the residue number and the donor and acceptor atom (O carbonyl oxygen; H amide hydrogen). [b] Assignment to ring-state types defined in this work and based on turn description or circular similarity of torsion angles (*cf.* Appendix A4); parentheses indicate tentative assignments. [c] Less populated than 3,4 and 2,3 turn; no unequivocal assignment possible (no torsion published). [d] PDB ID: 1NPO (dimer; different tail conformations for OT). [e] Circsim (*ref vs. clop*) = 68 %. [f] Circsim (*ref vs.* tws_{helix}) = 51 %. [g] Circsim (*ref vs.* tws_{helix}) = 52 %. [h] PDB ID: 1XY(1,2). [i] Cell dimension, space group, density. [j] PDB ID: 2MGO, circsim (*ref vs. saddle*) = 74%. Abbreviations: see p. xii.

The data in Table 2.3 may be summarised: NMR experiments of OT in DMSO suggest β-turns centred at residues 3 and 4 stabilised by hydrogen bonds (Tyr^2O-Asn^5H, Tyr^2O-Cys^6H, and/or Asn^5O-Tyr^2H). Nonpolar solvents like DMSO appear to favour a compact conformation with *folded* tail (7,8 β-turn and hydrogen bond Cys^6O-Gly^9H) but conformational flexibility is not excluded and interconverting β-turn types are assumed. In polar solvents, such as H_2O, the transannular hydrogen bonds are not evident inferring high conformational flexibility. However, in 1993, Kato *et al.*[87] pointed out that in DMSO none of the published NMR structures would be consistent with all of their observed NOE data and that "*for small peptides such as oxytocin […], problems of molecular flexibility or multiple conformers are very serious"*. The existence of multiple low-energy conformations was further suggested by early MM calculations[81,94] and Liwo *et al.*[96] in 1989 proposed a classification of ring conformations for dOT and Deamino-Arg^8-vasopressin (dAVP) by analogy with cyclohexane conformations; this was not developed further.[i]

Arg^8-Vasopressin. AVP differs from OT in positions 3 and 8 (*cf.* Scheme 2.1). Ile^3 becomes Phe^3, giving the possibility of π-π interaction with the neighbouring residue Tyr^2 and the hydrophobic Leu^8 changes to the sterically demanding hydrophilic residue Arg^8. Conformational diversity also applies to AVP (see Table 2.4). The X-ray structure of AVP bound to its carrier NP has yet to be published. However, an X-ray structure of AVP bound to trypsin exists (PDB ID: 1YF4, Fig. 2.3).[97]

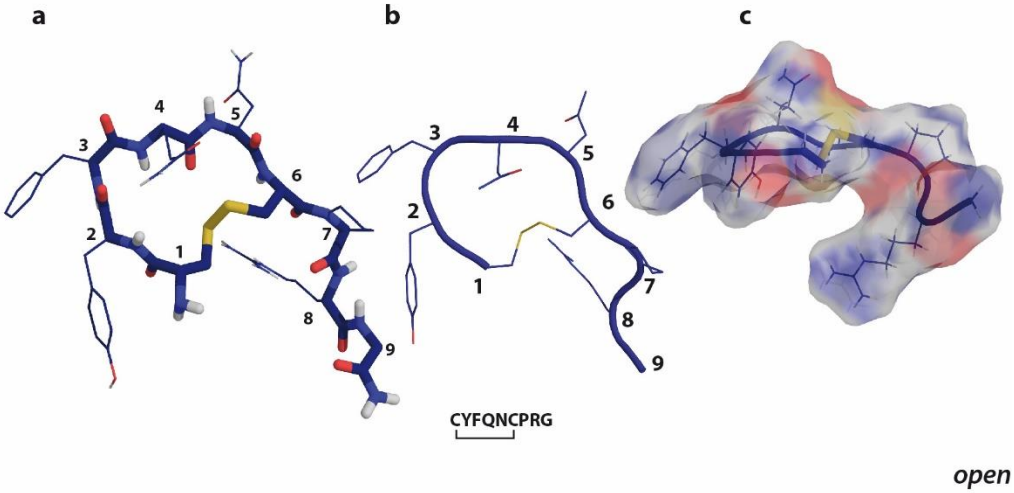


Figure 2.3a-c Arg^8-vasopressin, PDB ID: 1YF4. Crystal structure of AVP bound to enzyme trypsin (trypsin not shown). The backbone shape resembles the ring-state type *open* (= *lasso)*. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, residue numbers labelled; (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) surface.

[i] They assigned low-energy (MM calculations) ring conformations (defined by Cα positions) of dOT and dAVP to cyclohexane conformations (*boat*, *chair*, *twist*, *sofa* with 26 subcategories). The idea was not developed further, superseded by the clearer sequential notation with secondary-structure elements.

This shows an *unfolded* (*open*) conformation significantly different to the *folded* β-turn conformations found for OT_{NP} (*cf.* Fig. 2.2). However, the X-ray structure of the NP complex of Lys^8-VP (LVP, PDB ID: 1JK4, porcine vasopressin)[98] is very similar to OT with β-turns at 3,4. NMR experiments for AVP (see Table 2.4) suggest conformations with β-turns at residues 3,4 and/or 4,5 of different fast interchanging types and possible transannular hydrogen bonds, depending on the polarity of the solvent. AVP, like OT, is suggested to be a flexible molecule able to adopt multiple conformations[99,100] even more flexible than OT.[101]

Table 2.4 Ring and tail conformation of AVP (literature review)

Method	Ring conformation	Turns and hydrogen bonds [a]	Ring-state type assignment [b]	Reference
		AVP in H_2O		
MD+NMR (23 µs)	**70 % : 30 %** ***folded*:*open***	***folded conformations*** **(1)** 3,4,5 multiple turn (3,4 β-I + 2O5H, 2O6H) **(2)** 3,4,5 multiple turn (3,4 β-II + 2O5H) ***open conformations*** **(3)** open distorted 4,5 β-VIII/I **(4)** open, no classical turns Tail: *extended* and *folded*	***folded*** **(1)** *saddle* **(2)** *tws* ***open*** **(3)** *clop* **(4)** *open*	Haensele 2014[1(Chap. 4)] Haensele 2016[2(Chap. 5)]
REMD (50 ns)	*folded*[i]	**(1)** 3,4 β-III + 2O5H, 2O6H Tail: *extended* and *folded*	**(1)** *saddle*	Yedvabny 2014[102]
NMR	(*)	**(1)** open 3,4 β-II + 4,5 β-III' **(2)** open 3,4 β-II + 4,5 β-I'[ii]	**(1)** (*tws*) **(2)** (*tws*)	Sikorska 2008[103]
EDMC, MD [c] (400 ps)	*	**(1)** 3,4 β **(2)** 4,5 β **(3)** 2,3 β [d]	**(1)** (*tws*) **(2)** (*saddle*/*clop*) **(3)** (*open*)	Liwo 1996[81]
		AVP in DMSO [g]		
NMR	(*)	**(1)** 3,4 β-II (or I') **(2)** 3,4 β-II' **(3)** 3,4 β-I Tail: ≥ 2 conformers	**(1)** *tws* **(2)** – **(3)** *saddle*	Schmidt 1991[99]
NMR	(*)	**(1)** 3,4 β + 2O5H	**(1)** -	Walter 1974[104]
		AVP in micelles [e]		
NMR (DPD)	*folded*	**(1)** 3,4 β-II (or VII) + 4,5 β-I' (or IV) + 2O6H 70 % 6,7 β-I + 6O8H	**(1)** (*tws*)	Lubecka 2015[105]
NMR (SDS)	*folded*	**(1)** 3,4 β-II' + 4,5 β-I + 3O6H + 5,6 β-IV *cis*-peptide bond Cys^1-Tyr^2 Tail: γ or β-turn	**(1)** -	Rodziewicz 2008[100]
		AVP-trypsin complex		
X-ray	*open*	**(1)** 2O4H + 4O6H[iii]	**(1)** *open*	Ibrahim [97]

*Conformational flexibility or dynamic equilibrium of multiple conformations. (*)Conformational flexibility suggested. [a] Hydrogen bonds are denoted by the residue number and the donor and acceptor atom (O carbonyl oxygen; H amide hydrogen). [b] Assignment to ring-state types defined in this work and based on turn description or circular similarity of torsion angles (*cf.* Appendix A4); parentheses indicate tentative assignments. [c] With "hydration-shell". [d] Less populated than 3,4 and 4,5 turn; no unequivocal assignment possible (no torsion published). [e] Spherical aggregation of lipid molecules, "membrane mimic". Abbreviations: see p. xii.

[i] Only the ring-state type *saddle* was found (both for OT and AVP). The REMD simulation appears not to be converged and conclusions have to be considered with caution.

[ii] Remark: Long-scale MD simulations show that in solution a fluctuation of ±30° around an ideal turn torsions can be assumed. Thus, β-I' (+60° +30° +90° 0°) and β-III' (+60° +30° +60° +30°) turns are not distinguishable. Sikorska´s two conformations, (**1**) and (**2**), belong to the same conformational main type.

[iii] Cannot be verified: 2O4H = 5.0 Å, 4O6H = 3.6 Å (PyMOL)

In 1996, Liwo *et al.*[81] published the first comprehensive MD study (EDMC and Monte Carlo, total simulation time 400 ps) of OT and AVP and proposed conformations for both AVP and OT with β-turns centred at 2,3, 3,4 and 4,5. They predicted a prevalence of 3,4 and 4,5-turns for AVP, and 3,4 and 2,3-turns for OT.

Further experimental methods, *e.g.* CD and Raman spectroscopy, complement the structure elucidation of OT and AVP. The main results are listed in Table 2.5 and can be summarised as follows:

1) The ring-closing disulphide bridge is suggested to adopt conformations of right-handed chirality and a dihedral angle around ±90° (stddev 30°). (2) Non-covalent attractive interactions (π-π stacking) of Tyr^2 and Phe^3 in AVP are very likely. (3) The C-terminal tails of AVP and OT are more mobile than the ring. (4) AVP and OT show proline *cis/trans* isomerisation with approximately 5-10 % *cis*-Pro^7.

Table 2.5 Additional structural properties for OT, AVP, dOT, and UII (literature review)

Peptide	Conformational data	Meth. [a]	Reference
	Flexibility		
OT in H_2O	• OT is less flexible than VP (LVP)	A	Gryczynski 1991[101]
OT in H_2O	• flexible backbone	B, C	Hruby 1978[106]
	Secondary structure		
OT in H_2O	• β-turn like (ring)	C	Tu 1978[107]
AVP in H_2O	• β-turn, β-sheet, and random-coil bands	C	Podstawka 2006[108]
$UII_{(4-11)}$	• disordered conformers of random coil, turn and β-structures	C	Carotenuto 2004[109]
	Disulphide bridge		
OT in H_2O	• right-handedness, g-g-g > g-g-t, t-g-t	F,G	Pazderkova 2012[110]
	• right-handed chirality, ∢CSSC = 110-115°	B, C	Hruby 1978[106]
	• OT (and dOT) ∢CSSC = right-handed helical, distorted	B, E	Urry 1968[111]
	• g-g-g	C	Tu 1978[107]
OT in DMSO	• ∢CSSC = ±90° (stddev 30°)	B, C	Maxfield 1977[112]
AVP in H_2O	• right-handedness, g-g-g > g-g-t, t-g-t	F,G	Pazderkova 2012[110]
	• g-g-g and t-g-t	C	Podstawka 2006[108]
	• g-g-g possible	B, C	Tu 1979[113]
	Tail		
OT in H_2O	• above ring	A	Gryczynski 1991[101]
	• no evidence for noncovalent ring/tail interaction	D	Cowburn 1983[114]
	• higher flexibility than ring	D	Deslaurier 1974[115]
OT in DMSO	• more flexible than ring	D	Bhaskaran 1992[86]
AVP in H_2O	• no evidence for noncovalent ring/tail interaction	D	Cowburn 1983[114]
	• *folded* (above ring)	B	Fric 1975[116]
AVP in DMSO	• higher mobility than ring (AVP, OT); higher mobility than in OT	D	Walter 1974[104]
	π-π interaction		
AVP in H_2O	• π-π stacking of Tyr^2 and Phe^3	A	Szmacinski 1996[117]
	• possible but no major interaction	D	Cowburn 1983[114]
	• π-π stacking of Tyr^2 and Phe^3	B	Fric 1975[116]
	• higher local rigidity at Tyr^2 counts for π-π stacking	B	Fric 1975[116]
AVP in DMSO	• no π-π stacking of Tyr^2 and Phe^3	D	Schmidt 1991[99]

Table 2.5 continued

Peptide	Conformational data	Meth. [a]	Reference
	Pro^7 *cis/trans* isomerisation		
AVP in H_2O	• ~5 % *cis*	D	Sikorska 2008[103]
	• ~9 % *cis*	D	Larive 1992[118]
	• isomerisation *via* twisted Cys^6-Pro^7 imide bond	D	Larive 1993[119]
OT in H_2O	• ~10 % *cis*	D	Larive 1992[118]
	• ~0 % *cis*	D	Glasel 1973[120]
UII in H_2O	• ~11 % *cis*	D	Haensele [unpublished]
	Tyr^2		
OT in H_2O	• more shielded from solvent than in LVP	A	Gryczynski 1991[101]
	• exposed to solvent	C	Tu 1978[107]
AVP in H_2O	• exposed to solvent	B, C	Tu 1979[113]
	Sidechain conformation		
OT, AVP in H_2O	• no significant differences	D	Cowburn 1983[114]

[a] Spectroscopy methods: Fluorescence Anisotropy (A), Circular Dichroism (B), Raman (C), Nuclear Magnetic Resonance (D), UV (E), Vibrational Circular Dichroism (F), Raman Optical Activity (G).

Urotensin and Urotensin-Related Peptide Whereas OT and AVP are veterans in the field of structure determination, UII and URP are relatively new research objects. In 1999, UII was detected in mammals, including humans,[121-123] and four years later, its paralogue URP was identified[124] (*cf.* Introduction of Paper 3 (Chap. 6)). UII has a 6-membered cyclic ring and a disulphide bridge in common with OT, dOT and AVP but a very different sequence and its 4-residue tail is in N-terminal position instead of the C-terminal tail of AVP and OT/dOT (*cf.* Scheme 2.1). URP has the same ring sequence as UII but lacks the multi-residue tail. Conformational data are rare compared to OT and AVP and to date there are no X-ray structures of UII or URP. Structure descriptions deduced from spectroscopy experiments vary from distinct single conformations[125,126] with preferred turn centres at residues Lys and Tyr (8,9 for UII and 5,6 for URP) to unstructured/flexible[109,127] (*cf*. Table 2.6). Residues 8,9 of UII and 5,6 of URP correspond to centres 4,5 in OT, dOT, AVP, and CT.

Table 2.6 Ring and tail conformations of UII and URP (literature review)

Method	Ring conformation	Turns and hydrogen bonds [a]	Ring-state type assignment [b]	Reference
		UII in H_2O		
MD+NMR (37.8 µs)	**28 % : 72 %** ***folded:open***	***folded conformations*** **(1)** 7,8,9 (7,8 β-I) **(2)** 7,8,9 (7,8 β-II) **(3)** 6,7,8 (5-9 helix) **(4)** 7,8,9 (6-10 p-sheet) **(5)** 6,7,8 (6,7 β-III') ***open conformations*** **(6)** 8,9 β-I/VIII **(7)** 8,9 β-II **(8)** 6,7 β-I	***folded*** **(1)** *folded-I* **(2)** *folded-IVb2* **(3)** *inv-folded* **(4)** *folded-II* **(5)** *folded-III* ***open*** **(6)** *omega-I* **(7)** *omega-II* **(8)** *scoop + lasso*	Haensele 2017[3] (Chap. 6)
NMR	*open*	**(1)** widened 7,8,9 γ + 8,9,10 γ + possible: 7O9H + 8O9H	**(1)** *omega*	Lescot 2007[126]
CD+NMR	(*)	**(*)** disordered conformers [c]		Carotenuto 2004[109]
NMR	unstructured	no classical turns, no hydrogen bonds		Flohr 2002[128]

Table 2.6 continued

Method	Ring conformation	Turns and hydrogen bonds [a]	Ring-state type assignment [b]	Reference
		UII in DMSO		
NMR	*unstructured*	no standard secondary structure Tail: 3,4 β-I possible	(1) *omega* or *folded-IVb2*[d]	Grieco 2002[129]
		UII in SDS		
NMR	*folded*	**(1)** β-hairpin (7,8-II') **(2)** flexible	**(1)** (*folded)* **(2)** -	Carotenuto 2004[109]
		URP in H_2O		
MD+NMR (22.8 μs)	**14 % : 86 %** ***folded:open***	***folded conformations*** **(1)** 4,5,6 γ **(2)** 2-7 antip. β-sheet (4,5 β-II) ***open conformations*** **(3)** 5,6 β-I/VIII **(4)** 5,6 β-II **(5)** 3,4 β-VIII	***folded*** **(1)** *hybrid* **(2)** *sheet* ***open*** **(3)** *omega-I* **(4)** *omega-II* **(5)** $lasso_{45pbr}$	Haensele 2017[3] (Chap. 6)
NMR	(*)	high structural flexibility, no hydrogen bonds		Brancaccio 2015[127]
NMR	*open*	**(1)** 4,5,6 γ' + 4O6H	**(1)** *omega-I* [e]	Chatenet 2004[127]
		URP in SDS		
NMR	*folded*	**(1)** β-hairpin (7,8-II')	**(1)** (*folded)*	Brancaccio 2015[127]

*Conformational flexibility or dynamic equilibrium of multiple conformations. (*) Conformational flexibility suggested. [a] Hydrogen bonds are denoted by the residue number and the donor and acceptor atom (O carbonyl oxygen; H amide hydrogen). [b] Assignment to ring-state types defined in this work and based on turn description or circular similarity of torsion angles (*cf.* Appendix A4); parentheses indicate tentative assignments. [c] UII(4-11) fragment. [d] Circsim (*ref vs. omega-I* or *folded-IVb2*) = 67 %. [e] Circsim (*ref vs. omega-I*) = 86 %. Abbreviations: see p. xii.

Carbetocin. CT resembles OT´s primary structure. The sulphur atom from Cys^1 of the disulphide bridge is replaced by a methylene group and the hydroxyl group of Tyr^2 is methylated. CT is an approved drug substitute for OT and a commercial peptide (Ferring Arzneimittel GmbH). It is the subject of several pharmacological studies[31,130,131] but X-ray or NMR data for CT are not publicly available. Here it is used to complement the MD studies of natural peptide hormones with a synthetic analogue.

Motivation of the Study

As has been shown, the peptides in the focus of this thesis are interesting because of their versatile physiological properties, which are closely related to their structure. However, even if numerous structural data are available, there is no consistent description of their conformational preferences due to their flexibility.

For this thesis, the conformational space of several peptides with the common motif of a 6-residue ring was extensively explored with unrestrained μs-scale MD simulations to identify their main conformational types. This finally led to a general classification in terms of *open* and *folded* ring-state types (Chap. 7). Subtypes of the two classes are defined by turn centres and hydrogen bonds,

and similarities between the peptides are pointed out. The tables of this chapter (Tables 2.3, 2.4, 2.6) anticipate the results of this classification by assigning the structures from the literature to the ring-state types defined in this work.

A consistent conformational description of these peptides may help clarify contradictory structure definitions found in the literature. Each main conformational type identified for the peptide hormones is a potential bioactive conformation that can be used directly for molecular docking to investigate the interaction with their cognate receptors and to design pharmacophores. The generic classification of conformations (Chap. 7) will facilitate the structural analysis of related peptides and the modelling of analogues for subsequent research toward defining their nature of modulation of their cognate receptors. With the advent of a better understanding of these peptides and their conformational preferences and clarification of interactions with their receptors this may lead to better non-peptide drugs with therapeutically useful modulatory properties.

A further aim of this thesis was to investigate the molecular flexibility of the peptides as expressed by their conformational equilibria. These equilibria are difficult to access experimentally if conformational interconversions are fast relative to the NMR timescale.[132,133] In this work, the equilibrium concentrations for AVP, UII, and URP in aqueous solution were determined *via* long-scale MD simulations combined with enhanced sampling methods. The *in silico* results were validated *via* statistical comparison of DFT-calculated chemical shifts with NMR experimental chemical shifts. The protocol was developed using AVP, publicly introduced in Paper 2 (Chap. 5) and subsequently applied to UII and URP in Paper 3 (Chap. 6) to predict their multiple-conformation equilibria in solution. The validation technique provides a method for the analysis of fast interconverting multi-conformational systems in general and may contribute methodologically to the research field of intrinsically disordered peptides. The determination of the conformational equilibria in solution defines the thermodynamic starting point for an allosteric signalling cascade during ligand/receptor interaction. It matters particularly if a minor populated “experimentally invisible” state initiates a signal transduction rather than an experimentally predominant conformation.

This research project was embedded in the European network project PeReNE (Peptide Research Network of Excellence)[337] as part of the Interreg IVA France (Channel) - England program 2007-2014 and the results of this work have been used *inter alia* by the groups of Prof. R. Bureau (University of Normandy, drug design), Prof. J. Essex (University of Southampton, development of unbiased enhanced sampling methods for intrinsically disordered peptides and proteins), and Prof. T. Clark (Friedrich-Alexander-Universität Erlangen-Nürnberg, metadynamics simulation of multi-allosteric

ligand-receptor reaction pathways). The multidisciplinary nature of this collaboration afforded a great opportunity for the resultant successful cross fertilization of ideas. Regular group meetings and arranged symposia progressed the project considerably.

In summary, there are a multitude of reasons that make the results reported in this thesis interesting for differing research areas (MD simulation and related analysis methods, NMR spectroscopy, pharmacological research) and the scientific contribution is further indicated by the successful publication of the results of this thesis as peer-reviewed papers.

METHODS

Chapter 3: Methodological Backgrounds

> "Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments."
>
> Press Release 09-Oct-2013 of the Royal Swedish Academy of Science, Nobel Prize Chemistry 2013 [134]

This chapter gives an overview of the principles and background of the methods used for this thesis. Reasons are given justifying the choice of methods. A research setup is outlined. Further on, more specific, methodological details are given in the subsequent chapters containing the publications.[1-3]

Research Setup

Acquisition of Structural Data. Methods to determine the structure of peptides and proteins have been addressed briefly in Chapter 2, primarily X-ray crystallography and NMR spectroscopy. However, X-ray crystallography is not suited to determining conformations in solution and NMR spectroscopy is limited if conformational interconversions are fast relative to the NMR timescale. MD simulations, however, enable the study of conformational interconversions on an atomistic level. Protein folding occurs on a timescale of microseconds[135] to seconds[136,137] or even longer,[138,139] and current computer power makes it possible to access simulation times of microseconds. Thus, long-scale simulations aim *inter alia* to reach realistic timescales. Here, the AMBER force field ff99SB[140] was used to study the conformational space of cyclic peptide hormones with explicit water solvation.

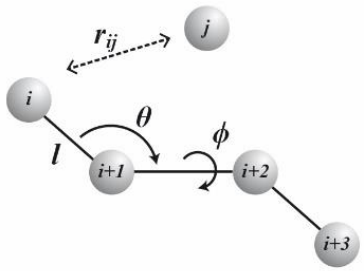
Analysis of Structural Data. The clustering of the conformational data was performed with DASH,[141] complemented with principal component analysis (PCA). For secondary-structure analyses, the AMBER tools *ptraj* and *cpptraj*[142,143] were used and the similarity of conformations was quantified by calculation of the circular similarity (see below).

Validation of Structural Data. NMR experiments of AVP, UII and URP in aqueous solution were performed to determine experimental chemical shifts; while chemical shifts for conformational

representatives of AVP, UII and URP were calculated using density functional theory (DFT) methods. The equilibrium populations for representative conformations were estimated *via* enhanced sampling with metadynamics simulations by Saleh and replica exchange simulations (REMD) by Essex and co-workers. The *in silico* results gathered from µs-scale MD simulations and enhanced sampling methods were compared statistically with the experimental results.

Molecular-Dynamics Simulations

What is MD simulation? Physicochemical Aspects. The potential energy of a molecule is a function of its conformation and the conformational space can be represented as a potential energy surface. This multi-dimensional energy surface can be described by a force field with contributions of bond lengths, bond angles, torsions, non-bonding forces and electrostatic forces. The energy contributions are parameterised and the potential energy is approximated by summation of harmonic potentials for bonds and angles, Fourier expansions for torsions (torsion potential), Lennard-Jones potentials for non-bonding forces (dispersion, van der Waals) and the Coulomb law for electrostatic contributions. This approach is called molecular mechanics (MM) and the resulting potential energy is the MM energy[i]. Equation (3.1) is a typical MM energy function as used by the AMBER force field ff99SB.[140]



$$E_{MM} = E_{bonded} + E_{non-bond} \tag{3.1}$$

$$= \sum_{bond} \frac{k_l}{2}(l - l_0)^2 + \sum_{angle} \frac{k_\Theta}{2}(\theta - \theta_0)^2 + \sum_{torsion} V_n\left(1 + cos(n\phi - \gamma_n)\right)$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N}\left\{4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \frac{1}{4\pi\varepsilon_0}\frac{q_i\, q_j}{r^2}\right\}$$

with E_{MM} molecular mechanics energy; $E_{bond/non\text{-}bond}$ energy terms for *bonded* and *non-bond* forces; *i,j* number of atom; *k* force constant; *l* bond length; suffix *0* equilibrium/minimum; *Θ* bond angle; V_n torsion force constant (amplitude); *Φ* torsion; γ_n phase (position of 1st maximum); *n* number of maxima; *ε* maximum attractive energy; r_{ij} distance; *σ* distance of no interaction; q_i, q_j partial charges at atom *i* and *j*, ε_0 electric constant

[i] in contrast to the evaluation of the potential energy by quantum mechanics

The *force field* is the dataset of parameters (*e.g.* k_l, l_0, Θ_0, ε in Eq. (3.1)) together with the energy functions used to calculate the potential energy. The parameterisation considers atom types (including hybridisation and atom charges), all possible atom-type combinations (*e.g.* bond lengths, bond angles, proper and improper torsions, non-bonded interactions, electrostatic interactions) and, if necessary, rules to estimate missing parameters. For textbooks and reviews on MD simulations, see *e.g.* references 144-147; additional physicochemical background knowledge is given in Appendix A6.

Molecular dynamics describes the possible motion of atoms within a molecule. This includes interconversions between different conformations. In classical mechanics, the atomic motion is determined by solving Newton´s laws of motion[148] (Eq. (3.2) and Appendix A6). This approach calculates the future position of an atom from its current and previous positions. It determines how changes in the potential energy ($\frac{\partial E_{MM}}{\partial r_i}$) are related to changes in position as a function of time (Eq. (3.3)).

$$\vec{F}_i = m_i \vec{a}_i \quad \text{and} \quad \vec{a}_i = \frac{\partial^2 \vec{r}_i}{\partial t^2} \tag{3.2}$$

$$-\frac{\partial E_{MM}}{\partial r_i} = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} \tag{3.3}$$

with i atom number, F_i force, m_i mass, a_i velocity, r_i position, t time

MD simulation, thus, explores the conformational space autonomously.[145] The results of MD simulations are time-trajectories of conformations and time-averaged populations of conformations. In theory, if the system were allowed to evolve indefinitely, all possible conformations would be sampled. Experimental methods, in contrast, result in ensemble averages. However, in the case of convergence, the time-average of populations ("MD equilibrium") should equal the experimentally observable ensemble average ("experimental equilibrium"). This axiom is called the ergodic hypothesis (*cf.* Appendix A6) and it is the fundamental reason why MD simulations should run for as long as possible.

Equilibrium populations and free energy are related *via* Eq. (3.4):

$$\Delta G = -RTlnK_{eq} = -RTln\frac{[P2]}{[P1]} \tag{3.4}$$

with ΔG difference of Gibbs free energy; R ideal gas constant (); T temperature; K_{eq} equilibrium constant; $P1$, $P2$ concentrations (populations)

However, if the sampling is insufficient, the energy cannot be deduced from the MD populations. Even with µs-scale simulation lengths, convergence cannot be taken for granted[149] and in this case, enhanced sampling methods are recommended, *e.g.* umbrella sampling,[150] metadynamics,[151]

replica-exchange molecular-dynamics simulations,[152] or solute tempering.[153] Enhanced sampling was performed for AVP, UII and URP[i] to supplement the long-scale MD simulations. Methodological details are given in Chapter 5 (Paper 2) and Chapter 6 (Paper 3). For additional information, reference is made to the literature.[150-152,154-161]

At this point, it must be mentioned that free energies and equilibrium populations cannot be deduced from potential energies under standard conditions (T≠0) because the entropy term is unknown (Eq. (3.5)):

$$G = H + TS = (U - pV) + TS \tag{3.5}$$

with *G* Gibbs free energy, *H* enthalpy, *T* absolute temperature, *S* entropy, *U* internal energy (here: E_{MM} potential energy), *p* pressure, *V* volume

Consequently, the global minimum of the potential energy surface need not be the highest populated (most stable) conformation.[145] Nevertheless, it is assumed that the low-energy regions of the hypersurface are the most populated.

Historical Aspects. The development of force fields and MD simulations started about 40 years ago with potential energy calculations by *e.g.* Scheraga[95,162] and Allinger *et al.*[163] The *consistent force field* (CFF) by Lifson and Warshel[164,165] is often regarded as the "*foundation of modern molecular modelling*".[166] One of the first protein MD simulations was published by McCammon, Karplus *et al.* in 1977.[167] Nowadays, several well recognised force fields are available and commonly used, *e.g.* AMBER,[168-170] CHARMM,[171,172] OPLS,[173,174] and GROMOS.[175] AMBER, CHARMM, and OPLS offer *all-atom* force fields. AMBER and CHARMM focus primarily on protein simulation, OPLS on the simulation of liquids. GROMOS was initially optimised for alkanes and it still uses *united-atom* force fields. Here, only the development of the AMBER force field, which was used in this work, will be described in detail. AMBER was introduced in 1981 as "a *general program for modelling molecules and their interaction*".[170] The first widely used AMBER force field, released in 1984,[169] considered only polarisable hydrogens explicitly and nonpolar hydrogens were parameterised as a unit with their bonding partners (*united-atom* force field). The architecture of force fields and programs was and is closely linked to the computer power available and in 1986, the first AMBER *all-atom* force field became available.[176] Further developments, including improved algorithms and protocols to extend the parameter set, led to the *ff94* or *Cornell* force field[177] in 1995. A weakness of ff94 (and ff99) was its bias in protein simulations towards the helical conformation.[178] This problem was addressed with the AMBER force field version ff99SB, released in 2006.[140] At the beginning of this project in 2011, ff99SB was commonly used as a standard force field that had proven reliable and

[i] by Essex *et al.* and Clark *et al.*

predictive to study the dynamics of proteins.[140,158,179] For this project, ff99SB was also chosen to ensure compatibility with the study of AVP-receptor interaction of Saleh and Clark.[28]

Strengths, Application and Limits of MD simulations. The strength of classical MD simulations based on force fields is the computational feasibility of calculating large systems, even in membrane environment with explicit solvation (*e.g.*[180-182]). The additive character of the potential functions enables extensive computational parallelisation and is the reason for the possibility of high-speed performance. MD simulation is used to study conformational changes (*e.g.* protein folding[182]), molecular recognition (*e.g.* ligand-receptor interaction,[28] DNA-protein interactions[183]), ion transport processes[184,185] and many other questions. Results are *inter alia* used for drug design.[186,187] Short restrained MD simulations are used routinely to refine conformations deduced from experimental methods (*e.g.* NMR spectroscopy).

The quality of the results correspond closely to the quality of the parameterisation.[149] Force field parameters are fitted either to experimental values (bond lengths, rotation barriers *etc*.) or to *ab initio* calculated values (*e.g.* partial charges). Test sets are chosen with respect to the system to be simulated. The AMBER force field, for example, is optimised for describing the secondary structure folding of proteins. Force fields tend to be biased in favour of potential energy minima, this aggravates the previously mentioned sampling problem.[160] A general weakness of the classical mechanics approach is that it does not allow the calculation of electronic processes (*e.g.* electron transfer, bond dissociation). Electron motion is significantly faster than atom motion so that atoms and electrons are assumed to move independently (Born-Oppenheimer approximation[188]). The charge distribution in standard force fields is defined by atom-fixed point charges and intramolecular electrostatic interactions between neighboured atoms (< 1-3 bonds) are omitted or scaled. The most accurate way to calculate electronic processes would be an all-atom *ab initio* quantum mechanical approach, but this is still not possible for large systems but mixed approaches do already exist. These methods combine the accuracy of quantum mechanics (QM) with the high-speed performance of molecular mechanics and the approach (QM/MM) was rewarded with the Nobel prize[i] for Chemistry in 2013.[134,166] Another approach uses polarisable force fields[189-191] but their use is not yet routine. Fixed-charge force fields at least should include explicit solvation for an optimum simulation of the electrostatic solute-solvent interactions. For a review, see *e.g.* Ponder *et al.*[192]

[i] Nobel prize Chemistry 2013: A. Warshel, J. Levitt, M. Karplus

Boundary Conditions for MD Simulations

Initial Conformation and Minimisation. A good starting point for an initial conformation of an MD simulation is an experimental template (*e.g.* X-ray or NMR structure), ideally, with complete coordinates or structural data (*e.g.* backbone torsions) to model a starting conformation. Other starting points may be created with short high-temperature MD runs or modelled from related analogues for which conformations are known. In any case, the initial conformation needs to be minimised before a simulation can be started. The aim of the minimisation is to optimise the initial conformation by releasing strains that would result in unacceptable energy gradients. For example, for AVP, the X-ray structure of the trypsin complex (PDB ID: 1YF4) was chosen, whereas for UII, a conformation was modelled, guided initially by published torsion angles deduced from NMR. Further conformations of UII were generated using high-temperature short-scale MD. A resulting minimised structure usually differs only slightly from the initial conformation. As an example, a superposition of an initial and minimised structure of UII is given in Figure 3.1.

Several minimisation methods are available (*e.g.* simplex, steepest descent, conjugate gradient, Hessian matrix) and the mathematical algorithms use energy differences, gradients or second order derivatives of the potential energy to find the closest energy minimum to the initial conformation. The minimisation methods differ in speed and accuracy.[158] A standard approach (used in this work) is to start the minimisation with steepest descent to find the right direction to the minimum quickly, followed by the slower but more accurate conjugate gradient method. An example for a typical minimisation is given in Appendix A7.

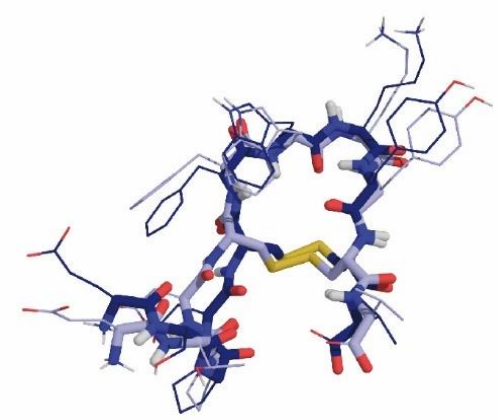


Figure 3.1 Superposition of an initial and minimised structure of UII (modelled from NMR data of URP,[125] $RMSD_{CA\text{-}backbone}$ = 0.963 Å)

Solvation. Explicit solvation requires accurate solvent models. One of the first computational models for liquid water was introduced as early as 1933 by Bernal and Fowler.[193] For force fields, the ST2 model by Stillinger *et al.*[194] was one of the first standard models for explicit solvation. It was

followed by Berendsen´s SPC[195] and Jorgensen´s TIP3P[196] water models as standard. The two models are quite similar and follow the concept of a rigid 3-site architecture (3 atoms with 3 non-polarisable point-charges). They are optimised for a good description of the bulk phase structure of water and a correct reproduction of thermodynamic properties (*e.g.* density and heat of vaporisation). Their interaction with the solute is mainly electrostatic.[192] TIP3P is still established as the standard model (default in AMBER 14) for explicit water solvation, although the 4-site model TIP4P, which uses an additional charge centre (pseudo-atom) is thought to be significantly better in simulating density and long-range electrostatics.[197] In the framework of this project, the TIP4P-Ew model[197,198] was employed, a re-parameterisation of the TIP4P model, optimised for the combination with Ewald methods (see below). There are numerous solvent models[199-202] but it is important to note that each force field needs an adaptation to the water model used. For ff99SB, this is the modified parameter set frcmod.tip4pew (used in this work) implemented in AMBER.

An ideal solvent should stretch indefinitely in all directions to enable free dynamics of the solute while ensuring a homogenous environment. For optimum computational performance, however, the number of solvent molecules should be as small as possible. How is this conflict solved? The solute is placed in the centre of a box, which is filled homogenously with solvent molecules and potential counterions. The ideal geometry of this box would be an asymmetric enlarged shell proportional to the surface of the solute. Practicable geometries are usually a cube or a truncated octahedron (Figure 3.2 shows an example of truncated octahedral solvation.).

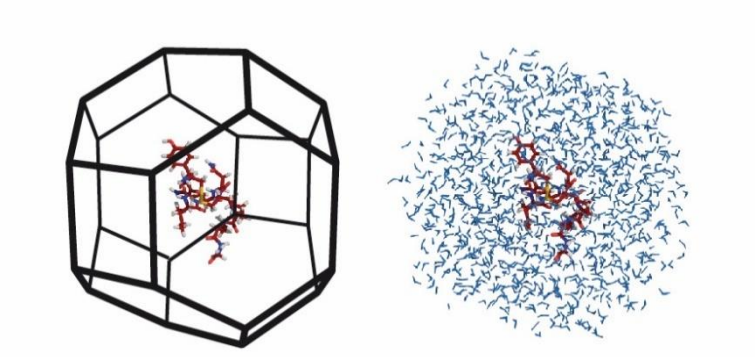


Figure 3.2 OT molecule in a truncated octahedral water box. Left side: schematic view. Right side: particle view.

Periodic boundary conditions imitate the “infinite” expansion of the solvent *via* imaging the solute-solvent box in all Cartesian dimensions. If an atom leaves the centre box during a simulation step, it is mirrored back into the centre box on the opposite side. This is a mathematical trick, resulting in no error in potential energy as long as the half box dimension is larger than the cut-off for non-

bonded molecular interactions. This cut-off is usually set to 8-10 Å.[i] The particle-mesh Ewald (PME) method[203,204] is a modified form of the Ewald summation algorithm[205] to evaluate the quantities of large periodic systems, here the potential energy in periodic boundary environment. PMEMD and PMEMD.CUDA[179,206] are implementations of PME in AMBER optimised for high-speed parallel performance on CPUs and GPU[ii] (see Appendix A7). For the peptides studied here, the truncated octahedral water box was chosen, which leads to an average of 97 % solvent atoms in each simulation.

Besides explicit solvation, several methods have been developed for implicit solvation, primarily the generalised Born[207,208] and the Poisson-Boltzmann models[209-211] for protein force fields. The solvent is represented by a dielectric continuum instead of individual water molecules, which makes simulations faster (less atoms). Results are reasonable for macroscopic values (*e.g.* solvation energies, pK_s estimation, redox potentials), but tend to overemphasise salt-bridges[212] and cannot simulate atomistic solvent-solute interactions (*e.g.* water bridges). The reaction-field (RF) approach[213,214] is an alternative to the PME method. It uses a combination of explicit solvation and a dielectric continuum after a certain cut-off distance to simulate solvation, which makes the technique fast. PME, however, is more widely used and RF is not the standard method in AMBER.

Temperature, Pressure, Density. The MD simulations of the peptides here were performed under standard conditions of 300 K[iii] and periodic boundary pressure conditions with adjustable volume to ensure the appropriate solvent density (target $\backsim$1 g cm^{-3} for water). The average fluctuation of the solvent density and temperature should be as small as possible. Constant temperature (which corresponds to a constant kinetic energy) and pressure are ensured *via* a Berendsen coupling algorithm (weak coupling to an external bath).[215]

Simulation Lengths. As already noted, one of the aims of MD simulation is to study protein dynamics on a realistic timescale and to converge ideally to a thermodynamic equilibrium.[216] The limits for this lie in the hardware and software.[217] The two develop mutually, and simulations are currently expanding to the µs-scale.[27,181,182,218] Of course, the term *long-scale simulations* in the title of this thesis for µs-scale MD simulations is relatively and perhaps in the near future, µs-*scale* simulations will be *short-scale*. Nevertheless, during the time of this research project, the computational requirement for µs-scale MD simulations in terms of necessary CPU time was still high. The real runtime of MD simulations for small systems[iv] need not be faster than for large

[i] Default cutoff in AMBER 10 and AMBER 14 is 8 Å
[ii] CPU Central Processing Unit; GPU Graphics Processing Unit
[iii] default in AMBER (ff99SB parameters are optimised for 300 K)
[iv] *e.g.* the peptides here

systems[i] surprisingly. Small systems cannot be parallelised as effective as large ones, which profit more from high-performance multicore supercomputers. Most MD simulations in this work were performed with AMBER 10[168,219,220] on an 8-node cluster of Intel CPUs (Xeon E5462) with an average production time of 41 days to simulate 1 µs of peptide. In 2015, several jobs were performed on NVIDIA GPUs (Tesla K20c and C2075) with the AMBER 14[206] GPU (CUDA) version[221-223] with a 4-fold better performance (9 d/µs). A summary of the performances of MD simulations in this work is given (Appendix A7).

Beside the hardware, the length of the simulation time-step is an intrinsic time-restricting factor. Time-step lengths cannot be chosen arbitrarily for a reliable simulation. The maximum time step should be 10 times shorter than the fastest motion to be studied.[145] In an all-atom MM force field, this would be the vibrations of X-H bonds (10 fs to 10 ps), requiring a simulation time-step of < 1 fs. The SHAKE algorithm[224] allows a time step of 2 fs to be used by constraining the C-H bonds during a simulation step followed by a short relaxation of C-H bonds, which reduces the total runtime. The SHAKE algorithm was used for all MD simulations in this work.

Simulation Tools

The AMBER software package includes special simulation tools (LEaP, SANDER) to prepare and run an MD simulation.[168] Examples of how to set up the initial parameters and run a minimisation and simulation are given in Appendix A7.

Structure Analysis

The result of an MD simulation is the time evolution (trajectory) of atom coordinates. The motion of the simulated peptides may be visualised as a video clip (an example is given online as Supporting Information of Paper 1). A video observation is similar to an experiment and provides a qualitative description of the peptide dynamics: for example, different fluctuations of tail and ring become clear, main interconversions of the backbone can be observed and the dynamics of intramolecular hydrogen bonds can be visualised.

[i] *e.g.* protein including membrane and explicit solvation

The methods and tools used in this thesis for quantitative analysis and characterisation of dynamics and conformations are described below.

Clustering. Before it is possible to describe the characteristics of conformations produced by the MD simulation, it is necessary to cluster them, grouping similar conformations together. Classical clustering methods use a pairwise metric to compare Cartesian coordinates. The algorithms are numerous[225] but their performance depends on the square of the number of data points, causing them to slow down exponentially the data performance with increasing data volume. DASH,[141] in contrast, clusters torsion ensembles (*e.g.* the $\Phi\Psi$ backbone torsions) using a sequential algorithm following the time-evolved trajectory of conformations. This enables the program to process large data volumes produced by long-scale MD simulations without an exponential decrease in performance. At the beginning of this project, the performance and consistency of DASH was tested against *average linkage* and *means*, two standard classical-clustering methods implemented in AMBER *ptraj* (for details, see Appendix A7). All these cluster methods are able to determine the main conformational types of the peptides studied but DASH proved to have a significantly better performance (see Appendix A7). DASH results in a state trajectory of accurate concordance with the time-evolved root mean square deviation (RMSD) coordinates of a MD simulation without requesting a predefined number of clusters. Classical clustering methods cannot easily distinguish between frequent intermediate conformations and long-lived main conformations. In DASH, however, a conformation needs to persist for a minimum lifetime to be considered as a relevant conformation (cluster centre). These advantages made DASH the optimum cluster method for the long-scale MD simulations in this thesis. During this project, the workflow of the DASH program applied to AMBER trajectories was automated as a Perl script called *amberDASH*. This requires only a few inputs to automatically extract torsion angles from the AMBER coordinate trajectory, to run the DASH analysis, and to produce coordinate files of representative states (cluster centres) in PDB format. Details are given in Appendix A7. DASH versions 2.10b1 to 2.11b2[226] have been used during this project.

Ptraj, Cpptraj. After identifying the main conformations their characteristics need to be determined. In principle, any structural data that change significantly between representative conformations can be taken as characteristics for a state. This ranges from basic geometric data (*e.g.* interatomic distances, angles, or torsions) to coarser scale properties (*e.g.* RMSD, radii of gyration, secondary structure propensities, or hydrogen-bond populations). For the characterisation of cyclic peptides, the determination of secondary structure propensities and hydrogen-bond populations of distinct conformational types identified with DASH proved very

useful. RMSD trajectories were used to generate 2D visualisations of the MD simulations, whereas radius of gyration, distance and torsion trajectories were used for supplementary monitoring of the dynamics of motions or for the identification of key parameters, *e.g.* key torsions for interconversions.

The programs used here to extract the above properties and to perform the necessary analyses were *ptraj* and *cpptraj* (extended C++ version of *ptraj*) included in AmberTools.[142,143] Secondary structure and hydrogen bond analyses will be explained in more detail; the standard analysis routines are described in the AmberTools user manuals.

The secondary structure analysis (*secstruct*) uses the DSSP[i] method of Kabsch and Sander[227] to identify secondary structure motifs such as turns, sheets or helices and to calculate their populations. DSSP defined β-turns by the Lewis distance criterion[228] for $C\alpha_i$ and $C\alpha_{i+3}$ (r < 7Å) rather than ideal torsion angles and a high populated hydrogen bond.[229]

The analysis tool *hbond* in *ptraj* tracks distances and angles of atom triplets. To analyse the intramolecular hydrogen bonds of the cyclic peptides, the triplets were defined by carbonyl oxygens as acceptor atom (O), hydrogen atom (H) and amide nitrogens as hydrogen-donor atom (N) with a distance cutoff of $r_{O\text{-}N}$ = 3.5 Å and an angle cutoff of 120° (deviation from a linear O-H-N configuration).

Principal Component Analysis (PCA). PCA is a data reduction method that projects high-dimensional datasets with many variables onto a small number of new variables that describe most of the variability in the data. It calculates the eigenvectors (principal components, PCs) of the covariance matrix. PCs with an eigenvalue > 1 are usually assumed to contain a significant amount of the variance in the system.[230] The first PC points in the direction of maximum variability (highest variance); the second and following PCs give orthogonal directions of decreasing variance. Here, the data to be analysed were the conformations with their $\Phi\Psi$ torsions as variables. 2D and 3D plots of the conformations in relation to the significant PCs allow groups with common properties (like clusters) to be visualised. For further reading, reference is made to *e.g.* [231-233]

PCA was used in two ways:

(i) The first objective was to investigate whether the *overall* conformations of the cyclic peptides could be characterised solely by their *ring* conformations (ring-state types). For this, principal components of the *overall* torsion space were calculated and a 3D plot of the first three PCs was drawn with each conformation colour coded according to its ring-state type. If each visible cluster in the PCA plot were assigned a unique colour, this would indicate that the ring-state types do

[i] Define Secondary Structure of Proteins

characterise the overall conformation. In addition, this would show that clustering by independent methods (PCA and DASH) provides equivalent results.

(ii) The second objective was to analyse the correlations between torsions in the ring and the tail. The weights (squared PC coefficients) of the descriptive variables (torsions) measure their contribution to each PC. If significant PCs are loaded equally strongly with ring and tail torsions, then correlation of ring and tail conformations can be assumed and independent motion is unlikely. For further details and examples, see Chapters 4 (Paper 1) and 6 (Papers 3).

In the early stages of the project, PCA was performed with the online application SARcaddle,[234] later a dedicated PCA routine was implemented in DASH[226] (an example output is shown in Appendix A7).

Circular Similarity. The consistency of assignments to conformational types was confirmed by calculation of torsion similarities using the program *dashsim*. The algorithm is explained in the Supporting Information of Paper 3 (Appendix A3, p S10) and a brief description of the functionality of *dashsim* is given in Appendix A7.

NMR Spectroscopy

Dynamic molecular processes cover a wide range of timescales ranging from bond vibrations of femto- or nanoseconds to conformational interconversion processes like protein folding lasting up to several seconds or even minutes. NMR is classically the method of choice to study molecular dynamics experimentally.[136] However, there is a "blind spot"[i] between fast timescale dynamics < 10 µs and slow timescale dynamics > 10 ns, which is difficult or not accessible to common NMR techniques.[136] Fast conformational interconversions that fall into this gap are only observable as averaged ensemble with a single set of signals under standard conditions.[133] In the literature (*cf.* Tables 2.3 to 2.6), the structure of the peptides of this thesis is characterised ambiguously both as single-conformation and unstructured in aqueous solution suggesting fast conformational equilibria within the timescale of this gap.

In the framework of this thesis, the experimental chemical shifts of different nuclei (^{1}H, ^{13}C, ^{15}N) were determined for AVP, UII, and URP with standard 1D- and 2D-NMR techniques. The NMR experimental data served to validate the *in silico* results of the MD simulations. The validation technique is explained in Chapter 5 (Paper 2). Experimental details are given in Chapters 5 and 6 (Papers 2 and 3 and the corresponding Appendices A2 and A3).

[i] An illustration is *e.g.* given by Palmer *et al.*[129] (Fig. 1a)

DFT Calculations of NMR Chemical Shifts

Being able to calculate NMR observables is of interest for many reasons. For example, accurate chemical-shift predictions can facilitate NMR assignments and re-assignments, allow diastereomers to be distinguished, confirm suggested structures and enable the study of conformational processes.[235] In this work, NMR chemical shifts were used to evaluate the *in silico* determined conformational equilibria of the cyclic peptides.

DFT is a widely used theoretical approach to calculate atomic and molecular properties, including magnetic properties such as NMR chemical shifts.[236] Simplified, it uses the electron density as basic function for quantum mechanical calculations instead of a complicated all-electron wavefunction. This makes the approach applicable for larger systems and has the additional advantage of the electron density being an experimental observable. Nevertheless, the cyclic peptides in this work still represent a large system for DFT. The level of theory for the DFT calculations was B3LYP/6-31G(d). B3LYP[237,238] is a popular hybrid functional for exchange and correlation energy expressions that has proven successful for many applications and includes a contribution from Hartree-Fock exchange.[239] 6-31G(d)[240] is a split-valence-plus-polarisation basis set to calculate the electronic wave function. It is relatively small, yet still appropriate to give accurate results while being computationally feasible for the cyclic peptides here.

Nuclear magnetic resonances arise due to the interaction of an external magnetic field with the magnetic moment of nuclei with unpaired spin (*e.g.* ^{1}H, ^{15}N, ^{13}C). The electrons close to the nuclei affect the external magnetic field and the effective local magnetic field varies depending on this *shielding*[241,242] (Appendix A6). Thus, chemical shifts of the NM resonances are caused by varying electron distribution related to the local conformation around the nuclei. An increase of the local magnetic field (shielding) effects an up-field shift and a decrease (deshielding) is followed by a downfield shift. Two well-established techniques to calculate nuclear magnetic shieldings within DFT are IGLO[243] (*Individual Gauges for Localised Orbitals*) and GIAO[244,245] (*Gauge-Invariant Atomic Orbital*). GIAOs are known to give more accurate results with small basis sets than IGLO[244] and show a fast convergence of calculated chemical shieldings.[246] Here, the standard implementation of GIAO in Gaussian09[247] was used at the level of theory mentioned above, representing approximately the minimum DFT level for reliable NMR observables.[248,249] For the quantum-mechanical evaluation of the relationship of structure and nuclear magnetic shielding, the fundamental theory of magnetic properties and an in depth discussion of density-functional theory, the reader is referred to the specialised literature.[235,239,250-253]

Solvent effects were simulated with the common polarizable continuum model (PCM)[254] for water representing an implicit solvation. The calculations of the magnetic shielding tensors for the peptide nuclei were preceded by DFT geometry optimisation (consistent DFT level). Linear regression parameters to convert the absolute isotropic nuclear magnetic shielding (σ, dimensionless) into chemical shifts (δ, ppm) were obtained by correlation of well-established chemical shifts of small organic molecules with DFT calculated magnetic shieldings at the same level of theory as used for the calculation of the peptides. Linear regression against several reference compounds provides a better error cancellation than simple referencing to only one NMR standard (*e.g.* DSS, TMS[i]). In this way, NMR chemical shifts (^{1}H, ^{13}C, ^{15}N) have been calculated for AVP, UII and URP and methodological details are given in Chapters 5 and 6 and the Supporting Information (Appendices A2 and A3).

Statistical Evaluation

The objective of the computational simulations in this thesis was to generate a realistic description of the structure and dynamics of the cyclic peptides. Single conformations and equilibrium mixtures are models for the "real" conformation and the hypothesis is that an equilibrium mixture of relevant conformations will yield a better description than any single conformation.
To test this hypothesis, experimental observables, *e.g.* NMR chemical shifts, were compared with the corresponding values calculated from the models. The model that fits best is assumed to describe the "real situation" most accurately. For this, different error metrics were used and the fundamentals of these statistic methods are explained.

Linear regression.[255-257] Before analysing error metrics, a scatter plot of experimental data against calculated data was drawn to visualise their correspondence. The relation between the two sets of data can be expressed mathematically with a simple linear regression (Eq. (3.6) and (3.7)):

$$y = mx + a \qquad (3.6)$$

m slope; a intersection

$$y' = m'x \qquad (3.7)$$

m' slope; intersection = 0

The regression straight line is the line to which all points are positioned as closely as possible. If the intersection of this straight line is set to the origin (Eq. (3.7)), different models can be compared directly. A measure of the agreement between the model and the experimental values is the

[i] DSS= 4,4-dimethyl-4-silapentane-1-sulfonic acid, $(CH_3)_3Si\text{-}(CH_2)_3\text{-}SO_3H$; TMS= tetramethylsilane, $Si(CH_3)_4$

coefficient of determination (R^2). Its ideal value is 1 and the worst case 0. In this work, diagrams were plotted within Microsoft® Excel® 2013 using the standard least squares method for simple linear regression to determine R^2.

Error metrics. Metrics are a measure of the differences between pairs of values. In the case of *model vs. experiment*, ideally, the pairs of values (*e.g.* the NMR chemical shifts for particular atoms) should be identical. The smaller the error metrics, the higher the accuracy of the model. Standard error metrics used in this project were the mean signed error (MSE), mean unsigned error (MUE), and root mean square error (RMSE). In addition, two new metrics were defined, the *weighted* RMSE (WRMSE) and the *coefficient of distinctiveness* Δ_σ.

The MSE[i], Eq. (3.8), is the mean of all individual pair differences and indicates mean systematic deviations for the entire dataset.

$$MSE = \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)}{N} \tag{3.8}$$

$\hat{y}_i$ calculated value (*e.g.* chemical shift); y_i experimental observable; i atom; N total number of atoms

The MUE, Eq. (3.9), is more significant than the MSE because it is based on absolute pair differences.

$$MUE = \frac{\sum_{i=1}^{N}|\hat{y}_i - y_i|}{N} \tag{3.9}$$

$\hat{y}_i$ calculated value (*e.g.* chemical shift); y_i experimental observable; i atom i; N total number of atoms

The RMSE, Eq. (3.10), also known as RMS deviation, gives the root of the mean of all squared differences. In contrast to the MUE, it weights large deviations more than small ones.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}} \tag{3.10}$$

$\hat{y}_i$ calculated value (*e.g.* chemical shift); y_i experimental observable; i atom; N total number of atoms

To enhance the significance of the error metrics further, the *weighted* RMSE (WRMSE) was introduced. It not only punishes large individual errors but also weights the dependence on conformation by introducing the standard deviation of the models (σ_i). Errors of particular values with large standard deviation are weighted more strongly than those that depend less strongly on conformation. σ_i is large for values that show strong deviations between different conformations. The WRMSE is given in Eq. (3.11).

$$WRMSE = RMSE \cdot \sqrt{\frac{\sigma_i}{\bar{\sigma}}} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \sigma_i}{\sum_{i=1}^{N}\sigma_i}}, \sigma_i = \sqrt{\frac{\sum_{j=1}^{M}(\hat{y}_{ij} - \bar{\hat{y}}_{ij})^2}{M}} \tag{3.11}$$

σ_i standard deviation of calculated values i of all models; $\bar{\sigma}$ average or arithmetic mean of σ_i; $\hat{y}_{ij}$ calculated value for atom i and model j; $\bar{\hat{y}}_{ij}$ average or arithmetic mean of calculated value for atom i of all models; i atom; j model; M total number of models

[i] Note: MSE is also used as acronym for the mean squared error

The second new metric, the *coefficient of distinctiveness* (Δ_σ), was designed to estimate the significance of MUEs. As has been explained for the WRMSE, the standard deviation of model values (σ_i) is proportional to the diversity of conformations. Thus, σ_i is used to estimate the distinctiveness of the overall error metrics by weighting particular MUEs of atoms with high conformational diversity between models less than others (*cf.* Eq. (3.12)). Ideally, models should differ enough (large σ_i) to make the decision for "the best" model significant. The limiting error value of $\Delta_\sigma \leq 1$ was introduced to characterise a model that is able to discriminate between different conformations. A detailed discussion is given in Chapter 5 (Paper 2).

$$\Delta_\sigma = \frac{\sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{\sigma_i}}{N} \tag{3.12}$$

Δ_σ Coefficient of distinctiveness; $|\hat{y}_i - y_i|$ particular MUE for values of atom i; σ_i standard deviation of calculated values i of all models; $\hat{y}_i$ calculated value for atom i; i atom; N total number of atoms

Methodological details of how to model the conformational equilibria and an in-depth discussion of the evaluation technique are given with the results for AVP in Chapter 5 (Paper 2).

Fitting methods. In theory, statistical methods could also be used to estimate the equilibrium conformations *via* statistical fitting of the calculated variables (NMR chemical shifts) to experimental values. A statistical approach was tested for AVP (Chap. 5, Paper 2). The methods applied were partial least squares regression (PLS) and bagged multiple linear regression (MLR). The first is related to PCA and uses variances in its regression approach while the latter uses linear combinations of descriptors (here chemical shifts of different conformations) to define the best model. However, the results have little predictive power if the majority of calculated variables are highly correlated, as they are in the case of the calculated NMR spectra. Thus, in this research, the relative populations of relevant conformations were determined *via* enhanced sampling methods.

DP4 probability.[258] Goodman and co-workers offer an easy to use Java-applet to test NMR chemical shift assignments. The required input is a set of experimental chemical shifts (^{1}H, ^{13}C) and the corresponding calculated chemical shifts of diastereomers. Their comparison of calculated and observable shifts results in a probability of correct assignment, called DP4 probability. In this work, the input was modified by using different conformations rather than diastereomers. The application was used to test whether the *in silico* predicted chemical shifts of the conformational equilibria were assigned the highest DP4 probability. For details, see Chapters 5 and 6 and the corresponding Supporting Information.

Summary

In summary, most methods used in this thesis were standard applications to guarantee a maximum of compatibility for cooperative research projects (*e.g.* standard MD force field, standard DFT parameters). Within the standard methods, the state-of-the-art parameters were chosen where possible to achieve maximum accuracy (*e.g.* TIP4P-Ew water instead of TIP3P, µs-scale run-times instead of ns-scale, *ab initio* DFT optimised conformations instead of MD minimised). For analyses beyond established methods, new methods and metrics were tested, applied and developed (*e.g.* DASH, Δ_σ).

RESULTS AND DISCUSSION

Chapter 4: Conformation and Dynamics of Arg^8-Vasopressin in Solution (Paper 1)

The results in this section have been published in:

Haensele E, Banting L, Whitley DC, Clark T. Conformation and dynamics of 8-Arg-vasopressin in solution. J Mol Model. 2014;20(11):2485(17).[1]

The paper is given as postprint.

Foreword

The application of new protocols and methods to known systems is a classical scientific approach to establish their reliability. Long-scale MD simulations[i] have not been reported for cyclic peptide hormones to date and DASH as a high-performance clustering methods for long trajectories has been tested here. AVP is a prime example for structural research in general and for bioactive flexible peptides in particular. As it is one of the first synthesised peptides,[259] a multitude of structural data can be found in the literature, as has been outlined in the Introduction (Chap. 2). This is a comfortable situation to test the reliability of the extended timescale MD simulations and to test the performance and accuracy of the analysis method DASH (*cf.* Appendix A7).

Paper 1 reports the results of 11 µs unrestrained MD simulation of AVP with the AMBER force field ff99SB and explicit water solvation. Three of the main conformations identified (*saddle, twisted saddle* and *open*) resemble known data (*cf.* Table 2.4), which shows the method is able to reproduce known conformations. A further, previously unknown, main conformation was also found and characterised (the *clinched open* conformation, shown in Fig. 4.1). Based on the data of the 11 µs simulate, it was assumed to be a minority population. However, the extended MD simulation to 23 µs (*cf.* Paper 2) showed an increased population of this conformation and the enhanced sampling studies (by Dr. Saleh) identified the *clinched open* conformation as the second most frequent main conformation of AVP.

[i] which is currently µs-scale

The clustering method DASH demonstrated excellent performance on long trajectories. Conformations were not only clustered for the complete peptide sequence (*overall* conformations) but also separately for main motifs (ring and tail) which led to a classification of main conformations of AVP based on conformational ring types. The paper explains in detail the conformational clustering of AVP and presents an optimised protocol for the analysis and clustering of long-scale MD simulations of flexible peptides.

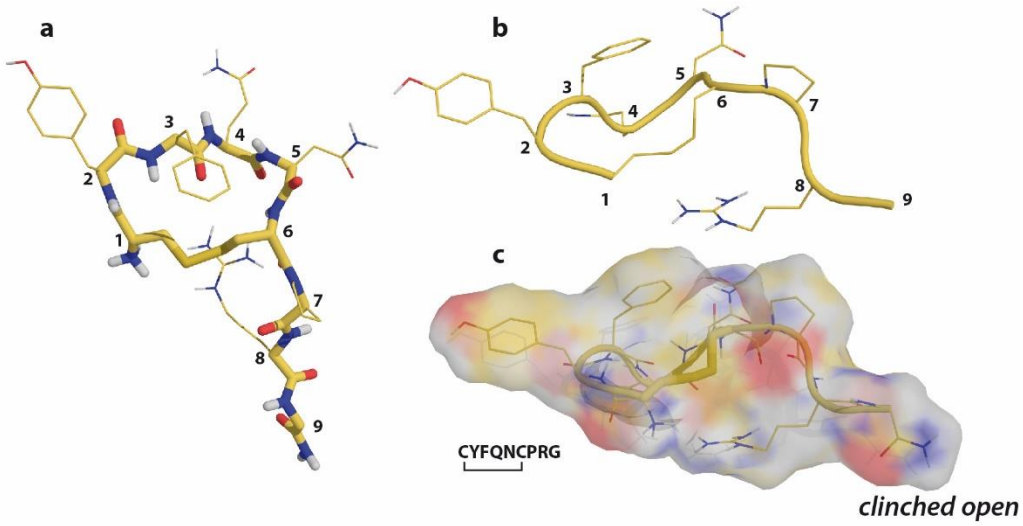


Figure 4.1a-c AVP representative for the ring-state type *clinched open.* Transannular hydrogen bonds are not significantly populated. Turns are centred at residues 4,5 (open β-turn type VIII/I). The ring-state type *clinched open* of AVP corresponds to the ring-state type *omega* of UII and URP. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, residues labelled; (**b**) backbone (cartoon), sidechains (lines), nonpolar hydrogens hidden, residues labelled; (**c**) surface.

Contribution of Authors

The results are the product of a joint research project between the University of Portsmouth (UK) and the FAU Erlangen-Nürnberg (D) within the framework of the European "Peptide Research Network of Excellence" (PeReNE).

MD simulations, data analyses and protocol optimisations were performed by Haensele.

Principal component analyses were performed by Prof. Clark.

Dr. Whitley extended the routines of DASH and improved the usability of the application. The source code of *amberDASH* was written (DW) based on an idea of Haensele, which facilitates the DASH-clustering of AMBER trajectories (see Appendix A7).

Linked Appendices: A1: Reprint Supporting Information Paper 1; A7: Hardware and Software.

Postprint of Paper 1

Haensele E, Banting L, Whitley DC, Clark T. Conformation and Dynamics of 8-Arg-Vasopressin in Solution. J Mol Model. 2014;20(11):2485(17).[i]

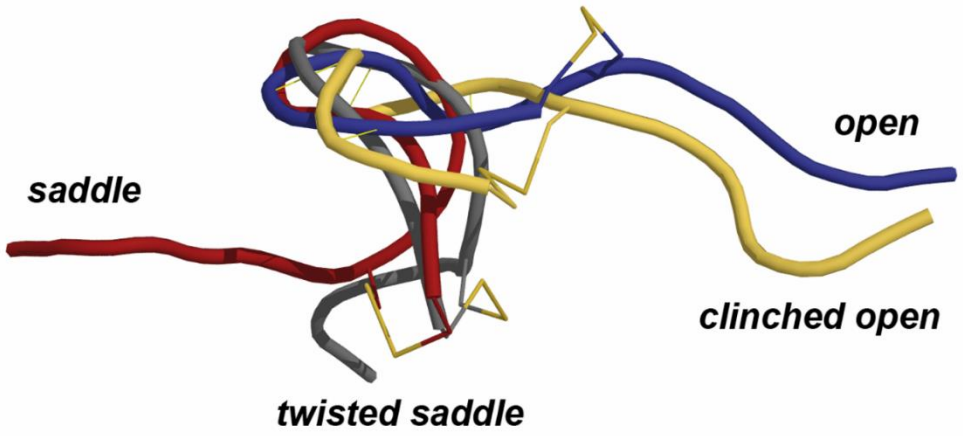


Table of Content Graphic
(Representative conformations of AVP)

Abstract

Arginine-vasopressin has been subjected to a long (11 μs) molecular-dynamics simulation in aqueous solution. Analysis of the results by DASH and principal components analyses reveals four main ring conformations that move essentially independently of the faster-moving tail region. Two of these conformations (labelled *saddle*) feature well defined β-turns in the ring and conserved transannular hydrogen bonds, whereas the other two (*open*) feature neither. The conformations have been identified and defined and are all of sufficient stability to be considered candidates for biologically active conformations in in their cognate receptors.

Keywords. Vasopressin –Molecular Dynamics – DASH Analysis – Peptides – Principal Component Analysis

[i] Elke Haensele,[a] Lee Banting,[a] David C. Whitley,[a] and Timothy Clark[a, b,✉]
[a] Centre for Molecular Design, School of Pharmacy and Biomedical Sciences; University of Portsmouth, Portsmouth PO1 2DT (United Kingdom); [b] Computer-Chemie-Centrum and Interdisciplinary Center for Molecular Materials; Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen (Germany)

Introduction

8-Arginine-vasopressin (AVP, also known simply as vasopressin (VP), antidiuretic hormone (ADH) or argipressin), one of the first biologically active peptides to be synthesised by du Vigneaud in 1954,[259] is a nonapeptide with a six-membered cyclic moiety (Cys^1-Tyr^2-Phe^3-Gln^4-Asn^5-Cys^6) closed by a Cys^1-Cys^6 disulphide bridge, and an α-amidated three residue tail (Pro^7-Arg^8-Gly^9-NH_2).

AVP is a neurohypophyseal hormone and belongs to the vasopressin family of the evolutionary lineage vasotocin-vasopressin. Vasopressin-like hormones are found in all vertebrates, with AVP being the mammalian form. They all possess a basic amino acid, such as arginine or lysine, in position eight and are all involved in water homeostasis (for reviews see *inter alia* [20,37,44]).

AVP is synthesised in the magnocellular neurons of the posterior pituitary gland[260] complexed with neurophysin, its carrier protein.[98] The function of NP is to target, package and store AVP before release into the bloodstream.[44] The receptors activated by AVP belong to the transmembrane G-protein coupled receptor superfamily.[20]

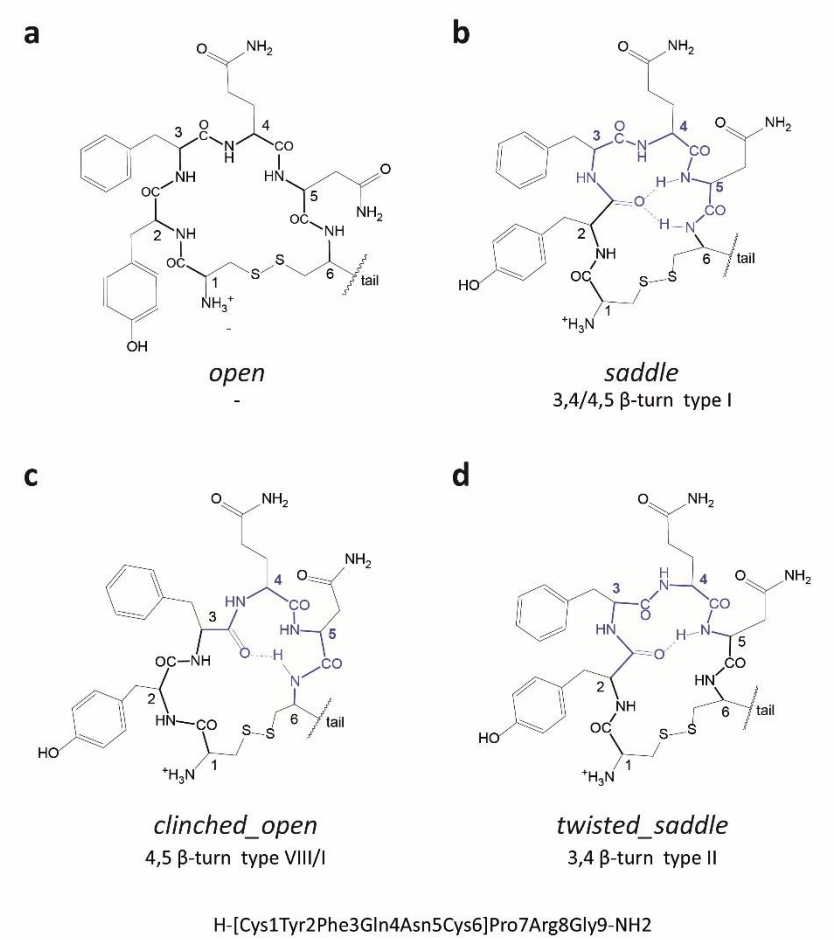
Once secreted into the blood stream, AVP is implicated in myriad physiological functions within the endocrine and neurocrine systems. Examples of its hormone function in addition to water homeostasis[20,37,44] include regulation of blood pressure,[261,262] antipyretic[263] and analgesic effects.[264] AVP acts as secretagogue for adrenocorticotropin,[36,41,44] glucagon and insulin.[265] The peptide is thought to mediate social and sexual behaviour, especially aggression, anxiety and pair-bonding.[42] Furthermore, AVP is believed to enhance memory and facilitate learning[44] and to be involved in the pathophysiology of clinical disorders such as autism,[266] and may even play a role in circadian rhythm misalignments, like jet lag.[267]

Lowered AVP release in humans effects an increased blood sodium concentration (hypernatremia), excessive urine production (polyuria) and thirst. This may in turn lead to diabetes insipidus treatable by administration of AVP and AVP analogues.[268] In contrast, heightened AVP release causes hyponatremia, which may result in brain diseases and lung cancer[269,270] and can be treated with AVP-receptor antagonists.[271] AVP can be used in emergency medicine as an alternative to epinephrine in the event of cardiac arrest.[39]

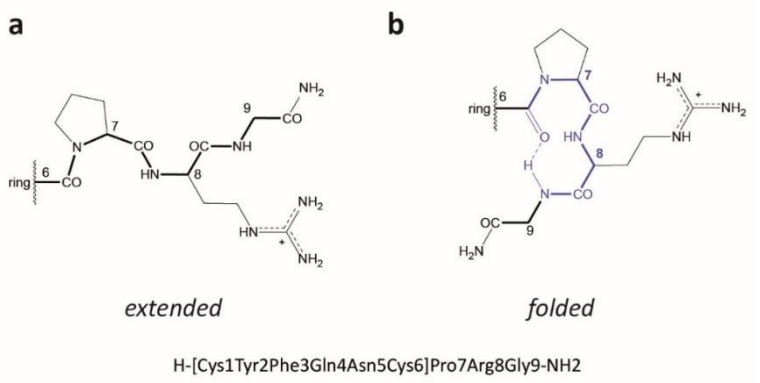
To date, the only fully resolved crystal structure of AVP is as part of a trypsin complex (PDB ID: 1YF4).[97] This structure contains a remarkably different backbone conformation to those found for the closely related peptide hormones 8-Lys-vasopressin (PDB ID: 1JK)[98] and oxytocin (PDB ID: 1NPO)[77] in their NP-complexes in the solid state.

The conformational characteristics of the peptide structures in the physiologically relevant neurophysin-complexes are a saddle-like ring with β-turns involving residues 3,4/4,5 and a high occurrence of transannular hydrogen bonds, primarily between Tyr^2O and Asn^5NH

(*cf*. Scheme 4.1b). The tripeptide tail is only resolved in the OT-NP complex (PDB ID: 1NPO) where it is extended or folded and possibly stabilized by a hydrogen bond Cys^{6}O-Gly^{9}NH (*cf*. Scheme 4.2b).



Scheme 4.1a-d Main conformational types of the cyclic part of AVP. (**a**) *open*: no intramolecular hydrogen bonds and no classical β-turn types; (**b**) *saddle*: β-turn type I centred at 3,4/4,5 and stabilised by a transannular hydrogen bond from Tyr^{2}O to Asn^{5}NH and Cys^{6}NH; (**c**) *clinched open*: minor propensity for β-turns type VIII or I centred at 4,5; (**d**) *twisted saddle*: β-turn type II centred at 3,4 with hydrogen bond Tyr^{2}O to Asn^{5}NH



Scheme 4.2a,b Main conformational types of the N-terminal tail of AVP. (**a**) *extended* tail: no turns, no significantly populated hydrogen bonds; (**b**) *folded* tail: β-turn type II centred at residues 7 and 8, hydrogen bond from Cys6O to Gly9NH

NMR studies suggest rapid interchange between the β-turn conformations of AVP in solution, although a folded (*saddle*) geometry appears to be maintained.[99] The polarity of the solvent seems only to affect formation of intramolecular hydrogen bonds. In DMSO, a hydrogen bond is indicated between Tyr^2O and Asn^5NH^{99} but apparently not in water.[103] Studies in sodium dodecyl sulphate (SDS) micelles suggest the lipophilic regions of the ring interact with a membrane, while the hydrophilic tail is exposed to the aqueous phase. Again, in this study the cyclic backbone of the AVP ring attached to the micelles appears similar to the NP-complexed form.[100]

These *saddle*-like conformations with a strongly puckered ring and the β-turns mentioned above have been confirmed computationally as "low-energy conformations" *inter alia* by Liwo *et al*.[81] *via* Monte Carlo and molecular-dynamics simulations.

In contrast, the conformation of AVP within the trypsin complex (PDB ID: 1YF4) is characterised by an unfolded, more planar ring conformation, here designated as *open*, with no significant internal hydrogen bonds and an *extended* tail (*cf*. Scheme 4.1a). AVP is an efficient inhibitor of trypsin,[97] although this is not known to be a true physiological function of AVP. The *open* conformation adopted in this trypsin complex can nevertheless be regarded as a bioactive conformation.

To our knowledge, little attention has been paid to an *open* conformation or its potential role in receptor binding with the vasopressin-receptor V2R.[272]

The V2R agonist-binding-pocket, common to all VP and OT receptor types, is located in a cleft within the transmembrane (TM) domains and AVP has been proposed to be almost completely buried within the receptor channel.[272,273] The hydrophobic ring residues (Cys^1-Tyr^2-Phe^3) are predicted to interact with residues of the TM-helices to activate signal transduction, while the tail points outside the TM-core, interacting with an extracellular loop *via* its hydrophilic residue Arg^8. The interaction between Arg^8 and the extracellular loops is also thought to be a key in receptor recognition.[20,22]

Current models for interactions of peptide hormones with their receptors suggest multi-step mechanisms in which the peptide first contacts the cell membrane and then diffuses to the receptor until it finally finds its position to trigger receptor activities.[24,25] These events are probably accompanied by conformational changes of the ligand and concomitant allosteric effects on the receptors.[63] A flexible ligand exists in solution as an equilibrium involving several conformations of differing bioactivities. A conformation that has not yet been recognised with "slow" experimental techniques, such as NMR, might nevertheless be the important conformation for triggering biological effects such as receptor recognition and activation or inhibition.[63,70,274]

Thus, we have now investigated the conformational dynamics of this peptide in solution in depth with modern computational methods and analysis tools with special regard to the *open*

conformation, which is evident in the largely ignored 1YF4 X-ray structure of AVP and is significantly different from the known *saddle* conformation.

Molecular-dynamics simulations have proven to be an accurate tool for describing the atomistic details of the conformational dynamics of biological systems in solution (*e.g.*[225]). Rapidly developing computational methods, increasing computational performance and improved force fields now make it possible to reveal new structural aspects of systems such as AVP, especially because microsecond simulations are now possible for a peptide of this size.

We now report an unrestrained 11 µs MD simulation of the AVP-1YF4-peptide in explicit water solvent at 300 K using AMBER 10[219] and a detailed analysis of the resulting conformational space with several analysis tools contained in Ptraj[142] and DASH[141] - a fast conformational analysis tool for MD simulations developed especially for long trajectories for which classical clustering algorithms scale poorly.

Methods

Molecular-Dynamics Simulation

The AMBER 10 program suite[219] was used to optimise geometries and for the MD simulations. The X-ray structure of AVP from the trypsin complex (PDB ID: 1YF4)[97] was chosen as the initial conformation. The peptide was placed in a truncated octahedron water box (box size (XYZ) = 38.97 $Å^3$) using the TIP4P-Ew water model.[197,275] Two chloride counterions were added to neutralise the system. The simulation system consisted of a total of 4,792 atoms, including 1,162 4-site water molecules and 142 AVP atoms.

The system was optimised using 500 steps of steepest-descent optimisation followed by 8,945 of conjugated-gradient minimisation at constant volume.

Molecular-dynamics simulations were carried out using the AMBER ff99SB force field[140] under constant temperature (T = 300 K, Berendsen coupling[215] of 1.0 ps to an external heat bath) and constant pressure (p = 1 atm) periodic boundary conditions with a non-bonded cut off of 8 Å. The SHAKE[224] algorithm was employed for hydrogen atoms with a simulation time step of 2 fs. Energies were calculated using the Particle Mesh Ewald method[203] and coordinate 'snapshots' were written every picosecond. AVP was simulated in explicit water at 300 K for 11 µs.

DASH Analysis

Conformational clustering was performed with DASH, Version 2.10.[141] DASH is a fast conformational analysis tool for MD simulations developed especially for long trajectories for which classical pairwise distance-metric clustering algorithms (Cα) scale poorly. It analyses time series of torsion angles, *e.g*. the trajectories of the $\Phi\Psi$ dihedral angles of the protein/peptide backbone during the MD simulation. The result is a time series of DASH states called a DASH state trajectory. A DASH state is simply an ensemble of torsion angles that is representative for a main conformation (equivalent to a conformational cluster). No predetermined number of states is required, in contrast to clustering algorithms that use a similarity matrix, such as those implemented in AMBER tools.[142] A conformation must persist for a minimum number of time steps before it is identified as a DASH state, which gives an accurate representation of significant conformational changes. The DASH software is released under the terms of the GNU General Public License and can be downloaded from the University of Portsmouth website.[276]

Principal Component Analysis

The principal component analysis was conducted using the dihedral angles extracted from the simulation (11,000 snapshots) using SAR-caddle.[234] Kaiser's eigenvalue-one test[230] was used to determine the number of significant PCs. Weights are simply the squares of the coefficients of the torsional angles in the relevant PC.

Further details of the calculations and analyses are given in the Supporting Information (Appendix A1).

Results and Discussion

Ring Conformations

Trajectory (Transitions). An 11 µs MD simulation of AVP in solution reveal the high conformational flexibility and fluctuation of this peptide (see SI Video S1). Figure 4.2a shows the trajectory of conformational changes of the Cα-backbone atoms 1 to 9 of AVP as root mean square deviation from the minimised starting conformation (PDB ID: 1YF4). Average RMSD values of distinct time windows from the trajectory are given in Table 4.1. Significant RMSD changes indicate significant conformational changes, but despite a high fluctuation, there are only few substantial RMSD changes during the 11 µs MD simulation. The most obvious transition is at 1.46 µs. Limiting the RMSD calculation either to the ring (Fig. 4.2b) or to the tail Cα-atoms (Fig. 4.2d) shows that the major overall transition of the peptide (Fig. 4.2a) corresponds to a change of the ring conformation. The radius of gyration of the ring system (Fig. 4.2c) reveals further distinct transitions between differently folded ring conformations at 5.90, 6.43 and 7.19 µs. The tail, however, fluctuates with a much higher frequency, apparently between two conformational states that are distributed evenly over the simulation. The video clip (SI, Video S1) suggests that these two tail states may be assigned to an *extended* state (Scheme 4.2a), in which the tail points away from the ring, and a *folded* state (Scheme 4.2b) in which the tail turns toward the lower face of the ring. The high frequency of transitions between the two tail states indicates a high flexibility of the tail, significantly higher than the ring.

Table 4.1 Average root mean square deviations (avRMSD) and average radii of gyration (avRadGyr) for significant trajectory time windows and backbone Cα alignments of Arg^8-vasopressin

Alignment		Trajectory time window [µs]				
		0-1.46 *open*	1.46-5.90 *saddle*	5.90-6.43 *variants*	6.43-7.19 *clinched open*	7.1-11.00 *twisted open*
avRMSD [Å]	*overall* [a]	1.825	2.930	2.683	2.274	2.756
	ring [b]	0.950	1.807	1.765	1.592	1.717
	tail [c]	0.991	1.157	1.102	1.164	1.068
avRadGyr [Å]	*ring*	4.077	3.278	3.569	3.870	3.500

[a] Cα 1-9 alignment; [b] Cα 1-6 alignment; [c] Cα 7-9 alignment. Abbreviations: see p. xii.

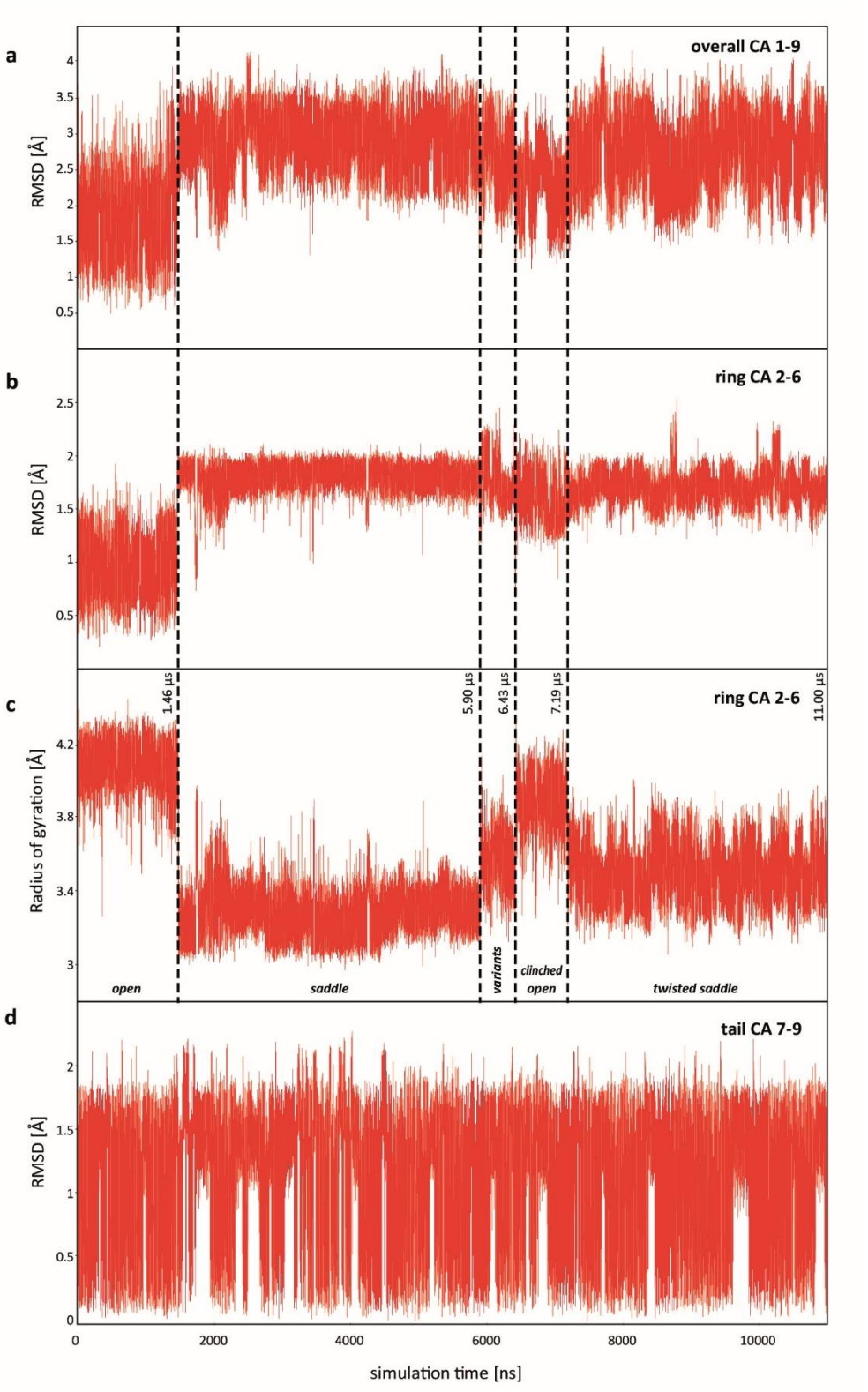


Figure 4.2a-d Root mean square deviations and radius of gyration (RadGyr) of Arg^8-vasopressin during 11 µs MD simulation (Reference: minimised initial MD structure, AVP_{1YF4}). (**a**) RMSD of Cα-backbone atoms 1to 9 (*overall*); (**b**) RMSD of Cα-backbone atoms 2 to 6 (*ring*); (**c**) Radius of gyration of Cα-backbone atoms 2 to 6 (*ring*); (**d**) RMSD of Cα-backbone atoms 7 to 9 (*tail*). Dotted lines indicate significant changes of the RMSD/RadGyr and mark time windows of different ring conformations (denoted as *open, saddle, variants, clinched open* and *twisted saddle*)

DASH State Analysis. A DASH analysis of all 16 $\Phi\Psi$ dihedral angles (T16) of the AVP backbone (Cα 2 to 9) during the 11 µs MD simulation results in 35 conformational states. Every DASH state represents, like a cluster, a conformation that is representative for an ensemble of similar backbone conformations. The 35 overall states can be clustered into four groups of states with common structural characteristics for the cyclic part of the peptide (Fig. 4.3a-d) and a fifth group of states ("*variants*", Appendix A1 Table S1 and Fig. S1) that does not match one of the main groups. This fifth group occurs between 5.90 and 6.43 µs. A detailed table of the sequence of DASH states (DASH state trajectory) during the 11 µs simulation is available as Supplementary Material (Appendix A1 Table S2).

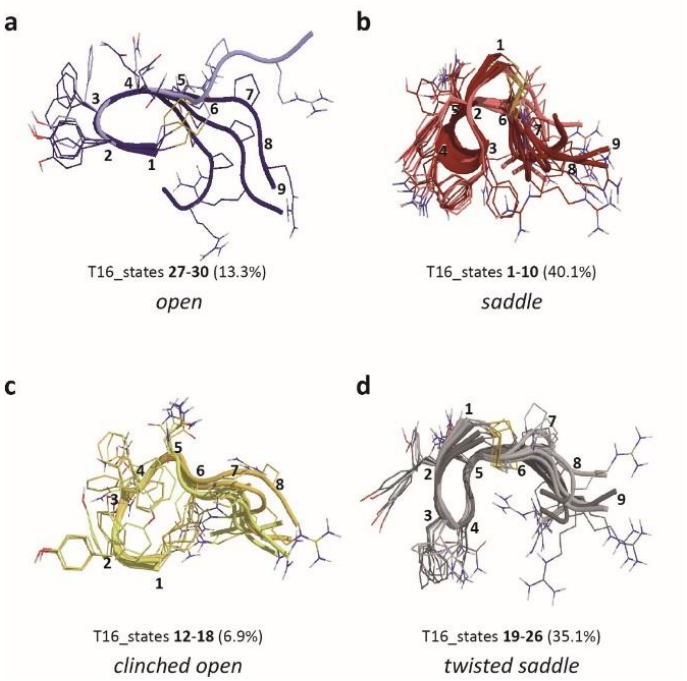


Figure 4.3a-d Main overall conformations of Arg[8]-vasopressin: Representative states in water resulting from a DASH state analysis of backbone dihedrals $\Phi\Psi$ 2 to 9. Absolute populations for every group of conformations are given in parentheses and refer to 11 µs MD. **(a)** Representatives with *open* ring conformation; **(b)** representatives with *saddle* ring conformation; **(c)** representatives with *clinched open* ring conformation; **(d)** representatives with *twisted saddle* ring conformation. Depiction: backbone = cartoon, side chains = lines, representatives are only labelled for the main populated state of each group. Residues are only labelled for each major populated state

DASH Ring-State Analysis. As the RMSD trajectories suggest that the tail movements do not affect the main ring conformation, we first focused the DASH analysis on the ring dihedrals $\Phi\Psi$ 2 to 6 (T10). Each DASH state now represents an ensemble of similar ring-backbone conformations shown in Figure 4.4 and Table 4.2. In order to distinguish between DASH overall states and DASH ring states, DASH overall states are denoted as T16 and DASH ring states as T10.

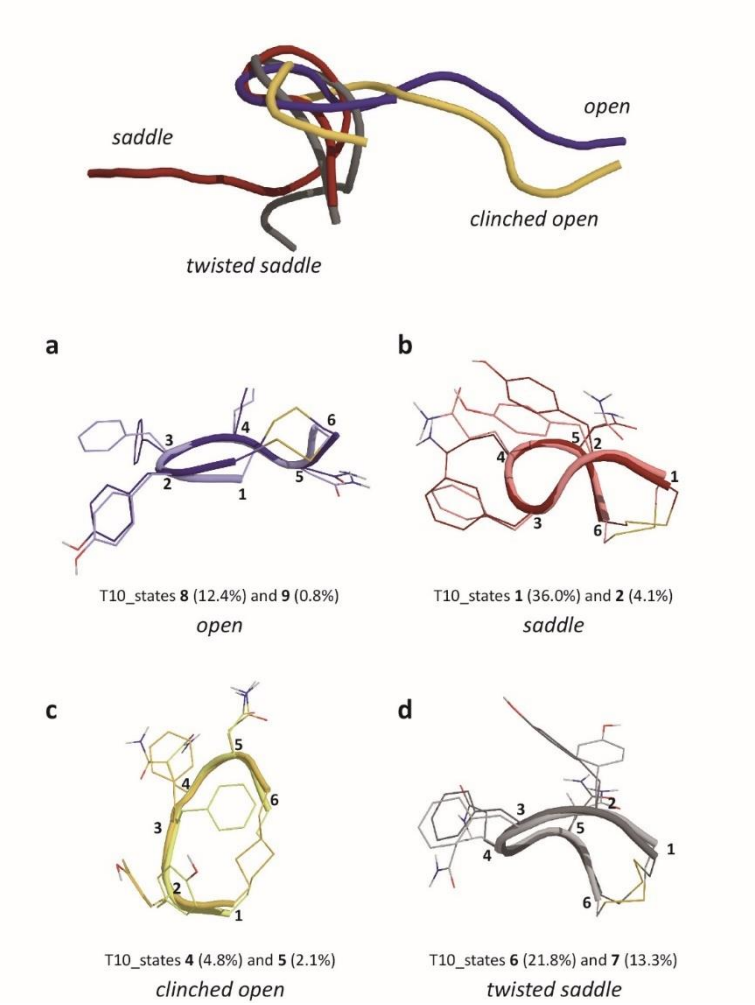


Figure 4.4a-d Main *ring* conformations of Arg^8-vasopressin. Main representative ring states in water resulting from a DASH state analysis of backbone dihedrals $\Phi\Psi$ 2 to 6. Absolute population of each state during 11µs MD are given in parentheses. (**a**) *open* ring states; (**b**) *saddle* ring states; (**c**) *clinched open* ring states; (**d**) *twisted saddle* ring states. Depiction: backbone = cartoon, sidechains = lines. Residues are only labelled for the major populated state each and the N-terminal tail is not shown in (**a**) to (**d**) for clarity. For illustration, a ring alignment of the backbone cartoons of the 4 main ring state (T10_8,1,4,6) including the tail, is shown above (**a**) to (**d**).

The initial 35 overall states are now reduced to twelve ring states. These ring states can be assigned clearly to the main time windows of the trajectory between the transitions at 1.46, 5.90, 6.43, and 7.10 µs (Fig. 4.2a-c). Furthermore, analysing the T10 and T16 DASH state trajectories (Fig. 4.5 and Appendix A1 Table S2), shows that each overall state can be assigned to a distinct ring state (see Table 4.2).

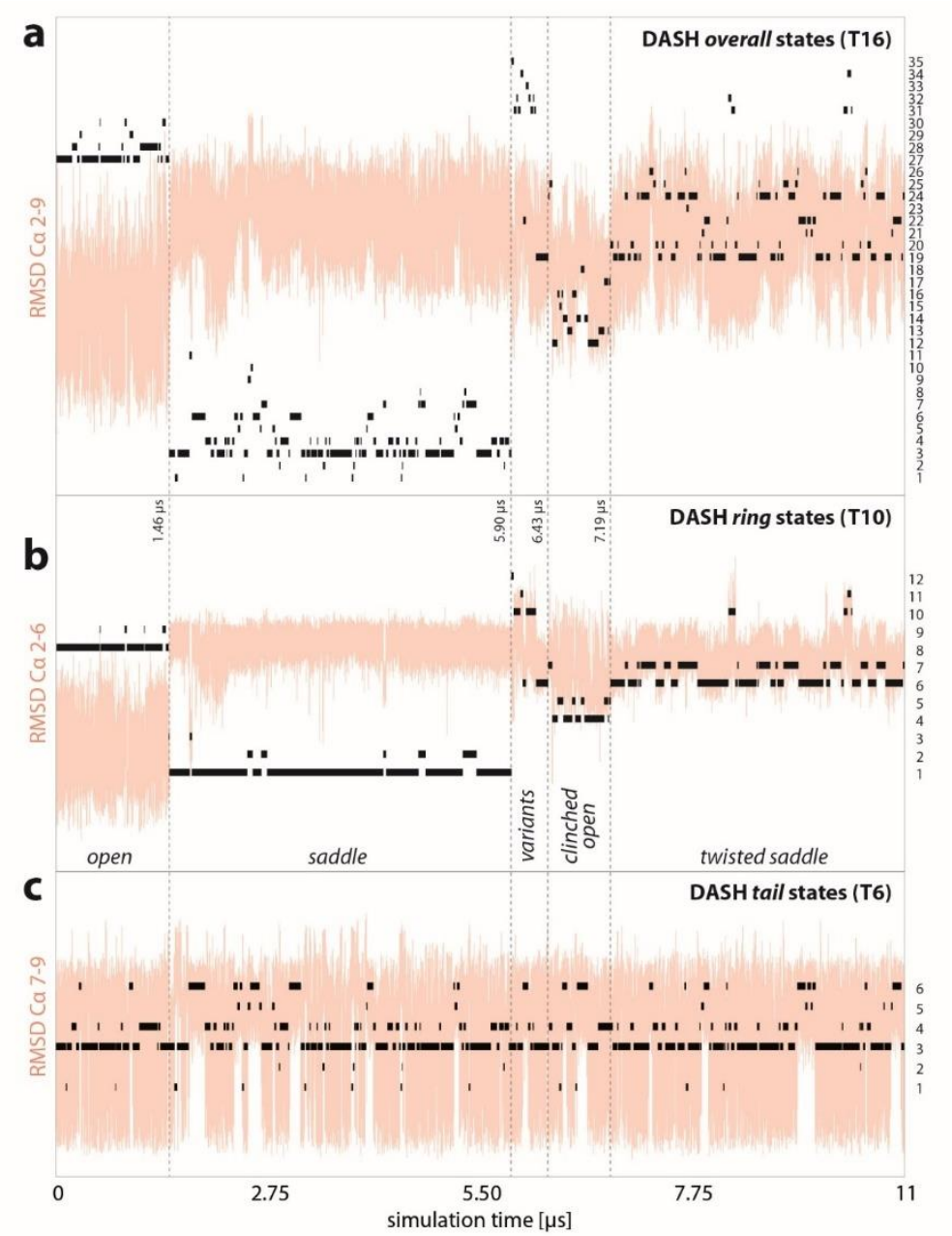


Figure 4.5a-c DASH state trajectories for (**a**) *overall* (T16), (**b**) *ring* (T10), and (**c**) *tails* (T6) states. For a better understanding, the corresponding RMSD trajectories for *overall* (Cα 2 to 9), *ring* (Cα 2-6) and *tail* (Cα 7-9) alignments are shown in the background. States are numbered consecutively on the second y-axis, thus every horizontal line is the trajectory of a single DASH state and illustrates its individual distribution during the simulation. The transitions between time windows of main ring conformations are marked with vertical dashed lines.

Table 4.2 Representative states of the main overall and ring conformations of Arg[8]-vasopressin [§]

T16 State	State population (T16) abs [%]	rel [%]	T10 State	State population (T10) abs [%]	rel [%]	Conformational characteristics β-turn type [a]	Turn centre	H bonds [b]
				open (0 to 1.455 µs = 1.455 µs)				
27	8.62	64.74	**8**	12.40	93.75			
28	3.25	24.60	**8**					
29	0.65	4.95	**8**					
30	0.75	5.70	**9**	0.83	6.25			
total	*13.28*	*100.00*		*13.23*	*100.00*	**no classical turns**	**2,3**	**(Tyr[2]OGln[4]NH)**
				saddle (1.455 to 5.900 µs = 4.445 µs)				
1	0.95	2.34	**1**	35.97	89.02			
2	0.74	1.82	**1**					
3	19.65	48.63	**1**					
4	7.88	19.51	**1**					
5	1.22	3.03	**1**					
6	5.60	13.87	**1**					
7	3.07	7.60	**2**	4.10	10.13			
8	0.35	0.87	**2**					
9	0.35	0.87	**2**					
10	0.25	0.62	**2**					
total	*40.06*	*99.15*		*40.07*	*99.16*	**I/(I)**	**3,4/4,5**	**Tyr[2]OAsn[5]NH, Tyr[2]OCys[6]NH**
				clinched open (6.429 to 7.187 µs = 0.758 µs)				
12	1.87	27.18	**4**	4.80	69.66	**(VIII)**		
13	1.41	20.45	**4**					
14	1.45	20.98	**4**					
15	0.23	3.30	**5**	2.09	30.34	**I**		
16	0.89	12.93	**5**					
17	0.62	8.97	**5**					
18	0.43	6.20	**5**					
total	*6.89*	*100.00*		*6.89*	*100.00*	**(VIII) /I**	**4,5**	**(Phe[3]OCys[6]NH)**
				twisted saddle (7.187 to 11.000 µs = 3.813 µs)				
19	14.04	36.32	**6**	21.8	57.62			
20	3.24	9.34	**6**					
21	0.85	2.44	**6**					
22	3.81	9.91	**6**					
23	0.26	0.76	**7**	13.33	37.11			
24	9.84	27.98	**7**					
25	2.07	5.01	**7**					
26	1	2.86	**7**					
total	*35.10*	*94.62*		*35.13*	*94.73*	**II**	**3,4**	**Tyr[2]OAsn[5]NH**
Σtotal	***95.33***			***95.32***				

[§] Listed are the population and conformational characteristics of the main overall states (T16) and ring states (T10) of AVP (minor and transient state variants: see Table S1). Absolute populations refer to 11 µs MD (100 %). Relative populations refer to the main time windows of each conformational group (*open, saddle, clinched open* or *twisted saddle*). Characteristics of each ring conformation are given by β-turn types, turn centres, and transannular hydrogen bonds (Hbonds). T16 = overall states defined by $\Phi\Psi$ 2 to 9; T10 = ring states defined by $\Phi\Psi$ 2 to 6. [a] Parentheses indicate distorted versions of ideal β-turn types. [b] Hydrogen bonds in parentheses are only populated 20-40 %. Abbreviations: see p. xii.

In other words, each overall state can be considered as a main ring conformation combined with a distinct tail-conformation, as will be discussed in detail below. Table 4.2 shows absolute and relative populations of overall and ring states and how they correspond. Absolute populations refer to the total simulation time of 11 µs and relative populations refer to the individual lengths of a conformational time window. The main ring conformations and the corresponding main windows are denoted as (a) *open* (0 to 1.46 µs), (b) *saddle* (1.46 to 5.90 µs), (c) *clinched open* (6.43 – 7.19 µs)

and (d) *twisted saddle* (7.19 to 11µs) to reflect common structural characteristics. The fifth window identified on the RMSD plot between 5.90 and 6.43 µs contains variants of the main ring conformations and will not be discussed in detail here. This work is focused on AVP´s main conformational states, which correspond to the four main trajectory windows. Other, short-lived states are observed during the simulation, but do not play a significant role and will only be defined in the Supporting Information (Appendix A1).

Populations of the Conformations. The four main ring conformations (*open*, *saddle*, *clinched open* and *twisted saddle*) are present for more than 95 % of the simulation time. As a result of the DASH ring analysis, every conformational group is represented by two ring states (T10), a main state with a relative population of up to 94 % and a less populated state. Both major and minor states are present for 95 to 100 % of the relevant time window. Figure 4.4 shows the Cα 1 to 6 alignment of the ring states for each main ring conformation. Frequent interconversions between the representative states occur within each main conformational window. The DASH state mean angles (Table 4.3 and Appendix A1 Fig. S2) reveal that the major and minor ring states of a distinct conformational group (*open, saddle, twisted saddle, clinched open*) differ significantly (> 60°) for only one torsional angle. This is *Φ* 6 for the *open* (72 %) and *saddle* (66°) states and *Ψ* 5 for the *clinched open* states (73°).[i] The maximum torsion difference for *twisted saddle* states is *Φ* 6 = 52°.

Table 4.3 DASH state mean angles (*ΦΨ*) of the main ring states (T10) of Arg[8]-vasopressin

T10 state	Tyr[2] *Φ*	Tyr[2] *Ψ*	Phe[3] *Φ*	Phe[3] *Ψ*	Gln[4] *Φ*	Gln[4] *Ψ*	Asn[5] *Φ*	Asn[5] *Ψ*	Cys[6] *Φ*	Cys[6] *Ψ*
open										
8	-112.54	134.53	55.31	3.41	-135.33	152.15	-75.09	124.68	-127.16	148.39
stddev	*37.81*	*18.46*	*9.08*	*31.34*	*23.92*	*18.2*	*18.32*	*32.08*	*31.88*	*23.33*
9	-98.98	129.37	56.09	0.76	-135.73	153.49	-66.41	113.78	-55.29	126.93
stddev	*54.64*	*26.48*	*9.27*	*31.72*	*23.88*	*23.1*	*31.29*	*80.23*	*61.52*	*40.84*
saddle										
1	-80.2	143.87	-62.88 [a]	-21.36 [a]	-86.73 [a]	-7.38 [a]	-113.37	-27.13	-126.42	133.12
stddev	*20.52*	*12.37*	*9.44*	*13.4*	*17.2*	*16.94*	*21.14*	*22.14*	*20.16*	*33.23*
2	-84.29	147.09	-57.99 [a]	-27.01 [a]	-85.13 [a]	-7.63 [a]	-122.03	-6.72	-60.49	142.38
stddev	*23.05*	*13.93*	*10.95*	*15.59*	*17.53*	*16.62*	*20.55*	*41.47*	*32.18*	*24.51*
clinched open										
4	-95.37	-19	-101.27	156.57	-67.65 [b]	-19.06 [b]	-112.46 [b]	86.89 [b]	-117.42	145.84
stddev	*28.15*	*22.75*	*29.46*	*14.56*	*16.85*	*23.84*	*28.66*	*61.3*	*36.21*	*21.54*
5	-90.52	-18.35	-116.2	151.18	-68.06 [b]	-20.5 [b]	-88.17 [b]	14.01 [b]	-82.72	144.88
stddev	*28.3*	*18.64*	*30.65*	*13.16*	*22.02*	*26.74*	*20.39*	*33.03*	*29.6*	*16.17*
twisted saddle										
6	-86.02	162.33	-52.48 [a]	127.66 [a]	55.04 [a]	12.34 [a]	-107.29	-7.44	-122.17	144.18
stddev	*29.44*	*13.88*	*16.16*	*14.69*	*9.01*	*21.14*	*29.86*	*48.29*	*28.23*	*23.53*
7	-115.65	174.87	-52.78 [a]	129.79 [a]	57.39 [a]	8.38 [a]	-114.1	-16.45	-70.67	148.3
stddev	*24.26*	*19.63*	*19*	*13.91*	*8.24*	*20.56*	*25.07*	*29.84*	*19.33*	*13.72*

[a,b] Torsions corresponding to β-turns: [a] turn propensity > 80 %, [b] turn propensity 40-65 %. Ideal *ΦΨ* (i+1, i+2): β-turn type I (-60°, -30°, -90°, 0°), type II (-60°, +120°, +80°, 0°), type VIII (-60°, -30°, -120°, 120°).[227,277] stddev= standard deviation.

[i] Sentence rephrased for a better understanding.

These result in different disulphide-bridge conformations for each state. Although the disulphide-bridge torsions were not included in the DASH ring-state analysis, their conformations are probably characteristic (see Fig. 4.4a-d, disulphide-bridges are shown as lines). RMSD differences between states of the same ring conformation are small, ≤ 0.25 Å, in comparison to RMSD differences between states of different ring conformations, 0.9 to 2.2 Å (Appendix A1 Table S3). The *saddle* and the *twisted saddle* ring conformations are the most populated structures, with absolute populations of 40 and 35 %, respectively. The *open* conformations, *open* and *clinched open*, occur only 13 % and 7 % of the time.

Secondary Structure and Hydrogen Bonds. The secondary structure was determined by means of ring-internal turn propensities, turn types and hydrogen bonds. Turn propensities and hydrogen-bond occupancies were calculated using AmberTools, and turn types were identified by comparing the DASH mean-angles with ideal β-turn type torsions. Turn propensities and hydrogen-bond populations are given in Tables 4.4 and 4.5 and the torsion-angle ensembles for every main DASH ring state in Table 4.3 and the results are illustrated in Scheme 4.1.

Table 4.4 Turn propensities [%] for the main ring conformations of Arg[8]-vasopressin

Turn centre residue	Main ring conformation / Trajectory time-window			
	open	*saddle*	*clinched open*	*twisted saddle*
Cys[1]	0.00	0.00	0.00	0.00
Tyr[2]	20.20	0.00	0.00	0.18
Phe[3]	20.20	94.10	0.03	90.57
Gln[4]	0.08	93.93	*46.28*	93.80
Asn[5]	0.07	89.33	*46.28*	*61.93*
Cys[6]	0.00	2.53	0.00	0.01
Pro[7]	3.80	18.57	20.99	10.82
Arg[8]	3.80	17.32	20.99	10.82
Gly[9]	0.00	0.00	0.00	0.00

Abbreviations: see p. xii.

Table 4.5 Occupancies of intramolecular hydrogen bonds (%) and corresponding turn centres for the main ring conformations of Arg[8]-vasopressin

H-bond		Main ring conformation / Trajectory time-window				Turn centre residues
O…	…HN	*open*	*saddle*	*cl.open*	*tw.saddle*	
Cys[1]	Gln[4]	12.03	0.00	0.00	2.24	2, 3
Tyr[2]	Asn[5]	0.00	95.70	0.00	82.60	3, 4
Tyr[2]	Cys[6]	0.00	83.19	0.00	37.28	3, 4, 5
Tyr[2]	Gln[4]	38.57	2.04	0.01	0.02	3
Phe[3]	Cys[6]	0.00	4.86	27.93	0.04	4, 5
Phe[3]	Asn[5]	0.13	2.41	10.21	23.83	4
Gln[4]	Cys[6]	8.78	0.06	18.67	1.60	5
Asn[5]	Tyr[2]	0.00	0.22	0.00	6.59	3, 4
Cys[6]	Gly[9]	2.20	10.82	12.11	5.67	7, 8

The *saddle* (Scheme 4.1b, Fig. 4.4b) and related *twisted saddle* (Scheme 4.1d, Fig. 4.4d) ring conformations are the most highly populated, occurring for 75 % of 11 µs (Table 4.2). Both feature

a highly populated (more than 90 %) turn at residues Phe^3 and Gln^4. The *saddle* is characterised by a further turn centred at Asn^5 (89 %). This turn also occurs in the *twisted saddle* but is less highly populated (62 %). The DASH-state mean-angles (Table 4.3 and Appendix A1 Table S4) reveal a β-turn type I centred at 3,4 for the *saddle* conformation in addition to a slightly distorted β-turn type I centred at 4,5. These β-turns are stabilised by highly populated hydrogen bonds (83-96 %) between the carbonyl-oxygen of Tyr^2 (Tyr^2O) and the amide-hydrogen of Asn^5 (Asn^5NH), and between Tyr^2O and Cys^6NH. In the *twisted saddle* ring conformation, however, only the Tyr^2O-Asn^5NH hydrogen bond is highly populated (83 %) and the 3,4 centred turn is now a classical β-turn type II. The difference between a β-turn type I and type II is simply the orientation of the central peptide bond 3,4.

The *twisted saddle* conformation shows a slight tendency to form a Tyr^2NH-Asn^5O hydrogen bond (7 % occupancy). A Tyr2NH-Asn5O hydrogen bond has also been suggested on the basis of NMR experiments.[99] A rearrangement from *saddle* to *twisted saddle* or changing the 3,4 β-turn from type I to type II twists the ring making it more open (*cf*. radius of gyration, Fig. 4.2c), whereas the main orientation of the side chains 2 to 4 remains unchanged. This may be necessary to facilitate AVP entry and/or fit into different GPCR pockets. There is no direct transition between *saddle* and *twisted saddle* in the 11 μs MD, suggesting that interconversion of the two ring conformations may occur *via* one or more conformational intermediates, *e.g*. the *clinched open* conformation or the variants that are observed between 5.90 and 6.43 μs.

The relatively sparsely populated *clinched open* ring conformation (Scheme 4.1c, Fig. 4.4c) is significantly less folded than the two *saddle* conformations but more than the *open*. The most highly populated intramolecular hydrogen bond is Phe^3O-Cys^6NH (28 %, Table 4.5) and turns centred at Gln^4 and Asn^5 occur about half of the time (46 %, Table 4.4). Loosely defined turns are thus more likely than ideal β-turns. The DASH ring-state mean-angles show a very high fluctuation of $\Psi 5$ (standard deviation = ±61°, Table 4.3) so unambiguous turn-type assignment is not possible. The *clinched open* ring structure can be classified as a flexible ring conformation with a tendency to form a β-turn type VIII or type I centred at 4,5.

Finally, the *open* ring conformation, the starting conformation for the simulation, shows none of the ideal turn structures (Scheme 4.1a, Fig. 4.4a). There is a slight tendency to centre ring turns at residues Tyr^2 and Phe^3 (20 %, Table 4.5) accompanied by sparsely populated hydrogen-bonding interactions of Tyr^2O and Gln^4NH (39 %), and Cys^1O and Gln^4NH (12 %). Summarising, this ring conformation can be readily classified as *open* as it has no significantly populated intramolecular hydrogen bonds and no defined β-turns.

Transition Key Torsions. A torsion angle is defined as a key torsion if its value changes significantly (> 90°) from one ring conformation to another. Figure 4.6 shows the differences of the mean torsions for the main ring conformations *open*, *saddle*, *clinched open* and *twisted saddle* represented by the main ring states (T10, Table 4.2). Only dihedrals *Φ* 2, *Φ* 5 and *ΦΨ* 6 do not show large differences, all other torsions can be qualified as key torsions for interconversions between the main ring conformations. A complete list of key-torsion angle differences between main ring conformations is given in Table S5 (Appendix A1).

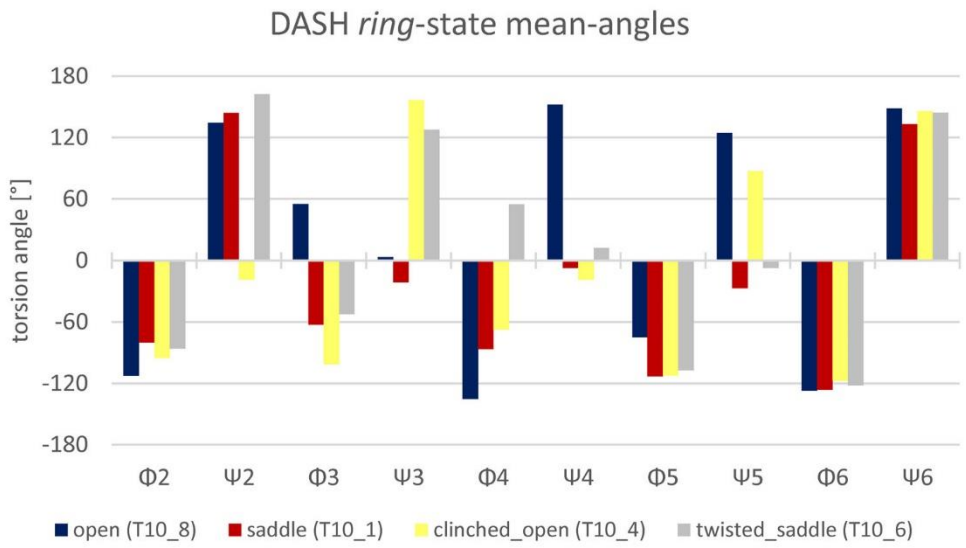
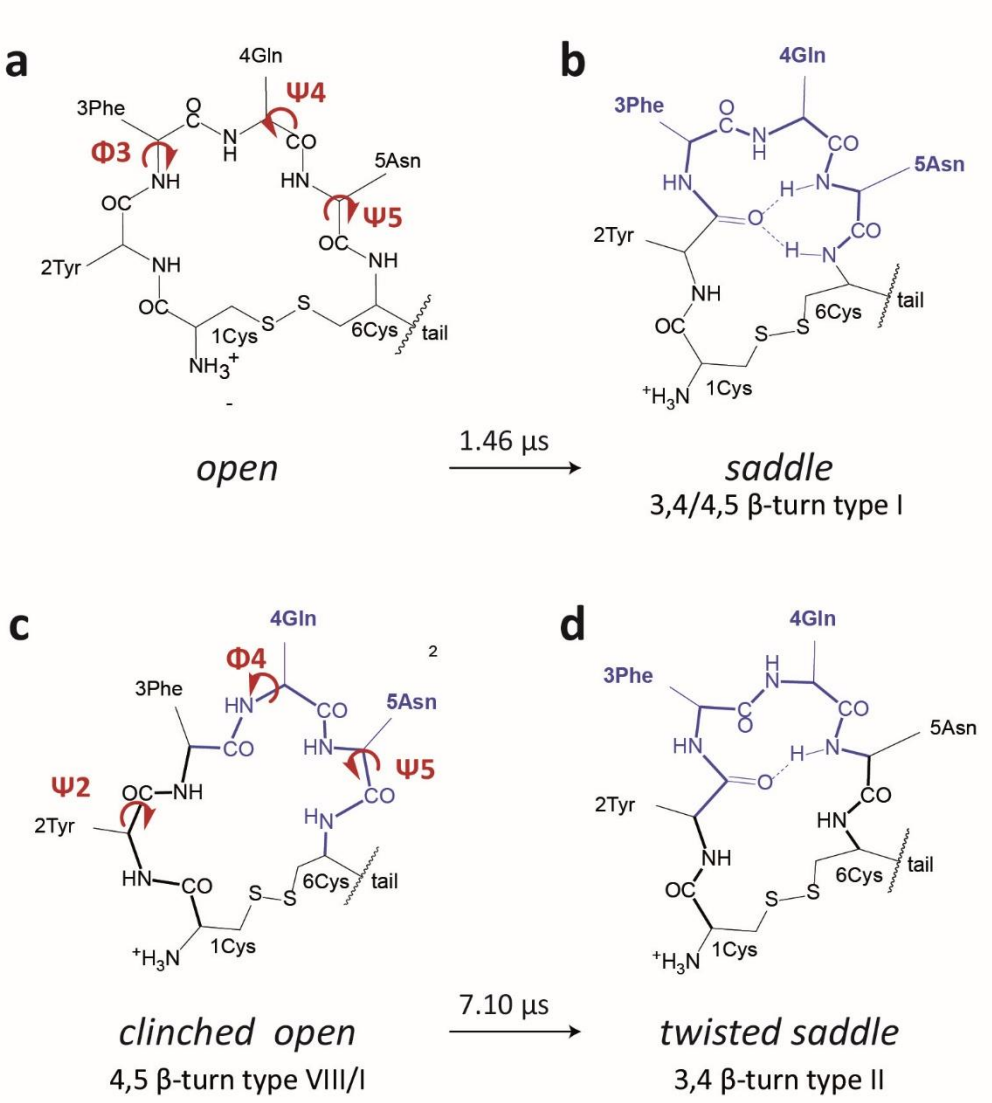


Figure 4.6 DASH state mean angles (*ΦΨ*) of the main ring conformations of Arg[8]-vasopressin. *open, saddle, clinched open,* and *twisted saddle* represented by the main ring states T10_8, 1, 4, and 6

Direct transitions only occurred between (i) *open* and *saddle* and (ii) *clinched open* and *twisted saddle*. The key torsions for these transitions are (i) *Φ* 3, *Ψ* 4, *Ψ* 5 (Scheme 4.3a,b) and (ii) *Ψ* 2, *Φ* 4, *Ψ* 5 (Scheme 4.3c,d). Changes of these torsions correlate with rotations of the corresponding peptide bonds and the relative orientation of carbonyl oxygens and amide hydrogens, and elucidate the mechanism of interconversions. For example, to convert *open* to *saddle*, the Tyr^2 carbonyl-oxygen and the amide hydrogens of Asn^5 and Cys^6 should point into the ring. Torsions *Φ* 3, *Ψ* 4 and *Ψ* 5 are the key torsions responsible for turning these atoms into the ring and thus to enable the characteristic intramolecular hydrogen bond to be formed. To interconvert from *clinched open* to *twisted saddle*, the hydrogen bond Phe^3O-Cys^6NH must be replaced by one between Tyr^2O and Asn^5NH. This is accomplished by rotating *Ψ* 2 and *Φ* 4, which turns Tyr^2O into the ring displacing Phe^3O.

A concomitant rotation of $\Psi 5$ causes Cys^6NH to turn thereby weakening the hydrogen bond between Phe^3O and Cys^6NH. These ring interconversions have so far proved too complex for their thermodynamics to be determined by simple umbrella sampling and are therefore now being investigated using dual-topology thermodynamic integration.



Scheme 4.3a-d Key torsions for the interconversion of the main ring conformations of AVP. (**a, b**) interconversion *open* to *saddle* at 1.46 µs (11 µs MD); (**c, d**) interconversion *clinched open* to *twisted saddle* at 7.10 µs (11 µs MD)

Disulphide Bridge. One remaining important feature of the ring conformations is the chiral disulphide dihedral $\chi 3$ ($\sphericalangle Cys^2C\beta$-$Cys^2S$-$Cys^6S$-$Cys^6C\beta$), Figure 4.7 shows the dynamics of this torsion. The disulphide bridge adopts two main conformations for $\chi 3$ with average values of either +88.9° (g) or -86.6° (g'). Interconversions between these two states do not necessarily correspond to transitions between different time windows of the ring conformations but are rather frequent

independent occurrences. Each main ring conformation can exhibit conformations of the disulphide bridge conformations with negative and positive dihedrals. Transitions between these disulphide conformations occur independently of the main ring conformation. The positive torsion angle is favoured by 78.2 % to 21.8 % and is consistent with experimental evidence (see *e.g.*[110]). The simulation suggests that g/g' transitions are more frequent for *open* ring conformations than for *saddle*.

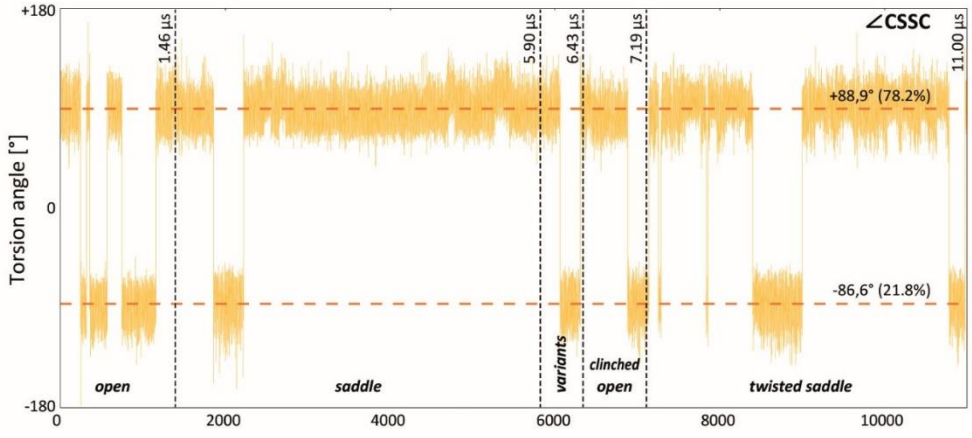


Figure 4.7 Trajectory of the disulphide-bridge torsion Cys2χ3 (∠CSSC). Horizontal dashed lines = average disulphide bridge torsions, vertical dashed lines = transitions between time windows of main ring conformations.

Tail Conformations

As described above, the RMSD trajectory of the Cα 7 to 9 segment (Fig. 4.2d) suggests two equally distributed main conformations for the tail, whereas the DASH analysis of the *tail* dihedrals *ΦΨ* 7 to 9 allows these dynamic conformations to be classified in detail. The results are given in Table 4.6, Scheme 4.2 and Figure 4.8. There are six distinct tail states (T6) that reveal two major tail conformations, (i) an *extended* tail conformation with no significant turns, and (ii) a tail conformation with a 7,8 β-turn type II, here denoted as *folded*. Each main conformation is represented by two DASH states, differing in torsions *ΦΨ* 9 (Appendix A1 Fig. S3). These torsions are only responsible for the orientation of the C-terminal $CONH_2$-group and do not affect the *extended* or *folded* conformation.

AVP favours the *extended* conformation of the tail significantly with an absolute population of 81 % during the simulation *vs.* 17 % for the *folded* 7,8 β-turn type II conformation. The preference for the

extended tail conformation is most likely due to the bulky residue Arg[8], which causes steric clashes when the tail is *folded*.

Two further transient conformations can be identified, a hybrid tail conformation (absolute population 2.0 %), which is not completely *extended* but has no defined folding, and a 7,8 β-turn type I structure (absolute population 0.8 %).

Table 4.6 Population and distribution of tail conformations[§]

T6 state [a]	Tail state population (T6)				
	(0-11.0 µs) abs [%]	***open*** (0-1.46 µs) rel [%]	***saddle*** (1.46-5.90 µs) rel [%]	***clinched open*** (6.43-7.19 µs) rel [%]	***twisted saddle*** (7.19-11.00 µs) rel [%]
			extended		
3	61.44	69.22	56.02	37.99	68.34
4	19.59	24.39	20.32	30.35	15.10
total	*81.03*	*93.61*	*76.34*	*68.34*	*83.44*
			7,8 β-turn type II		
5	2.52	0.00	3.87	0.00	2.75
6	13.61	4.91	15.31	26.12	12.51
total	*16.13*	*4.91*	*19.18*	*26.12*	*15.26*
			7,8 β-turn type I		
2	0.83	0.00	1.82	0.00	0.28
			distorted turn		
1	2.00	1.48	2.66	5.54	1.02
Σtotal	*100.00*	*100.00*	*100.00*	*100.00*	*100.00*

[§]The first column contains the DASH states (T6) that represent the tail conformations of Arg[8]-vasopressin during 11 µs MD simulation. Absolute populations refer to the total simulation time of 11 µs. Relative populations refer to the time windows of the main ring conformation (*open, saddle, clinched open,* and *twisted saddle*). T6 = tail states defined by *ΦΨ* 7 to 9. Abbreviations: see p. xii.

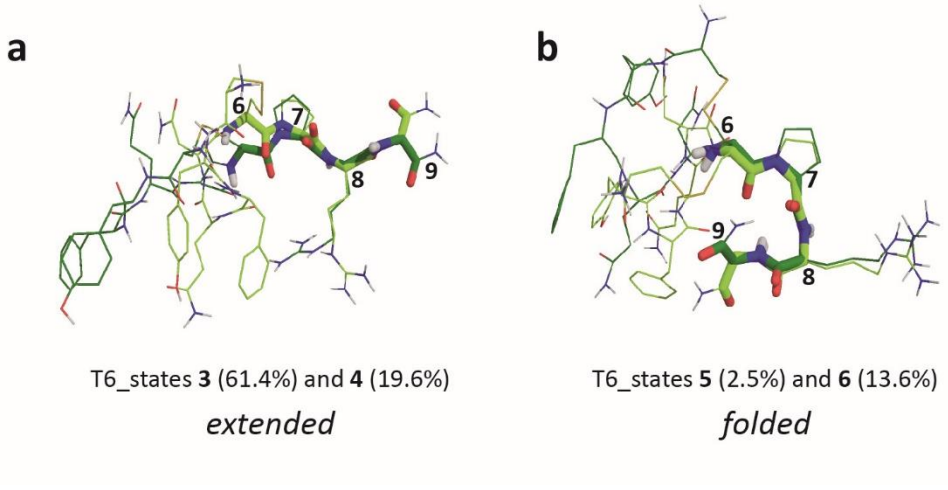


Figure 4.8a,b *Tail* conformations of Arg[8]-vasopressin. Main representative tail states resulting from a DASH state analysis of backbone dihedrals *ΦΨ* 7 to 9. Absolute population are given in parentheses. (**a**) *extended* tail conformations; (**b**) *folded* tail conformations with β-turn centred at residues 7 and 8. Depiction: tail = sticks, ring and side chains = lines. Residues are only labelled for each major populated state.

Figure 4.5 shows the DASH state trajectories for all *overall* states (T16), *ring* states (T10) and *tail* states (T6). There are 176 transitions between the six T6 tail states but only 77 transitions between the twelve T10 ring states, confirming that the tail is significantly more flexible than the ring.

The most striking result, however, is that the tail states do not correlate directly with ring states in terms of transitions or formation of distinct conformational groups. In fact, similarly to the two states of the disulphide bridge, all tail states are distributed evenly over the entire simulation independently of the ring conformation. This is shown convincingly by a principal component analysis of the torsion angles throughout the simulation. As is shown in Figure 4.9a, there are six significant principal components (PCs) according to the eigenvalue-one test, of which PCs 5 and 6 have eigenvalues just barely larger than unity and are therefore at best marginally non-trivial. Figure 4.9b shows the weights (squared coefficients) of the contributions of the individual torsions to PCs 1 and 2 and Figure 4.9c those of PCs 3 and 4.The former are clearly localised on the ring and the latter on the tail. The contributions of ring torsions in PCs 3 and 4 and of tail fluctuations in 1 and 2 are very small. Interactive 3D-plots of the first four PCs are given as HTML-pages in the Online Supporting Material. Interestingly, these show that the two twisted saddle states, although very similar, are clearly separated in PC-space. This separation is mostly due to PC2.

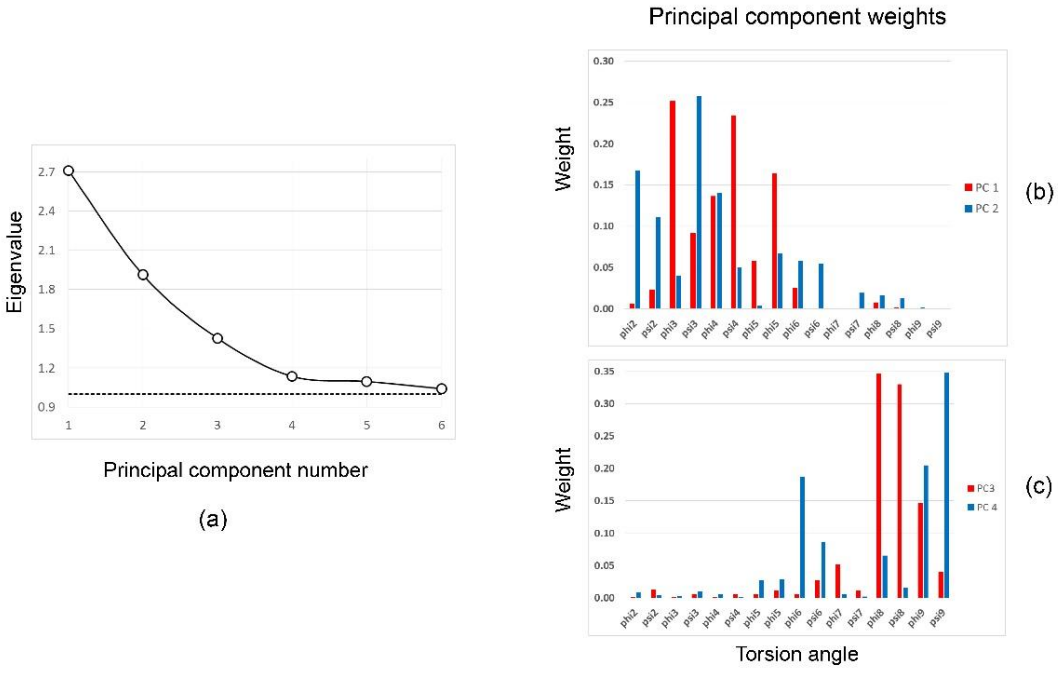


Figure 4.9a-c Summaries of the principal component analysis of the torsion angles during the simulation. (**a**) Eigenvalue plot; (**b**) weights of the torsion angles for principal components 1 and 2; (**c**) weights of the torsion angles for principal components 3 and 4

Thus, every *overall* state can be described in a modular manner as a combination of a *ring* and a *tail* state. The matrix of all main-state combinations found in this simulation is given in Table 4.7. Although the conformational changes of the ring and tail are not correlated, the relative populations of *extended* and *folded* tail states vary depending on the ring conformation (see Table 4.6). The highest preference for a *folded* tail is found for the *clinched open* ring (26 % relative population), the lowest (5 %) for an *open* ring conformation. The *saddle* and *twisted saddle* ring conformations are found together with a *folded* tail 20 and 15 % of their occurrence, respectively. Again, the sterically demanding Arg[8] residue is most likely responsible for these preferences; the propensity to form a folded tail depends on the available space.

Table 4.7 *Overall* state = *ring* state + *tail* state [§]

Overall states (T16) (T6 X T10)				Tail states (T6)			
				extended		*7,8 β-turn II*	
				61.4 %	19.6 %	13.6 %	2.5 %
				3	**4**	**6**	**5**
Ring states (T10)	*open*	12.4 %	**8**	**27**	**28**	**29**	
		0.8 %	**9**	**30**	**28**		
	saddle	36.0 %	**1**	**3**	**4**	**6**	**5**
		4.1 %	**2**	**7**	**8**	**10**	**9**
	clinched open	4.8 %	**4**	**12**	**13**	**14**	
		2.1 %	**5**	**16**	**17**	**18**	
	twisted saddle	21.8 %	**6**	**19**	**20**	**22**	**21**
		13.3 %	**7**	**24**	**25**	**26**	

[§] A matrix of overall states (T16) as combination of the ring and tail states (T10 and T6) that represent the main ring and tail conformations (*open, saddle, clinched open, twisted saddle, extended,* and *folded*). The states result from DASH analyses of torsions *ΦΨ* 2 to 9 (T16, overall states), 2 to 6 (T10, ring states) and 7 to 9 (T6, tail states). The assignment is based on the corresponding DASH state trajectories. PDBs of all states are provided in the Online Supplementary Material.

Conclusions

A single long (11 μs) simulation of AVP in water has revealed details of its conformational behaviour and possible biologically active conformations. Conformational changes on the MD timescale are frustratingly slow, so that, even from the long simulation, we cannot estimate the free-energy

difference between the ring conformations from their concentrations. However, the conformational rearrangements are clearly fast on the NMR timescale, in agreement with the experimental results.

The simulation reveals four distinct ring conformations that are essentially independent of the faster tail motions. The *saddle* and *twisted saddle* ring conformations exhibit β-turns centred at residues 3,4/4,5 as expected from experiments and are fixed by transannular hydrogen bonds. The alternative *open* and *clinched open* conformation do not feature transannular H-bridges. The *saddle* structure identified in the simulation corresponds closely to that found in crystal structure 1JK4.

The simulation is quite consistent with Sikorska and Rodziewicz-Motowidlo's NMR results.[103] They suggest two main conformations, both with 3,4 β II-turns. One is proposed to exhibit a 4,5 β III'-turn and the other a type I'-turn at this position. Our simulations also reveal turns at 3,4 and 4,5 to be dominant in aqueous solution. The 3,4 β-turn type II is found in our *twisted saddle* conformation, but only with a sparsely populated 4,5 turn; a significantly high turn propensity at residues 4 and 5 is found in our *saddle* conformation but here in combination with a β-turn type I at residues 3 and 4. The two studies agree well about the tail conformation, which we found to be approximately 80 %[i] *extended*.

The *open* structure featured in the simulation corresponds closely to the AVP conformation found in the crystal structure of the trypsin complex (PDB ID: 1YF4) and features neither a well characterised β-turn nor conserved transannular hydrogen bonds. The *clinched open* conformation identified in the simulation is apparently new and probably represents an intermediate minority conformation involved in inter-*saddle* rearrangements.[ii]

In general, the simulation is compatible with the known experimental data, which allows us to be confident about its accuracy, even though it is limited to 11 μs and exhibits only a few transitions between major rings states. Above all, the main conformations found can all be considered as candidates for biologically active conformations in different receptors as they are clearly easily accessible thermodynamically. We are now carrying out extensive thermodynamic integration studies to define the thermodynamics of the major conformations in solution.

Technically, DASH has proven to be an extremely useful and effective analysis tool for such simulations. In particular, its beneficial scaling helps to analyse such long simulations. The finding that the movements of the ring and the tail are largely independent facilitates the analysis considerably.

[i] Note: The convergence of the tail conformations was further confirmed by the extended 23μs MD of AVP (see Chapter 5 and Appendix A4 Additional Analysis)

[ii] The *clinched open* state turned out to be major populated proved by the extended 23μs MD and metadynamics simulations (*cf.* Chap. 5)

The conformational distribution demonstrated in this work can now serve as a basis for comparison with those simulated for AVP docked into receptor pockets and for extended simulations of NMR and circular-dichroism spectra. Above all, however, MD simulations have proven once more to be useful, and perhaps the most powerful, tools for analysing the conformational behaviour of peptide hormones of comparable size to AVP.

Acknowledgments. This work was supported by the European project "Peptide Research Network of Excellence" PeReNE as part of the Interreg IVA France (Channel) – England 2007-2014 program (Interreg EU). We thank Jonathan Essex (University of Southampton, UK) and Ronan Bureau (University of Caen, France) for helpful discussions and Harald Lanig (University of Erlangen, Germany) for support with the simulations. Work in Erlangen was supported by the Deutsche Forschungsgemeinschaft as part of Graduiertenkolleg 1910 "*Medicinal Chemistry of Selective GPCR Ligands*".

Chapter 5: Deciphering NMR-Data for Conformational Equilibria: Arginine-Vasopressin (Paper 2)

The results in this section have been published in:

Haensele E, Saleh N, Read C M, Banting L, Whitley D C, Clark T. Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine-Vasopressin. J Chem Inf Model. 2016;56(9):1798-807.[2]

The paper is given as postprint.

Foreword

As has been shown in the previous chapter, long-scale MD simulations are able to identify the main conformational types of a flexible peptide. AVP demonstrates four main ring conformations *saddle*, *twisted saddle*, *clinched open* and *open* that are combined with two tail conformations *extended* and *folded*.[1] However, what are the relative populations of these conformations in solution? Unfortunately, even the extension of the 11 µs unrestrained MD simulation to 23 µs[i] showed no converged sampling of the main ring conformations. Thus, metadynamics enhanced sampling studies were performed to determine the relative populations.
These calculations predict a ratio of approximately 70 % folded (*saddle, twisted saddle*) conformations and 30 % open (*clinched open, open*) conformations for AVP in solution.[2] As this is a purely *in silico* result, validation by comparison with experiment was required. The results are published in Paper 2 (see postprint below).

Deciphering Technique. The paper gives a detailed description of the statistical analysis of the linear regression of calculated and experimental data of AVP and discusses the significance of results in depth. However, the evaluation of the calculated conformational equilibria is not the only result, the protocol or "deciphering technique" itself is even more important a result. Thus, an illustration shall be given in advance. Figure 5.1 shows the logical flowchart of the evaluation technique. The numbering follows the explanation in the paper (see Postprint of Paper 2). The technique performs best for ^{1}H chemical shifts of AVP and holds promise of being extended to

[i] 23 µs AVP MD simulation was equivalent to a CPU time of 17,000 hours (almost 2 years net computation time). For performance of AMBER simulations, see Appendix A7

flexible molecules in general to decipher their experimental NMR data for conformational equilibria in solution.

The workflow to determine and evaluate fast conformational equilibria can be explained as follows: On the experimental side (green), NMR experiments are performed to gain experimental data (*e.g.* chemical shifts, NOE distances, coupling constants). On the *in silico* side (blue), the corresponding data are calculated for representative *single conformations* and for the predicted *mixture of main conformations*. To this, unrestrained long-scale MD simulations (**1**) are combined with enhanced sampling (**2**) and DFT calculations (**3a**). Finally, experimental data and calculated data of *single conformations* (**3a**) and *equilibrium mixtures* (**3b**) are directly compared *via* linear regression (4) and the model with the highest accordance to experiments is chosen as the best assignment.

As anticipated, the best assignment for the AVP conformations to the experimental data was a 70:30 equilibrium of *folded* and *open (unfolded)* conformations as predicted *in silico*.

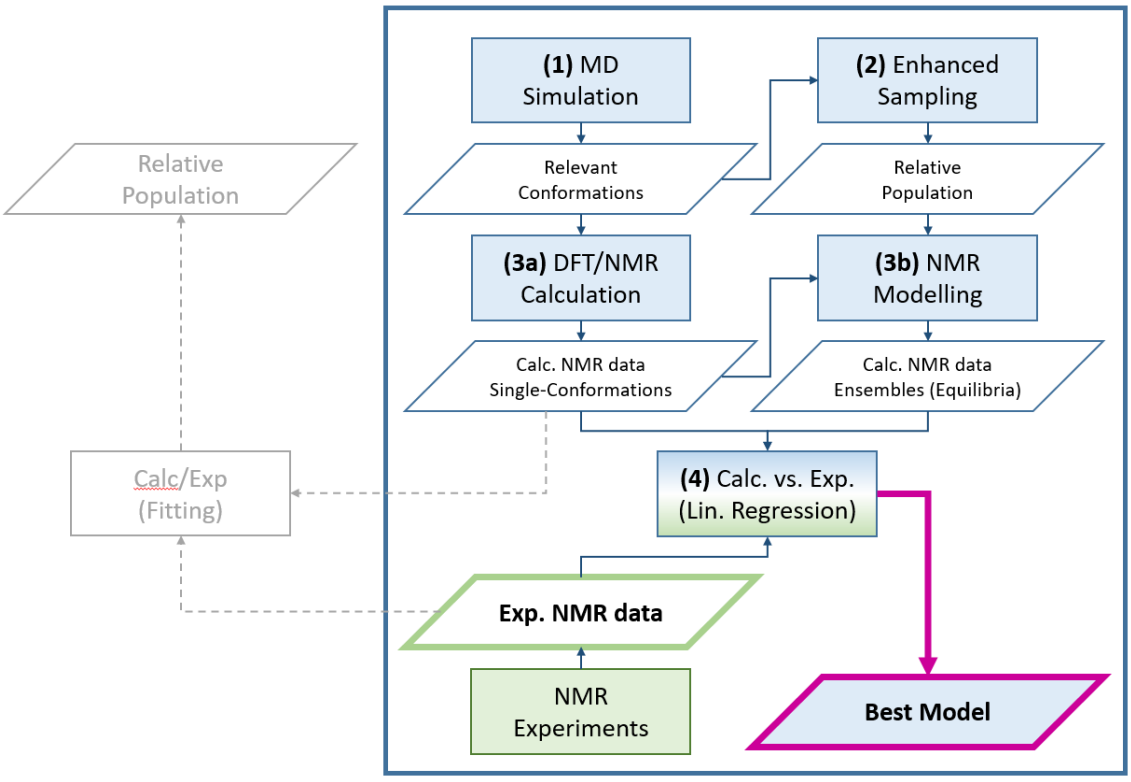


Figure 5.1 Technique to decipher NMR-data of flexible peptides in solution (flowchart). Right side (blue frame): Deciphering technique. Direct comparison of calculated spectra for relevant single-conformations (**3a**) and conformational ensembles (**3b**). Relevant conformations are identified with unrestrained long-scale MD simulations (**1**) and relative populations are calculated from enhanced sampling methods, *e.g.* metadynamics or replica exchange MD simulations (**2**). Left side (grey): Reality check. Relative populations resulting from least square fits of calculated single-conformations and experimental NMR-data.

Contribution of Authors

The results are the product of a joint research project of the University of Portsmouth (UK) and the FAU Erlangen-Nürnberg (D).

Metadynamics simulations of relevant conformations of AVP were performed by Dr. Saleh in Prof. Clark´s group. Representative conformations were chosen by Haensele.

NMR experiments were carried out by Dr. Read. Spectra assignment and analysis was performed by Haensele. Experimental interatomic distances were derived from NOESY spectra (Dr. Read). The corresponding interatomic distances in the relevant conformations were extracted from MD trajectories by Haensele.

DFT optimisations and calculations of NMR shielding tensors were performed by Prof. Clark with coordinate files of representative conformations of AVP from the long-scale MD simulation (Haensele). Further data processing, *e.g.* shielding/shift conversion, was done by Haensele.

Long-scale MD simulations, conformational analyses (DASH), NMR-modelling and statistical evaluation were performed by Haensele.

DP4 probabilities were calculated by Dr. Whitley.

Comparative calculations of PLS regression and bagged MLR *via* SARcaddle were performed by Prof. Clark. He also introduced the two novel error metrics WRMSE and Δ_σ to access the significance of statistical models. The according statistical analyses and application of the new error metrics were calculated by Haensele.

Linked Appendices: A2: Reprint Supporting Information Paper 2, A4: Additional Analysis.

Postprint of Paper 2

Haensele E, Saleh N, Read CM, Banting L, Whitley DC, Clark T. Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine–Vasopressin. J Chem Inf Model. 2016;56(9):1798-807.[i]

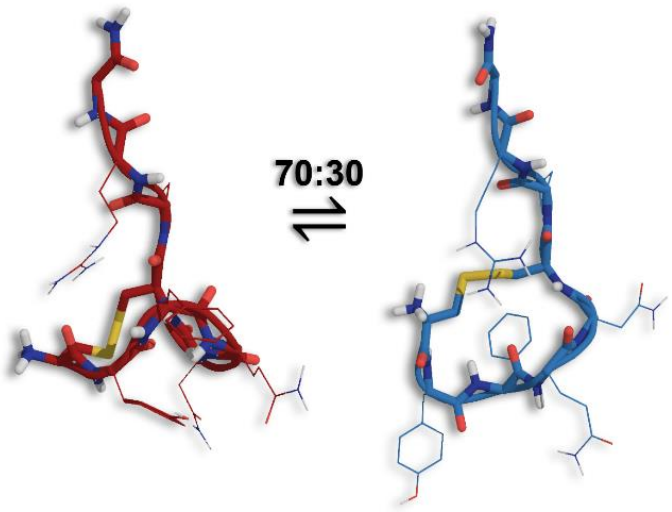


Table of Content Graphic
(Equilibrium of *folded* and *open (unfolded)* conformations of AVP)

Abstract

Arginine vasopressin (AVP) has been suggested by molecular-dynamics (MD) simulations to exist as a mixture of conformations in solution. The ^{1}H and ^{13}C NMR chemical shifts of AVP in solution have been calculated for this conformational ensemble of the ring conformations (identified from a 23 µs molecular-dynamics simulation). The relative free energies of these conformations were calculated using classical metadynamics simulations in explicit water. Chemical shifts for representative conformations were calculated using density-functional theory. Comparison with experiment and analysis of the results suggests that the ^{1}H chemical shifts are most useful for assigning equilibrium concentrations of the conformations in this case. ^{13}C chemical shifts distinguish less clearly between conformations and the distances calculated from the nuclear Overhauser effect do not allow the

[i] Elke Haensele,[a,b] Noureldin Saleh, [a] Christopher M. Read, [c] Lee Banting, [b] David C. Whitley, [b] and Timothy Clark [a,✉]
[a] Computer-Chemie-Centrum der Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany; [b] School of Pharmacy and Biomolecular Sciences, University of Portsmouth, Portsmouth PO1 2DT, United Kingdom; [c] School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DY, United Kingdom

conformations to be assigned clearly. The ^{1}H chemical shifts can be reproduced with a standard error of less than 0.24 ppm (< 2.2 ppm for ^{13}C). The combined experimental and theoretical results suggest that AVP exists in an equilibrium of approximately 70 % *saddle*-like and 30 % *clinched open* conformations. Both newly introduced statistical metrics designed to judge the significance of the results and Smith and Goodman's DP4 probabilities are presented.

Introduction

Many biologically important molecules, especially peptide hormones, exist as an equilibrium mixture of two or more conformations in solution.[278,279] Identifying these conformations and their relative free energies is important because, as long as the conformations in solution are competitive in energy then each is a candidate as the biologically active conformation, which need not be the same in all receptors.

X-ray crystallography usually only provides single snapshots that give little insight into dynamic equilibria, so that NMR spectroscopy becomes the experimental method of choice. Unfortunately, the most common technique used to determine structures in solution, using the nuclear Overhauser effect[280] is often not sufficient to determine even a single structure uniquely, and even less so for conformational equilibria. In this respect, the r^{-6} distance dependence of the NOE (r is the internuclear distance) prevents simple averaging of the structures and renders interpretation more difficult, even when MD simulations are used as the basis for ensemble calculations.[281]

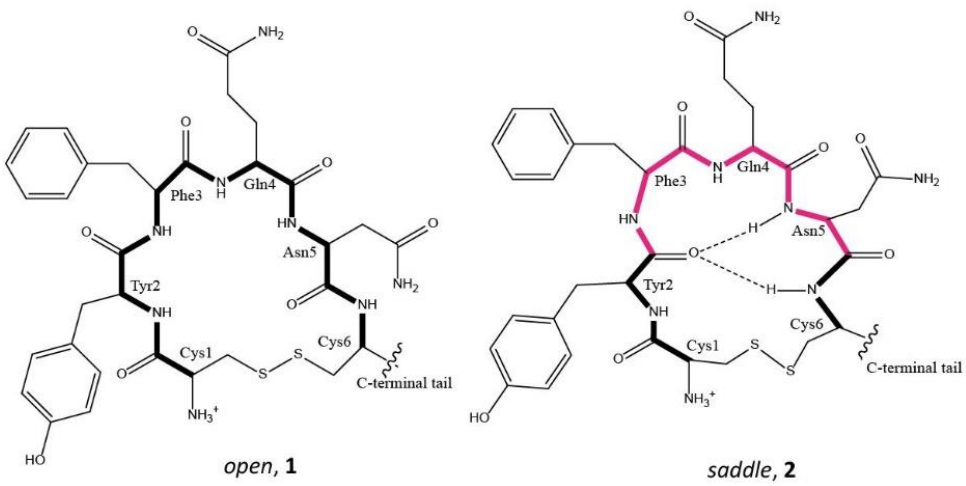
Chemical shifts are not often used to determine conformations in solution because they are not directly related to interatomic distances. A reliable technique for calculating chemical shifts for a given geometry is thus needed and density-functional theory (DFT) calculations now provide such a technique at a reasonable computational cost.[282,283] When a regression equation was used to convert atomic screening to chemical shifts, accuracies of ±0.15 ppm and ±2.2 were obtained for ^{1}H and ^{13}C chemical shifts, respectively.[284] Unfortunately, what might naively be considered the most informative chemical shifts in peptides and proteins, those of the acidic ($pK_a \approx 15$) amide NH-protons often involved in hydrogen bonds, are also strongly affected by exchange phenomena in aqueous solution.[285,286] These effects increase their chemical shifts compared to those calculated for the pure NH protonation state in continuum water. The inclusion of explicit water molecules in the DFT calculations can improve the results,[285,286] but in the case of vasopressin, a nonapeptide, this would lead to extensive sampling problems and make the technique computationally intractable. A further difficulty is that the superficially attractive technique of calculating the chemical shifts of the possible conformations in the equilibrium and fitting a linear combination to

the experimental chemical shifts by regression lacks predictive power because the calculated chemical shifts of the conformations are strongly correlated, so that least-squares fits are seldom unique. This means that, although the fitted results are good, the coefficients of the individual conformations may not necessarily be meaningful because of their strongly correlated chemical shifts. This problem is most visible in bagging regression models, where the coefficients obtained in the different component models vary widely, but is also inherent in partial least squares models, where it is less obvious. These problems have been addressed by Smith and Goodman,[258,287] who used chemical shifts exclusively to distinguish between pairs of diastereomers and proposed improved metrics to overcome the fitting problem. Unfortunately, most of their metrics were designed to assist assignment of spectra to pairs of chiral molecules for which both experimental spectra are available. However, their DP4 probabilities[258] are applicable in the present case, as demonstrated by Nazarski *et al.*,[288] but even using these probabilities as a metric does not solve the problem of linearly dependent descriptors. We have therefore resorted to MD simulations to avoid the fitting problem. We have investigated the use of MD simulations and DFT chemical-shift calculations combined with NMR experiments to assign the conformational equilibrium in solution for 8-arginine-vasopressin (AVP), a flexible peptide hormone.

AVP is the human form of vasopressin, a peptide hormone of the vasopressin family. Vasopressin-related peptides, which include vasopressin, oxytocin, urotensin II and a variety of other non-human tocins, are G-protein coupled receptor ligands that share the common feature of a six-residue ring closed by a disulphide bridge. Although the peptides are very closely related, the conformation of the six-residue ring differs in X-ray crystal structures of AVP (1YF4),[97] 8-lysine-vasopressin (1JK4)[98] and oxytocin (1NPO),[77] suggesting that multiple bioactive conformations may be operative, depending on the binding site.

The ring conformations for these peptide hormones can be classified broadly into *open* and *saddle*[i] types, shown in Scheme 5.1. The *open* ring conformations, **1**, such as that found in PDB-entry 1YF4, do not feature transannular hydrogen bonds and exhibit a flat, *open* ring structure. In contrast, the *saddle* conformations, **2**, (PDB entries 1JK4 and 1NPO) feature a ring that is folded with possible transannular hydrogen bonds, resulting in a *saddle*-like shape that features well-defined β-turns at residues 3 and 4 and/or 4 and 5.

[i] Author´s note: later in this project, the more general notation *folded* was used



Scheme 5.1 The *open* and *saddle* conformational types for AVP. The ring backbone bonds are shown as broad lines and the β-turns in magenta.

NMR studies of AVP[99,103] have concentrated on the *cis*/*trans*-isomerisation across the Cys[6]-Pro[7] peptide bond and have assumed only *folded* (*saddle*) ring conformations. The *trans*-isomer predominates in solution, although the *cis*-isomer can be identified in the NMR spectrum. It will not be discussed here because the *cis*/*trans*-interconversion is slow on the NMR timescale. Recent extensive MD simulations[1] suggest that AVP exists in an equilibrium between several interconverting ring conformations in aqueous solution. The NMR studies summarised in Table 1 of Reference 103 indicate that the ring can adopt diverse structures, all of which, however, have been interpreted as containing well-defined turns, as found in the *saddle* conformation. Exact knowledge of the ring conformational equilibrium is, however, important, as the biologically relevant conformations of AVP have not been identified. We therefore now report a combined theoretical (MD simulations, DFT modelling) and NMR study of the conformational equilibrium of AVP in aqueous solution that compares chemical shifts and interatomic distances calculated without experimental input with data derived from experiments.

Methods

Complete computational and experimental details are given in the Supporting Information (Appendix A2); the procedure will only be described briefly here. Measured chemical-shift and NOE data are compared directly with those predicted essentially without experimental input. These predictions are based on:

1) Identifying the Relevant Conformations of AVP in Solution from Extended Timescale, Unconstrained MD Simulations. Our previous[1] 11 μs MD simulation of AVP in solution was extended to 23 μs to improve sampling. Even this simulation, however, proved insufficient to deduce equilibrium concentrations of individual conformations, as identified using DASH.[141] We therefore, used the conformations identified in the 23 μs simulation to define the path variable for subsequent metadynamics simulations.[151]

2) Calculating the Relative Free Energies of these Conformations in Solution Using Metadynamics. The single path variable used for the metadynamics is simply a numerical assignment to one of the five most prevalent conformations found in the long MD simulation. These conformational assignments are made using the root-mean-square deviation from the individual conformations. This criterion allowed over 90 % of the frames from the 23 μs simulation to be assigned. In order to make the collective variable as "physical" as possible, the numbering of the conformations was chosen so that the 23 μs simulation exhibited transitions between adjacently numbered conformations, thus ensuring that paths between neighbouring conformations exist.

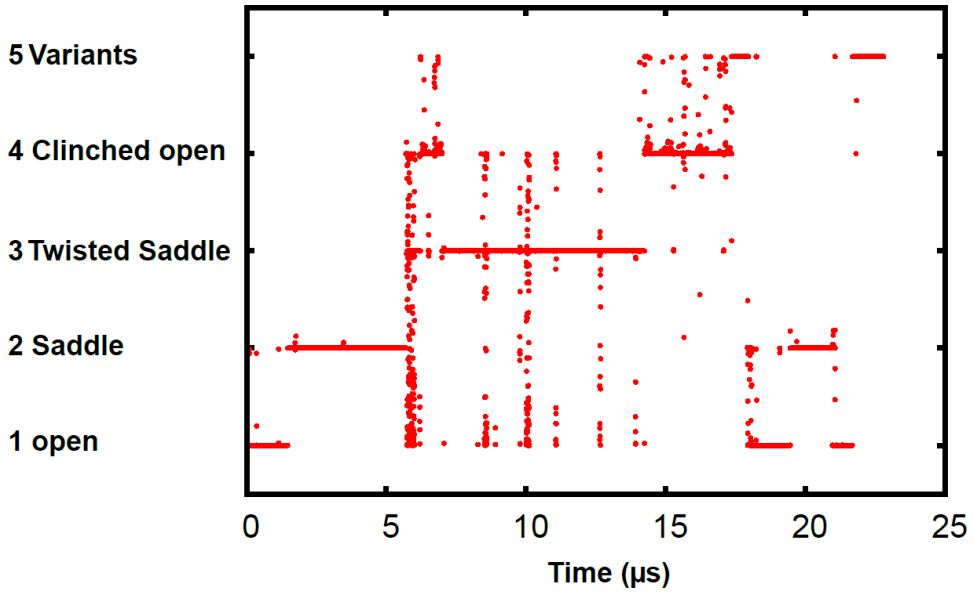


Figure 5.2 The numerical order of the conformations used in the metadynamics collective variable. The conformational assignments are plotted against simulation time for the five most populated DASH states observed in the 23 μs unconstrained simulation. The conformations 1-4 can interconvert as follows: 1-2, 2-3, 3-4. The direct 4-5 interconversion is also seen but conformations 5 were not included (see text).

Figure 5.2 shows the numerical assignment of the conformations. The *variants* cluster of conformations, which proved to be least stable and only occurred after the original 11 μs simulation, was not included in the further analysis (for details, see the Supporting Information, Appendix A2).

3) Calculating Geometries for Cluster Centres and NMR Chemical Shifts with DFT. Cluster centres for the four most populated ring conformations (including two different tail conformations for *saddle* and *clinched open* to give a total of six representative structures) were taken from the 23 μs simulation and optimised with Gaussian09[247] at the B3LYP[237,238]/6-31G(d)[240] level using the standard polarizable continuum model (PCM) for water.[254] The optimised geometries are given in the Supporting Information (Appendix A2). Dispersion corrections were not used, as we do not expect them to be appropriate for PCM calculations in a polar solvent. Note that this neglect of dispersion corrections can only affect the DFT-optimised geometries because the relative DFT energies are not used in the analysis. Relative free energies include dispersion because they were obtained exclusively from force-field based simulations with explicit solvent. Magnetic shieldings were calculated on the optimised structures using the gauge-independent atomic orbital (GIAO) technique[244] at the B3LYP/6-31G(d) level with PCM water. The regression technique for converting calculated isotropic magnetic shielding to chemical shifts in solution[284] was extended to enable B3LYP/6-31G(d) calculations with PCM-water to reproduce ^{1}H and ^{13}C chemical shifts in D_2O relative to (3-trimethylsilyl)propane sulphonic acid (DSS). Details of the training set and the results are given in the Supporting Information (Appendix A2). The regression equations are:

$$\delta(^{1}H) = -0.9912\sigma_H + 32.05$$

$$\delta(^{13}C) = -1.0833\sigma_C + 203.97 \qquad (5.1)$$

where δ is the chemical shift and σ the calculated isotropic atomic magnetic shielding, both in ppm. The root-mean-square deviations from experiment for the training set are 0.18 ppm for ^{1}H and 1.96 ppm for ^{13}C.

Chemical shifts for each optimised cluster-centre conformation were calculated using Eq. (5.1) and ensemble chemical shifts (denoted as *equilibrium* in the following) obtained by linear combination of the individual shifts according to the calculated equilibrium concentrations. $^{1}H^{N}$ chemical shifts were not included, as in practice, these are subject to wide variation by hydrogen bonding, pH and solvent-based environmental changes and are generally not reproduced well by calculations on a single protonation state.

4) Direct Comparison of Experimental and Calculated Spectra or Measurements. The ensemble NMR spectra calculated in step (**3**) can be compared with experimental data. We have assigned the ^{1}H, and ^{13}C chemical-shifts almost fully, in two different aqueous solution conditions at pH 4.7 and pH 6.0. The former pH is that given on dissolving the peptide in H_2O and the latter was chosen to be compatible with the MD simulations. To complete the set of known experimental NMR data we

report for the first time ^{15}N shifts at natural abundance. NOESY and TOCSY NMR spectra gave NOEs and facilitated assignment (see the Supporting Information, Appendix A2, for details).

Both the quality of fit between the calculated and experimental parameters and whether the fit for the calculated equilibrium mixture of conformations is better than that for any of the individual contributing conformations serve to validate the approach. This is often not a straightforward analysis problem,[258,287,288] so that we have defined two statistical metrics below that are designed to test the significance of the differences in correlations of the chemical shifts calculated for individual conformations with the experimental data.

Results and Discussion

Unconstrained Molecular Dynamics

A 23 μs unrestrained MD simulation of Arg^8-vasopressin was performed with explicit water-solvation at 300 K using the AMBER ff99SB force field[133] (details are given in the Supporting Information, Appendix A2). The conformational space was clustered using DASH[141] and compared with the conformations (clusters) found in the first 11 μs of the simulation.[1] These main clusters, *open, saddle, clinched open,* and *twisted saddle*, dominated the simulation (Fig. 5.3).

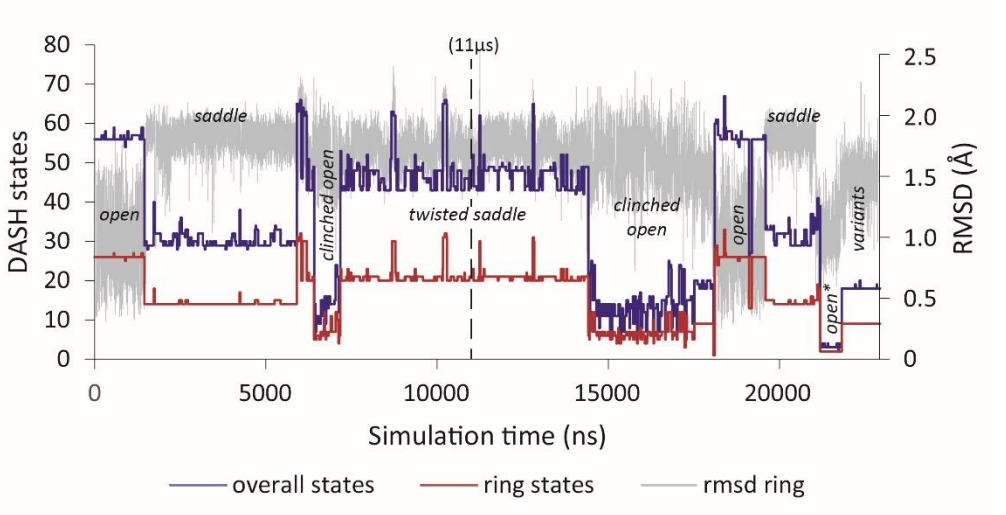


Figure 5.3 RMSD of $C\alpha_{1\text{-}6}$ (grey), and the corresponding clusters of *ring* and *overall* conformations of 23 μs unrestrained MD simulation of Arg^8-vasopressin. The main clusters (ring conformations) are labelled.

They have been described in detail.[1] Even after 23 μs, the simulation exhibited too few interconversions between the main clusters to estimate reliable equilibrium populations directly. Thus, we chose the representatives (cluster centres) of the four main clusters to calculate their free

energies and relative populations with metadynamics. A fifth cluster, *variants*[i], which occurred for the first and only time at the end of the 23 μs simulation, was also added to the selection. A description of this cluster of conformations is given in the Supporting Information (Appendix A2). The conformational clusters *open, saddle, clinched open* and *twisted saddle* represent 86.4 % of all conformations found for AVP in the simulation, and *variants* 7.4 % to give a total of 93.8 % that can be assigned to the five clusters. The rest are transient states not discussed here further. We showed[1] previously that the tail moves independently of the ring conformation of AVP, adopting either *folded* or *extended* conformations, which interconvert frequently and rapidly. Thus, it is possible to take the individual populations of these tail conformations directly from the 23 μs MD simulation.

Metadynamics

A well-tempered metadynamics simulation[289] using four walkers converged within 200 ns to give the relative free energies of the five conformations shown in Table 5.1.

Table 5.1 Equilibrium populations and relative free energies (ΔΔG) from the metadynamics simulation[§]

	saddle	*clinched open*	*twisted saddle*	*open*	*variants*
ΔΔG (kcal mol^{-1})	0.0	0.5	3.0	2.0	3.5
% at equilibrium (5 conformations)	68.5	29.5	0.4	1.4	0.2
% at equilibrium (4 conformations)	68.7±3.9	29.5±4.0	0.4±0.1	1.4±0.5	-

[§]The ΔΔG values are converged to approximately ± 0.2 kcal mol^{-1}. The equilibrium concentrations are given at 298K. Errors are based on ± 0.2 kcal mol^{-1} energetic uncertainty and are given as ± one standard deviation.

These results can be compared with those obtained by least squares fitting the calculated NMR chemical shifts to observations, although the latter, as outlined above, may not be significant. The comparison therefore serves at best as a rough test as to whether the equilibrium concentrations obtained from the simulations are similar to those that would give the best fit. Figure 5.4 shows the equilibrium concentrations calculated from free-energy differences obtained in the metadynamics simulations and those obtained by fitting two different regression models to the experimental chemical shifts.

[i] Author´s note: subsequently denoted as *clinched open*$_{45pbr}$

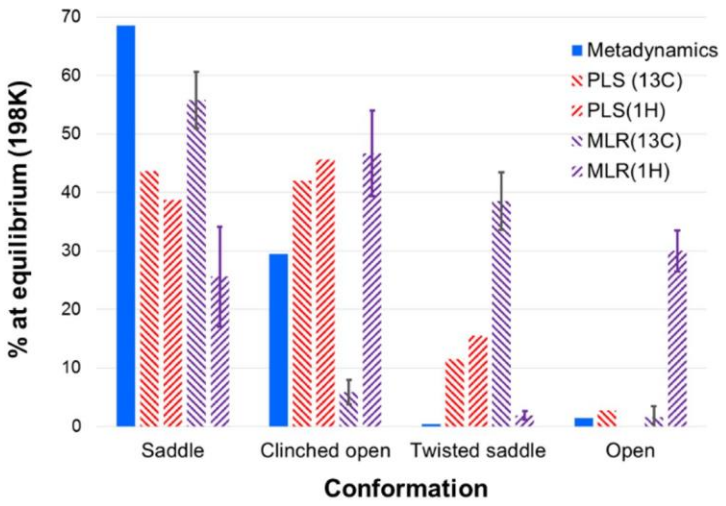


Figure 5.4 Calculated equilibrium concentrations (%, 298K) for the *saddle*, *clinched open*, *open* and *twisted saddle* conformations. The fitted values are taken from partial least squares (PLS) and bagged multiple linear regression (MLR) fits. The variants conformations are not found to be significant. Bagged MLR and PLS calculations were performed with SAR-caddle.[234] The error bars given for the bagged MLR results are the standard deviations of five fitting runs.

As the calculated chemical shifts for the individual conformations correlate strongly, fitting does not yield a robust statistical model, as demonstrated by the scatter in the fitted results. Whereas the regression models differ as to whether the *saddle* or *clinched open* conformation is the most prevalent in the solution equilibrium, the metadynamics results indicate that the population of the *saddle* conformation is highest. The fitted equilibrium concentration can serve, however, as a reality check for the metadynamics results. The metadynamics equilibrium is quite compatible with the optimum PLS-fits for this dataset (Fig. 5.4), which is encouraging, and we emphasise once more that, in contrast to the regression data, those calculated for the metadynamics equilibrium use essentially no experimental data. The exception is the standard set of chemical shifts used to obtain Eq. (5.1) to convert shielding to ppm. However, the training dataset (given in the Supporting Information, Appendix A2) only contains small organic molecules, which can be considered independent of AVP. The conformations were identified from the 23 μs MD-simulation, the chemical shifts calculated for B3LYP/6–31G(d)-optimised geometries and the equilibrium calculated from the free energies obtained from the metadynamics simulations.

Figure 5.5 shows the B3LYP/6-31G(d) (in PCM water) optimised structure of the major *saddle* conformation. The C-terminal tail adopts two conformations.[1,99] The *extended* conformation, which positions the guanidinium moiety of Arg[8] close to the ring was present in the 23 μs MD simulation for approximately 73 % of the occurrence time for the *saddle* conformation (Fig. 5.5a). The *folded* tail conformation (Fig. 5.5b) makes up the remaining 27 %. In this case, error estimates are difficult because probable errors depend on how well the simulation has converged, which is unknown. We

estimate from the length and convergence of the simulation that the above concentrations have uncertainties of at most ±5 %. The equilibrium between these two tail conformations is fast on the simulation timescale, so that we can refine the calculation of the NMR chemical shifts by treating the *saddle* conformation as a 73:27 mixture of the two conformations shown in Figure 5.5. The *clinched open* conformation is treated similarly (63 % *extended*: 37 % *folded*, see the Supporting Information, Appendix A2). This results in some improvements in the agreement between calculations and experiment, as shown in Table 5.2 below.

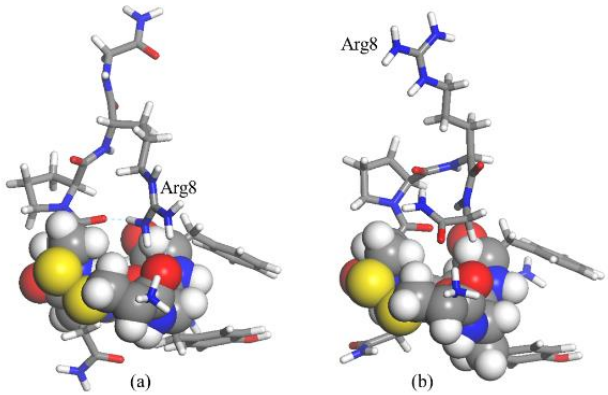


Figure 5.5a,b Optimised structures of the *saddle* conformation obtained at the B3LYP/6-31G(d) level in PCM water solvent. The ring atoms as spheres: (**a**) the *extended* tail conformation, (**b**) the *folded* equivalent.

Figure 5.6 shows plots of the results of the final computational model compared with experiment for both ^{13}C and ^{1}H chemical shifts.

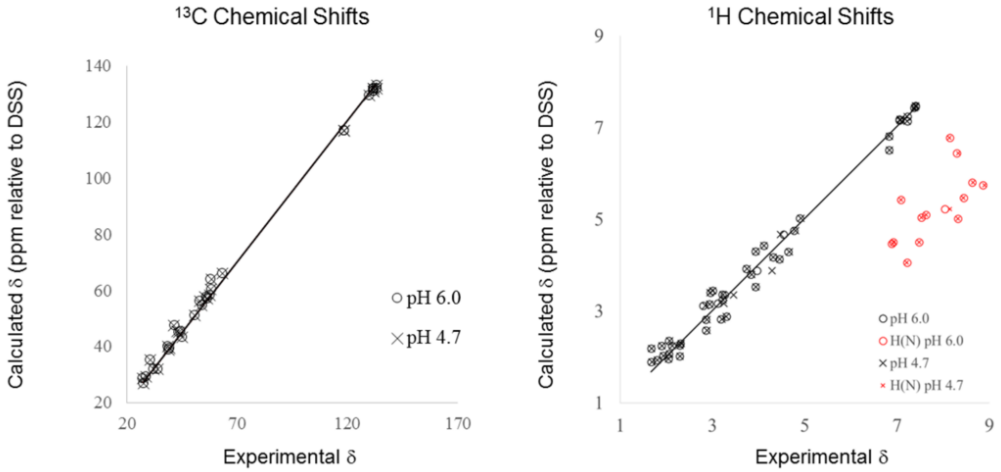


Figure 5.6 Plots of the calculated *vs*. experimental ^{13}C and ^{1}H chemical shifts using the equilibrium model for both ring and tail conformations. δ (ppm) are relative to DSS (3-(trimethylsilyl)propane sulphonic acid). The NH-protons are outliers due to hydrogen bonding and exchange effects.[285,286]

Is the Statistical Analysis for the Calculated Equilibrium Significant?

This question has been addressed several times in the literature.[288,290,291] These studies have been summarised by Smith and Goodman,[258,287] who also proposed improved metrics for judging the goodness of fit between calculated and experimental chemical shifts. As outlined above, many of their metrics were designed to assist assignment of spectra to pairs of diastereomers for which both experimental spectra are available and are inapplicable in this case. We have resorted to conventional metrics such as mean signed (MSE) and unsigned error (MUE), coefficient of determination (R^2) and root-mean-square error (RMSE) as a specific test of the significance of the conclusions, but have also defined a weighted RMSE (WRMSE) in the spirit of Smith and Goodman and have used their DP4 probability[258] as an additional check.

The WRMSE is defined as;

$$WRMSE = \frac{\sqrt{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \sigma_i}}{\sqrt{\sum_{i=1}^{N} \sigma_i}} \tag{5.2}$$

where $\hat{y}_i$ and y_i are the predicted and observed chemical shifts for atom i, respectively, and σ_i is the standard deviation of the calculated chemical shifts for atom i over all conformations.

WRMSE is equivalent to RMSE if all σ_i are equal and otherwise weights the contributions of the atoms that display a wide range of chemical shifts between the conformations more heavily than those with little variation.

A second specific test of the significance of the conclusions is the mean absolute error expressed in units of the standard deviation over all conformations, Δ_σ:

$$\Delta_\sigma = \frac{\sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{\sigma_i}}{N} \tag{5.3}$$

Δ_σ expresses the significance of the MUE in terms of the total spread of calculated chemical shifts for the individual conformations. Ideally $\Delta_\sigma \leq 1$ indicates that on average the deviation between experimental and calculated results is below the standard deviation between the different conformations; the model can discriminate between conformations. We arbitrarily assign a limit of $\Delta_\sigma \leq 1.5$ to indicate reliable discrimination between conformations. The results are shown in Table 5.2.

Table 5.2 Statistics of the comparison of ^{13}C and ^{1}H chemical shifts for AVP at pH 6.0 and 4.7 in aqueous solution[§]

Conformation Ring	Tail	MSE	MUE	RMSE	WRMSE	Δ_σ	R^2
		^{13}C, pH 6.0					
saddle	*extended*	0.87	1.69	2.33	2.74	1.40	0.9965
	folded	0.52	1.75	2.52	3.18	1.26	0.9958
	equilibrium	0.78	1.62	2.23	2.68	1.32	0.9968
clinched open	*extended*	0.74	2.27	3.15	3.75	1.71	0.9936
	folded	0.78	2.18	2.94	3.48	1.71	0.9943
	equilibrium	0.76	2.16	2.98	3.56	1.65	0.9942
twisted saddle	*extended*	0.73	1.70	2.23	2.66	1.42	0.9969
open	*extended*	1.18	2.49	3.72	5.24	1.93	0.9807
Equilibrium	*extended*	0.84	1.55	2.19	2.50	1.34	0.9969
Equilibrium	*equilibrium*	0.78	1.46	2.12	2.45	1.26	0.9971
		^{13}C, pH 4.7					
saddle	*extended*	0.95	1.73	2.37	2.74	1.47	0.9964
	folded	0.60	1.76	2.55	3.19	1.31	0.9957
	equilibrium	0.85	1.66	2.26	2.68	1.39	0.9967
clinched open	*extended*	0.82	2.36	3.28	3.93	1.79	0.993
	folded	0.85	2.28	3.09	3.67	1.80	0.9938
	equilibrium	0.83	2.26	3.13	3.74	1.74	0.9937
twisted saddle	*extended*	0.80	1.78	2.32	2.75	1.50	0.9966
open	*extended*	1.25	2.56	3.77	5.29	2.00	0.9904
Equilibrium	*extended*	0.91	1.63	2.28	2.58	1.43	0.9967
Equilibrium	*equilibrium*	0.85	1.54	2.20	2.54	1.35	0.9969
		^{1}H, pH 6.0					
saddle	*extended*	0.06	0.22	0.31	0.37	1.02	0.9706
	folded	0.02	0.31	0.38	0.42	1.43	0.9571
	equilibrium	0.05	0.23	0.29	0.33	1.03	0.9748
clinched open	*extended*	0.05	0.22	0.28	0.30	1.13	0.9773
	folded	-0.02	0.33	0.43	0.51	1.57	0.9441
	equilibrium	0.02	0.20	0.25	0.26	1.11	0.9800
twisted saddle	*extended*	-0.03	0.42	0.58	0.79	1.62	0.9486
open	*extended*	0.01	0.30	0.48	0.68	1.09	0.9674
Equilibrium	*extended*	0.05	0.20	0.26	0.30	0.93	0.9793
Equilibrium	*equilibrium*	0.04	0.18	0.23	0.25	0.93	0.9832
		^{1}H, pH 4.7					
saddle	*extended*	0.04	0.23	0.32	0.37	1.03	0.9692
	folded	0.00	0.31	0.38	0.43	1.43	0.9562
	equilibrium	0.02	0.23	0.29	0.34	1.04	0.9735
clinched open	*extended*	0.02	0.22	0.29	0.31	1.12	0.9753
	folded	-0.04	0.33	0.43	0.51	1.54	0.9446
	equilibrium	0.00	0.21	0.26	0.27	1.12	0.9789
twisted saddle	*extended*	-0.03	0.42	0.58	0.79	1.62	0.9496
open	*extended*	0.01	0.30	0.48	0.68	1.09	0.9672
Equilibrium	*extended*	0.03	0.21	0.27	0.31	0.95	0.9777
Equilibrium	*equilibrium*	0.02	0.19	0.24	0.26	0.94	0.9820

[§] The best performing model is indicated in bold for each parameter. The amide protons are omitted, as outlined in the text.

Surprisingly, the ^{1}H chemical shifts give the clearest and most consistent picture; they indicate that the experimental ^{1}H shifts are best in agreement with the equilibrium model that uses metadynamics free-energy differences for the ring conformations and equilibrium concentrations from the unconstrained simulation for the faster tail equilibrium. This model is quite consistently the best for ^{13}C; only Δ_σ at pH 4.7 indicates the *saddle* conformation with a *folded* tail to fit better than the calculated equilibrium. However, WRMSE is always larger than RMSE and Δ_σ

approximately 1.3, so that we must conclude that the ^{13}C chemical shifts are not sensitive enough to conformation to allow us to assign values to the conformational equilibrium unequivocally.
The situation for the ^{1}H chemical shifts is clearer; with the exception of the MSE, all metrics indicate that the model that uses the metadynamics free energies for the ring conformations and the distributions of the tail conformations from the 23 μs unconstrained simulation matches the experimental data best. Most importantly, in contrast to the ^{13}C results, WRMSE is close to RMSE for the equilibrium models and Δ_σ is less than one.

The DP4 probabilities lead to exactly the same conclusions as the metrics reported in Table 5.2. The conformational model that considers the equilibrium distributions of both the ring and the tail fits the experimental data best and ^{1}H chemical shifts allow firmer conclusions than ^{13}C. However, the DP4 probabilities also allow tentative conclusions to be reached from the ^{13}C chemical shifts; the equilibrium conformational model gives a 60-75 % probability of being correct, although this probability is close to 100 % for ^{1}H. Table 5.3 shows that Smith and Goodman's DP4 probabilities[258] provide very strong support for these conclusions.

Table 5.3 DP4 probabilities for the AVP conformations at pH 6.0 and 4.7 in aqueous solution [§]

Conformation		pH 6			pH 4.7		
Ring	Tail	^{13}C	^{1}H	^{13}C and ^{1}H	^{13}C	^{1}H	^{13}C and ^{1}H
saddle	*extended*	1.7	0.0	0.0	4.7	0.0	0.0
	folded	0.0	0.0	0.0	0.1	0.0	0.0
clinched open	*extended*	0.0	0.0	0.0	0	0.0	0.0
	folded	0.0	0.0	0.0	0	0.0	0.0
twisted saddle	*extended*	0.3	0.0	0.0	0.2	0.0	0.0
open	*extended*	0.0	0.0	0.0	0.0	0.0	0.0
saddle	*equilibrium*	4.2	0.0	0.0	12.7	0.0	0.0
clinched open	*equilibrium*	0.0	0.2	0.0	0.0	0.5	0.0
Equilibrium	*extended*	19.8	0.2	0.0	19.1	0.2	0.1
	equilibrium	**74.0**	**99.6**	**100.0**	**63.2**	**99.3**	**99.9**

[§] The best performing model is indicated in bold. The probabilities were calculated using the data from the Supporting Information (Appendix A2) with the DP4 app.[258] The amide protons are omitted, as outlined in the text.

As outlined above, the differences in the statistical metrics would not be as convincing if they were based on a fitting procedure. However, as the identification of possible conformations, the calculation of equilibrium concentrations and the chemical-shift calculations are all *ab initio*, in the sense that they are completely independent of experimental data (with the exception of the regression equations (5.1)), we consider the quality of the agreement between experimental and calculated chemical shifts to be significant. RMSEs lower than 0.24 ppm for ^{1}H (without ^{1}NH) and 2.2 for ^{13}C are as good as, or better than, those reported previously using a variety of techniques,[280,281,283,284] and these values are only slightly larger than the standard errors obtained for the training set of small molecules (0.18 and 1.96 ppm for ^{1}H and ^{13}C, respectively). In order to

strengthen these conclusions, we have carried out a sensitivity analysis to see how sensitive WRMSE and Δ_σ are to the equilibrium concentrations. For this analysis, we used both a binary mixture of the majority *saddle* and *clinched open* conformations (WRMSE andΔ_σ) and the full equilibrium with four components (WRMSE' and Δ_σ').

Sensitivity Analysis

Figure 5.7 shows the dependence of WRMSE and Δ_σ on the percentage of the *saddle* conformation in the binary mixture. Both react quite sensitively to the concentrations at equilibrium and exhibit clear minima. For ^{13}C, the two curves correspond closely with a common minimum at the metadynamics values of 70 % *saddle* and 30 % *clinched open*. The two metrics agree less well for the ^{1}H data; WRMSE gives a minimum at approximately 35 % *saddle* and Δ_σ at approximately 60 %. As three of four metrics give minima close to the metadynamics prediction, we feel that Figure 5.7 supports our conclusions.

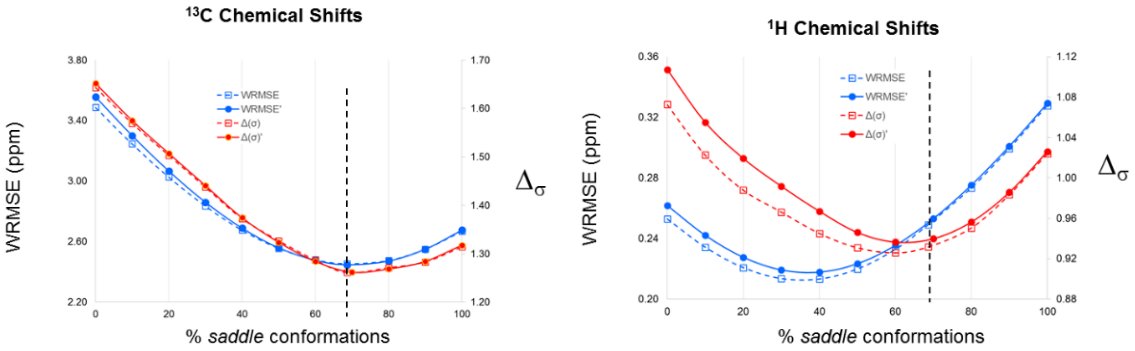


Figure 5.7 The dependence of WRMSE and Δ_σ on the concentrations in mixtures of *saddle* and *clinched open* conformations at pH 6.0. The vertical dashed lines indicate the metadynamics equilibrium. WRMSE and Δ_σ refer to the binary mixture and WRMSE' and Δ_σ' to the four-component equilibrium. The corresponding plots for pH 4.7 are very similar.

Nuclear Overhauser Effect

An independent check of the conformational assignment compares the interatomic distances provided by nuclear Overhauser effect (NOE) data with those obtained from the simulations (details of the calculations are given in the Supporting Information, Appendix A2). The correlation obtained for the observed NOEs and the statistics of the agreement between experiment and simulation are shown in Table 5.4.

Table 5.4 Statistics of the comparison of calculated and observed NOE distances for AVP [§]

Conformation		pH 4.7				pH 6.0			
Ring	Tail	MSE	MUE	RMSE	R^2	MSE	MUE	RMSE	R^2
saddle	*extended*	0.33	0.56	0.74	0.549	-0.12	0.36	0.45	**0.553**
	folded	**-0.33**	**0.52**	**0.68**	**0.622**	-0.16	0.37	0.48	0.084
	equilibrium	**0.33**	0.55	0.71	0.575	-0.11	0.32	**0.39**	0.176
clinched open	*extended*	0.41	0.56	0.81	0.370	**-0.06**	**0.31**	**0.39**	**0.553**
	folded	-0.38	0.56	0.80	0.417	-0.07	0.32	0.41	0.514
	equilibrium	0.40	0.56	0.80	0.395	**-0.06**	0.31	**0.39**	0.543
twisted saddle	*extended*	0.37	0.55	0.75	0.527	-0.13	0.35	0.45	0.340
open	*extended*	0.32	0.56	0.72	0.551	-0.11	0.35	0.44	0.337
Equilibrium	*extended*	0.36	0.54	0.72	0.533	-0.11	0.32	**0.39**	0.366
Equilibrium	*equilibrium*	0.36	0.53	0.71	0.557	-0.12	0.33	0.40	0.344

[§] pH 6.0 and 4.7 in aqueous solution The best performing model is indicated in bold for each parameter. Details of the derivation of both experimental and simulated distances are given in the Supporting Information (Appendix A2).

At pH 4.7, highest R^2 (0.622) is found for the saddle conformation with folded tail but this model is not favoured clearly by any other metric. The metadynamics equilibrium considering the tail conformation is always close to the best values found but the differences are not significant. All conformations perform similarly (there are, for instance, five conformations with an RMSE of 0.39 Å). The *saddle* and *clinched open* conformations with the extended tail conformation give the best coefficients of determination (0.553) but the data are in general inconclusive. The small number of NOE distances available at pH 6.0 also does not allow a definitive conformational determination but tend towards *clinched open* with the extended tail conformation. Thus, the NOE simulations are compatible with the chemical-shift results but not definitive. These results illustrate the difficulties pointed out by Zagrovic and van Gunsteren[281] that NOE studies can, in fact, often be ambiguous; especially for highly flexible structures where intramolecular hydrogen bond distances may "average" by fast equilibria of different conformations.

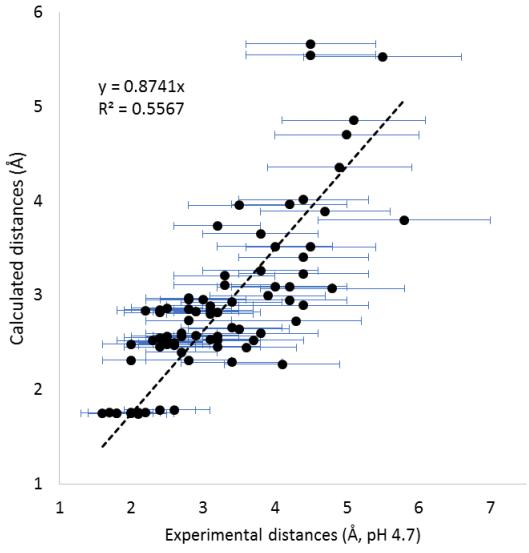


Figure 5.8 Plots of the calculated *vs.* experimental interatomic distances at pH 4.7 (*cf.* Appendix A2 Fig. S3)

Figure 5.8 shows the correlation between experimental at pH 4.7 and r^{-6} time-averaged interatomic distances for the metadynamics equilibrium.

Conclusions

We have reported an attempt to assign conformations for the equilibrium structures of AVP in aqueous solution by simulating the equilibrium and comparing calculated chemical shifts directly with experiment. This procedure avoids fitting and uses only minimal unconnected experimental data to parameterise the regression equation for the calculated chemical shifts. Our models reproduce the experimental data very well (RMSE < 0.24 ppm for ^{1}H and < 2.2 ppm for ^{13}C) but the question remains as to whether the agreement is significant enough to allow conclusions about the equilibrium mixture of conformations.

The proton NMR results present the strongest argument, even though amide protons cannot be included because they are shifted from the calculations for the pure NH-protonation state by exchange. The ^{13}C data are reproduced well, but the diagnostic metrics are not as clear, indicating that the ^{13}C chemical shifts are less sensitive to conformation than ^{1}H and therefore less suitable for our purpose.

The calculated equilibrium concentrations are, however, comparable to those found for an optimal fit, so that we can be confident that they are close to reality, although the regression models suffer from strongly correlated descriptors.

We conclude that the conformational equilibrium for AVP in aqueous solution consists of approximately 70 % *saddle*, 30 % *clinched open* conformations and that the free-energy penalty for *clinched open* as a biologically active conformation is approximately 0.5 kcal mol^{-1}.

It is conceivable that the folded, *saddle*-like type of conformations comprises a higher amount of *twisted saddle* than predicted by metadynamics. In fact, the representative conformations of *saddle* and *twisted saddle* are closely related; they only differ in the turn type of the β-turn at residues 3 and 4. This is also reflected in a very high correlation of the ^{13}C chemical shifts for *saddle* and *twisted saddle* (R^2 = 0.997) in contrast to ^{1}H (R^2 = 0.949). A similar sensitivity analysis to that shown in Figure 5.7 indicates that the ^{13}C data are compatible with mixtures from 10 % to 50 % *twisted saddle* in the *saddle*-like component of the equilibrium and the ^{1}H data with approximately 70:30 *saddle* : *twisted saddle*. We are currently unable to resolve this discrepancy between long unbiased simulation and metadynamics. In any case, all data are consistent with the conservative conclusion that the equilibrium consists of 70 % *saddle*-type and 30 % *open*-type conformations (Scheme 5.1).

One important result of this work is to show that modern MD-simulations and DFT calculations provide data that can be compared directly with experiment without fitting. In this respect, as suggested by Smith and Goodman,[258,287] chemical shifts prove to be more useful than NOEs and, surprisingly, in this example ^{1}H chemical shifts present a clearer picture than ^{13}C, as also found by Nazarski *et al*.[288] Smith and Goodman's DP4 probabilities[258] suggest very clearly that, of those considered, our equilibrium model agrees best with experiment.

The methodology used does not require the unconstrained MD simulation to be long enough to be able to determine equilibrium concentrations. Its function is to identify the conformations (and the transitions between them) for subsequent determination of the free energy differences, here metadynamics simulations. For this reason, and for economy of computer time, we have used the cluster centres for each conformation, rather than calculating shifts for a large number of snapshots in an ensemble model.

Acknowledgments. This work was supported by the Deutsche Forschungsgemeinschaft within GRK 1910 Medicinal Chemistry of Selective GPCR ligands and by a large grant of computer time on SuperMUC at the Leibniz Rechenzentrum München (project pr94to).

Associated Content. The Supporting Information is available free of charge on the ACS Publication website at DOI: 10.1021/acs.jcim.6b00344. Computational and experimental (NMR) details and Gaussian Archive Entries for the B3LYP/6-31G(d)-optimised geometries (PDF).

Chapter 6: Conformation and Dynamics of Human Urotensin II and Urotensin-Related Peptide in Aqueous Solution (Paper 3)

The results in this section have been published in:

Haensele E, Mele N, Miljak M, Read C M, Whitley D C, Banting L, Delépée C, Sopkova-de Oliveira Santos J, Lepailleur A, Bureau R, Essex J W, Clark T. Conformation and Dynamics of Human Urotensin II and Urotensin-Related Peptide in Aqueous Solution. J Chem Inf Model. 2017. doi: 10.1021/acs.jcim.6b00706.[3]

The paper is given as postprint.

Foreword

The results in Chapter 4 (Paper 1) proved that long-scale MD simulations, in combination with DASH, are suited to identify and characterise the main conformational types of cyclic peptide hormones with the example of AVP. In Chapter 5 (Paper 2), a technique was introduced to determine *in silico* conformational equilibria including the validation *via* comparison of DFT calculated chemical shifts with experimental shifts. The significance of the results was tested intensively and discussed for AVP. Consequently, the protocol was applied to UII and URP, for which only few and seemingly contradictory conformational data are available (*cf.* Table 2.6). The results are given in this chapter (Paper 3) and may be outlined as follows:

Firstly, the conformational space and dynamics of UII and URP was explored extensively using unrestrained long-scale molecular-dynamics simulations, different force fields and ion concentrations, enhanced sampling with replica exchange MD simulations, DASH clustering and principal component analysis (PCA). This resulted in the classification of the main conformations for UII and URP, a tentative explanation of distinct behaviour of UII and URP ascribed to possible ring/tail interaction in UII and the *in silico* prediction of their conformational equilibria in solution.

Parallel, NMR experiments were performed for UII and URP in aqueous solution at different pH to gain comparison values for the evaluation of the *in silico* equilibria. Hitherto unpublished ^{15}N chemical shifts of UII and URP were determined and a complete assignment of the *cis*-Pro^3 isomer of UII was possible.

Finally, the equilibrium populations of *open* and *folded* conformations of UII and URP, as predicted by REMD, were confirmed by statistical evaluation following the technique described in Chapter 5 (Paper 2).

The conformational equilibria were identified as 72 % *folded* and 28 % *open* conformations for UII and 86 % *folded :* 14 % *open* for URP, respectively.

Contribution of Authors

The results are the product of a joint research project of the Universities of Portsmouth (UK), Southampton (UK), Caen (F), and the FAU Erlangen-Nürnberg (D) supported by the European "Peptide Research Network of Excellence" (PeReNE).

REMD simulations were carried out by Miljak and Mele of Prof. Essex´s group.

NMR experiments were carried out by Dr. Read. Spectra assignments and analysis were performed by Haensele.

Two 1 µs CHARMM MD simulations were contributed by Delépée of Prof. Bureau´s group. Torsion trajectories were re-analysed by Haensele to ensure the consistency of analysis methods. New cluster centres were used as initial conformations for further long-scale MDs by Haensele.

The program DASH was extended with a routine to perform principal component analyses by DW. A modified version was developed by DW to analyse REMD trajectories. He also wrote the program *dashsim* to compare conformations based on circular similarities of torsions (*cf.* Appendix A7). Circular similarities were calculated by Haensele to ensure compatibility of conformational assignments from the differing sources.

DFT optimisations and calculations of NMR shielding tensors were performed by Prof. Clark. Further data processing, *e.g.* shielding/shift conversion, NMR modelling and statistical evaluation was done by Haensele.

The project was managed by Haensele (comprehensive data-processing and junction of results) supported and supervised by Dr. Banting and Prof. Clark.

Linked Appendices: A3: Supporting Information Paper 3; A7: Hardware and Software.

Postprint of Paper 3

Haensele E, Mele N, Miljak M, Read C M, Whitley D C, Banting L, Delépée C, Sopkova-de Oliveira Santos J, Lepailleur A, Bureau R, Essex J W, Clark T. Conformation and Dynamics of Human Urotensin II and Urotensin-Related Peptide in Aqueous Solution. J Chem Inf Model. 2017:298-310.[i]

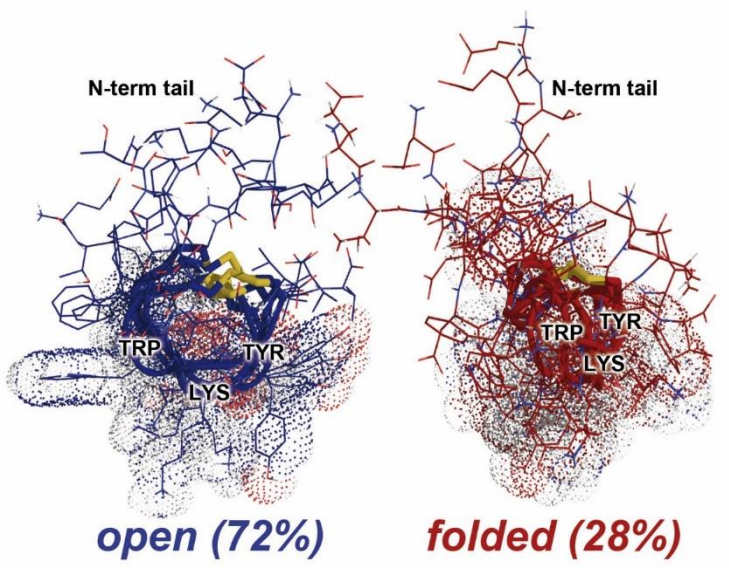


Table of Content Graphic
(Equilibrium of *open (unfolded)* and *folded* conformations of UII)

Abstract

Conformation and dynamics of the vasoconstrictive peptides human urotensin II and urotensin-related peptide have been investigated by both unrestrained and enhanced-sampling molecular-dynamics simulations and NMR spectroscopy. These peptides are natural ligands of the G-protein coupled urotensin II receptor (UTR) and have been linked to mammalian pathophysiology. UII and URP cannot be characterised by a single structure but exist as an equilibrium of two main classes of ring conformations, *open* and *folded*, with rapidly interchanging subtypes . The *open* states are characterised by turns of various types centred at K^8Y^9 or F^6W^7 predominantly with no or only

[i] Elke Haensele,[a] Nawel Mele,[b] Marija Miljak,[b] Christopher M. Read,[c] David C. Whitley,[a] Lee Banting,[a] Carla Delépée,[d] Jana Sopkova-de Oliveira Santos,[d] Alban Lepailleur,[d] Ronan Bureau,[d] Jonathan W. Essex,[b] and Timothy Clark[e,✉]
[a] School of Pharmacy and Biomedical *Sciences* and [c] School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DT, United Kingdom; [b] School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom; [d] Normandie Université, CS 14032 Caen Cedex 5, France, Centre d'Etudes et de Recherche sur le Médicament de Normandie (CERMN, EA 4258, FR CNRS 3038 INC3M SF 4206 ICORE), UFR des Sciences Pharmaceutiques, Université de Caen Basse−Normandie (UNICAEN), F-14032 Caen, France; [e] Computer-Chemie-Centrum and Interdisciplinary Center for Molecular Materials, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany

sparsely populated transannular hydrogen bonds. The *folded* conformations show multiple turns stabilised by highly populated transannular hydrogen bonds comprising centres $F^6W^7K^8$ or $W^7K^8Y^9$. Some of these conformations have not been characterised previously. The equilibrium populations that are experimentally difficult to access were estimated by replica-exchange MD simulations and validated by comparison of experimental NMR data with chemical shifts calculated with density-functional theory. UII exhibits approximately 72 % *open* : 28 % *folded* conformations in aqueous solution. URP shows very similar ring conformations as UII but differs in an *open:folded* equilibrium shifted further toward *open* conformations (86:14) possibly arising from the absence of folded N-terminal tail/ring interaction. The results suggest that the different biological effects of UII and URP are not caused by differences in ring conformations but rather by different interactions with the UTR.

Introduction

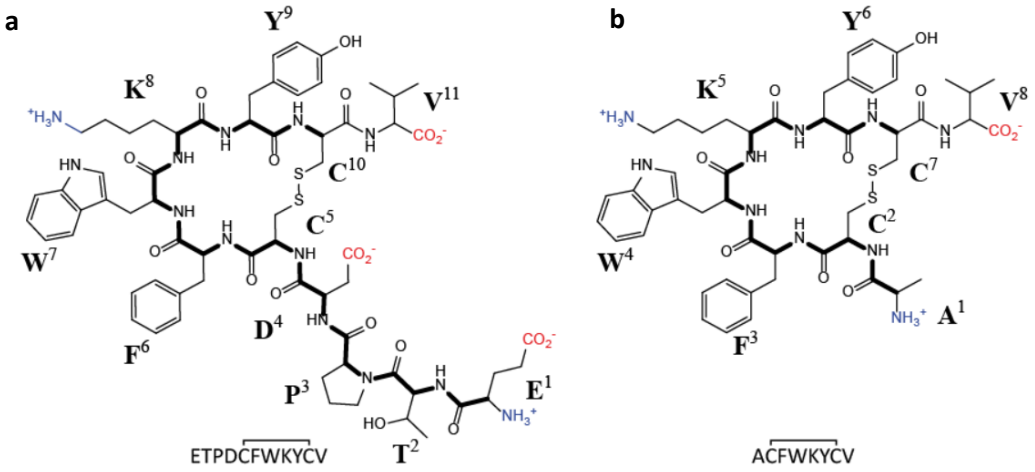
The neuropeptide urotensin II was originally found in the urophysis of teleost fishes.[292] A human homologue[293] of the orphan receptor GPR14[294] (a G-protein coupled receptor that is very similar to the somatostatin receptor first isolated from rats) was identified in 1999.[121-123] UII is the natural ligand of this receptor, now called the urotensin II receptor (UTS2R, UTR).

All vertebrate isoforms of UII show a highly conserved C-terminal sequence: a cyclic 6-residue moiety (CFWKYC) closed by a disulphide bridge and flanked by valine as extra-annular residue (Scheme 6.1a).[295] The length of the N-terminus of human UII is four residues but this is species variable, so that the total peptide length ranges from 11 residues for human UII up to 17 for hamster UII.[23,295-297] Urotensin-related peptide (URP) is a paralog of UII.[48] It has the same C-terminal cyclic moiety as UII but the extra-annular N-terminus of UII is replaced by a single alanine at position 1 in URP (Scheme 6.1b).[124] The 6-membered ring closed by a disulphide bridge is a common motif with other hormone peptides, such as Arg^8-vasopressin and Leu^8-oxytocin.

UII is the most potent vasoconstrictive natural peptide known[293] and both UII and URP are thought to be involved in important physiological processes such as cardiovascular regulation, and endocrine and behavioural effects.[21,23,48,295] Consequently, they are linked to a multitude of pathophysiological processes such as atherosclerosis, heart failure, and many more.[21,23,48,295,298]

Although UII and URP show similar potency at the UTR[124,125,299] and apparently have overlapping binding sites,[300] their signalling outcomes may, nevertheless, differ.[21] UII can behave as an almost irreversible UTR agonist, and the two peptides can affect astrocyte activity differently.[301,302] The

effects of UII or URP are often not conserved across species[43,23,303] and may even be opposite (vasoconstrictive and vasodilative) within the same species[304]



Scheme 6.1a,b (**a**) Human urotensin II and (**b**) urotensin-related peptide

In summary, the urotensinergic system is far from being well understood. Multiallosteric interactions of receptor and ligands or biased agonism that ultimately trigger different functions have been hypothesised.[47]

Biological activity studies have shown that the ring sequence $UII_{(4-11)}$ is necessary to retain full agonistic potency[299,305] and that the motif WKY is essential for receptor activation.[128,305] An intact bridge also seems essential[299,306,307] but need not be a disulphide.[306] However, recently, the first acyclic peptide agonist for UTR has been described, a UII analogue still suggesting WKY as receptor activating motif.[308]

Nuclear magnetic resonance studies in water[109,128] and dimethyl sulphoxide,[129] supported by circular dichroism (CD) spectroscopy,[109] have been interpreted to indicate an unstructured form for human UII with no classical turns or intramolecular hydrogen bonds. However, Lescot *et al.*[126] inferred, from NMR studies, a widened 7,8,9 γ-turn and a 8,9,10 γ-turn with close W^7O-Y^9H^N and K^8O-$C^{10}H^N$ distances for the human UII conformation in water, thus localizing a turn centre in the ring at residues K^8 and Y^9. All NMR investigations show the N-terminal tail to be more flexible than the ring. URP has been suggested from the NMR experiments by Chatenet *et al.*[125] to have an inverse 4,5,6 γ-turn centred at K^5 in water with the intramolecular hydrogen bond W^4O-Y^6H^N. NMR experiments by Brancaccio *et al.*,[127] however, suggest structural flexibility in aqueous solution and a high similarity of URP and UII ring conformations. Carotenuto *et al.*[109] made NMR studies of UII and the smaller URP-like version, $UII_{(4-11)}$, in sodium dodecyl sulphate (SDS) micelles mimicking a cell-surface environment. They found two slowly exchanging states: one specified as β-hairpin with a β-turn type II' centred at W^7 and K^8 and another weakly populated, apparently, with a more

flexible and random structure. The highly structured state was suggested to be the active conformation in the receptor-binding pocket. Analogous experiments for URP in SDS micelles suggested a very similar structure.[127]

We now report unrestrained molecular-dynamics simulations of human UII and URP with the AMBER ff99SB force field on extended timescales (see Table S1 and Figs. S1-S6 of the Supporting Information, SI, Appendix A3). These simulations are designed to investigate the conformational space of the peptides as completely as possible. To rule out small force-field artefacts that might become important for such small peptides, we have also performed additional unrestrained microsecond-scale MD simulations with the CHARMM c36b2 force field. These simulations revealed no significant difference between the conformations obtained with the two force fields, so that we concentrate on the AMBER results, which are more extensive. Replica-exchange molecular-dynamics simulations have been used to improve the conformational sampling and to obtain thermodynamic information. The results are compared with NMR-spectroscopic experiments and a statistical model of the conformational equilibrium in aqueous solution is given.

Methods

Molecular-Dynamics Simulations

MD simulations of the peptides UII and URP were performed with AMBER 10,[168,219] AMBER 14 CUDA,[206,221-223] and CHARMM c36b2.[171] AMBER calculations used the ff99SB force field.[140] Comparison simulations with CHARMM parameter set[171] were used to rule out force-field artefacts. REMD simulations were performed with AMBER. All simulations were carried out with unrestrained distances and explicit water solvation. Further simulation details are given in the Supporting Information (pp S3-S8, Appendix A3).

Conformational Analysis

Conformational clustering of the backbone dihedrals (*overall* states) was performed with DASH.[141,309] Additional sub-clustering of the ring and tail conformations led to a classification of UII and URP conformations in terms of distinct *ring-state types*. As representatives, the overall conformations of highest similarity to each ring-state type were chosen, equivalent to cluster centres (Appendix A3 Table S2). Hydrogen-bond populations and secondary structure motifs of characteristic conformations were calculated from corresponding sections of the MD trajectories using AmberTools with default settings.[142,206,219] Consistency of type assignments of states from different simulations was ensured by comparing the circular similarities of ring torsions, turn

propensities and Cα alignments. Further details are given in the Supporting Information (p S9, Appendix A3).

Principal Component Analysis

A possible correlation of ring and tail motions was analysed with principal component analysis implemented in DASH.[309] Torsion weights were calculated from the coefficients of the relevant principal components. The number of significant PCs was determined by Kaiser´s eigenvalue-one test.[230] PC clustering was visualised *via* 3D-scatter plots of the three most significant principal components colour-coded according to the assigned DASH states in SAR-caddle.[234] Further details are given in the Supporting Information (p S13-S14, Appendix A3).

NMR

NMR spectra were recorded for human UII and URP at pH 3.0/3.5 and pH 6.0 in H_2O and D_2O on a Varian Inova 600 MHz spectrometer. Proton resonance assignments were achieved using 2D ^{1}H-^{1}H total chemical shift correlation spectroscopy (TOCSY)[310] and ^{1}H-^{1}H nuclear Overhauser effect spectroscopy (NOESY) NMR spectra.[311] Resonance assignments of carbon and nitrogen at natural abundance were achieved through standard ^{13}C-^{1}H gradient heteronuclear single quantum coherence (gHSQC) and ^{15}N-^{1}H gHSQC experiments.[136,312,313] Details of sample preparation and NMR experiments are given in the Supporting Information (pp S15-19, Appendix A3).

Density-Functional Theory Calculations on Representative Conformations

The geometries of representative conformations for UII and URP derived from the DASH analysis were first optimised at the B3LYP[237,238,314,315]/6-31G(d)[240,316-319] level with Gaussian 09, Revision C.01.[247] Water solvation was simulated with the default Polarizable Continuum Model (PCM) using the integral equation formalism variant (IEFPCM).[254] The DFT-optimised structures were then used to calculate the magnetic shielding tensors in solution at the same level of theory and converted to ^{1}H, ^{13}C, and ^{15}N chemical shifts using regression formulas based on standard sets of chemical shifts and calculated values. The regression formulae and calculated chemical shifts are given in the Supporting Information (Appendix A3 pp S20-24, Fig. S7, Tables S9-S11).

Equilibrium Models and Experimental Evaluation

Free energies and relative populations (equilibrium models) for the representative conformations of UII and URP were calculated from extended REMD simulations. For each peptide, three simulations of 500 ns were performed starting from different initial conformations (UII: *omega-I*$_{open}$*, folded-I, lasso; URP: omega-I*$_{open}$*, omega-II*, *lasso*). ^{1}H chemical shifts for the equilibria were

calculated *via* linear combination of the calculated shifts for the representative conformations according to the populations suggested by REMD. The calculated shifts of representatives and conformational equilibria were then compared by linear regression with our experimental data for nonexchangeable ^{1}H chemical shifts of UII and URP in aqueous solution at pH 6.0 and pH 3.5, respectively. We have recently published details of chemical-shift comparisons for the closely related vasopressin and have suggested statistical metrics for judging whether conformational equilibria suggested by simulations are consistent with experiment.[2] Here, we used REMD to determine equilibrium populations, rather than the metadynamics. This substitution is tested here.

Further details are given in the Supporting Information (Appendix A3 pp S25-S29, Figs. S8-S9, Tables S13-S15).

Results and Discussion

Conformations of Urotensin II

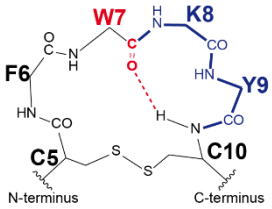
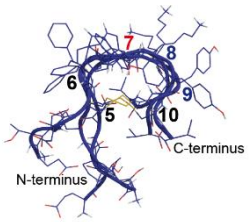
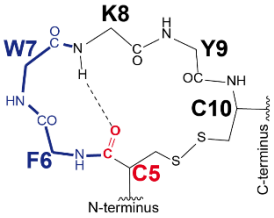
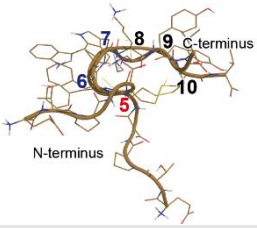
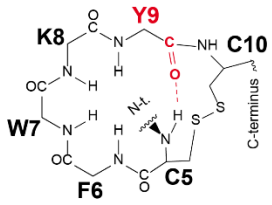
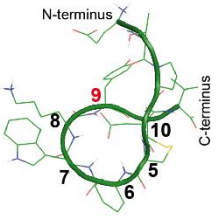
In total, 35 μs of unrestrained MD simulations with the AMBER ff99SB force field supplemented with 1.3 μs CHARMM c36b2 trajectories were used to explore the conformational space of UII (Tables S1-S2 of the SI, Appendix A3). The conformational analysis led to the classification summarised in Table 6.1. UII exhibits two main types of ring states, unfolded *open* and *saddle*-like *folded* ring conformations, which are subdivided into a total of 11 subtypes, each defined by its main turn centre. Secondary structure propensities and populations of transannular hydrogen bonds are given in Tables 6.2 and 6.3.

***Open* Ring-State Types.** Turns in this class are centred at residues $K^{8}Y^{9}$ or $F^{6}W^{7}$ (Table 6.3) with turns fluctuating around ideal β-turn angles (Table S3 of the SI, Appendix A3). The majority of these turns have no or only sparsely populated transannular O_i-H_{i+3} hydrogen bonds (Table 6.2). Only type *scoop* (6,7 β-I) and *omega-I*$_{hbond}$ (8,9 β-I) exhibit significant transannular hydrogen-bond populations but the latter frequently interconverts with the open *omega-I*$_{open}$ state (8,9 β-VIII) resulting in an average population of 44.3 % equivalent to an open turn. Additionally, a ring state was found with no defined β-turns in the ring (*circle*), a loop structure closed by hydrogen bond $W^{7}O$-$C^{5}H^{N}$. The interpreted structures based on NMR studies of UII in aqueous solution resemble the *open* ring-state types (*e.g.*, turn centres at residues 8,9[47] or no transannular hydrogen bonds[109]). Furthermore, the open *omega* conformations of UII show significant similarities to the *clinched open* states of the

related peptide Arg[8]-vasopressin (AVP)[1] (Table 6.4). The *clinched open* conformation of AVP, however, is only populated approximately 30 % in aqueous solution.[2]

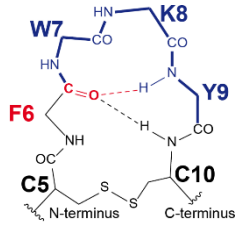
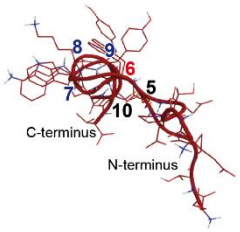
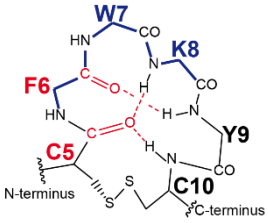
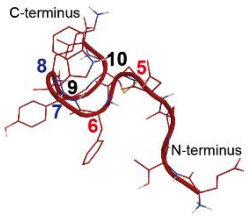
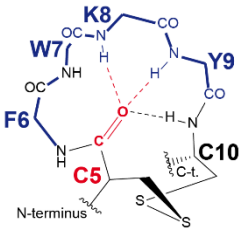
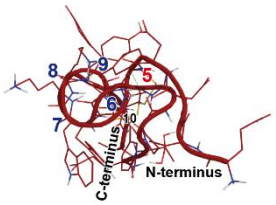
***Folded* Ring-State Types.** The second main cluster comprises *saddle*-like ring conformations with multiple turns, centred either at residues $F^6W^7K^8$ or $W^7K^8Y^9$ (Tables 6.1 and 6.3). This class shows highly populated transannular hydrogen bonds that stabilise the *folded* conformations of the ring (Table 6.2). Subtype *folded-I* (turns centred at $W^7K^8Y^9$ comprising a 7,8 β-I turn) corresponds to the *saddle* state of AVP; subtype *folded-IVb2* (a peptide-bond rotamer of *folded-I* with a 7,8 β-II turn) is equivalent to the *twisted saddle* state of AVP. Interestingly, for AVP, the folded *saddle* conformation is the most highly populated in aqueous solution,[2,103] whereas for UII a folded conformation (β-hairpin centred at W^7K^8) has only been identified experimentally in SDS micelles.[109] The SDS conformation resembles the *folded* conformations found in our MD simulations.

Table 6.1 Classification of ring conformations of UII [§]

Ring-state type	Turn type	H-bond	Subtype	Cartoon	ID [a]
		Open			
		Turns centred at KY (8,9)			
	8,9 β-I	^{7}O-^{10}H	*omega-I$_{hbond}$*		**2**
	8,9 β-VIII	open	*omega-I$_{open}$*		**1**
	8,9 β-II	open	*omega-II*		**3**
		Turns centred at FW (6,7)			
	6,7 β-I	^{5}O-^{8}H	*scoop*		**5**
	6,7 β-I + 4,5 β-I (N-term)	open	*lasso*		**4**
		Loop without defined turn centres			
	(5-9 loop)	^{9}O-^{4}H,^{5}H	*circle*		**10**

[§] Ring-state types are characterised by their turn centres (blue) and the donor oxygen for transannular hydrogen-bond interactions (red). Side chains are indicated by the 1-letter code of the residue. Turn types and corresponding hydrogen bonds populated > 70 % are listed.
[a] Mean torsion angles (Appendix A3 Table S3) and coordinate files of representatives are given in the SI (ID = ID of representative).

Table 6.1 continued

Ring-state type	Turn type	H-bond	Subtype	Cartoon	ID [a]
***Folded (saddle*-like)**					
Multiple turns centred at FWK (6,7,8) or WKY (7,8,9)					
	7,8,9 (7,8 β-I)	^{6}O-(^{9}H,^{10}H)	*folded-I*		**6**
	7,8,9 (7,8 β-II)	^{6}O-^{9}H	*folded-IVb2*		**7**
	6,7,8 (5-9 3_{10}-helix) + 4,5 β-I N-term)	^{6}O-^{9}H, ^{5}O-(^{8}H,^{10}H) ^{3}O-^{6}H	*inv-folded*		**11**
	7,8,9 (6-10 parallel sheet) + 4,5 β-I (N-term)	^{5}O-(^{8}H,^{9}H,^{10}H) ^{3}O-^{6}H	*folded-II*		**8**
	6,7,8 (6,7 β-III') + 4,5 β-I (N-term)	^{5}O-(^{8}H, ^{9}H) ^{3}O-^{6}H	*folded-III*		**9**

[§] Ring-state types are characterised by their turn centres (blue) and the donor oxygen for transannular hydrogen-bond interactions (red). Side chains are indicated by the 1-letter code of the residue. Turn types and corresponding hydrogen bonds populated > 70 % are listed.
[a] Mean torsion angles (Appendix A3 Table S3) and coordinate files of representatives are given in the SI (ID = ID of representative).

Table 6.2 Hydrogen-bond populations and corresponding turn centres of UII ring-state types [§]

Hydrogen bonds		Conformation (ring-state type)							Turn centre
		Open							
		Ω-I$_{hbond}$	*Ω -I$_{open}$*	*Ω-I$_{av}$**	*Ω -II*	*lasso*	*scoop*	*circle*	
W^{7}O	C^{10}H	88.1	18.8	44.3	6.0	0.0	0.0	0.0	8,9
C^{5}O	K^{8}H	0.0	0.0	0.0	0.0	12.4	73.8	0.0	6,7
W^{7}O	Y^{9}H	9.8	8.5	9.9	0.0	0.7	70.4	0.0	8
Y^{9}O	C^{5}H	0.0	0.0	0.0	0.0	0.0	0.0	96.3	(9-5 loop)
Y^{9}O	D^{4}H	0.0	0.0	0.0	0.0	0.0	0.0	92.4	(9-4 loop)
		Folded							
		folded-I	*folded-IVb2*	*inv-folded*	*folded-II*	*folded-III*			
F^{6}O	Y^{9}H	95.8	73.9	95.8	0.0	0.1			7,8
F^{6}O	C^{10}H	63.6	10.5	0.0	0.0	0.1			7,8,9
C^{5}O	K^{8}H	0.0	2.4	96.1	77.2	83.7			6,7
C^{5}O	Y^{9}H	0.0	0.1	0.2	99.4	98.2			6,7,8
C^{5}O	C^{10}H	0.0	0.0	96.9	89.3	0.3			(5-10)
P^{3}O	F^{6}H	0.9	0.2	68.0	84.3	85.1			4,5
T^{2}O	W^{7}H	0.0	0.0	0.0	0.0	61.6			(2-7)

[§] Hydrogen-bond populations are relative to the lifetime of the ring-state type; only those hydrogen bonds are listed that were found to be populated >50 % for at least one ring-state subtype; hydrogen bonds > 70 % (presumably involved in classical turns) are shown in bold. *Average hydrogen-bond population for the frequently interconverting subtypes *Ω-I$_{hbond}$* and *Ω-I$_{open}$* (*cf.* Fig. S1 of the SI, Appendix A3); Ω = *omega.*

Table 6.3 Secondary-structure populations (%) [a] for ring-state types of UII

Ring-state type	UII Residue									Motif [b]
	T[2]	P[3]	D[4]	C[5]	F[6]	W[7]	K[8]	Y[9]	C[10]	
Open										
omega-l_{open}	0.00	1.21	1.24	0.66	0.62	0.12	*27.12*	*27.16*	1.23	T
omega-l_{hbond}	0.00	24.14	24.15	0.02	0.00	0.00	**77.96**	**78.51**	*25.05*	T
omega-$l_{average}$	0.10	4.46	4.52	0.17	0.02	0.00	*47.41*	*47.71*	15.95	T
omega-II	0.35	2.22	1.88	0.02	0.02	0.00	*48.69*	*48.70*	0.52	T
scoop	0.00	3.04	3.04	0.00	**86.91**	**86.94**	0.15	7.79	7.79	T
lasso	0.00	0.05	*53.99*	*56.26*	21.61	18.36	1.92	1.49	0.00	T
circle	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	T
Folded										
folded-I	1.02	10.04	11.83	2.50	0.09	**75.10**	**75.22**	*67.70*	3.20	T
folded-IVb2	0.00	4.96	5.22	0.29	0.51	**78.78**	**86.33**	*68.73*	9.36	T
inv-folded	0.00	2.14	*59.19*	*59.19*	5.83	6.84	7.84	**99.97**	**89.60**	T
	0.00	0.54	20.44	23.31	**92.64**	**93.16**	**92.15**	0.00	0.00	H
folded-II	0.00	0.00	**95.07**	**95.07**	*42.71*	**100.00**	**100.00**	**98.87**	0.00	T
	0.00	0.00	0.00	0.00	*57.29*	0.00	0.00	0.00	*56.47*	P
folded-III	24.93	*43.33*	*63.97*	*64.83*	**99.90**	**99.65**	**99.69**	13.04	0.44	T

[a] Populations > 75 % (classical turns) and > 25 % (potentially open turn) are shown in bold and italics, respectively (for notation of secondary-structure elements, see SI, Appendix A3).[b] T = turn, P = parallel sheet, H = 3_{10}-helix.

Table 6.4 Similarity of ring torsions of $UII_{(5-10)}$, $URP_{(2-8)}$, and $AVP_{(1-6)}$

Conformation (ring-state type)[a]						Circular similarity [b]		Turn type		
UII		URP		AVP		UII/URP	UII/AVP	UII	URP	AVP
Open										
1	*omega-l_{open}*	**3r**	*omega-l_{open}*	**12**	*cl.open*	0.95	0.88	8,9 β-VIII	5,6 β-VIII	4,5 β-$VIII_{dist}$/I
2	*omega-l_{hbond}*	**1r**	*omega-l_{hbond}*	**12**	*cl.open*	0.99	0.83	8,9 β-I	5,6 β-I	4,5 β-$VIII_{dist}$/I
3	*omega-II*	**2r**	*omega-II*		-	0.93	-	8,9 β-II	5,6 β-II	4,5 β-II
4	*lasso*	**6r**	*($lasso_{45pbr}$)*	**27**	*(open)*	0.55[c]	0.55[d]	6,7 β-I	3,4 β-$VIII_{dist}$	2,3
5	*scoop*	-	-		-	-	-	6,7 β-I	-	-
10	*circle*	-	-		-	-	-	(5-9 loop)	-	-
Folded										
6	*folded-I*	-	-	**3**	*saddle*	-	0.93	7,8,9 (7,8 β-I)	-	3,4,5 (3,4 β-I)
7	*folded-IVb2*	**4r**	*hybrid*	**19**	*tw.saddle*	0.89	0.95	7,8,9 (7,8 β-II)	4,5,6 γ	3,4,5 (3,4 β-II)
		5r	*sheet*		-	0.67	-	-	4,5 (ap.sheet β-II)	-
11	*inv-folded*	-	-		-	-	-	6,7,8 (3_{10}-helix)	-	-
8	*folded-II*	-	-		-	-	-	7,8,9 (p.sheet)	-	-
9	*folded-III*	-	-		-	-	-	6,7,8 (6,7 β-III')	-	-

[a] Coordinate files of UII representative (UII 1 to 11, URP 1r to 6r) are given in the SI (Appendix A3); coordinate files of AVP representatives (T16_3,12,19,27) have been published previously[2]. [b] Circular similarity of corresponding ring torsions (1.00 = identical; for methodological details see SI, Appendix A3). [c] $RMSD_{CA\text{-}ring}$ = 0.714 Å. [d] $RMSD_{CA\text{-}ring}$ = 0.218 Å (AVP_{open} is a peptide-bond rotamer of UII_{lasso} which has the same backbone shape but a different peptide bond orientation at residues 2,3). Abbreviations: UII = human urotensin II, URP = urotensin-related peptide, AVP = Arg^8-vasopressin (representative T16 states[1]), *cl.open = clinched open*, ap = antiparallel, p = parallel, dist = distorted, pbr = peptide-bond rotamer, *inv = inverse*.

Are Tail and Ring Conformation of Urotensin II Mutually Dependent? As described above, the structure of UII can be characterised by its ring conformation and by treating the N-terminus as an additional residue. A principal-component analysis (PCA) of the overall torsion space supports this approach. It clusters the overall conformations of UII in accordance to the ring-state types clustered with DASH[141] (Fig. 6.1).

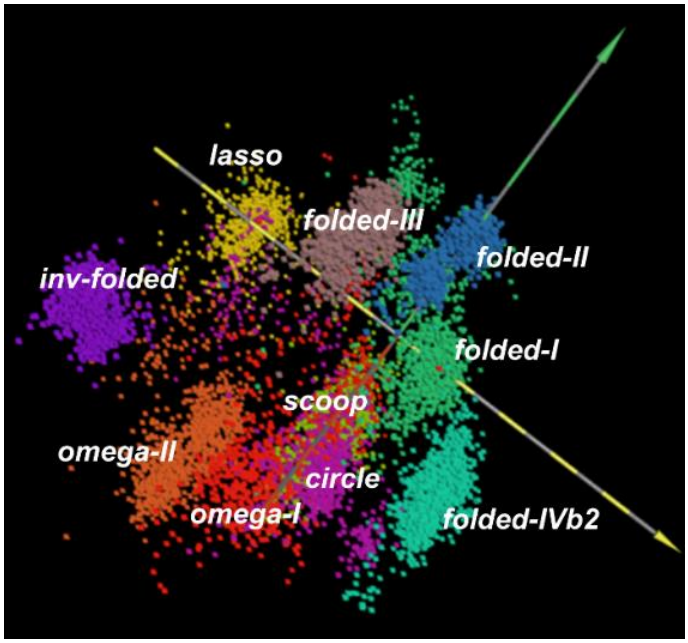


Figure 6.1 PCA clusters of UII conformations. 3D-scatter plot of the three main PCs of the overall backbone torsion space of UII. Each dot represents a conformational snapshot of UII from the MD simulations. Conformations are colour-coded by DASH ring-state types. PCA confirms that DASH clustering of ring conformations is suitable for characterizing the overall structure of UII.

Nevertheless, the tail remains of special interest, as it is the only structural difference between UII and URP. DASH clustering (Figs. S1-S6 of the SI, Appendix A3) reveals that the basic conformation of the N-terminal tail is *extended* or *folded* with the majority of *folded* tail-conformations caused by a single turn centred at either P^3D^4 or D^4C^5 of turn types β-I/VIII or II, as shown in Figure 6.2.

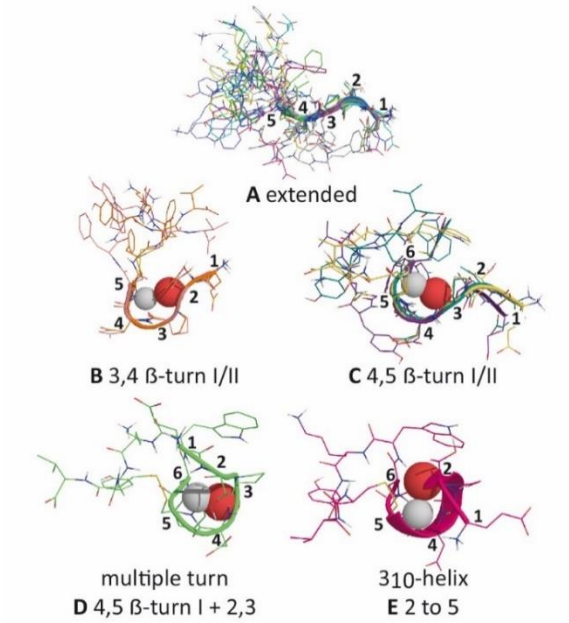


Figure 6.2a-e Tail-state types of UII. Hydrogen and oxygen atoms of hydrogen bonds are represented as spheres.

The relative populations of *extended* and *folded* tail states in the MD simulations vary significantly (*cf.* Figs. S1-S6 of the SI, Appendix A3). Some ring-state types show frequent interconversions of *extended* and *folded* tail states, others none or few; and the *extended:folded* ratio for some types is not consistent between simulations. This raises the question as to whether tail and ring states might be mutually dependent. A qualitative answer is given by analysing the weights of ring and tail torsions of the main significant PCs for each type of ring conformation (Appendix A3 Table S4). If both ring and tail torsions are significantly loaded on one PC, correlation can be assumed. The results are summarised in Table 6.5. Few ring-state types (*folded-I* and *folded-IVb2*) show unambiguously that ring and tail torsions are not correlated, whilst *omega-I* types show uncorrelated ring/tail motions only if the tail is exclusively *extended* (Appendix A3 Fig. S1). For all other types, the PCA results suggest interdependence of ring and tail conformations. This contrasts with AVP, where the tail (the three C-terminal residues) moves essentially independently of the ring.[1] A tentative explanation is the longer tail in UII of four residues facilitates interactions with the ring (*e.g.* by hydrogen bonding). Mutual dependence of ring and tail conformations is a dynamic property that differentiates UII from URP (no tail) and could modulate different bioactivities.

Table 6.5 Relative populations (%),[a] interconversion frequencies [b] and correlation [c] of *extended* and *folded* tail conformation for UII ring-state types

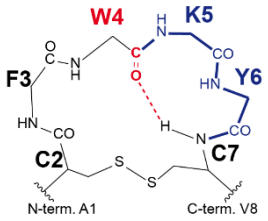
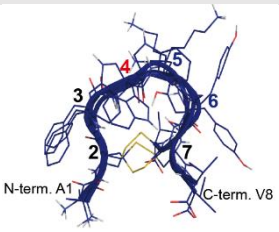
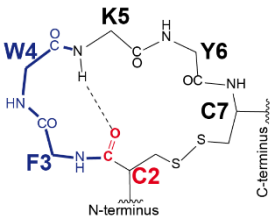
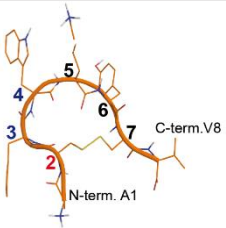
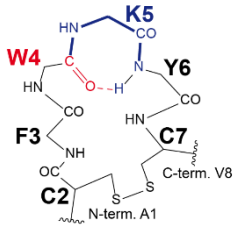
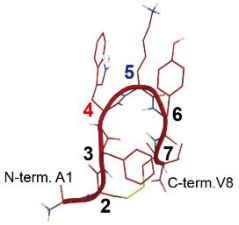
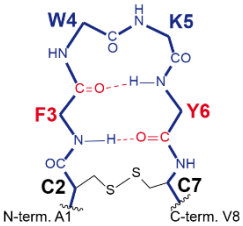
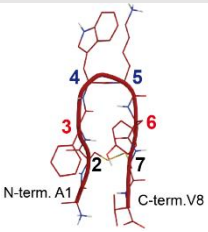
Ring-state type	Correlation ring/tail	Tail conformation [d] *extended* (%)	*folded* (%)	Interconversion *extended/folded*	MD [e]
		open			
omega-I$_{hbond/open}$	no	100.0 (**A**)	-	-	I
	yes	38.1 (**A**)	61.9 (**B**)	*few*	III
	yes	37.7 (**A**)	62.3 (**C**)	*frequent*	IV
omega-II	yes	100.0 (**A**)	-	-	III
	yes	61.9 (**A**)	38.1 (**C**)	*frequent*	XI
scoop	yes	100.0 (**A**)	-	-	III
lasso	yes	40.8 (**A**)	59.2 (**C**)	*frequent*	IV
circle	yes	100.0 (**A**)	-	-	IV
		folded			
folded-I	no	88.6 (**A**)	11.4 (**B**)	*few*	II
folded-IVb2	no	90.2(**A**)	9.80 (**B**)	*few*	III
inv-folded	yes	10.7 (**A**)	89.3 (**C**)	*few*	XI
folded-II	yes	-	100.0 (**C**)	-	V
folded-III	yes	-	100.0 (**D,E,C**) [f]	-	V

[a] Populations are relative to the length of analysed sections occupied by single ring-state types in the MD simulations listed. [b] *cf.* DASH tail-state trajectories. [c] Qualitative results from the overall torsion space PCA: If relevant PCs (Eigenvalue > 1.0) correspond to both ring and tail torsions, then correlation was assumed (for details, see SI, Appendix A3). [d] Turn types (Fig. 6.2) are in parentheses. [e] MD = MD simulation (DASH ring and tail-state trajectories are given in Figs. S1-S6 of the SI, Appendix A3). [f] 40.9 % (D) + 32.9 % (E) + 26.2 % (C).

Conformations of Urotensin-Related Peptide

In total, 22.8 µs MD were analysed for URP (Table S1 of the SI, Appendix A3). In the MD simulations, the majority of URP conformations (98.4 %) belong to the *open* class of *omega* ring-state types (Table 6.6 and Appendix A3 Table S3) with the turn centred at residues K^5 and Y^6 and a circular similarity of more than 90 % to the *omega* states of UII (Table 6.4).

Table 6.6 Classification of ring conformations of URP[§]

Open					
Ring-state type	**Turn type**	**H-bond**	**Subtype**	**Cartoon**	**ID [a]**
Turns centred at KY (5,6)					
	5,6 β-I	^{4}O-^{7}H	*omega-I$_{hbond}$*		1r
	5,6 β-VIII	open	*omega-I$_{open}$*		3r
	5,6 β-II	open	*omega-II*		2r
Turns centred at FW (3,4)					
	3,4 β-VIII	open	*lasso$_{45pbr}$*		6r
***Folded* (*saddle*-like)**					
Ring-state type	**Turn type**	**H-bond**	**Subtype**	**Cartoon**	**ID [a]**
Turns centred at K (5)					
	4,5,6 γ	^{4}O-^{6}H	*hybrid*		4r
Turns centred at WK (4,5)					
	2-7 antip. β-sheet (4,5 β-II)	^{6}O-^{3}H (^{3}O-^{6}H)[b]	*sheet*		5r

[§]Ring-state types are characterised by their turn centres (blue) and the donor oxygen for transannular hydrogen-bond interactions (red). Side chains are indicated by the single-letter code of the residue. Turn types and corresponding hydrogen bonds populated > 70 % are listed. [a] Mean torsion angles (Appendix A3 Table S3) and coordinate files of representatives are given in the SI (ID = ID of representative). [b] 48 % population.

A high similarity of UII and URP ring conformation was postulated also by Brancaccio *et al*. based on their NMR studies.[127] Hydrogen-bond populations at Y^4O-C^7H^N and turn propensities at K^5Y^6 of URP´s *omega type* resembles the data of the corresponding UII conformations (Tables 6.2 and 6.3).

Conformations with turns different to K^5Y^6 are only found as transient states with low absolute populations. There is a variant of the UII *lasso* type with a type VIII β-turn centred at F^3W^4. Two further transient states are comparable with the *folded* conformations of UII. One (denoted as *sheet*) forms an antiparallel β-sheet with a β-II turn at W^4K^5, the other (denoted as *hybrid*) exhibits a γ-turn at $W^4K^5Y^6$ and shows 89 % similarity to the ring torsions of the *folded-IVb2* state of UII. The *sheet* type resembles the postulated single-conformer structure of URP in SDS micelle solution.[127] The *hybrid* type is reminiscent of Chatenet´s NMR-based single-conformer description of URP in aqueous solution.[125]

Determination of UII and URP Equilibrium Populations

Most of the ring-state types described above exhibit significant lifetimes during MD simulation and, therefore, represent candidates for the main conformations in solution. However, interconversions are too infrequent to derive equilibrium populations directly from the MD simulations. We therefore performed extended REMD simulations of UII and URP to determine the relative population of the states and, hence, to calculate their free energies. NMR experiments were carried out to validate these *in silico* equilibria *via* comparison of calculated and experimental chemical shifts using the statistical metrics reported previously.[2]

NMR Experiments. ^1H, ^{13}C, and ^{15}N chemical shifts could be assigned for UII and URP in H_2O at pH 3.0/3.5 and 6.0, with the exception of C and N atoms without directly bonded protons and some rapidly exchangeable H^N atoms at pH 6.0. Our ^1H chemical shifts of UII and URP agree well with those already published[109,125-127] and are complemented by our results for ^{13}C and ^{15}N shifts at the different pH values. The experimental shift lists are given in the Supporting Information (Appendix A3 Tables S5-S8). The pH was varied to see if changing the protonation state induces significant conformational changes. A change to acidic pH values protonates charged carboxylic acid-containing residues (E^1, D^3, and the C-terminal V^{11} in UII; the C-terminal V^8 in URP) and this can affect the local electronic structure, as seen by changes in NMR chemical shifts of these residues and their immediate neighbours. The UII peptide is more affected by pH, changing its protonation state from -1 at pH 6.0 to +2 at pH 3.0, whereas URP only changes from +1 at pH 6.0 to +2 at pH 3.0. However, these pH-dependent changes are small compared to those that occur if the solvent is changed from water to an SDS micelle containing aqueous solution, with no buffer added.[109,127] A significant conformational change such as that found in SDS micelles[109,127] can be excluded. Thus, it can be assumed that the most highly populated conformations of UII and URP at pH 6.0 resemble the published NMR structures in aqueous solution. We eschewed a further classical structure determination using experimental nuclear Overhauser effect (NOE) distances or

coupling constants and focused on determining conformational equilibrium concentrations *via* ^{1}H chemical shifts, which proved to be most efficient for vasopressin.[2] In this context, it is important to note that, while observed NMR chemical shifts represent the time average of the shifts of all structures in a dynamic equilibrium, this is not true of distances derived from NOE peaks. This is because the distance-dependence of the NOE depends on the inverse sixth power (r^{-6}),[320] so that simply averaging the distance (r) will yield incorrect results. Thus, short contacts that occur infrequently can give rise to significant NOE peaks, even though the time-averaged interatomic distance may be large.

For the same reason, NOE peaks that result from several different conformations in equilibrium can masquerade as a single fictitious conformation. A second set of resonances representing a minor population (~10 % of the total) was also observed in the UII NMR spectra. This was identified as the *cis*-Pro^{3} isomer of UII and fully sequentially assigned. As the *cis/trans* conversion in peptides is known to be slow on the NMR timescale[321,322] it will not contribute to fast equilibria and is not discussed here.

Conformational Equilibrium of Urotensin II. The relative populations for the representative conformations of UII from three REMD simulations (with different initial conformations) are given in Table 6.7. This table covers approximately 80 % of the conformational REMD snapshots, the remaining 20 % (circular similarity of ring torsions < 65 %) are transients that cannot be assigned unambiguously to the representatives. All three REMD simulations predict a similar ratio of *open* to *folded* conformations and thus, the simulations can be assumed converged for these main conformational types. Unfortunately, the population of the individual subtypes of *open* and *folded* has not converged and differs strongly between the three REMD simulations (Table 6.7). However, convergence would necessitate significantly longer simulation times, which are currently unobtainable.

A statistical comparison of the calculated and experimental chemical shifts of UII at pH 6 is given in Table 6.8. All *open:folded* equilibria of UII correspond better to the experimental values than any single conformation. The best agreement was found for equilibrium REMD-I, predicting a ratio of 72 % *open* and 28 % *folded* conformations for UII in aqueous solution. A plot of the predicted *vs.* experimental shifts is shown in Figure 6.3. Correlation of calculated and experimental ^{15}N chemical shifts also confirms the ratio of 72:28 *open* to *folded* as the equilibrium that gives the best agreement, although the number of shifts is very small (Table S14 of the SI, Appendix A3).

The correlation of calculated ^{13}C chemical shifts with experimental shifts is satisfactory for the equilibria but gives the best fit for the *omega-I$_{open}$* conformations (Table S13 of the SI, Appendix A3). However, the correlation within the calculated sets of ^{13}C shifts is too high to give unambiguously

distinguishable models (Appendix A3 Fig. S8). This was also found for AVP[2] and is further discussed in the Supporting Information (Appendix A3).

Table 6.7 Relative free energies (ΔΔG, kcal mol^{-1}) [a] and relative populations (%) [b] of representative conformations for UII from REMD simulations

UII representatives		REMD simulations (UII)							
		REMD-I [c]		REMD-II		REMD-III		stddev [d]	
Conformation	ID [e]	ΔΔG	pop%	ΔΔG	pop%	ΔΔG	pop%	ΔΔG	pop%
				open					
omega-I$_{open}$	**1**	0.39	15.19	1.08	8.72	1.09	8.98	±0.33	±2.99
omega-I$_{hbond}$	**2**	0.41	14.76	1.45	4.68	1.19	7.69	±0.44	±4.22
omega-II	**3**	1.04	5.07	2.21	1.29	1.55	4.10	±0.48	±1.61
lasso	**4**	0.00	29.75	0.00	54.11	0.00	56.73	±0.00	±12.15
scoop	**5**	1.43	2.67	3.08	0.30	3.37	0.20	±0.85	±1.14
circle	**10**	1.12	4.53	2.16	1.39	2.08	1.68	±0.47	±1.42
Σ open			***72.0***		***70.5***		***79.4***		
				folded					
folded-I	**6**	1.67	1.76	1.71	3.00	2.00	1.82	±0.15	±0.57
folded-IVb2	**7**	2.28	0.63	3.01	0.34	3.13	0.28	±0.38	±0.15
inv-folded	**11**	0.35	16.39	1.02	9.67	0.75	15.96	±0.28	±3.07
folded-II	**8**	1.21	3.89	1.34	5.58	1.84	2.56	±0.27	±1.24
folded-III	**9**	1.02	5.37	0.95	10.92	-	0.00	±0.04	±4.46
Σ folded			***28.0***		***29.5***		***20.6***		

[a] Average standard deviation of all ΔΔG is 0.37 kcal mol^{-1}. [b] Total population of assigned representatives: REMD-I 82%, II 77%, III 87%. [c] The REMD-I equilibrium gives the best agreement with experiment. [d] stddev = standard deviation. [e] Coordinate files are available as Online Supporting Material. ID = ID of representative.

Table 6.8 Statistical error values (ppm), coefficients of distinctiveness (Δ_σ), and determination (R^2) for the linear regression of calculated and experimental ^{1}H chemical shifts of UII in aqueous solution at pH 6.0[a]

UII representatives and equilibria (*open:folded*)	MSE	MUE	RMSD	WRMSE	Δ_σ	R^2
omega-I$_{open}$	-0.09	0.38	0.51	0.56	1.11	0.9338
omega-I$_{hbond}$	-0.02	0.31	0.42	0.46	0.99	0.9556
omega-II	0.03	0.33	0.43	0.46	1.02	0.9533
lasso	0.03	0.29	0.35	0.38	0.96	0.9682
scoop	0.03	0.41	0.50	0.54	1.26	0.9383
circle	0.00	0.30	0.40	0.42	0.95	0.9622
folded-I	0.04	0.32	0.39	0.43	1.06	0.9627
folded-IVb2	0.11	0.32	0.39	0.40	1.01	0.9661
inv-folded	0.06	0.34	0.42	0.44	1.14	0.9547
folded-II	0.05	0.40	0.49	0.55	1.15	0.9355
folded-III	-0.04	0.37	0.45	0.50	1.19	0.9456
Equilibrium REMD-I (72:28)	0.01	0.21	0.26	0.27	0.75	0.9824
Equilibrium REMD-II (70:30)	0.01	0.22	0.28	0.29	0.78	0.9799
Equilibrium REMD-III (79:21)	0.02	0.23	0.29	0.30	0.81	0.9791

[a] Best results are shown in bold. MSE = Mean Square Error, MUE = Mean Unsigned Error, RMSD = Root Mean Square Deviation, WRMSE = Weighted Root MSE, Δ_σ = coefficient of distinctiveness,[2] R^2 = coefficient of determination.

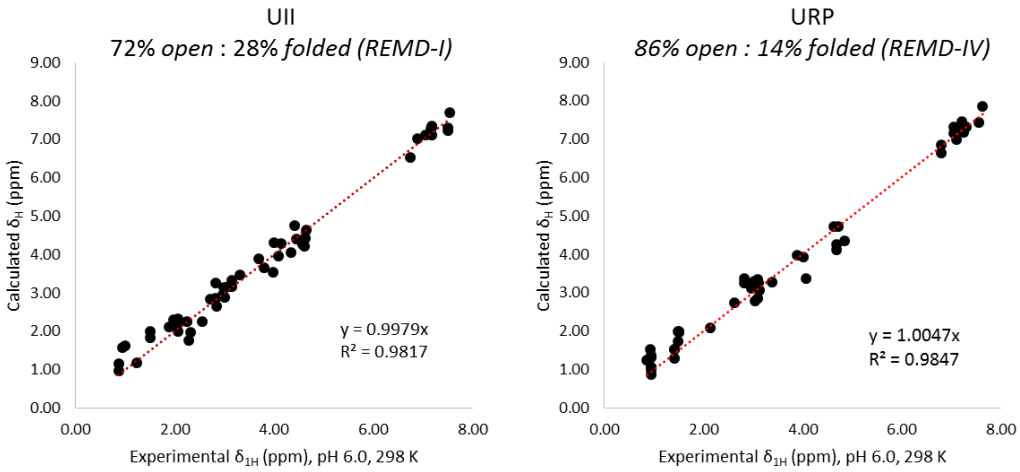


Figure 6.3 Linear regression of calculated ^{1}H chemical shifts for the best predicted equilibria of *open* and *folded* conformations of UII and URP against experimental chemical shifts of nonexchangeable ^{1}H of UII and URP in aqueous solution at pH 6.0, 298 K

Smith and Goodman have proposed the so-called DP4-metric, which they designed especially to discriminate between conformations on the basis of the agreement between calculated and experimental NMR chemical shifts.[258] The DP4 probability is based on Bayes' theorem and is intended to provide an objective assessment of how likely it is that a given diastereomer (or in our case equilibrium distribution of conformations, is correct based on calculated and experimental chemical shifts. In our case the DP4 probabilities for both ^{13}C and ^{1}H shifts help confirm that the chemical shift ensemble resulting from equilibrium REMD-I (72:28) has the highest probability of being a correct assignment (Table S15 of the SI, Appendix A3) in comparison to the single conformations or the equilibria REMD-II and –III. Finally, the dependence of DP4 ("best-fit probability") on variations of the *open:folded* ratio also results in a clear maximum for an equilibrium at approximately 70:30 (Fig. 6.4), in accordance with our prediction.

Besides the experimental shifts of UII at pH 6, a second set of experimental shifts at pH 3 was measured and compared with the calculated shifts. The statistical metrics (data not shown) are extremely close to those at pH 6, which suggests conformational independence of UII for different protonation states (+2 at pH 3, -1 at pH 6).

The seemingly contradictory experimental single-conformer interpretations of UII´s structure in H_2O (no classical turns[109] *vs.* widened 7,8,9+8,9,10 γ-turns[126]) are more precisely a fast (on the NMR timescale) equilibrium of major *open* and minor *folded* ring conformations, rather than any single conformation. A *folded* conformation has so far only been proposed from NMR experiments in SDS micelles, and was suggested to be the bioactive conformation in the UII receptor (UTR).[109] Our results indicate that the proposed bioactive *folded*-type conformations already exist in aqueous

solution to a significant extent, hidden in the fast equilibrium and that, if it is the bioactive conformation, it is selected by preferential binding to the receptor from the conformational ensemble.

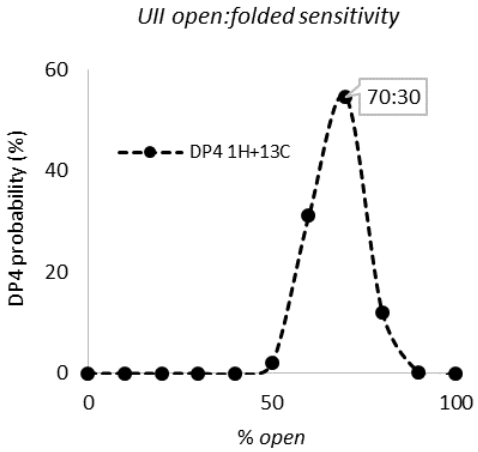


Figure 6.4 Dependence of DP4 probabilities on the *open:folded* ratio of UII. *Open* and *folded* subtype mixtures correspond to the relative concentrations of the 11-component equilibrium REMD-I. The maximum probability (most likely ratio) is approximately 70:30 *open*:*folded*.

Conformational Equilibrium of Urotensin-Related Peptide. Three REMD simulations of URP starting from different initial conformations gave the relative free energies and populations listed in Table 6.9. The representatives cover approximately 70 % of all REMD conformations. The remaining 30 % (circular similarity of ring torsions < 65 %) are transient conformations that could not be assigned unambiguously. The overall ratio of *open:folded* conformations from different REMD simulations are again similar and can be regarded as converged.

Table 6.9 Relative free energies (ΔΔG, kcal mol^{-1}) [a] and relative populations (%) [b] of representative conformations for URP from three different REMD simulations

URP representatives		REMD simulations (URP)							
		REMD-IV [c]		REMD-V		REMD-VI		stddev [d]	
Conformation	ID [e]	ΔΔG	%	ΔΔG	%	ΔΔG	%	ΔΔG	%
				open					
omega-I_{open}	**3r**	0.34	18.92	1.38	5.80	0.45	19.70	±0.47	±6.38
omega-I_{hbond}	**1r**	0.08	29.73	0.49	26.09	0.33	24.24	±0.17	±2.28
omega-II	**2r**	0.00	33.78	0.00	59.42	0.00	42.42	±0.00	±10.65
lasso	**6r**	1.26	4.05	1.79	2.90	1.32	4.55	±0.24	±0.69
Σ open			***86.5***		***94.2***		***90.9***		
				folded					
sheet	**5r**	0.71	10.14	1.38	5.80	1.73	2.27	±0.42	±3.22
hybrid	**4r**	1.36	3.38	-	0.00	1.08	6.82	±0.14	±2.78
Σ folded			***13.5***		***5.8***		***9.1***		

[a] Average standard deviation 0.29 kcal mol^{-1}. [b] Total population of assigned representatives: REMD-IV 74 %, V 69 %, VI 66 %. [c] REMD-IV equilibrium gives the best agreement with experiment. [d] stddev = standard deviation. [e] Coordinate files are available as SI (Appendix A3). ID = ID of representative.

The model that agrees best with experiment is the equilibrium from REMD-IV (calculated ^{1}H chemical shifts for URP are given in Table S12 of the SI, Appendix A3) predicting a ratio of 86 % *open* and 14 % *folded* conformations for URP with a predominance of *omega* conformations (Table 6.10 and Fig. 6.3). This result is further supported by the DP4 assignment probabilities (Appendix 3 Table S15).

Table 6.10 Statistical error values (ppm), coefficients of distinctiveness (Δ_σ) and determination (R^2) for the linear regression of calculated and experimental ^{1}H chemical shifts of URP in aqueous solution at pH 6.0[a]

URP representatives and equilibria (*open:folded*)	MSE	MUE	RMSD	WRMSE	Δ_σ	R^2
omega-I$_{open}$	-0.02	0.27	0.37	0.43	1.02	0.9774
omega-I$_{hbond}$	-0.09	0.32	0.44	0.55	0.99	0.9624
omega-II	-0.11	0.40	0.53	0.64	1.20	0.9456
lasso	-0.08	0.41	0.52	0.64	1.26	0.9489
sheet	-0.05	0.28	0.38	0.43	1.01	0.9755
hybrid	-0.01	0.33	0.44	0.53	1.12	0.9666
Equilibrium REMD-IV (86:14)	-0.08	0.22	0.29	0.31	0.78	0.9847
Equilibrium REMD-V (94:6)	-0.10	0.29	0.38	0.44	0.91	0.9723
Equilibrium REMD-VI (91:9)	-0.08	0.25	0.31	0.35	0.84	0.9815

[a] Best results are shown in bold. MSE = Mean Square Error; MUE = Mean Unsigned Error; RMSD = Root Mean Square Deviation; WRMSE = Weighted Root MSE; Δ_σ = coefficient of distinctiveness; R^2 = coefficient of determination.

Equilibrium REMD-VI also performs better than any single conformation. Only equilibrium REMD-V fits worse than the *omega-I$_{open}$* conformation. It is noteworthy that the average ratio of the frequently interconverting conformations *omega-I$_{open}$* and *omega-I$_{hbond}$* in the long-scale MD simulations is 42:58. This resembles the relative populations of REMD-IV (39:61) and VI (45:55) but not REMD-V (18:82). Insufficient convergence of the *omega-I$_{open}$: omega-I$_{hbond}$* ratio may explain the poor performance of equilibrium REMD-V.

How Do the Conformational Equilibria of URP and UII Differ? Both exhibit predominantly *open* conformations in aqueous solution but UII shows a higher population of *folded* conformations (UII: 28 %, URP: 14 %). This result is consistent with the possible interdependence of ring and tail conformation in UII but not URP, and supports the hypothesis that the N-terminal tail facilitates the formation of *folded* ring conformations.

Conclusions

Conformation and dynamics of UII and URP in aqueous solution were explored and classified by combining computational and experimental methods. The two peptides exhibit similar ring conformations. The structures of both UII and URP in aqueous solution cannot be described by

single conformations. As found previously for Arg^8-vasopressin,[2] UII and URP exist in solution in a conformational equilibrium between *open* and *folded* (*saddle*-like) ring conformations and in combination with *extended* and *folded* tail conformations. In contrast to vasopressin, however, the ring and tail conformations of UII are not independent of each other, so that UII behaves differently to URP, as URP lacks the tail region. *Folded* (*saddle*-like) conformations of URP appear only transiently in unrestricted MD simulations and the equilibrium distribution of conformations that results from REMD simulations and agrees best with experimental ^{1}H chemical shifts is 86 % *open* to 14 % *folded*. The corresponding equilibrium for UII is 72 % *open*:28 % *folded*.

These data suggest that the free-energy penalty for a possible *folded* biologically active conformation is approximately 1.1 kcal mol^{-1} for URP but considerably smaller (approximately 0.6 kcal mol^{-1}) for UII, probably because of ring/tail interactions in UII. This difference may be significant in determining different effects of the two peptides on binding to the UII-receptor (UT2SR, UTR). The high similarity of ring conformations of UII and URP supports Brancaccio´s finding[127] that differences in the biological function are not related to differences in ring conformations. UII and URP show the same conformational main types as the structurally related GPCR-ligand Arg^8-vasopressin. However, both prefer *open*-type conformations in solution, in strong contrast to AVP (70 % *folded* conformations).

All thermodynamically accessible representative conformations of UII and URP can serve as templates for 3D ligand-based drug design or docking, the structural data are given in the Supporting Information (online).

The NMR data reported here supplement and complete published data. They include an almost complete assignment of the spectra of the *cis*-Pro^3 isomers of UII. We have developed a novel and robust procedure to extract conformational equilibria from NMR data by combining experiment with enhanced sampling simulations.

The protocol was developed on AVP[2] and tested here on UII and URP. It seems a powerful tool for exploring the conformational equilibria of intrinsically flexible peptides. In the case of UII and URP, we have used REMD to determine the calculated equilibrium concentrations, rather than the metadynamics procedure used for AVP. Future work will evaluate a variety of enhanced-sampling protocols in order to determine the most suitable for peptide conformational equilibria. The protocol tested and published[2] for Arg^8-vasopressin and based on proton chemical shifts also yields well-defined predictions for UII and URP, here using REMD to determine the calculated equilibrium concentrations.

Unfortunately, we have little information about the lifetimes of the individual conformations. The conformational equilibria are fast on the NMR timescale but too slow for us to be able to sample them adequately in unbiased simulations.

Associated Content. Supporting Information. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00706.
Details of MD simulations, conformational analysis, principal component analysis, NMR experiments, DFT calculations, REMD equilibrium models, ^{13}C linear regression, sensitivity analysis of metrics, ^{15}N linear regression, Tables of experimental and calculated ^{1}H, ^{13}C, ^{15}N chemical shifts. (PDF). Coordinate files of representatives. (ZIP)

Author Information. Corresponding Author. *E-mail: tim.clark@fau.de

Notes. The authors declare no competing financial interest.

Acknowledgments. This work was supported by the European project “Peptide Research Network of Excellence” PeReNE as part of the Interreg IVA France (Channel) – England 2007 2014 program (Interreg EU). NM, MM, and JWE are grateful to AstraZeneca for their support. We thank Helene Castel and Jérôme Leprince (University of Rouen, France) for helpful discussions. Isabelle Milazzo (University of Rouen, France) is acknowledged for providing NMR-structures of h-UII and URP for the CHARMM simulations.

Chapter 7: Related Peptides and General Conformational Classification

The results in this section are not published yet.

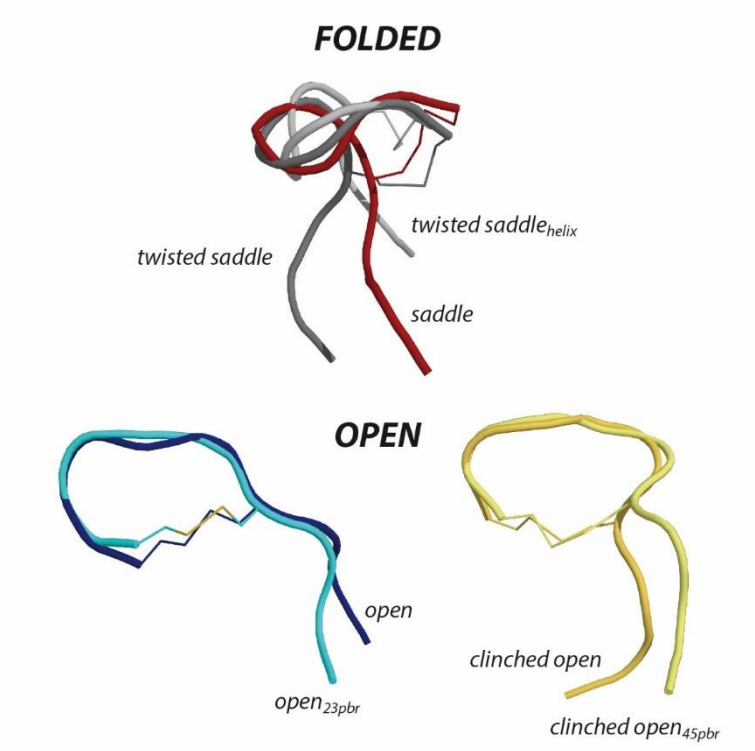


Table of Content Graphic
(Representative conformations of OT)

Foreword

The three preceding chapters showed how to characterise and study the conformation and dynamics of the cyclic peptide hormones AVP, UII and URP. For AVP, four representative conformations, *open, clinched open, saddle* and *twisted saddle,* were identified that can be classified as either *open,* without significant transannular hydrogen bonds, or *folded,* with *saddle*-like conformations that are stabilised by highly populated transannular hydrogen bonds and multiple turns comprising β-turns at residues Phe^3 and Gln^4. The four main conformations of AVP are highly similar to the *open* and *folded* conformations found for UII, *lasso, omega-I, folded-I* and *folded-IVb2* (*cf.* Table 6.4). This suggests that a conformational classification in terms of common *open* and *folded* ring-state types may be generally applicable to cyclic peptides similar to AVP. To test this concept, the conformations of the related peptides oxytocin, deamino-oxytocin, and the

analogue carbetocin were clustered by analysis of long-scale MD simulations following the same protocols as for AVP, UII and URP. Supplementary Information for this chapter is given in Appendix 8.

Molecular-Dynamics Simulations of OT, dOT and CT

Oxytocin. For OT, a total of 50 µs MD simulations were performed in four MD runs with different initial conformations: (i) a *saddle* conformation of OT from the X-ray structure, PDB ID: 1NPO (Fig. 2.2), (ii) an *open* conformation generated by a high-temperature (800K) short-scale MD simulation, (iii) a *clinched open* conformation modelled from the AVP representative T16_12 by mutating Phe^3 to Ile^3 and Arg^8 to Leu^8, and (iv) a *twisted saddle* conformation modelled in the same way from the AVP representative T16_19. The coordinate files of the representative conformations of AVP were given as Online Supporting Information to Paper 1.

Deamino-Oxytocin. The X-ray structure PDB ID: 1XY1 (Fig. 2.1) of dOT was used as starting conformation for a single 3 µs MD simulation. This conformation resembled the ring-state type *twisted saddle* of AVP (Fig. 4.4).

Carbetocin. Starting conformations for CT were homology-modelled from the *saddle* and *open* conformations (i) and (ii) of OT (see above), and each conformation was simulated for 5 µs.
Schemes of the primary structures of OT, dOT and CT are given in Chapter 2 (Scheme 2.1). Further simulation details of OT, dOT and CT are given in Table A8.1 (Appendix A8).

Dynamics and Conformation of OT, dOT and CT

Figures of the RMSD and DASH state trajectories of the MD simulations are given in Appendix A8 (Figs. A8.1 to A8.7). DASH *ring* states, corresponding *overall* states and representative conformations for OT, dOT and CT (Table A8.2) and the mean angles of representatives (Table A8.3) are listed in Appendix A8.

Oxytocin. Two of the four MD simulations, OT_MD-I (Fig. A8.1) and OT_MD-IV (Fig. A8.4), show a transition between the main classes of *open* and *folded* ring-state types. Simulations OT_MD-II (Fig. A8.2) and OT_MD-III (Fig. A8.3) only show interconversions between *open* conformations. This indicates kinetic trapping, as already observed in the simulations of AVP, UII and URP, and

necessitates enhanced sampling to access relative populations of conformation. The tail frequently interconverts between *extended* and *folded* conformations giving similar RMSD plots as for AVP. Thus, its motion can be assumed to be independent. Future PC analyses can be performed to substantiate this observation. *Extended* tail conformations are significantly favoured (absolute populations of approximately 80-90 %). The dynamics of OT seems very similar to AVP; however, the present results do not allow clear conclusions regarding the flexibility of OT compared to AVP. NMR modelling will show if the ratio of *open* and *folded* OT conformations differ from AVP.

In principle, the same representative conformations as for AVP were identified for OT (*open, clinched open, saddle,* and *twisted saddle*). In addition, some significantly populated variants of *open, clinched open* and twisted *saddle* emerged:

(i) a 4,5 peptide bond rotamer of *clinched open* denoted as *clinched open*$_{45pbr}$ (*clop*$_{45pbr}$) that was also found for AVP (Appendix A2 Fig. S2) but weakly populated,

(ii) a 2,3 peptide bond rotamer of the *open (unfolded)* ring-state type denoted as *open*$_{23pbr}$ (Fig. 7.1) resembling the *lasso* conformation of UII (Fig. 6.1), and

(iii) a 3_{10}-helix (1 to 5) variant (Fig. 7.2) of the *twisted saddle* conformation denoted as *twisted saddle*$_{helix}$ comparable to the *inverse-folded* conformation of UII but with different screw-sense.

In simulation OT_MD-III (Fig. A8.3), a 3,4 peptide-bond rotamer of *open*$_{23pbr}$ denoted as *open*$_{2334pbr}$ (Fig. 7.3) was identified as a transient species and will be mentioned here because it corresponds to the *lasso*$_{45pbr}$ conformation, a representative of URP.

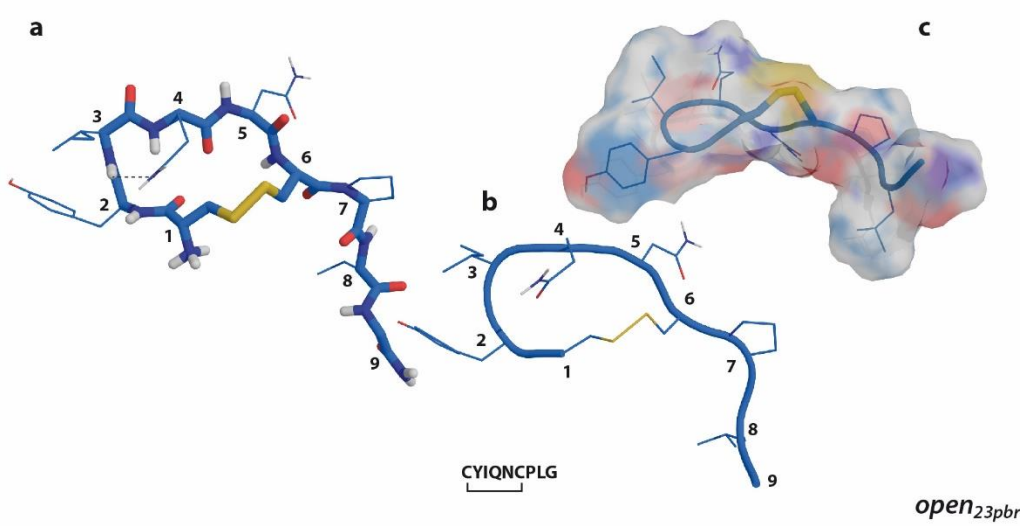


Figure 7.1a-c OT representative for the ring-state type *open*$_{23pbr}$. *Open*$_{23pbr}$ is a peptide bond rotamer of the *open* ring-state type enabling hydrogen-bonding interactions with the sidechain carbonyl-O of Gln^4. The ring-state type *open*$_{23pbr}$ of OT corresponds to the ring-state type *lasso* of UII. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, residue numbers labelled, hydrogen bonding interaction (dashed line); (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) surface.

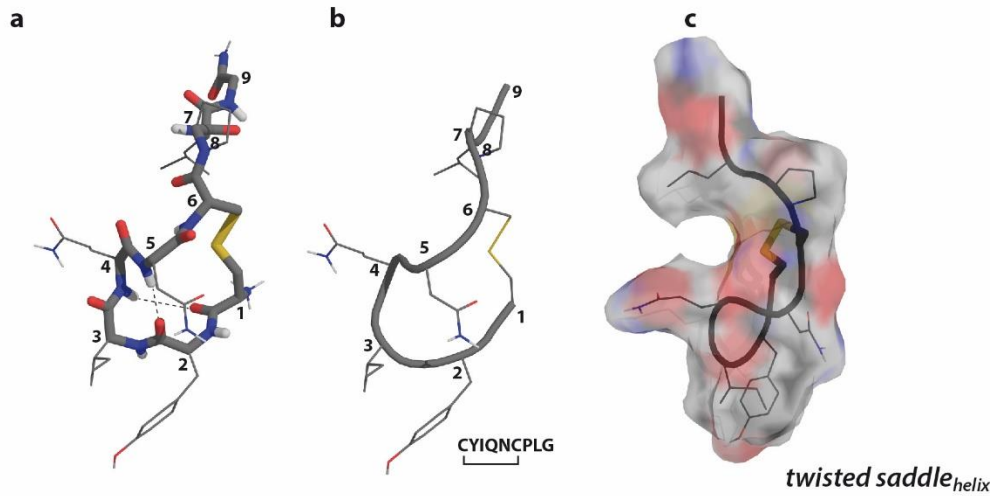


Figure 7.2a-c OT representative for the ring-state type *twisted saddle*$_{helix}$. *Twisted saddle*$_{helix}$ is a variant of the *twisted saddle* ring-state type with ~70 % circular similarity of ring torsions. The ring adopts a 3_{10}-helical form comprising residues 1 to 5 with hydrogen bonds Cys^{1}O-Gln^{4}H and Tyr^{2}O-Asn^{5}H. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, residue numbers labelled, hydrogen bonding interaction (dashed line); (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) surface.

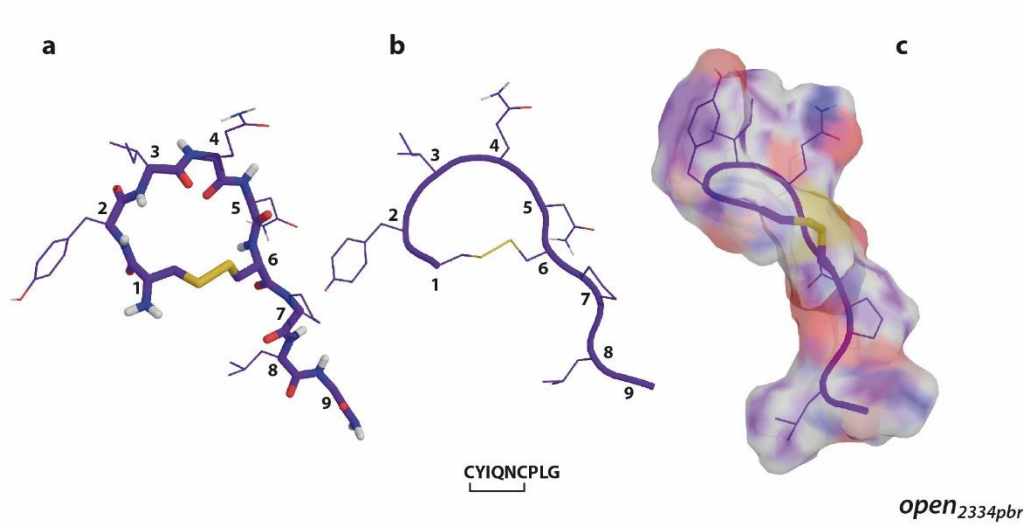


Figure 7.3a-c Transient variant *open*$_{2334pbr}$ of the ring-state type *open*$_{23pbr}$. *Open*$_{2334pbr}$ is a peptide bond rotamer of the *open*$_{23pbr}$ ring-state and corresponds to the conformation type *lasso*$_{45pbr}$ of URP. Depiction: (**a**) backbone (sticks), sidechains (lines), disulphide bridge (sticks), nonpolar hydrogens hidden, residue numbers labelled, hydrogen bonding interaction (dashed line); (**b**) backbone (cartoon), sidechains (lines), disulphide bridge (lines), nonpolar hydrogens hidden, residue numbers labelled; (**c**) structure.

Deamino-Oxytocin. For dOT, only 3 µs MD simulations have been performed so far. The dynamics resembles simulation MD-IV of OT and only shows interconversions between the conformational type *twisted saddle* and its helical variant *twisted saddle*$_{helix}$, except for a transient occurrence of a conformation, denoted as *open*$_{23var}$[i], that corresponds to the *scoop* type of UII. The higher frequency of interconversions may indicate a higher flexibility of dOT. This may mean that the N-terminal NH_3^+ has a stabilising effect on the ring conformation. dOT is a superagonist compared to OT suggesting that flexibility may correspond to activity.

Carbetocin. The MD simulation, starting with the folded *saddle* conformation, soon interconverted to the ring-state type *open* followed by interconversions to *open*$_{23pbr}$ and *clinched open*$_{45pbr}$. The second MD simulation, starting with an *open* conformation, behaved inversely and the initial *open* conformation interconverted to the *saddle* type after 3 µs. Hence, CT exhibits three of the four main ring-state types of AVP and OT, *saddle*, *clinched open* and *open,* only the *twisted saddle* type was not found yet.

[i] a hybrid of OT´s ring-state types *open*$_{23pbr}$ and *clinched open*

Hydrogen-Bond Populations and Secondary-Structure Propensities. Hydrogen-bond populations and secondary-structure propensities of the representative conformations of OT, dOT and CT are given in Table 7.1. Values for AVP are added from Table 4.4 and 4.5 for comparison. Only populations higher than 10 % are listed. The peptides of a particular ring-state type show similar hydrogen-bond and secondary-structure populations. This confirms that the two properties are suitable as diagnostic characteristics for ring-state types. It can nicely be seen how *folded* ring-state types exhibit highly populated transannular hydrogen bonds and well-defined turns in the ring, whereas the *open* ring-state types show no or only weakly populated β-turn centres and hydrogen bonds. The ring-state type variants, *clinched open*$_{45pbr}$ and *open*$_{23pbr}$ can position the residue Gln^4 above the ring, which enables occasional interactions of the sidechain carboxy-O with amide protons of the ring.

Table 7.1 Hydrogen-bond populations (%) and secondary-structure propensities (%) of representative ring-state types of cyclic peptide hormones (AVP, OT, dOT, CT)

Ring-state type	Hydrogen bonds >10 %[a]				Turn propensities >10 %[b] Turn centres					Ring Conformation (Structure Example)
saddle										**(Fig. 2.2)**
	2O5H	**2O6H**	**6O9H**		**3**	**4**	**5**	**7**	**8**	
AVP	95.7	83.2	10.8		94.1	93.9	89,3	18.6	17.3	3,4,5 multiple turn 3,4 β-I (7,8 β-II)
OT	96.5	87.5	11.5		95.9	95.9	92.7	19.3	18.3	
CT	97.3	92.3	11.5		98.4	98.4	96.3	10.7	10.4	
twisted saddle										**(Fig. 2.1)**
	2O5H	**2O6H**	**3O5H**		**3**	**4**	**5**	**7**	**8**	
AVP	82.6	37.3	23.8		90.6	93.8	61.9	10.8	10.8	3,4,5 multiple turn 3,4 β-II
OT	72.5	52.3	32.2		90.4	97.1	74.4	16.0	16.0	
dOT	73.7	60.3	24.9		95.6	97.0	79.2	-	-	
twisted saddle$_{helix}$										**(Fig. 7.2)**
	1O4H	**2O5H**	**6O9H**		**2**	**3**	**4**	**7**	**8**	
OT	55.18	85.67	10.5		-	31.3	29.3	12.8	12.8	
					68.0	68.0	68.0			3,4 turn within 3_{10}-helix
dOT	42.0	88.53	38.6	35.6		53.9	49.9	28.3	28.3	3_{10} helix (1 to 5)
					35.0	35.0	35.0			
clinched open										**(Fig. 4.1)**
	3O6H	**4O6H**	**3O5H**		**4**	**5**		**7**	**8**	
AVP	27.9	18.7	10.2		46.3	46.3		21.0	21.0	open dist 4,5 β-VIII/I
OT	15.7	26.6	--		28.9	28.9		-	-	
clinched open$_{45pbr}$										**(Fig. A2.S1)**
	3O6H	**4O6H**	**3O5H**		**4**	**5**		**7**	**8**	
OT	-	-	-		37.8	37.8		-	-	open dist 4,5 β-II
CT	-	-	-		39.3	39.3		14.2	14.2	
	4OE3H	**4OE2H**	**4OE4H**							
CT	38.0	32.3	23.3							4OE Gln^4/ ring Hbonds
open										**(Fig. 2.3)**
	1O4H	**2O4H**	**4O6H**		**2**	**3**		**7**	**8**	
AVP	12.0	38.6	-		20.2	20.2		-	-	
OT	15.6	15.9	-		31.7	31.7		10.8	10.8	open, no classical turn
CT	19.7	16.7	10.2		41.8	41.8		12.1	12.0	
open$_{23pbr}$										**(Fig. 7.1)**
	4OE3H	**4OE2H**	**4OE4H**		**2**	**3**		**7**	**8**	
OT	42.6	36.5	29.6		-	-		-	-	open, no classical turn
CT	38.0	32.3	23.3		-	-		-	-	4OE Gln^4/ ring Hbonds

[a] Hydrogen-bond populations are relative to the lifetime of the ring-state type. Notation: Residue numbers of carbonyl O and amide H. Populations > 70 % are highlighted. [b] Secondary-structure propensities > 75 % indicate classical turns. Abbreviation: dist = distorted.

Comparison of AVP, OT, dOT and CT with UII and URP

UII and URP share the same motif of a cyclic 6-residue moiety with the peptide group of AVP, OT, dOT and CT. Their ring conformations can be classified similarly to the peptides as *open (unfolded)* or *folded*. However, their sequences, sequence lengths, the position of the ring within the sequence and the ratio of *open* and *folded* conformations differ significantly. The sequence similarity of AVP, OT, dOT and CT seems high enough to result in the same representative conformations (*saddle, twisted saddle, clinched open, open* and their *variants*). It has been shown for both groups of peptides that particular ring-state types can be characterised by their turn types and hydrogen-bond populations and that turn centres differ distinctly for *open* and *folded* conformations. The question arises as to whether a general classification of conformations for peptide hormones with similar cyclic 6-residue motifs is possible based only on the positions of the turn centres (Fig. 7.4) in the ring independently of the sequence.

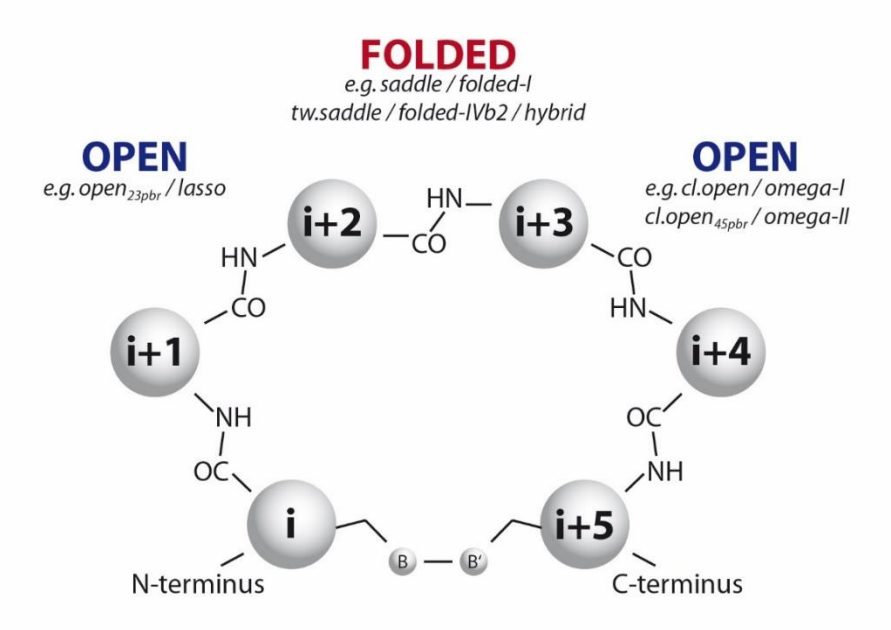


Figure 7.4 Conformational classification of peptide hormones with a 6-residue ring motif. *Folded* ring conformations show turns at residues i+2 and i+3, *open* (*unfolded*) ring conformations either at i+1,i+2 or i+3,i+4. The bridge is symbolised with B and B' (= S for disulphide bridges).

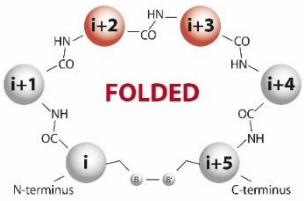
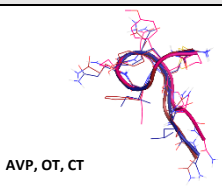
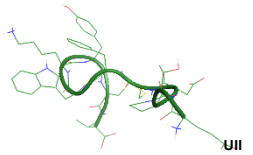
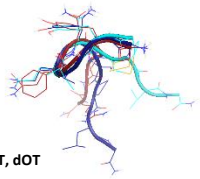
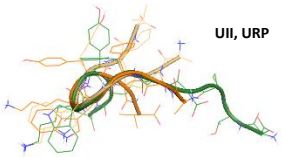
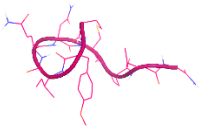
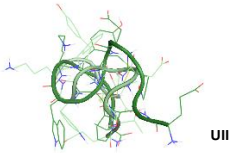
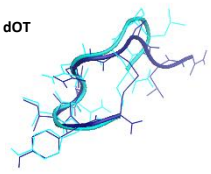
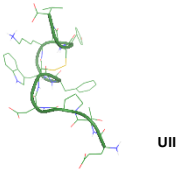
To answer this question, a comprehensive analysis of the circular similarity of ring torsions $\Phi\Psi$ i+1 to i+4 and Φ i+5 of all peptides studied in this work was performed (Table A8.4).

The striking result is the finding that almost all representative conformations of AVP-like peptides (AVP, OT, dOT, CT) can be assigned to representative conformations of UII and URP with ring torsion similarities of > 90 %. Ring-state subtypes *scoop, lasso*$_{45pbr}$ and *folded-II/III* of UII/URP were found for dOT and OT as transients. The ring-state subtype *open* of AVP, OT and CT was not found for UII

and URP but conformational variants were found (*open*$_{23pbr}$ = *lasso*, *open*$_{23var}$* = *scoop, open*$_{2334pbr}$ = *lasso*$_{45pbr}$). Helical variants were identified for UII, OT and dOT, however, with different screw-senses.

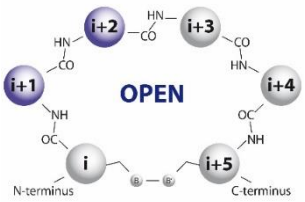
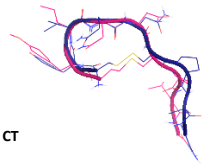
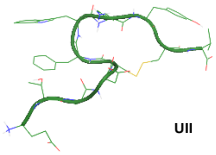
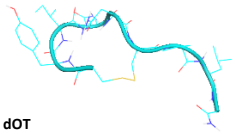
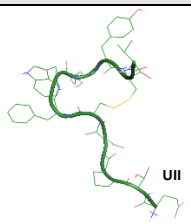
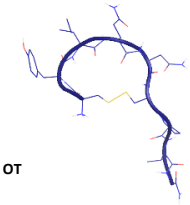
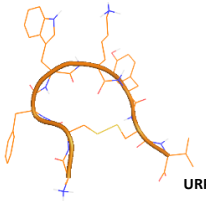
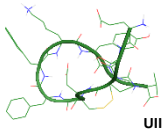
In Tables 7.2 to 7.4, the assignment of representative ring-state types of AVP, OT, dOT and CT to representatives of UII and URP is illustrated. Table 7.2 includes the ring-state types of the *folded* class that all comprise residues i+2 and i+3 as turn centres. Tables 7.3 and 7.4 provide the ring-state types of the *open* class subdivided by conformations with turns centres at i+1,i+2 and i+3,i+4.

Table 7.2 *Folded* (*saddle*-like) ring-state types with turns centred at i+2,i+3: a tabular assignment of representative conformations of AVP, OT, CT, dOT (i=1) to UII (i=5) and URP (i=2)

Ring-state types with turns centred at i+2,i+3				
				
saddle	***folded-I***	**Turn Type**	**Hbonds [a] > 70 %**	**Circular Similarities**
		i+2,i+3 β-I, multiple turn i+1 to i+5/6	$^{i+1}$O-($^{i+4}$H,$^{i+5}$H)	AVP/UII 93 % OT/UII 94 % CT/UII 95 %
twisted saddle	***folded-IVb2, hybrid, sheet***	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		i+2,i+3 β-II	$^{i+1}$O-$^{i+4}$H	AVP/UII 95 % OT/UII 94 % dOT/UII 91 % AVP/URP 91 % OT/URP 92 % dOT/URP 91 % UII/URP 89 %
saddle$_{var}$*	***folded-II (folded-III)[b]***	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		parallel sheet i+1 to i+5	^{i}O-($^{i+3}$H,$^{i+4}$H,$^{i+5}$H)	CT/UII 96 %
twisted saddle$_{helix}$	***(inv-folded)***	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		3_{10}-helix i to i+4	$^{i+1}$O-$^{i+4}$H (^{i}O-$^{i+3}$H)[c] (^{i}O-$^{i+5}$H)[c]	(OT/dOT 37 %) (dOT/UII 37 %)

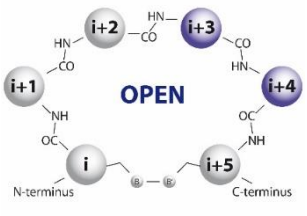
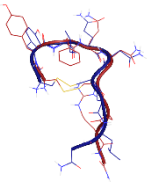
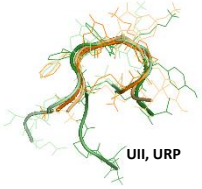
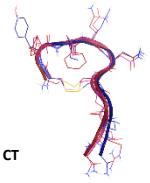
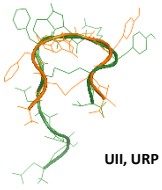
* Only found as transient conformation. [a] Hydrogen-bond population. [b] *Folded-III* can be regarded as tail-variant of *folded-II*. [c] Only UII.

Table 7.3 *Open (unfolded)* ring-state types with turns centred at i+1,i+2: a tabular assignment of representative conformations of AVP, OT, CT, dOT (i=1) to UII (i=5) and URP (i=2)

Ring-state types with turns centred at i+1,i+2				
				
***open*$_{23pbr}$**	***lasso***	**Turn Type**	**Hbonds [a] > 70 %**	**Circular Similarities**
		i+1,i+2 (β-I)	-	OT/UII 92 % CT/UII 91 %
open*$_{23var}$	***scoop***	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		i+1,i+2 β-I	(O_iH_{i+3})[b]	OT/UII 72 %
open*$_{2334pbr}$	***lasso*$_{45pbr}$**	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		i+1,i+2 β-VIII	-	OT/URP 94 %
open		**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
	-	(i+1,i+2, undefined)	-	-
	circle	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
-		(-)	-	-

*Only found as transient conformation. [a] Hydrogen-bond population. [b] UII.

Table 7.4 *Open (unfolded)* ring-state types with turns centred at i+3,i+4: a tabular assignment of representative conformations of AVP, OT, CT, dOT (i=1) to UII (i=5) and URP (i=2)

Ring-state types with turns centred at i+3,i+4				
				
clinched open	***omega-I***	**Turn Type**	**Hbonds [a] > 70 %**	**Circular Similarities**
		i+3,i+4 β-I/VIII	-[b]	AVP/UII 88 % OT/UII 86 % AVP/URP 88 % OT/URP 86 %
clinched open$_{45pbr}$	***omega-II***	**Turn Type**	**Hbonds > 70 %**	**Circular Similarities**
		i+3,i+4 β-II	-	AVP/UII 96 % OT/UII 95 % CT/UII 95 % AVP/URP 95 % OT/URP 94 % CT/URP 95 %

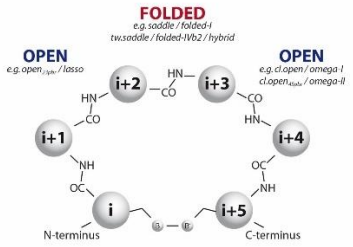
[a] Hydrogen-bond population. [b] Related to *omega-I$_{av}$*

Conformational Classification of Cyclic Peptide Hormones

(with 6-residue ring moiety closed by a disulphide bridge)

The following general conformational classification is suggested based on the occurrence of highly similar ring conformations for all peptides studied in this work and the previous finding that the ring conformations are the principal conformational characteristic of these peptides (*cf.* Chaps. 4 and 6):

1. **Main Conformational Classes: *Open* and *Folded.*** The ring conformations can be generally classified as either *folded* with highly populated transannular hydrogen bonds or *open* (*unfolded*) without significantly populated transannular hydrogen bonds. *Folded* (*saddle*-like) conformations always comprise residues i+2 and i+3 as turn centres, often as part of a multiple turn with highly populated transannular hydrogen bond(s) and well-defined β-turns. *Unfolded* (*open)* conformations exhibit turn centres at residues i+1,i+2 or i+3,i+4 and show no or only weakly populated transannular hydrogen bonds.



(Figure A7.4)

2. **Ring-State Types.** The main classes of *open* (*unfolded*) and *folded* (*saddle*-like) ring conformations comprise several conformational subtypes that are characterised by their turn centres, turn types and the population of transannular hydrogen bonds. Representative subtypes are denoted as ring-state types. Four relevant ring-state types have been identified for AVP:

   saddle, twisted saddle, clinched open, open.

   The MD simulations of OT revealed several variants of these ring-state types, *e.g.*:

   twisted saddle$_{helix,}$ *clinched open*$_{45pbr}$, *open*$_{23pbr}$, *open*$_{2334pbr}$

   These variants are mainly peptide-bond rotamers of the four main ring-state types of AVP with similar backbone shapes.[i] Highly similar ring-state types were also found for UII, *e.g.*:

   folded-I (= saddle), folded-IVb2 (= twisted saddle), omega-I (= clinched open), omega-II (= clinched open$_{45pbr}$*), lasso (= open*$_{23pbr}$*)* and *lasso*$_{45pbr}$ *(= open*$_{2334pbr}$**, open)*

[i] The rotation of a peptide bond leads to significantly different backbone torsions but only insignificant changes of the Cα coordinates; thus the backbone shape remains similar.

An illustrated summary of the representative ring-state types of all peptides studied in this work is given in Tables 7.2 to 7.4. Detailed descriptions are given in the associated chapters.

3. **Tail Conformations: *Extended* and *Folded.*** The N- or C-terminal tails can be regarded as additional residue of the ring with individual conformation. The 3- and 4-membered tails of the investigated peptides adopt two main conformations, *extended* or *folded*. The *extended* form is usually favoured, most likely to minimise sterically hindrance. UII´s ring-state types, however, exhibits in part higher populated *folded* tail conformations. *Folded* tails comprise β-turns that include one of the bridge residues probably induced by the Pro-residues of the tails (*cf*. Scheme 4.2). Figures of representative tail conformations for AVP and UII are given in Figures 4.8 and 6.2.

4. **Overall Conformations.** Overall conformations can be described as combination of ring and tail conformations (*cf.* Table 4.7):

   overall conformation = *ring* conformation + *tail* conformation

   For the majority of *overall* conformations, the tail moves essentially independently of the ring with more frequent interconversions but a correlation of ring and tail conformations cannot be excluded, as has been found for UII. It can be assumed that the probability of ring-tail interaction increases with tail length.

FINAL CONCLUSIONS AND OUTLOOK

Summary and Conclusions

Extensive µs-scale MD simulations were performed to study the conformational space of peptide hormones with a cyclic 6-residue moiety as putative bioactive motif.

Conformational Classification. A general classification of conformations of peptides with the above-mentioned ring motif has been proposed (Chap. 7) based on the identification of main conformations for AVP, OT, dOT, CT, UII and URP. A detailed description of representative conformations for each peptide studied was given (Chaps. 4-7) and their similarity was discussed (Chap. 7).

All peptides studied exhibit similar ring conformations, independent of their primary sequence. They can be assigned to two main classes of ring conformations: *open* and *folded.* The *open* class comprises largely *unfolded* ring conformations with turns centred at i+1,i+2 or i+3,i+4 and no significantly populated transannular hydrogen bonds (Tables 7.3-7.4). The second class comprises strongly *folded* conformations and turns centred at residues i+2,i+3 with highly populated transannular hydrogen bonds (Table 7.2). The short N-terminal tail of AVP (3 residues) moves essentially independently of the ring and the same can be assumed for OT, dOT and CT. Peptides with these short N-terminal tails clearly prefer *extended* tail conformations. UII, however, exhibits several *overall* conformations where ring and tail (4 residues, C-terminal) are correlated, which might account for the different bioactivities of UII and URP (Chap. 5). The *ring* conformation is the main conformational characteristic of the cyclic peptides, which was proven by PCA; *overall* conformations can be regarded as modular combinations of *ring* and *tail* conformations (Chap. 4).

Conformational Equilibria in Aqueous Solution. The frequency of interconversions between main conformations in the MD simulations was too low to deduce equilibrium populations. The AMBER force field showed a clear bias towards main potential minima. For this reason, enhanced sampling was performed for AVP, UII and URP in the research groups of Prof. Clark (AVP, metadynamics) and Prof. Essex (UII/URP, replica exchange MD) that resulted in *in silico* approximations for equilibrium concentrations and confirmed the selection of representative conformations resulting from long-scale MD simulations. A technique was developed and introduced to validate the *in silico* equilibria *via* direct comparison of DFT-calculated ^{1}H chemical shifts with the experimentally observable NMR chemical shifts (Chap. 5). Accordingly, the best approximation for AVP is an equilibrium of 70 % *folded* (*saddle*-like) and 30 % *open* (*unfolded*) conformations; for UII and URP, in contrast, *open*

conformations are favoured with *open:folded* ratio of 72:28 (UII) and 86:14 (URP). The results help explain the seemingly contradictory structural descriptions in the literature of either *single-conformations* and *flexible, disordered* conformations (a literature review was given in Tables 2.3 to 2.6): *Single-conformation* descriptions are in accord with the major conformations of the *in silico* equilibria, which is the folded *saddle* conformation for AVP and OT and the open *omega-I (= clinched open)* conformation for UII and URP; a structural description as *flexible* or *disordered* indicates fast interconversions of multiple conformations, which is confirmed by the *in silico* predictions in this thesis.

DASH. The analysis method *DASH*, used for clustering, proved to perform excellently for processing the demanding volumes of data and the user interface was optimised for application to AMBER trajectories (*amberDASH*).

NMR. Experimental ^{15}N chemical shifts of AVP, UII and URP and the assignment of ^{1}H, ^{13}C and ^{15}N chemical shifts of *cis*-Pro^{3}-UII from this work (Appendices A2 and A3) complement published NMR data. ^{1}H and ^{13}C chemical shifts of AVP, UII and URP are in accord with NMR data in the literature.[103,109,127]

Relevance and Outlook

Development of Validation Technique. The validation technique served well to predict the equilibrium populations of the peptides in this study and promises to be generally applicable to interpret NMR spectra of intrinsically flexible peptides or peptide sequences. For NMR-modelling of cyclic peptides, structurally related to the peptides in this work, it is recommended to choose a maximum number of representative subtypes, refined by their *overall* states with *folded* and *extended* tails, to model the main classes of *open* and *folded* conformations with maximum accuracy. However, the method needs to be developed further and several problems need to be addressed, *e.g.*: (i) The convergence problem of MD simulations. Even enhanced sampling does not guarantee convergence. However, the surprisingly consistent REMD-results for the ratio of main conformational types (*open* and *folded*) of UII and URP despite the non-convergence of the sub-state populations are promising. It suggests that reliable results can be expected as soon as the main conformational types are converged. Further convergence will refine the accuracy of the model but will not alter the magnitude of the ratio. Currently, enhanced-sampling protocols are tested for flexible peptides in the groups of Prof. Clark and Prof. Essex to find the most suitable one.

(ii) The significance of models needs further improvement, *e.g.* by model refinement *via* clustering of sidechain conformations. (iii) The reliability and general applicability of the method should be evaluated by further examples. Presently, OT is being processed.

Drug Design and Docking Calculations. The results are relevant for drug design and docking calculations as each thermodynamically stable conformer can be a bioactive conformation in terms of multi-allosteric interactions with the receptor. Even less populated conformations can be extracted continuously from a solution equilibrium (population shift). Thus, all representative conformations in this work (Chaps. 4-7) may offer templates for drug design and docking calculations and the general classification may help to predict main conformations of similar peptides. For drug design, one representative for each turn-centre in the ring should be chosen (*cf*. Fig. 7.4), *e.g.* a *saddle (folded-I)*, an *open*$_{23pbr}$ *(lasso)* and a *clinched open (omega-I)* conformation. For docking simulations, at least one *open* and one *folded* conformation should be considered, *e.g.* an *open*$_{23pbr}$ and a *saddle (folded-I)* conformation. The conformational classification can also be used to define a reasonable selection of starting conformations for MD simulations of structurally related peptides *via* homology modelling.

Subsequent Simulations. AVP representatives of this work have already been established as initial conformations and reference for subsequent research projects. Saleh *et al.,* for example, studied the binding free energy of the interaction between AVP and its renal receptor V2R with molecular-dynamics simulations and metadynamics enhanced-sampling and suggested a three-site mechanism.[28] The simulations revealed atomistic details of the bioactive conformations of AVP, yet unpublished: According to their simulations, AVP undergoes multiple conformational interconversions along the binding path. A *vestibule* and an *intermediate site* of V2R required *folded* conformations of AVP. However, the ligand interconverted to *open* conformations (mainly *clinched open*) frequently when crossing the barrier between *intermediate* and *orthosteric site*. The final conformation in the binding pocket found was a *saddle* conformation.

Flexibility. In the Introduction of this thesis it was mentioned that flexibility and even disorder may affect bioactivity. For the peptides investigated in this work, the ratio of *open* and *folded* conformations could serve as measure for their flexibility in terms of "*the higher the population of folded conformations - the lower the disorder*" or "*the higher the population of open conformations - the higher the flexibility*". This already characterises URP as being more flexible than UII and might be a further factor related to different bioactivities towards UTR. From experiments, it is known that AVP is more flexible[101] than OT and a partial agonist of OTR. Calculations of OT equilibria are currently in progress. If the conformational equilibrium of OT is shifted to *folded* conformations

relative to AVP, its lower flexibility would be confirmed and could be related to AVP´s agonism towards V2R. The same applies for dOT, a known super-agonist for OTR. The known order of agonistic activity (dOT > OT > AVP) towards OTR could be compared with their *folded:open* ratio of conformations in relation to flexibility. UII and URP, finally, show a significant preference of *open* conformations in solution in contrast to AVP, suggesting a significantly higher flexibility. Future simulations of ligand-receptor interactions may show if this notable property difference is implicated in different reaction paths.

Further Possible Projects. Even if the main conformational types of the 6-membered cyclic peptides seem to be quite independent of the residues, a correlation of the primary sequence to the *open:folded* conformational equilibrium, and thus the flexibility, is very likely. This putative relation could be studied by calculation of further *open:folded* equilibria of sequence mutants.

The present MD trajectories still provide plenty material for the study of further atomistic details. For example, in Chapter 4 (Paper 1), key-torsions for conformational interconversions were analysed for AVP. These may help understand the mechanisms of conformational interconversions, so that they can be used to define reaction coordinates. This approach could be pursued for the other peptides. Another example is the bridge, which was discussed for AVP (Chap. 4). Main conformational types of the bridges and a potential mutual interdependence with ring conformations should be analysed for all peptides to complete the conformational classification. As already mentioned, the clustering of sidechain conformations deems it necessary to improve the accuracy of *in silico* equilibrium models and, for this, suitable analysis protocols need to be tested.

A further project, not completed yet, studies the *cis/trans* isomerisation of the proline residues in AVP and UII. First results for AVP indicate that a *trans/cis*-Pro^{7} interconversion *via* a single-path rotation of the peptide bond does not affect the ring conformation. Results may give insight into the implication of Pro *cis/trans* mutation for bioactivity.

.

APPENDICES

A 1: Reprint Supporting Information Paper 1

The Supporting Information is available on the Springer Link Publications website at DOI:10.1007/s00894-014-2485-0.

Supporting Information

Conformation and Dynamics of 8-Arg-Vasopressin in Solution

Elke Haensele[a], Lee Banting[a], David C. Whitley[a], and Timothy Clark*[,a,b]

[a]E. Haensele, Dr. L. Banting, Dr. D. C. Whitley, Prof. Dr. T. Clark

Centre for Molecular Design, School of Pharmacy and Biomedical Sciences

University of Portsmouth, Portsmouth PO1 2DT (United Kingdom)

[b]Prof. Dr. T. Clark

Computer-Chemie-Centrum and Interdisciplinary Center for Molecular Materials

Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen (Germany)

Fax: (+49) 9131 8526565

E-mail: tim.clark@port.ac.uk

Supporting Information

Supplementary Tables

Table S 1. Population and conformational characteristics of variants

Absolute populations refer to 11 µs (100%) and correspond to the absolute lifetime of each state during the whole simulation. The conformational characteristic of each ring state is given by its turns and hydrogen bonds (Hbonds).

T16 State	T10 State	State population abs [ns]	abs [%]	Conformational characteristics β-turn type[a]	Turn center[b]	H bonds
			intermediate saddle			
			(trajectory time-windows 1455-1466 and 1733-1759 ns)			
			hybrid conformation of T10_states 1 (*saddle*) and 8 (*open*)			
11	**3**	38	0.3	(VIII)	(3,4)	no Hbonds
			variant I			
			(trajectory time-windows 5930-6017, 6090-6218, 8713-8804, 10209-10254, 10306-10320 ns			
31,32,33	**10**	365	3.3	III' (or I')	3,4	2O5NH
			variant II			
			(trajectory time-windows 6017-6048, 10254-10306 ns)			
34	**11**	88	0.8	III' (or I')	3,4/2,3	2O5NH/1O4NH
			variant III			
			(trajectory time-window 5900-5930 ns)			
			possibly a hybrid conformation of T10_states 4 (*clinched open*) and 6 (*twisted saddle*)			
35	**12**	30	0.3	(II)	2,3	no Hbonds
Σtotal		521	4.7			

Abbreviations: abs = absolute; O = carbonyl oxygen; NH = amide hydrogen; Hbond = hydrogen bond; T16 = overall states defined by Φ/Ψ 2 to 9; T10 = ring states defined by Φ/Ψ 2 to 6. [a] parenthesis indicate distorted turns; [b] residues (i+1)/(i+2)

Supporting Information

Table S 2. DASH state trajectory of ring, overall and tail states

Ring (T10) and overall states (T16) are listed in their sequential order during the 11µs MD simulation with their individual time-windows (bout lengths) and transition points (TS). States are grouped in correspondence to their common ring conformation. Conformational characteristics are given in terms of turn centers, β-turn types and hydrogen bonds. Parenthesis indicate either distorted ideal turns or only low populated hydrogen bond (<< 50% on 11 µs MD). Tail states (T6) are not listed explicitly, but the percentage population of each possible tail conformation is given relative to main ring conformations (*open*, *saddle*, *clinched saddle*, and *twisted saddle*).

Bout [ns]	TS [ns]	max occ	Turn	Hbond	T10 states	T16 states	T6[a] states
				open (0-1455 ns)			
566	566	*	(2,3 dist)	(2O4H)	8	27,28,29	
11	577				9	30	
311	887				8	27,28	
27	914				9	30	93.6 % extended
227	1140				8		4.9 % 7,8-b-turn type II
10	1150				9	27,28,29	1.5 % distorted turn
227	1377				8		
44	1420	*	(2,3 dist)	(2O4H)	9	30	
35	1455				8	27	
				saddle (1455-5900 ns)			
11	1466				3	3,1	
267	1733				1		
27	1759	*	(3,4 VIII)	no Hbonds	3	11	
717	2476				1	1,3,4,5,6	
72	2548				2	9,10	
111	2659				1	5,6	76.3 % extended
75	2733				2	7	19.2 % 7,8-b-turn type II
1507	4240	*	3,4 I / 4,5 I	2O5H / 2O6H	1	1,2,3,4,5,6	2.7 % distorted turn
37	4276				2	7	1.8 % 7,8-b-turn type I
422	4698				1	1,2,3,4	
90	4788				2	7,8	
483	5271				1	3,4,5,6	
178	5448	*	3,4 I / 4,5 I	2O5H / 2O6H	2	7,8	
452	5900				1	2,3,4	
				variants (5900-6429 ns)			
30	5930	*	2,3 (II)		12	35	
87	6017				10	31,32	
32	6048				11	34	
42	6090				6	22	
128	6218	*	3.4 III' (or I')		10	31,32,33	
160	6378				6	19	
51	6429				7	24,25	
				clinched open (6429-7187 ns)			
71	6499				4	12	
74	6573	*	4,5 I	(3O6H)	5	15,16	
117	6689				4	14,13	68.3 % extended
42	6731				5	16	26.1 % 7,8-b-turn type II
70	6801				4	14	5.5 % distorted turn
47	6848				5	18	

Supporting Information

Bout [ns]	TS [ns]	max occ	Turn	Hbond	T10 states	T16 states	T6[a] states	
257	7105	*	4,5 (VIII)	(3O6H)	4	12,13,14		
47	7151				5	17		
14	7165				4	13		
22	7187				5	22		
				twisted saddle (7187-11000 ns)				
186	7372				6	19.2		
37	7409				7	24		
114	7523				6	19,20		
14	7537				7	24		
12	7548				6	19		
12	7560				7	24		
22	7581				6	19		
190	7771				7	24,25,26		
107	7878				6	19,20		
90	7968				7	24,25		
89	8057				6	19,20		
255	8312				7	23,24,26		
402	8713	*	3,4 II	2O5H	6	19,20,21,22		
91	8804				10	31,32		
26	8830				6	20		
17	8847				7	24		
233	9080				6	19		
14	9094				7	24	83.4 %	extended
12	9106				6	20	15.3 %	7,8-b-turn type II
155	9260				7	24,25	1.2 %	distorted turn
65	9325				6	19	0.3 %	7,8-b-turn type I
26	9350				7	24		
80	9430				6	19,20		
194	9624	*	3,4 II	2O5H	7	24,25,26		
320	9943				6	19,21,22		
39	9982				7	24		
52	10034				6	19		
141	10174				7	24		
35	10209				6	19,20		
45	10254				10	31		
52	10306	*	3,4 III' (or I')		11	34		
14	10320				10	31		
111	10430				6	19,20		
92	10522				7	24,25,26		
109	10631				6	19,20		
120	10751				7	24		
227	10978				6	19,21,22		
23	11001				7	24		

ns= bout length [ns]; TS= transition point [ns] on 11µs MD trajectory; Hbonds= hydrogen bonds. *occurrence with maximum lifetime during 11µs MD. [a] T6 states: 3, 4 = extended, 5, 6 = 7,8 β-turn type II, 2 = 7,8 β-turn type I, 1 = distorted turn

Supporting Information

Table S 3. RMSD differences [Å] between representative ring states (T10)

This table supplements Figure 3 ("Main ring conformations of 8-Arg-vasopressin")

T10 ring states		8	9	1	2	4	5	6	7
		open	*open*	*saddle*	*saddle*	*cl.open*	*cl.open*	*tw.saddle*	*tw.saddle*
open	8	0	0.233	1.004	1.050	1.812	1.833	1.311	1.378
open	9		0	0.938	1.019	1.713	1.806	1.132	1.200
saddle	1			0	0.169	2.193	2.145	0.946	0.833
saddle	2				0	2.234	2.167	1.0744	0.966
cl.open	4					0	0.173	1.661	1.903
cl.open	5						0	1.682	1.920
tw.saddle	6							0	0.251
tw.saddle	7								0

T10 state = representative ring conformation resulting from a DASH analysis of backbone dihedrals Φ/Ψ 2 to 6; cl.open = clinched open; tw.saddle = twisted saddle

Table S 4. Dihedral angles Φ/Ψ and β-turns

Listed are the mean torsion angle ensembles for each representative ring conformation and the corresponding ideal β-type torsions.

T10[a] state	Torsion angle [°] Tyr[2] Φ	Tyr[2] Ψ	Phe[3] Φ	Phe[3] Ψ	Gln[4] Φ	Gln[4] Ψ	Asn[5] Φ	Asn[5] Ψ	Cys[6] Φ	Cys[6] Ψ	Turns center	Turns type
open (13.2%)												
8	-112.54	134.53	55.31	3.41	-135.33	152.15	-75.09	124.68	-127.16	148.39		
StdDev	37.81	18.46	9.08	31.34	23.92	18.2	18.32	32.08	31.88	23.33		
9	-98.98	129.37	56.09	0.76	-135.73	153.49	-66.41	113.78	-55.29	126.93		
StdDev	54.64	26.48	9.27	31.72	23.88	23.1	31.29	80.23	61.52	40.84		
saddle (40.1%)[b]												
1	-80.2	143.87	-62.88	-21.36	-86.73	-7.38	-113.37	-27.13	-126.42	133.12		
StdDev	20.52	12.37	9.44	13.4	17.2	16.94	21.14	22.14	20.16	33.23		
			-60	-30	-90	0					3,4	type I
					-60	-30	-90	0			4,5	(type I)
2	-84.29	147.09	-57.99	-27.01	-85.13	-7.63	-122.03	-6.72	-60.49	142.38		
StdDev	23.05	13.93	10.95	15.59	17.53	16.62	20.55	41.47	32.18	24.51		
			-60	-30	-90	0					3,4	type I
					-60	-30	-90	0			4,5	(type I)
clinched open (6.9%)												
4	-95.37	-19	-101.27	156.57	-67.65	-19.06	-112.46	86.89	-117.42	145.84		
StdDev	28.15	22.75	29.46	14.56	16.85	23.84	28.66	61.3	36.21	21.54		
					-60	-30	-120	120			4,5	type VIII
5	-90.52	-18.35	-116.2	151.18	-68.06	-20.5	-88.17	14.01	-82.72	144.88		
StdDev	28.3	18.64	30.65	13.16	22.02	26.74	20.39	33.03	29.6	16.17		
					-60	-30	-90	0			4,5	type I
twisted saddle (35.1%)												
6	-86.02	162.33	-52.48	127.66	55.04	12.34	-107.29	-7.44	-122.17	144.18		
StdDev	29.44	13.88	16.16	14.69	9.01	21.14	29.86	48.29	28.23	23.53		
			-60	120	80	0					3,4	type II
7	-115.65	174.87	-52.78	129.79	57.39	8.38	-114.1	-16.45	-70.67	148.3		
StdDev	24.26	19.63	19	13.91	8.24	20.56	25.07	29.84	19.33	13.72		
			-60	120	80	0					3,4	type II
intermediate saddle (0.3%)												
3	-68.56	163.21	-72.91	-0.62	-125.2	146.92	26.03	63.46	-105.36	135.68		
StdDev	13.39	18.34	10.4	21.2	24.04	15.76	51.49	53.77	36.37	21.37		
			-60	-30	-120	120					3,4	(type VIII)
variant I (3.3%)												
10	-83.42	142.17	50.15	31.15	53.69	20.44	-102.4	67.7	-109.12	150.49		
StdDev	35.83	15.34	8.86	19.44	16.39	18.56	27.79	60.36	34.19	20.15		
	-90	120	60	30	60	30	-90	60	-120	(120,0)	3,4	type III'
			60	30	90	0					3,4	(type I')
variant II (0.8%)												
11	-72.18	142.55	49.74	32.67	56.77	17.51	-122.61	118.01	-94.72	143.74		
StdDev	25.09	13.26	7.79	11.24	6.91	20.65	26.37	21.41	22.96	21.72		
	-60	120	60	30	60	30	-120	120	-90	120	3,4	type III'
			60	30	90	0					3,4	(type I')
	-60	120	80	0							2,3	(type II)
variant III (0.3%)												
12	-82.3	158.41	60.66	170.49	-76.27	-19.49	-113.78	103.37	-79.86	137.76		
StdDev	24.16	18.87	7.43	21.95	17.98	22.45	29.57	68.16	30.38	15.54		
	-60	120	80	0	-80	0	-120	120	-60	120	2,3	(type II)

Abbreviations: StdDev = standard deviation; [a] T10 state = representative ring conformation resulting from a DASH analysis of backbone dihedrals Φ/Ψ 2 to 6; [b] Absolute population on 11 µs MD;

Supporting Information

Table S 5. Key torsions and torsion differences of main ring conformations of 8-Arg-vasopressin

The main ring conformations are represented by the major DASH ring state each.

	T10 state n	key torsion	*open* Δtors [n]-[8]	key torsion	*saddle* Δtors [n]-[1]	key torsion	*cl.open* Δtors [n]-[4]
saddle	**1**	Φ3	-118.19				
		Ψ4	-159.53				
		Ψ5	-151.81				
cl.open	**4**	Ψ2	-153.53	Ψ2	162.87		
		Φ3	-156.58	Ψ3	177.93		
		Ψ3	153.16	Ψ5	114.02		
		Ψ4	-171.21				
tw.saddle	**6**	Φ3	-107.79	Ψ3	149.02	Ψ2	181.33
		Ψ3	124.25	Φ4	141.77	Φ4	122.69
		Φ4	190.37			Ψ5	-94.33
		Ψ4	-139.81				
		Ψ5	-132.12				

Δtors = torsion difference of key torsions in [°] between interconverting ring conformations, n = state number, T10 = DASH ring state analysis of ring torsions Φ/Ψ 2 to 9, cl.open = clinched open, tw.saddle = twisted saddle

Supplementary Figures

Figure S 1. Conformational variants

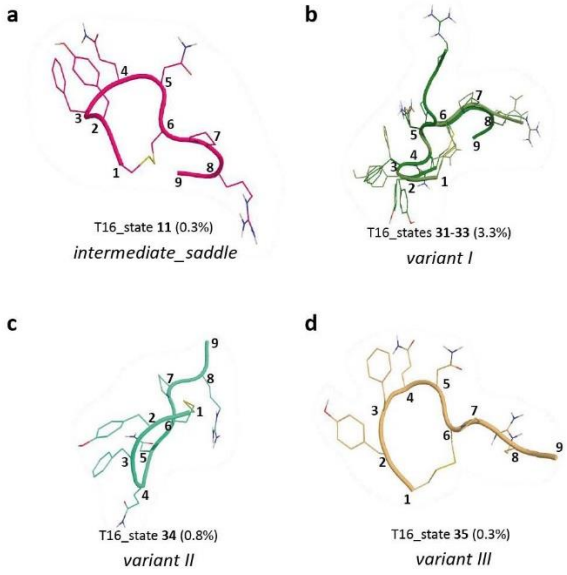


Figure S 1. Conformational variants of 8-Arg-vasopressin during 11µs MD that are not exactly assignable to a main conformational group (open, saddle, clinched open or twisted saddle). The states result from a DASH state analysis of backbone dihedrals Φ/Ψ 2 to 9. Populations are given in parenthesis and refer to 11 µs MD simulation.

Supporting Information

Figure S 2. DASH torsion-angle ensembles for the main ring conformations

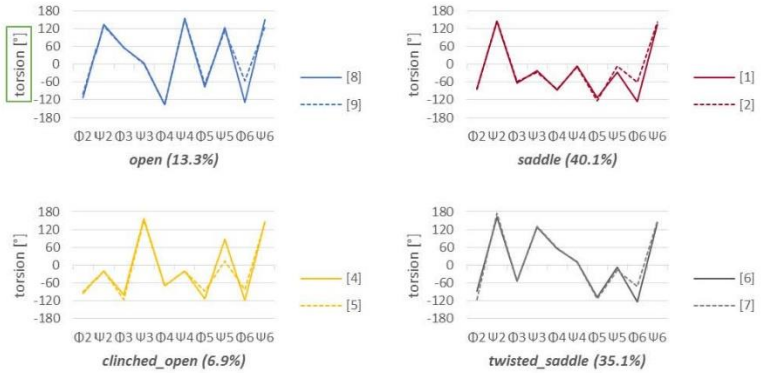


Figure S 2. DASH torsion angle ensembles for the main ring conformations of 8-Arg-vasopressin *open*, *saddle*, *clinched open* and *twisted saddle*. Each conformational group is represented by two ring states resulting from a DASH analysis of ring torsions Φ/Ψ 2 to 6 (T10). The minor state is depicted as dashed line.

Figure S 3. DASH torsion-angle ensembles for all tail conformations

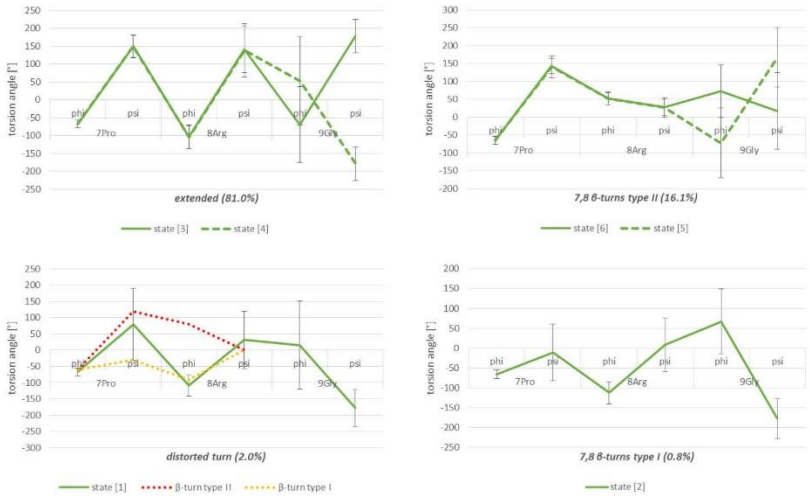


Figure S 3. DASH torsion angle ensembles for all tail conformations of 8-Arg-vasopressin on 11 µs MD. Extended and β-turn type II tail conformations are represented by two ring states each, distorted and type I by one single state. The states resulting from a DASH analysis of ring torsions Φ/Ψ 7 to 8. Minor states are depicted as dashed lines. In addition the i+1, i+2 torsions of an ideal β-turn type II (red) and type I (yellow) are given as dotted lines in the diagram of the distorted turn conformation.

Supporting Information

Supplementary Information: Methods

DASH state analysis

To classify the relevant conformational states, the data density of the 11 µs MD trajectory was reduced to 22,000 snapshots (2 snapshot/ ns). To determine the overall states, torsion angles phipsi 2 to 9 were analysed. Referring to the total number of 16 torsions analysed, this analysis setup was called T16. The default bout length (time-window) of 20 frames was chosen meaning that a torsion angle ensemble has to persist/exist a minimum time of here 10 ns on the 11µs MD trajectory to be considered as representative conformational state. The DASH analysis was run within AmberDASH, an interface that extracts torsions angles from AMBER netcdf trajectories followed by a DASH analysis and a final extraction of representatives for each state in PDB format. Representative ring conformations were determined by reducing the 16 overall torsions to the 10 ring torsions phipsi 2 to 6 using the same parameters as for the T16 analysis. This setup was denoted T10 referring to the number of analyzed torsions. Finally a separate analysis of 6 tail torsions phipsi 7 to 9 was made (T6), again with consisten parameters. As representative of a the DASH state, the frame/snapshot with the highest similarity to a given DASH state mean angle ensemble was chosen and output as PDB structure.

Ptraj

Trajectories of conformational data like root mean square deviations (RMSD) and radii of gyration (RadGyr) were calculated using Ptraj, the analysis tools of AMBER tools[31] within the AMBER program package[30].

Hbond-analysis

Hydrogen bonding interactions between all backbone amide N and H atoms , and all carbonyl O atoms were calculated via Ptraj Hbond analysis. This analysis measures distances and angles between triplets of atoms [31] and calculates the procentual occurance of hydrogen bonds over a considered simulation period. Here, a hydrogen bond is defined by a maximum OH distance of 3.5 Å and a O..H..N cutoff angle of 120°. A total input data sets was created taking every 100th frame of the 11 µs trajectory. Trajectory time-windows referring to representative conformational ring states were analyzed separately.

Secondary structure analysis

The secondary structure was calculated for the atom selection backbone Cα-atoms 1-9 using the DSSP method (define secondary structure of proteins) of Kabsch & Sander[43] via Ptraj. Every trajectory time-window referring to a conformational ring state was analyzed separately taking every 100th frame.

Supporting Information

Animated Multimedia Application

Video S 1. 11 µs Molecular Dynamics of 8-Arg-Vasopressin in Water at 300K

The simulation starts with the open 1YF4 conformation. Significant transitions of ring conformations are at 1.4 µs (*open/saddle*), 5.9 µs (*saddle/variants*), 6.4 µs (*variants/clinched open*), and 7.1 µs (*clinched open/twisted saddle*). Depiction: backbone Cα 1-9 = cartoon; side chains = lines; $Tyr^2O/Phe^3O/Asn^5NH$ = spheres (O red, H white); water molecules are not shown.

File: AVP_VideoClip.MP4

Interactive 3D-plots of Principle Components

Interactive 3D-plots of the first four PCs are given as HTML-pages in the Supporting Material. Simply download and unzip the folder below and doubleclick the html-files to open in your standard browser.

Zip-Folder: torsions_3D_PCA_Plots.zip

DASH States (PDB)

T16 overall states

Representative overall states of 8-Arg-Vasopressin

Zip-Folder: T16_states.zip

T10 ring states

Representative ring states of 8-Arg-Vasopressin

Zip-Folder: T10_states.zip

T6 tail states

Representative ring states of 8-Arg-Vasopressin

Zip-Folder: T10_states.zip

A 2: Reprint Supporting Information Paper 2

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00344.

Supporting Information

Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine-Vasopressin

*Elke Haensele,[‡,†] Noureldin Saleh,[‡] Christopher M. Read,[§] Lee Banting,[†] David C. Whitley,[†] and Timothy Clark.[‡,]**

[‡] Computer-Chemie-Centrum der Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany.

[†] School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth PO1 2DT, United Kingdom.

[§] School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DY, United Kingdom.

e-mail: Tim.Clark@fau.de

Table of Contents

Supporting Information

Computational details

Long-scale molecular dynamics simulation

Table S1 Parameters for the unrestrained long-scale molecular-dynamics simulation of Arg[8]-vasopressin

Force field	ff99SB[1] (Amber 10, Amber 14 CUDA)[2]
Initial conformation	open (PDB ID: 1YF4);[3] neutralized with 2 Cl^- (ions08.lib, frcmod.ionsjc_tip4pew[4, 5])
Solvation	explicit, TIP4PEw,[6] truncated octahedral box
Temperature and Pressure	T= 300K, p= 1 bar (Berendsen coupling,[7] 1.0 ps external heath bath)
Minimization	8,945 steps: 500 steps steepest-descent followed by conjugate-gradient method
Molecular dynamics	2 fs time steps, SHAKE algorithm,[8] 8.0 Å non-bonded cut-off, Particle Mesh Ewald method,[9] periodic boundary conditions
Simulation time	23,000 ns

Representative conformations of Arg[8]-vasopressin

Analysis of the conformational space of 11 µs MD simulation revealed 4 main ring conformations for AVP: *open, saddle, clinched open, twisted saddle*. A detailed description of the conformations has been published previously.[10]

The most populated overall state of each cluster (ring-state type, main ring conformation) was chosen as representative for further DFT and NMR calculations; *saddle* and *clinched open* each of with *extended* and *folded* tail. The mean torsions of the representatives are given in Table S2.

Table S2 Mean backbone torsions and standard deviations (±) of representative conformationas of Arg[8]-vasopressin

		$saddle_{ext}$		$saddle_{fold}$		$cl.open_{ext}$		$cl.open_{fold}$		*tw.saddle*		*open*	
		state3[a]		state6		state12		state14		state19		state27	
Tyr^2	phi	-80.9	±21.4	-78.5	±18.5	-101.2	±28.8	-86.8	±25.4	-84.6	±29.0	-112.0	±37.9
	psi	144.2	±12.8	143.2	±11.6	-16.4	±23.3	-22.9	±21.1	162.1	±13.1	134.6	±18.8
Phe^3	phi	-63.0	±9.5	-62.7	±9.1	-99.9	±32.8	-106.6	±26.0	-52.0	±16.8	54.9	±12.0
	psi	-20.7	±13.6	-22.5	±12.5	157.4	±13.5	154.5	±16.0	127.3	±15.1	4.4	±31.3
Gln^4	phi	-86.5	±17.4	-89.2	±16.3	-66.4	±18.0	-69.4	±11.3	55.0	±8.4	-135.9	±23.8
	psi	-7.6	±17.8	-4.7	±16.5	-18.2	±25.7	-21.8	±18.7	12.4	±21.1	151.6	±19.2
Asn^5	phi	-113.3	±21.2	-113.8	±21.0	-112.1	±27.8	-109.9	±30.4	-106.1	±29.4	-75.4	±19.8
	psi	-27.4	±21.8	-26.4	±23.3	74.5	±62.7	92.8	±57.7	-8.7	±47.3	124.7	±31.9
Cys^6	phi	-126.3	±20.0	-126.6	±21.0	-111.1	±34.9	-117.3	±38.9	-120.8	±28.3	-128.8	±30.4
	psi	132.7	±34.5	139.8	±23.1	145.9	±19.2	145.6	±20.1	143.8	±24.2	149.0	±22.7
Pro^7	phi	-67.8	±11.3	-64.6	±11.1	-66.3	±11.1	-64.8	±11.5	-67.7	±11.2	-67.5	±11.5
	psi	-63.0	±11.3	-62.7	±11.3	-99.9	±32.8	-106.6	±26.0	-52.0	±16.8	53.3	±12.0
Arg^8	phi	-20.8	±13.8	150.3	±12.5	150.3	±16.0	150.3	±16.0	127.3	±15.1	4.4	±31.3
	psi	-86.5	±17.4	-89.2	±16.3	-69.4	±18.0	-69.4	±11.3	55.0	±8.4	-135.9	±23.8

[a]PDB files of the representative states are available under these designations from the SI (ESM4) of our previous publication (http://dx.doi.org/10.1007/s00894-014-2485-0); Abbreviations: clinched (cl.); twisted (tw.); ext (extended); fold (folded)

Supporting Information

After extending the MD-simulation to 23 µs, the conformational space was again clustered using DASH[11]. The resulting representatives were assigned to the former representatives already defined for the first 11 µs of the MD simulation via circular similarity of backbone torsions in order to ensure consistency. From the 23 µs MD simulation the representatives of the five most populated ring state types were chosen in order to calculate their free energies and populations. The first four clusters were the already identified main conformations from the 11 µs MD; the additional conformation (*variants*) was seen for the first time after 11 µs. *Variants* is a *clinched open* variant; a rotamer of the peptide-bond between residue Gln^4 and Asn^5 of the *clinched open* conformation (Figure S1). *Variants* was added to the selection, although low populated in the simulation, because it occurred only once and at the end of the 23 µs simulation with no following interconversion. Thus, it could not be ruled out that this cluster might be another main cluster of AVP. However, the thermodynamic calculations (metadynamics) showed that the conformation is the least stable (*cf.* main text) and it was not considered further for DFT-NMR-calculations.

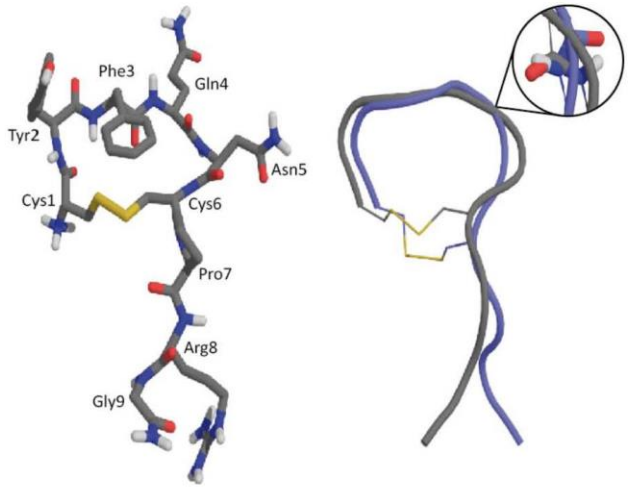


Figure S1 Representative for the ring-state type *variants* (grey). Left: stick depiction of *variants*. Right: cartoon depiction of *variants and* ring alignment with the representative for *clinched open* (blue).
Variants is a 4,5 peptide-bond rotamer of *clinched open* as illustrated in the zoom. The conformation *variants* occurred for the first time at the end of 23 µs MD simulation of Arg^8-vasopressin and is expected to be populated insignificantly in aqueous solution according to thermodynamic calculations.

Cartesian coordinates of the B3LYP/6-31G(d) optimized geometries are given below as Gaussian Archive Entries.

Supporting Information

Metadynamics simulations

We used a combination of metadynamics,[12, 13] in its well-tempered variant (WT),[14] the multiple-walker technique[15] and the path collective variable (PCV)[16, 17] to determine the free-energy differences between AVP conformations (identified from the 23 μs trajectory) in water. Four main ring conformers identified previously[10] and one new found in the extended 23 μs MD simulation were used as starting geometries for the metadynamics simulations. The PCV used was a numerical assignment to the most similar AVP conformer based on the RMSD of the backbone atoms of the ring residues. An analysis of the unbiased 23 μs MD trajectory using this PCV showed that 90% of the frames can be uniquely assigned to one of the five ring conformers. We were therefore able to use this single PCV for the metadynamics simulation.

The simulation boxes and topologies used for the AMBER MD simulations were converted to GROMACS[18] format. The simulation configuration used the same water model, temperature and thermodynamic ensemble as the reference unbiased simulation. Particle mesh Ewald (PME) was used to treat electrostatic interactions, using a cut-off distance of 1.0 nm and each of the five models were equilibrated for 20 ns. Gaussian hills with an initial height of 0.6 kcal mol^{-1} applied every ps were used. The Gaussian functions were rescaled in the WT scheme using a bias factor of 10. The metadynamics simulations were performed using GROMACS with the PLUMED plug-in.[19]

Supporting Information

Calculation of chemical shifts and correlation with experimental values

Representative structures were taken from the 23 µs MD simulation and optimized using the B3LYP hybrid density functional,[20] the 6-31G(d) basis set[21] and the default polarizable continuum model (PCM) for water solvent using Gaussian09.[22] These optimized structures were then used to calculate ^{1}H and ^{13}C chemical shifts using the GIAO formalism,[23] again with PCM-water. The chemical shifts were obtained from the calculated magnetic shielding using the correlation formulae given in the main text. These formulae were obtained by linear regression of a training set of experimental chemical shifts[24] and the corresponding shielding values at level B3LYP/6-31G* with PCM-water (Figure S2).

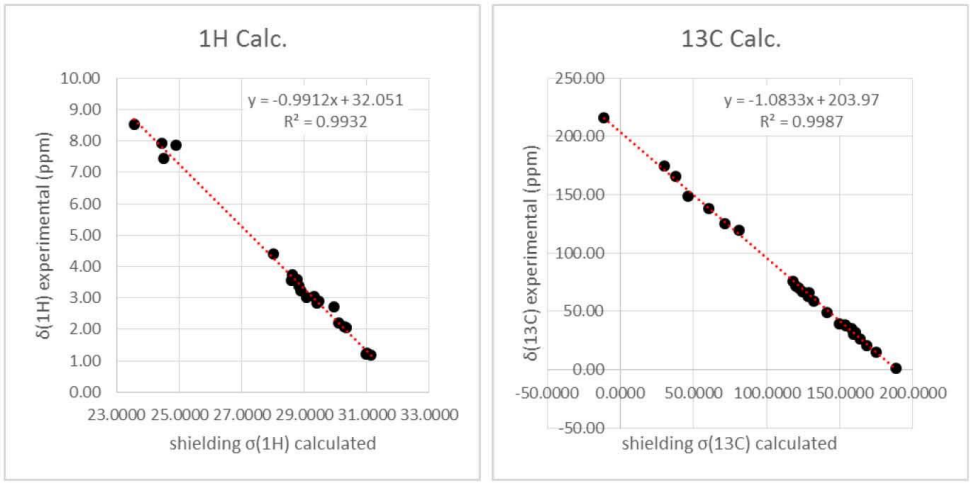


Figure S2 Linear regression of magnetic shielding at level B3LYP/6-31G(d) with PCM water and a training set of experimental chemical shifts (ppm) for ^{1}H and ^{13}C

Table S3-4 show the calculated ^{13}C and ^{1}H chemical shifts (δ ppm) for the individual conformations and the equilibrium mixtures (Eq. 1-4) calculated from the metadynamics free-energy differences, both assuming a single (*extended*) tail conformation and using the *extended:folded* equilibrium determined from the 23 µs MD simulation for each representative conformation.

$$\delta_{saddle_{eq.}} = 0.7314 \times \delta_{saddle_{ext}} + 0.2686 \times \delta_{saddle_{fold}} \quad (1)$$

$$\delta_{cl.open_{eq.}} = 0.6263 \times \delta_{cl.open_{ext}} + 0.3737 \times \delta_{cl.open_{fold}} \delta_{open} \quad (2)$$

$$\delta_{equilibrium_{ext}} = 0.6865 \times \delta_{saddle_{ext}} + 0.2951 \times \delta_{cl.open_{ext}} + 0.0043 \times \delta_{tw.saddle} + 0.0141 \times \delta_{open} \quad (3)$$

$$\delta_{equilibrium_{eq.}} = 0.6865 \times (0.7314 \times \delta_{saddle_{ext}} + 0.2686 \times \delta_{saddle_{fold}}) + 0.2951 \times (0.6263 \times \delta_{cl.open_{ext}} + 0.3737 \times \delta_{cl.open_{fold}}) + 0.0043 \times \delta_{tw.saddle} + 0.0141 \times \delta_{open} \quad (4)$$

Supporting Information

Table S3 Calculated ^{13}C chemical shifts (B3LYP/6-31G(d), PCM water) of Arg^{8}-vasopressin

		Calculated ^{13}C chemical shifts (ppm)									
		Individual conformations								Metadynamics equilibrium	
			Saddle			*Clinched open*		*Twisted saddle*	*Open*	Equilibrium	**Equilibrium**
res	atom	*ext*	*fold*	*equil*[a]	*ext*	*fold*	*equil*[b]	*ext*	*ext*	*ext*[c]	***equilibrium***[d]
Cys^{1}	C^{α}	57.90	57.27	57.73	56.81	58.08	57.28	57.23	58.08	57.58	57.60
Cys^{1}	C^{β}	43.53	43.38	43.49	49.46	49.98	49.66	46.52	45.04	45.32	45.34
Tyr^{2}	C^{α}	63.12	63.82	63.30	65.73	65.80	65.76	58.02	63.64	63.87	64.01
Tyr^{2}	C^{β}	37.72	39.58	38.22	40.42	40.46	40.44	39.97	42.43	38.59	38.94
Tyr^{2}	$C^{\delta1}$	133.17	131.80	132.80	134.21	134.62	134.37	135.89	134.23	133.51	133.30
Tyr^{2}	$C^{\delta2}$	131.08	133.13	131.63	133.02	132.52	132.83	133.95	133.37	131.70	132.02
Tyr^{2}	$C^{\varepsilon1}$	118.14	117.08	117.85	115.46	114.56	115.13	115.79	115.83	117.31	117.01
Tyr^{2}	$C^{\varepsilon2}$	118.09	116.37	117.63	115.90	115.61	115.79	117.39	116.41	117.42	117.07
Phe^{3}	C^{α}	62.93	62.34	62.77	54.22	56.81	55.19	62.29	62.98	60.36	60.54
Phe^{3}	C^{β}	37.98	38.02	37.99	41.37	44.82	42.66	40.51	39.87	39.02	39.40
Phe^{3}	$C^{\delta1}$	131.31	131.04	131.23	134.45	133.75	134.19	131.16	130.93	132.23	132.10
Phe^{3}	$C^{\delta2}$	131.63	131.33	131.55	133.42	131.86	132.84	133.24	134.32	132.20	131.98
Phe^{3}	$C^{\varepsilon1}$	131.22	131.40	131.27	130.95	131.08	131.00	131.95	130.27	131.13	131.18
Phe^{3}	$C^{\varepsilon2}$	131.67	131.34	131.58	130.01	131.21	130.46	131.21	129.58	131.15	131.22
Phe^{3}	C^{ζ}	129.93	129.83	129.90	129.57	128.92	129.32	128.48	128.24	129.79	129.70
Gln^{4}	C^{α}	57.93	56.76	57.61	61.15	58.78	60.26	61.77	55.39	58.86	58.38
Gln^{4}	C^{β}	30.23	27.62	29.53	28.74	27.92	28.44	25.60	40.73	29.92	29.35
Gln^{4}	C^{γ}	34.65	29.18	33.18	29.41	29.45	29.42	33.47	33.86	33.09	32.08
Asn^{5}	C^{α}	58.15	56.35	57.67	53.20	52.10	52.79	53.70	53.33	56.61	56.15
Asn^{5}	C^{β}	41.48	43.06	41.90	36.42	35.29	36.00	42.79	40.79	39.98	40.15
Cys^{6}	C^{α}	55.48	54.06	55.10	53.99	55.27	54.47	55.96	52.63	55.00	54.88
Cys^{6}	C^{β}	45.80	49.45	46.78	50.39	48.02	49.50	46.15	52.71	47.25	47.67
Pro^{7}	C^{α}	66.11	66.08	66.10	66.59	66.73	66.64	65.76	66.07	66.25	66.26
Pro^{7}	C^{β}	31.87	32.15	31.95	31.93	32.19	32.03	31.61	31.67	31.88	31.96
Pro^{7}	C^{γ}	26.37	27.70	26.73	27.51	27.54	27.52	26.92	26.81	26.72	26.97
Pro^{7}	C^{δ}	51.28	51.17	51.25	50.72	51.23	50.91	49.90	50.72	51.10	51.13
Arg^{8}	C^{α}	56.93	58.60	57.38	56.10	58.71	57.08	56.21	55.59	56.66	57.26
Arg^{8}	C^{β}	36.06	32.65	35.14	36.20	33.57	35.22	36.46	35.94	36.10	35.18
Arg^{8}	C^{γ}	30.23	26.10	29.12	28.70	27.17	28.13	27.37	28.20	29.74	28.81
Arg^{8}	C^{δ}	44.58	44.90	44.66	46.36	45.84	46.16	44.45	46.34	45.13	45.13
Gly^{9}	C^{α}	42.47	44.61	43.04	42.59	46.31	43.98	42.80	42.54	42.51	43.31

[a]Eq. 1; [b]Eq. 2; [c]Eq. 3; [d]Eq. 4; Abbreviations: extended (ext); folded (fold); equilibrium (equil)

Supporting Information

Table S4 Calculated ^{1}H chemical shifts (B3LYP/6-31G(d), PCM water) of Arg^{8}-vasopressin

		Calculated ^{1}H chemical shifts (ppm)									
		Individual conformations								Metadynamics equilibrium	
			Saddle		*Clinched open*			*Twisted saddle*	*Open*	Equilibrium	**Equilibrium**
res	atom	*ext*	*fold*	*equil*[a]	*ext*	*fold*	*equil*[b]	*ext*	*ext*	*extendedt*[c]	***equilibrium***[d]
Cys^{1}	H^{α}	3.87	3.91	3.88	3.63	4.26	3.86	4.40	4.10	3.80	3.88
Cys^{1}	$H^{\beta a}$	3.39	3.69	3.47	3.37	2.65	3.10	2.84	3.10	3.38	3.35
Cys^{1}	$H^{\beta b}$	3.05	2.74	2.97	3.32	4.20	3.65	3.49	2.85	3.13	3.17
Tyr^{2}	H^{α}	4.52	4.26	4.45	3.95	3.81	3.90	4.87	4.40	4.35	4.29
Tyr^{2}	$H^{\beta a}$	3.14	3.20	3.16	3.11	2.87	3.02	3.12	3.12	3.13	3.12
Tyr^{2}	$H^{\beta b}$	3.99	2.75	3.66	2.95	2.64	2.84	2.75	2.98	3.66	3.40
Tyr^{2}	$H^{\delta 1}$	7.37	7.24	7.33	7.33	5.73	6.73	7.48	7.18	7.36	7.15
Tyr^{2}	$H^{\delta 2}$	7.25	6.95	7.17	7.25	7.09	7.19	7.54	7.37	7.25	7.18
Tyr^{2}	$H^{\varepsilon 1}$	7.10	6.41	6.91	6.72	6.32	6.57	6.74	6.71	6.98	6.81
Tyr^{2}	$H^{\varepsilon 2}$	6.52	6.42	6.49	6.56	6.49	6.54	6.94	6.73	6.54	6.51
Phe^{3}	H^{α}	4.64	5.02	4.75	4.48	4.71	4.57	4.03	3.52	4.58	4.67
Phe^{3}	$H^{\beta a}$	3.68	3.71	3.69	3.12	2.62	2.94	2.72	2.47	3.49	3.44
Phe^{3}	$H^{\beta b}$	2.73	2.74	2.73	3.19	3.20	3.19	2.63	3.90	2.88	2.88
Phe^{3}	$H^{\delta 1}$	7.08	7.10	7.09	7.15	7.41	7.25	7.49	7.10	7.11	7.14
Phe^{3}	$H^{\delta 2}$	7.07	7.17	7.09	7.57	7.56	7.57	7.26	7.61	7.23	7.24
Phe^{3}	$H^{\varepsilon 1}$	7.37	7.42	7.38	7.66	7.56	7.63	7.56	7.34	7.46	7.45
Phe^{3}	$H^{\varepsilon 2}$	7.43	7.47	7.44	7.52	7.60	7.55	7.51	7.30	7.46	7.47
Phe^{3}	H^{ζ}	7.39	7.40	7.39	7.48	7.64	7.54	7.44	7.24	7.41	7.43
Gln^{4}	H^{α}	4.45	4.79	4.54	4.12	4.27	4.17	3.28	4.26	4.34	4.42
Gln^{4}	$H^{\beta a}$	2.05	1.38	1.87	2.60	1.55	2.21	1.76	1.25	2.20	1.96
Gln^{4}	$H^{\beta b}$	2.26	2.52	2.33	1.84	2.40	2.05	2.37	2.37	2.14	2.25
Gln^{4}	$H^{\gamma a}$	2.00	1.81	1.95	2.12	2.31	2.19	1.51	2.19	2.03	2.02
Gln^{4}	$H^{\gamma b}$	2.25	2.20	2.23	2.23	2.38	2.29	0.50	2.66	2.24	2.25
Asn^{5}	H^{α}	4.73	4.93	4.78	4.71	4.61	4.68	4.97	5.10	4.73	4.76
Asn^{5}	$H^{\beta a}$	2.75	2.36	2.64	2.50	2.32	2.44	2.71	2.63	2.67	2.58
Asn^{5}	$H^{\beta b}$	2.63	2.84	2.68	3.31	2.84	3.14	2.39	2.50	2.83	2.81
Cys^{6}	H^{α}	5.09	5.05	5.08	5.03	4.72	4.91	4.66	4.43	5.06	5.02
Cys^{6}	$H^{\beta a}$	2.78	2.48	2.70	3.18	2.96	3.10	3.23	2.86	2.90	2.82
Cys^{6}	$H^{\beta b}$	2.97	3.19	3.03	3.41	3.41	3.41	3.72	3.34	3.11	3.15
Pro^{7}	H^{α}	4.16	4.07	4.14	4.20	4.00	4.12	4.19	4.16	4.17	4.13
Pro^{7}	$H^{\beta a}$	1.95	2.12	1.99	2.05	2.16	2.09	1.93	2.03	1.98	2.02
Pro^{7}	$H^{\beta b}$	2.23	2.37	2.27	2.37	2.36	2.37	2.23	2.34	2.27	2.30
Pro^{7}	$H^{\gamma a}$	2.26	2.47	2.32	2.39	2.50	2.43	2.25	2.17	2.30	2.35
Pro^{7}	$H^{\gamma b}$	2.03	2.09	2.05	2.11	2.12	2.11	1.99	2.01	2.06	2.07
Pro^{7}	$H^{\delta a}$	3.95	3.80	3.91	3.90	4.10	3.97	3.66	3.68	3.93	3.92
Pro^{7}	$H^{\delta b}$	3.79	3.69	3.76	3.74	4.14	3.89	3.54	3.60	3.77	3.80
Arg^{8}	H^{α}	4.31	3.86	4.19	4.33	3.80	4.13	4.41	4.32	4.32	4.18
Arg^{8}	$H^{\beta a}$	1.88	2.38	2.02	1.46	2.29	1.77	1.68	1.39	1.75	1.93
Arg^{8}	$H^{\beta b}$	2.51	1.71	2.29	2.21	1.98	2.13	2.05	2.20	2.41	2.24
Arg^{8}	$H^{\gamma a}$	1.86	2.00	1.90	1.80	1.97	1.86	1.98	1.84	1.84	1.89
Arg^{8}	$H^{\gamma b}$	2.19	2.28	2.22	2.06	2.23	2.12	2.19	1.83	2.15	2.18
Arg^{8}	$H^{\delta a}$	3.38	3.17	3.32	3.40	3.55	3.45	3.16	3.28	3.38	3.36

Supporting Information

Arg^{8}	$H^{\delta b}$	3.27	3.09	3.22	3.33	3.18	3.27	3.32	3.40	3.29	3.24
Gly^{9}	$H^{\alpha a}$	4.37	4.57	4.43	4.43	3.28	4.00	4.38	4.39	4.39	4.30
Gly^{9}	$H^{\alpha b}$	3.26	4.34	3.55	3.29	3.81	3.48	3.27	3.24	3.27	3.53
Tyr^{2}	H^{N}	6.03	5.52	5.90	5.21	5.48	5.31	8.92	6.62	5.81	5.75
Phe^{3}	H^{N}	4.93	5.17	4.99	5.40	6.28	5.73	4.90	5.90	5.08	5.22
Gln^{4}	H^{N}	4.97	4.62	4.88	5.26	5.30	5.28	5.59	6.03	5.08	5.02
Gln^{4}	$H^{\varepsilon 21}$	4.52	4.45	4.50	4.39	4.41	4.39	4.25	4.44	4.48	4.47
Gln^{4}	$H^{\varepsilon 22}$	5.18	4.82	5.08	4.92	4.81	4.88	4.48	6.34	5.12	5.04
Asn^{5}	H^{N}	6.49	6.97	6.62	6.39	5.58	6.09	6.40	5.53	6.44	6.44
Asn^{5}	$H^{\delta 21}$	4.54	4.40	4.50	4.50	4.40	4.47	4.86	5.03	4.54	4.50
Asn^{5}	$H^{\delta 22}$	5.20	5.04	5.16	4.98	4.86	4.93	5.17	5.39	5.14	5.10
Cys^{6}	H^{N}	6.98	7.32	7.07	5.98	6.36	6.12	6.85	6.12	6.67	6.78
Arg^{8}	H^{N}	5.75	6.01	5.82	5.69	5.92	5.78	5.80	5.62	5.73	5.81
Arg^{8}	H^{ε}	3.83	4.07	3.90	4.62	4.00	4.39	4.39	4.67	4.08	4.05
Gly^{9}	H^{N}	4.88	6.60	5.35	4.99	7.12	5.79	4.97	4.95	4.92	5.47
$(Gly^{9})NH_2$	H^{N1}	4.76	7.23	5.42	4.77	6.67	5.48	4.76	4.78	4.76	5.43
$(Gly^{9})NH_2$	H^{N2}	4.43	4.88	4.55	4.43	4.29	4.38	4.41	4.43	4.43	4.50

[a]Eq. 1; [b]Eq. 2; [c]Eq. 3; [d]Eq. 4; Abbreviations: extended (ext); folded (fold); equilibrium (equil)

Calculation of interatomic distances and correlation with experimental values

To calculate interatomic distances, the longest sections of the 23 µs MD trajectory of Arg^{8}-vasopressin that were occupied entirely by a distinct ring state were chosen (278 ns $saddle_{ext}$; 212 ns $saddle_{fold}$, 136 ns *clinched* $open_{ext}$; 67 ns *clinched* $open_{fold}$, 191 ns *twisted saddle* and 220 ns *open*). The individual distance-trajectories corresponding to the experimental NOE distances were extracted from each representative MD-section. The equilibrium distances are calculated as the 1/6 power means of the distances within each state, weighted according to the distribution given by the metadynamics simulations (Eq. 5-8).

$$r_{equilibrium_{ext}} = (0.6865 \times r_{saddle_{ext}}^{1/6} + 0.2951 \times r_{cl.open_{ext}}^{1/6} + 0.0043 \times r_{tw.saddle}^{1/6} + 0.0141 \times r_{open}^{1/6})^{6} \quad (5)$$

$$r_{equilibrium_{eq.}} = \left[0.6865 \times (0.7314 \times r_{saddle_{ext}}^{1/6} + 0.2686 \times r_{saddle_{fold}}^{1/6}) + 0.2951 \times (0.6263 \times r_{cl.open_{ext}}^{1/6} + 0.3737 \times r_{cl.open_{fold}}^{1/6}) + 0.0043 \times r_{tw.saddle}^{1/6} + 0.0141 \times r_{open}^{1/6}\right]^{6} \quad (6)$$

The distances for the main states (*saddle* and *clinched open*) were refined by taking the relative populations of *extended* and *folded* tail conformations (Eq. 5-6) into consideration.

$$r_{saddle_{eq.}} = (0.7314 \times r_{saddle_{ext}}^{1/6} + 0.2686 \times r_{saddle_{fold}}^{1/6})^{6} \quad (7)$$

$$r_{cl.open_{eq.}} = (0.6263 \times r_{cl.open_{ext}}^{1/6} + 0.3737 \times r_{cl.open_{fold}}^{1/6})^{6} \quad (8)$$

Supporting Information

The results of the statistical analysis of the correlation between calculated and experimental distances are given in the main text. Figure S3 shows the plot of calculated vs. experimental interatomic distances at pH 6.0 and pH 4.7 and calculated distances are listed in Table S5. Mean unsigned errors (MUE) and root mean square deviations (RMSD) are the same order of magnitude as the experimental error limits (pH 6.0 ±0.5 Å, pH 4.7 ±0.7 Å) for all individual conformations and equilibrium distances calculated from metadynamics. The results are discussed in the main text. At pH 6.0, the number of experimental distances is decreased due to proton exchange. As long as only a small number of experimental distances are available and if the experimental error limits are relatively higher than the differences between representative conformations, the linear regression remains insignificant.

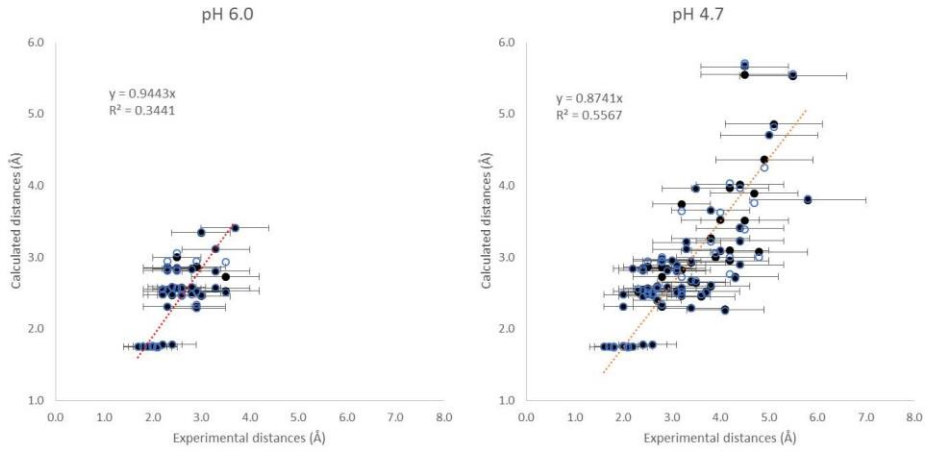


Figure S3 Linear regression of calculated equilibrium and experimental NOE distances at pH 6.0 (left) and pH 4.7 (right). Open blue circles indicate the equilibrium conformations with extended tail. The arrow bars show the error limits of the experimental NOE constraints.

Supporting Information

Table S5 Calculated (r^{-6} weighted) interatomic distances of conformations of Arg^{8}-vasopressin in solution (TIP4PEw water)

				Interatomic distances (Å)					
				Individual conformations				Metadynamics equilibrium	
res	atom	res	atom	*Saddle*[a] *equilibrium*	*Clinched open*[b] *equilibrium*	*Twisted saddle extended*	*Open extended*	Equilibrium[c] *extended*	**Equilibrium[d]** ***equilibrium***
Cys^{1}	H^{α}	Cys^{1}	$H^{\beta a}$	2.49	2.66	2.61	2.98	2.53	2.55
Cys^{1}	H^{α}	Cys^{1}	$H^{\beta b}$	2.94	2.69	2.55	2.51	2.94	2.85
Cys^{1}	$H^{\beta a}$	Cys^{1}	$H^{\beta b}$	1.76	1.75	1.75	1.75	1.76	1.76
Cys^{1}	H^{α}	Tyr^{2}	H	2.26	2.28	2.33	2.30	2.26	2.27
Cys^{1}	H^{α}	Phe^{3}	H	6.30	4.00	6.39	6.16	5.55	5.53
Cys^{1}	$H^{\beta a}$	Cys^{6}	H	5.44	5.88	4.76	4.30	5.68	5.55
Cys^{1}	$H^{\beta a}$	Cys^{6}	H^{α}	5.94	5.09	4.95	5.22	5.71	5.66
Tyr^{2}	H	Tyr^{2}	H^{α}	2.88	2.93	2.89	2.93	2.89	2.89
Tyr^{2}	H^{α}	Tyr^{2}	$H^{\beta a}$	2.53	2.50	2.49	2.59	2.56	2.52
Tyr^{2}	H^{α}	Tyr^{2}	$H^{\beta b}$	2.90	2.64	3.01	2.78	2.81	2.82
Tyr^{2}	H^{α}	Tyr^{2}	H^{δ}*($H^{\delta 1}$)	4.03	3.83	4.25	3.46	4.03	3.96
Tyr^{2}	$H^{\beta a}$	Tyr^{2}	$H^{\beta b}$	1.75	1.74	1.75	1.75	1.75	1.75
Tyr^{2}	$H^{\beta a}$	Tyr^{2}	H^{δ}*($H^{\delta 1}$)	3.24	3.22	3.57	2.91	3.22	3.23
Tyr^{2}	$H^{\beta b}$	Tyr^{2}	H^{δ}*($H^{\delta 1}$)	2.66	2.88	2.41	3.17	2.70	2.73
Tyr^{2}	$H^{\delta 1}$	Tyr^{2}	H^{ε}*($H^{\varepsilon 1}$)	2.48	2.47	2.48	2.49	2.48	2.48
Tyr^{2}	H	Phe^{3}	H	4.63	2.30	4.57	4.36	3.81	3.80
Tyr^{2}	H^{α}	Phe^{3}	H	2.28	3.38	2.38	2.20	2.57	2.57
Tyr^{2}	$H^{\beta a}$	Phe^{3}	H	3.97	4.12	3.90	3.94	3.96	4.01
Phe^{3}	H	Phe^{3}	H^{α}	2.77	2.91	2.78	2.23	2.80	2.80
Phe^{3}	H	Phe^{3}	$H^{\beta a}$	3.18	2.89	2.49	3.30	3.00	3.09
Phe^{3}	H	Phe^{3}	$H^{\beta b}$	2.47	2.88	2.48	3.78	2.62	2.60
Phe^{3}	H^{α}	Phe^{3}	$H^{\beta a}$	2.40	2.57	2.45	2.77	2.47	2.46
Phe^{3}	H^{α}	Phe^{3}	$H^{\beta b}$	2.60	2.78	2.98	2.66	2.68	2.66
Phe^{3}	H^{α}	Phe^{3}	H^{δ}*($H^{\delta 1}$)	4.04	3.59	3.18	3.41	3.75	3.89
Phe^{3}	$H^{\beta a}$	Phe^{3}	$H^{\beta b}$	1.74	1.75	1.75	1.75	1.74	1.74
Phe^{3}	$H^{\beta a}$	Phe^{3}	H^{δ}*($H^{\delta 1}$)	2.95	3.01	2.88	3.06	3.00	2.97
Phe^{3}	$H^{\beta b}$	Phe^{3}	H^{δ}*($H^{\delta 1}$)	3.10	3.05	3.28	3.00	3.10	3.09
Phe^{3}	$H^{\delta 1}$	Phe^{3}	H^{ε}*($H^{\varepsilon 1}$)	2.47	2.47	2.47	2.48	2.47	2.47
Phe^{3}	$H^{\varepsilon 1}$	Phe^{3}	H^{ζ}	2.48	2.48	2.48	2.48	2.48	2.48
Phe^{3}	H^{α}	Gln^{4}	H	3.49	2.38	2.15	3.06	3.11	3.11
Phe^{3}	$H^{\beta a}$	Gln^{4}	H	4.11	3.61	4.13	4.17	3.96	3.96
Phe^{3}	$H^{\beta b}$	Gln^{4}	H	3.86	3.44	4.19	4.12	3.65	3.74
Gln^{4}	H	Gln^{4}	H^{α}	2.92	2.82	2.23	2.94	2.88	2.89
Gln^{4}	H	Gln^{4}	$H^{\beta a}$	2.58	2.76	3.49	3.11	2.67	2.64
Gln^{4}	H	Gln^{4}	$H^{\beta b}$	3.39	2.97	3.94	3.27	3.22	3.26
Gln^{4}	H	Gln^{4}	$H^{\gamma *}$($H^{\gamma a}$)	3.49	3.21	3.37	3.75	3.41	3.40
Gln^{4}	H^{α}	Gln^{4}	$H^{\beta a}$	2.89	2.70	2.95	2.68	2.81	2.83
Gln^{4}	H^{α}	Gln^{4}	$H^{\beta b}$	2.55	2.58	2.52	2.67	2.57	2.56
Gln^{4}	H^{α}	Gln^{4}	$H^{\gamma *}$($H^{\gamma a}$)	2.90	3.24	2.68	3.12	3.05	3.00
Gln^{4}	$H^{\beta a}$	Gln^{4}	$H^{\beta b}$	1.76	1.75	1.76	1.76	1.76	1.76
Gln^{4}	$H^{\beta a}$	Gln^{4}	$H^{\gamma *}$($H^{\gamma a}$)	2.84	2.86	2.92	2.94	2.84	2.85
Gln^{4}	$H^{\varepsilon 1}$	Gln^{4}	$H^{\varepsilon 2}$	1.75	1.75	1.75	1.75	1.75	1.75
Gln^{4}	H^{α}	Cys^{6}	H	4.62	4.85	4.02	5.80	4.70	4.70

Supporting Information

				Interatomic distances (Å)					
				Individual conformations				Metadynamics equilibrium	
res	atom	res	atom	*Saddle*[a] *equilibrium*	*Clinched open*[b] *equilibrium*	*Twisted saddle extended*	*Open extended*	Equilibrium[c] *extended*	**Equilibrium**[d] ***equilibrium***
Asn^5	H	Asn^5	H^{α}	2.96	2.94	2.94	2.86	2.95	2.95
Asn^5	H	Asn^5	$H^{\beta}*(H^{\beta a})$	2.86	3.10	2.81	2.63	2.94	2.93
Asn^5	H^{α}	Asn^5	$H^{\beta}*(H^{\beta a})$	2.54	2.49	2.50	2.49	2.52	2.52
Asn^5	$H^{\delta 1}$	Asn^5	$H^{\delta 2}$	1.75	1.75	1.75	1.75	1.75	1.75
Asn^5	H	Cys^6	H	2.17	3.49	2.49	4.30	2.51	2.53
Asn^5	H^{α}	Cys^6	H	3.53	2.58	3.32	2.26	3.22	3.21
Cys^6	H	Cys^6	H^{α}	2.97	2.92	2.94	2.93	2.95	2.95
Cys^6	H	Cys^6	$H^{\beta a}$	2.83	2.80	3.19	2.70	2.73	2.82
Cys^6	H	Cys^6	$H^{\beta b}$	3.68	3.17	3.25	3.65	3.62	3.52
Cys^6	H^{α}	Cys^6	$H^{\beta a}$	2.93	2.72	2.61	2.96	2.94	2.86
Cys^6	H^{α}	Cys^6	$H^{\beta b}$	2.44	2.70	2.57	2.54	2.48	2.51
Cys^6	$H^{\beta a}$	Cys^6	$H^{\beta b}$	1.75	1.75	1.75	1.75	1.75	1.75
Cys^6	H	Pro^7	$H^{\delta a}$	4.78	5.03	5.02	4.96	4.81	4.86
Cys^6	H	Pro^7	$H^{\delta b}$	4.20	4.74	4.72	4.62	4.25	4.36
Cys^6	H^{α}	Pro^7	$H^{\delta a}$	2.50	2.58	2.66	2.63	2.51	2.52
Cys^6	H^{α}	Pro^7	$H^{\delta b}$	2.49	2.36	2.34	2.36	2.48	2.45
Pro^7	H^{α}	Pro^7	$H^{\beta a}$	2.84	2.84	2.84	2.84	2.84	2.84
Pro^7	H^{α}	Pro^7	$H^{\beta b}$	2.31	2.31	2.31	2.31	2.31	2.31
Pro^7	H^{α}	Pro^7	$H^{\gamma *}$	3.66	3.65	3.65	3.65	3.65	3.65
Pro^7	$H^{\beta a}$	Pro^7	$H^{\beta b}$	1.78	1.78	1.78	1.79	1.78	1.78
Pro^7	$H^{\beta a}$	Pro^7	$H^{\gamma *}$	2.45	2.46	2.46	2.46	2.45	2.45
Pro^7	$H^{\beta b}$	Pro^7	$H^{\gamma *}$	2.50	2.49	2.49	2.49	2.50	2.50
Pro^7	$H^{\gamma}*$	Pro^7	$H^{\delta}*$	2.29	2.29	2.29	2.29	2.29	2.29
Pro^7	$H^{\delta b}$	Pro^7	$H^{\delta b}$	1.78	1.78	1.79	1.78	1.78	1.78
Pro^7	H^{α}	Arg^8	H	2.33	2.26	2.38	2.36	2.34	2.31
Arg^8	H	Arg^8	H^{α}	2.75	2.68	2.93	2.94	2.93	2.73
Arg^8	H	Arg^8	$H^{\beta a}$	2.94	2.98	2.77	2.75	2.77	2.95
Arg^8	H	Arg^8	$H^{\beta b}$	3.50	3.55	3.34	3.34	3.39	3.51
Arg^8	H^{α}	Arg^8	$H^{\beta a}$	2.84	2.86	2.81	2.81	2.83	2.84
Arg^8	H^{α}	Arg^8	$H^{\beta b}$	2.57	2.57	2.60	2.59	2.59	2.57
Arg^8	$H^{\beta a}$	Arg^8	$H^{\beta b}$	1.75	1.75	1.76	1.76	1.75	1.75
Arg^8	$H^{\beta a}$	Arg^8	$H^{\gamma}*$	2.57	2.56	2.57	2.58	2.56	2.57
Arg^8	$H^{\beta b}$	Arg^8	$H^{\gamma}*$	2.60	2.59	2.60	2.61	2.59	2.60
Arg^8	$H^{\gamma}*$	Arg^8	$H^{\delta}*$	2.40	2.40	2.40	2.39	2.40	2.40
Gly^9	H	Gly^9	$H^{\alpha 1,2}$	2.49	2.50	2.50	2.50	2.50	2.50
Gly^9	$H^{\alpha 1,2}$	Gly^9	$H^{N1,2}$	3.07	3.09	3.00	3.00	3.00	3.07
Gly^9	H^{N1}	Gly^9	H^{N2}	1.75	1.75	1.75	1.75	1.75	1.75

[a]Eq. 7; [b]Eq. 8; [c]Eq. 5; [d]Eq. 6

Supporting Information

Experimental details

Sample preparation for NMR

Arg^8-vasopressin was obtained from Bachem (UK) Ltd as the trifluoroacetate salt of the chemically synthesized peptide, having a purity (by HPLC) of >96%. Mass spectrometry of the synthesized material gave a molecular mass of 1084.55 Da, in close agreement to the calculated molecular mass of 1086.26 Da for the reduced form of the peptide.

Samples of 5.0 mg dry weight were dissolved in 320 µl of 90% H_2O/ 10% D_2O to give a peptide concentration of 14.4 mM. The pH of the sample was measured to be 4.7 and NMR spectra were recorded without adjustment. In addition the sample was dried by lyophilization, then redissolved in 320 µl of 20 mM potassium phosphate buffer (pH 6.5) in 90% H_2O/ 10% D_2O and NMR spectra were recorded at a pH measured as 6.0. NMR spectra of Arg^8-vasopressin in D_2O at both pH 4.7 and pH 6.0 were recorded at least 2h after redissolving the extensively dried samples in 99.9% D_2O (Sigma Aldrich).

NMR experiments

NMR spectroscopy was performed on a Varian Inova 600 MHz spectrometer, equipped with 5-channels, a 5 mm triple resonance ($^1H/^{13}C/^{15}N$) coldprobe and actively shielded pulse field z-axis gradients.

Proton resonance assignments were achieved using a combination of 2D ^1H-^1H total chemical shift correlation spectroscopy (TOCSY),[25] and ^1H-^1H nuclear Overhauser effect spectroscopy (NOESY) NMR spectra.[26] Spectra were acquired as 2048 complex points, with 32 transients for each of 512 increments and a spectral width of 10.0 ppm in both dimensions. Mixing times of 60 and 75 ms for the TOCSY experiment and 200 and 300 ms for the NOESY experiment were used. Water suppression was achieved through use of the watergate 3919 sequence.[27]

Resonance assignments for carbon and nitrogen at natural abundance were obtained through the use of gradient heteronuclear single quantum coherence (gHSQC) experiments. A standard ^{13}C-^1H gHSQC NMR spectrum,[28, 29] was acquired as 1024 complex points in t2 (observe ^1H dimension) and 280 increments in t1 (indirect ^{13}C dimension) using 32 transients over spectral widths of 6000.60 Hz (10.0 ppm) and 21114.68 Hz (140.0 ppm) respectively. The transmitter offsets were initially set to the water position in the ^1H and to 70 ppm in the ^{13}C dimension, but other combinations of offset and sweep width were later used to focus on the aliphatic and aromatic regions. A 2D sensitivity enhanced ^{15}N-^1H gHSQC NMR spectrum[28, 30, 31] was acquired as 2048 complex points in t2 (observe ^1H dimension) and 128 increments in t1 (indirect ^{15}N dimension) using 32 transients over spectral widths of 6000.60 Hz (10.0 ppm) and 2431.06 Hz (40.0 ppm) respectively. The transmitter offsets were set to the water position in the ^1H and to 120 ppm in the ^{15}N dimension. States–TPPI quadrature detection was employed in the ^{15}N-dimension.[32]

Supporting Information

Spectral processing and format conversion was performed using NMRPipe[33] and visualized with NMRView[34]. Arg^8-vasopressin spectra were assigned using Analysis v2.0.7 from the CcpNMR software suite.[35] Proton and ^{13}C chemical shifts were referenced to 3-trimethyl silyl propane sulfonic acid (DSS) and ^{15}N chemical shifts were referenced to an external liquid ammonia. The ^{1}H, ^{13}C and ^{15}N chemical shifts of the major populated *trans*-Pro^7 isomer of Arg^8-vasopressin in H_2O/pH 4.7, D_2O/pH 4.7, H_2O/pH 6.0, D_2O/pH 6.0 are given in Table S6-9. The volumes of assigned peaks were determined using the box sum method in Analysis with an r^{-6} distance calibration against the fixed distance between the Tyr^2 H^{δ} and H^{ε} atoms. A 20% change (the default) in the calculated target distance was taken. These experimentally derived distances are listed in Table S10.

Supporting Information

Experimental ^{1}H, ^{13}C and ^{15}N chemical shifts

Table S6 Experimental NMR chemical shifts (δ ppm) of Arg^{8}-vasopressin in H_2O at pH 6.0/ 298 K

Residue	H^N	N	H^α	C^α	Others
Cys^{1}	-	-	3.97	56.18	3.27:$H^{\beta a}$; 3.12:$H^{\beta b}$ 44.23:C^{β}
Tyr^{2}	8.57$^{?}$	123.79$^{?}$	4.64	58.02$^{?}$	2.80$H^{\beta a}$; 2.96:$H^{\beta b}$; 7.05:H^{δ}*; 6.83:H^{ϵ}* 39.10:C^{β}; 133.34:C^{δ}*; 118.46:C^{ϵ}*
Phe^{3}	8.04	122.55$^{?}$	4.54	58.37	3.01:$H^{\beta a}$; 3.31:$H^{\beta b}$; 7.23:H^{δ}*; 7.40:H^{ϵ}*; 7.37:H^{ζ} 39.42:C^{β}; 131.99:C^{δ}*; 131.81:C^{ϵ}*; 130.17:C^{ζ}
Gln^{4}	8.32	120.80	4.12	57.82	2.05:$H^{\beta a}$; 2.12:$H^{\beta b}$; 2.29:H^{γ}*; 6.89:$H^{\epsilon a}$; 7.53:$H^{\epsilon b}$ 28.66:C^{β}; 33.99:C^{γ}; 114.24:N^{ϵ}
Asn^{5}	8.30	118.20	4.77	-	2.86:H^{β}*; 6.92:$H^{\delta a}$; 7.63:$H^{\delta b}$ 38.74:C^{β}; 114.53:N^{δ}
Cys^{6}	8.15	122.39	4.89	-	3.18:$H^{\beta a}$; 2.93:$H^{\beta b}$ 41.69:C^{β}
trans-Pro^{7}	-	-	4.45	63.52	1.93:$H^{\beta a}$; 2.31:$H^{\beta b}$; 2.06:H^{γ}*; 3.73:$H^{\delta a}$; 3.83:$H^{\delta b}$ 32.16:C^{β}; 27.60:C^{γ}; 50.79:C^{δ}
Arg^{8}	8.63	123.97	4.32	56.45	1.80:$H^{\beta a}$; 1.90:$H^{\beta b}$; 1.67:H^{γ}*; 3.22:H^{δ}*; 7.22:H^{ϵ} 30.76:C^{β}; 27.28:C^{γ}; 43.46:C^{δ}; 86.76:N^{ϵ}
Gly^{9}	8.45	113.01	3.93*	45.06	
NH_2^{10}	-	-	-	-	7.09:H^{N1}; 7.48:H^{N2}; 109.16:N

$^{?}$tentative assignment; *degenerate atoms

Table S7 Experimental NMR chemical shifts (δ ppm) of Arg^{8}-vasopressin in D_2O at pH 6.0/ 298 K

Residue	H^N	N	H^α	C^α	Others
Cys^{1}	-	-	3.98	56.01	3.28:$H^{\beta a}$; 3.12:$H^{\beta b}$ 44.20:C^{β}
Tyr^{2}	-	-	4.64	58.02	2.81:$H^{\beta a}$; 2.97:$H^{\beta b}$; 7.06:H^{δ}*; 6.83:H^{ϵ}* 39.11:C^{β}; 133.35:C^{δ}*; 118.38:C^{ϵ}*
Phe^{3}	8.05	-	4.54	58.33	3.01:$H^{\beta a}$; 3.31:$H^{\beta b}$; 7.23:H^{δ}*; 7.40:H^{ϵ}*; 7.37:H^{ζ} 39.37:C^{β}; 131.99:C^{δ}*; 131.81:C^{ϵ}*; 130.18 C^{ζ}
Gln^{4}	-	-	4.12	57.75	2.05:$H^{\beta a}$; 2.13:$H^{\beta b}$; 2.29:H^{γ}*; 6.89:$H^{\epsilon a}$; 7.53:$H^{\epsilon b}$ 28.60:C^{β}; 33.93:C^{γ}
Asn^{5}	8.31	-	4.78	53.08	2.86:H^{β}*; 6.92:$H^{\delta a}$; 7.63:$H^{\delta b}$ 38.66:C^{β}
Cys^{6}	8.15	-	4.90	54.27	3.19:$H^{\beta a}$; 2.93:$H^{\beta b}$ 41.67:C^{β}
trans-Pro^{7}	-	-	4.46	63.51	1.94:$H^{\beta a}$; 2.32:$H^{\beta b}$; 2.07:H^{γ}*; 3.74:$H^{\delta a}$; 3.84:$H^{\delta b}$ 32.16:C^{β}; 27.61:C^{γ}; 50.80:C^{δ}
Arg^{8}	8.62	-	4.32	56.36	1.81:$H^{\beta a}$; 1.91:$H^{\beta b}$; 1.68:H^{γ}*; 3.23:H^{δ}* 30.72:C^{β}; 27.28:C^{γ}; 43.34:C^{δ}
Gly^{9}	8.46	-	3.92*	44.93	
NH_2^{10}	-	-	-	-	-

*degenerate atoms

Supporting Information

Table S8 Experimental NMR chemical shifts (δ ppm) of Arg^8-vasopressin in H_2O at pH 4.7/ 298 K

Residue	H^N	N	H^α	C^α	Others
Cys^1	-	-	4.29	55.39	3.46:$H^{\beta a}$; 3.25:$H^{\beta b}$ 42.87:C^β
Tyr^2	8.90	125.26	4.67	-	2.85:$H^{\beta a}$; 2.95:$H^{\beta b}$; 7.07:H^δ*; 6.84:H^ϵ* 39.25:C^β; 133.36:C^δ*; 118.49:C^ϵ*
Phe^3	8.15	123.04	4.48	58.63	3.01:$H^{\beta a}$; 3.30:$H^{\beta b}$; 7.22:H^δ*; 7.40:H^ϵ*; 7.38:H^ζ 39.31:C^β; 131.96:C^δ*; 131.81:C^ϵ*; 130.17:C^ζ
Gln^4	8.32	121.11	4.12	57.93	2.05:$H^{\beta a}$; 2.13:$H^{\beta b}$; 2.30:H^γ*; 6.90:$H^{\epsilon a}$; 7.53:$H^{\epsilon b}$ 28.64:C^β; 33.96:C^γ; 114.18:N^ϵ
Asn^5	8.33	118.14	4.80	-	2.88:H^β*; 6.93:$H^{\delta a}$; 7.63:$H^{\delta b}$ 38.83:C^β; 114.55:N^δ
Cys^6	8.14	122.09	4.92	-	3.21:$H^{\beta a}$; 2.94:$H^{\beta b}$ 41.28:C^β
trans-Pro^7	-	-	4.46	63.48	1.94:$H^{\beta a}$; 2.32:$H^{\beta b}$; 2.07:H^γ*; 3.75:$H^{\delta a}$; 3.85:$H^{\delta b}$ 32.18:C^β; 27.60:C^γ; 50.80:C^δ
Arg^8	8.65	124.06	4.32	56.50	1.81:$H^{\beta a}$; 1.91:$H^{\beta b}$; 1.69:$H^{\gamma a}$; 1.68:$H^{\gamma b}$; 3.23:H^δ*; 7.21:H^ϵ 30.75:C^β; 27.28:C^γ; 43.46:C^δ; 86.76:N^ϵ
Gly^9	8.44	112.97	3.94*	45.04	
NH_2^{10}	-	-	-	-	7.09:H^{N1}; 7.48:H^{N2}; 109.17:N

?tentative assignment; *degenerate atoms

Table S9 Experimental NMR chemical shifts (δ ppm) of Arg^8-vasopressin in D_2O at pH 4.7/ 298 K

Residue	H^N	N	H^α	C^α	Others
Cys^1	-	-	4.29	55.22	3.46:$H^{\beta a}$; 3.25:$H^{\beta b}$ 42.83:C^β
Tyr^2	-	-	4.66	58.12	2.85:$H^{\beta a}$; 2.96:$H^{\beta b}$; 7.07:H^δ*; 6.84:H^ϵ* 39.27:C^β; 133.37:C^δ*; 118.42:C^ϵ*
Phe^3	-	-	4.47	58.59	3.02:$H^{\beta a}$; 3.30:$H^{\beta b}$; 7.21:H^δ*; 7.40:H^ϵ*; 7.37:H^ζ 39.28:C^β; 131.96:C^δ*; 131.81:C^ϵ*; 130.18:C^ζ
Gln^4	8.33	-	4.12	57.85	2.05:$H^{\beta a}$; 2.13:$H^{\beta b}$; 2.30:H^γ*; 6.90:$H^{\epsilon a}$; 7.54:$H^{\epsilon b}$ 28.58:C^β; 33.89:C^γ
Asn^5	8.30	-	4.81	53.04	2.89:H^β*; 6.94:$H^{\delta a}$; 7.63:$H^{\delta b}$ 38.75:C^β
Cys^6	-	-	4.92	53.97	3.21:$H^{\beta a}$; 2.93:$H^{\beta b}$ 41.26:C^β
trans-Pro^7	-	-	4.46	63.49	1.94:$H^{\beta a}$; 2.33:$H^{\beta b}$; 2.07:H^γ*; 3.75:$H^{\delta a}$; 3.85:$H^{\delta b}$ 32.18:C^β; 27.61:C^γ; 50.81:C^δ
Arg^8	8.65	-	4.31	56.41	1.82:$H^{\beta a}$; 1.91:$H^{\beta b}$; 1.68:H^γ*; 3.23:H^δ* 30.71:C^β; 27.28:C^γ; 43.34:C^δ
Gly^9	8.45	-	3.93*	44.91	
NH_2^{10}	-	-	-	-	7.48:H^N*

*degenerate atoms

Supporting Information

Experimental NOE distances

Table S10 Experimental NOE distances of Arg^8-vasopressin

				NOE distances (Å)					
				pH 4.7			pH 6.0		
res	atom	res	atom	Constraint	Error limits		Constraint	Error limits	
				r_{exp}	+	-	r_{exp}	+	-
Cys^1	H^{α}	Cys^1	$H^{\beta a}$	2.4	0.5	0.5	2.3	0.5	0.5
Cys^1	H^{α}	Cys^1	$H^{\beta b}$	2.8	0.6	0.6	2.3	0.5	0.5
Cys^1	$H^{\beta a}$	Cys^1	$H^{\beta b}$	2.0	0.3	0.4	2.0	0.2	0.4
Cys^1	H^{α}	Tyr^2	H	4.1	0.8	0.8	-	-	-
Cys^1	H^{α}	Phe^3	H	5.5	1.1	1.1	-	-	-
Cys^1	$H^{\beta a}$	Cys^6	H	4.5	0.9	0.9	-	-	-
Cys^1	$H^{\beta a}$	Cys^6	H^{α}	4.5	0.9	0.9	-	-	-
Tyr^2	H	Tyr^2	H^{α}	4.4	0.9	0.9	-	-	-
Tyr^2	H^{α}	Tyr^2	$H^{\beta a,b}$	2.3	0.5	0.5	2.2	0.4	0.4
Tyr^2	H^{α}	Tyr^2	$H^{\beta b,a}$	2.4	0.5	0.5	2.3	0.5	0.5
Tyr^2	H^{α}	Tyr^2	H^{δ}*	4.2	0.8	0.8	-	-	-
Tyr^2	$H^{\beta a}$	Tyr^2	$H^{\beta b}$	1.8	0	0.4	2.1	0.3	0.4
Tyr^2	$H^{\beta a}$	Tyr^2	H^{δ}*	4.4	0.9	0.9	-	-	-
Tyr^2	$H^{\beta b}$	Tyr^2	H^{δ}*	4.3	0.9	0.9	-	-	-
Tyr^2	$H^{\delta 1}$	Tyr^2	H^{ε}*	2.5	0.5	0.5	2.6	0.5	0.5
Tyr^2	H	Phe^3	H	5.8	1.2	1.2	-	-	-
Tyr^2	H^{α}	Phe^3	H	3.2	0.6	0.6	3.3	0.7	0.7
Tyr^2	$H^{\beta a}$	Phe^3	H	4.4	0.9	0.9	-	-	-
Phe^3	H	Phe^3	H^{α}	3.1	0.6	0.6	3.3	0.7	0.7
Phe^3	H	Phe^3	$H^{\beta a}$	4.2	0.8	0.8	-	-	-
Phe^3	H	Phe^3	$H^{\beta b}$	3.8	0.8	0.8	-	-	-
Phe^3	H^{α}	Phe^3	$H^{\beta a}$	3.2	0.6	0.6	3.0	0.6	0.6
Phe^3	H^{α}	Phe^3	$H^{\beta b}$	3.4	0.7	0.7	-	-	-
Phe^3	H^{α}	Phe^3	H^{δ}*	4.7	0.9	0.9	-	-	-
Phe^3	$H^{\beta a}$	Phe^3	$H^{\beta b}$	2.1	0.4	0.4	2.1	0.4	0.4
Phe^3	$H^{\beta a}$	Phe^3	H^{δ}*	2.8	0.6	0.6	-	-	-
Phe^3	$H^{\beta b}$	Phe^3	H^{δ}*	4.0	0.8	0.8	-	-	-
Phe^3	$H^{\delta 1}$	Phe^3	H^{ε}*	2.6	0.5	0.5	2.6	0.5	0.5
Phe^3	$H^{\varepsilon 1}$	Phe^3	H^{ζ}	2.0	0.3	0.4	2.2	0.4	0.4
Phe^3	H	Gln^4	H	-	-	-	3.7	0.7	0.7
Phe^3	H^{α}	Gln^4	H	3.3	0.7	0.7	3.3	0.7	0.7
Phe^3	$H^{\beta a}$	Gln^4	H	3.5	0.7	0.7	-	-	-
Phe^3	$H^{\beta b}$	Gln^4	H	3.2	0.6	0.6	-	-	-
Gln^4	H	Gln^4	H^{α}	3.1	0.6	0.6	-	-	-
Gln^4	H	Gln^4	$H^{\beta a}$	3.5	0.7	0.7	-	-	-
Gln^4	H	Gln^4	$H^{\beta b}$	3.8	0.8	0.8	-	-	-
Gln^4	H	Gln^4	H^{γ}*($H^{\gamma a}$)	4.4	0.9	0.9	-	-	-
Gln^4	H^{α}	Gln^4	$H^{\beta a}$	2.9	0.6	0.6	2.5	0.5	0.5
Gln^4	H^{α}	Gln^4	$H^{\beta b}$	2.7	0.5	0.5	2.5	0.5	0.5
Gln^4	H^{α}	Gln^4	H^{γ}*($H^{\gamma a}$)	3.9	0.8	0.8	2.5	0.5	0.5
Gln^4	$H^{\beta a}$	Gln^4	$H^{\beta b}$	2.1	0.3	0.4	-	-	-

Supporting Information

NOE distances (Å)									
				pH 4.7			pH 6.0		
res	atom	res	atom	Constraint	Error limits		Constraint	Error limits	
				r_{exp}	+	-	r_{exp}	+	-
Gln^4	$H^{\beta a}$	Gln^4	$H^{\gamma*}$ ($H^{\gamma a}$)	2.4	0.5	0.5	2.5	0.5	0.5
Gln^4	$H^{\varepsilon 1}$	Gln^4	$H^{\varepsilon 2}$	1.8	0.0	0.4	1.8	0.1	0.4
Gln^4	H^{α}	Cys^6	H	5.0	1.0	1.0	-	-	-
Gln^4	H^{α}	Asn^5	H	-	-	-	3.0	0.6	0.6
Asn^5	H	Asn^5	H^{α}	3.0	0.6	0.6	-	-	-
Asn^5	H	Asn^5	H^{β}*	3.4	0.7	0.7	-	-	-
Asn^5	H^{α}	Asn^5	H^{β}*	3.2	0.6	0.6	2.9	0.6	0.6
Asn^5	$H^{\delta 1}$	Asn^5	$H^{\delta 2}$	1.7	0.1	0.3	1.7	0	0.3
Asn^5	H	Cys^6	H	3.1	0.6	0.6	3.5	0.7	0.7
Asn^5	H^{α}	Cys^6	H	3.3	0.7	0.7	-	-	-
Cys^6	H	Cys^6	H^{α}	2.8	0.6	0.6	-	-	-
Cys^6	H	Cys^6	$H^{\beta a}$	3.2	0.6	0.6	-	-	-
Cys^6	H	Cys^6	$H^{\beta b}$	4.0	0.8	0.8	-	-	-
Cys^6	H^{α}	Cys^6	$H^{\beta a}$	2.5	0.5	0.5	2.9	0.6	0.6
Cys^6	H^{α}	Cys^6	$H^{\beta b}$	2.5	0.5	0.5	2.8	0.6	0.6
Cys^6	$H^{\beta a}$	Cys^6	$H^{\beta b}$	2.0	0.3	0.4	2.0	0.3	0.4
Cys^6	H	Pro^7	$H^{\delta a}$	5.1	1.0	1.0	-	-	-
Cys^6	H	Pro^7	$H^{\delta b}$	4.9	1.0	1.0	-	-	-
Cys^6	H^{α}	Pro^7	$H^{\delta a}$	3.7	0.7	0.7	-	-	-
Cys^6	H^{α}	Pro^7	$H^{\delta b}$	3.6	0.7	0.7	-	-	-
Pro^7	H^{α}	Pro^7	$H^{\beta a}$	2.2	0.4	0.4	2.3	0.5	0.5
Pro^7	H^{α}	Pro^7	$H^{\beta b}$	2.0	0.2	0.4	2.3	0.5	0.5
Pro^7	H^{α}	Pro^7	$H^{\gamma*}$	3.8	0.8	0.8	-	-	-
Pro^7	$H^{\beta a}$	Pro^7	$H^{\beta b}$	2.6	0.5	0.5	2.4	0.5	0.5
Pro^7	$H^{\beta a}$	Pro^7	$H^{\gamma*}$	2.4	0.5	0.5	2.4	0.5	0.5
Pro^7	$H^{\beta b}$	Pro^7	$H^{\gamma*}$	2.6	0.5	0.5	2.6	0.5	0.5
Pro^7	H^{γ}*	Pro^7	H^{δ}*	3.4	0.7	0.7	2.9	0.6	0.6
Pro^7	$H^{\delta b}$	Pro^7	$H^{\delta b}$	2.4	0.5	0.5	2.2	0.4	0.4
Pro^7	H^{α}	Arg^8	H	2.8	0.6	0.6	2.9	0.6	0.6
Arg^8	H	Arg^8	H^{α}	2.8	0.6	0.6	3.5	0.7	0.7
Arg^8	H	Arg^8	$H^{\beta a}$	4.2	0.8	0.8	-	-	-
Arg^8	H	Arg^8	$H^{\beta b}$	4.5	0.9	0.9	-	-	-
Arg^8	H^{α}	Arg^8	$H^{\beta a}$	3.1	0.6	0.6	2.8	0.6	0.6
Arg^8	H^{α}	Arg^8	$H^{\beta b}$	2.9	0.6	0.6	2.8	0.6	0.6
Arg^8	$H^{\beta a}$	Arg^8	$H^{\beta b}$	2.2	0.4	0.4	2.1	0.4	0.4
Arg^8	$H^{\beta a}$	Arg^8	H^{γ}*	2.5	0.5	0.5	2.4	0.5	0.5
Arg^8	$H^{\beta b}$	Arg^8	H^{γ}*	2.7	0.5	0.5	2.6	0.5	0.5
Arg^8	H^{γ}*	Arg^8	H^{δ}*	2.7	0.5	0.5	-	-	-
Gly^9	H	Gly^9	$H^{\alpha 1,2}$	2.5	0.5	0.5	-	-	-
Gly^9	$H^{\alpha 1,2}$	Gly^9	$H^{N1,2}$	4.8	1.0	1.0	-	-	-
Gly^9	H^{N1}	Gly^9	H^{N2}	1.6	0.1	0.3	1.9	0.1	0.4

Supporting Information

Gaussian archive entries for the B3LYP/6-31G(d)-optimized geometries.

***Saddle*, folded:**

1\1\FAU-CCC-CCDH171\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\09-F
eb-2016\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP saddle with f
olded tail conformation\\2,1\N,-1.4938370366,-0.6015217359,12.69638104
71\H,-0.9903782545,-1.5308378805,12.7504989922\H,-0.8120518054,0.15181
10275,12.8326961217\H,-1.864119398,-0.5033399292,11.7454588633\C,-2.60
5792741,-0.5379161939,13.7155775749\H,-3.1272650232,0.4092089179,13.56
57956249\C,-3.5365011596,-1.7319516155,13.4178273726\H,-4.0811751524,-
1.5399616501,12.4899417702\H,-2.9239088163,-2.6256043932,13.2846564448
\S,-4.7416176025,-2.0836913771,14.7794126001\C,-1.9074363272,-0.570973
569,15.0898536924\O,-0.8862951443,-1.2401765695,15.2457431489\N,-2.440
4259855,0.191763833,16.0614987713\H,-3.2752259189,0.737835376,15.88877
86607\C,-1.81556349,0.2788706725,17.3797275877\H,-0.8062490779,0.68372
96834,17.2494270715\C,-2.6475139559,1.237607738,18.2609920639\H,-2.787
5771881,2.1610416502,17.6836944533\H,-3.6393370486,0.7975285728,18.412
084954\C,-2.0120234757,1.5600552325,19.5973358275\C,-0.777570942,2.222
2731772,19.6750796848\H,-0.2544535776,2.507786744,18.7653090771\C,-0.2
013346913,2.542564407,20.9031226035\H,0.7573791642,3.054800283,20.9387
612406\C,-0.8653551264,2.2135865057,22.091514115\O,-0.3614814147,2.503
4483726,23.326783623\H,0.4931870747,2.9546035973,23.2300859509\C,-2.10
47138479,1.5665240756,22.034757754\H,-2.6222336217,1.3324771725,22.959
8151025\C,-2.6623940603,1.2428382407,20.7983366683\H,-3.6314359279,0.7
501133008,20.7687275015\C,-1.7067274053,-1.1312443015,17.9970141985\O,
-2.6443626237,-1.9348709493,17.9161253668\N,-0.5555920276,-1.415201642
7,18.6453494755\H,0.1663074328,-0.7047273746,18.6843521394\C,-0.168134
9355,-2.7860094021,18.9854756442\H,-0.537534714,-3.4365174647,18.18462
97966\C,1.3726669538,-2.9263203407,19.0328974749\H,1.5725283117,-3.973
7235764,19.2774503306\H,1.7561772452,-2.749113278,18.0226104799\C,2.06
15343458,-2.0068742125,20.0207203673\C,2.0609839625,-2.2959298728,21.3
953418302\H,1.5876309794,-3.2069291943,21.7519664203\C,2.6739387047,-1
.4344992468,22.3054729819\H,2.6668246519,-1.6806729184,23.3636938642\C
,3.2997791612,-0.2665222104,21.8586034026\H,3.7791063265,0.4020902253,
22.5682300898\C,3.3146070867,0.0283748682,20.4948489335\H,3.8083755581
,0.9269438007,20.1350873444\C,2.70216177,-0.8381541307,19.5841647418\H
,2.7389460203,-0.6115654988,18.5203585808\C,-0.829064977,-3.3772836854
,20.242780506\O,-0.6402312765,-4.5649587972,20.5072987981\N,-1.6492466
899,-2.5741577546,20.9682027117\H,-1.7082217143,-1.5920724623,20.72512
62504\C,-2.5423014686,-3.1180508137,21.9838685387\H,-2.0725706153,-4.0
278007727,22.3632085842\C,-2.7519176721,-2.1493289437,23.1593931226\H,
-3.211200055,-1.2239997959,22.7918853644\H,-3.4696748711,-2.618930618,
23.8358922006\C,-1.4710140576,-1.8002429392,23.9206241204\H,-0.7059466
155,-1.3986527806,23.2451290556\H,-1.6876597,-0.9989111824,24.63940870
11\C,-0.8807494765,-2.9750428888,24.7001135588\O,-1.4510693245,-4.0599
639721,24.8169051415\N,0.3422637719,-2.7377648967,25.2492507726\H,0.72
04032898,-3.4234539713,25.8900005475\H,0.723703374,-1.8030035166,25.29
55736956\C,-3.9075817202,-3.5680319732,21.4111438628\O,-4.7114370312,-
4.1550173662,22.1368177205\N,-4.1346624855,-3.2780298514,20.1042144663
\H,-3.4352202799,-2.7548987318,19.585618415\C,-5.3639930836,-3.6369881
053,19.4152832151\H,-6.024910324,-4.0886938439,20.154357237\C,-6.05510
38056,-2.3819830441,18.8376025408\H,-6.1158433098,-1.6408114656,19.643
3182007\H,-5.4511319711,-1.9439134656,18.0372511848\C,-7.4881686262,-2
.6719956463,18.3923706592\O,-8.2783092774,-3.2763225652,19.1167914113\
N,-7.8278221938,-2.2027472811,17.1669773814\H,-8.7790383209,-2.3175369
864,16.8423413805\H,-7.1901959993,-1.67480329,16.5887016921\C,-5.14852
67525,-4.7190357307,18.3321312567\O,-5.9325317146,-5.6632134075,18.234
2804478\N,-4.0892121103,-4.540785298,17.4929332952\H,-3.4457399182,-3.
7644007415,17.6274043324\C,-3.7506141093,-5.5281018133,16.484120581\H,
-4.5304351037,-6.2863551843,16.535687226\C,-3.6623321009,-4.9553085587
,15.0574077918\H,-2.8167264157,-4.2677961541,14.9859798717\H,-3.480058
9647,-5.7733933854,14.3531005866\S,-5.1637652684,-4.0927309583,14.4239
349699\C,-2.3579907822,-6.1139905787,16.7999939994\O,-1.4001404382,-5.
3419392298,16.9790470217\N,-2.2136163182,-7.4511441477,16.8514741095\C
,-3.2203930696,-8.4895605436,16.5400175803\H,-3.8156211585,-8.20848842
51,15.6690730919\H,-3.8918135316,-8.630964882,17.3951743856\C,-2.36119
79398,-9.7324872905,16.2828067978\H,-2.017148384,-9.7339270737,15.2442
730277\H,-2.9125821018,-10.6576202271,16.4683888395\C,-1.1692443826,-9
.5464465061,17.234240822\H,-0.2879260796,-10.1212109134,16.9381095061\
H,-1.4373210873,-9.844709345,18.2527934803\C,-0.9068147959,-8.02175105
91,17.2113194834\H,-0.6116049203,-7.6454465069,18.1952487167\C,0.15448
3021,-7.6250674196,16.1732111395\O,-0.0798547885,-7.6570532438,14.9601
654681\N,1.3444592008,-7.2534937567,16.7005121883\H,1.4274408752,-7.24

Supporting Information

85036015,17.7086027115\C,2.5229786673,-6.8087824372,15.9517310657\H,3.
1395847589,-6.2927669933,16.6974354844\C,3.3972826753,-7.9481405505,15
.3931836215\H,4.3169858817,-7.4850403446,15.0204550582\H,3.6793493526,
-8.583563539,16.2412655589\C,2.7785773507,-8.8078169378,14.2842398386\
H,1.8647521376,-9.2947649909,14.6374663678\H,2.5065000127,-8.169077955
4,13.4374922454\C,3.7688136265,-9.8771905086,13.8186136687\H,4.6919860
591,-9.4092104322,13.4542050558\H,4.0258198598,-10.5351477023,14.65801
28772\N,3.1726922716,-10.6813251871,12.7438235291\H,2.2788137323,-10.3
7999347,12.3778290395\C,3.7479322491,-11.7394793342,12.1692084122\N,4.
9599930663,-12.1498613985,12.5617732133\H,5.4117490229,-11.7450380218,
13.3683095963\H,5.3470088616,-13.0155039561,12.2129797708\N,3.10501300
22,-12.411507185,11.2000604547\H,2.2735681473,-12.0230103239,10.776976
483\H,3.5961741641,-13.1063802632,10.6548133481\C,2.2100891808,-5.7216
554212,14.9008818785\O,2.9334176821,-5.5776821502,13.9158899051\N,1.17
33350483,-4.8881731464,15.1822028085\H,0.5541471777,-5.085936346,15.96
3319407\C,0.8168717704,-3.7991957615,14.2975529267\H,0.4155025021,-2.9
653757221,14.875506144\H,1.7187641795,-3.4532564342,13.7845240038\C,-0
.2132817309,-4.1249627755,13.2171219534\O,-0.6702757754,-3.2130221091,
12.4976405632\N,-0.5796689504,-5.4026585811,13.0691652248\H,-0.2675266
131,-6.1444916948,13.6949608566\H,-1.2667189567,-5.6266958827,12.36040
94465\\Version=ES64L-G09RevD.01\State=1-A\HF=-4313.5927691\RMSD=3.591e
-09\RMSF=3.036e-06\Dipole=13.4501097,-4.2290357,-17.452117\Quadrupole=
-127.3751545,152.9779912,-25.6028367,-153.0985891,-9.8813411,143.23000
92\PG=C01 [X(C46H67N15O12S2)]\\@

***Saddle*, extended:**

1\1\FAU-CCC-CCDH171\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\23-A
pr-2015\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP_10us_T16_3_sa
ddle\\2,1\N,-5.2438321567,-5.6490704753,-16.3542613855\H,-5.5936123741
,-5.6194867886,-17.3196982834\H,-4.2388925628,-5.4434193036,-16.389696
2299\H,-5.7092717337,-4.8938011157,-15.8344195378\C,-5.5245781916,-6.9
806887529,-15.6947633282\H,-5.1779390577,-7.7451622323,-16.3917974115\
C,-4.7497007695,-7.0391067826,-14.3699953292\H,-3.6876353832,-6.851902
245,-14.5479475133\H,-5.143333902,-6.2883430372,-13.6806020209\S,-4.95
25760774,-8.7199371575,-13.6174522588\C,-7.051774079,-7.022519435,-15.
4805573825\O,-7.6089593165,-6.0544907688,-14.9434026512\N,-7.691413028
6,-8.1086850389,-15.9204083914\H,-7.15767149,-8.861476026,-16.33834847
14\C,-9.1434609539,-8.2707895853,-15.850703176\H,-9.5983708043,-7.3735
894984,-16.2824037378\C,-9.5173112668,-9.5157327937,-16.6978849388\H,-
9.031071766,-9.379277505,-17.672557498\H,-9.0758349435,-10.4016055616,
-16.2275114132\C,-11.0032207558,-9.7232081954,-16.8923445562\C,-11.637
8648783,-10.8658104557,-16.3810211422\H,-11.0422509135,-11.6080582817,
-15.8534932167\C,-13.003503726,-11.0833309846,-16.5627430592\H,-13.488
2196155,-11.9665270216,-16.1551223304\C,-13.7632349429,-10.1445558123,
-17.2707798571\O,-15.0986020453,-10.3939985696,-17.4150740099\H,-15.51
23524365,-9.6620986448,-17.9011716291\C,-13.1460397028,-9.0062965525,-
17.8039453512\H,-13.7314215119,-8.2823733684,-18.365695679\C,-11.77949
48419,-8.8053061424,-17.6144167232\H,-11.315542432,-7.9219373531,-18.0
479052846\C,-9.5967405498,-8.3894782115,-14.3797016025\O,-8.8337005022
,-8.8140549592,-13.5018836335\N,-10.8673823352,-8.0136888415,-14.11131
07724\H,-11.4571064426,-7.6872336609,-14.8691635824\C,-11.3476882406,-
7.8424569001,-12.7387450498\H,-10.5165651648,-7.4352823044,-12.1550641
052\C,-12.4981965238,-6.8097296457,-12.6651579457\H,-12.7506351129,-6.
7215750472,-11.6038436473\H,-12.092712098,-5.8433370038,-12.9821523956
\C,-13.7334006281,-7.1452510183,-13.4759368123\C,-14.6592150372,-8.097
4042677,-13.0176740417\H,-14.4972659528,-8.5846372311,-12.0595575162\C
,-15.7923888754,-8.4095361546,-13.7692261839\H,-16.5003014471,-9.14473
8668,-13.3962357461\C,-16.0241052785,-7.7711943798,-14.9918621357\H,-1
6.9091161752,-8.0121115841,-15.5740235545\C,-15.118416223,-6.814938218
6,-15.4521600425\H,-15.2954536166,-6.3029735791,-16.3939513091\C,-13.9
829446161,-6.5044822609,-14.6975706893\H,-13.2930043885,-5.742914132,-
15.0550247731\C,-11.6892085524,-9.1423550239,-11.9872817625\O,-11.8743
425706,-9.0891676901,-10.7717517816\N,-11.7188591452,-10.3000844549,-1
2.6975622374\H,-11.6370808614,-10.2567181232,-13.7076084038\C,-11.7956
375268,-11.5985233056,-12.0392761616\H,-12.2517952741,-11.4188293553,-
11.0609558448\C,-12.6513227593,-12.609222876,-12.8197455566\H,-12.2494
335828,-12.7453284728,-13.8298370122\H,-12.5529254999,-13.5714761726,-
12.3087662447\C,-14.1286788577,-12.2088499152,-12.9049751457\H,-14.557
1810979,-12.1199012636,-11.9002325468\H,-14.2251356241,-11.2229118877,
-13.3781245518\C,-14.9314478253,-13.1911720002,-13.7519022182\O,-14.49
88357541,-13.6521426661,-14.8123793019\N,-16.1578607399,-13.5115716488
,-13.2751516107\H,-16.7536075213,-14.1238645196,-13.8166284057\H,-16.5
179287827,-13.1414909668,-12.4078056841\C,-10.4074945114,-12.210737152
8,-11.7311839683\O,-10.3322737665,-13.2712608892,-11.110937336\N,-9.32

Supporting Information

97540584,-11.5160345524,-12.1774072926\H,-9.4780086904,-10.6599440727,
-12.7009430424\C,-7.9635200384,-11.9752918257,-11.9890587593\H,-8.0293
19482,-12.9762661512,-11.5611641328\C,-7.1946910342,-12.0173121028,-13
.3325401676\H,-7.1541582613,-11.0171630548,-13.770066191\H,-6.16908419
36,-12.3483219697,-13.1356807214\C,-7.8756015796,-12.9349328128,-14.34
51902503\O,-8.6192309814,-12.4994700529,-15.2278291764\N,-7.6356407107
,-14.2580567544,-14.187564745\H,-8.0737846834,-14.9183431663,-14.81604
71502\H,-6.9936084718,-14.6175853385,-13.4957886861\C,-7.1892310303,-1
1.1421468423,-10.9428825122\O,-6.5765957924,-11.7031115191,-10.0352298
682\N,-7.2014861167,-9.7901538361,-11.1170879648\H,-7.7935133544,-9.37
32304752,-11.8311918714\C,-6.6100208541,-8.8951457605,-10.1323138753\H
,-6.0173889475,-9.5274517724,-9.4739439546\C,-5.7110280082,-7.81051706
44,-10.747054146\H,-6.2801670908,-7.1379447389,-11.3907026509\H,-5.280
3419659,-7.2056705306,-9.9411634413\S,-4.2506392382,-8.4325927305,-11.
6831496068\C,-7.7591078363,-8.203358775,-9.3701570951\O,-8.5421511667,
-7.4537917616,-9.9850158165\N,-7.9202176724,-8.4648884395,-8.064675439
3\C,-7.0399898429,-9.2726007907,-7.1880631605\H,-5.9949272361,-8.98090
42852,-7.3144342891\H,-7.1406804158,-10.3343534051,-7.4374787963\C,-7.
5593202313,-8.9573538026,-5.7783651119\H,-7.0429711432,-8.0777906074,-
5.3821196467\H,-7.3978923077,-9.7900950474,-5.0894885593\C,-9.04680697
83,-8.6422112753,-5.9976870977\H,-9.4884637614,-8.0559689085,-5.187584
8624\H,-9.6271359605,-9.5649869551,-6.1011956978\C,-9.0617160244,-7.88
62028897,-7.3459703896\H,-9.9785435505,-8.0691289553,-7.9133648856\C,-
8.8717063525,-6.3683559595,-7.1637130908\O,-7.7569774383,-5.8481199926
,-7.1108972137\N,-10.0332833278,-5.67303761,-7.0645675997\H,-10.907221
8015,-6.1818095582,-6.9826609674\C,-10.1072026736,-4.2759235101,-6.661
0432904\H,-9.0880581415,-3.9507169097,-6.4344316362\C,-10.7323427033,-
3.3686525045,-7.7522125802\H,-11.623117456,-3.8742710103,-8.1461547679
\H,-11.0896253774,-2.4545601032,-7.2642503114\C,-9.809308724,-2.951529
566,-8.9065348162\H,-10.3728234145,-2.25068606,-9.5357515237\H,-8.9481
354688,-2.3987165396,-8.5085650298\C,-9.3098972482,-4.0979728641,-9.79
09902416\H,-8.5810686701,-4.7144091024,-9.2578514257\H,-10.1438883228,
-4.7412967762,-10.0921795566\N,-8.6939047148,-3.5396161745,-11.0039377
516\H,-8.4822621397,-2.5500413039,-11.0109205966\C,-8.3929992618,-4.23
76244058,-12.1018913514\N,-8.5191138566,-5.5696920106,-12.1294311461\H
,-8.5274978346,-6.1421562167,-11.2805095189\H,-8.2872107,-6.0429594403
,-12.9976400969\N,-7.9570130934,-3.6001763649,-13.2040419614\H,-8.1429
477906,-2.6151449259,-13.3298175196\H,-7.7789098529,-4.1491042245,-14.
0369149485\C,-10.9950080061,-4.1933165636,-5.4080343992\O,-11.93588223
42,-4.9712643471,-5.2475256702\N,-10.6756495596,-3.2056353013,-4.53335
96156\H,-9.9571941926,-2.5468501103,-4.804918686\C,-11.5897075466,-2.7
885649613,-3.4950750781\H,-12.1556716258,-3.6651754814,-3.1681520646\H
,-11.0243808297,-2.414573019,-2.636120128\C,-12.5754636271,-1.71324973
17,-3.9924968712\O,-12.6594270079,-1.3983425274,-5.1772295753\N,-13.33
34717341,-1.1585542887,-3.0185064536\H,-13.2740466164,-1.4462978671,-2
.0517802751\H,-14.0404658222,-0.4800677065,-3.2678130753\\Version=EM64
L-G09RevC.01\State=1-A\HF=-4313.5939836\RMSD=4.734e-09\RMSF=5.834e-06\
Dipole=8.4695476,12.8897308,-8.0012343\Quadrupole=-12.8841068,-43.2331
016,56.1172084,44.9917112,-88.8808838,-2.3231117\PG=C01 [X(C46H67N15O1
2S2)]\\@

***Clinched open*, folded:**

1\1\FAU-CCC-CCDH172\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\09-F
eb-2016\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP clinched open
 with folded tail conformation\\2,1\N,-4.6682334799,8.0993140401,13.31
31705382\H,-5.4455683387,7.8609298261,13.9367649039\H,-3.9137849027,8.
5642922186,13.8529211236\H,-4.287253024,7.2210351736,12.9444587875\C,-
5.0582736647,9.047799715,12.2053997212\H,-5.3511784186,8.4449303268,11
.3445834089\C,-6.2289566695,9.9295751679,12.6766037168\H,-7.0091715035
,9.2969004206,13.1077449991\H,-5.8921985738,10.6599018823,13.417708400
4\S,-7.0099476305,10.8208125414,11.2489347616\C,-3.7699108136,9.862774
0526,11.947072403\O,-2.9370252525,9.9615305362,12.8496246607\N,-3.6551
614883,10.4664836688,10.7498184442\H,-4.3053031726,10.2279471715,10.00
92802115\C,-2.4239397414,11.1858245286,10.3940607696\H,-1.5733278493,1
0.5682264845,10.6947407768\C,-2.3834567456,11.4005907394,8.8590818851\
H,-2.6811898805,10.4512210157,8.396441109\H,-3.1375293703,12.145397884
7,8.580511477\C,-1.0223423136,11.8056399072,8.3340509134\C,-0.73756151
34,13.1341960957,7.9930310539\H,-1.5048631877,13.8952011658,8.10496788
87\C,0.5139498014,13.5053464789,7.5061868163\H,0.7251846618,14.5373857
964,7.2427681646\C,1.5147702398,12.5400543621,7.3508047987\O,2.7250664
709,12.9549325292,6.8673814166\H,3.3240159891,12.1928134868,6.80623601
2\C,1.2490496235,11.2060571255,7.6827860293\H,2.0199679729,10.44904631
83,7.5567171748\C,-0.00968363,10.8514307881,8.1663365243\H,-0.20332481

Supporting Information

23,9.8091956483,8.4107661371\C,-2.2573846566,12.5043864906,11.17770635
79\O,-1.1327157912,12.9177546425,11.4490449676\N,-3.4040991886,13.1686
824344,11.479775856\H,-4.290686099,12.691286644,11.3619228817\C,-3.431
7071847,14.3165196902,12.3662821095\H,-2.4125741241,14.4644901369,12.7
333542365\C,-3.9262670584,15.6144466131,11.6591180015\H,-4.9742993887,
15.4735231636,11.3697353027\H,-3.9068613291,16.4209949098,12.402431037
6\C,-3.1122427775,16.0364254437,10.4554185252\C,-3.6274224787,15.89896
25091,9.1599342777\H,-4.6130301916,15.4613719437,9.0207759923\C,-2.899
4417098,16.3347276396,8.0491151438\H,-3.3210814243,16.2296398646,7.053
001413\C,-1.6382080258,16.9098229347,8.22026056\H,-1.0729192225,17.255
4921415,7.3590849411\C,-1.1104798182,17.0443720008,9.5080274139\H,-0.1
317835238,17.4943204874,9.6514692509\C,-1.8437150951,16.6136853996,10.
6143139681\H,-1.4312505417,16.7385786474,11.6131399887\C,-4.3932043631
,14.0037319264,13.5284958171\O,-5.2893281039,13.1654274857,13.40810032
66\N,-4.2611667208,14.7683678712,14.6440577408\H,-3.4666108482,15.3900
04613,14.7322659759\C,-5.2087890894,14.6531526134,15.7520611742\H,-5.3
844292729,13.5944510175,15.9533050968\C,-4.6736809637,15.3262476099,17
.0278315136\H,-4.4643826201,16.3831768723,16.8247937515\H,-5.475673131
2,15.2899622246,17.7701004532\C,-3.4243978184,14.6618900475,17.6153033
574\H,-2.6089163754,14.610630812,16.8841023719\H,-3.0492723824,15.2836
552509,18.4387145236\C,-3.6941925394,13.2659629408,18.1797870601\O,-4.
8311622434,12.8464223525,18.4006806993\N,-2.5849203604,12.5238664493,1
8.4252653948\H,-2.6862918005,11.6369956969,18.9006330501\H,-1.65342815
77,12.9036971634,18.3318217957\C,-6.5943947396,15.2219831457,15.383888
2557\O,-7.6208099417,14.7493910912,15.8818836014\N,-6.6011086652,16.27
197177,14.5272429111\H,-5.7278703555,16.5317715418,14.086984881\C,-7.8
34811949,16.8430568549,13.9856479062\H,-8.5869745613,16.8105520262,14.
7759472663\C,-7.608321904,18.2849119199,13.5468277475\H,-7.1617648998,
18.8490547293,14.3743719965\H,-6.9069317085,18.3204513481,12.706662018
5\C,-8.9233146417,18.9511597396,13.1465406458\O,-10.0176327374,18.4392
950454,13.3870313009\N,-8.79433576,20.1443687558,12.5194529329\H,-9.62
60884214,20.6693368131,12.2851278979\H,-7.8970596321,20.5815400148,12.
3646418039\C,-8.3032911123,15.9591870406,12.806661352\O,-8.0788615555,
16.2484922794,11.6310435825\N,-8.9136680393,14.8114913885,13.197297656
8\H,-8.9742090959,14.6276348889,14.1953187656\C,-9.2059515061,13.72669
02771,12.2628461332\H,-8.4924147025,13.8152005888,11.4429312296\C,-9.0
198543131,12.3888264403,12.9869809439\H,-8.0915924424,12.4016438827,13
.5620507188\H,-9.8567152059,12.2127636137,13.6706456984\S,-8.997115672
,10.9222409011,11.86617347\C,-10.6603650972,13.8913923766,11.767990087
1\O,-11.592376258,13.7927932489,12.5823855818\N,-10.8691348271,14.2049
80273,10.4741728999\C,-9.8858569296,14.2072194102,9.3682761444\H,-9.23
433685,13.3335137938,9.4330444917\H,-9.2712439035,15.1125487793,9.4199
0964\C,-10.7748882654,14.1738573792,8.1201870553\H,-11.0476583535,13.1
394620093,7.889840425\H,-10.2751124271,14.60362596,7.2482858808\C,-12.
0211299134,14.9655247257,8.5440956449\H,-12.9031179603,14.7368890817,7
.9400986978\H,-11.8368171466,16.0423058254,8.4734992806\C,-12.21740746
12,14.5877237678,10.0317883336\H,-12.5545035018,15.4454900943,10.62219
95844\C,-13.2006561815,13.4233775548,10.2248373371\O,-12.8984280776,12
.2622674603,9.9317100835\N,-14.402122238,13.7884625339,10.7325686793\H
,-14.5303480043,14.7662189527,10.9563804071\C,-15.5037201196,12.892880
8653,11.0984997896\H,-16.1457625645,13.504902242,11.743130042\C,-16.39
52795153,12.4425970659,9.9242384393\H,-17.2652922404,11.9451076858,10.
3662189911\H,-16.7646015001,13.3507553497,9.4341727641\C,-15.758776714
,11.5027481905,8.8921492902\H,-14.889999737,11.9685043151,8.4153671067
\H,-15.3903046263,10.6029697236,9.3931241967\C,-16.7481678482,11.04950
22783,7.8148502498\H,-16.2920667211,10.273194696,7.188474528\H,-17.644
1645883,10.6284734729,8.2826155541\N,-17.1645612293,12.1824019987,6.97
55374188\H,-16.4823169182,12.9141027307,6.8182579454\C,-18.3034290738,
12.2597927825,6.2819015618\N,-19.1330461587,11.2110002926,6.217907566\
H,-18.8004316725,10.2750782633,6.3998588584\H,-20.0295308186,11.293737
5201,5.7592567127\N,-18.6242873386,13.4024517914,5.6571799586\H,-18.19
39272157,14.2762130013,5.9271122494\H,-19.4514034455,13.4591171818,5.0
800779077\C,-15.0663174955,11.7182555966,12.000855228\O,-15.7361364557
,10.6850844169,12.0491030467\N,-13.9816869185,11.9224565565,12.7890243
235\H,-13.3956339526,12.7449733928,12.6699409801\C,-13.5056416607,10.8
744533884,13.6692773544\H,-12.8198105603,11.321392088,14.3947718291\H,
-14.342446348,10.4418269488,14.2228168529\C,-12.7686729154,9.704517182
8,12.9976678454\O,-12.5320672351,8.6842689788,13.6463701879\N,-12.3833
909922,9.8871476241,11.7159987309\H,-12.6016210048,10.722883909,11.178
6756797\H,-11.8827108197,9.1376263655,11.2579064837\\Version=ES64L-G09
RevD.01\State=1-A\HF=-4313.5754649\RMSD=7.546e-09\RMSF=2.142e-06\Dipol
e=-9.3182369,-3.6441124,-8.9227906\Quadrupole=207.4827637,-101.2244415
,-106.2583222,-49.7858048,259.7104929,20.5083893\PG=C01 [X(C46H67N15O1
2S2)]\\@

Supporting Information

***Clinched open*, extended:**

1\1\FAU-CCC-CCDH172\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\23-A
pr-2015\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP_10us_T16_12_c
l_open\\2,1\N,-9.3435783672,-12.4949768434,2.5153303585\H,-9.694391987
4,-13.1779261134,1.8211119347\H,-10.137759484,-11.9119649439,2.7978609
187\H,-8.9856452812,-12.9745120635,3.3534778048\C,-8.2632699003,-11.69
26504466,1.8506910731\H,-8.2705617264,-10.696151436,2.294555961\C,-6.8
837520222,-12.3633918603,2.0156760588\H,-6.8257688226,-13.2805771336,1
.4253666584\H,-6.1235026876,-11.6700761369,1.6497814328\S,-6.485045714
3,-12.7753040285,3.7803917189\C,-8.6556394425,-11.663084404,0.35477259
34\O,-9.442834375,-12.5063183906,-0.0764313241\N,-8.0492226972,-10.737
2055623,-0.4108728957\H,-7.5003032481,-10.0074763586,0.0288637766\C,-8
.3507957576,-10.6218631595,-1.8434661608\H,-9.433694848,-10.5390852481
,-1.9716805145\C,-7.661840712,-9.3504767688,-2.3978935499\H,-7.9734032
038,-8.5126777275,-1.7615128041\H,-6.5768601371,-9.4567555351,-2.28250
46561\C,-8.0037903426,-9.049051218,-3.8410226336\C,-7.1201788121,-9.36
99683064,-4.880816833\H,-6.1669696601,-9.8395773539,-4.6499570604\C,-7
.4329860119,-9.0948366248,-6.2103359879\H,-6.7420075617,-9.3422812609,
-7.0105116634\C,-8.6527322826,-8.4854237716,-6.5252188215\O,-8.9120654
424,-8.234211565,-7.8438523201\H,-9.7824642485,-7.810723148,-7.9250062
723\C,-9.5474381556,-8.1532546391,-5.5008769667\H,-10.4929458973,-7.67
13905694,-5.7387562383\C,-9.2174120299,-8.4340247781,-4.1758667164\H,-
9.9184248048,-8.1618094597,-3.390081482\C,-7.9472897027,-11.8945969526
,-2.6186675536\O,-8.6491823811,-12.322762511,-3.5294435316\N,-6.751616
3134,-12.4378376066,-2.2581203596\H,-6.3144371123,-12.0968330395,-1.41
11013975\C,-6.3193791154,-13.7513446303,-2.7012723588\H,-6.9059386829,
-13.9880814192,-3.5917154799\C,-4.8126599214,-13.7527597315,-3.0966154
778\H,-4.5268960777,-14.7826970217,-3.3387146214\H,-4.7378963989,-13.1
738562702,-4.0231704128\C,-3.8724796663,-13.1723240148,-2.0585817905\C
,-3.2588287184,-13.9830613242,-1.0933461534\H,-3.4353038451,-15.054145
3147,-1.0852429002\C,-2.4022355179,-13.436853137,-0.1354771533\H,-1.94
12037425,-14.0951724657,0.594847959\C,-2.1428733201,-12.0639265999,-0.
1266055002\H,-1.4741977041,-11.6375694373,0.6163378135\C,-2.7404271543
,-11.24458692,-1.0869710781\H,-2.5372938635,-10.1770244323,-1.09739859
01\C,-3.594362353,-11.7961886362,-2.0452108713\H,-4.0433967374,-11.153
2434477,-2.798350334\C,-6.6427508728,-14.7677921034,-1.5894493133\O,-6
.5630657494,-14.4727911578,-0.3979463884\N,-7.012278106,-16.0193101805
,-1.9926502619\H,-7.1994018351,-16.1682057147,-2.9764826851\C,-7.63629
98611,-16.9347958272,-1.0375372291\H,-8.2666373312,-16.3383700024,-0.3
69427189\C,-8.5358652487,-17.9685670983,-1.7429211058\H,-9.0253071056,
-18.5665353424,-0.9703519707\H,-9.3284720359,-17.4212269878,-2.2658368
579\C,-7.8254396097,-18.9174571044,-2.7161826912\H,-7.4189631193,-18.3
857266926,-3.5839140537\H,-6.9729044273,-19.4045836409,-2.2240303836\C
,-8.7649118842,-20.0314588911,-3.1839296409\O,-9.6809402273,-20.456616
7731,-2.4800959841\N,-8.5050124718,-20.5252631753,-4.419591129\H,-9.02
84653765,-21.3261840502,-4.7472817129\H,-7.7162153618,-20.2184058851,-
4.9702468876\C,-6.6375840192,-17.6282262506,-0.0990620009\O,-7.0490308
996,-18.1234921974,0.9647070557\N,-5.3589543873,-17.7054604215,-0.5044
245002\H,-5.1172286238,-17.240955376,-1.3722607612\C,-4.3009557169,-18
.3614104429,0.265520001\H,-4.7787534082,-19.1234226302,0.8827481359\C,
-3.2953772326,-19.0202638288,-0.6756855734\H,-3.8298845266,-19.6482281
705,-1.3975808388\H,-2.7508043772,-18.2562755949,-1.2411611748\C,-2.30
10559253,-19.8796248561,0.1032133477\O,-2.3989200201,-20.0518117631,1.
3192942622\N,-1.3177563658,-20.4363532038,-0.6409733303\H,-0.640170377
3,-21.0383353717,-0.1929320713\H,-1.2527900089,-20.3061849217,-1.64044
23222\C,-3.6030480579,-17.3276931838,1.1823896445\O,-2.6058377048,-16.
7036880636,0.8193450618\N,-4.2088472346,-17.141917444,2.3792195914\H,-
4.9799334821,-17.7382954429,2.6494156477\C,-3.6788976679,-16.273853645
2,3.4161793005\H,-2.6079239531,-16.1601954914,3.2251400695\C,-4.275234
3936,-14.8464946642,3.3677731086\H,-3.7684768342,-14.1964532142,4.0843
464315\H,-4.1274428566,-14.4481764025,2.3615235454\S,-6.0877466545,-14
.8212676908,3.7642276456\C,-3.922198901,-16.9751102738,4.7669324101\O,
-4.9049537338,-17.7254094348,4.8982743606\N,-3.0589303322,-16.74619717
79,5.7693332339\C,-1.8511459651,-15.8918696519,5.7476761376\H,-2.04501
73323,-14.9424541443,5.2459325535\H,-1.0386023152,-16.4072246258,5.221
2714726\C,-1.5310031459,-15.7103510971,7.2352770505\H,-2.1226387188,-1
4.8842657012,7.6412278942\H,-0.472656068,-15.4957707205,7.4025111856\C
,-1.980206716,-17.0403618387,7.857997365\H,-2.1628410488,-16.974068880
1,8.9335508018\H,-1.2268289904,-17.8173586823,7.6919519787\C,-3.262114
3132,-17.4010834079,7.0707191659\H,-3.3593601658,-18.4802260698,6.9216
061887\C,-4.536336925,-16.8643396032,7.7487488756\O,-4.9078245324,-15.
6983620451,7.6147857195\N,-5.1819152691,-17.7800265345,8.5098901124\H,

Supporting Information

-4.7975947571,-18.7102628522,8.6263391978\C,-6.41003242,-17.5121331267
,9.239051148\H,-6.5302751141,-16.4267456375,9.2848365681\C,-7.62963612
49,-18.158284796,8.5359824771\H,-7.4172954959,-19.2310759924,8.4344304
03\H,-8.5036501134,-18.0746748481,9.1931667268\C,-7.9319421489,-17.547
4565544,7.1611505699\H,-8.3745383606,-16.5510507656,7.2808301856\H,-7.
0018940286,-17.4183505993,6.5973501471\C,-8.8906961675,-18.4408858717,
6.3628231391\H,-8.4364326524,-19.4186969684,6.1733725401\H,-9.81091706
39,-18.6157205937,6.9276914232\N,-9.2954638628,-17.8743669169,5.071257
6459\H,-10.1336120417,-17.3079585699,5.0534173659\C,-8.6935905986,-18.
1021815836,3.8988040903\N,-7.4969263392,-18.7222051623,3.844490038\H,-
6.7924842302,-18.5158197539,4.5476857932\H,-7.1183854593,-18.845852347
7,2.9094047276\N,-9.2884682641,-17.7594583712,2.7525739524\H,-10.26939
78288,-17.5226164764,2.7150135435\H,-8.7469676304,-17.8072159832,1.889
1350585\C,-6.2798588702,-18.1163065079,10.6428512316\O,-5.6651680774,-
19.1686044428,10.8211063758\N,-6.900297617,-17.4276340844,11.634309282
1\H,-7.4758691547,-16.6374292213,11.3734456094\C,-7.1414388091,-18.018
0229179,12.9305920101\H,-6.3232757073,-18.711403165,13.1440339298\H,-7
.1304883899,-17.2387529481,13.6988635488\C,-8.4753234882,-18.789031182
,12.9738280657\O,-9.1410270253,-19.0047519619,11.9642805376\N,-8.83855
29665,-19.2022184436,14.2103740872\H,-8.2616531748,-19.0482034484,15.0
253856991\H,-9.6715021997,-19.7648823499,14.3193959866\\Version=EM64L-
G09RevC.01\State=1-A\HF=-4313.5925\RMSD=4.425e-09\RMSF=2.938e-06\Dipol
e=-4.2029982,2.790374,2.1439386\Quadrupole=-50.5993335,-28.3678385,78.
967172,-45.4992104,-34.5696073,1.9057859\PG=C01 [X(C46H67N15O12S2)]\\@

***Twisted saddle*, extended:**

1\1\FAU-CCC-CCDH173\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\23-A
pr-2015\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP_10us_T16_19_t
w_saddle\\2,1\N,6.1686704246,7.3143387257,-9.0049356261\H,6.2464486616
,7.0336633398,-9.9968301557\H,5.7786842663,6.5201520252,-8.4846886719\
H,7.1046279972,7.4980372998,-8.6285305788\C,5.2619704319,8.5217194046,
-8.9562854103\H,4.8465765776,8.5965721412,-7.9506218898\C,6.1083113637
,9.7577187113,-9.30753096\H,6.8292900672,9.9488660217,-8.5091476943\H,
6.6402670504,9.5923032143,-10.2485309433\S,5.0259532188,11.2533136141,
-9.4807022099\C,4.1845937524,8.2296067293,-10.0260006242\O,4.547806786
5,7.7282614454,-11.0946008937\N,2.9344861728,8.5616722227,-9.681267642
6\H,2.7628448765,8.9023152245,-8.7239647619\C,1.8075514336,8.444611962
,-10.5958956128\H,2.1421517805,7.8230968338,-11.4318538812\C,0.6261057
682,7.734436115,-9.8679485066\H,1.0757375305,6.9018978986,-9.315233947
5\H,0.2031188209,8.4142971504,-9.1203589226\C,-0.4764921979,7.20312220
71,-10.7622793046\C,-1.7572573522,7.7736152701,-10.7582864448\H,-1.959
2971356,8.6354491118,-10.1272657604\C,-2.781175117,7.2685155079,-11.56
18238894\H,-3.7674038676,7.7259129528,-11.5405023952\C,-2.535242593,6.
1693008044,-12.3912652802\O,-3.4884893434,5.6243728648,-13.2019124264\
H,-4.3193891404,6.1173783273,-13.1003313265\C,-1.2622169969,5.58331912
7,-12.4083493121\H,-1.0825727588,4.7270654187,-13.0511547541\C,-0.2528
79153,6.0973100565,-11.6008452398\H,0.7254097731,5.622203492,-11.61636
39265\C,1.4552197467,9.849000539,-11.1519444706\O,2.0940408393,10.8564
627898,-10.821750848\N,0.4210594744,9.9164403827,-12.012025794\H,-0.12
09957964,9.0817172873,-12.2018124635\C,-0.0402056301,11.1832364292,-12
.5794147029\H,0.7795796467,11.6219773572,-13.1587061823\C,-1.234508577
3,10.872671172,-13.5074813744\H,-0.9105082139,10.0772763725,-14.191322
0257\H,-2.0404875873,10.4603754617,-12.8912003976\C,-1.7550666352,12.0
434055904,-14.3161152504\C,-0.9521632747,12.6685458545,-15.2840089848\
H,0.0705317054,12.3312531994,-15.4361131257\C,-1.4565198249,13.7050378
51,-16.069256914\H,-0.8222444926,14.1737286774,-16.816551577\C,-2.7813
736156,14.1260983848,-15.9108205579\H,-3.1819481989,14.9158457285,-16.
5410197989\C,-3.5895295564,13.5118716365,-14.9510085854\H,-4.622785239
4,13.8236823966,-14.8248600303\C,-3.074572036,12.4841450229,-14.155011
8869\H,-3.7108840037,12.007253571,-13.4138163624\C,-0.4550572185,12.15
33249846,-11.4469118388\O,-1.2679842464,11.8241154289,-10.5843956494\N
,0.1394659204,13.3742185679,-11.4827228789\H,0.8236571015,13.55389942,
-12.2046164434\C,-0.1482233007,14.4526014188,-10.5318127874\H,0.599087
253,15.2268645981,-10.7463968732\C,-1.5413113039,15.07337958,-10.69916
14819\H,-2.3018964068,14.2977948145,-10.5642350193\H,-1.6826608052,15.
803199899,-9.899821288\C,-1.7164899908,15.769132283,-12.051476999\H,-0
.8758225059,16.4569577777,-12.2254366429\H,-1.704855553,15.047153562,-
12.8742414937\C,-2.9863579968,16.6179612254,-12.0960904342\O,-3.394283
1936,17.2469920497,-11.1184262775\N,-3.6026775383,16.6698608748,-13.30
35781465\H,-4.4533671758,17.2085372281,-13.3988960867\H,-3.3247043537,
16.0755542357,-14.0721994406\C,0.1141091075,14.0386690055,-9.062922771
5\O,-0.5404892668,14.5091779877,-8.1377828084\N,1.1899563344,13.2149646
282,-8.87135928\H,1.5432983442,12.7108298051,-9.6820831962\C,1.4736978

Supporting Information

455,12.6296270061,-7.5663144636\H,0.8022258138,13.1294054514,-6.860781
1748\C,1.1443171023,11.1149627178,-7.636706688\H,0.0529635942,11.02061
21666,-7.6589970084\H,1.5246401477,10.7576007347,-8.5916287725\C,1.706
8357188,10.167203993,-6.5852786112\O,2.5310284635,9.2936172363,-6.9037
890289\N,1.1737326302,10.2204551975,-5.3464743503\H,1.6044182346,9.662
4953176,-4.6199132405\H,0.6476796291,11.0291732312,-5.0448283274\C,2.8
938486735,12.925338712,-7.0512938189\O,3.3139766146,12.3650675594,-6.0
356789023\N,3.603872435,13.8563202652,-7.7222679986\H,3.2217378008,14.
3115292581,-8.5441260645\C,4.8795621816,14.3486593707,-7.2385469858\H,
4.8005367027,14.5030440481,-6.156357749\C,6.0391054488,13.3382009289,-
7.4384059974\H,6.9836316057,13.7718119146,-7.0977698729\H,5.8374144728
,12.4452659039,-6.8447436482\S,6.3664203438,12.8231653599,-9.177369641
\C,5.1556040372,15.6777345581,-7.9573660789\O,4.6501763437,15.89332166
76,-9.06810784\N,5.9869705804,16.5566877034,-7.3609433675\C,6.58606141
95,16.4780932281,-6.0116613314\H,6.9599745225,15.4763319161,-5.7927928
892\H,5.8350979395,16.7390002796,-5.2556787163\C,7.7123725065,17.51699
79921,-6.0677083893\H,8.6148576427,17.0588934331,-6.4831955992\H,7.949
3678104,17.9175132402,-5.0788698206\C,7.1677553651,18.5809320691,-7.03
26373483\H,7.9492394216,19.2051895615,-7.473686601\H,6.4591168028,19.2
403337129,-6.5204914186\C,6.4102021428,17.7543991154,-8.0968578396\H,5
.534379806,18.2837564076,-8.4820930648\C,7.3196874824,17.3404490664,-9
.2701511479\O,8.0674400292,16.363013482,-9.1987064017\N,7.2408499864,1
8.1442061791,-10.3563386326\H,6.6417045534,18.961712109,-10.3555445375
\C,8.0321474726,17.9425523002,-11.5588945571\H,8.9033233328,17.3444067
909,-11.2786300375\C,7.2343734979,17.2101928864,-12.6672307047\H,6.398
6972121,17.8536567571,-12.9712765143\H,7.884502922,17.1003439761,-13.5
438128205\C,6.7006883436,15.8470427112,-12.2142114136\H,7.5264494742,1
5.2091735405,-11.8791753068\H,6.0310454116,15.9865340673,-11.357935469
7\C,5.9346889831,15.1393838761,-13.3323688127\H,5.1634446162,15.803473
699,-13.7438677856\H,6.6209512385,14.8669750775,-14.1425572244\N,5.304
4320859,13.920549972,-12.8070245024\H,5.2686565694,13.8159870182,-11.7
99209256\C,4.6711815782,13.0082950418,-13.5501894391\N,4.6603643309,13
.1166895704,-14.8866432513\H,4.984010484,13.9558879441,-15.3439745156\
H,4.1104950076,12.4820405318,-15.4482409547\N,4.0453148877,11.97700177
48,-12.9677212706\H,4.1229909486,11.793327238,-11.9723866553\H,3.67041
08262,11.2260429449,-13.5293904958\C,8.4543571899,19.3225538005,-12.07
95253168\O,7.716251286,20.299737373,-11.9447981197\N,9.6621926574,19.3
713424748,-12.6941967443\H,10.1551055807,18.5017162025,-12.8495932966\
C,10.0641476394,20.5113200682,-13.4849670067\H,9.6505831339,21.4114910
813,-13.0218096017\H,11.154646231,20.5984728243,-13.4727345877\C,9.553
587063,20.4173150476,-14.9363976986\O,8.7593972366,19.5528482329,-15.2
968046416\N,10.0528752357,21.3659614475,-15.7628596571\H,10.6747198243
,22.0941619843,-15.440432973\H,9.724240805,21.4127513519,-16.717917836
7\\Version=EM64L-G09RevC.01\State=1-A\HF=-4313.5804546\RMSD=3.550e-09\
RMSF=4.995e-06\Dipole=9.4610691,-7.3547137,-3.7051286\Quadrupole=-5.21
16244,29.149813,-23.9381885,41.9635208,27.3753322,-55.7098266\PG=C01 [
X(C46H67N15O12S2)]\\@

***Open*, extended:**

1\1\FAU-CCC-CCDH173\FOpt\RB3LYP\6-31G(d)\C46H67N15O12S2(2+)\CLARK\24-A
pr-2015\0\\# b3lyp/6-31g(d) opt name=clark scrf=pcm\\AVP_10us_T27_open
\\2,1\N,0.4475246287,-3.1779434367,-3.5701330464\H,-0.5341706603,-2.88
25080235,-3.5394377666\H,0.630961857,-3.882061537,-2.8241126635\H,0.60
74771251,-3.6323651944,-4.4749533667\C,1.4175750482,-2.0519324222,-3.3
157507722\H,1.8408115906,-1.7474456493,-4.2748450585\C,0.7357973228,-0
.8704294347,-2.6217975522\H,0.2391520442,-1.1935284468,-1.7028613464\H
,1.501583425,-0.1383286191,-2.3667457184\S,-0.4811951036,-0.038392031,
-3.7510888461\C,2.5086337441,-2.6800778896,-2.403645719\O,2.2747360951
,-3.7630876102,-1.859017045\N,3.6330024686,-1.9669722602,-2.2553773879
\H,3.7734897951,-1.1160030541,-2.7988812368\C,4.7570159515,-2.44729703
67,-1.4427791763\H,4.3336708144,-3.0453028042,-0.6330636844\C,5.717857
0215,-3.3266355651,-2.2863890204\H,5.1115897125,-4.1249763025,-2.72746
94894\H,6.1025607642,-2.7124058126,-3.1072051567\C,6.8572977336,-3.912
2661588,-1.4804746611\C,8.1236575811,-3.3085158824,-1.4665770637\H,8.2
967516351,-2.4143144137,-2.0603180018\C,9.1699249165,-3.8324500002,-0.
7099168111\H,10.149706574,-3.3645238628,-0.7098458205\C,8.9635759022,-
4.9855950103,0.056655234\O,10.0226139205,-5.4610247138,0.7780304264\H,
9.751831831,-6.260301344,1.2588757938\C,7.7080002354,-5.6048876844,0.0
548078818\H,7.5442005547,-6.5055563399,0.6421351785\C,6.6721104291,-5.
0674377862,-0.7086302563\H,5.7045114237,-5.564372769,-0.7067273919\C,5
.4841973261,-1.2092021754,-0.8973177933\O,5.8874285518,-0.3325023244,-
1.6711781046\N,5.6360006746,-1.1372611073,0.4415144527\H,5.3328351703,
-1.9204926778,1.0054567213\C,6.3056454754,-0.0202965873,1.1259249479\H

Supporting Information

,6.1366505233,-0.2021023071,2.1921903792\C,7.8316923788,0.0209268586,0
.8708586026\H,8.1630928554,-1.0190952775,0.7755552417\H,8.0181890184,0
.5142110809,-0.0851182503\C,8.617664545,0.6949402828,1.9799764421\C,8.
8219673415,0.035003442,3.2015852701\H,8.413172207,-0.9630895414,3.3476
731897\C,9.5513371764,0.6351427926,4.2297456731\H,9.6988733243,0.10558
93717,5.1673967908\C,10.0975902378,1.9084563233,4.0494740187\H,10.6693
82467,2.3764385655,4.846361945\C,9.9064980668,2.5732280985,2.835918752
5\H,10.3295230468,3.5630851556,2.6846858852\C,9.1709164602,1.971258312
9,1.8125700572\H,9.0115375646,2.4978119205,0.8766700221\C,5.6121259489
,1.3267896745,0.7833752105\O,6.2483926483,2.3465746359,0.5257592758\N,
4.2598270574,1.2752701586,0.818880499\H,3.7893093495,0.4254779758,1.10
88545255\C,3.3717247555,2.3717979928,0.4793353525\H,3.8748552506,3.318
0358983,0.6993801244\C,2.9389690218,2.3139860095,-1.0095353207\H,2.469
0344454,1.3334938809,-1.1529120127\H,2.1647718774,3.0623098184,-1.2060
843099\C,4.0900179819,2.4896144533,-2.0163410839\H,4.3674472743,3.5476
548957,-2.0800637794\H,4.9685514935,1.9297612287,-1.6975243441\C,3.642
6896543,2.0361554422,-3.4008172474\O,2.7531971144,2.6076639299,-4.0285
714959\N,4.2504975086,0.8978073052,-3.863063117\H,4.1017844827,0.65245
05137,-4.8351277785\H,5.0959788629,0.5558184017,-3.4151218523\C,2.1020
884507,2.2013174563,1.3329459598\O,1.8364903835,1.1199678347,1.8584310
224\N,1.2855934288,3.2881432317,1.3987591035\H,1.4984661566,4.08220083
07,0.8047143689\C,-0.1174972069,3.1488890822,1.7652845525\H,-0.2310346
94,2.1502989969,2.1865891728\C,-0.5299344147,4.1815351025,2.8293320033
\H,0.2296797856,4.176215614,3.6202939735\H,-0.5527454927,5.1915456728,
2.4079397698\C,-1.8632763933,3.8241517706,3.4735021903\O,-2.2034632306
,2.636340791,3.6251099513\N,-2.6283668384,4.8463215437,3.8784395015\H,
-3.4986149451,4.6616492843,4.3616619151\H,-2.3513342799,5.8096666714,3
.7503216914\C,-0.9502227728,3.2834431883,0.4699787581\O,-0.612368158,4
.0752978656,-0.4130176333\N,-2.0221969429,2.464951224,0.3763759172\H,-
2.3069723152,1.8985106556,1.1687588278\C,-2.875563971,2.3924754074,-0.
7960159985\H,-2.8110632028,3.3482698423,-1.3212158436\C,-2.461338766,1
.2242902853,-1.7318036293\H,-2.2615448956,0.3376054787,-1.1257549022\H
,-3.2586223275,0.9910466781,-2.4403727783\S,-0.9373094554,1.6855799768
,-2.6801100439\C,-4.3164837332,2.1162629512,-0.3240896762\O,-4.5012526
706,1.5374297738,0.7559171037\N,-5.3247983585,2.4675891667,-1.14257154
21\C,-5.2375697502,3.1700694204,-2.4412659491\H,-4.3787038897,2.832587
7782,-3.0243546128\H,-5.1396179145,4.2494348818,-2.2733074876\C,-6.570
7649481,2.8236083035,-3.1140483819\H,-6.4844782152,1.8629176533,-3.631
1787976\H,-6.8716387024,3.5822117659,-3.8407227307\C,-7.5444275529,2.7
048132111,-1.9319238325\H,-8.439637445,2.1228735552,-2.1658679576\H,-7
.8662696838,3.6961649824,-1.5962719125\C,-6.6941635956,2.0454411875,-0
.8212507711\H,-6.970506235,2.4066312154,0.1733814727\C,-6.7804043101,0
.5045499748,-0.8411424942\O,-6.0489909969,-0.1826236574,-1.554489992\N
,-7.7271660559,-0.0068071046,-0.0180201572\H,-8.3406601696,0.610695250
2,0.5006633364\C,-7.9648897054,-1.4295355952,0.1622268142\H,-7.4405838
189,-1.9487606443,-0.6440292191\C,-7.4479723582,-1.911138949,1.5417185
554\H,-7.9737394792,-1.3248195035,2.3075731259\H,-7.7509903149,-2.9547
001548,1.6879376702\C,-5.9311609079,-1.7527911339,1.7138324089\H,-5.39
87196437,-2.5230228803,1.1427464207\H,-5.6118920103,-0.7859345168,1.31
03955172\C,-5.5285193128,-1.8386955815,3.1936681367\H,-6.1001515623,-1
.1182022868,3.7875804772\H,-5.742945873,-2.8320825398,3.5970134613\N,-
4.1043217452,-1.5821568139,3.4343772738\H,-3.4744395234,-2.3733653049,
3.4212752543\C,-3.5760672573,-0.3776112172,3.6782175028\N,-4.326985093
8,0.7376331286,3.5691066237\H,-4.9791612042,0.7840660111,2.7939234501\
H,-3.8156436785,1.611934258,3.7073049088\N,-2.3080376119,-0.2626342817
,4.0892700438\H,-1.7663206603,-1.0649586896,4.3769230718\H,-1.87118036
13,0.6543387416,4.0912541412\C,-9.4766109097,-1.6817850965,0.098904131
2\O,-10.2709349096,-0.859522806,0.5569318343\N,-9.8503989609,-2.860130
0053,-0.461419201\H,-9.1238481642,-3.5204160982,-0.7062994789\C,-11.18
8703029,-3.3810458135,-0.3048644933\H,-11.8779494515,-2.5338635057,-0.
252231273\H,-11.4569390693,-3.978647742,-1.1813739079\C,-11.3327789761
,-4.2241102079,0.9775129169\O,-10.4478548607,-4.2812379008,1.827466556
1\N,-12.5099676514,-4.8842109293,1.0786428708\H,-13.2417231451,-4.7972
810981,0.3874456221\H,-12.7097539461,-5.4147439632,1.9158729715\\Versi
on=EM64L-G09RevC.01\State=1-A\HF=-4313.5915701\RMSD=3.128e-09\RMSF=3.2
56e-06\Dipole=-8.6815536,-6.1404563,0.8977326\Quadrupole=51.4497445,-1
2.421363,-39.0283815,-12.5836972,-7.2798524,53.9027695\PG=C01 [X(C46H6
7N15O12S2)]\\@

Supporting Information

References

1. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins-Structure Function and Bioinformatics* **2006,** 65, 712-725.
2. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C., An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2013,** 3, 198-210.
3. Syed Ibrahim, B.; Pattabhi, V., Trypsin Inhibition by a Peptide Hormone: Crystal Structure of Trypsin-Vasopressin Complex. *J Mol Biol* **2005,** 348, 1191-1198.
4. Joung, I. S.; Cheatham, T. E., Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *The Journal of Physical Chemistry. B* **2008,** 112, 9020-9041.
5. Joung, I. S.; Cheatham, T. E., Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. *The Journal of Physical Chemistry. B* **2009,** 113, 13279-13290.
6. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T., Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *The Journal of chemical physics* **2004,** 120, 9665-9678.
7. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular Dynamics with Coupling to an External Bath *Journal of Chemical Physics* **1984,** 81, 3684-3690.
8. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* **1977,** 23, 327-341.
9. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **1993,** 98, 10089-10092.
10. Haensele, E.; Banting, L.; Whitley, D. C.; Clark, T., Conformation and Dynamics of 8-Arg-Vasopressin in Solution. *Journal of Molecular Modeling* **2014,** 20, 2485(17).
11. Salt, D. W.; Hudson, B. D.; Banting, L.; Ellis, M. J.; Ford, M. G., DASH: A Novel Analysis Method for Molecular Dynamics Simulation Data. Analysis of Ligands of PPAR-Gamma. *J Med Chem* **2005,** 48, 3214-3220.
12. Barducci, A.; Bonomi, M.; Parrinello, M., Metadynamics. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2011,** 1, 826-843.
13. Laio, A.; Parrinello, M., Escaping Free-Energy Minima. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, 12562-12566.
14. Barducci, A.; Bussi, G.; Parrinello, M., Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys Rev Lett* **2008,** 100, 1-4.
15. Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M., Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *Journal of Physical Chemistry B* **2006,** 110, 3533-3539.
16. Juraszek, J.; Saladino, G.; van Erp, T. S.; Gervasio, F. L., Efficient Numerical Reconstruction of Protein Folding Kinetics with Partial Path Sampling and Pathlike Variables. *Phys Rev Lett* **2013,** 110, 108106(5).
17. Branduardi, D.; Gervasio, F. L.; Parrinello, M., From A to B in Free Energy Space. *Journal of Chemical Physics* **2007,** 126, 054103(10).
18. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013,** 29, 845-854.
19. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New Feathers for an Old Bird. *Computer Physics Communications* **2014,** 185, 604-613.
20. Becke, A. D., Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics* **1993,** 98, 5648-5652.
21. Ditchfield, R.; Hehre, W. J.; Pople, J. A., Self-Consistent Molecular-Orbital Methods .9. Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *Journal of Chemical Physics* **1971,** 54, 724-728.
22. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalman, G.; Barone, V.; Mennucci, B.; Petersson, M.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.;

Supporting Information

Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision C.01*, Revision C.01; Gaussian, Inc.: Wallingford CT, 2010.

23. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum Mechanical Continuum Solvation Models. *Chem Rev* **2005,** 105, 2999-3093.

24. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A., NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J Org Chem* **1997,** 62, 7512-7515.

25. Bax, A.; Davis, D. G., Mlev-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. *Journal of Magnetic Resonance* **1985,** 65, 355-360.

26. Kumar, A.; Ernst, R. R.; Wuthrich, K., A Two-Dimensional Nuclear Overhauser Enhancement (2d Noe) Experiment for the Elucidation of Complete Proton-Proton Cross-Relaxation Networks in Biological Macromolecules. *Biochemical and Biophysical Research Communications* **1980,** 95, 1-6.

27. McKay, R. T., Recent Advances in Solvent Suppression for Solution NMR: A Practical Reference. *Annual Reports on Nmr Spectroscopy, Vol 66* **2009,** 66, 33-76.

28. Kay, L. E.; Keifer, P.; Saarinen, T., Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *Journal of the American Chemical Society* **1992,** 114, 10663-10665.

29. John, B. K.; Plant, D.; Hurd, R. E., Improved Proton-Detected Heteronuclear Correlation Using Gradient-Enhanced Z and Zz Filters. *Journal of Magnetic Resonance Series A* **1993,** 101, 113-117.

30. Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M., Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation Nmr-Spectroscopy. *Journal of Magnetic Resonance* **1991,** 93, 151-170.

31. Kontaxis, G.; Stonehouse, J.; Laue, E. D.; Keeler, J., The Sensitivity of Experiments Which Use Gradient Pulses for Coherence-Pathway Selection. *Journal of Magnetic Resonance Series A* **1994,** 111, 70-76.

32. States, D. J.; Haberkorn, R. A.; Ruben, D. J., A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in 4 Quadrants. *Journal of Magnetic Resonance* **1982,** 48, 286-292.

33. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J Biomol NMR* **1995,** 6, 277-293.

34. Johnson, B. A.; Blevins, R. A., Nmr View - a Computer-Program for the Visualization and Analysis of NMR Data. *J Biomol NMR* **1994,** 4, 603-614.

35. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **2005,** 59, 687-696.

A 3: Reprint Supporting Information Paper 3

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00706.

Supporting Information

Conformation and Dynamics of Human Urotensin II and Urotensin Related Peptide in Aqueous Solution

Elke Haensele,[a] Nawel Mele,[b] Marija Miljak,[b] Christopher M. Read,[c] David C. Whitley,[a] Lee Banting,[a] Carla Delépée,[d] Jana Sopkova-de Oliveira Santos,[d] Alban Lepailleur,[d] Ronan Bureau,[d] Jonathan W. Essex,[b] and Timothy Clark[e,]*

[a]School of Pharmacy and Biomedical Sciences, and [c]School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DT, United Kingdom

[b]School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom

[d]Normandie Université, CS 14032 Caen Cedex 5, France, Centre d'Etudes et de Recherche sur le Médicament de Normandie (CERMN, EA 4258, FR CNRS 3038 INC3M SF 4206 ICORE), UFR des Sciences Pharmaceutiques, Université de Caen Basse–Normandie (UNICAEN), F-14032 Caen, France

[e]Computer-Chemie-Centrum and Interdisciplinary Center for Molecular Materials, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany

*Corresponding author: tim.clark@fau.de

Supporting Information

Table of contents

Supporting Information

MD simulations

Methodological details

Amber long-scale MD simulation. Amber long-scale (> 5 µs) simulations were started with (a) peptide conformations modelled corresponding to experimental data (MD-I,-II,-IXa),[1, 2] (b) random or minor populated states from short-term (12ns), high-temperature (400, 550, 700 K) Amber MD simulations (MD-IXb,-IXc,-IXd), (c) Amber REMD simulations (MD-V,-XI) and (d) CHARMM simulations (MD-III,-IV). The TIP4P-Ew water model[3, 4] with a truncated octahedral water box was used. The system was neutralized with Na^+ for urotensin-II (UII) and Cl^- for urotensin-related peptide (URP), either by simple charge equalization or with multiple counterions (Na^+ and Cl^-) to mimic physiological ion concentration. Energy was minimized at constant volume and the method was switched after 500 steps *steepest descent* to 9,500 steps *conjugated gradient*. Production runs were performed at constant temperature (T = 300 K, Berendsen coupling[5] of 1.0 ps to an external heat bath) and constant pressure (p = 1 atm) periodic boundary conditions with a non-bonded cut off of 8 Å. The SHAKE[6] algorithm was employed for hydrogen atoms with a simulation time step of 2 fs. Electrostatic energies were calculated using the Particle Mesh Ewald (PME) method[7] and coordinate 'snapshots' were written every 1 or 10 picosecond.

CHARMM simulation. Further simulations (MD-VI,-X) were carried out using CHARMM c36b2[8] with parameter set 36.[8] Initial structures for UII and URP originated from the NMR structures by Chatenet and Leprince et al.[2, 9] Each peptide was surrounded by a cubic box of TIP3P water molecules. The systems were neutralized by adding seven sodium and six chloride ions to the UII waterbox and four sodium and five chloride ions to the URP waterbox. All systems were simulated in the canonical ensemble (NVT) using periodic boundary conditions. The van der Waals interactions were brought smoothly to zero at 13 Å using a switching function while the electrostatic contribution was calculated using the PME summation method[7]. The system was heated to 300 K in 120 ps and equilibrated for 1 ns with velocity rescaling. The temperature of the systems was coupled to a Berendsen thermostat[5] using a coupling constant of 5 ps. The SHAKE[6] algorithm was applied to the hydrogens thus allowing an integration time-step of 2 fs.

REMD simulation. REMD was first proposed by Sugita and Okamoto.[10] In this approach a number of simulations are performed in parallel at different temperatures in a canonical ensemble. Periodically, pairs of replicas are exchanged following the Metropolis criteria based on the temperature and the potential energy of each replica (see Equation 1). Each replica is simulated for a period of time after the swapping, thus allowing a replica movement in temperature space. In a successful exchange, the

Supporting Information

two replicas swap their temperatures and a scaling factor involving the previous and the new target temperatures then rescales the associated velocities of all the atoms.

$$Probability\ of\ accepting\ the\ swap(i,j) = \begin{cases} 1, if\ \Delta \leq 0 \\ e^{-\Delta}, if\ \Delta > 0' \end{cases} \quad (1)$$

with $\Delta = \left({}^{1}/_{kT_i} - {}^{1}/_{kT_j}\right) * (V_j - V_i)$, k = Boltzmann constant, V = potential energy, T = temperature

$$Rescaling\ assignment\ for\ the\ replica\ i: v_{inew} = \sqrt{\frac{T_{inew}}{T_{iold}}} * v_{iold} \quad (2)$$

The mobility of the replicas in temperature space is governed by a range of factors including the time between swaps, the thermostat parameters, the temperature distribution of the replicas and the size of the system.[11, 12]

Three different configurations of the peptides UII (*lasso*, *folded* and *omega*) and URP (*lasso*, *omega-I*$_{open}$ and *omega-II*), extracted from the long-scale MD studies, were simulated for 500 ns using the PMEMD module in AMBER 12.[13] The temperature range was generated using the online generator http://folding.bmc.uu.se/remd/ with an overall expected acceptance ratio among replicas between 25-35 % and provided 64 replicas from 298 K to 543 K.[14] The Amber ff99SB force field was used with explicit TIP3P water model. The initial structures were solvated in a cubic box with periodic boundary conditions and neutralized with 1 Na^+ for UII and 1 Cl^- for URP. The Particle Mesh Ewald method was used for long-range interactions using a 10 Å cutoff. Bonds involving hydrogen were constrained using the SHAKE algorithm with a tolerance of 0.00001 Å. REMD simulations were performed in the NVT ensemble using a Langevin thermostat for the temperature coupling with a collision frequency of 1 ps^{-1}. 200 ps of NVT simulation was used to equilibrate the initial state to the desired temperature for each replica, following a rescaling of the velocities. Using these equilibrated replicas, 500 ns of REMD simulation was performed on each replica, resulting in 32 µs of molecular dynamics (REMD-I,-II,-III,-IV,-V,-VI). All exchanges between neighboring replicas were allowed every 2 ps in the NVT ensemble.

A summary of simulation details is given in **Table S1**. RMSD trajectories of long-scale (> 5 µs) MD simulations for UII are shown in Figure S1 to Figure S6.

Supporting Information

Table S1 Summary of MD simulation details

Simulation	Time (μs)	Initial conformation	Resulting ring-state types	NAtoms (WAT)*
UII (11 residues, 181 atoms, charge -1)				
MD: Amber ff99sb/ TIP4PEw/ trunc.oct. / 1Na^+/ 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
MD-I	5	*omega-I open*	*omega-I*	6154 (1493)
MD-II	5	*folded-I*	*folded-I*	5754 (1393)
MD-III	10	*lasso*	*lasso, omega-II, omega-I, scoop, folded-IVb2*	6642 (1615)
MD-V	5	*folded-II*	*folded-II, folded-III*	4338 (1039)
MD-XI	5	*inv-folded*	*inv-folded, lasso, omega-II*	5402 (1305)
MD: Amber ff99sb/ TIP4PEw/ trunc.oct./ 7Na^+, 6Cl^-/ 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
MD-IV	5	*omega-I open*	*omega-I, circle, lasso*	10082 (2472)
MD: CHARMM c36b2/ TIP3P/ cubic/ 7Na^+, 6Cl^-/ NVT/ 300K/ 13Å cutoff/ PME/ PCB				
MD-VI	1.3	*omega-I open*	*omega-I, lasso, scoop*	6899 (2235)
REMD: REMD: Amber ff99sb/ TIP3P / cubic/ 1Na^+/ 298K/ 1bar/ 10Å cutoff/ PME/ PBC/ Shake				
REMD-I	0.5	*omega-I open*	*omega-I/II, lasso, scoop, circle, folded-I/II/III/IVb2, inv-folded*	6632 (2150)
REMD-II	0.5	*folded-I*	*omega-I/II, lasso, scoop, circle, folded-I/II/III/Ivb2, inv-folded*	6476 (2098)
REMD-III	0.5	*lasso*	*omega-I/II, lasso, scoop, circle, folded-I/II/Ivb2, inv-folded*	6845 (2221)
total	*37.8*			
URP (8 residues, 136 atoms, charge +1)				
MD: Amber ff99sb/ TIP4Pew/ trunc.oct./ 1Cl^-/ 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
MD-Ixa	5	*URP omega-I open*	*omega-I, sheet, hybrid*	4153 (1004)
MD-IXb	5	*URP omega-I hbond*	*omega-I*	3633 (874)
MD-IXc	5	*URP 4O6H*	*omega-I*	3965 (957)
MD-IXd	5	*URP antip. ß-sheet*	*omega-I*	5021 (1221)
MD: CHARMM c36b2/ TIP3P/ cubic/ 4Na^+, 5Cl^-/ NVT/ 300K/ 13Å cutoff/ PME/ PCB				
MD-X	1.3	*omega-I open*	*omega-I, omega-II, $lasso_{45pbr}$*	4648 (1501)
REMD: Amber ff99sb/ TIP3P/ cubic/ 1Cl^-/ 298K/ 1bar/ 10Å cutoff/ PME/ PBC/ Shake				
REMD-IV	0.5	*URP omega-I open*	*omega-I/II, lasso, folded (sheet, hybrid)*	5561 (1808)
REMD-V	0.5	*URP omega-II*	*omega-I/II, lasso, folded (sheet, hybrid)*	6278 (2047)
REMD-VI	0.5	*URP lasso*	*omega-I/II, lasso, folded (sheet, hybrid)*	5948 (1937)
total	*22.8*			

*NAtoms: total number of atoms; WAT: number of water molecules

Supporting Information

Trajectories

Figure S1 to **Figure S6** show RMSD and DASH-state trajectories for UII of all long-scale (> 5 µs) Amber MD simulations. The long-scale Amber MD simulations for URP resulted in only one major populated ring-state type (*omega-I*) and hence not shown. State populations and assignment of representatives are given in Table S2.

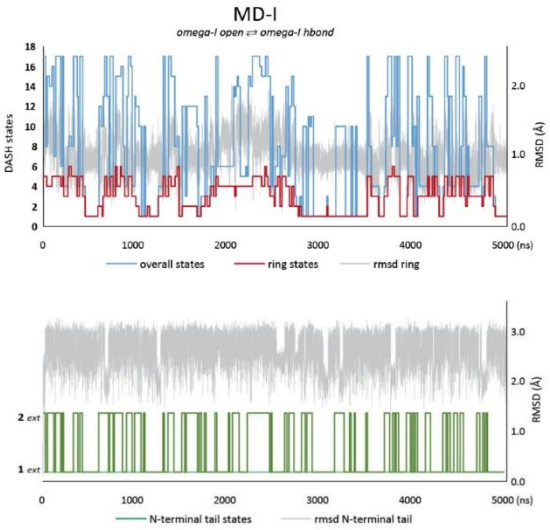


Figure S1 RMSD and DASH trajectories of simulation MD-I (5 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation: *omega-I*$_{open}$. Frequent interconversions of ring-state types *omega-I*$_{open}$ (T10_3,5,6) and *omega-II*$_{hbond}$ (T10_1,2,4). Despite multiple ring-state interconversions (variants of C-terminus, disulfide bridge and tail conformations), the unfolded ring shape *omega* persists during the complete MD simulation. Overall states are combinations of *omega-I* ring states with two variants of *extended* N-termini. *Omega-I* matches the *clinched open* conformation of AVP.

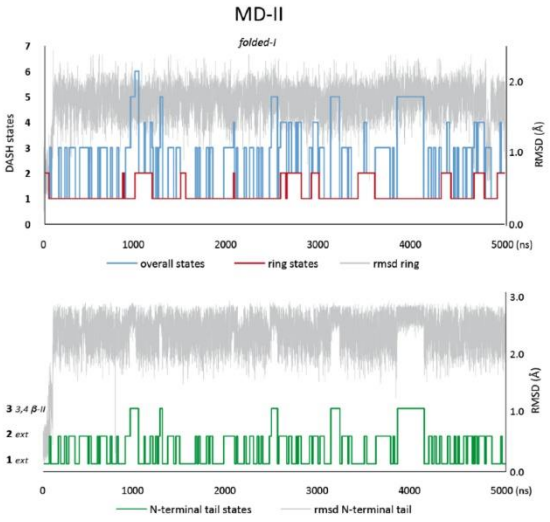


Figure S2 RMSD and DASH trajectories of simulation MD-II (5 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation: *folded-I*. The main ring-state type *folded-I* persists for the complete MD simulation comprising two C-terminus variants. Overall states are combinations of *folded-I* ring states with *extended* or *folded* N-termini. *Folded-I* matches the *saddle* conformation of AVP.

Supporting Information

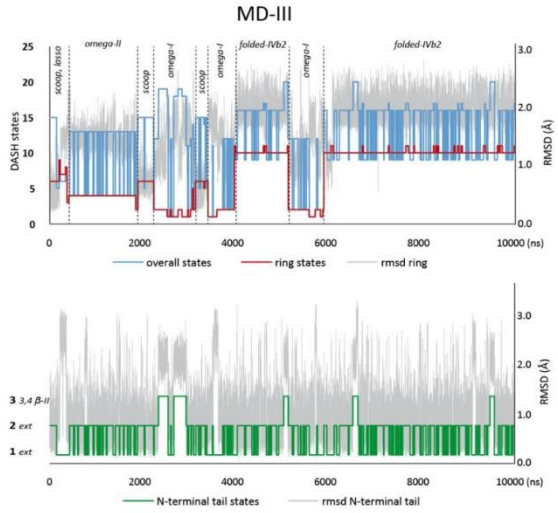


Figure S3 RMSD and DASH trajectories of simulation MD-III (10 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation *lasso*. Main ring-state types are labelled. For approximately 4 µs, interconversions of *unfolded/open* ring-state types take place (*scoop, lasso, omega-II, omega-I*) before the ring conformation changes to the *folded* variant *folded-VIb2*. MD-III also shows *omega* states with folded N-terminus (in contrast to MD-I). The overall states are combinations of the ring states and mainly *extended* N-termini. *Folded-IVb2* matches the *twisted-saddle* conformation of AVP; *omega-II* is a 8,9 peptide-bond rotamer of *omega-I* that matches the *clinched open* conformation of AVP and its variants.

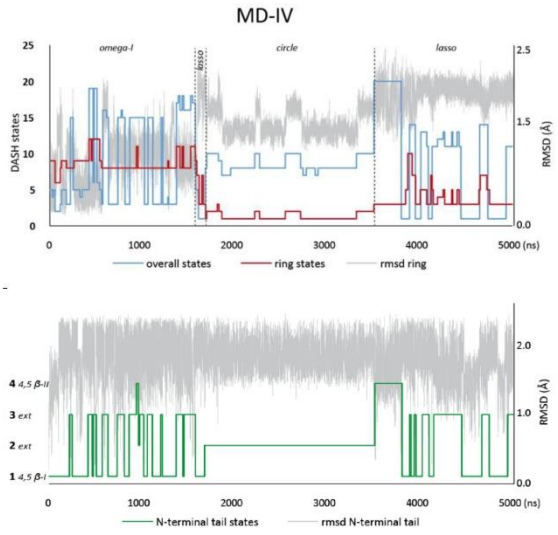


Figure S4 RMSD and DASH trajectories of simulation MD-IV (5 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation *omega-I*. Main ring-state types are labelled. The simulation only shows *unfolded* states with few interconversions of main ring-state types. MD-IV is the only Amber simulation mimicking physiological ion concentration and the *circle* conformation was only found under these conditions. Tail and ring state are strongly correlated for the *circle* type. *Omega* and *lasso* states show frequent interconversions of *folded* and *extended* N-termini. Note: *omega-I* in MD-I showed only *extended* tail states. This suggests an influence of counter-ions on the population of overall conformation in the MD simulations. However, as relative populations cannot be deduced from the MD simulations (due to sparse interconversions), this effect has not been investigated further here. *Lasso* is similar to AVP´s *open* ring-state type.

Supporting Information

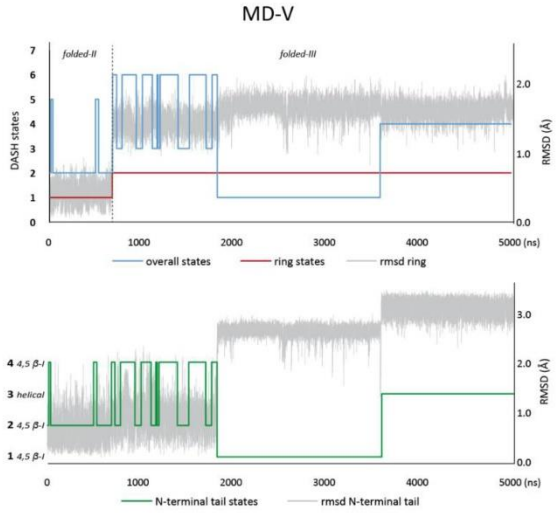


Figure S5 RMSD and DASH trajectories of simulation MD-V (5 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation *folded-II*. Main ring-state types are labelled. The simulation shows only one interconversion of the ring conformation from *folded-II* to *folded-III*. The ring-state types *folded-II/III* comprise several overall conformations with different folded N-termini. Tail-conformations are strongly correlated with the ring conformations. *Folded-II* is similar (circular similarity = 0.75) to the *saddle* conformation of AVP.

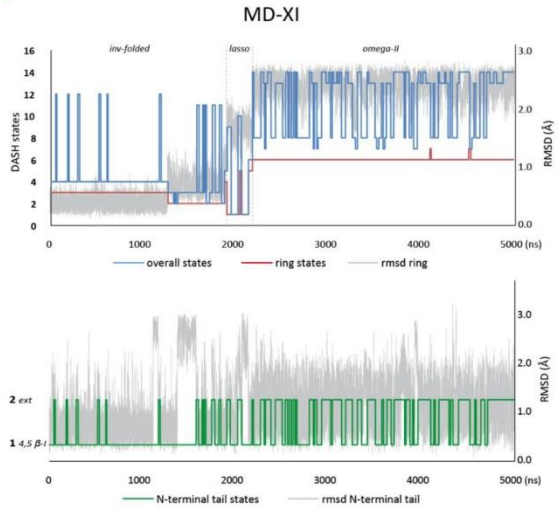


Figure S6 RMSD and DASH trajectories of simulation MD-XI (5 µs)
Trajectories of DASH states (overall, ring, N-terminal-tail) and RMSD (C^{α} 5-10, ring; C^{α} 1-5, tail) of UII. Initial conformation *inv-folded*. Main ring-state types are labelled. The simulation shows an interconversion from the *folded* ring-state type *inv-folded* to the *unfolded/open* states *lasso* and *omega-II*. No similar conformation to *inv-folded* has been found for AVP, yet.

Supporting Information

Conformational analysis

Clustering

Conformational clustering was performed by analyzing backbone ϕψ dihedral angles with DASH[15-17]. The conformations of UII were clustered using the whole structure to obtain overall states, and separately as ring and N-terminal tail states. Overall states are defined by the torsion range ϕψ 2-10 (T18), ring states by ψ 5, ϕψ 6-9, ϕ 10 (T10), and N-tail states by ψ 1, ϕψ 2-4, ψ 5 (T8). The ring states were further grouped into ring-state types**.** States assigned to the same ring-state type show identical turn centers and highly similar backbone cartoons but may comprise several subtypes with different turn types or hydrogen bond populations (each ring-state subtype may further comprise different disulfide bridge conformations and C/N-termini orientations). To analyze the N-tail states, the tail torsions of all long-scale MDs (MD-I to V, and XI) were clustered *via* DASH and grouped according to their secondary structure motifs. Relative populations of tail-state types for each ring-state type were calculated *via* their DASH state distribution in ring-state type sections. A list of all identified DASH states is given in **Table S2**. The DASH state trajectories are given in Figure S1 to **Figure S6** together with the RMSD trajectories. The most populated overall states corresponding to characteristic ring states (except for MD VI where only ring states were available) were taken as representatives for the ring-state types. Mean ring torsions of these representatives are given in **Table S3**. Finally, snapshots of the MD trajectory with maximum similarity to the representative states were extracted to provide 3D structures of the representatives. The coordinate files are available as supplementary files.

Supporting Information

Circular similarity

The consistency of the state assignment resulting from different simulations was ensured by comparison of the circular similarity of ring torsions. Circular similarity is defined as

$$S(\boldsymbol{x},\boldsymbol{y}) = 1 - D(\boldsymbol{x},\boldsymbol{y})/180\sqrt{n} \quad (3)$$

where

$$D(\boldsymbol{x},\boldsymbol{y}) = \sqrt{d(x_1,y_1)^2 + \cdots + d(x_n,y_n)^2} \quad (4)$$

is the distance between two states. Each Dash state is represented by the vector of mean torsion angles $x = (x_1,\ldots,x_n)$ and the distance between two angles (in degrees) is

$$d(x_i,y_i) = \min(|x_i - y_i|, 360 - |x_i - y_i|) \quad (5)$$

S(x,y) lies in [0, 1], with a value of 1 for identical states, and 0 for dissimilar states. Cosine similarity is defined as

$$S(\boldsymbol{x},\boldsymbol{y}) = \cos(\theta) = \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|} = \frac{\sum_{i=1}^{n} x_i \times y_i}{\sqrt{\sum_{i=1}^{n}(x_i)^2} \times \sqrt{\sum_{i=1}^{n}(y_i)^2}} \quad (6)$$

S(x,y) lies in [-1,1], with 1 meaning identical, 0 dissimilar and -1 opposite similarity. In cases of marginal torsion similarity, assignments were based on backbone CA atoms alignment and root mean square deviation (RMSD).

Notation of secondary structure elements

Turns were labelled by their turn centers, residues *i+1* and *i+2* (*e.g.* "8,9 ß-turn type-I", meaning a turn from residue 7 (*i*) to 9 (*i+3*) centered at residues 8 and 9 with ϕψ-angles for a type-I turn at residue 8 (*i+1*) and 9 (*i+2*)). The assignment to a distinct turn type was made if the trajectory of the ϕψ dihedrals (*i+1, i+2*) showed a continued fluctuation around the ideal torsion values (*cf.* **Table S3**).

β-turns were denoted as either *open* or classical (*hydrogen-bonded* or *hbond)* depending on the population of the hydrogen bond from O_i to NH_{i+3} and the turn propensity at turn centers *i+1* and *i+2*, with the following criteria:

a) *Classical* β-*turn* (*hbond*): hydrogen-bond population > 70±10%, turn propensity > 75%

b) *Open* β-*turn (open)*: hydrogen-bond population < 50±10%, turn propensity < 75%

The *open* β-*turn* is in accord with the more recent definition by Lewis et al.[18] using a distance criterion of < 7.0 Å for $C\alpha_i$-$C\alpha_{i+3}$ rather than the postulated hydrogen bond (classical definition by Ramachandran and Venkatachalam[19]).

Supporting Information

DASH states of UII and URP

Table S2. DASH states of UII and URP[a]

MD	Ring state		Overall state			Representative		Ring-state type
	T10	Pop (%)	T18	Pop (%)	CircSim T18 vs. T10	ID[b]	CircSim Rep vs. T10	
					UII			
I	**5**	22.5	**17***	13.1	0.99	**1**	0.99	*omega-l_{open}*
	3	21.8	**4**	16.2	0.99	**1**	0.93	*omega-l_{open}*
	6	2.2	**9**	1.3	0.99	**1**	0.90	*omega-l_{open}*
	1	29.3	**1***	23.1	1.00	**2**	1.00	*omega-l_{hbond}*
	4	13.6	**15**	5.0	0.97	**2**	0.84	*omega-l_{hbond}*
	2	10.6	**3**	1.4	0.99	**2**	0.92	*omega-l_{hbond}*
II	**1**	78.1	**1***	36.1	0.99	**6**	0.99	*folded-I*
	2	21.9	**2**	13.3	0.99	**6**	0.90	*folded-I*
III	**2**	15.3	**19**	2.7	0.98	**1**	0.94	*omega-l_{open}*
	4	14.4	**13***	11.2	0.98	**3**	0.98	*omega-II*
	1	7.3	**18**	2.0	0.99	**2**	0.93	*omega-l_{hbond}*
	3	0.4	**3**	0.4	1.00	**4**	0.91	*lasso*
	6	7.8	**15**	6.3	0.99	**5**	0.93	*scoop*
	9	0.2	**8**	0.3	1.00	**5**	0.58	*scoop-var2*
	5	0.3	**14**	0.3	1.00	**5**	0.59	*scoop-var1*
	10	47.8	**16***	2.2	1.00	**7**	1.00	*folded-IVb2*
	11	4.9	**17**	0.9	0.97	**7**	0.84	*folded-IVb2*
	7	1.1	**6**	1.1	1.00	**5**	0.53	*Na-helix*
	8	0.4	**7**	0.3	1.00	**5**	0.55	*Na-helix*
IV	**8**	19.5	**15**	6.95	0.99	**2**	0.98	*omega-l_{hbond}*
	11	2.5	**18**	1.15	0.97	**2**	0.88	*omega-l_{hbond}*
	9	6.5	**5**	5.91	0.99	**1**	0.92	*omega-l_{open}*
	12	2.1	**19**	1.5	0.97	**1**	0.99	*omega-l_{open}*
	6	1.2	**2**	1.16	1.00	**1**	0.69	*omega-l_{open}*
	3	22.7	**1***	13.34	0.99	**4**	0.99	*lasso*
	5	4.1	**13**	2.33	0.97	**4**	0.95	*lasso*
	10	0.9	**20**	5.87	0.90	**4**	0.88	*lasso*
	7	3.2	**14**	3.01	0.97	**4**	0.77	*lasso*
	4	1.3	**12**	1.24	1.00	**4**	0.57	*lasso-var*
	1	25.1	**8***	21.93	1.00	**10**	1.00	*circle*
	2	11.0	**10**	10.63	1.00	**10**	0.72	*circle-var*
V	**1**	13.5	**2***	12.6	1.00	**8**	1.00	*folded-II*
	2	86.5	**1***	35.3	0.99	**9**	0.99	*folded-III*
VI	**3**	55.3	**7**	37.9	0.99	**2**	0.86	*omega-l_{hbond}*
	4	12.4	**5**	12.0	1.00	**2**	0.67	*omega-l_{hbond}*
	5	17.7	**6**	17.9	1.00	**1**	0.91	*omega-l_{open}*
	1	6.3	**1**	6.3	1.00	**4**	0.88	*lasso*
	2*	8.3	**3**	7.5	0.99	**5**	1.00	*scoop*
XI	**6**	55.7	**14**	19.46	0.99	**3**	0.98	*omega-II*
	7	0.7	**7**	3.24	0.98	**3**	0.94	*omega-II*
	3	25.3	**4***	23.43	1.00	**11**	0.99	*inv-folded*
	2	12.2	**3**	8.14	0.99	**11**	0.74	*inv-folded*
	1	4.2	**1**	3.03	0.99	**4**	0.97	*lasso*
	5	1.3	**6**	0.86	0.97	**4**	0.78	*lasso*
	4	0.5	**5**	0.57	1.00	**4**	0.89	*lasso*
					URP			
	T10	**Pop (%)**	**T14**	**Pop (%)**	**T14 vs. T10**	**ID**	**Rep vs. T10**	
IXa	1	58.29	**1***	30.02	0.97	**1r**	0.97	*omega-l_{hbond}*
	2	36.30	**4***	35.58	1.00	**3r**	1.00	*omega-l_{open}*
	3	2.81	**5***	2.81	1.00	**4r**	1.00	*hybrid*
	4	2.60	**6***	2.60	1.00	**5r**	1.00	*sheet*
IXb	1	63.80	**1,2**	63.73	0.97	**1r**	0.98	*omega-l_{hbond}*
	2	36.20	**3**	36.27	1.00	**3r**	0.99	*omega-l_{open}*
IXc	1	51.88	**1**	32.50	0.98	**1r**	0.98	*omega-l_{hbond}*
	2	48.12	**3**	48.30	1.00	**3r**	0.99	*omega-l_{open}*
IXd	1	56.12	**1**	32.90	0.98	**1r**	0.98	*omega-l_{hbond}*
	2	43.88	**3**	44.22	1.00	**3r**	0.98	*omega-l_{open}*
X	1	13.93	**1**	13.93	1.00	**1r**	0.86	*omega-l_{hbond}*
	2	16.76	**2**	16.76	1.00	**3r**	0.94	*omega-l_{open}*
	4	66.91	**4**	66.91	1.00	**2r**	1.00	*omega-II*
	3	2.40	**3***	2.40	1.00	**6r**	1.00	*lasso$_{45pbr}$*

* T18/T14 states chosen as representatives. [a] Ring states (T10), corresponding overall states (T18) and representatives of all long-scale MD simulations. Listed are the populations relative to MD simulation times and the similarities of ring torsions. Coordinate files of the representatives are available as supplementary files. [b] ID of representative.

Ring torsions of UII and URP representatives

Table S3 Ring torsions of the representative DASH states for UII and URP[a]

							UII						
Ring-state type	**ID[b]**	**$C^5\Psi$**	**$F^6\Phi$**	**$F^6\Psi$**	**$W^7\Phi$**	**$W^7\Psi$**	**$K^8\Phi$**	**$K^8\Psi$**	**$Y^9\Phi$**	**$Y^9\Psi$**	**$C^{10}\Phi$**	**Turn type**	**ideal**
omega-I open	**1**	154.83	-83.82	-12.26	-111.33	166.81	-65.25*	-24.57*	-134.32*	131.66*	-125.02	open 8,9 β-VIII	-60°,-30°,-120°,+120°
std dev		*19.85*	*23.42*	*22.84*	*28.45*	*13.9*	*12.27*	*16.4*	*17.69*	*27.35*	*24.64*		
omega-I hbond	**2**	134.44	-126.75	10.58	-102.00	158.07	-59.98*	-13.17*	-95.59*	-6.35*	-124.15	8,9 β-I	-60°,-30°,-90°,0°
std dev		*14.61*	*16.19*	*20.79*	*29.75*	*10.36*	*17.59*	*26.48*	*24.98*	*18.51*	*21.14*		
omega-II	**3**	145.15	-81.52	-15.15	-111.26	159.42	-73.19*	153.41*	54.80*	33.84*	-95.12	open 8,9 β-II	-60°,+120°,+80°,0°
std dev		*11.87*	*16.96*	*20.11*	*27.85*	*11.45*	*13.58*	*10.52*	*8.46*	*14.93*	*21.24*		
lasso	**4**	15.06	-77.66*	-28.46*	-119.47*	-12.57*	-111.90	157.70	-72.59	132.29	-129.70	open 6,7 β-I	-60°,-30°,-90°,0°
std dev		*37.19*	*27.00*	*14.54*	*20.04*	*22.42*	*30.75*	*15.22*	*14.9*	*16.01*	*23.35*		
scoop	**5**	131.08	-59.68*	-39.31*	-99.71*	9.14*	75.12	-12.23	-116.95	125.99	-119.38	6,7 β-I	-60°,-30°,-90°,0°
std dev		*23.84*	*10.79*	*10.21*	*11.47*	*17.13*	*10.18*	*57.34*	*54.92*	*34.13*	*35.1*		
folded-I	**6**	143.34	-104.63	138.18	-57.40*	-26.51*	-72.79*	-16.95*	-129.24	-11.58	-118.97	7,8 β-I	-60°,-30°,-90°,0°
std dev		*29.46*	*30.66*	*15.58*	*13.41*	*12.85*	*15.63*	*13.48*	*16.01*	*27.44*	*29.46*		
folded-IVb	**7**	150.13	-69.04	158.25	-52.46*	126.54*	53.30*	15.23*	-93.38	-25.19	-112.04	7,8 β-II	-60°,+120°,+80°,0°
std dev		*15.51*	*16.51*	*10.85*	*17.33*	*16.29*	*8.99*	*21.67*	*22.97*	*17.46*	*20.51*		
folded-II	**8**	-9.69	-66.59	137.54	56.81*	7.28*	-115.82*	-47.2*	-149.04*	-16.75*	-113.94	(multiple turn)	
std dev		*8.39*	*13.67*	*13.98*	*7.84*	*21.32*	*24.52*	*13.04*	*12.18*	*19.43*	*24.63*		
folded-III	**9**	16.21	50.18*	28.16*	53.92*	12.33*	-131.79*	-22.02*	-118.18	24.93	-87.49	(multiple turn, 6,7 β-III')	+60°, +30°, +60°,+30°
std dev		*9.66*	*8.58*	*9.63*	*7.23*	*18.91*	*18.53*	*14.45*	*14.12*	*13.64*	*16.91*		
circle	**10**	27.94	-143.87	-21.53	-138.06	-39.58	-135.59	-47.95	-142.81	145.86	-145.90	(loop)	
std dev		*9.85*	*10.77*	*12.59*	*13.91*	*17.21*	*13.68*	*17.07*	*12.86*	*11.10*	*11.94*		
inv-folded	**11**	0.05	-65.79*	-25.58*	-61.99*	-21.41*	-117.49*	22.84*	54.86*	18.75*	54.59*	(multiple turn)	
std dev		*9.82*	*9.48*	*11.29*	*10.10*	*11.02*	*11.25*	*9.32*	*8.15*	*10.70*	*7.98*		
							URP						
Ring-state type	**ID**	**$C^2\Psi$**	**$F^3\Phi$**	**$F^3\Psi$**	**$W^4\Phi$**	**$W^4\Psi$**	**$K^5\Phi$**	**$K^5\Psi$**	**$Y^6\Phi$**	**$Y^6\Psi$**	**$C^7\Phi$**	**Turn type**	**ideal**
omega-I open	**3r**	144.82	-112.69	-9.94	-113.22	168.37	-61.36*	-24.06*	-129.43*	133.87*	-125.04	open 5,6 β-VIII	-60°,-30°,-120°,+120°
std dev		*17.04*	*29.49*	*23.49*	*29.24*	*14.25*	*16.90*	*18.78*	*19.66*	*33.20*	*30.06*		
omega-I hbond	**1r**	140.05	-124.10	8.70	-103.66	157.23	-59.67*	-13.75*	-95.18*	-7.29*	-127.26	5,6 β-I	-60°,-30°,-90°,0°
std dev		*13.91*	*18.00*	*20.79*	*28.96*	*9.86*	*14.95*	*25.57*	*26.02*	*13.91*	*27.13*		
omega-II	**2r**	135.34	-102.91	-3.40	-117.26	158.25	-66.06*	144.59*	68.19*	43.16*	-75.91	open 5,6 β-II	-60°,+120°,+80°,0°
std dev		*19.76*	*23.01*	*25.13*	*35.34*	*16.75*	*21.69*	*14.39*	*13.08*	*18.89*	*28.52*		
lasso	**6r**	126.55	-109.98*	-12.55*	-132.39*	136.74*	54.15	-167.73	-86.43	129.84	-114.06	open 3,4 β-VIII	-60°,-30°,-120°,+120°
std dev		*20.18*	*25.75*	*20.18*	*24.68*	*16.00*	*34.20*	*50.95*	*35.36*	*15.37*	*24.13*		
hybrid	**4r**	164.74	-89.11	149.79	-70.56*	142.40*	57.70*	-23.58*	-113.44	2.58	-113.03	4,5 β-II	-60°,+120°,+80°,0°
std dev		*13.91*	*35.37*	*55.82*	*32.98*	*16.30*	*10.91*	*31.06*	*25.81*	*42.48*	*32.09*		
sheet	**5r**	150.47*	-110.99*	-151.64*	-73.84*	97.87*	57.40*	-8.54*	-139.35*	136.65*	-129.53*	2-7 antip.-sheet (4,5 β-II)	-60°,+120°,+80°,0°
std dev		*18.30*	*34.75*	*84.03*	*27.82*	*49.50*	*9.06*	*39.45*	*21.60*	*34.64*	*21.67*		

[a] Torsions of turn centers are marked in accordance to the secondary structure analysis. The ideal turn-type torsions are given in the final column for comparison. std dev = standard deviation. Coordinate files of representatives are given as supplementary files. [b] ID of representative.

Supporting Information

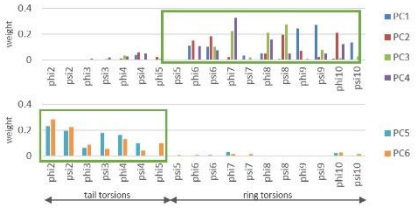
Principal component analysis

Overall torsion-trajectories were prepared from sections of the MD trajectories that were occupied exclusively by a single ring-state type. The correlation of ring and tail states was analyzed with principal component analysis (PCA) by comparing the weights of tail and ring torsions of the significant PCs with eigenvalue > 1.00 (Table S4). If the significant PCs correspond mainly to either tail or ring torsions, then the dynamics of ring and tail conformations can be regarded as independent, while significant weightings of both ring and tail torsions on a leading PC indicates that ring and tail conformations affect each other.

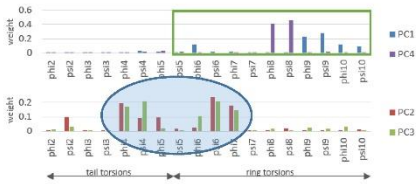
Table S4 Principal component weights of ring and tail torsions

Correlation of ring and tail torsions

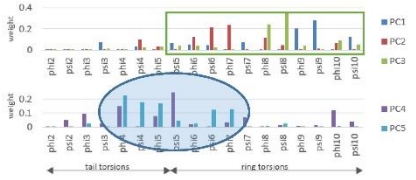
omega-I (MD I, 0-5000 ns, 100% ext): no



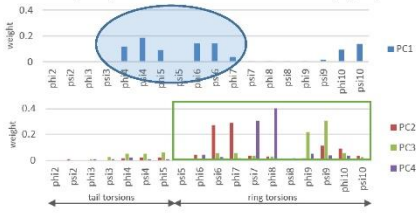
omega-I (MD-III, 2261-3024 ns, 38:62 ext:fold): yes



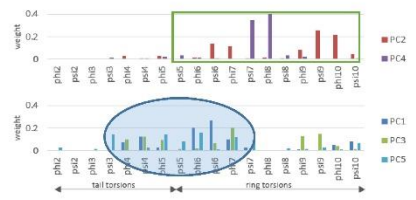
omega-I (MD-IV, 0-1588 ns, 38:62 ext:fold): yes



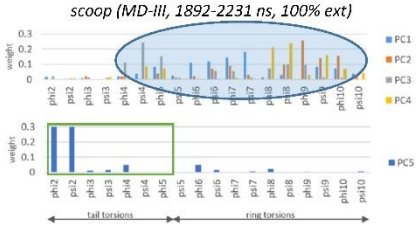
omega-II (MD-III, 408-1847 ns, 100% ext): (yes)



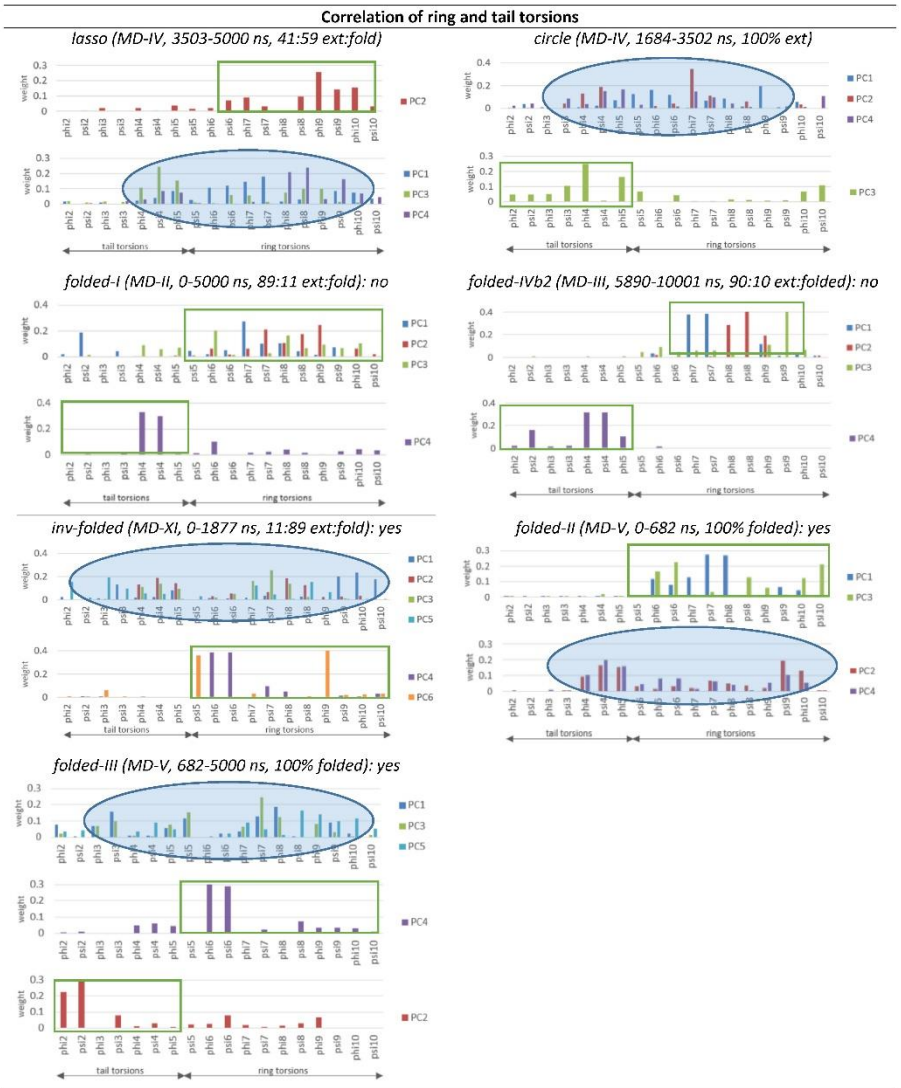
omega-II (MD-XI, 2179-5001, 62:38 ext:fold): yes



scoop (MD-III, 1892-2231 ns, 100% ext)



Supporting Information



Ψ 5 to Ψ 10 are ring torsions, ϕ 2 to ϕ 5 are tail torsions. PCs primarily loaded with either ring or tail torsions, or correlated ring/tail torsions are plotted separately. Correlated torsions are marked blue, uncorrelated green.

Supporting Information

NMR

Sample preparation

Human UII and URP were obtained from Bachem (UK) Ltd as the trifluoroacetate salt of the chemically synthesized peptide, each having a purity (by HPLC) of >95%. Mass spectrometry of the synthesized material gave molecular masses of 1388.3 and 1018.44 Da for UII and URP respectively, in close agreement to the calculated molecular masses of 1388.60 and 1017.26 Da for the reduced forms of the UII and URP peptides. Samples of 5.0 mg dry weight were dissolved in 320 µl of 90% H_2O/10% D_2O to give peptide concentrations of 11.25 mM (UII) and 14.80 mM (URP) respectively. The pH of the samples was measured to be 3.0/3.5 and NMR spectra were recorded without adjustment. In addition, samples were dried by lyophilization, then redissolved in 320 µl of 20 mM potassium phosphate buffer (pH 6.5) in 90% H_2O/10% D_2O and NMR spectra were recorded at a pH measured as 6.0. NMR spectra of UII and URP in D_2O at both pH 3.0/3.5 and pH 6.0 were recorded at least 2h after redissolving the extensively dried samples in 99.9% D20 (Sigma Aldrich).

NMR experiments

NMR spectroscopy was performed on a Varian Inova 600 MHz spectrometer, equipped with 5-channels, a 5 mm triple resonance ($^1H/^{13}C/^{15}N$) coldprobe and actively shielded pulse field z-axis gradients. ^1H-^1H TOCSY and ^1H-^1H NOESY spectra were acquired as 2048 complex points, with 32 transients for each of 512 increments and a spectral width of 12.0 ppm in both dimensions. Mixing times of 75 and 90 ms for the TOCSY experiment and 200 and 300 ms for the NOESY experiment were utilized. Water suppression was achieved through use of the watergate sequence. A ^{13}C-^1H gHSQC spectrum was acquired as 1024 complex points in the observe ^1H dimension and 280 increments in the indirect ^{13}C dimension using 32 transients over spectral widths of 12.0 and 140 ppm for the ^1H and ^{13}C dimensions respectively. The ^{13}C transmitter offset was initially set at 70 ppm, but other combinations of offset and sweep width were used to focus onto the aliphatic and aromatic regions of the spectrum. A ^{15}N-^1H gHSQC[20-22] spectrum was recorded with spectral widths of 12.0 ppm with 1024 complex points in the observe ^1H dimension and 30 ppm with 128 increments in the indirect ^{15}N dimension respectively. The ^{15}N transmitter offset was set to 120 ppm. A ^{15}N-^1H gHSQC spectrum of the URP sample was additionally recorded with the ^{15}N transmitter offset to 50 ppm in order to confirm the Lys^5 $H^{\zeta*}$ and N^{ζ} chemical shifts. In all experiments, States-TPPI quadrature detection was employed in the indirect dimension.[23]

Spectral processing and format conversion was performed using NMRPipe[24] and visualized with NMRView.[25] Spectra obtained for the UII and URP peptides were assigned using Analysis v2.3.1 from

Supporting Information

the CcpNMR software suite.[26, 27] Proton and ^{13}C chemical shifts were referenced to 3-trimethyl silyl propane sulfonic acid (DSS) and ^{15}N chemical shifts were referenced to an external liquid ammonia.

A second set of resonances representing a minor population (~10% of the total) was also observed in the UII NMR spectra. Downfield chemical shift changes in the Pro^3C^β and an upfield shift of Pro^3C^γ that are diagnostic of a *cis*-Pro conformation rather than the *trans*-Pro conformation was found[28, 29]. A $\Delta\beta\gamma$ = 4.7 ppm for the *trans*-Pro^3 $\Delta\beta\gamma$ = 9.69 ppm for the *cis*-Pro^3 conformations agrees closely with the statistical analyses of Schubert et al.[30] and Shen and Bax[31]. In addition there was the expected strong Thr^2H^α to Pro^3H^α NOE in the *cis*-Pro conformation as opposed to the strong Thr^2H^α to Pro^3H^δ NOEs of the *trans*-Pro conformation[32]. A minor UII conformation due to *cis*-isomerization of Pro^3 was thus identified and fully sequentially assigned.

The ^1H, ^{13}C and ^{15}N chemical shifts of the major populated *trans*-Pro^3 and minor populated *cis*-Pro^3 isomers of UII at pH 3.0, and pH 6.0 are given in **Table S5** and **Table S6.** The ^1H, ^{13}C and ^{15}N chemical shifts of URP at pH 3.5, and pH 6.0 are given in **Table S7** and **Table S8**.

Supporting Information

Experimental chemical shifts (^{1}H, ^{13}C, ^{15}N)

Table S5 ^{1}H, ^{13}C, ^{15}N NMR chemical shifts (ppm) of UII in H_2O/ D_2O at pH 3.0/ 298 K. [a]

Residue	H^N	N	H^α	C^α	Others
Glu[1]	-	-	4.67	52.20	2.17:$H^{\beta*}$; 2.53:$H^{\gamma*}$ 28.94:C^β; 32.24:C^γ
Thr[2]	8.78	122.61	4.67	60.09	4.18:H^β; 1.30:$H^{\gamma2*}$ 69.92:C^β; 21.76:C^γ
trans-Pro[3]	-	-	4.39	63.54	1.91:$H^{\beta a}$; 2.30:$H^{\beta b}$; 2.03:$H^{\gamma*}$; 3.74:$H^{\delta a}$; 3.87:$H^{\delta b}$ 32.32:C^β; 27.58:C^γ; 51.31:C^δ
Asp[4]	8.52	120.47	4.54	53.19	2.82:$H^{\beta*}$ 38.26:C^β
Cys[5]	8.00	121.00	4.59	55.75	2.87:$H^{\beta a}$; 3.15:$H^{\beta b}$ 41.94:C^β
Phe[6]	7.90	125.41	4.57	57.73	2.61:$H^{\beta a}$; 2.93:$H^{\beta b}$; 7.00:$H^{\delta*}$; 7.23:$H^{\varepsilon*}$; 7.24:H^ζ 39.21:C^β; 131.84:$C^{\delta*}$; 131.50:$C^{\varepsilon*}$; 130.00:C^ζ
Trp[7]	7.55	122.36	4.74	57.15	3.06:$H^{\beta a}$; 3.37:$H^{\beta b}$; 7.23:$H^{\delta1}$; 10.28:$H^{\varepsilon1}$; 7.59:$H^{\varepsilon3}$; 7.56:$H^{\zeta2}$; 7.23:$H^{\zeta3}$; 7.31:$H^{\eta2}$ 30.07:C^β; 127.49:$C^{\delta1}$; 121.14:$C^{\varepsilon3}$; 115.09:$C^{\zeta2}$; 122.45:$C^{\zeta3}$; 125.04:$C^{\eta2}$ 132.10:$N^{\varepsilon1}$
Lys[8]	7.92	122.04	3.97	58.15	1.55:$H^{\beta*}$; 0.98:$H^{\gamma a}$; 1.05:$H^{\gamma b}$; 1.55:$H^{\delta*}$; 2.88:$H^{\varepsilon*}$; 7.53:$H^{\zeta*}$ 32.71:C^β; 24.63:C^γ; 29.20:C^δ; 42.19:C^ε 34.79:N^ζ
Tyr[9]	7.65	117.72	4.63	57.48	3.07:$H^{\beta*}$; 7.10:$H^{\delta*}$; 6.79:$H^{\varepsilon*}$ 38.12:C^β; 133.42:$C^{\delta*}$; 118.36:$C^{\varepsilon*}$
Cys[10]	7.99	122.55	4.73	55.79	3.07:$H^{\beta*}$ 43.10:C^β
Val[11]	8.11	122.99	4.20	62.02	2.19:H^β; 0.97:$H^{\gamma a*}$; 0.97:$H^{\gamma b*}$ 32.99:C^β; 20.32:$C^{\gamma a}$; 21.33:$C^{\gamma b}$
Glu[1]	-	-	4.14	55.33	2.17:$H^{\beta*}$; 2.58:$H^{\gamma*}$ 28.92:C^β; 32.13:C^γ
Thr[2]	8.59	120.67	4.46	59.48	4.06:H^β; 1.21:$H^{\gamma2*}$ 70.87:C^β; 21.45:C^γ
cis-Pro[3]	-	-	4.82	63.27	2.13:$H^{\beta a}$; 2.40:$H^{\beta b}$; 1.84:$H^{\gamma a}$; 1.96:$H^{\gamma b}$; 3.53:$H^{\delta a}$; 3.62:$H^{\delta b}$ 34.68:C^β; 24.99:C^γ; 50.26:C^δ
Asp[4]	8.74	121.45	4.63	53.29	2.88:$H^{\beta*}$ ():C^β
Cys[5]	8.02	121.35	4.63	55.40	2.87:$H^{\beta a}$; 3.14:$H^{\beta b}$ 42.73: C^β
Phe[6]	8.04	125.07	4.57	(57.73)	2.68:$H^{\beta a}$; 2.93:$H^{\beta b}$; 7.04:$H^{\delta*}$; ():$H^{\varepsilon*}$; 7.28:H^ζ 39.28:C^β; 131.85:$C^{\delta*}$; ():$C^{\varepsilon*}$; ():C^ζ
Trp[7]	7.62	122.88	4.71	(57.15)	():$H^{\beta a}$; 3.36:$H^{\beta b}$; ():$H^{\delta1}$; 10.27:$H^{\varepsilon1}$; ():$H^{\varepsilon3}$; ():$H^{\zeta2}$; ():$H^{\zeta3}$; ():$H^{\eta2}$ ():C^β; ():$C^{\delta1}$; ():$C^{\varepsilon3}$; ():$C^{\zeta2}$; ():$C^{\zeta3}$; ():$C^{\eta2}$ ():$N^{\varepsilon1}$
Lys[8]	7.83	122.04	3.92	58.22	1.52:$H^{\beta*}$; 0.92:$H^{\gamma a}$; 0.97:$H^{\gamma b}$; 1.52:$H^{\delta*}$; 2.87:$H^{\varepsilon*}$; ():$H^{\zeta*}$ ():C^β; ():C^γ; ():C^δ; ():C^ε ():N^ζ
Tyr[9]	7.70	117.87	4.65	(57.48)	3.05:$H^{\beta a}$; 3.11:$H^{\beta b}$; 7.14:$H^{\delta*}$; 6.83:$H^{\varepsilon*}$ ():C^β; 133.32:$C^{\delta*}$; 118.27:$C^{\varepsilon*}$
Cys[10]	8.00	122.42	4.75	57.31	3.02:$H^{\beta*}$ 43.13:C^β
Val[11]	8.12	(122.99)	4.20	(62.02)	2.17:H^β; 1.08:$H^{\gamma a*}$; 1.08:$H^{\gamma b*}$ 32.25:C^β; 20.33:$C^{\gamma a}$; 21.33:$C^{\gamma b}$

[a] Upper half *trans*-Pro[3] isomer of UII; lower half *cis*-Pro[3] isomer of UII. SL1, MS1:A *trans*, MS1:C cis. Some resonances, indicated by (), are not assigned because they are too close to resonances of the main isomer (*trans*-Pro[3]); Resonances not observed are indicated with -.

Supporting Information

Table S6 ^{1}H, ^{13}C, ^{15}N NMR chemical shifts (ppm) of UII in H_2O/ D_2O at pH 6.0/ 298 K [a]

Residue	H^N	N	H^α	C^α	Others
Glu[1]	-	-	4.13	55.79	2.10:$H^{\beta a}$; 2.11:$H^{\beta b}$; 2.33:$H^{\gamma a}$; 2.36:$H^{\gamma b}$ 30.49:C^β; 35.89:C^γ
Thr[2]	8.41	121.84	4.68	59.99	4.19:H^β; 1.27:$H^{\gamma 2}$* 69.87:C^β; 21.77:C^γ
trans-Pro[3]	-	-	4.39	63.59	1.92:$H^{\beta a}$; 2.29:$H^{\beta b}$; 2.02:H^γ*; 3.73:$H^{\delta a}$; 3.85:$H^{\delta b}$ 32.29:C^β; 27.57:C^γ; 51.27:C^δ
Asp[4]	8.39	121.71	4.46	54.71	2.60:H^β* 41.09:C^β
Cys[5]	8.05	120.67	4.58	55.78	2.88:$H^{\beta a}$; 3.19:$H^{\beta b}$ 41.37:C^β
Phe[6]	7.99	125.18	4.50	58.41	2.76:$H^{\beta a}$; 2.87:$H^{\beta b}$; 6.94:H^δ*; 7.22:H^ε*; 7.23:H^ζ 39.10:C^β; 131.76:C^δ*; 131.54:C^ε*; 130.02:C^ζ
Trp[7]	7.64	121.79	4.70	57.07	3.12:$H^{\beta a}$; 3.36:$H^{\beta b}$; 7.23:$H^{\delta 1}$; 10.25:$H^{\varepsilon 1}$; 7.58:$H^{\varepsilon 3}$; 7.55:$H^{\zeta 2}$; 7.21:$H^{\zeta 3}$; 7.28:$H^{\eta 2}$ 29.90:C^β; 127.57:$C^{\delta 1}$; 121.18:$C^{\varepsilon 3}$; 115.02:$C^{\zeta 2}$; 122.41:$C^{\zeta 3}$; 124.99:$C^{\eta 2}$ 132.03:$N^{\varepsilon 1}$
Lys[8]	7.73	121.57	4.04	57.46	1.54:$H^{\beta *}$; 0.98:$H^{\gamma a}$; 1.04:$H^{\gamma b}$; 1.54:H^δ*; 2.87:H^ε* 33.02:C^β; 24.53:C^γ; 29.53:C^δ; 42.07:C^ε -:N^ζ
Tyr[9]	7.71	118.41	4.62	57.53	3.03:$H^{\beta a}$; 3.05:$H^{\beta b}$; 7.10:H^δ*; 6.79:H^ε* 38.13:C^β; 133.36:C^δ*; 118.35:C^ε*
Cys[10]	8.14	123.80	4.65	55.71	3.02:$H^{\beta a}$; 3.19:$H^{\beta b}$ 42.88:C^β
Val[11]	7.72	126.83	4.02	63.85	2.10:H^β; 0.91:$H^{\gamma a}$*; 0.91:$H^{\gamma b}$* 33.44:C^β; 20.31:$C^{\gamma a}$; 21.73:$C^{\gamma b}$
Glu[1]	-	-	4.07	55.85	2.08:$H^{\beta a}$; 2.13:$H^{\beta b}$; 2.42:$H^{\gamma a}$; ():$H^{\gamma b}$ ():C^β; 35.98:C^γ
Thr[2]	8.20	120.94	4.48	59.70	4.07:H^β; 1.19:$H^{\gamma 2}$* 70.60:C^β; 21.21:$C^{\gamma 2}$
cis-Pro[3]	-	-	4.83	63.37	2.16:$H^{\beta a}$; 2.35:$H^{\beta b}$; 1.84:$H^{\gamma a}$; 1.95:$H^{\gamma b}$; 3.52:$H^{\delta a}$; 3.60:$H^{\delta b}$ 34.58:C^β; 24.90:C^γ; 50.25:C^δ
Asp[4]	8.56	123.61	4.55	54.66	2.68:H^β* 41.39:C^β
Cys[5]	8.08	120.68	4.58	(55.78)	2.92:$H^{\beta a}$; 3.21:$H^{\beta b}$ 42.22:C^β
Phe[6]	7.94	124.93	4.48	(58.41)	2.75:$H^{\beta a}$; 2.85:$H^{\beta b}$; 6.94:H^δ*; 7.22:H^ε*; 7.23:H^ζ ():C^β; ():C^δ*; 131.54:C^ε*; ():C^ζ
Trp[7]	7.77	122.50	4.64	57.23	3.12:$H^{\beta a}$; 3.36:$H^{\beta b}$; ():$H^{\delta 1}$; 10.25:$H^{\varepsilon 1}$; ():$H^{\varepsilon 3}$; ():$H^{\zeta 2}$; ():$H^{\zeta 3}$; ():$H^{\eta 2}$ ():C^β; ():$C^{\delta 1}$; ():$C^{\varepsilon 3}$; ():$C^{\zeta 2}$; ():$C^{\zeta 3}$; ():$C^{\eta 2}$ 131.66:$N^{\varepsilon 1}$
Lys[8]	7.64	121.57	3.98	57.57	1.53:$H^{\beta *}$; 0.97:$H^{\gamma a}$; 1.03:$H^{\gamma b}$; ():H^δ*; 2.75:H^ε*; 2.86:H^ε* ():C^β; ():C^γ; ():C^δ; ():C^ε -:N^ζ
Tyr[9]	7.75	118.50	4.62	(57.53)	3.01:$H^{\beta a}$; 3.08:$H^{\beta b}$; ():H^δ*; ():H^ε* ():C^β; ():C^δ*; ():C^ε*
Cys[10]	8.20	123.80	4.65	(55.71)	2.87:$H^{\beta a}$; 3.20:$H^{\beta b}$ 42.88:C^β
Val[11]	7.75	126.82	4.02	(63.85)	2.09:H^β; 0.91:$H^{\gamma a}$*; 0.91:$H^{\gamma b}$* ():C^β; ():$C^{\gamma a}$; ():$C^{\gamma b}$

[a] Upper half *trans*-Pro[3] isomer of UII; lower half *cis*-Pro[3] isomer of UII. SL2, MS2:A *trans*, MS2:B,C cis. Some resonances are not assigned because they are too close to resonances of the main isomer (*trans*-Pro[3]); these resonances are in parentheses (). Resonances not observed are marked with dash -.

Supporting Information

Table S7 ^{1}H, ^{13}C, ^{15}N NMR chemical shifts (ppm) of URP in H_2O/ D_2O at pH 3.5/ 298 K [a]

Residue	H^N	N	H^α	C^α	Others
Ala[1]	-	-	4.03	51.83	1.41:H^{β}* 19.59:C^{β}
Cys[2]	8.61	120.59	4.65	55.61	2.99:$H^{\beta a}$; 3.12:$H^{\beta b}$ 42.94:C^{β}
Phe[3]	8.25	126.40	4.64	57.55	2.61:$H^{\beta a}$; 2.96:$H^{\beta b}$; 7.05:H^{δ}*; 7.26:H^{ε}*; 7.24:H^{ζ} 39.50:C^{β}; 131.84:C^{δ}*; 129.99:C^{ε}*; 127.45:C^{ζ}
Trp[4]	7.60	123.09	4.75	57.11	3.07:$H^{\beta a}$; 3.39:$H^{\beta b}$; 7.24:$H^{\delta 1}$; 10.31:$H^{\varepsilon 1}$; 7.64:$H^{\varepsilon 3}$; 7.58:$H^{\zeta 2}$; 7.25:$H^{\zeta 3}$; 7.32:$H^{\eta 2}$ 30.10:C^{β}; 131.54:$C^{\delta 1}$; 121.19:$C^{\varepsilon 3}$; 115.07:$C^{\zeta 2}$; 122.45:$C^{\zeta 3}$; 125.07:$C^{\eta 2}$ 132.09:$N^{\varepsilon 1}$
Lys[5]	7.89	122.18	3.88	58.54	1.51:H^{δ}*; 0.91:$H^{\gamma a}$; 0.98:$H^{\gamma b}$; 1.51:H^{δ}*; 2.86:H^{ε}*; 7.54:H^{ζ}* 32.45:C^{β}; 24.67:C^{γ}; 29.18:C^{δ}; 42.05:C^{ε} 34.80:N^{ζ}
Tyr[6]	7.66	117.32	4.68	57.23	3.05:$H^{\beta a}$; 3.14:$H^{\beta b}$; 7.12:H^{δ}*; 6.80:H^{ε}* 38.33:C^{β}; 133.46:C^{δ}*; 118.34:C^{ε}*
Cys[7]	8,01	122.19	4.80	56.10	3.09:H^{β}* 42.95:C^{β}
Val[8]	8.10	123.25	4.20	62.36	2.19:H^{β}; 0.97:$H^{\gamma a}$*; 0.98:$H^{\gamma b}$* 33.16:C^{β}; 20.30:$C^{\gamma a}$; 21.41:$C^{\gamma b}$

[a] SL1, MS1:A. Resonances not observed are marked with dash -.

Table S8 ^{1}H, ^{13}C, ^{15}N NMR chemical shifts (ppm) of URP in H_2O/ D_2O at pH 6.0/ 298 K

Residue	H^N	N	H^α	C^α	Others
Ala[1]	-	-	4.02	51.86	1.41:H^{β}* 19.69:C^{β}
Cys[2]	-	-	4.69	55.66	2.97:$H^{\beta a}$; 3.09:$H^{\beta b}$ 43.31:C^{β}
Phe[3]	8.29	126.17	4.63	57.54	2.62:$H^{\beta a}$; 2.97:$H^{\beta b}$; 7.05:H^{δ}*; 7.25:H^{ε}*; 7.22:H^{ζ} 39.53:C^{β}; 131.89:C^{δ}*; 129.97:C^{ε}*; 127.40:C^{ζ}
Trp[4]	7.67	123.17	4.72	57.13	3.11:$H^{\beta a}$; 3.39:$H^{\beta b}$; 7.22:$H^{\delta 1}$; 10.29:$H^{\varepsilon 1}$; 7.63:$H^{\varepsilon 3}$; 7.56:$H^{\zeta 2}$; 7.23:$H^{\zeta 3}$; 7.30:$H^{\eta 2}$ 29.88:C^{β}; 131.52:$C^{\delta 1}$; 121.19:$C^{\varepsilon 3}$; 115.05: :$C^{\zeta 2}$ 122.44:$C^{\zeta 3}$; 125.05:$C^{\eta 2}$ 132.07:$N^{\varepsilon 1}$
Lys[5]	7.86	121.81	3.88	58.39	1.50:H^{δ}*; 0.86:$H^{\gamma a}$; 0.93:$H^{\gamma b}$; 1.49:H^{δ}*; 2.83:H^{ε}* 32.46:C^{β}; 24.66:C^{γ}; 29.26:C^{δ}; 42.03:C^{ε} -:N^{ζ}
Tyr[6]	7.64	117.77	4.69	57.15	3.02:$H^{\beta a}$; 3.13:$H^{\beta b}$; 7.10:H^{δ}*; 6.80:H^{ε}* 38.47:C^{β}; 133.47:C^{δ}*; 118.31:C^{ε}*
Cys[7]	8.12	122.99	4.84	56.22	3.03:$H^{\beta a}$; 3.10:$H^{\beta b}$ 43.38:C^{β}
Val[8]	7.85	127.08	4.06	63.85	2.13:H^{β}; 0.94:$H^{\gamma a}$*; 0.94:$H^{\gamma b}$* 33.65:C^{β}; 20.31:$C^{\gamma a}$; 21.76:$C^{\gamma b}$

[a] SL2, MS2:A. Resonances not observed are marked with dash -.

Supporting Information

DFT calculations

^{1}H regression formula

Conversion of magnetic shielding at level B3LYP/6-31G(d)+PCM$_{water}$ to chemical shifts[33]

$$\delta(^{1}H) = -0.9912\sigma_H + 32.05 \quad (7)$$

$\delta(^{1}H)$ = ^{1}H chemical shift (ppm); σ_H = B3LYP/6-31G(d)-SCRF isotropic shielding

^{13}C regression formula

Conversion of magnetic shielding at level B3LYP/6-31G(d)+PCM$_{water}$ to chemical shifts[33]

$$\delta(^{13}C) = -1.0833\sigma_C + 203.97 \quad (8)$$

$\delta(^{13}C)$ = ^{13}C chemical shift (ppm); σ_C = B3LYP/6-31G(d)-SCRF isotropic shielding

^{15}N regression formula

A new correlation was calculated for ^{15}N chemical shifts based on experimental and calculated values for ammonia, ammonium, tetramethyl ammonium, tetramethyl urea, dimethylformamide, nitro-methane and nitrate (**Figure S7**).[34] The resulting regression equation is:

$$\delta(^{15}N) = -0.985\sigma_N + 254.23 \quad (9)$$

$\delta(^{15}N)$ = ^{15}N chemical shift (ppm); σ_C = B3LYP/6-31G(d)-SCRF isotropic shielding

where $\delta(^{15}N)$ is the ^{15}N chemical shift (ppm) and σ_N is the B3LYP/6-31G(d)-SCRF calculated isotropic shielding. This equation gives a mean unsigned error of 5.9 ppm and a root-mean-square deviation of 7.1 ppm for the seven reference compounds.

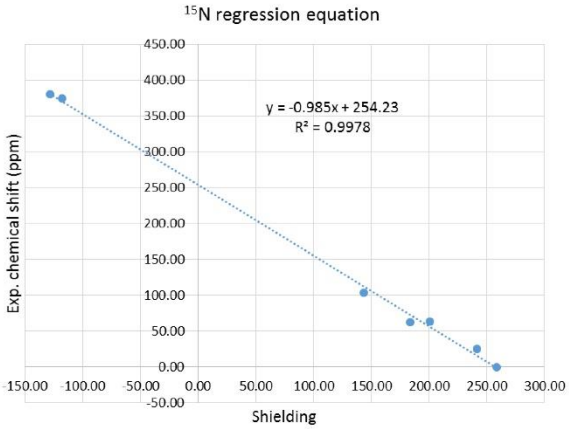


Figure S7 Regression formula for ^{15}N
Correlation between calculated isotropic magnetic shielding at B3LYP/6-31G(d)-SCRF level and a training set of experimental standard ^{15}N chemical shifts and

The calculated chemical shifts are given in Table S9 to **Table S12**.

Supporting Information

Calculated chemical shifts (^{1}H,^{13}C, ^{15}N)

Table S9 Calculated ^{1}H chemical shifts (ppm) of UII [a]

Atoms		Representatives						folded conformations				UII Equilibria open:folded		
		omega-I open	*open conformations omega-I hbond*	*omega-II*	*lasso*	*scoop*	*circle*	*folded-I*	*folded-IVb*	*inv-folded*	*folded-II*	*folded-III*	*72:28 REMD-I*	*70:30 REMD-II*	*79:21 REMD-III*
GLU[1]	HA	4.21	4.21	3.21	3.94	3.28	5.02	3.78	3.82	3.72	3.67	3.96	3.96	3.96	3.92
GLU[1]	HB2	2.35	2.42	1.86	2.47	1.78	1.89	1.53	2.43	1.72	2.41	1.93	2.20	2.29	2.28
GLU[1]	HB3	2.04	2.02	1.94	1.71	1.80	1.47	2.63	1.69	2.42	1.70	2.57	1.99	1.96	1.90
GLU[1]	HG2	1.51	1.47	2.40	1.65	2.73	1.42	2.01	2.43	2.01	1.60	2.54	1.77	1.79	1.71
GLU[1]	HG3	1.94	1.61	2.67	2.11	2.37	2.43	1.78	1.93	1.63	1.95	2.07	1.96	2.03	2.00
THR[2]	HA	4.39	4.56	3.93	4.28	4.55	4.26	4.37	4.33	4.65	4.57	4.91	4.44	4.45	4.37
THR[2]	HB	5.24	4.33	4.08	3.90	4.44	4.31	4.58	4.66	3.97	3.93	4.69	4.28	4.19	4.09
THR[2]	HG21	1.34	1.48	1.16	1.01	1.24	1.46	1.50	1.28	1.00	0.94	1.31	1.18	1.12	1.09
PRO[3]	HD2	5.00	3.74	3.72	3.40	3.95	3.81	3.67	3.67	3.97	3.83	3.81	3.88	3.73	3.70
PRO[3]	HD3	3.49	3.66	4.29	3.42	3.77	4.85	3.67	3.71	3.73	3.64	3.61	3.66	3.57	3.56
PRO[3]	HG2	3.08	2.12	2.11	1.99	1.97	2.06	2.35	2.39	2.57	2.39	1.94	2.30	2.19	2.22
PRO[3]	HB2	2.38	1.96	2.02	1.89	2.32	2.27	2.53	2.11	2.39	2.22	1.80	2.11	2.03	2.05
PRO[3]	HB3	2.22	2.30	2.30	2.25	1.90	2.43	2.26	2.28	2.29	2.38	1.87	2.25	2.24	2.27
PRO[3]	HA	4.15	3.91	4.03	4.06	4.11	4.24	4.13	4.23	4.17	4.17	3.60	4.06	4.06	4.08
ASP[4]	HA	4.76	5.28	4.51	4.70	4.25	4.77	4.13	3.69	4.75	4.58	4.31	4.74	4.69	4.74
ASP[4]	HB2	2.18	1.96	2.08	2.52	2.14	2.83	2.33	2.12	2.08	2.61	1.95	2.26	2.37	2.36
CYS[5]	HA	4.66	4.37	4.19	4.58	3.55	3.52	4.76	4.29	3.87	3.97	4.88	4.34	4.51	4.41
CYS[5]	HB2	2.17	2.88	3.06	2.43	4.00	3.11	2.55	2.87	2.65	2.08	3.62	2.65	2.61	2.51
CYS[5]	HB3	3.58	3.19	3.19	3.42	2.35	3.57	2.84	3.42	3.51	2.27	3.04	3.32	3.33	3.39
PHE[6]	HA	4.80	4.42	3.76	4.68	3.83	4.55	4.74	3.92	4.08	4.00	3.48	4.39	4.46	4.52
PHE[6]	HB2	3.67	1.49	2.76	2.91	2.04	3.26	3.62	2.93	2.89	3.09	3.21	2.83	2.99	2.88
PHE[6]	HB3	2.41	3.23	1.88	2.61	2.86	2.82	2.84	2.92	3.48	2.85	3.74	2.87	2.86	2.76
PHE[6]	HD*	7.08	7.09	6.28	6.94	6.65	7.51	7.79	7.66	7.09	6.42	7.47	7.02	7.07	6.97
PHE[6]	HE*	7.33	7.31	7.13	7.48	7.31	7.45	7.75	7.60	7.29	6.16	7.47	7.33	7.41	7.38
PHE[6]	HZ	7.30	7.41	7.21	7.37	7.32	7.23	7.56	7.44	7.35	7.13	7.37	7.34	7.40	7.35
TRP[7]	HA	5.03	4.73	3.85	5.09	4.46	4.89	4.43	4.17	4.10	3.88	3.85	4.64	4.75	4.80
TRP[7]	HB2	3.32	2.61	3.05	3.17	3.62	3.17	3.69	2.80	3.04	3.95	3.69	3.16	3.27	3.14
TRP[7]	HB3	3.44	2.82	3.25	3.74	3.58	3.50	2.82	2.97	3.69	3.32	3.50	3.46	3.60	3.58
TRP[7]	HD1	6.61	7.20	7.34	6.98	7.34	7.42	6.36	7.01	7.64	6.31	7.50	7.10	7.08	7.06
TRP[7]	HZ2	7.44	7.34	7.40	7.27	7.33	7.36	7.29	7.37	7.16	7.28	7.12	7.29	7.31	7.28
TRP[7]	HH2	7.32	7.32	7.31	7.28	7.26	7.27	7.20	7.24	7.03	7.25	7.05	7.24	7.28	7.25
TRP[7]	HZ3	7.29	7.41	7.37	7.33	8.04	7.25	7.05	7.25	7.04	7.17	7.06	7.28	7.30	7.28
TRP[7]	HE3	7.82	7.72	7.99	7.73	8.15	7.70	7.45	7.64	7.47	7.50	7.44	7.70	7.71	7.70
LYS[8]	HA	3.99	3.95	3.82	4.49	3.62	4.33	4.19	3.22	4.73	4.53	4.52	4.30	4.45	4.40
LYS[8]	HB*	1.91	1.98	1.81	1.73	2.25	2.18	1.21	2.14	1.73	1.92	1.61	1.83	1.76	1.78
LYS[8]	HG2	2.07	1.64	1.49	1.69	1.56	0.75	0.96	0.85	1.33	1.46	1.42	1.58	1.61	1.62
LYS[8]	HG3	2.05	1.82	1.88	1.89	1.82	0.89	0.29	0.45	0.56	1.89	2.20	1.62	1.75	1.63
LYS[8]	HD*	2.23	2.05	1.99	2.25	1.84	1.36	1.51	1.81	1.64	1.91	1.83	2.00	2.08	2.08
LYS[8]	HE*	3.66	3.53	3.66	3.22	3.49	2.30	3.14	3.21	2.64	3.73	3.52	3.26	3.29	3.21

Supporting Information

Atoms		Representatives										UII Equilibria		
		open conformations					folded conformations					open:folded		
		omega-I open	*omega-I hbond*	*omega-II*	*lasso*	*scoop*	*circle*	*folded-I*	*folded-IVb*	*inv-folded*	*folded-II*	*folded-III*	*72:28 REMD-I*	*70:30 REMD-II*	*79:21 REMD-III*
TYR[9]	HA	4.50	4.68	5.02	4.04	4.16	4.76	4.62	4.46	3.39	4.69	4.75	4.26	4.24	4.11
TYR[9]	HB2	3.01	3.80	2.79	3.08	2.45	2.90	2.59	3.06	3.22	3.50	2.57	3.14	3.08	3.14
TYR[9]	HB3	2.58	2.57	3.49	2.65	3.57	2.99	3.63	3.40	3.22	2.66	3.61	2.88	2.87	2.79
TYR[9]	HD*	7.38	7.23	7.73	6.83	7.42	6.84	7.09	7.32	7.01	7.42	7.36	7.12	7.08	7.00
TYR[9]	HE*	6.68	6.49	6.69	6.44	6.68	6.53	6.55	6.57	6.40	6.76	6.76	6.53	6.56	6.48
CYS[10]	HA	4.46	4.78	4.07	4.09	4.70	4.30	4.50	4.14	3.46	4.80	4.17	4.21	4.19	4.11
CYS[10]	HB2	2.10	3.27	3.28	2.85	2.86	2.87	3.03	2.81	3.43	4.10	2.52	2.95	2.92	2.96
CYS[10]	HB3	3.17	3.72	3.64	3.07	2.52	2.60	3.00	2.72	3.10	2.59	3.19	3.16	3.12	3.13
VAL[11]	HA	3.47	3.53	3.46	3.36	3.13	3.36	3.52	2.98	4.08	3.89	3.47	3.55	3.51	3.52
VAL[11]	HB	2.34	2.32	2.40	2.34	1.76	2.32	2.43	1.52	2.35	2.47	2.32	2.33	2.36	2.35
VAL[11]	HG1*	1.04	1.11	1.27	1.18	1.82	1.15	1.06	0.72	1.10	1.19	1.22	1.15	1.17	1.15
VAL[11]	HG2*	0.77	0.95	0.77	0.92	0.96	0.61	0.74	0.78	1.45	1.15	1.07	0.98	0.98	0.98

[a] Mag. shielding σ_H at level B3LYP/6-31G(d) with PCM-water; regression formula $\delta_H = -0.9912\sigma_H + 32.05$; equilibria deduced from REMD simulations.

Table S10 Calculated ^{13}C chemical shifts (ppm) of UII [a]

Atoms		Representatives										UII Equilibria		
		open conformations					folded conformations					open:folded		
		omega-I open	*omega-I hbond*	*omega-II*	*lasso*	*scoop*	*circle*	*folded-I*	*folded-IVb*	*inv-folded*	*folded-II*	*folded-III*	*72:28 REMD-I*	*70:30 REMD-II*	*79:21 REMD-III*
GLU[1]	CA	58.99	60.64	60.05	62.54	59.66	58.15	62.55	61.41	61.94	62.20	60.21	61.08	61.71	61.79
GLU[2]	CB	32.60	33.12	31.74	34.61	31.03	32.50	32.04	33.82	34.41	34.82	35.76	33.74	34.33	34.08
GLU[3]	CG	29.66	30.65	38.52	36.28	38.39	35.85	36.11	34.26	36.34	36.25	38.06	34.69	35.65	35.34
THR[2]	CA	61.25	62.38	63.13	62.63	61.49	59.91	60.37	61.01	60.46	60.40	60.64	61.66	61.84	62.01
THR[3]	CB	70.89	72.72	73.87	72.07	75.49	71.19	71.42	71.51	75.63	75.76	72.30	72.86	72.57	72.73
THR[4]	CG2	22.99	21.78	20.08	21.12	15.98	19.96	22.64	22.80	19.86	19.78	19.70	20.97	20.97	21.07
PRO[3]	CD	48.90	51.39	51.20	50.58	52.86	51.54	49.56	49.52	49.96	50.11	50.02	50.41	50.32	50.41
PRO[4]	CG	26.59	28.58	28.13	28.61	28.17	28.15	27.92	26.94	27.42	27.33	26.21	27.85	27.94	28.16
PRO[5]	CB	31.11	32.90	33.39	32.54	33.09	34.11	31.54	33.20	32.51	32.50	30.86	32.40	32.25	32.48
PRO[6]	CA	65.97	67.84	66.78	66.90	66.45	66.97	64.91	65.30	65.69	65.48	65.39	66.51	66.44	66.62
ASP[4]	CA	54.26	56.62	54.55	61.33	59.60	56.54	61.59	58.42	61.90	61.16	60.74	59.00	60.32	60.06
ASP[5]	CB	42.92	45.51	42.86	44.89	44.55	45.41	48.06	39.63	45.13	44.79	44.79	44.65	44.81	44.77
CYS[5]	CA	59.44	59.61	54.62	57.78	58.29	59.80	54.57	55.72	58.68	59.38	55.05	58.25	57.77	58.10
CYS[6]	CB	43.07	50.96	48.68	44.22	50.81	51.24	48.69	48.47	44.99	42.94	48.29	46.17	45.21	45.14
PHE[6]	CA	57.35	61.00	63.98	60.01	64.52	63.52	58.97	62.63	65.74	64.78	65.30	61.65	61.32	61.11
PHE[7]	CB	38.16	40.27	41.64	40.81	41.55	44.76	44.40	40.71	41.51	41.25	37.30	40.58	40.44	40.82
PHE[8]	CD1	132.36	130.44	131.81	132.56	132.44	131.15	130.90	135.08	132.22	130.56	131.20	131.91	132.08	132.20
PHE[9]	CE1	130.28	130.42	129.97	129.79	130.21	130.75	131.82	129.88	130.29	129.63	130.56	130.19	130.07	130.02
PHE[10]	CZ	128.78	128.90	128.94	128.03	129.00	128.48	130.28	129.73	128.98	128.51	128.40	128.62	128.39	128.42
TRP[7]	CA	57.37	60.58	66.09	55.13	61.64	56.68	62.57	61.96	62.40	60.81	67.50	59.33	58.38	57.70
TRP[8]	CB	31.95	33.41	29.11	32.64	27.98	36.32	28.42	32.02	29.32	25.08	25.19	31.20	30.92	31.74

TRP[9]	CD1	126.48	126.12	126.01	125.01	124.52	126.78	125.61	127.13	136.44	121.53	130.01	127.56	126.71	127.05
TRP[10]	CZ2	112.12	112.64	112.37	111.42	111.68	112.33	111.61	111.75	112.02	111.62	111.99	111.95	111.70	111.73
TRP[11]	CH2	123.77	123.64	123.87	122.97	123.97	123.28	123.97	124.16	121.29	123.28	120.52	122.92	122.71	122.90
TRP[12]	CZ3	121.14	120.50	121.89	120.78	122.64	120.80	121.37	121.72	119.19	120.35	119.35	120.58	120.51	120.59
TRP[13]	CE3	120.04	120.89	119.03	121.67	119.99	119.99	122.11	120.35	120.58	120.09	118.84	120.67	120.93	121.11
LYS[8]	CA	61.01	62.40	59.89	55.73	56.76	60.67	58.73	63.53	57.22	60.53	58.86	58.69	57.50	57.41
LYS[9]	CB	33.85	34.25	34.25	40.88	22.66	37.20	35.44	30.04	31.40	35.85	37.41	35.75	37.99	37.60
LYS[10]	CG	25.16	26.83	25.57	25.61	19.43	24.60	26.37	26.49	24.64	24.02	25.18	25.29	25.39	25.46
LYS[11]	CD	29.22	30.13	28.70	29.84	23.17	29.28	29.26	29.91	30.86	30.71	29.82	29.72	29.89	29.91
LYS[12]	CE	46.57	47.24	48.28	43.63	43.86	45.47	47.60	47.45	44.27	42.49	47.70	45.31	44.72	44.55
TYR[9]	CA	58.63	57.76	64.65	62.80	61.92	58.84	61.01	61.37	64.12	59.96	59.08	61.18	61.67	62.15
TYR[10]	CB	44.46	38.19	34.46	39.96	41.41	43.46	43.59	40.83	34.62	39.28	40.71	39.51	39.89	39.26
TYR[11]	CD1	132.25	133.24	132.47	133.62	132.66	133.81	134.62	132.59	132.02	135.06	132.91	133.06	133.34	133.22
TYR[12]	CE1	116.27	116.48	116.16	115.34	116.15	115.23	114.43	115.65	115.61	115.98	115.90	115.81	115.58	115.59
CYS[10]	CA	53.48	56.43	60.60	54.25	57.30	55.88	55.56	60.26	60.91	56.09	57.28	56.32	55.53	55.79
CYS[11]	CB	48.09	52.14	49.64	52.76	49.97	49.58	51.47	48.24	50.74	55.59	45.99	50.95	51.40	51.82
VAL[11]	CA	65.12	64.11	63.43	64.61	65.02	64.16	64.33	64.57	63.31	63.83	63.61	64.25	64.33	64.33
VAL[12]	CB	36.01	37.17	36.89	35.93	38.07	37.28	36.74	38.34	35.76	36.85	37.45	36.42	36.27	36.12
VAL[13]	CG1	22.28	23.16	22.73	22.65	20.93	23.44	23.10	20.87	20.52	22.66	23.28	22.35	22.52	22.34
VAL[13]	CG2	16.77	16.21	16.91	16.53	21.86	16.37	16.66	21.10	16.41	16.65	16.88	16.71	16.61	16.55

[a] Mag. shielding σ_C at level B3LYP/6-31G(d) with PCM-water; regression formula $\delta_{13C} = -1.0833\sigma_H + 203.97$; equilibria deduced from REMD simulations.

Table S11 Calculated ^{15}N chemical shifts (ppm) of UII [a]

Atoms		Representatives											UII Equilibria open:folded		
		open conformations						folded conformations							
		omega-I open	*omega-I hbond*	*omega-II*	*lasso*	*scoop*	*circle*	*folded-I*	*folded-IVb*	*inv-folded*	*folded-II*	*folded-III*	*72:28 REMD-I*	*70:30 REMD-II*	*79:21 REMD-III*
THR[2]	N	113.89	110.68	120.84	116.17	116.99	130.34	114.66	117.47	124.64	124.24	117.21	117.66	117.31	117.51
ASP[4]	N	121.00	127.44	124.83	127.28	125.72	122.62	125.01	121.63	127.15	124.78	129.16	125.89	126.60	126.40
CYS[5]	N	121.25	121.39	121.46	111.81	119.34	115.30	107.34	117.54	109.62	113.36	106.05	114.87	112.41	113.49
PHE[6]	N	114.56	126.79	126.65	120.79	129.69	115.97	118.96	127.96	131.03	133.20	123.88	123.40	122.55	122.81
TRP[7]	N	121.31	116.02	115.63	117.06	113.60	120.95	124.70	131.69	117.59	120.94	112.55	117.80	117.46	117.73
TRP[7]	NE1	126.30	127.22	127.62	125.73	127.62	125.46	128.49	125.37	142.56	123.87	143.84	129.89	129.46	128.66
LYS[8]	N	120.55	121.21	124.27	118.59	111.07	118.90	116.85	119.28	116.84	117.79	112.83	118.73	118.04	118.86
TYR[9]	N	115.26	110.05	113.49	122.31	116.15	123.37	115.07	121.38	115.25	111.82	114.90	116.78	118.71	118.85
CYS[10]	N	122.83	118.30	114.92	124.49	127.55	134.83	118.69	115.75	112.36	121.44	121.26	120.96	122.18	121.50
VAL[11]	N	125.98	125.13	116.39	126.02	136.15	134.25	128.21	134.59	115.61	129.35	127.94	124.67	125.48	124.20

[a] Mag. shielding σ_N at level B3LYP/6-31G(d) with PCM-water; regression formula $\delta_{15N} = -0.985\sigma_N + 254.23$; equilibria deduced from REMD simulations.

Supporting Information

Table S12 Calculated ^{1}H chemical shifts (ppm) of URP [a]

Atoms		Representatives						UII Equilibria		
		open conformations				*folded conformations*		*open:folded*		
		omega-I open	*omega-I hbond*	*omega-II*	*lasso*	*hybrid*	*sheet*	*86:14 REMD-IV*	*94:6 REMD-V*	*91:9 REMD-VI*
ALA[1]	HA	3.88	3.94	3.88	3.94	4.14	4.16	3.93	3.91	3.92
ALA[1]	HB1	1.54	1.62	1.44	1.61	1.82	1.28	1.55	1.52	1.53
ALA[1]	HB2	1.74	1.30	1.67	1.14	1.55	1.48	1.53	1.56	1.56
ALA[1]	HB3	1.57	1.13	1.21	1.27	1.94	1.22	1.33	1.25	1.31
CYS[2]	HA	4.06	3.91	4.04	4.22	4.26	5.01	4.07	4.03	4.06
CYS[2]	HB2	3.60	3.33	2.82	3.30	3.14	2.70	3.17	3.03	3.14
CYS[2]	HB3	2.61	3.55	2.37	2.87	2.97	2.78	2.86	2.74	2.77
PHE[3]	HA	5.04	4.72	4.58	4.74	4.81	4.62	4.74	4.66	4.73
PHE[3]	HB2	2.56	1.48	3.88	3.83	3.52	2.27	2.82	3.15	2.97
PHE[3]	HB3	3.72	3.31	2.67	2.57	3.18	3.08	3.12	2.92	3.07
PHE[3]	HD1	7.11	7.19	7.43	7.52	7.68	7.38	7.33	7.37	7.33
PHE[3]	HE1	7.26	7.27	7.42	7.42	7.41	7.41	7.34	7.37	7.35
PHE[3]	HZ	7.34	7.14	7.22	7.26	7.43	7.32	7.24	7.22	7.24
PHE[3]	HE2	7.46	7.20	6.90	7.16	7.53	7.31	7.18	7.06	7.15
PHE[3]	HD2	7.50	6.95	7.17	7.25	7.43	6.89	7.19	7.15	7.20
TRP[7]	HA	5.16	4.72	4.57	4.49	4.78	4.50	4.74	4.65	4.73
TRP[7]	HB2	3.26	2.62	3.93	2.83	3.31	3.04	3.28	3.48	3.37
TRP[7]	HB3	3.67	2.72	3.21	3.75	3.29	4.07	3.21	3.13	3.23
TRP[7]	HD1	7.58	7.22	7.73	7.09	7.30	7.31	7.46	7.54	7.51
TRP[7]	HZ2	7.44	7.34	7.49	7.30	7.35	7.49	7.41	7.44	7.43
TRP[7]	HH2	7.33	7.29	7.33	7.29	7.29	7.23	7.31	7.32	7.32
TRP[7]	HZ3	7.43	7.34	7.29	7.27	7.24	7.32	7.33	7.31	7.33
TRP[7]	HE3	8.31	7.65	7.60	8.47	7.56	8.28	7.80	7.67	7.80
LYS[8]	HA	3.92	4.04	4.35	3.22	2.46	3.43	3.91	4.10	3.99
LYS[8]	HB2	1.99	1.95	2.02	1.88	1.07	2.40	1.90	1.94	1.93
LYS[8]	HB3	1.43	2.16	1.96	2.61	2.11	2.20	1.97	2.01	1.95
LYS[8]	HG2	-0.31	1.62	1.83	2.06	0.31	1.14	1.19	1.57	1.25
LYS[8]	HG3	0.79	1.96	1.87	1.70	0.66	0.75	1.52	1.75	1.56
LYS[8]	HD2	1.57	2.07	1.88	2.12	1.63	0.36	1.81	1.91	1.83
LYS[8]	HD3	1.44	2.11	2.41	2.07	1.20	1.59	1.97	2.19	2.03
LYS[8]	HE2	2.95	3.71	3.43	3.66	2.71	3.04	3.34	3.44	3.35
LYS[8]	HE3	2.96	3.57	3.43	3.66	3.18	2.21	3.32	3.43	3.34
TYR[9]	HA	4.39	4.67	3.70	4.28	4.75	4.57	4.28	4.07	4.19
TYR[9]	HB2	2.88	3.82	3.34	2.77	2.72	2.84	3.29	3.39	3.29
TYR[9]	HB3	2.89	2.57	3.47	2.47	3.96	3.32	3.10	3.20	3.12
TYR[9]	HD1	7.13	7.24	7.27	6.99	7.34	7.62	7.24	7.25	7.24
TYR[9]	HE1	6.50	6.48	6.77	6.35	6.89	6.94	6.63	6.67	6.64
TYR[9]	HE2	6.85	6.68	7.01	7.23	6.83	6.64	6.86	6.91	6.89
TYR[9]	HD2	7.40	7.23	6.50	6.67	7.39	7.11	7.00	6.80	6.93
CYS[10]	HA	4.50	4.83	3.67	4.22	3.93	5.08	4.27	4.05	4.19
CYS[10]	HB2	2.10	3.17	2.83	3.20	2.21	2.91	2.75	2.85	2.74
CYS[10]	HB3	2.95	3.86	3.43	2.63	3.04	2.62	3.37	3.47	3.36
VAL[11]	HA	3.28	3.56	3.23	3.17	3.18	3.60	3.34	3.31	3.32
VAL[11]	HB	2.37	2.32	1.83	1.64	1.61	1.99	2.05	1.97	2.03
VAL[11]	1HG1	1.07	1.13	1.00	0.94	0.91	1.18	1.05	1.03	1.04
VAL[11]	2HG1	1.17	0.99	0.88	0.83	1.78	1.52	1.08	0.98	1.04
VAL[11]	3HG1	0.91	0.95	1.89	1.88	0.83	1.03	1.29	1.53	1.38
VAL[11]	1HG2	0.93	1.15	1.74	0.87	0.84	1.80	1.29	1.46	1.34
VAL[11]	2HG2	1.01	1.03	0.86	0.91	0.78	1.25	0.94	0.91	0.93
VAL[11]	3HG2	0.75	0.64	0.96	1.35	1.20	1.05	0.87	0.89	0.88

[a] Mag. shielding σ_H at level B3LYP/6-31G(d) with PCM-water; regression formula $\delta_{1H} = -0.9912\sigma_H + 32.05$; equilibria deduced from REMD simulations.

Supporting Information

REMD equilibrium models

UII and URP equilibrium equations

UII and URP models were built by the following linear combinations of the calculated chemical shifts of representative conformations. The models were applied to experimental ^{1}H (UII/URP, main paper) and ^{15}N (UII) chemical shifts.

Equilibrium UII REMD-I (72:28)

$$\delta_{Eq.VIIa} = 0.1519\delta_{omega\text{-}I_{open}} + 0.1476\delta_{omega\text{-}I_{hbond}} + 0.0507\delta_{omega\text{-}II} + 0.2975\delta_{lasso} + 0.0267\delta_{scoop} + 0.0453\delta_{circle} + 0.0176\delta_{folded\text{-}I} + 0.0063\delta_{folded\text{-}IVb2} + 0.1639\delta_{inv\text{-}folded} + 0.0389\delta_{folded-II} + 0.0537\delta_{folded\text{-}III} \quad (10)$$

Equilibrium UII REMD-II (70:30)

$$\delta_{Eq.VIIb} = 0.0872\delta_{omega\text{-}I_{open}} + 0.0468\delta_{omega\text{-}I_{hbond}} + 0.0129\delta_{omega\text{-}II} + 0.5411\delta_{lasso} + 0.0030\delta_{scoop} + 0.0139\delta_{circle} + 0.0300\delta_{folded\text{-}I} + 0.0034\delta_{folded\text{-}IVb2} + 0.0967\delta_{inv\text{-}folded} + 0.0558\delta_{folded-II} + 0.1092\delta_{folded\text{-}III} \quad (11)$$

Equilibrium UII REMD-III (79:29)

$$\delta_{Eq.VIIc} = 0.0898\delta_{omega\text{-}I_{open}} + 0.0769\delta_{omega\text{-}I_{hbond}} + 0.0410\delta_{omega\text{-}II} + 0.5673\delta_{lasso} + 0.0020\delta_{scoop} + 0.0168\delta_{circle} + 0.0182\delta_{folded\text{-}I} + 0.0028\delta_{folded\text{-}IVb2} + 0.1596\delta_{inv\text{-}folded} + 0.0256\delta_{folded-II} + 0.0000\delta_{folded\text{-}III} \quad (12)$$

Equilibrium URP REMD-IV (86:14)

$$\delta_{Eq.VIIIa} = 0.1892\delta_{omega\text{-}I_{open}} + 0.2973\delta_{omega\text{-}I_{hbond}} + 0.3378\delta_{omega\text{-}II} + 0.0405\delta_{lasso} + 0.0338\delta_{sheet} + 0.1014\delta_{hybrid} \quad (13)$$

Equilibrium URP REMD-V (94:6)

$$\delta_{Eq.VIIIb} = 0.0580\delta_{omega\text{-}I_{open}} + 0.2609\delta_{omega\text{-}I_{hbond}} + 0.5942\delta_{omega\text{-}II} + 0.0290\delta_{lasso} + 0.0001\delta_{sheet} + 0.0579\delta_{hybrid} \quad (14)$$

Equilibrium URP REMD-VI (91:9)

$$\delta_{Eq.VIIIc} = 0.1970\delta_{omega\text{-}I_{open}} + 0.2424\delta_{omega\text{-}I_{hbond}} + 0.4242\delta_{omega\text{-}II} + 0.0455\delta_{lasso} + 0.0227\delta_{sheet} + 0.0682\delta_{hybrid} \quad (15)$$

Supporting Information

Linear regression of UII ^{13}C chemical shifts and sensitivity analysis

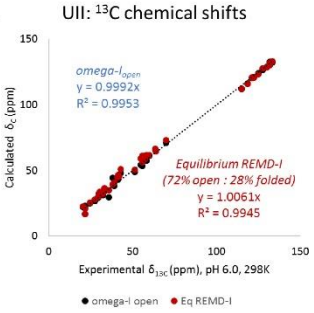
Linear regression of UII calculated and experimental ^{13}C chemical shifts. The ^{13}C statistics favor the conformation *omega-I$_{open}$* to be most similar to the experiment (pH6) followed by the equilibrium REMD-I which matches best for ^{1}H-metrics (**Table S13**). However, the ^{13}C-models show less distinctiveness (Δ_σ) than ^{1}H and the calculated chemical shifts are stronger correlated (R^2_{13C}=0.9970; R^2_{1H}=0.9676). This already makes the ^{13}C-models less reliable than the ^{1}H-models (**Figure S8**).

Analysis of ^{1}H and ^{13}C metrics-sensitivities. Nevertheless, a sensitivity analysis (examining the dependence of metrics on variation of the equilibrium ratio) has been made to test if the ^{1}H equilibrium-models underestimate the concentration of *open*-type conformations especially the *omega-I$_{open}$* conformation. This question is further interesting because *omega-I$_{open}$* resembles the conformation that has been suggested for UII in H_2O by Lescot et al.[35] (taking into account the different descriptions of turn-types). **Figure S9** shows the dependence of the metrics WRMSE and Δ_σ on the concentration of *open* conformations in *open:folded* mixtures (^{1}H, ^{13}C) and *omega-I$_{open}$* conformations in *omega-I$_{open}$:folded* mixtures (^{1}H). The *open:folded* models are mixtures of the 11 representatives with relative subtype concentrations corresponding to the equilibrium REMD-I. The plots for ^{1}H metrics (**Figure S9**a) show defined minima for *open:folded* mixtures with no preference for 100% *open* much less 100% *omega-I$_{open}$* conformations. Absolute error values are lower for the ^{1}H models if all 11 representatives are used. This makes them more reliable than the models with *omega-I$_{open}$* as sole representative for the *open* type. The highest distinctiveness (Δ_σ) is found for *open:folded* concentrations around 70:30 which resembles the REMD-prediction; the ^{1}H-WRMSE minimum of the 11-component model is slightly shifted to lower concentrations of *open* conformations. If the *open* type is solely represented by the *omega-I$_{open}$* conformation, the minimum shifts even more to lower concentrations and an *open:folded* ratio (20:80) that is no longer in accordance with the majority of metrics. The analogous analysis for ^{13}C (**Figure S9**b) shows a poor coefficient of distinctiveness for all *open:folded* ratios. Nevertheless, all minima suggest the optimum *open:folded* ratio is approximately 70:30 independent of the subtype mixture of *open* conformations.

Supporting Information

Table S13 Statistical metrics for the linear regression of calculated[a] and experimental[b] ^{13}C chemical shifts of UII [c]

UII representatives and equilibria (*open:folded*)	MSE	MUE	RMSD	WRMSE	Δ_σ	R^2
omega-I open	**-0.37**	**2.03**	**2.54**	**2.77**	**1.53**	**0.9953**
omega-I hbond	-1.39	2.54	3.39	3.78	1.71	0.9916
omega-II	-1.33	2.67	3.46	3.98	1.77	0.9913
lasso	-1.19	2.57	3.34	3.82	1.84	0.9917
scoop	-0.53	3.08	4.04	5.10	1.92	0.9893
circle	-1.44	2.51	3.35	3.88	1.70	0.9917
folded-I	-1.48	2.59	3.52	3.94	1.73	0.9911
folded-IVb2	-1.42	2.42	3.05	3.38	1.68	0.9935
inv-folded	-1.34	2.91	3.78	4.35	1.85	0.9902
folded-II	-1.06	2.78	3.70	4.17	1.86	0.9898
folded-III	-1.07	2.89	3.61	4.23	1.94	0.9903
Equilibrium REMD-I (72:28)	-1.13	2.15	2.74	2.94	1.57	0.9945
Equilibrium REMD-II (70:30)	-1.14	2.20	2.90	3.17	1.62	0.9938
Equilibrium REMD-III (79:21)	-1.17	2.20	2.91	3.16	1.63	0.9938



[a] GIAO, B3LYP/6-31G*, PCM water. [b] H_2O/D_2O, pH 6.0, 298 K. [c] The best results are shown in bold. MSE= Mean standard error, MUE= Mean unsigned error, RMSD= Root mean square deviation, WRMSE= weighted root mean square error, Δσ= coefficient of distinctiveness (<= 1.0 indicates that the average deviation between experimental and calculated values is less than the standard deviation between different conformations)[33]

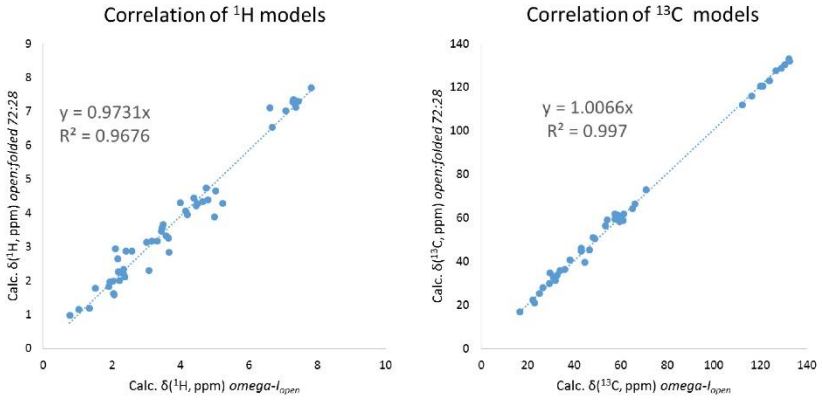


Figure S8 Correlation of ^1H and ^{13}C models
Linear regression of calculated ^1H and ^{13}C chemical shifts of models "REMD-equilibrium I (*open:folded* 72:28)" and "*omega-*I_{open}". The ^{13}C chemical shifts of the equilibrium model show a high correlation (similarity) to those of the single, which makes the ^{13}C-models less distinctive.

Supporting Information

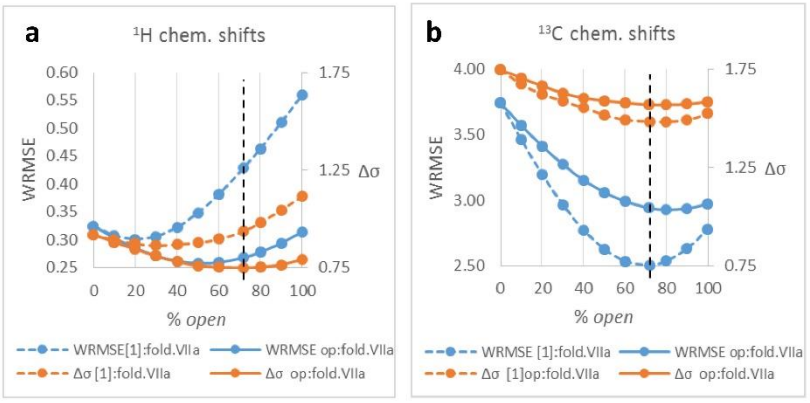


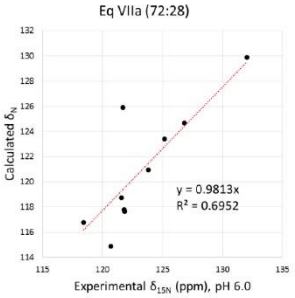
Figure S9 Metrics sensitivity.
Sensitivity analysis of the dependence of WRMSE and Δ_σ on different *open:folded* ratio of UII: (**a**) ^{1}H chemical shifts, (**b**) ^{13}C chemical shifts. Solid lines *open:folded* mixtures; dashed lines *omega-I$_{open}$:folded* mixtures; [**1**]= *omega-I$_{open}$* conformation; dashed vertical line = REMD-equilibrium I (72:28); WRMSE= weighted root mean square error; Δ_σ= coefficient of distinctiveness. Subtype concentrations of *open* and *folded* type are scaled from REMD-I concentrations

Linear regression of UII ^{15}N chemical shifts

Linear regression of the experimental ^{15}N chemical shifts (**Table S6**) with calculated shifts (**Table S11**) gives the same result as the ^{1}H regression. In both cases, the best performing model is the equilibrium REMD-I predicting a ratio of 72:28 for *open* and *folded* conformations of UII in aqueous solution (**Table S14**). However, the ^{15}N regression was only based on a data set of 10 experimental values which is reflected in bad coefficients of determinations compared to the ^{1}H models. ^{15}N modelling thus confirms the ^{1}H results but is not suitable to be used as a stand-alone method.

Table S14 Statistical metrics for the linear regression of calculated[a] and experimental[b] ^{15}N chemical shifts of UII [c]

^{15}N, pH6, CRC training set						
Representatives and equilibria (open:folded)	**MSE**	**MUE**	**RMSD**	**WRMSE**	**Δσ**	**R²**
omega-I open	3.09	3.21	4.72	5.01	**0.60**	0.2917
omega-I hbond	2.96	4.57	5.66	5.58	0.99	0.3964
omega-II	2.77	4.39	5.39	5.72	0.88	0.1321
lasso	2.36	4.39	4.98	4.99	0.96	0.1710
scoop	0.99	5.31	6.04	5.98	1.14	0.3982
circle	-0.82	5.75	6.63	7.07	1.06	0.0204
folded-I	3.59	5.10	6.01	6.14	1.05	0.4319
folded-IVb2	**0.12**	4.80	5.64	6.00	0.87	0.1426
inv-folded	2.12	7.04	7.82	8.30	1.39	0.3741
folded-II	1.30	4.51	5.19	5.36	0.93	0.3563
folded-III	2.42	6.49	7.83	7.86	1.43	0.4687
Equilibrium REMD-I (72:28)	2.32	**3.15**	**3.40**	**3.34**	0.69	**0.6952**
Equilibrium REMD-II (70:30)	2.36	3.40	4.02	3.95	0.76	0.5590
Equilibrium REMD-III (79:21)	2.38	3.41	3.82	3.79	0.74	0.5361



[a] GIAO, B3LYP/6-31G*, PCM water. [b] H_2O/D_2O, pH 6.0, 298 K. [c] The best results are shown in bold. MSE= Mean standard error, MUE= Mean unsigned error, RMSD= Root mean square deviation, WRMSE= weighted root mean square error, Δσ= coefficient of distinctiveness (<= 1.0 indicates that the average deviation between experimental and calculated values is less than the standard deviation between different conformations)[33]

Supporting Information

DP4 probabilities for UII and URP assignments

The DP4 probability by Smith and Godman[36] is a probability that the assignment of a set of calculated shifts to one set of experimental shifts is correct. The application is available under http://www-jmg.ch.cam.ac.uk/tools/nmr/DP4/.

Table S15 lists the DP4 probabilities separately for ^{13}C and ^{1}H assignments and for the recommended combination of ^{1}H and ^{13}C for the single conformations and the REMD equilibria of UII and URP. ^{1}H and $^{1}H/^{13}C$ DP4 probabilities favor unequivocally the *open:folded* ratio of 72:28 for UII and 86:14 for URP to be the best possible assignment in accordance with our predictions based on REMD-populations. DP4 probabilities only using ^{13}C chemical shifts are not unambiguous. The possible reasons have already been discussed for our metrics (too low distinctiveness because of too high correlation of the compared models).

Table S15 DP4 probabilities for correct assignment of calculated to experimental chemical shifts of single conformations and equilibrium mixtures of UII and URP

	DP4 probability (%)		
UII representatives and equilibria (*open:folded*)	**^{13}C**	**^{1}H**	**$^{13}C+^{1}H$**
omega-I_{open}	**24.6**	0.0	0.0
omega-I_{hbond}	0.0	0.0	0.0
omega-II	0.0	0.0	0.0
lasso	0.0	0.0	0.0
scoop	0.0	0.0	0.0
circle	0.0	0.0	0.0
folded-I	0.0	0.0	0.0
folded-IVb2	0.0	0.0	0.0
inv-folded	0.0	0.0	0.0
folded-II	0.0	0.0	0.0
folded-III	0.0	0.0	0.0
Equilibrium REMD-I (72:28)	**68.5**	**100.0**	**100.0**
Equilibrium REMD-II (70:30)	**3.5**	0.0	0.0
Equilibrium REMD-III (79:21)	**3.4**	0.0	0.0
URP representatives and equilibria (open:folded)	**^{13}C**	**^{1}H**	**$^{13}C+^{1}H$**
omega-I_{open}	0.0	0.0	0.0
omega-I_{hbond}	0.0	0.0	0.0
omega-II	7.6	0.0	0.0
lasso	0.0	0.0	0.0
sheet	0.0	0.0	0.0
hybrid	1.5	0.0	0.0
Equilibrium REMD-IV (86:14)	**16.8**	**100.0**	**100.0**
Equilibrium REMD-V (94:6)	**19.9**	0.0	0.0
Equilibrium REMD-VI (91:9)	**54.2**	0.0	0.0

3D structural data

Coordinate files in PDB format of representative conformations of UII and URP are attached in the supplementary file: Representative_conformations_of_UII_and_URP.zip

Supporting Information

References

1. Carotenuto, A.; Grieco, P.; Campiglia, P.; Novellino, E.; Rovero, P., Unraveling the active conformation of urotensin II. *J Med Chem* **2004,** 47, 1652-61.
2. Chatenet, D.; Dubessy, C.; Leprince, J.; Boularan, C.; Carlier, L.; Ségalas-Milazzo, I.; Guilhaudis, L.; Oulyadi, H.; Davoust, D.; Scalbert, E.; Pfeiffer, B.; Renard, P.; Tonon, M.-C.; Lihrmann, I.; Pacaud, P.; Vaudry, H., Structure–activity relationships and structural conformation of a novel urotensin II-related peptide. *Peptides* **2004,** 25, 1819-1830.
3. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* **1983,** 79, 926-935.
4. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T., Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *The Journal of chemical physics* **2004,** 120, 9665-9678.
5. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular Dynamics with Coupling to an External Bath *Journal of Chemical Physics* **1984,** 81, 3684-3690.
6. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* **1977,** 23, 327-341.
7. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **1993,** 98, 10089-10092.
8. Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry* **2009,** 30, 1545-614.
9. Leprince, J.; Chatenet, D.; Dubessy, C.; Fournier, A.; Pfeiffer, B.; Scalbert, E.; Renard, P.; Pacaud, P.; Oulyadi, H.; Segalas-Milazzo, I.; Guilhaudis, L.; Davoust, D.; Tonon, M. C.; Vaudry, H., Structure-activity relationships of urotensin II and URP. *Peptides* **2008,** 29, 658-73.
10. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters* **1999,** 314, 141-151.
11. Penev, E. S.; Lampoudi, S.; Shea, J. E., TiREX: Replica-exchange molecular dynamics using TINKER. *Computer Physics Communications* **2009,** 180, 2013-2019.
12. Kofke, D. A., On the acceptance probability of replica-exchange Monte Carlo trials. *Journal of Chemical Physics* **2002,** 117, 6911-6914.
13. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J Comput Chem* **2005,** 26, 1668-1688.
14. Patriksson, A.; van der Spoel, D., A temperature predictor for parallel tempering simulations. *Phys Chem Chem Phys* **2008,** 10, 2073-7.
15. Salt, D. W.; Hudson, B. D.; Banting, L.; Ellis, M. J.; Ford, M. G., DASH: A Novel Analysis Method for Molecular Dynamics Simulation Data. Analysis of Ligands of PPAR-Gamma. *J Med Chem* **2005,** 48, 3214-3220.
16. Ellis, M. J.; Hudson, B.; Salt, D. DASH 1.0. www.port.ac.uk/research/cmd/software (accessed Oct 15th, 2014)
17. Ellis, M. J.; Hudson, B.; Salt, D.; Whitley, D. C. DASH 2.11. www.port.ac.uk/research/cmd/software (accessed Mar 17th, 2016)
18. Lewis, P. N.; Momany, F. A.; Scheraga, H. A., Chain reversals in proteins. *Biochim Biophys Acta* **1973,** 303, 211-29.

Supporting Information

19. Ramachandran, G. N.; Venkatachalam, C. M., Stereochemical criteria for polypeptides and proteins. IV. Standard dimensions for the cis-peptide unit and conformation of cis-polypeptides. *Biopolymers* **1968,** 6, 1255-62.
20. Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M., Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation Nmr-Spectroscopy. *Journal of Magnetic Resonance* **1991,** 93, 151-170.
21. Kay, L. E.; Keifer, P.; Saarinen, T., Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *Journal of the American Chemical Society* **1992,** 114, 10663-10665.
22. Kontaxis, G.; Stonehouse, J.; Laue, E. D.; Keeler, J., The Sensitivity of Experiments Which Use Gradient Pulses for Coherence-Pathway Selection. *Journal of Magnetic Resonance Series A* **1994,** 111, 70-76.
23. States, D. J.; Haberkorn, R. A.; Ruben, D. J., A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in 4 Quadrants. *Journal of Magnetic Resonance* **1982,** 48, 286-292.
24. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J Biomol NMR* **1995,** 6, 277-293.
25. Johnson, B. A.; Blevins, R. A., Nmr View - a Computer-Program for the Visualization and Analysis of NMR Data. *J Biomol NMR* **1994,** 4, 603-614.
26. Wassenaar, T. A.; van Dijk, M.; Loureiro-Ferreira, N.; van der Schot, G.; de Vries, S. J.; Schmitz, C.; van der Zwan, J.; Boelens, R.; Giachetti, A.; Ferella, L.; Rosato, A.; Bertini, I.; Herrmann, T.; Jonker, H. R. A.; Bagaria, A.; Jaravine, V.; Güntert, P.; Schwalbe, H.; Vranken, W. F.; Doreleijers, J. F.; Vriend, G.; Vuister, G. W.; Franke, D.; Kikhney, A.; Svergun, D. I.; Fogh, R. H.; Ionides, J.; Laue, E. D.; Spronk, C.; Jurkša, S.; Verlato, M.; Badoer, S.; Dal Pra, S.; Mazzucato, M.; Frizziero, E.; Bonvin, A. M. J. J., WeNMR: Structural Biology on the Grid. *Journal of Grid Computing* **2012,** 10, 743-767.
27. CcpNmr Analysis, version 2.2.2. http://www.ccpn.ac.uk/software/analysis (accessed Mar 13th, 2015)
28. Dorman, D. E.; Bovey, F. A., Carbon-13 magnetic resonance spectroscopy. Spectrum of proline in oligopeptides. *J Org Chem* **1973,** 38, 2379-2383.
29. Siemion, I. Z.; Wieland, T.; Pook, K. H., Influence of the distance of the proline carbonyl from the beta and gamma carbon on the 13C chemical shifts. *Angew Chem Int Ed Engl* **1975,** 14, 702-3.
30. Schubert, M.; Labudde, D.; Oschkinat, H.; Schmieder, P., A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics. *Journal of Biomolecular NMR* **2002,** 24, 149-54.
31. Shen, Y.; Bax, A., Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *Journal of Biomolecular NMR* **2010,** 46, 199-204.
32. Wüthrich, K., *NMR of Proteins and Nucleic Acids*. Wiley: New York; Chichester, 1986; p 292.
33. Haensele, E.; Saleh, N.; Read, C. M.; Banting, L.; Whitley, D. C.; Clark, T., Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine-Vasopressin. *Journal of Chemical Information and Modeling* **2016**.
34. Bruno, T. J.; Svoronos, P. D. N., *CRC Handbook of Basic Tables for Chemical Analysis, Third Edition*. CRC Press: 2010.
35. Lescot, E.; Sopkova-de Oliveira Santos, J.; Dubessy, C.; Oulyadi, H.; Lesnard, A.; Vaudry, H.; Bureau, R.; Rault, S., Definition of new pharmacophores for nonpeptide antagonists of human urotensin-II. Comparison with the 3D-structure of human urotensin-II and URP. *J Chem Inf Model* **2007,** 47, 602-12.
36. Smith, S. G.; Goodman, J. M., Assigning Stereochemistry to Single Diastereoisomers by GIAO NMR Calculation: The DP4 Probability. *Journal of the American Chemical Society* **2010,** 132, 12946-12959.

A 4: Additional Analyses

Convergence of Tail Conformations for AVP

The high frequency of *folded/extended* interconversions during the 11 µs MD simulation of AVP suggested that the ratio *folded:extended* states had converged. Further evidence for this was provided by a comparative analysis of the populations of the T6 tail states after 11 and 23 µs. Independent of the simulation time, the populations of *folded* and *extended* states showed no significant changes (Table A4.1).

Table A4.1 Populations of tail conformations from 11 and 23 µs MD simulation of AVP

Tail state (T6)	Tail state population (%)	
	0-11µs[a]	0-23µs
	Extended Tail	
3	61.4	61.2
4	19.6	20.4
Total	*81.0*	*81.6*
	Folded Tail	
	7,8 β-turn type II	
5	2.5	1.5
6	13.6	14.4
	7,8 β-turn type I	
1	2.0	1.4
	distorted	
2	0.8	1.2
Total	*19.0*	*18.4*

T6 = torsions $\Phi\Psi$ 7 to 9. [a] published[1]

Data Consistency AVP (11 µs) vs. AVP (23 µs)

For the NMR-modelling,[2] representative overall states with *extended* and *folded* tail were chosen for the *saddle* and *clinched open* conformation, and the exact ratio of these states were calculated to refine the conformational equilibria. Two questions arise:

(i) Is the ratio of the two single representative states (one with *extended*, the other with *folded* tail) in accordance with the *extended/folded* tail-populations deduced from all overall states (23 µs)?

(ii) Are the populations of *extended* and *folded* tail states deduced from 23 µs (all overall states) in accordance with the previously published populations[1] deduced from 11 µs?

Table A4.2 lists the populations of *extended* and *folded* tail conformations deduced from all overall states (23 µs) and *tail* states (11 µs). The populations are sufficiently consistent to draw the following conclusions:

(i) The representatives used for NMR modelling are suitable to represent the conformational subtypes of the ring conformations *saddle* and *clinched open* with *extended* or *folded* tail.

(ii) The populations of tail conformations in relation to the ring conformations deduced from 23 µs are in accordance with those published previously for 11 µs.

Table A4.2 Population of tail conformations for the four main ring conformations of AVP

	Tail conformation					
	All *overall* states (T16)		**All *tail* states (T6)**		**Representatives (T16)**	
Ring conformation	**23µs MD**		**11µs MD**		**23µs MD**	
	ext	*fold*	*ext*	*fold*	*ext*	*fold*
saddle	77.06	22.94	76.34 [a]	23.66 [a]	73.14 [b]	26.86 [b]
cl.open	71.87	28.13	68.34 [a]	31.66 [a]	62.63 [b]	37.37 [b]
tw.saddle	89.28	10.72	83.44 [a]	16.56 [a]	82.54	17.46
open	90.15	9.85	93.61 [a]	6.39 [a]	87.99	12.01

[a] data published in reference[1]; [b] data published in reference[2]

Circular Similarity: Representative Conformations vs. Literature

Oxytocin.

Table A4.3 Circular similarity of OT representatives (this work) and conformations from the literature [§]

OT [a]	Lippens[92] [b]	Ward[93] [c]	Nikiforovich[94] [d]
saddle	0.55	0.45	0.45
tws	0.63	0.42	0.43
twshelix	0.46	0.51	0.52
int.saddle	0.46	0.42	0.42
$open_{23pbr}$	0.45	0.40	0.40
open	0.30	0.36	0.37
clop	0.68	0.45	0.45
$clop_{45pbr}$	0.60	0.44	0.44

[§] Maximum similarities are highlighted. [a] Representative conformations of OT from a total of 50 µs MD simulations (EH, this work). [b] OT conformation bound to NP, NMR. [c] Conformer "MD, Table 8" (lowest energy conformation predicted by MD simulation). [d] Conformer "I, Table 5" (potential energy minimised conformation).

Urotensin II, Urotensin-Related Peptide.

Table A4.4 Circular similarity of UII and URP representatives (this work) and conformations from the literature [§]

UII [a]	ID [b]	Grieco[129] [c]	URP [a]	ID [b]	Chatenet[125] [d]
folded-I	**6**	0.56			
folded-IVb2	**7**	0.67	*hybrid*	**4r**	0.57
			sheet	**5r**	0.55
folded-II	**8**	0.40			
folded-III	**9**	0.37			
circle	**10**	0.46			
inv.folded	**11**	0.41			
$omega\text{-}I_{open}$	**1**	0.54	$omega\text{-}I_{open}$	**3r**	0.86
$omega\text{-}I_{hbond}$	**2**	0.67	$omega\text{-}I_{hbond}$	**1r**	0.73
omega-II	**3**	0.54	*omega-II*	**2r**	0.62
lasso	**4**	0.47	*lasso*	**6r**	0.59
scoop	**5**	0.62			

[§] Maximum similarities are highlighted. [a] Representative conformations of UII from a total of 35 µs MD simulations (EH, Paper 3). [b] ID of representative conformation, *cf.* Table 6.1 and 6.6, Chapter 6 (Paper 3). [c] UII in DMSO, NMR. [d] URP in H_2O, NMR.

A 5: Classical Turn Types

A turn can be defined by the turn centres involved and the $\Phi\Psi$ torsion angles at these residues. The sequence of torsions defines the turn type. All residues involved in a turn are numbered consecutively (i, i+1, i+2, i+3…). Ideal turn types as defined by Venkatachalam *et al.*[277] and Richardson *et al.*[58] are given in Table A5.1 (for further turn types, see *e.g.*[323]). A turn is denoted as *classical* if the neighbour residues of the turn centres form a hydrogen bond. However, approximately 25 % of all β-turns do not possess a hydrogen bond as stipulated by Venkatachalam. Therefore, Lewis *et al.*[228] suggested a distance maximum of 7 Å for $C\alpha_i$ to $C\alpha_{i+3}$ as a new criterion to define a β-turn. If there is no hydrogen bond and the dihedrals fluctuate around ideal torsion angles, the turn can be denoted as *open* turn. The notation for turns and other secondary-structure elements used in this thesis is explained in the Supporting Information of Paper 3 (Appendix A3, p S10).

Table A5.1 Ideal turn types

Turn type	Torsion				Distance $C\alpha_{i,i+3}$
	Φ_{i+1}	Ψ_{i+1}	Φ_{i+2}	Ψ_{i+2}	
β-I	-60°	-30°	-90°	0°	$d_{\alpha\alpha}$ < 7 Å
β-III (≈I) [a]	-60°	-30°	-60°	-30°	$d_{\alpha\alpha}$ < 7 Å
β-VIII	-60°	-30°	-120°	120°	$d_{\alpha\alpha}$ < 7 Å
β-II	-60°	120°	80°	0°	$d_{\alpha\alpha}$ < 7 Å
β-I’	+60°	+30°	+90°	0°	$d_{\alpha\alpha}$ < 7 Å
β-III’ (≈I’) [a]	+60°	+30°	+60°	+30°	$d_{\alpha\alpha}$ < 7 Å
β-II’	+60°	-120°	-80°	0°	$d_{\alpha\alpha}$ < 7 Å
3_{10}-helix [b]	-49°	-26°	-49°	-26°	$(O_{i-3}H_i)_n$
$\gamma_{classical}$	+75°	-65°			
$\gamma_{inverse}$	-75°	+65°			

[a] Under dynamic conditions (stddev ±30° assumed). [b] Right-handed, $\Phi+\Psi\approx$-75°. i = residue number.

A 6: Physical Laws and Definitions

This chapter is thought to supplement the physical laws and quantities used in this work with some background knowledge. General sources have been textbooks of physics[324], textbooks of chemical physics[325] and special lexica[326]. Topics are given alphabetical.

Coulomb´s Law. Electrostatic force:

$$F = k_e \frac{q_1\, q_2}{r^2} = \frac{1}{4\pi\varepsilon_0}\frac{q_1\, q_2}{r^2}$$

k_e Coulomb constant, $q_1 q_2$ charges of atom 1 and 2, r distance between atom 1 and 2, ε_0 electric constant

Ergodic Hypothesis. The ergodic hypothesis comes from the statistical mechanics and says that the time-average equals the ensemble-average if the trajectory of a dynamic system has passed every possible point (in its phase space).[326]

Force and Force Constant. A physical quantity that changes the state of a body, *e.g.* the position of a body and thus, his potential energy. A force constant in this context defines the shape of the harmonic potential that is used to approximate the potential energy of the system.

$$\vec{F} = m\vec{a}$$

m mass, $\vec{a}$ velocity

Gibb´s Free Energy.

$$G = H - TS = U + pV - TS$$

with G Gibb´s free energy, H free enthalpy, T temperature, S entropy, U potential energy (inner energy), p pressure, V volume

In biochemistry, the free standard enthalpy is defined as

$$\Delta G'^{0}_{m} = -RTlnK_{eq}, \qquad R = N_A k_B$$

$$R = 8.3144598\ \frac{J}{K\ mol} = 0.001987204\ \frac{kcal}{K\ mol}$$

$$N_A = 6.022140857 \cdot 10^{23}\,\frac{1}{mol}, \quad k_B = 1.38064852 \cdot 10^{-23}\,\frac{J}{K}$$

with R ideal gas constant, T absolute temperature, K_{eq} equilibrium constant, N_AAvogadro constant, k_B Boltzmann constant

Biochemical standard conditions: 25°C (298.15K), p = 1 atm, a_i = 1 (chemical activity), pH = 7

Hook´s Law.

$$V(l) = \frac{k}{2}(l - l_0)^2$$

$V(l)$ potential energy relative to position, k force constant, l position, l_0 origin

Lennard-Jones Potential. Approximation of the non-bonding interactions (energy) between uncharged atoms. Attractive interactions are *e.g.* van-der-Waal´s forces or permanent dipole-dipole interactions.

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \qquad \text{and} \qquad r_{min} = \sqrt[6]{\sigma}$$

ε depth of potential minimum [J], σ distance where $V(r) = 0$, r distance between atoms; $\left(\frac{\sigma}{r}\right)^{12}$ Pauli repulsion; $\left(\frac{\sigma}{r}\right)^{6}$ attractive long-range term (v.d.Waals force or dispersion force)

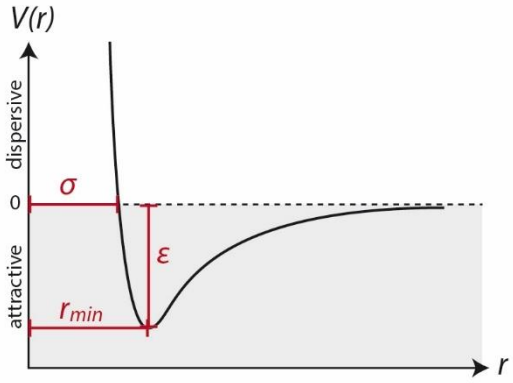


Figure A6.1 Lennard-Jones Potential

Newton´s Laws of Motion.[148]

(1) "*Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus illud a viribus impressis cogitur statum suum mutare*"[i]

(2) *„Mutationem motus proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimitur."*[ii]

$$\dot{\vec{v}} \propto \vec{F} \text{ (Euler´s form: } \vec{F} = m\vec{a})$$

(3) *"Actioni contrariam semper et aequalem esse reactionem: sive corporum duorum actiones in se mutuo semper esse aequales et in partes contrarias dirigi." (action = reaction)*[iii]

$$\vec{F}_{A\to B} = \vec{F}_{B\to A}$$

(4) (Superposition principle)

$$\vec{F}_{res} = \vec{F}_1 + \vec{F}_1 + \cdots + \vec{F}_n$$

[i] "Every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed." (Trägheitsprinzip, Impulserhaltungsgesetz)

[ii] "The alteration of motion is ever proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed." (Aktionsprinzip)

[iii] "To every action there is always opposed an equal reaction: or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts."

Nuclear Magnetic Shielding Tensor.[246] The nuclear magnetic shielding (dimensionless unit) is a 2nd order property of the electronic energy under the influence of a magnetic field. The shielding tensor is antisymmetric. Exchange of α and β results different quantities.

$$\sigma_{\alpha\beta} = \frac{\delta^2 E}{\delta\mu_\alpha \delta B_\beta}$$

E total electronic energy of the molecule; B external magnetic field; μ magnetic moment of nucleus.

Root Mean Square Deviation. A measure of the geometrical difference between two conformations.

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \delta_i^2}$$

$RMSD$ root mean square deviation, N number of atoms; δ coordinate difference of atom i relative to intial conformation

A 7: Hardware and Software

This chapter gives background information on hardware and software used in this project. It comprises the sections *Simulation Tools*, *Analysis Tools*, and *Graphic Tools*.

Simulation Tools

LEaP (AmberTools)

LEaP (Link Edit and Parm) is a program for the preparation of topology and initial coordinate files. It can be used *via* a graphics interface or in text mode. Detailed instructions are given in the AmberTools manuals.[142,143]

Example. Preparation of topology and initial coordinate files for the unrestrained simulation MD-XI of UII based on coordinates of a significantly populated state (*inv-folded*) found in REMD simulation REMD-I.[3]

LEaP input file

```
verbosity 0
#
source leaprc.ff99SB
loadoff ions08.lib
mod3 = loadamberparams frcmod.tip4pew
jc_ions = loadamberparams frcmod.ionsjc_tip4pew
prot = loadpdb 11_UIIre_21440_state1.pdb
bond prot.5.SG prot.10.SG
addions prot Na+ 0
solvateoct prot TIP4PEWBOX 8.0
saveamberparm prot uii_invf.top uii_invf.crd
```

Input file giving the instructions to build a topology file uii_invf.top and an initial coordinate file uii_invf.crd of structure 11_UIIre_21440_state1.pdb with counterions (addions…) to get a total charge of 0, solvation with a truncated octahedral box of TIP4P-Ew water and a non-bonded cut-off of 8 Å (solvateoct…), a disulphide linkage of residues 5 and 10 (bond…) and the parameters for force field ff99SB (source…) with matching water model (mod3…) and ion parameters (jc_ions…).

Command line: tleap -f input_file

SANDER

SANDER is the main program within the AMBER package to run energy minimisations and MD simulations. Here, the PMEMD implementation for high-performance parallel processing was used. Example files for energy minimisation and MD simulation are given below.

Minimisation Example. Input file for energy minimisation.

```
Minimisation of UII inverted-
folded state
 &cntrl
  imin   = 1,
  maxcyc = 10000,
  ncyc   = 500,
  ntb    = 1, iwrap=1,
  ntr    = 0,
  cut    = 8,
  ntpr=50, ntwr=50, ntwx=50
```

Parameter setup to perform a minimisation (*imin*=1) of maximum 10,000 steps (*maxcyc*) with 500 steps steepest descent (*ncyc*) continued by conjugated gradient minimisation under periodic boundary conditions and constant volume (*ntb*=1); no restraints (*ntr*=0); restart file is written back to the original solvation box (*iwrap*=1); non-bonded cutoff is 8 Å (*cut*=8, default for PME); results are written to the output every 50th step the (*ntpr*=50, = every 0.1 ps); every 50th step a restart file is written (*ntwr*); every 50th step coordinates are written to a coordinate file (*ntwx*=50).

Command line (GPU): pmemd.cuda –O –i min.in –c min.crd –o min.out –r min.rst –x min.trj –p uii_invf.top
(with –i input, -c coordinate file, -o output file, -r restart file, -x trajectory file, -p topology)
Command line (CPU, 8 cores): mpirun -np 8 pmemd -O -i min.in -c min.crd -o min.out -r min.rst -x min.trj –p uii_invf.top
(with –np number of cores)

MD Simulation Example.

Input 1, unrestrained MD simulation starting from a minimised conformation

```
Heating to 300K

&cntrl
  irest=0, ntx=1, imin=0, nmropt=0, nstlim=125000000, dt=0.002,
  nsnb=25, cut=8.0,
  ntt=1, tautp=1.0, tempi=0.0, temp0=300.0,
  ntc=2, ntf=2, ntb=2,
  ntp=1, taup=1.0, pres0=1.0,
  ntpr=5000, ntwr=5000, ntwx=5000, ntr=0, iwrap=1, ioutfm=1,
 &end
```

Instructions to start a new (*irest*=0), unrestrained (*ntr*=0) MD simulation; reading the initial file without velocities (*ntx*=1); no minimisation (*imin*=0); 125,000,000 MD-steps to be performed (*nstlim,* =250 ns); maximum time step 0.002 ps (*dt*); default frequency of non-bonded list updates (*nsnb*=25); 8 Å non-bonded cutoff (*cut*); constant temperature *via* weak coupling (*ntt*=1) of 1.0 ps to external bath (*tautp*=1.0); intital temperature 0 K (*tempi*), reference temperature default (*temp0*=300.0); use SHAKE algorithm (*ntc*=2); omit force evaluation involving H-atoms if SHAKE is on (*ntf*=2); constant pressure periodic boundary conditions (*ntb*=2) with isotropic position scaling (*ntp*=1), default relaxation time in ps (*taup*=1.0), and 1 bar (default) reference pressure (*pres0*); results (*ntpr*), coordinates (*ntwr*), and restart files (*ntwx*) are written to outputs every 5000th step (=10 ns); restart file is written back to the original solvation box (*iwrap*=1); NETCDF format for trajectory files output files (*ioutfm*=1)

Command line (GPU): pmemd.cuda –O –i md1.in –c md1.crd –o md1.out –r md1.rst –x md1.trj –p uii_invf.top
(with –i input file 1, -c coordinate file 1 (= min.crd), -o output file 1, -r restart file 1, -x trajectory file 1, -p topology)
Command line (CPU, 8 cores): mpirun -np 8 pmemd -O -i md1.in -c md1.crd -o md1.out -r md1.rst -x md1.trj –p uii_invf.top
(with –np number of cores)

MD Simulation Example.

Input 2, subsequent restarts

```
production, T = 300K

&cntrl
  irest=1, ntx=5, imin=0, nmropt=0, nstlim=125000000, dt=0.002,
  nsnb=25, cut=8.0,
  ntt=1, tautp=1.0, tempi=0.0, temp0=300.0,
  ntc=2, ntf=2, ntb=2,
  ntp=1, taup=1.0, pres0=1.0,
  ntpr=5000, ntwr=5000, ntwx=5000, ntr=0, iwrap=1, ioutfm=1,
 &end
```

Instructions to continue (irest=1) a MD simulation using coordinates, velocities and box size from an external input file (here the NETCDF trajectory file of the preceding MD simulation run); all other parameters are identical with Input 1.

Command line (GPU): pmemd.cuda –O –i md2.in –c md2.crd –o md2.out –r md2.rst –x md2.trj –p uii_invf.top
(with –i input file 2, -c coordinate file 2 (= md1.crd), -o output file 2, -r restart file 2, -x trajectory file 2, -p topology)
Command line (CPU, 8 cores): mpirun -np 8 pmemd -O -i md2.in -c md2.crd -o md2.out -r md2.rst -x md2.trj –p uii_invf.top
(like GPU command, with –np number of cores)

AMBER 10 on CPUs (Central Processing Unit)

From 2011 to 2016, most MD simulations were performed with the AMBER_10 PMEMD implementation of SANDER, Release 10-14[219] on eight *Harpertown Intel Xeon E5462* cores (@ 2.83 GHz), a multi-node cluster of the Computer-Chemie-Centre of the FAU Erlangen-Nürnberg (for performance, see Table A7.1).

AMBER 14 on GPUs (Graphics Processing Unit)

In 2015 and 2016, MD simulations were also performed with the AMBER_14 PMEMD.CUDA (*Computer Unified Device Architecture Language*) implementation of SANDER, Release 14[206,221,222] on a *4,800 MB NVIDIA Tesla K20c* graphic card with 2496 CUDA cores (@ 0.71 GHz) and a *5,375 MB NVIDIA C2075* graphic card with 448 CUDA cores (@1.15 GHz), two single node machines of the Computer-Chemie-Centre of the FAU Erlangen-Nürnberg (for performance, see Table A7.1).

Performance of AMBER 10 (CPU) and AMBER 14 (GPU) MD simulations

Table A7.1 Performance of AMBER 10 (CPU) and AMBER 14 (GPU) on unrestrained MD simulations of cyclic peptides

	System Size			Performance				Reference	
Peptide	NAtoms [a]	% water [b]	µs [c]	d/µs (CPU) [d]	ns/d (CPU) [d]	d/µs (GPU) [e]	ns/d (GPU) [e]	MD-ID[f]	Thesis
AVP	142	4792	97.0	23.0	38.2	26.2	7.5	133.3	MinMD_1YF4	Chapter 4,5[1,2]
UII	181	6154	97.0	5.0	48.8	20.5	-	-	UII_MD-I	Chapter 6[3]
	181	5754	96.8	5.0	45.1	22.2	-	-	UII_MD-II	Chapter 6[3]
	181	6642	97.3	10.0	34.6	28.9	10.4	96.2	UII_MD-III	Chapter 6[3]
	181	10082	98.1	5.0	88.4	11.3	15.9	62.9	UII_MD-IV	Chapter 6[3]
	181	4338	95.8	5.0	-	-	7.3	-	UII_MD-V	Chapter 6[3]
	181	5402	96.6	5.0	-	-	9.5	-	UII_MD-XI	Chapter 6[3]
URP	136	4153	96.7	5.0	30.7	32.6	-	-	URP_MD-IXa	Chapter 6[3]
	136	3633	96.2	5.0	28.6	35.0	-	-	URP_MD-IXb	Chapter 6[3]
	136	3965	96.5	5.0	29.6	33.8	-	-	URP_MD-IXc	Chapter 6[3]
	136	5021	97.3	5.0	38.8	25.8	-	-	URP_MD-IXd	Chapter 6[3]
OT	136	4029	96.6	15.0	29.9	33.4	9.1	109.9	OT_MD-I	Chapter 7
	136	4985	97.3	15.0	38.5	26.0	6.5	153.8	OT_MD-II	Chapter 7
	136	6241	97.8	10.0	46.3	21.6	10.6	94.3	OT_MD-III	Chapter 7
	136	5121	97.3	10.0	42.0	23.8	8.4	119.0	OT_MD-IV	Chapter 7
dOT	133	3961	96.6	3.0	29.5	33.9	-	-	dOT_MD-I	Chapter 7
CT	138	4446	96.9	5.3	32.7	30.6	-	-	CT_MD-I	Chapter 7
	138	6802	98.0	5.0	58.5	17.1	-	-	CT_MD-II	Chapter 7
av	***151***	***5307***	***97.0***	-	***41.3***	***26.4***	***9.5***	***109.09***		
total	-	-	-	***141.3***	-	-	-	-		

[a] Total number of atoms (peptide, water, ions); [b] Percentage water atoms; [c] Simulation time; [d] 8 CPU cores @ 2.83 GHz (Intel Xeon E5462 Harpertown); [e] 2496 CUDA cores @ 0.71 GHz (NVIDIA Tesla K20c) and 448 CUDA cores @ 1.15 GHz (NVIDIA C2075); [f] ID or working title of corresponding MD simulation

Analysis Tools

DASH

Principles and Parameters. A conformation can be defined by the sequence of *ΦΨ* torsions on which a DASH analysis is based. The torsions are extracted from the coordinate files of the MD simulations with *ptraj* (or *cpptraj*) and gathered as a torsion trajectory. This torsion trajectory is the input for the DASH program. DASH clusters the torsion space as a time series of DASH states. A DASH state is characterised by mean torsion angles (cluster centre). The representative of such a mean torsion ensemble is the torsion-trajectory snapshot with the highest similarity to the mean torsions. To visualise the representative, the corresponding frame of the original AMBER trajectory with the Cartesian coordinates can be extracted. The main parameters for DASH are the bout length (l) and the number of steps (n) (analysed frames) that define the minimum lifetime of a state:

$$state\ lifetime = \frac{bout\ length\ (l)}{time\ steps\ (n)} * simulation\ time$$

Reducing the minimum state lifetime will increase the number of final states by means of sub-clustering main conformations. A state refinement can be achieved if the number of time steps is increased while the minimum lifetime is held constant. In this thesis, a bout length of l = 20 was chosen as standard for DASH analyses, which equals a minimum lifetime of 10 ns for a distinct ensemble of mean torsions to be considered as DASH state. This gives a reasonable number of DASH states to classify main conformations with maximum performance.

Performance and Accuracy. A comparison of the performance of the classical pairwise-metrics cluster methods *average-linkage* and *means* against DASH is shown in Table A7.2. For these calculations, the ring conformations were clustered by the Cα and S atoms of the ring (*average-linkage* and *means*) and the dihedral angles *ΦΨ* 2-6 and 1χ2, 1χ3, 6χ2 (DASH), respectively. The first five microseconds of the MD simulation of AVP[1,2] were chosen as dataset. *Average-linkage* is a hierarchical bottom-up algorithm starting with individual conformations that are iteratively merged to a defined final number of clusters. *Means* is a refinement algorithm starting with a predefined number of seeds that are resorted (refined). Both methods require a predefined number of clusters. DASH is a sequential analysis method and does not need any predefined number of clusters.
The RMSD trajectory of the ring (Cα, S) and the DASH state trajectories of the analysis of 10,000 and 5,000,000 frames are shown in Figure A7.1. The cluster results of *average-linkage* and *mean* are shown in Figures A7.2 to A7.4.

A significant conformational transition occurs at 1.46 µs (Fig. A7.1), which is the interconversion from ring type *open* to *saddle*. All methods tested reproduce these two main conformational ring types but differ in the description of subtypes. *Means* had difficulty identifying unambiguously populated (>80 %) cluster bins within the *open* section. Cluster bins of *average-linkage* were more highly populated (Fig. A7.4) but the method showed poor performance compared to *means,* even with small datasets of 5,000 time steps. DASH performed almost 30 times faster than *means* when 10,000 time steps were analysed. Even when every single frame of the simulation was analysed (= 5 million conformations), the run time of DASH was less than one hour. Classical cluster methods cannot process such data volumes (*cf. means*: doubling the data volume from 5,000 to 10,000 time steps already quadrupled the runtime).

Table A7.2 Performance of clustering methods DASH, Average-Linkage and Means

Method	Time steps	Defined Clusters	Final Clusters	Main Clusters [a]	Runtime [min]	Cluster Trajectories
			Ring Alignment [b]			
Means	5,000	5	5	3	13	Fig. A7.2
Av.Linkage	5,000	5	5	2	45	Fig. A7.4
Means	10,000	5	5	3	58	Fig. A7.3
DASH	10,000	-	10	6	2	Fig. A7.1
DASH	5,000,000	-	16	5	55	Fig. A7.1

[a] Population > 5 %. [b] Ring and disulphide bridge torsions (T13ss: $\Phi\Psi$ 2-6, 1χ2, 1χ3, 6χ2) or ring and sulphur atoms (:1-6@CA, S*).

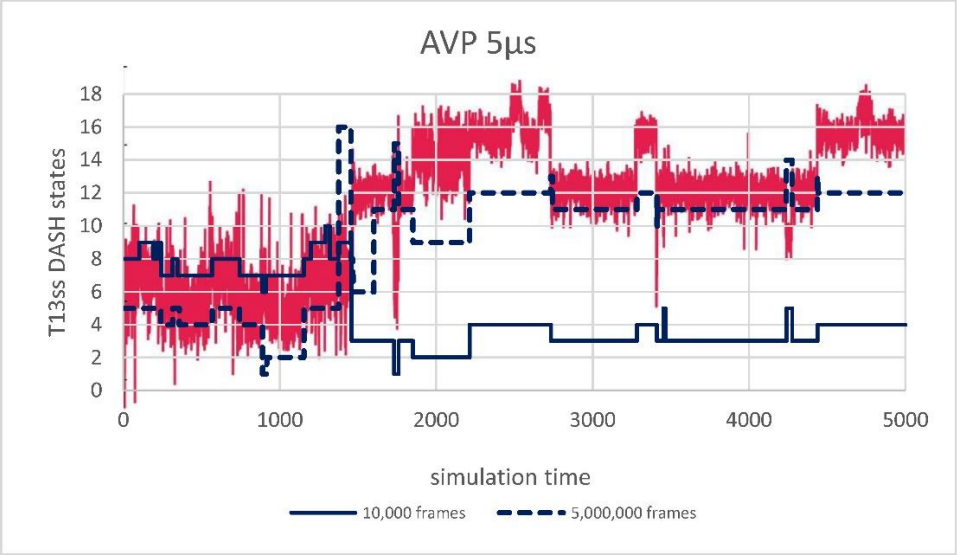


Figure A7.1 RMSD and DASH state trajectory of 5 µs MD simulation of AVP

```
#                       Condensed Map
#
#    5000 points divided into 50 windows, each window contains 100 points.
#
#
#Clustering: divide 5000 points into 5 clusters
#Cluster    0: has    94 points, occurrence 0.019
#Cluster    1: has  3530 points, occurrence 0.706
#Cluster    2: has   568 points, occurrence 0.114
#Cluster    3: has   606 points, occurrence 0.121
#Cluster    4: has   202 points, occurrence 0.040
#Cluster    0   .... 111.  .11.
#Cluster    1   .             3XX9XXXXXXXXXXXXXXX9XXXXXXXXXXXXXXX
#Cluster    2     4594 39782...
#Cluster    3   7932.373...4665  .               .
#Cluster    4   1.21...2.2.311.  .
```

Figure A7.2 Means clustering AVP 5,000snaps(1 snap/ns)|:1-6@CA,S*|5µs|cl05|50 win

```
#                       Condensed Map
#
#    5000 points divided into 50 windows, each window contains 100 points.
#
#
#Clustering: divide 10000 points into 5 clusters
#Cluster    0: has   200 points, occurrence 0.020
#Cluster    1: has  7061 points, occurrence 0.706
#Cluster    2: has  1121 points, occurrence 0.112
#Cluster    3: has  1282 points, occurrence 0.128
#Cluster    4: has   336 points, occurrence 0.034
#Cluster    0   ......11....11.
#Cluster    1   .             3XX9XXXXXXXXXXXXXXX9XXXXXXXXXXXXXXX
#Cluster    2     3594.39782...
#Cluster    3   7942.383 ..4675  .               .
#Cluster    4   1.11...1.1.221.  .
```

Figure A7.3 Means clustering AVP 10,000snaps(2 snap/ns)|:1-6@CA,S*|5µs|cl05|50 win

```
#                       Condensed Map
#
#    5000 points divided into 50 windows, each window contains 100 points.
#
#
#Clustering: divide 5000 points into 5 clusters
#Cluster    0: has  1336 points, occurrence 0.267
#Cluster    1: has   104 points, occurrence 0.021
#Cluster    2: has    70 points, occurrence 0.014
#Cluster    3: has    20 points, occurrence 0.004
#Cluster    4: has  3470 points, occurrence 0.694
#Cluster    0   9999X8989897896
#Cluster    1   .... ..1.1 21..
#Cluster    2   .            ..  3. ..     .   .. 1
#Cluster    3     .. . ....
#Cluster    4                3XX69X99XXXXX9XXX99X8XXXXXXXXXXXXXXX
```

Figure A7.4 Average-Linkage clustering AVP 5,000snaps(1 snap/ns)|:1-6@CA,S*|5µs|cl05|50 win

AmberDASH

Initial versions of the Perl script *amberDASH* by David Whitley in 2013 were programmed based on specifications by EH to simplify the workflow of DASH analyses on AMBER trajectories. It can be used in combination with MDASH,[327] the current version of DASH.

Documentation.

```
AMBERDASH2(1)                User Contributed Perl Documentation                AMBERDASH2(1)

NAME
   amberDASH - Dash interface for AMBER trajectories

SYNOPSIS
   amberdash [options] seed [-- dash_options]
      Run dash on trajectories generated by the AMBER molecular dynamics package (http://ambermd.org/).

OPTIONS
   -help
      Print help message and exit.
   -version
      Print version number and exit.
   -debug
      Print progress messages on stderr.
   -keep
      Keep all the intermediate files that are normally deleted when the program exits.
   -keep-dash-input
      Keep the dash input file seed.dash.in which is normally deleted when the program exits.
   -keep-ptraj-input
      Keep the ptraj input file seed.ptraj.in which is normally deleted when the program exits.
   -no-dash
      Generate (and keep) the dash input file seed.dash.in but do not run dash.
   -progress
      Print output from the ptraj command on stdout. Useful for monitoring progress when reading large trajectories.
   -snap
      Write PDB files containing snapshots representing the dash states.
   -backbone r1:r2
      Analyse the sequence of backbone torsion angles from residue r1 to residue r2.

DASH OPTIONS
   The dash_options are described in the dash documentation. The dash flags -N (number of frames) and -T (number of torsions) are
   not required; they are supplied automatically by amberDASH.

INPUT FILES
   Several input files are required, specified by the seed prefix, in order to identify the topology, the torsion angles and the trajectory.
   The topology is always specified by an AMBER topology file seed.top. The torsion angles may be specified either by the -backbone
   option or by a file seed.tor. The -backbone option takes precedence. The trajectory may be specified either as a single AMBER
   trajectory file seed.trj, or by a sequence of trajin commands in a text file seed.trajin. If seed.trajin exists, it is used and seed.trj is
   ignored; otherwise seed.trj is used. The seed.trajin approach is necessary if the trajectory spans several files or a subset of the
   trajectory is required via start, stop and offset arguments to trajin.If snapshots are required (-snap) and the trajectory is specified
   in seed.trajin, the trajin commands in seed.trajin must include start, stop and offset fields. Otherwise the script is unable to locate
   the representative dash states in the trajectory and the snapshots are omitted.

   seed.top
      An AMBER topology file corresponding to the trajectory to be analysed.
   seed.trajin
      A text file containing trajin commands to extract the trajectory to be analysed. Any lines not containing trajin commands are
      ignored.
   seed.trj
      An AMBER trajectory file.
   seed.tor
      A text file defining the torsion angles to be analysed. Each torsion angle is specified by a whitespace-separated line with five
      fields:
            name mask1 mask2 mask3 mask4
      Here name is an identifier and mask1, ..., mask4 are AMBER atom masks defining the torsion angle. Lines starting with '#' are
      treated as comments and ignored. This file must be prepared manually by the user.

OUTPUT FILES
```

seed.dash.out
The dash output file.
seed.ptraj.out
The output from the ptraj command to extract the torsions.
If -snap is specified, the following files are written for each dash state:
seed.stateN.frame
The PDB file containing the representative frame frame for dash state N.
seed.ptraj.stateN.out
The output from the ptraj command to generate the PDB file for dash state N.
If -keep is specified, the following intermediate files are retained:
seed.ptraj.in
The input file for the ptraj command to extract the torsions.
seed.name
The torsion angles for each torsion name.
seed.dash.in
The dash input file obtained by joining the torsion angle files.
If -keep-ptraj and -snap are specified, the following files are retained:
seed.ptraj.stateN.in
The input file for the ptraj command to generate the PDB file for dash state N.

INSTALLATION
The programs dash and either ptraj or cpptraj are required. If they are not on the PATH their full pathnames must be specified at the top of the *amberDASH* script. cpptraj reads large trajectories faster than ptraj, while ptraj calculates statistics for the torsion angles.

NOTE
All the output files will be clobbered by the next run of the script for the same seed.

REFERENCE
D. W. Salt, B. D. Hudson, L. Banting, M. J. Ellis and M. G. Ford
DASH: A novel analysis method for molecular-dynamics simulation data.
Analysis of ligands of PPAR-gamma, J. Med. Chem., 48, 3214-3220, 2005.

ACKNOWLEDGEMENT
This work was funded by the Peptide Research Network of Excellence (PeReNE) (http://www.perene-project.eu/) as part of the INTERREG IV A France (Channel) England Programme (http://www.interreg4a-manche.eu/).

AUTHOR
David Whitley, University of Portsmouth <david.whitley@port.ac.uk>.

perl v5.18.1 2015-12-08 AMBERDASH2(1)

Command line (GPU): amberdash2.pl [options] seed [-- dash_options]
Options: -help (print help message and exit); -version (print version number and exit); -debug (print progress messages on stderr); -keep (keep all intermediate files); -keep-dash-input (keep the dash input file); -keep-ptraj-input (keep the ptraj input file); -no-dash (keep the dash input file but do not run dash); -progress (print output from the ptraj command on stdout); -snap (write PDB files containing snapshots representing the dash states); -backbone r1:r2 (analyse the sequence of backbone torsion angles from residue r1 to residue r2). For full documentation, use the command "perldoc amberdash2.pl".

Example. DASH ring-state analysis

Command-line with torsion trajectory output labelled with ring states for further analysis with SARcaddle (PCA)
amberdash2.pl -keep-dash-input -progress -snap seed -- -L UIIonvf_5us_T10_dash-label.in

Input: AMBER trajectory (coordinates in netcdf format): seed.trajin

```
trajin /delta3/haensele/Urotensin II/MinMD_invfolded/md1.trj  1  25000 50
trajin /delta3/haensele/Urotensin II/MinMD_invfolded/md2.trj  1  75000 50
…
```

Input with definition of torsions to be analysed: seed.tor

```
# Ring
psi5 :5@N :5@CA :5@C :6@N
phi6 :5@C :6@N :6@CA :6@C
psi6 :6@N :6@CA :6@C :7@N
phi7 :6@C :7@N :7@CA :7@C
psi7 :7@N :7@CA :7@C :8@N
phi8 :7@C :8@N :8@CA :8@C
psi8 :8@N :8@CA :8@C :9@N
phi9 :8@C :9@N :9@CA :9@C
psi9 :9@N :9@CA :9@C :10@N
phi10 :9@C :10@N :10@CA :10@C
```

Input topology: seed.top (AMBER topology file used for the MD simulation)

Output file seed.dash.in (torsion trajectory defined by seed.tor) = Input for DASH:

```
-3.0233 -67.6670 -1.3856 -66.3812 -17.5568 -132.0548 20.6964 55.5236 25.8836 54.8955
-11.6154 -62.8468 -14.5486 -59.1280 -33.9189 -116.4681 32.1009 53.9973 29.7835 44.3236
9.3207 -74.3252 -10.2949 -56.7112 -46.4321 -97.9060 18.3238 56.7269 8.3868 57.5924
...
```

Output file seed.dash.out (result of DASH analysis)

```
DASH, version 2.11b5
Thu Mar 17 00:55:26 2016

[TRAJECTORY]
file   : /delta3/haensele/Urotensin II/MinMD_invfolded/analysis_dash/seed.dash.in
angles : 10
frames : 10001

[OPTIONS]
winsize  : 11
binsize  : 4
runlen   : 3
fmax     : 2.4
smin     : 48
boutlen  : 20
smooth   : 40
roughen  : 20

[ANGLE_1]
maxima      : 2, 142
states      : 2 = [-180, -108), 1 = [-108, 72), 2 = [72, 180)
transitions : 6
reassigned  : 27 (0.27%)
...

[ANGLE_10]
maxima      : -78, 54
states      : 1 = [-180, -12), 2 = [-12, 168), 1 = [168, 180)
transitions : 2
reassigned  : 0 (0.00%)

[SUMMARY]
combined states : 9
final states    : 7
transitions     : 11
reassigned      : 59 (0.59%)

[DASH_STATES]
[1]     1 1 1 1 1 1 2 1 1 1
[2]     1 1 1 2 1 1 1 1 1 1
[3]     1 1 1 2 1 1 1 1 1 2
[4]     1 1 1 2 1 1 2 1 1 1
[5]     2 1 1 1 1 1 2 1 1 1
[6]     2 1 1 1 2 1 2 1 1 1
[7]     2 1 1 2 2 1 2 1 1 1

[DASH_STATE_DISTRIBUTION]
State     Frames   %Frames   Rep.Frame      RMSD
[1]          417      4.17        4154      9.44
[2]         1224     12.24        2787      6.24
[3]         2529     25.29        2047      2.86
[4]           54      0.54        3804      7.21
[5]          133      1.33        3880     11.35
[6]         5575     55.74        4976      3.86
[7]           69      0.69        7167      4.99

[DASH_STATE_MEAN_ANGLES]
[1]       19.86  -80.89  -29.57 -121.01  -15.64 -119.19  164.54  -68.14  137.85 -118.79
[2]        3.07  -62.26  -27.11  -63.66  -21.35 -111.00   16.93   55.07   69.18  -87.63
[3]        1.34  -65.52  -24.70  -60.19  -24.83 -115.24   21.23   54.60   22.34   53.25
[4]       -1.14  -65.61   -3.27  -72.72  -17.43 -126.77  177.84  -72.28  116.59 -143.00
[5]      124.33 -137.89  -24.28 -118.99  -17.33 -112.83  166.75  -68.30  136.26 -111.99
[6]      143.77  -88.28   -8.57 -116.07  161.23  -73.54  150.85   56.00   35.06  -84.96
[7]      139.90  -94.13  -16.13  -80.20  155.79  -69.47  144.34   53.60   30.74  -88.51

[DASH_STATE_STANDARD_DEVIATIONS]
```

```
[1]        47.49   37.56   17.02   22.98   19.76   28.18   17.96   16.85   15.29   28.91
[2]        12.91   10.15   11.28   11.48   11.28   15.99   12.17    9.38   44.66   37.14
[3]         9.75    9.46   10.23    8.69   10.31   13.72    9.46    8.13   10.63    8.02
[4]        12.24    8.58   13.50   13.65   11.45   18.49   18.27   17.56   30.59   27.57
[5]        36.05   26.81   17.82   27.51   18.28   30.35    9.79   15.67   23.01   33.10
[6]        15.52   22.76   21.09   27.75   15.77   14.73   11.84    8.52   17.66   21.31
[7]        11.87   20.41   19.11   24.05   11.51   13.09   10.96    7.99   17.21   21.36

[DASH_STATE_TRAJECTORY]
State   Frames  Cumulative
[3]       2529        2529
[2]       1224        3753
[4]         54        3807
[1]        276        4083
[5]         44        4127
[1]        141        4268
[5]         89        4357
[6]       3836        8193
[7]         26        8219
[6]        815        9034
[7]         43        9077
[6]        924       10001

[DASH_STATE_TRANSITIONS]
11

[DASH_STATE_BOUTS]
[1]     276 141
[2]     1224
[3]     2529
[4]     54
[5]     44 89
[6]     3836 815 924
[7]     26 43

[CPU_TIME]
Input : 0.2s
Dash  : 0.1s
Total : 0.3s
```

Output labelled with ring states for further analysis with SARcaddle (PCA).

```
-3.0233 -67.6670 -1.3856 -66.3812 -17.5568 -132.0548 20.6964 55.5236 25.8836 54.8955 3
-11.6154 -62.8468 -14.5486 -59.1280 -33.9189 -116.4681 32.1009 53.9973 29.7835 44.3236 3
9.3207 -74.3252 -10.2949 -56.7112 -46.4321 -97.9060 18.3238 56.7269 8.3868 57.5924 3
…
```

Principal Component Analysis (PCA) in DASH

Recent versions of DASH include a routine to calculate principal components of the torsion angle trajectory.

Example. Analysis of the correlation of torsions for distinct ring-state types of UII.

Input: Overall torsion trajectories extracted from sections of MD simulations exclusively occupied by a distinct ring-state type.

Command line: dash2 –p –i <input> -o <output1> -L <arg>

(with –p calculate principal components; -i specify input file (torsion trajectory); -o specify output file; -L write input data with state labels to file <arg>)

Output (PCA part)

```
DASH, version 2.11b5
Mon Jul 13 12:43:21 2015

[TRAJECTORY]
file   : /delta3/haensele/Urotensin II/MinMD_ALL/analysis_pca/T18_trajectory_state-
type_sections/omega-I_UIInmr_1-10001.pca.in
```

angles : 18
frames : 10001

[OPTIONS]
winsize : 11
binsize : 4
runlen : 3
fmax : 2.4
smin : 48
boutlen : 20
smooth : 40
roughen : 20
…

[PCA_SUMMARY]

	PC1	PC2	PC3	PC4	PC5	PC6	…
Variance	2.3517	1.8056	1.5347	1.2906	1.1824	1.1521	…
Explained	0.1306	0.1003	0.0853	0.0717	0.0657	0.0640	…
Cumulative	0.1306	0.2310	0.3162	0.3879	0.4536	0.5176	…

[PCA_COEFFICIENTS]

PC1	PC2	PC3	PC4	PC5	PC6	PC7	…
-0.0202	0.0432	-0.0033	0.0225	-0.4801	0.5317	0.0803	…
0.0549	-0.0299	-0.0554	0.0213	0.4403	-0.4711	0.2742	…
0.0047	-0.0374	0.0027	0.0924	-0.2532	-0.2940	0.4537	…
0.0109	-0.0227	-0.0789	0.1385	-0.4225	-0.2298	0.5141	…
-0.0073	0.0961	-0.1842	0.1565	-0.4011	-0.3613	-0.2844	…
-0.1945	0.2383	0.0257	-0.2245	0.3108	0.2065	0.3194	…
-0.0114	0.1415	-0.0975	0.0485	0.0525	0.3118	0.3782	…
0.0229	0.0427	-0.0332	-0.0016	-0.0354	-0.0865	-0.3111	…
-0.3313	0.3904	-0.0299	-0.3284	-0.0636	-0.0771	0.0360	…
0.3171	-0.4295	0.3210	-0.2718	-0.0816	-0.0045	0.0530	…
0.0027	0.1420	-0.4721	0.5708	0.1741	0.1104	-0.0477	…
0.1829	0.0027	-0.1343	-0.0157	0.0418	0.1224	0.0859	…
0.2235	0.2234	0.4589	0.3963	0.0606	0.0168	0.0336	…
-0.0742	-0.4396	-0.5239	-0.2192	0.0195	0.0531	0.0349	…
0.4940	0.2683	-0.0156	-0.0738	0.0092	0.0197	0.0093	…
-0.5209	-0.1515	0.2783	0.2276	-0.0037	0.0076	-0.0043	…
0.0898	0.4576	-0.1185	-0.3495	-0.1452	-0.1615	-0.0703	…
-0.3677	0.0538	0.1545	-0.0428	-0.0598	-0.1242	-0.0664	…

[PCA_WEIGHTS]

PC1	PC2	PC3	PC4	PC5	PC6	PC7	…
0.0004	0.0019	0.0000	0.0005	0.2305	0.2828	0.0065	…
0.0030	0.0009	0.0031	0.0005	0.1939	0.2219	0.0752	…
0.0000	0.0014	0.0000	0.0085	0.0641	0.0864	0.2058	…
0.0001	0.0005	0.0062	0.0192	0.1785	0.0528	0.2643	…
0.0001	0.0092	0.0339	0.0245	0.1609	0.1305	0.0809	…
0.0378	0.0568	0.0007	0.0504	0.0966	0.0427	0.1020	…
0.0001	0.0200	0.0095	0.0024	0.0028	0.0972	0.1430	…
0.0005	0.0018	0.0011	0.0000	0.0013	0.0075	0.0968	…
0.1098	0.1524	0.0009	0.1078	0.0040	0.0059	0.0013	…
0.1006	0.1845	0.1030	0.0738	0.0067	0.0000	0.0028	…
0.0000	0.0202	0.2229	0.3258	0.0303	0.0122	0.0023	…
0.0335	0.0000	0.0180	0.0002	0.0017	0.0150	0.0074	…
0.0499	0.0499	0.2106	0.1570	0.0037	0.0003	0.0011	…
0.0055	0.1932	0.2744	0.0481	0.0004	0.0028	0.0012	…
0.2441	0.0720	0.0002	0.0055	0.0001	0.0004	0.0001	…
0.2714	0.0229	0.0774	0.0518	0.0000	0.0001	0.0000	…
0.0081	0.2094	0.0140	0.1222	0.0211	0.0261	0.0049	…
0.1352	0.0029	0.0239	0.0018	0.0036	0.0154	0.0044	…

[PCA_CENTROID]

ANGLE1	ANGLE2	ANGLE3	ANGLE4	ANGLE5	ANGLE6	ANGLE7	…
-95.2916	138.7798	-67.9813	126.0465	-94.2493	42.5560	-114.4368	…

Dashsim

Dashsim is a C++ script written by David Whitley to enable direct comparison of DASH states in the output files of different DASH analyses. A comparison of torsion angle sections from any source is possible (similar to RMSD alignments) if the number of torsions is identical and the DASH output format is used. A sample input is given below. The script calculates the circular similarity of the compared torsion angles. Further details on the algorithm and principle of the method is given in Appendix A3 (p S10). Command line: dashsim file1 file2

Example. Comparison of representative conformations of OT from long-scale MD simulations with torsion angles from molecular mechanics calculations published 1991 by Ward *et al.*[93]

Input file 1: torsion angles of representative conformations for OT (dash output format)

```
DASH, output version 2.10b1

[TRAJECTORY]
file   : various
angles : 10
frames : undefined

# Ring torsions of OT representatives
# OT T10= phi2 psi2 phi3 psi3 phi4 psi4 phi5 psi5 phi6 psi6

[DASH_STATE_MEAN_ANGLES]
[1ot]    -93.96  146.36  -62.67  -16.64  -90.07   -5.74 -121.42  -22.99 -130.13  150.68
[2ot]   -111.29  -30.18 -121.25   -6.9  -114.38  151.34  -68.46  107.67 -129.87  139.68
[3ot]   -104.82  128.47   48.21   19.89 -138.34  149.28  -77.04  121.74 -130.4   145.9
[4ot]    -95.3   -19.66 -106.67  157.29  -67.66  -22.47 -100.22   67.45 -108.45  141.44
[5ot]    -89.74  -19.86 -101.05  158.29  -71.09  148.76   51.7    47.5   -76.36  139.84
[6ot]    -88.3   161.35  -58.33  136.92   56.08    9.61  -110.41 -15.15 -118.39  144.64
[7ot]    -87.3   137.46   44.39   32.68   55.37   25.79  -84.82   17.51 -110.11  152.76
[99ot]   -76.0   161.57  7 1.98   -0.22 -128.51  148.58   51.64   34.96 -113.58  133.73
```

Input file 2: torsion angles from molecular mechanics calculations published by Ward *et al.* 1991 [93] (dash output format)

```
[TRAJECTORY]
file   : article Ward 1991
angles : 10
frames : undefined

# ring torsions of OT from Ward 1991 Ward DJ, Chen Y, Platt E, Robson B (1991) Development and testing of
protocols for computer-aided design of peptide drugs, using oxytocin. J Theor Biol 148 (2):193-227; Conf
1-8 from energy calculation (Molecular Mechanic calculation), Conf MD lowest energy conformaer from MD
simulation
# OT T10= phi2 psi2 phi3 psi3 phi4 psi4 phi5 psi5 phi6 psi6

[DASH_STATE_MEAN_ANGLES]
[Conf 1] -126.99  170.04  -58.95  -35.89  -58.67  -47.01 -165.59  121.96 -138.09  120.08
[Conf 2]  -54.91  -34.93  -44.82  -35.08 -105.06  -68.55 -168.21  131.66 -165.67  141.06
[Conf 3]  -55.06  -35.08  -45.21  -34.92 -104.99  -68.65 -168.27  132.02 -165.96  140.12
[Conf 4] -117.91 -177.95  -71.74   74.33 -156.29 -168.00  -67.47   88.85 -119.80  150.36
[Conf 5]  -70.10 -174.63  -75.46   81.97 -131.70 -131.35  -61.16  -32.02  -52.39  143.96
[Conf 6] -115.99  126.80   64.33 -114.65 -112.70    5.54  -95.67 -175.72 -161.32  158.51
[Conf 7] -107.63 -106.92  -76.27  -63.21 -105.71  -84.70 -104.21 -140.42  -97.01  -57.10
[Conf 8]  -70.32  134.04  -39.05  110.04  163.08 -176.74  -63.35   85.57  -94.85  151.89
[Conf MD] -62.51  -42.11  -84.06  147.02  -75.71  -51.86  -87.57   82.32 -105.80  137.47
```

Output circular similarity representative conformations of OT *vs.* MM Ward *et al.* [93]

```
reading file OT_T16T10.dashsim.in
angles = 10
reading DASH_STATE_MEAN_ANGLES
states = 8

reading file OT_Ward1991_T10.dashsim.in
angles = 10
reading DASH_STATE_MEAN_ANGLES
states = 10

Circular Similarity Matrix
A = representative conformations of OT from long-scale MD (EH)
B = conformations of OT by MM calculation (Ward 1991)
```

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
A2	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
A3	0.37	0.37	0.37	0.37	0.36	0.36	0.36	0.36	0.37	0.37
A4	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
A5	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44
A6	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.42	0.43	0.43
A7	0.52	0.52	0.52	0.52	0.51	0.51	0.51	0.51	0.52	0.52
A8	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42

Result: Maximum similarity 51-52 % of Ward B1-10i to OT A7 (*twisted saddle*)

Graphic Tools

Figures and posters were prepared with PyMOL 1.3 (educational product),[328] POV-Ray 3.6.2,[329] Gnuplot 5.0,[330] Chimera 1.10,[331] ChemBioDraw Ultra 13,[332] Microsoft 2013 Excel,[333] Adobe Photoshop CS5,[334] and Adobe Illustrator CS5.[335]

i The similarity between Ward´s conformations is 99-100%. Thus all Ward conformations belong to the same ring-state type.

A 8: Supporting Information Chapter 7

MD Simulations and Dynamics

Simulation Parameters

Table A8.1 Summary of MD simulation details of OT, dOT and CT[§]

Simulation	Time (µs)	Initial conformation	Resulting ring-state types	NAtoms (WAT)*
OT (9 residues, 136 atoms, charge +1)				
MD: AMBER ff99sb/ TIP4PEw/ trunc.oct. / 1Cl^-/ 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
OT_MD-I	15	*saddle*	*saddle, open$_{23pbr}$*	4029 (973)
OT_MD-II	15	*open*	*open, open$_{23pbr}$, cl.open, cl.open$_{45pbr}$*	4985 (1212)
OT_MD-III	10	*cl.open*	*cl.open, cl.open$_{45pbr}$*	6241 (1526)
OT_MD-IV	10	*tw.saddle*	*tw.saddle, tw.saddle$_{helix}$, cl.open*	5121 (1246)
dOT (9 residues, 133 atoms, charge 0)				
MD: AMBER ff99sb/ TIP4PEw/ trunc.oct. / 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
dOT_MD	3	*tw.saddle*	*tw.saddle, tw.saddle$_{helix}$*	3961 (957)
CT (9 residues, 138 atoms, charge 0)				
MD: AMBER ff99sb/ TIP4PEw/ trunc.oct. / 300K/ 1bar/ 8Å cutoff/ PME/ PBC/ Shake				
CT_MD-I	5	*saddle*	*saddle, open, open$_{23pbr}$, cl.open$_{45pbr}$*	4446 (1077)
CT_MD-II	5	*open*	*open, saddle*	6802 (1666)

[§]Force-field parameters for dOT and CT were modified within ff99SB using semi-empirical partial charges and force constants of similar atom combinations from *e.g.* alanine (-CH3), threonine (CH2-O-H) and methionine (-CH2-S-). *NAtoms: total number of atoms; WAT: number of water molecules.

RMSD and DASH State Trajectories

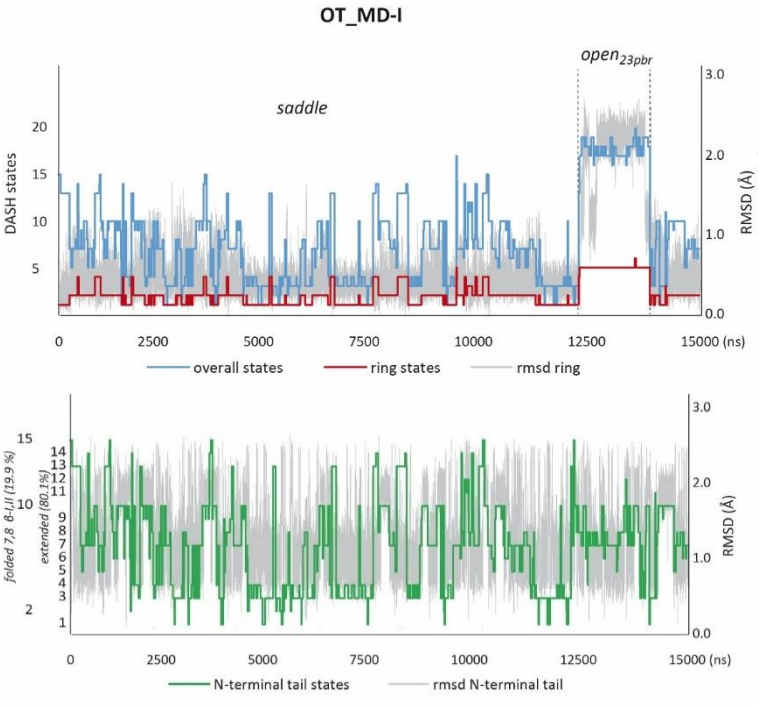


Figure A8.1 RMSD and DASH trajectories of simulation **OT_MD-I** (15µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T6 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *open*. Main ring-state types are labelled. The simulation shows two main ring-state types, *saddle* and *open$_{23pbr}$* (= 2,3 peptide bond rotamer of ring-state type *open*) with direct *open/folded* interconversion. The tail exhibits frequent interconversions predominated by *extended* (80.1 %$_{15µs}$) conformations; *folded* tail-state types (19.9 %$_{15µs}$, 7,8 β-II and β-I). Overall states **T16_7**, **T16_10** and **T16_17** were chosen as representatives for the ring-state types *saddle$_{ext}$*, *saddle$_{fold}$* and *open$_{23pbr(ext)}$* (subscript *ext* = *extended tail*, *fold* = *folded tail*).

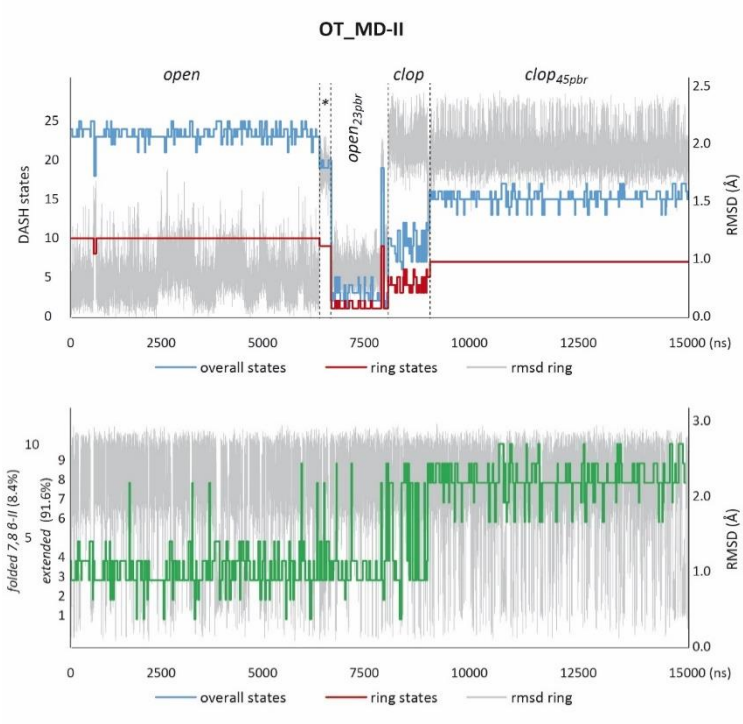


Figure A8.2 RMSD and DASH trajectories of simulation **OT_MD-II** (15µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T6 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *open*. Main ring-state types are labelled. The simulation only shows *open* conformations of ring-state types *open, open$_{23pbr}$, clinched open (clop)* and *cl.open$_{45pbr}$ (clop$_{45pbr}$)*. The section labelled with * is occupied by the transient state *intermediate saddle* (*cf.* AVP, Appendix A1). The tail exhibits frequent interconversions predominated by *extended* (91.6 %$_{15µs}$) conformations; *folded* tail-state types (8.4 %$_{15µs}$, 7,8 β-II). The RMSD trajectory shows a higher frequency of *extended* and *folded* tail conformations indicating a state lifetime < 10 ns for *folded* tail conformations. Overall states **T16_23** and **T16_25** of OT_MD-II were chosen as representatives for the ring-state types *open$_{ext}$* and *open$_{fold}$*; overall state T16_17 of OT_MD-II was chosen as representative for the ring-state type *clinched open$_{45pbr.fold}$* (subscript *ext* = *extended tail*, *fold* = *folded tail*).

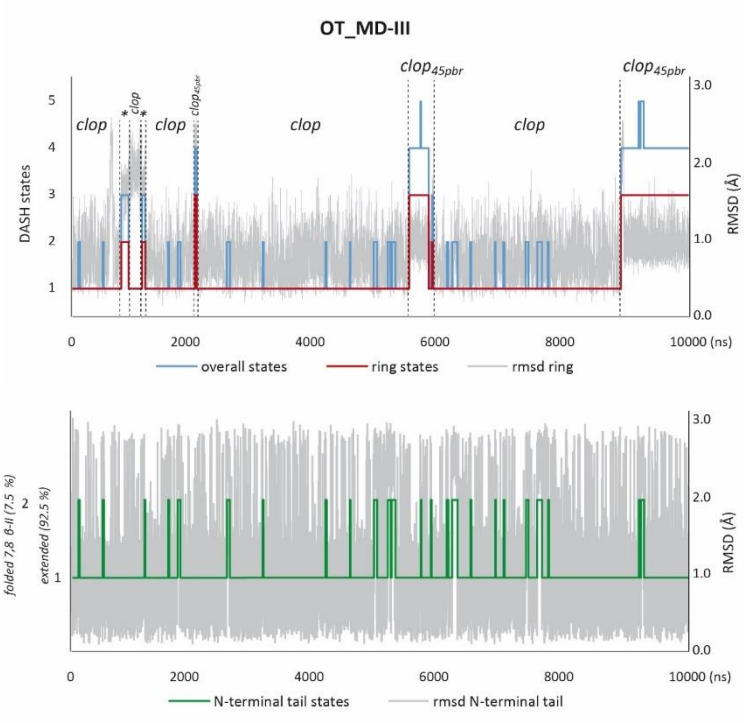


Figure A8.3 RMSD and DASH trajectories of simulation **OT_MD-III** (10 µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T6 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *clinched open*. Main ring-state types are labelled. The simulation only shows *open* conformations of ring-state type *clinched open (clop), clinched open$_{45pbr}$ (clop$_{45pbr}$)* and a transient *open* variant, *open$_{2334pbr}$* resembling the URP *lasso$_{45pbr}$* representative. The tail exhibits frequent interconversions predominated by *extended* (92.5 %$_{10µs}$) conformations; *folded* tail-state types (7.5 %$_{10µs}$, 7,8 β-II). The RMSD trajectory shows a higher frequency of *extended* and *folded* tail conformations indicating a state lifetime < 10 ns for *folded* tail conformations. Overall states **T16_1** and **T16_2** of OT_MD-III were chosen as representatives for the ring-state types *clinched open$_{ext}$* and *clinched open$_{fold}$*; overall state **T16_4** of OT_MD-III was chosen as representative for the ring-state type *clinched open$_{45pbr.ext}$* (subscript *ext* = *extended tail*, *fold* = *folded tail*).

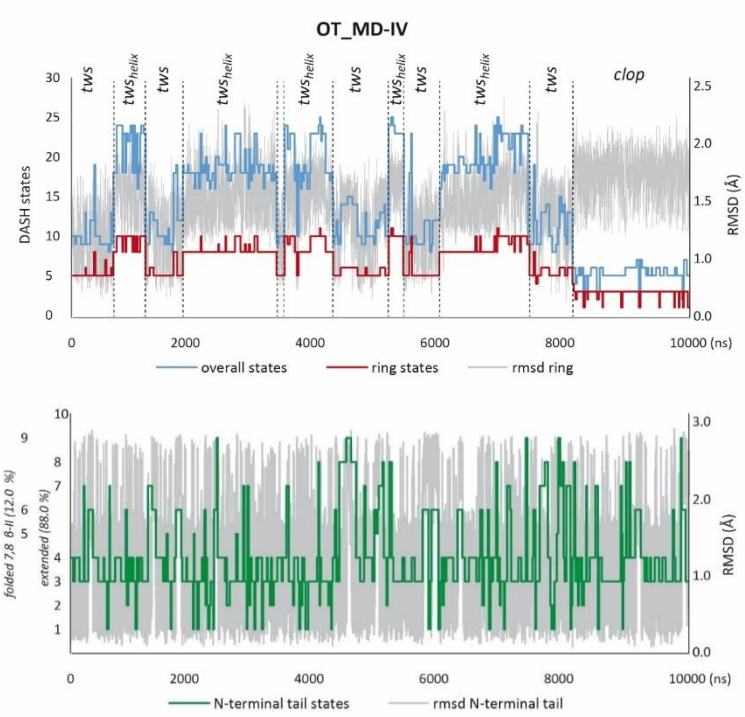


Figure A8.4 RMSD and DASH trajectories of simulation **OT_MD-IV** (10 µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T6 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *twisted saddle*. Main ring-state types are labelled. The first 8 µs, the simulation shows frequent interconversions within the *folded* ring-state types *twisted saddle (tws)* and the 3_{10}-helical variant *twisted saddle*$_{helix}$ *(tws*$_{helix}$*)*, then interconverts to the *open* ring-state type *clinched open (clop)*. The tail exhibits frequent interconversions between *extended* (88.0 %$_{10µs}$) and *folded* tail-state types (12.0 %$_{10µs}$, 7,8 β-II). The RMSD trajectory shows a higher frequency of *extended* and *folded* tail conformations indicating a state lifetime < 10 ns for *folded* tail conformations. Overall states **T16_9** and **T16_12** of OT_MD-IV were chosen as representatives for the ring-state types *twisted saddle*$_{ext}$ and *twisted saddle*$_{fold}$; overall state **T16_18** and **T16_20** of OT_MD-IV were chosen as representatives for the ring-state types *twisted saddle*$_{helix,ext}$ and *twisted saddle*$_{helix,fold}$ (subscript *ext* = *extended tail*, *fold = folded tail*).

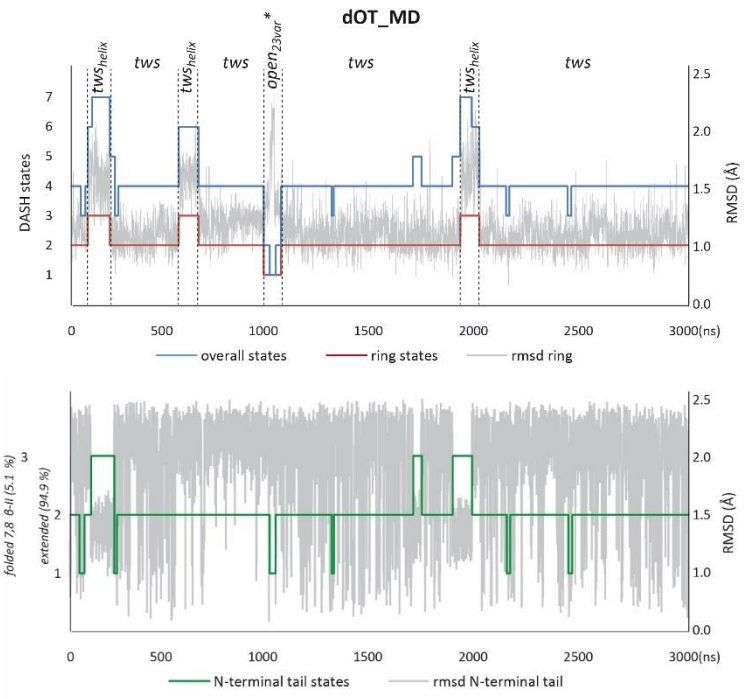


Figure A8.5 RMSD and DASH trajectories of simulation **dOT_MD** (3 µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T8 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation twisted saddle (PDB ID: 1XY1). Main ring-state types are labelled. dOT shows occasional interconversion to the 310-helical variant twisted saddlehelix (twshelix). After 1µs a scoop-like transient was identified (denoted as twisted saddlehelix-var) with an *open/folded* hybrid structure that might be an intermediate for *open/folded* inter-conversions. The tail exhibits frequent interconversions between extended (94.9 %$_{3µs}$) and folded tail-state types (5.1 %$_{3µs}$ 7,8 β-II). Within the sections of the *saddle* ring-state types, inter-conversions seem to be independent of the ring conformation (highly frequent interconversions), whereas for the *twisted saddle* sections a folded tail-conformations seems to be favoured that may suggest correlation of ring and tail conformation. Overall states **T16_4** and **T16_7** of dOT_MD were chosen as representatives for the ring-state types twisted saddle and twisted saddlehelix.

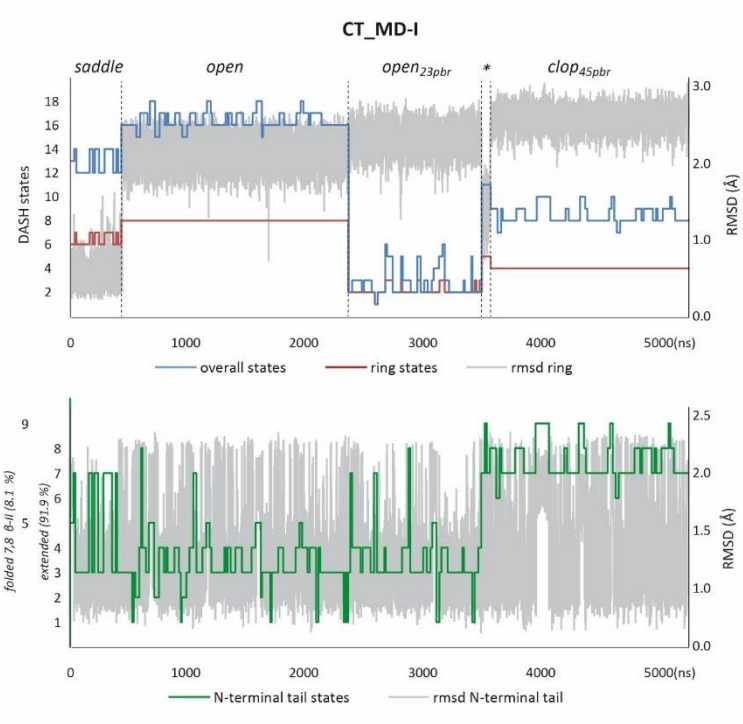


Figure A8.6 RMSD and DASH trajectories of simulation **CT_MD-I** (5.25 µs).

Trajectories of DASH states (T18 overall blue, T10 ring red, T8 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *saddle*. Main ring-state types are labelled. After 428 ns the initially *folded* conformation *saddle* interconverts step-wise to *open (unfolded)* ring-state types *open, open*$_{23pbr}$ and *clinched open*$_{45pbr}$ *(clop*$_{45pbr}$*)*. The interconversion from *open*$_{23pbr}$ to *clop*$_{45pbr}$ passes the transient state *intermediate saddle* (*). The tail exhibits frequent interconversions between *extended* (91.9 %$_{5.25µs}$) and *folded* tail-state types (8.1 %$_{5.25µs}$ 7,8 β-II). The RMSD trajectory shows a higher frequency of *extended* and *folded* tail conformations indicating a state lifetime < 10 ns for *folded* tail conformations. Overall states **T16_2** and **T16_8** of CT_MD-I were chosen as representatives for the ring-state types *open*$_{23pbr}$ and *clinched open*$_{45pbr}$ that resemble the ring-state types *lasso* and *omega-II* of UII/URP.

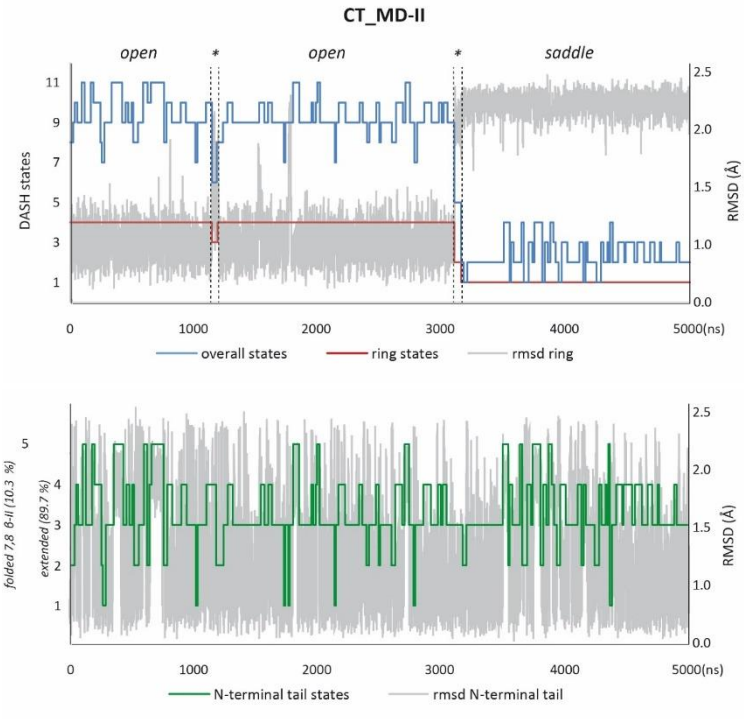


Figure A8.7 RMSD and DASH trajectories of simulation **CT_MD-II** (5 µs).

Trajectories of *DASH* states (T18 overall blue, T10 ring red, T8 C-terminal tail green) and RMSD (Cα 1-6, ring; Cα 6-9, tail). Initial conformation *open*. Main ring-state types are labelled. After 3.2 µs the initially *unfolded* conformation *open* interconverts to the ring-state type *saddle (folded).* The transient state (*) at 3.1 µs is a *saddle*-variant with high similarity to the *folded-II* type of UII (96 %). The transient state (*) at 1.2 µs is a variant of the *intermediate saddle* ring-state type and a conformational hybrid of *open* and *folded* ring-state types. The tail exhibits frequent interconversions between *extended* (89.7 %$_{5µs}$) and *folded* tail-state types (10.3 %$_{5.25µs}$ 7,8 β-II). The RMSD trajectory shows a higher frequency of extended and folded tail conformations indicating a state lifetime < 10 ns for folded tail conformations. Overall states **T16_9** and **T16_2** of CT_MD-II were chosen as representatives for the ring-state types *open* and *saddle* of CT that resemble the ring-state types *lasso*$_{56pbr}$ and *folded-I* of UII/URP.

Conformations and Circular Similarity

DASH States and Representative Conformations

Table A8.2 Absolute populations and circular similarities of DASH *ring* states, corresponding *overall* states and representatives of the MD simulations of OT, dOT and CT[§]

MD		Ring state [a]		Overall state [b]			Representative [c]		Ring-state type
		T10	Pop (%)	T16	Pop (%)	circsim T16 *vs.* T10	ID [b]	circsim T10 *vs.* Rep	
OT	**OT_MD-I**	**2**	45.5	**7**	17.7	1.00	**I_T16_7**	1.00	*saddle$_{ext}$*
				10	15.3	0.99	**I_T16_10**	0.99	*saddle$_{fold}$*
		1	33.7	**3**	20.8	1.00	**I_T16_7**	0.92	*saddle*
		4	8.8	**13**	6.3	1.00	**I_T16_7**	0.88	*saddle*
		3	0.9	**11**	0.4	1.00	**I_T16_7**	0.67	*saddle*
		Total	*88.9*					*ext:fold [d] =*	*77.9 : 22.1*
		5	10.9	**17**	4.4	0.98	**I_T16_17**	0.98	*open$_{23pbr}$*
		6	0.2	**20**	0.2	1.00	**I_T16_17**	0.59	*open$_{2356pbr}$*
		Total	*11.1*					*ext:fold [d] =*	*100:0*
	OT_MD-II	**7**	41.9	**15**	25.7	1.00	**III_T16_4**	0.99	*clop$_{45pbr}$*
		10	39.8	**23**	19.4	1.00	**II_T16_23**	1.00	*open$_{ext}$*
				25	4.8	0.99	**II_T16_20**	0.99	*open$_{fold}$*
		8	0.4	**18**	0.3	1.00	**II_T16_23**	0.70	*open*
		Total	*40.2*					*ext:fold [d] =*	*88.0 : 12.0*
		2	1.9	**4**	1.0	0.99	**I_T16_17**	0.96	*open$_{23pbr}$*
		4	2.3	**9**	1.5	0.99	**III_T16_1**	0.97	*clop*
		5	2.0	**11**	1.6	1.00	**III_T16_1**	0.93	*clop*
		3	1.9	**7**	1.4	0.99	**III_T16_1**	0.98	*clop*
		6	0.6	**12**	0.6	1.00	**III_T16_1**	0.94	*clop*
		Total	*6.8*						
		9	2.2	**19**	1.6	1.00	-	-	*int.saddle*
	OT_MD-III	**1**	83.8	**1**	77.1	1.00	**III_T16_1**	1.00	*clop$_{ext}$*
				2	6.7	0.98	**III_T16_2**	0.98	*clop$_{fold}$*
								ext:fold [d] =	*92.0 : 8.0*
		3	14.4	**4**	13.7	1.00	**III_T16_4**	1.00	*clop$_{45pbr.ext}$*
				5	0.7	0.98	**II_T16_17**	0.98	*clop$_{45pbr.fold}$*
								ext:fold [d] =	*95.3 : 4.7*
		2	1.8	**3**	1.8	1.00	**III_T16_4**	0.62	*open$_{2334pbr}$*
	OT_MD-IV	**8**	28.4	**18**	12.5	0.99	**IV_T16_18**	0.99	*tws$_{helix.ext}$*
				20	2.0	1.00	**IV_T16_20**	1.00	*tws$_{helix.fold}$*
		10	15.7	**23**	10.5	1.00	**IV_T16_18**	0.84	*tws$_{helix}$*
		9	1.7	**21**	1.3	0.97	**IV_T16_18**	0.84	*tws$_{helix}$*
		11	0.7	**25**	0.8	1.00	**IV_T16_18**	0.82	*tws$_{helix}$*
		Total	*46.5*					*ext:fold [d] =*	*95.8 : 4.2*
		5	26.5	**9**	11.7	0.99	**IV_T16_9**	0.99	*tws$_{ext}$*
				12	5.2	0.98	**IV_T16_12**	0.98	*tws$_{fold}$*
		6	7.7	**13**	3.8	1.00	**IV_T16_9**	0.90	*tws*
		7	0.2	**(9,13)**	-	-	**IV_T16_9**	0.84	*tws*
		Total	*34.4*					*ext:fold [d] =*	*76.5 : 23.5*
		3	16.4	**5**	9.6	0.99	**III_T16_1**	0.98	*clop*
		1	1.5	**3**	0.6	0.96	**III_T16_1**	0.92	*clop*
		2	0.8	**4**	0.2	0.80	**III_T16_1**	0.85	*clop*
		Total	*18.7*					*ext:fold [d] =*	*91.4:8.6*
		4	0.5	**(6,7)**	-	-	**III_T16_1**	0.64	*open$_{2334pbr}$*
dOT	**dOT_MD**	**2**	87.3	**4**	81.0	1.00	**T16_4**	1.00	*tws*
		Total	*87.3*					*ext:fold [d] =*	*95.9 : 4.1*
		3	9.9	**7**	4.7	0.96	**T16_7**	0.96	*tws$_{helix_fold}$*
		Total	*9.9*					*ext:fold [d] =*	*52.0:48.0*
		1	2.9	**2**	1.9	0.97	-	-	*open$_{var}$*
CT	**CT_MD-I**	**8**	36.7	**16**	20.4	1.00	**II_T16_9**	1.00	*open*
		Total	*36.7*					*ext:fold [d] =*	*92.7:7.3*
		4	32.1	**8**	18.8	1.00	**I_T16_8**	1.00	*clop$_{45pbr}$*
		Total	*32.1*					*ext:fold [d] =*	*88.1:11.9*
		2	17.9	**2**	10.8	1.00	**I_T16_2**	1.00	*open$_{23pbr}$*
		3	3.2	**5**	2.44	0.99	**I_T16_2**	0.95	*open$_{23pbr}$*
		Total	*21.1*					*ext:fold [d] =*	*95.7:4.3*
		5	1.44	**11**	1.44	1.00	-	-	*int.saddle*
		1	0.46	**1**	0.46	1.00	-	-	*open$_{2356pbr}$*
	CT_MD-II	**1**	59.7	**1**	20.5	1.00	**II_T16_1**	1.00	*saddle*
		Total	*59.7*					*ext:fold [d] =*	*90.5:9.5*
		4	37.1	**9**	34.3	1.00	**II_T16_9**	1.00	*open*
		Total	*37.1*					*ext:fold [d] =*	*87.7:12.3*
		2	1.8	**5**	1.1	1.00	-	-	*saddle$_{var}$ (= folded-II)*
		3	1.5	**6**	0.8	1.00	-	-	*int.saddle$_{var}$*

[§] Representative states for each peptide are highlighted in green. For OT, representatives subtypes have been defined with *extended* and *folded* tail for subsequent NMR modelling. [a] Ring states are ordered by ring-state types and descending populations. [b] Overall states with maximum similarity of ring torsions to *ring* states. [c] Representative states are highlighted (green = *extended* tail; light green = *folded* tail). [d] Ratio of *extended* and *folded* tail conformations are calculated from the relative populations of overall states of the same ring-state type. Abbreviations: T10 = *ring* states defined *via ΦΨ* 2-6 (10 torsions); T16 = overall states defined *via ΦΨ* 2-9 (16 torsions); Pop = absolute state population relative to simulation time; circsim = circular similarity; *clop* = *clinched open*; *tws* = *twisted saddle*.

Mean Angles of Representative Conformations

Table A8.3 Mean angles of DASH states for representative and transient conformations of all peptides investigated §

	Φ_{i+1}	Ψ_{i+1}	Φ_{i+2}	Ψ_{i+2}	Φ_{i+3}	Ψ_{i+3}	Φ_{i+4}	Ψ_{i+4}	Φ_{i+5}	Representative
AVP										
saddle, tws clop, open	see Table S2 of Supporting Information Paper 2 (Appendix A2)									
*clop$_{45pbr}$**	-84.3	-10.9	-120.7	160.1	-74.1	148.6	52.8	42.3	-79.6	MD_23us_T16_18
	22.2	*17.9*	*27.9*	*11.5*	*14.1*	*11.1*	*8.6*	*16.7*	*21.0*	*stddev*
*int.saddle**	-68.6	163.2	-72.9	-0.6	-125.2	146.9	26.0	63.5	-105.4	MD_11us_T10_3
	13.4	*18.3*	*10.4*	*21.2*	*24.0*	*15.8*	*51.5*	*53.8*	*36.4*	*stddev*
OT										
saddle	-94.0	146.4	-62.7	-16.6	-90.1	-5.7	-121.4	-23.0	-130.1	MD-I_T16_7
	25.5	*17.2*	*10.1*	*16.1*	*18.5*	*17.2*	*20.5*	*23.5*	*23.6*	*stddev*
tws	-88.3	161.4	-58.3	136.9	56.1	9.6	-110.4	-15.2	-118.4	MD-IV_T16_9
	30.5	*16.8*	*20.7*	*20.4*	*8.3*	*25.8*	*30.4*	*40.3*	*28.8*	*stddev*
tws$_{helix}$	-87.3	137.5	44.4	32.7	55.4	25.8	-84.8	17.5	-110.1	MD-IV_T16_18
	42.6	*14.0*	*7.8*	*11.5*	*8.5*	*15.2*	*21.2*	*43.5*	*36.7*	*stddev*
clop	-95.3	-19.7	-106.7	157.3	-67.7	-22.5	-100.2	67.5	-108.5	MD-III_T16_1
	29.7	*24.4*	*28.8*	*15.0*	*24.5*	*23.7*	*29.8*	*52.9*	*40.1*	*stddev*
clop$_{45pbr}$	-89.7	-19.9	-101.1	158.3	-71.1	148.8	51.7	47.5	-76.4	MD-III_T16_4
	26.2	*17.6*	*28.3*	*11.4*	*25.8*	*20.2*	*8.9*	*17.4*	*23.2*	*stddev*
open	-104.8	128.5	48.2	19.9	-138.3	149.3	-77.0	121.7	-130.4	MD-II_T16_23
	37.6	*12.8*	*8.1*	*19.0*	*17.7*	*14.5*	*14.8*	*29.4*	*30.7*	*stddev*
open$_{23pbr}$	-111.3	-30.2	-121.3	-6.9	-114.4	151.3	-68.5	107.7	-129.9	MD-I_T16_17
	33.8	*53.7*	*28.6*	*20.5*	*31.5*	*17.0*	*44.4*	*40.0*	*27.6*	*stddev*
*open$_{2334pbr}$**	-114.2	-22.9	-132.7	149.3	60.1	171.4	-79.5	129.8	-128.5	MD-III_T16_3
	27.6	*16.3*	*18.8*	*22.3*	*8.3*	*31.3*	*21.7*	*22.6*	*27.1*	*stddev*
*int.saddle**	-76.0	161.6	-72.0	-0.2	-128.5	148.6	51.6	35.0	-113.6	MD-II_T10_9
	21.5	*18.8*	*18.7*	*18.7*	*21.6*	*17.8*	*19.5*	*29.0*	*33.9*	*stddev*
dOT										
tws	-98.6	158.9	-58.8	138.2	56.4	8.0	-123.0	-22.5	-91.3	MD_T16_4
	28.5	*13.7*	*11.3*	*10.6*	*8.0*	*21.0*	*26.1*	*32.5*	*32.8*	*stddev*
tws$_{helix}$	-67.7	132.7	44.0	32.1	54.7	22.9	-87.2	24.1	-105.9	MD_T16_7
	27.0	*12.1*	*7.5*	*10.7*	*8.4*	*13.8*	*21.9*	*40.4*	*37.9*	*stddev*
*open$_{23var}$**	-131.4	-19.9	-138.7	60.9	55.2	27.6	-115.0	25.5	-84.6	MD_T16_2
	32.1	*13.0*	*14.7*	*28.2*	*8.2*	*18.2*	*27.6*	*18.7*	*24.8*	*stddev*
CT										
saddle	-117.4	147.7	-61.6	-19.9	-86.1	-8.7	-120.2	-29.3	-120.2	MD-II_T16_01
	30.4	*13.8*	*10.2*	*16.1*	*17.6*	*16.8*	*20.6*	*18.0*	*29.4*	*stddev*
clop$_{45pbr}$	-87.4	-7.8	-110.4	157.4	-73.7	150.3	51.4	47.0	-74.3	MD-I_T16_8
	19.45	*19.44*	*27.54*	*10.01*	*12.12*	*10.74*	*8.42*	*15.42*	*19.51*	*stddev*
open	-90.3	124.8	48.1	24.5	-135.4	157.7	-74.6	118.1	-127.3	MD-II_T16_9
	37.0	*15.8*	*8.0*	*17.4*	*18.9*	*15.8*	*20.0*	*35.9*	*34.8*	*stddev*
open$_{23pbr}$	-124.3	-29.2	-129.7	-9.8	-119.2	159.1	-72.8	122.8	-125.5	MD-I_T16_2
	43.09	*17.14*	*18.98*	*19.54*	*28.66*	*11.31*	*18.24*	*32.05*	*31.47*	*stddev*
*int.saddle**	-82.0	164.3	-73.3	-0.4	-133.3	147.6	50.4	34.4	-90.0	MD-I_T16_11
	17.83	*37.6*	*16.34*	*35.72*	*21.61*	*14.24*	*9.38*	*23.04*	*32.64*	*stddev*
*int.saddle$_{var}$**	-104.7	129.5	48.9	19.6	-122.7	27.5	53.0	42.8	-95.6	MD-II_T16_6
	38.7	*13.1*	*6.7*	*22.7*	*20.5*	*48.9*	*9.3*	*34.4*	*37.3*	*stddev*
*saddle$_{var}$**	-60.7	129.1	49.1	10.3	-120.9	-33.9	-142.5	-29.6	-115.8	MD-II_T16_5
	21.2	*15.4*	*8.8*	*31.8*	*20.9*	*21.6*	*23.2*	*19.2*	*33.3*	*stddev*
UI										
folded-I, -IVb2 folded-II,-II inv-folded Ω-I,-II lasso, scoop circle	see Table S3 of Supporting Information Paper 3 (Appendix A3)									
URP										
hybrid, sheet Ω-I,-II lasso$_{45pbr}$	see Table S3 of Supporting Information Paper 3 (Appendix A3)									

§AVP (i=1), OT (i=1), dOT (i=1), CT (i=1), UII (i=5) and URP (i=2)

Circular Similarity of Representative and Transient* Conformations

Table A8.4 Circular similarity [§] of ring torsions [a] of representative and transient conformations of AVP, OT, dOT, CT, UII, and URP [b]

Circular Similarity of Ring Torsions		AVP						OT								
		saddle	*tws*	*clop*	*clop$_{45pbr}$**	*open*	*int.saddle**	*saddle*	*tws*	*tws$_{helix}$*	*clop*	*clop$_{45pbr}$*	*open*	*open$_{23pbr}$*	*open$_{2334pbr}$**	*int.saddle**
AVP	*saddle*	1.00	0.62	0.51	0.36	0.52	0.57	0.97	0.60	0.64	0.51	0.36	0.53	0.47	0.30	0.56
	tws	0.62	1.00	0.56	0.41	0.43	0.44	0.62	0.97	0.74	0.55	0.42	0.45	0.37	0.47	0.44
	clop	0.51	0.56	1.00	0.56	0.39	0.39	0.52	0.56	0.48	0.97	0.56	0.42	0.54	0.58	0.37
	*clop$_{45pbr}$**	0.36	0.41	0.56	1.00	0.40	0.54	0.36	0.41	0.37	0.57	0.96	0.42	0.58	0.60	0.54
	open	0.52	0.43	0.39	0.40	1.00	0.66	0.53	0.41	0.56	0.38	0.41	0.96	0.55	0.41	0.62
	*int.saddle**	0.57	0.44	0.39	0.54	0.66	1.00	0.57	0.43	0.49	0.40	0.54	0.67	0.61	0.40	0.93
OT	*saddle*	0.97	0.62	0.52	0.36	0.53	0.57	1.00	0.61	0.63	0.51	0.36	0.54	0.48	0.31	0.55
	tws	0.60	0.97	0.56	0.41	0.41	0.43	0.61	1.00	0.71	0.55	0.41	0.44	0.36	0.46	0.42
	tws$_{helix}$	0.64	0.74	0.48	0.37	0.56	0.49	0.63	0.71	1.00	0.47	0.38	0.57	0.39	0.40	0.48
	clop	0.51	0.55	0.97	0.57	0.38	0.40	0.51	0.55	0.47	1.00	0.57	0.41	0.54	0.59	0.37
	clop$_{45pbr}$	0.36	0.42	0.56	0.96	0.41	0.54	0.36	0.41	0.38	0.57	1.00	0.44	0.58	0.60	0.54
	open	0.53	0.45	0.42	0.42	0.96	0.67	0.54	0.44	0.57	0.41	0.44	1.00	0.57	0.42	0.63
	open$_{23pbr}$	0.47	0.37	0.54	0.58	0.55	0.61	0.48	0.36	0.39	0.54	0.58	0.57	1.00	0.56	0.58
	*open$_{2334pbr}$**	0.30	0.47	0.58	0.60	0.41	0.40	0.31	0.46	0.40	0.59	0.60	0.42	0.56	1.00	0.37
	*int.saddle**	0.56	0.44	0.37	0.54	0.62	0.93	0.55	0.42	0.48	0.37	0.54	0.63	0.58	0.37	1.00
dOT	*tws*	0.60	0.93	0.55	0.40	0.40	0.41	0.60	0.94	0.70	0.55	0.40	0.42	0.35	0.44	0.40
	tws$_{helix}$	0.63	0.73	0.48	0.37	0.55	0.49	0.63	0.71	0.96	0.47	0.38	0.56	0.39	0.40	0.47
	*open$_{23var}$**	0.52	0.59	0.67	0.50	0.38	0.38	0.52	0.59	0.55	0.67	0.51	0.39	0.55	0.62	0.37
CT	*saddle*	0.93	0.61	0.51	0.35	0.52	0.55	0.95	0.60	0.63	0.50	0.35	0.52	0.47	0.30	0.54
	clop$_{45pbr}$	0.37	0.42	0.56	0.98	0.42	0.55	0.36	0.42	0.38	0.57	0.97	0.44	0.58	0.60	0.55
	open	0.52	0.45	0.42	0.44	0.94	0.68	0.53	0.43	0.56	0.41	0.45	0.97	0.57	0.42	0.63
	open$_{23pbr}$	0.45	0.35	0.52	0.56	0.55	0.59	0.45	0.33	0.36	0.52	0.56	0.56	0.96	0.56	0.55
	*int.saddle**	0.55	0.44	0.36	0.54	0.61	0.92	0.55	0.43	0.48	0.37	0.54	0.62	0.58	0.38	0.96
	*int.saddle$_{var}$**	0.58	0.47	0.43	0.46	0.63	0.66	0.58	0.45	0.58	0.43	0.47	0.63	0.44	0.28	0.67
	*saddle$_{var}$**	0.76	0.55	0.47	0.31	0.54	0.46	0.76	0.53	0.62	0.47	0.32	0.54	0.38	0.25	0.48
UII	*folded-I*	0.93	0.62	0.53	0.36	0.52	0.54	0.94	0.61	0.65	0.53	0.35	0.53	0.48	0.33	0.52
	folded-IVb2	0.62	0.95	0.55	0.43	0.42	0.45	0.62	0.94	0.73	0.55	0.43	0.44	0.36	0.45	0.44
	folded-II	0.75	0.54	0.46	0.32	0.57	0.47	0.75	0.53	0.63	0.45	0.33	0.57	0.39	0.27	0.50
	folded-III	0.58	0.43	0.49	0.32	0.47	0.42	0.57	0.42	0.53	0.49	0.34	0.49	0.41	0.29	0.39
	inv-folded	0.43	0.31	0.42	0.50	0.36	0.50	0.43	0.30	0.37	0.44	0.51	0.38	0.49	0.28	0.49
	Ω-I$_{open}$	0.47	0.51	0.88	0.51	0.37	0.35	0.46	0.51	0.42	0.86	0.51	0.40	0.51	0.60	0.32
	Ω-I$_{hbond}$	0.57	0.62	0.83	0.57	0.38	0.41	0.57	0.63	0.51	0.84	0.57	0.40	0.50	0.52	0.40
	Ω-II	0.36	0.41	0.55	0.96	0.40	0.53	0.36	0.41	0.37	0.56	0.95	0.43	0.58	0.60	0.54
	lasso	0.45	0.35	0.52	0.56	0.55	0.60	0.45	0.34	0.37	0.51	0.56	0.57	0.92	0.56	0.56
	scoop	0.47	0.54	0.60	0.39	0.39	0.41	0.47	0.53	0.52	0.60	0.40	0.41	0.54	0.56	0.38
	circle	0.50	0.34	0.62	0.41	0.46	0.40	0.51	0.33	0.36	0.61	0.40	0.45	0.65	0.47	0.39
URP	*hybrid*	0.59	0.91	0.59	0.40	0.39	0.40	0.59	0.92	0.69	0.59	0.40	0.41	0.33	0.48	0.39
	sheet	0.52	0.70	0.62	0.39	0.45	0.41	0.53	0.69	0.62	0.61	0.41	0.47	0.45	0.55	0.39
	Ω-I$_{hbond}$	0.57	0.62	0.83	0.57	0.37	0.41	0.57	0.63	0.51	0.84	0.57	0.40	0.50	0.53	0.40
	Ω-I$_{open}$	0.46	0.51	0.88	0.50	0.37	0.35	0.47	0.51	0.42	0.86	0.50	0.40	0.51	0.61	0.32
	Ω-II	0.36	0.42	0.54	0.95	0.40	0.54	0.37	0.41	0.37	0.55	0.94	0.42	0.56	0.59	0.55
	lasso$_{45pbr}$	0.34	0.45	0.62	0.60	0.41	0.40	0.35	0.45	0.40	0.63	0.60	0.43	0.58	0.94	0.37

[§] Similarities > 0.65 are highlighted in green, > 0.90 in red. Transient states are marked with *. [a] $\Phi\Psi$ i+1 to i+4 and Φ i+5. [b] AVP (i=1), OT (i=1), dOT (i=1), CT (i=1), UII (i=5) and URP (i=2). Abbreviations: *int.saddle* = hybrid of *open* an *saddle*, *open$_{23var}$** =*scoop*-like hybrid of *open$_{23pbr}$* and *cl.open*, *saddle$_{var}$* = *folded-II, int.saddle$_{var}$* = hybrid of *int.saddle* and *saddle$_{var}$*.

Table A8.4 continued

Circular Similarity of Ring Torsions		dOT			UII										
		tws	tws_{helix}	$open_{23var}$*	*folded-I*	*folded-IVb2*	*folded-II*	*folded-III*	*inv-folded*	Ω-I_{open}	Ω-I_{hbond}	Ω-*II*	*lasso*	*scoop*	*circle*
AVP	*saddle*	0.60	0.63	0.52	0.93	0.62	0.75	0.58	0.43	0.47	0.57	0.36	0.45	0.47	0.50
	tws	0.93	0.73	0.59	0.62	0.95	0.54	0.43	0.31	0.51	0.62	0.41	0.35	0.54	0.34
	clop	0.55	0.48	0.67	0.53	0.55	0.46	0.49	0.42	0.88	0.83	0.55	0.52	0.60	0.62
	$clop_{45pbr}$*	0.40	0.37	0.50	0.36	0.43	0.32	0.32	0.50	0.51	0.57	0.96	0.56	0.39	0.41
	open	0.40	0.55	0.38	0.52	0.42	0.57	0.47	0.36	0.37	0.38	0.40	0.55	0.39	0.46
	*int.saddle**	0.41	0.49	0.38	0.54	0.45	0.47	0.42	0.50	0.35	0.41	0.53	0.60	0.41	0.40
OT	*saddle*	0.60	0.63	0.52	0.94	0.62	0.75	0.57	0.43	0.46	0.57	0.36	0.45	0.47	0.51
	tws	0.94	0.71	0.59	0.61	0.94	0.53	0.42	0.30	0.51	0.63	0.41	0.34	0.53	0.33
	tws_{helix}	0.70	0.96	0.55	0.65	0.73	0.63	0.53	0.37	0.42	0.51	0.37	0.37	0.52	0.36
	clop	0.55	0.47	0.67	0.53	0.55	0.45	0.49	0.44	0.86	0.84	0.56	0.51	0.60	0.61
	$clop_{45pbr}$	0.40	0.38	0.51	0.35	0.43	0.33	0.34	0.51	0.51	0.57	0.95	0.56	0.40	0.40
	open	0.42	0.56	0.39	0.53	0.44	0.57	0.49	0.38	0.40	0.40	0.43	0.57	0.41	0.45
	$open_{23pbr}$	0.35	0.39	0.55	0.48	0.36	0.39	0.41	0.49	0.51	0.50	0.58	0.92	0.54	0.65
	$open_{2334pbr}$*	0.44	0.40	0.62	0.33	0.45	0.27	0.29	0.28	0.60	0.52	0.60	0.56	0.56	0.47
	*int.saddle**	0.40	0.47	0.37	0.52	0.44	0.50	0.39	0.49	0.32	0.40	0.54	0.56	0.38	0.39
dOT	*tws*	1.00	0.70	0.60	0.61	0.91	0.53	0.42	0.30	0.50	0.63	0.40	0.32	0.51	0.32
	tws_{helix}	0.70	1.00	0.54	0.65	0.73	0.63	0.55	0.37	0.43	0.51	0.37	0.38	0.53	0.36
	$open_{23var}$*	0.60	0.54	1.00	0.54	0.58	0.42	0.42	0.43	0.61	0.68	0.50	0.52	0.72	0.53
CT	*saddle*	0.60	0.63	0.53	0.95	0.61	0.73	0.54	0.42	0.46	0.57	0.35	0.44	0.46	0.51
	$clop_{45pbr}$	0.41	0.38	0.51	0.36	0.43	0.33	0.33	0.51	0.51	0.57	0.95	0.56	0.39	0.41
	open	0.42	0.56	0.38	0.52	0.44	0.59	0.50	0.37	0.40	0.40	0.44	0.57	0.41	0.45
	$open_{23pbr}$	0.32	0.36	0.53	0.45	0.33	0.38	0.38	0.46	0.51	0.48	0.55	0.91	0.53	0.67
	*int.saddle**	0.41	0.48	0.38	0.52	0.45	0.49	0.39	0.52	0.32	0.40	0.54	0.56	0.38	0.38
	$int.saddle_{var}$*	0.43	0.57	0.37	0.57	0.47	0.65	0.52	0.54	0.38	0.46	0.46	0.43	0.37	0.40
	$saddle_{var}$*	0.53	0.63	0.42	0.75	0.55	0.96	0.70	0.42	0.42	0.51	0.33	0.38	0.41	0.43
UII	*folded-I*	0.61	0.65	0.54	1.00	0.61	0.75	0.56	0.43	0.50	0.58	0.36	0.45	0.49	0.53
	folded-IVb2	0.91	0.73	0.58	0.61	1.00	0.54	0.44	0.32	0.49	0.62	0.43	0.34	0.52	0.31
	folded-II	0.53	0.63	0.42	0.75	0.54	1.00	0.68	0.41	0.41	0.50	0.33	0.38	0.40	0.45
	folded-III	0.42	0.55	0.42	0.56	0.44	0.68	1.00	0.47	0.45	0.46	0.33	0.42	0.49	0.48
	inv-folded	0.30	0.37	0.43	0.43	0.32	0.41	0.47	1.00	0.38	0.42	0.49	0.47	0.41	0.46
	Ω-I_{open}	0.50	0.43	0.61	0.50	0.49	0.41	0.45	0.38	1.00	0.72	0.51	0.51	0.60	0.66
	Ω-I_{hbond}	0.63	0.51	0.68	0.58	0.62	0.50	0.46	0.42	0.72	1.00	0.56	0.46	0.53	0.54
	Ω-*II*	0.40	0.37	0.50	0.36	0.43	0.33	0.33	0.49	0.51	0.56	1.00	0.55	0.38	0.41
	lasso	0.32	0.38	0.52	0.45	0.34	0.38	0.42	0.47	0.51	0.46	0.55	1.00	0.54	0.65
	scoop	0.51	0.53	0.72	0.49	0.52	0.40	0.49	0.41	0.60	0.53	0.38	0.54	1.00	0.65
	circle	0.32	0.36	0.53	0.53	0.31	0.45	0.48	0.46	0.66	0.54	0.41	0.65	0.65	1.00
URP	*hybrid*	0.91	0.68	0.61	0.60	0.89	0.52	0.42	0.29	0.54	0.65	0.40	0.31	0.53	0.36
	sheet	0.67	0.62	0.63	0.55	0.66	0.45	0.34	0.32	0.62	0.52	0.40	0.44	0.71	0.51
	Ω-I_{hbond}	0.62	0.51	0.68	0.58	0.62	0.49	0.46	0.42	0.72	0.99	0.56	0.46	0.53	0.54
	Ω-I_{open}	0.50	0.43	0.62	0.50	0.49	0.40	0.42	0.37	0.95	0.73	0.50	0.50	0.59	0.67
	Ω-*II*	0.42	0.37	0.51	0.38	0.43	0.33	0.31	0.50	0.53	0.56	0.93	0.54	0.39	0.42
	$lasso_{45pbr}$	0.44	0.40	0.60	0.37	0.44	0.31	0.31	0.28	0.63	0.56	0.59	0.57	0.60	0.48

Table A8.4 continued

Circular Similarity of Ring Torsions		CT							URP					
		saddle	*clop*$_{45pbr}$	*open*	*open*$_{23pbr}$	*int.saddle**	*int.saddle*$_{var}$*	*saddle*$_{var}$*	*hybrid*	*sheet*	*Ω-I*$_{hbond}$	*Ω-I*$_{open}$	*Ω-II*	*lasso*$_{45pbr}$
AVP	*saddle*	0.93	0.37	0.52	0.45	0.55	0.58	0.76	0.59	0.52	0.57	0.46	0.36	0.34
	tws	0.61	0.42	0.45	0.35	0.44	0.47	0.55	0.91	0.70	0.62	0.51	0.42	0.45
	clop	0.51	0.56	0.42	0.52	0.36	0.43	0.47	0.59	0.62	0.83	0.88	0.54	0.62
	clop$_{45pbr}$*	0.35	0.98	0.44	0.56	0.54	0.46	0.31	0.40	0.39	0.57	0.50	0.95	0.60
	open	0.52	0.42	0.94	0.55	0.61	0.63	0.54	0.39	0.45	0.37	0.37	0.40	0.41
	*int.saddle**	0.55	0.55	0.68	0.59	0.92	0.66	0.46	0.40	0.41	0.41	0.35	0.54	0.40
OT	*saddle*	0.95	0.36	0.53	0.45	0.55	0.58	0.76	0.59	0.53	0.57	0.47	0.37	0.35
	tws	0.60	0.42	0.43	0.33	0.43	0.45	0.53	0.92	0.69	0.63	0.51	0.41	0.45
	tws$_{helix}$	0.63	0.38	0.56	0.36	0.48	0.58	0.62	0.69	0.62	0.51	0.42	0.37	0.40
	clop	0.50	0.57	0.41	0.52	0.37	0.43	0.47	0.59	0.61	0.84	0.86	0.55	0.63
	clop$_{45pbr}$	0.35	0.97	0.45	0.56	0.54	0.47	0.32	0.40	0.41	0.57	0.50	0.94	0.60
	open	0.52	0.44	0.97	0.56	0.62	0.63	0.54	0.41	0.47	0.40	0.40	0.42	0.43
	open$_{23pbr}$	0.47	0.58	0.57	0.96	0.58	0.44	0.38	0.33	0.45	0.50	0.51	0.56	0.58
	open$_{2334pbr}$*	0.30	0.60	0.42	0.56	0.38	0.28	0.25	0.48	0.55	0.53	0.61	0.59	0.94
	*int.saddle**	0.54	0.55	0.63	0.55	0.96	0.67	0.48	0.39	0.39	0.40	0.32	0.55	0.37
dOT	*tws*	0.60	0.41	0.42	0.32	0.41	0.43	0.53	0.91	0.67	0.62	0.50	0.42	0.44
	tws$_{helix}$	0.63	0.38	0.56	0.36	0.48	0.57	0.63	0.68	0.62	0.51	0.43	0.37	0.40
	open$_{23var}$*	0.53	0.51	0.38	0.53	0.38	0.37	0.42	0.61	0.63	0.68	0.62	0.51	0.60
CT	*saddle*	1.00	0.36	0.51	0.44	0.54	0.58	0.75	0.59	0.53	0.57	0.46	0.36	0.34
	clop$_{45pbr}$	0.36	1.00	0.45	0.56	0.55	0.47	0.32	0.40	0.39	0.56	0.50	0.95	0.60
	open	0.51	0.45	1.00	0.56	0.63	0.63	0.56	0.41	0.45	0.40	0.40	0.43	0.43
	open$_{23pbr}$	0.44	0.56	0.56	1.00	0.56	0.42	0.37	0.31	0.44	0.47	0.51	0.54	0.57
	*int.saddle**	0.54	0.55	0.63	0.56	1.00	0.67	0.47	0.40	0.39	0.40	0.31	0.55	0.37
	int.saddle$_{var}$*	0.58	0.47	0.63	0.42	0.67	1.00	0.64	0.43	0.42	0.45	0.37	0.47	0.28
	saddle$_{var}$*	0.75	0.32	0.56	0.37	0.47	0.64	1.00	0.52	0.44	0.51	0.40	0.33	0.29
UII	*folded-I*	0.95	0.36	0.52	0.45	0.52	0.57	0.75	0.60	0.55	0.58	0.50	0.38	0.37
	folded-IVb2	0.61	0.43	0.44	0.33	0.45	0.47	0.55	0.89	0.66	0.62	0.49	0.43	0.44
	folded-II	0.73	0.33	0.59	0.38	0.49	0.65	0.96	0.52	0.45	0.49	0.40	0.33	0.31
	folded-III	0.54	0.33	0.50	0.38	0.39	0.52	0.70	0.42	0.34	0.46	0.42	0.31	0.31
	inv-folded	0.42	0.51	0.37	0.46	0.52	0.54	0.42	0.29	0.32	0.42	0.37	0.50	0.28
	Ω-I$_{open}$	0.46	0.51	0.40	0.51	0.32	0.38	0.42	0.54	0.62	0.72	0.95	0.53	0.63
	Ω-I$_{hbond}$	0.57	0.57	0.40	0.48	0.40	0.46	0.51	0.65	0.52	0.99	0.73	0.56	0.56
	Ω-II	0.35	0.95	0.44	0.55	0.54	0.46	0.33	0.40	0.40	0.56	0.50	0.93	0.59
	lasso	0.44	0.56	0.57	0.91	0.56	0.43	0.38	0.31	0.44	0.46	0.50	0.54	0.57
	scoop	0.46	0.39	0.41	0.53	0.38	0.37	0.41	0.53	0.71	0.53	0.59	0.39	0.60
	circle	0.51	0.41	0.45	0.67	0.38	0.40	0.43	0.36	0.51	0.54	0.67	0.42	0.48
URP	*hybrid*	0.59	0.40	0.41	0.31	0.40	0.43	0.52	1.00	0.71	0.65	0.55	0.40	0.52
	sheet	0.53	0.39	0.45	0.44	0.39	0.42	0.44	0.71	1.00	0.52	0.63	0.41	0.58
	Ω-I$_{hbond}$	0.57	0.56	0.40	0.47	0.40	0.45	0.51	0.65	0.52	1.00	0.73	0.56	0.56
	Ω-I$_{open}$	0.46	0.50	0.40	0.51	0.31	0.37	0.40	0.55	0.63	0.73	1.00	0.53	0.64
	Ω-II	0.36	0.95	0.43	0.54	0.55	0.47	0.33	0.40	0.41	0.56	0.53	1.00	0.59
	lasso$_{45pbr}$	0.34	0.60	0.43	0.57	0.37	0.28	0.29	0.52	0.58	0.56	0.64	0.59	1.00

Coordinate Files

Coordinate files of representative and several transient conformations as discussed in the main text (*).

OT, saddle (FOLDED)

OT_MD-I_15us_T16_7

```
ATOM   1 N   CYX  1   -17.763 -7.989 13.506 0.00 0.00
ATOM   2 H1  CYX  1   -17.482 -8.951 13.627 0.00 0.00
ATOM   3 H2  CYX  1   -18.450 -7.690 12.828 0.00 0.00
ATOM   4 H3  CYX  1   -18.172 -7.808 14.412 0.00 0.00
ATOM   5 CA  CYX  1   -16.556 -7.260 13.262 0.00 0.00
ATOM   6 HA  CYX  1   -16.879 -6.225 13.148 0.00 0.00
ATOM   7 CB  CYX  1   -15.737 -7.314 14.453 0.00 0.00
ATOM   8 HB2 CYX  1   -16.345 -7.290 15.358 0.00 0.00
ATOM   9 HB3 CYX  1   -15.199 -8.260 14.497 0.00 0.00
ATOM  10 SG  CYX  1   -14.584 -5.900 14.696 0.00 0.00
ATOM  11 C   CYX  1   -15.840 -7.488 11.947 0.00 0.00
ATOM  12 O   CYX  1   -15.963 -8.582 11.446 0.00 0.00
ATOM  13 N   TYR  2   -15.203 -6.449 11.408 0.00 0.00
ATOM  14 H   TYR  2   -15.018 -5.551 11.832 0.00 0.00
ATOM  15 CA  TYR  2   -14.419 -6.591 10.216 0.00 0.00
ATOM  16 HA  TYR  2   -14.923 -7.237  9.498 0.00 0.00
ATOM  17 CB  TYR  2   -14.576 -5.255  9.468 0.00 0.00
ATOM  18 HB2 TYR  2   -15.642 -5.099  9.304 0.00 0.00
ATOM  19 HB3 TYR  2   -14.186 -4.475 10.123 0.00 0.00
ATOM  20 CG  TYR  2   -13.869 -5.237  8.074 0.00 0.00
ATOM  21 CD1 TYR  2   -14.449 -5.969  7.018 0.00 0.00
ATOM  22 HD1 TYR  2   -15.457 -6.341  7.122 0.00 0.00
ATOM  23 CE1 TYR  2   -13.778 -6.114  5.782 0.00 0.00
ATOM  24 HE1 TYR  2   -14.286 -6.612  4.969 0.00 0.00
ATOM  25 CZ  TYR  2   -12.539 -5.418  5.611 0.00 0.00
ATOM  26 OH  TYR  2   -11.816 -5.666  4.504 0.00 0.00
ATOM  27 HH  TYR  2   -12.268 -6.250  3.891 0.00 0.00
ATOM  28 CE2 TYR  2   -11.973 -4.629  6.623 0.00 0.00
ATOM  29 HE2 TYR  2   -11.076 -4.060  6.429 0.00 0.00
ATOM  30 CD2 TYR  2   -12.633 -4.601  7.888 0.00 0.00
ATOM  31 HD2 TYR  2   -12.184 -4.028  8.686 0.00 0.00
ATOM  32 C   TYR  2   -12.923 -6.962 10.265 0.00 0.00
ATOM  33 O   TYR  2   -12.236 -6.539 11.182 0.00 0.00
ATOM  34 N   ILE  3   -12.386 -7.642  9.253 0.00 0.00
ATOM  35 H   ILE  3   -13.003 -8.003  8.540 0.00 0.00
ATOM  36 CA  ILE  3   -11.096 -8.357  9.284 0.00 0.00
ATOM  37 HA  ILE  3   -11.134 -9.014 10.153 0.00 0.00
ATOM  38 CB  ILE  3   -10.924 -9.238  8.051 0.00 0.00
ATOM  39 HB  ILE  3   -10.028 -9.819  8.268 0.00 0.00
ATOM  40 CG2 ILE  3   -12.060 -10.267  8.132 0.00 0.00
ATOM  41 HG21 ILE  3   -13.041 -9.871  7.867 0.00 0.00
ATOM  42 HG22 ILE  3   -11.945 -11.044  7.376 0.00 0.00
ATOM  43 HG23 ILE  3   -11.976 -10.788  9.086 0.00 0.00
ATOM  44 CG1 ILE  3   -10.729 -8.522  6.731 0.00 0.00
ATOM  45 HG12 ILE  3   -11.679 -8.042  6.495 0.00 0.00
ATOM  46 HG13 ILE  3   -9.948 -7.770  6.842 0.00 0.00
ATOM  47 CD1 ILE  3   -10.483 -9.289  5.425 0.00 0.00
ATOM  48 HD11 ILE  3   -10.730 -8.562  4.652 0.00 0.00
ATOM  49 HD12 ILE  3   -9.442 -9.590  5.303 0.00 0.00
ATOM  50 HD13 ILE  3   -11.127 -10.162  5.315 0.00 0.00
ATOM  51 C   ILE  3   -9.879 -7.420  9.491 0.00 0.00
ATOM  52 O   ILE  3   -8.887 -7.893 10.105 0.00 0.00
ATOM  53 N   GLN  4   -9.952 -6.140  9.101 0.00 0.00
ATOM  54 H   GLN  4   -10.750 -5.763  8.609 0.00 0.00
ATOM  55 CA  GLN  4   -8.830 -5.199  9.427 0.00 0.00
ATOM  56 HA  GLN  4   -7.897 -5.734  9.602 0.00 0.00
ATOM  57 CB  GLN  4   -8.485 -4.280  8.377 0.00 0.00
ATOM  58 HB2 GLN  4   -9.304 -3.617  8.099 0.00 0.00
ATOM  59 HB3 GLN  4   -7.624 -3.744  8.776 0.00 0.00
ATOM  60 CG  GLN  4   -8.070 -5.019  7.068 0.00 0.00
ATOM  61 HG2 GLN  4   -7.410 -5.833  7.367 0.00 0.00
ATOM  62 HG3 GLN  4   -8.958 -5.430  6.589 0.00 0.00
ATOM  63 CD  GLN  4   -7.396 -4.183  6.050 0.00 0.00
ATOM  64 OE1 GLN  4   -7.062 -3.021  6.247 0.00 0.00
ATOM  65 NE2 GLN  4   -7.245 -4.743  4.901 0.00 0.00
ATOM  66 HE21 GLN  4   -6.786 -4.224  4.166 0.00 0.00
ATOM  67 HE22 GLN  4   -7.611 -5.680  4.816 0.00 0.00
ATOM  68 C   GLN  4   -9.072 -4.525 10.783 0.00 0.00
ATOM  69 O   GLN  4   -8.230 -3.717 11.236 0.00 0.00
ATOM  70 N   ASN  5   -10.121 -4.828 11.516 0.00 0.00
ATOM  71 H   ASN  5   -10.777 -5.543 11.239 0.00 0.00
ATOM  72 CA  ASN  5   -10.423 -4.293 12.843 0.00 0.00
ATOM  73 HA  ASN  5   -9.674 -3.538 13.082 0.00 0.00
ATOM  74 CB  ASN  5   -11.765 -3.453 12.813 0.00 0.00
ATOM  75 HB2 ASN  5   -11.851 -2.853 11.908 0.00 0.00
ATOM  76 HB3 ASN  5   -12.580 -4.131 12.562 0.00 0.00
ATOM  77 CG  ASN  5   -12.008 -2.641 14.057 0.00 0.00
ATOM  78 OD1 ASN  5   -13.080 -2.794 14.624 0.00 0.00
ATOM  79 ND2 ASN  5   -11.068 -1.944 14.640 0.00 0.00
ATOM  80 HD21 ASN  5   -11.322 -1.592 15.553 0.00 0.00
ATOM  81 HD22 ASN  5   -10.238 -1.750 14.099 0.00 0.00
ATOM  82 C   ASN  5   -10.422 -5.347 13.988 0.00 0.00
ATOM  83 O   ASN  5   -9.934 -5.077 15.097 0.00 0.00
ATOM  84 N   CYX  6   -10.795 -6.565 13.663 0.00 0.00
ATOM  85 H   CYX  6   -11.221 -6.776 12.772 0.00 0.00
ATOM  86 CA  CYX  6   -10.911 -7.689 14.622 0.00 0.00
ATOM  87 HA  CYX  6   -10.281 -7.506 15.493 0.00 0.00
ATOM  88 CB  CYX  6   -12.346 -7.900 15.107 0.00 0.00
ATOM  89 HB2 CYX  6   -12.942 -8.248 14.263 0.00 0.00
ATOM  90 HB3 CYX  6   -12.293 -8.842 15.654 0.00 0.00
ATOM  91 SG  CYX  6   -13.244 -6.638 16.100 0.00 0.00
ATOM  92 C   CYX  6   -10.419 -8.998 14.027 0.00 0.00
ATOM  93 O   CYX  6   -10.641 -9.286 12.802 0.00 0.00
ATOM  94 N   PRO  7   -9.942 -9.930 14.890 0.00 0.00
ATOM  95 CD  PRO  7   -9.416 -9.615 16.249 0.00 0.00
ATOM  96 HD2 PRO  7   -10.353 -9.579 16.806 0.00 0.00
ATOM  97 HD3 PRO  7   -8.803 -8.722 16.129 0.00 0.00
ATOM  98 CG  PRO  7   -8.532 -10.762 16.604 0.00 0.00
ATOM  99 HG2 PRO  7   -8.642 -10.981 17.666 0.00 0.00
ATOM 100 HG3 PRO  7   -7.525 -10.555 16.241 0.00 0.00
ATOM 101 CB  PRO  7   -9.044 -11.899 15.778 0.00 0.00
ATOM 102 HB2 PRO  7   -9.816 -12.499 16.260 0.00 0.00
ATOM 103 HB3 PRO  7   -8.224 -12.561 15.501 0.00 0.00
ATOM 104 CA  PRO  7   -9.445 -11.223 14.469 0.00 0.00
ATOM 105 HA  PRO  7   -8.527 -11.073 13.901 0.00 0.00
ATOM 106 C   PRO  7   -10.500 -12.058 13.696 0.00 0.00
ATOM 107 O   PRO  7   -11.702 -12.080 13.985 0.00 0.00
ATOM 108 N   LEU  8   -10.064 -12.859 12.687 0.00 0.00
ATOM 109 H   LEU  8   -9.081 -12.758 12.479 0.00 0.00
ATOM 110 CA  LEU  8   -10.833 -13.794 11.832 0.00 0.00
ATOM 111 HA  LEU  8   -11.899 -13.661 12.016 0.00 0.00
ATOM 112 CB  LEU  8   -10.468 -13.255 10.361 0.00 0.00
ATOM 113 HB2 LEU  8   -10.959 -12.287 10.260 0.00 0.00
ATOM 114 HB3 LEU  8   -9.403 -13.111 10.181 0.00 0.00
ATOM 115 CG  LEU  8   -10.861 -14.135  9.145 0.00 0.00
ATOM 116 HG  LEU  8   -10.448 -15.138  9.259 0.00 0.00
ATOM 117 CD1 LEU  8   -12.341 -14.043  8.985 0.00 0.00
ATOM 118 HD11 LEU  8   -12.619 -13.089  8.537 0.00 0.00
ATOM 119 HD12 LEU  8   -12.622 -14.843  8.301 0.00 0.00
ATOM 120 HD13 LEU  8   -12.852 -14.153  9.941 0.00 0.00
ATOM 121 CD2 LEU  8   -10.142 -13.606  7.912 0.00 0.00
ATOM 122 HD21 LEU  8   -9.054 -13.676  7.927 0.00 0.00
ATOM 123 HD22 LEU  8   -10.428 -14.197  7.042 0.00 0.00
ATOM 124 HD23 LEU  8   -10.512 -12.615  7.647 0.00 0.00
ATOM 125 C   LEU  8   -10.531 -15.316 12.150 0.00 0.00
ATOM 126 O   LEU  8   -9.477 -15.910 11.882 0.00 0.00
ATOM 127 N   GLY  9   -11.517 -16.021 12.576 0.00 0.00
ATOM 128 H   GLY  9   -12.296 -15.514 12.972 0.00 0.00
ATOM 129 CA  GLY  9   -11.503 -17.506 12.559 0.00 0.00
ATOM 130 HA2 GLY  9   -10.667 -17.978 13.074 0.00 0.00
ATOM 131 HA3 GLY  9   -12.433 -17.824 13.031 0.00 0.00
ATOM 132 C   GLY  9   -11.643 -18.138 11.061 0.00 0.00
ATOM 133 O   GLY  9   -11.906 -17.477 10.045 0.00 0.00
ATOM 134 N   NHE  10   -11.520 -19.441 10.990 0.00 0.00
ATOM 135 HN1 NHE  10   -11.554 -20.046 11.798 0.00 0.00
ATOM 136 HN2 NHE  10   -11.701 -19.891 10.105 0.00 0.00
TER  137      NHE  10
ATOM 138 Cl- Cl-  11     3.195 -13.426 12.937 0.00 0.00
TER  139      Cl-  1
END
```

OT, twisted saddle (FOLDED)

OT_MD-IV_10us_T16_9

```
ATOM   1 N   CYX  1     5.684 -13.412  1.766 0.00 0.00
ATOM   2 H1  CYX  1     6.619 -13.755  1.598 0.00 0.00
ATOM   3 H2  CYX  1     5.740 -12.870  2.616 0.00 0.00
ATOM   4 H3  CYX  1     5.412 -12.787  1.021 0.00 0.00
ATOM   5 CA  CYX  1     4.771 -14.527  1.903 0.00 0.00
ATOM   6 HA  CYX  1     4.908 -14.830  2.941 0.00 0.00
ATOM   7 CB  CYX  1     3.333 -14.184  1.706 0.00 0.00
ATOM   8 HB2 CYX  1     2.655 -15.012  1.914 0.00 0.00
ATOM   9 HB3 CYX  1     3.135 -13.370  2.404 0.00 0.00
ATOM  10 SG  CYX  1     2.988 -13.457  0.024 0.00 0.00
ATOM  11 C   CYX  1     5.161 -15.697  0.976 0.00 0.00
ATOM  12 O   CYX  1     5.637 -15.439 -0.154 0.00 0.00
ATOM  13 N   TYR  2     5.062 -16.924  1.453 0.00 0.00
ATOM  14 H   TYR  2     4.540 -17.010  2.313 0.00 0.00
ATOM  15 CA  TYR  2     5.483 -18.174  0.797 0.00 0.00
ATOM  16 HA  TYR  2     6.519 -18.010  0.499 0.00 0.00
ATOM  17 CB  TYR  2     5.326 -19.371  1.717 0.00 0.00
ATOM  18 HB2 TYR  2     5.822 -19.262  2.682 0.00 0.00
ATOM  19 HB3 TYR  2     4.253 -19.536  1.818 0.00 0.00
ATOM  20 CG  TYR  2     5.930 -20.655  1.171 0.00 0.00
ATOM  21 CD1 TYR  2     7.338 -20.673  0.950 0.00 0.00
ATOM  22 HD1 TYR  2     7.899 -19.799  1.246 0.00 0.00
ATOM  23 CE1 TYR  2     7.928 -21.807  0.316 0.00 0.00
ATOM  24 HE1 TYR  2     8.953 -21.685 -0.002 0.00 0.00
ATOM  25 CZ  TYR  2     7.065 -22.884 -0.085 0.00 0.00
ATOM  26 OH  TYR  2     7.603 -23.892 -0.815 0.00 0.00
ATOM  27 HH  TYR  2     6.912 -24.465 -1.156 0.00 0.00
ATOM  28 CE2 TYR  2     5.673 -22.851  0.261 0.00 0.00
ATOM  29 HE2 TYR  2     4.991 -23.627 -0.052 0.00 0.00
ATOM  30 CD2 TYR  2     5.099 -21.732  0.911 0.00 0.00
ATOM  31 HD2 TYR  2     4.026 -21.718  1.027 0.00 0.00
ATOM  32 C   TYR  2     4.646 -18.346 -0.488 0.00 0.00
ATOM  33 O   TYR  2     3.465 -17.962 -0.453 0.00 0.00
ATOM  48 HD11 ILE  3     4.355 -22.096 -5.158 0.00 0.00
ATOM  49 HD12 ILE  3     5.118 -23.333 -4.090 0.00 0.00
ATOM  50 HD13 ILE  3     3.831 -22.393 -3.505 0.00 0.00
ATOM  51 C   ILE  3     3.070 -20.071 -2.445 0.00 0.00
ATOM  52 O   ILE  3     3.073 -20.931 -1.567 0.00 0.00
ATOM  53 N   GLN  4     2.875 -19.710 -3.063 0.00 0.00
ATOM  54 H   GLN  4     2.079 -19.145 -3.893 0.00 0.00
ATOM  55 CA  GLN  4     0.641 -20.194 -2.820 0.00 0.00
ATOM  56 HA  GLN  4     0.033 -19.508 -3.410 0.00 0.00
ATOM  57 CB  GLN  4     0.468 -21.705 -3.216 0.00 0.00
ATOM  58 HB2 GLN  4     1.228 -22.392 -2.846 0.00 0.00
ATOM  59 HB3 GLN  4    -0.499 -22.013 -2.818 0.00 0.00
ATOM  60 CG  GLN  4     0.330 -21.763 -4.741 0.00 0.00
ATOM  61 HG2 GLN  4    -0.481 -21.079 -4.992 0.00 0.00
ATOM  62 HG3 GLN  4     1.234 -21.363 -5.200 0.00 0.00
ATOM  63 CD  GLN  4    -0.020 -23.141 -5.295 0.00 0.00
ATOM  64 OE1 GLN  4     0.797 -23.844 -5.845 0.00 0.00
ATOM  65 NE2 GLN  4    -1.294 -23.548 -5.245 0.00 0.00
ATOM  66 HE21 GLN  4    -1.458 -24.515 -5.488 0.00 0.00
ATOM  67 HE22 GLN  4    -1.983 -23.021 -4.729 0.00 0.00
ATOM  68 C   GLN  4     0.058 -19.960 -1.398 0.00 0.00
ATOM  69 O   GLN  4    -1.091 -20.320 -1.176 0.00 0.00
ATOM  70 N   ASN  5     0.663 -19.290 -0.407 0.00 0.00
ATOM  71 H   ASN  5     1.475 -18.761 -0.692 0.00 0.00
ATOM  72 CA  ASN  5    -0.017 -18.930  0.864 0.00 0.00
ATOM  73 HA  ASN  5    -0.808 -19.666  1.007 0.00 0.00
ATOM  74 CB  ASN  5     0.987 -19.123  1.969 0.00 0.00
ATOM  75 HB2 ASN  5     1.777 -19.850  1.780 0.00 0.00
ATOM  76 HB3 ASN  5     1.477 -18.158  2.102 0.00 0.00
ATOM  77 CG  ASN  5     0.359 -19.436  3.349 0.00 0.00
ATOM  78 OD1 ASN  5    -0.452 -20.342  3.493 0.00 0.00
ATOM  79 ND2 ASN  5     0.837 -18.944  4.499 0.00 0.00
ATOM  80 HD21 ASN  5     0.387 -19.218  5.361 0.00 0.00
ATOM  95 CD  PRO  7    -3.008 -13.580 -1.193 0.00 0.00
ATOM  96 HD2 PRO  7    -2.422 -12.669 -1.310 0.00 0.00
ATOM  97 HD3 PRO  7    -2.960 -13.892 -0.150 0.00 0.00
ATOM  98 CG  PRO  7    -4.479 -13.346 -1.359 0.00 0.00
ATOM  99 HG2 PRO  7    -4.718 -12.322 -1.073 0.00 0.00
ATOM 100 HG3 PRO  7    -5.018 -14.043 -0.718 0.00 0.00
ATOM 101 CB  PRO  7    -4.676 -13.648 -2.867 0.00 0.00
ATOM 102 HB2 PRO  7    -4.807 -12.742 -3.458 0.00 0.00
ATOM 103 HB3 PRO  7    -5.551 -14.247 -3.123 0.00 0.00
ATOM 104 CA  PRO  7    -3.411 -14.423 -3.337 0.00 0.00
ATOM 105 HA  PRO  7    -3.780 -15.420 -3.579 0.00 0.00
ATOM 106 C   PRO  7    -2.646 -13.805 -4.558 0.00 0.00
ATOM 107 O   PRO  7    -1.697 -13.027 -4.340 0.00 0.00
ATOM 108 N   LEU  8    -2.995 -14.255 -5.767 0.00 0.00
ATOM 109 H   LEU  8    -3.725 -14.944 -5.876 0.00 0.00
ATOM 110 CA  LEU  8    -2.133 -13.872 -6.926 0.00 0.00
ATOM 111 HA  LEU  8    -1.096 -13.702 -6.634 0.00 0.00
ATOM 112 CB  LEU  8    -2.169 -14.992 -7.966 0.00 0.00
ATOM 113 HB2 LEU  8    -3.202 -15.274 -8.171 0.00 0.00
ATOM 114 HB3 LEU  8    -1.666 -15.840 -7.501 0.00 0.00
ATOM 115 CG  LEU  8    -1.645 -14.699 -9.370 0.00 0.00
ATOM 116 HG  LEU  8    -2.218 -13.883 -9.810 0.00 0.00
ATOM 117 CD1 LEU  8    -0.224 -14.071 -9.246 0.00 0.00
ATOM 118 HD11 LEU  8    -0.093 -13.151 -8.676 0.00 0.00
ATOM 119 HD12 LEU  8     0.354 -14.860 -8.765 0.00 0.00
ATOM 120 HD13 LEU  8     0.168 -13.901 -10.249 0.00 0.00
ATOM 121 CD2 LEU  8    -1.650 -16.000 -10.242 0.00 0.00
ATOM 122 HD21 LEU  8    -2.638 -16.388 -10.489 0.00 0.00
ATOM 123 HD22 LEU  8    -1.201 -15.748 -11.203 0.00 0.00
ATOM 124 HD23 LEU  8    -1.015 -16.696 -9.695 0.00 0.00
ATOM 125 C   LEU  8    -2.735 -12.570 -7.503 0.00 0.00
ATOM 126 O   LEU  8    -3.894 -12.549 -7.882 0.00 0.00
ATOM 127 N   GLY  9    -1.867 -11.532 -7.595 0.00 0.00
```

ATOM 34 N ILE 3 5.151 -18.999 -1.563 0.00 0.00
ATOM 35 H ILE 3 5.957 -19.585 -1.399 0.00 0.00
ATOM 36 CA ILE 3 4.370 -19.277 -2.736 0.00 0.00
ATOM 37 HA ILE 3 3.978 -18.284 -2.954 0.00 0.00
ATOM 38 CB ILE 3 5.172 -19.828 -3.972 0.00 0.00
ATOM 39 HB ILE 3 4.711 -19.688 -4.950 0.00 0.00
ATOM 40 CG2 ILE 3 6.392 -18.920 -4.202 0.00 0.00
ATOM 41 HG21 ILE 3 7.030 -19.202 -3.365 0.00 0.00
ATOM 42 HG22 ILE 3 6.819 -19.129 -5.183 0.00 0.00
ATOM 43 HG23 ILE 3 6.014 -17.898 -4.204 0.00 0.00
ATOM 44 CG1 ILE 3 5.713 -21.294 -3.689 0.00 0.00
ATOM 45 HG12 ILE 3 5.853 -21.395 -2.613 0.00 0.00
ATOM 46 HG13 ILE 3 6.606 -21.542 -4.264 0.00 0.00
ATOM 47 CD1 ILE 3 4.706 -22.326 -4.152 0.00 0.00
ATOM 81 HD22 ASN 5 1.306 -18.056 4.387 0.00 0.00
ATOM 82 C ASN 5 -0.668 -17.623 0.768 0.00 0.00
ATOM 83 O ASN 5 -1.397 -17.269 1.689 0.00 0.00
ATOM 84 N CYX 6 -0.348 -16.903 -0.337 0.00 0.00
ATOM 85 H CYX 6 0.160 -17.351 -1.087 0.00 0.00
ATOM 86 CA CYX 6 -0.823 -15.510 -0.568 0.00 0.00
ATOM 87 HA CYX 6 -1.577 -15.279 0.184 0.00 0.00
ATOM 88 CB CYX 6 0.309 -14.574 -0.281 0.00 0.00
ATOM 89 HB2 CYX 6 0.097 -13.570 -0.649 0.00 0.00
ATOM 90 HB3 CYX 6 0.519 -14.554 0.788 0.00 0.00
ATOM 91 SG CYX 6 1.967 -14.925 -0.960 0.00 0.00
ATOM 92 C CYX 6 -1.516 -15.381 -1.958 0.00 0.00
ATOM 93 O CYX 6 -1.036 -16.003 -2.926 0.00 0.00
ATOM 94 N PRO 7 -2.532 -14.523 -2.160 0.00 0.00
ATOM 128 H GLY 9 -0.889 -11.720 -7.426 0.00 0.00
ATOM 129 CA GLY 9 -2.249 -10.239 -8.213 0.00 0.00
ATOM 130 HA2 GLY 9 -3.225 -9.847 -7.929 0.00 0.00
ATOM 131 HA3 GLY 9 -1.653 -9.409 -7.835 0.00 0.00
ATOM 132 C GLY 9 -2.289 -10.338 -9.717 0.00 0.00
ATOM 133 O GLY 9 -1.854 -11.329 -10.265 0.00 0.00
ATOM 134 N NHE 10 -2.777 -9.337 -10.401 0.00 0.00
ATOM 135 HN1 NHE 10 -3.018 -8.450 -9.983 0.00 0.00
ATOM 136 HN2 NHE 10 -2.931 -9.549 -11.376 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 3.249 -11.561 6.458 0.00 0.00
TER 139 Cl- 11
END

OT, twisted saddlehelix (FOLDED)

OT_MD-IV_10us_T16_18

ATOM 1 N CYX 1 -5.904 10.130 -12.698 0.00 0.00
ATOM 2 H1 CYX 1 -5.587 9.642 -13.524 0.00 0.00
ATOM 3 H2 CYX 1 -5.815 11.128 -12.829 0.00 0.00
ATOM 4 H3 CYX 1 -6.871 9.851 -12.614 0.00 0.00
ATOM 5 CA CYX 1 -5.212 9.777 -11.447 0.00 0.00
ATOM 6 HA CYX 1 -5.393 8.724 -11.231 0.00 0.00
ATOM 7 CB CYX 1 -3.696 9.943 -11.520 0.00 0.00
ATOM 8 HB2 CYX 1 -3.325 9.652 -10.538 0.00 0.00
ATOM 9 HB3 CYX 1 -3.399 9.213 -12.273 0.00 0.00
ATOM 10 SG CYX 1 -3.011 11.557 -11.991 0.00 0.00
ATOM 11 C CYX 1 -5.764 10.625 -10.325 0.00 0.00
ATOM 12 O CYX 1 -6.120 11.772 -10.387 0.00 0.00
ATOM 13 N TYR 2 -5.862 9.978 -9.176 0.00 0.00
ATOM 14 H TYR 2 -5.441 9.069 -9.047 0.00 0.00
ATOM 15 CA TYR 2 -6.463 10.550 -7.977 0.00 0.00
ATOM 16 HA TYR 2 -7.359 11.068 -8.319 0.00 0.00
ATOM 17 CB TYR 2 -6.933 9.396 -7.101 0.00 0.00
ATOM 18 HB2 TYR 2 -7.447 8.616 -7.663 0.00 0.00
ATOM 19 HB3 TYR 2 -6.092 8.789 -6.768 0.00 0.00
ATOM 20 CG TYR 2 -7.781 9.779 -5.910 0.00 0.00
ATOM 21 CD1 TYR 2 -9.099 10.262 -6.120 0.00 0.00
ATOM 22 HD1 TYR 2 -9.560 10.333 -7.094 0.00 0.00
ATOM 23 CE1 TYR 2 -9.709 10.838 -5.002 0.00 0.00
ATOM 24 HE1 TYR 2 -10.663 11.320 -5.159 0.00 0.00
ATOM 25 CZ TYR 2 -9.142 10.719 -3.752 0.00 0.00
ATOM 26 OH TYR 2 -9.802 11.288 -2.667 0.00 0.00
ATOM 27 HH TYR 2 -10.752 11.422 -2.702 0.00 0.00
ATOM 28 CE2 TYR 2 -7.834 10.174 -3.540 0.00 0.00
ATOM 29 HE2 TYR 2 -7.331 10.129 -2.585 0.00 0.00
ATOM 30 CD2 TYR 2 -7.183 9.582 -4.638 0.00 0.00
ATOM 31 HD2 TYR 2 -6.253 9.068 -4.444 0.00 0.00
ATOM 32 C TYR 2 -5.629 11.591 -7.135 0.00 0.00
ATOM 33 O TYR 2 -4.403 11.500 -6.964 0.00 0.00
ATOM 34 N ILE 3 -6.351 12.630 -6.759 0.00 0.00
ATOM 35 H ILE 3 -7.279 12.761 -7.135 0.00 0.00
ATOM 36 CA ILE 3 -5.894 13.863 -6.047 0.00 0.00
ATOM 37 HA ILE 3 -6.593 14.640 -6.357 0.00 0.00
ATOM 38 CB ILE 3 -5.908 13.667 -4.469 0.00 0.00
ATOM 39 HB ILE 3 -5.379 14.520 -4.045 0.00 0.00
ATOM 40 CG2 ILE 3 -7.364 13.638 -4.045 0.00 0.00
ATOM 41 HG21 ILE 3 -7.997 12.996 -4.658 0.00 0.00
ATOM 42 HG22 ILE 3 -7.434 13.400 -2.984 0.00 0.00
ATOM 43 HG23 ILE 3 -7.768 14.647 -4.121 0.00 0.00
ATOM 44 CG1 ILE 3 -5.001 12.417 -4.092 0.00 0.00
ATOM 45 HG12 ILE 3 -4.071 12.598 -4.632 0.00 0.00
ATOM 46 HG13 ILE 3 -5.405 11.468 -4.443 0.00 0.00
ATOM 47 CD1 ILE 3 -4.539 12.292 -2.695 0.00 0.00
ATOM 48 HD11 ILE 3 -4.204 11.265 -2.554 0.00 0.00
ATOM 49 HD12 ILE 3 -3.645 12.897 -2.548 0.00 0.00
ATOM 50 HD13 ILE 3 -5.313 12.570 -1.978 0.00 0.00
ATOM 51 C ILE 3 -4.511 14.331 -6.577 0.00 0.00
ATOM 52 O ILE 3 -3.671 14.653 -5.770 0.00 0.00
ATOM 53 N GLN 4 -4.354 14.485 -7.932 0.00 0.00
ATOM 54 H GLN 4 -5.132 14.184 -8.501 0.00 0.00
ATOM 55 CA GLN 4 -3.141 15.007 -8.546 0.00 0.00
ATOM 56 HA GLN 4 -3.178 14.985 -9.636 0.00 0.00
ATOM 57 CB GLN 4 -2.923 16.464 -8.144 0.00 0.00
ATOM 58 HB2 GLN 4 -3.123 16.725 -7.105 0.00 0.00
ATOM 59 HB3 GLN 4 -1.916 16.802 -8.387 0.00 0.00
ATOM 60 CG GLN 4 -3.766 17.299 -9.093 0.00 0.00
ATOM 61 HG2 GLN 4 -3.410 17.168 -10.115 0.00 0.00
ATOM 62 HG3 GLN 4 -4.815 17.058 -8.920 0.00 0.00
ATOM 63 CD GLN 4 -3.630 18.778 -8.828 0.00 0.00
ATOM 64 OE1 GLN 4 -2.527 19.375 -8.680 0.00 0.00
ATOM 65 NE2 GLN 4 -4.762 19.511 -8.779 0.00 0.00
ATOM 66 HE21 GLN 4 -4.679 20.474 -8.488 0.00 0.00
ATOM 67 HE22 GLN 4 -5.611 18.963 -8.795 0.00 0.00
ATOM 68 C GLN 4 -1.803 14.211 -8.249 0.00 0.00
ATOM 69 O GLN 4 -0.726 14.705 -8.410 0.00 0.00
ATOM 70 N ASN 5 -2.020 12.916 -7.791 0.00 0.00
ATOM 71 H ASN 5 -2.970 12.574 -7.819 0.00 0.00
ATOM 72 CA ASN 5 -0.920 12.060 -7.325 0.00 0.00
ATOM 73 HA ASN 5 -0.171 12.703 -6.862 0.00 0.00
ATOM 74 CB ASN 5 -1.458 11.042 -6.242 0.00 0.00
ATOM 75 HB2 ASN 5 -0.630 10.688 -5.628 0.00 0.00
ATOM 76 HB3 ASN 5 -2.132 11.623 -5.612 0.00 0.00
ATOM 77 CG ASN 5 -2.230 9.878 -6.726 0.00 0.00
ATOM 78 OD1 ASN 5 -2.402 9.610 -7.920 0.00 0.00
ATOM 79 ND2 ASN 5 -2.675 9.086 -5.770 0.00 0.00
ATOM 80 HD21 ASN 5 -3.220 8.283 -6.051 0.00 0.00
ATOM 81 HD22 ASN 5 -2.276 9.092 -4.841 0.00 0.00
ATOM 82 C ASN 5 -0.182 11.374 -8.517 0.00 0.00
ATOM 83 O ASN 5 0.623 10.518 -8.249 0.00 0.00
ATOM 84 N CYX 6 -0.429 11.730 -9.836 0.00 0.00
ATOM 85 H CYX 6 -1.153 12.411 -10.014 0.00 0.00
ATOM 86 CA CYX 6 0.489 11.277 -10.938 0.00 0.00
ATOM 87 HA CYX 6 1.388 10.875 -10.471 0.00 0.00
ATOM 88 CB CYX 6 -0.217 10.318 -11.942 0.00 0.00
ATOM 89 HB2 CYX 6 0.492 9.774 -12.567 0.00 0.00
ATOM 90 HB3 CYX 6 -0.768 9.602 -11.332 0.00 0.00
ATOM 91 SG CYX 6 -1.383 11.066 -13.161 0.00 0.00
ATOM 92 C CYX 6 1.038 12.556 -11.675 0.00 0.00
ATOM 93 O CYX 6 0.358 13.577 -11.531 0.00 0.00
ATOM 94 N PRO 7 2.104 12.519 -12.454 0.00 0.00
ATOM 95 CD PRO 7 2.978 11.360 -12.596 0.00 0.00
ATOM 96 HD2 PRO 7 2.538 10.402 -12.875 0.00 0.00
ATOM 97 HD3 PRO 7 3.559 11.300 -11.676 0.00 0.00
ATOM 98 CG PRO 7 3.956 11.887 -13.691 0.00 0.00
ATOM 99 HG2 PRO 7 3.550 11.420 -14.588 0.00 0.00
ATOM 100 HG3 PRO 7 4.995 11.635 -13.477 0.00 0.00
ATOM 101 CB PRO 7 3.887 13.433 -13.705 0.00 0.00
ATOM 102 HB2 PRO 7 4.034 13.985 -14.634 0.00 0.00
ATOM 103 HB3 PRO 7 4.565 13.907 -12.995 0.00 0.00
ATOM 104 CA PRO 7 2.410 13.640 -13.363 0.00 0.00
ATOM 105 HA PRO 7 2.256 14.605 -12.881 0.00 0.00
ATOM 106 C PRO 7 1.636 13.638 -14.687 0.00 0.00
ATOM 107 O PRO 7 1.556 12.647 -15.395 0.00 0.00
ATOM 108 N LEU 8 0.947 14.753 -14.970 0.00 0.00
ATOM 109 H LEU 8 1.165 15.510 -14.337 0.00 0.00
ATOM 110 CA LEU 8 0.015 14.887 -16.072 0.00 0.00
ATOM 111 HA LEU 8 -0.388 13.904 -16.319 0.00 0.00
ATOM 112 CB LEU 8 -1.080 15.794 -15.461 0.00 0.00
ATOM 113 HB2 LEU 8 -0.602 16.634 -14.956 0.00 0.00
ATOM 114 HB3 LEU 8 -1.611 15.198 -14.719 0.00 0.00
ATOM 115 CG LEU 8 -2.156 16.354 -16.440 0.00 0.00
ATOM 116 HG LEU 8 -1.615 16.793 -17.279 0.00 0.00
ATOM 117 CD1 LEU 8 -3.086 15.210 -16.856 0.00 0.00
ATOM 118 HD11 LEU 8 -3.765 14.982 -16.034 0.00 0.00
ATOM 119 HD12 LEU 8 -3.769 15.584 -17.619 0.00 0.00
ATOM 120 HD13 LEU 8 -2.459 14.379 -17.176 0.00 0.00
ATOM 121 CD2 LEU 8 -2.956 17.486 -15.718 0.00 0.00
ATOM 122 HD21 LEU 8 -3.961 17.456 -16.139 0.00 0.00
ATOM 123 HD22 LEU 8 -2.984 17.230 -14.659 0.00 0.00
ATOM 124 HD23 LEU 8 -2.416 18.432 -15.761 0.00 0.00
ATOM 125 C LEU 8 0.702 15.536 -17.213 0.00 0.00
ATOM 126 O LEU 8 1.271 16.635 -17.110 0.00 0.00
ATOM 127 N GLY 9 0.619 14.981 -18.412 0.00 0.00
ATOM 128 H GLY 9 0.070 14.134 -18.416 0.00 0.00
ATOM 129 CA GLY 9 1.299 15.415 -19.590 0.00 0.00
ATOM 130 HA2 GLY 9 2.300 15.716 -19.279 0.00 0.00
ATOM 131 HA3 GLY 9 1.294 14.614 -20.329 0.00 0.00
ATOM 132 C GLY 9 0.642 16.625 -20.250 0.00 0.00
ATOM 133 O GLY 9 -0.498 17.019 -19.956 0.00 0.00
ATOM 134 N NHE 10 1.242 17.156 -21.296 0.00 0.00
ATOM 135 HN1 NHE 10 2.196 16.864 -21.453 0.00 0.00
ATOM 136 HN2 NHE 10 0.815 17.969 -21.716 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 -3.502 9.387 10.061 0.00 0.00
TER 139 Cl- 11
END

OT, clinched open (OPEN)

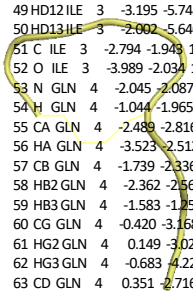
OT_MD-III_10us_T16_1

ATOM 1 N CYX 1 -9.131 -10.448 9.870 0.00 0.00
ATOM 2 H1 CYX 1 -8.320 -10.528 9.274 0.00 0.00
ATOM 3 H2 CYX 1 -8.691 -10.654 10.755 0.00 0.00
ATOM 4 H3 CYX 1 -9.364 -9.466 9.829 0.00 0.00
ATOM 5 CA CYX 1 -10.289 -11.307 9.548 0.00 0.00
ATOM 6 HA CYX 1 -10.978 -11.135 10.376 0.00 0.00
ATOM 7 CB CYX 1 -10.936 -10.957 8.205 0.00 0.00
ATOM 8 HB2 CYX 1 -11.147 -9.896 8.077 0.00 0.00
ATOM 9 HB3 CYX 1 -10.191 -11.158 7.435 0.00 0.00
ATOM 10 SG CYX 1 -12.405 -11.976 7.806 0.00 0.00
ATOM 11 C CYX 1 -9.902 -12.780 9.704 0.00 0.00
ATOM 12 O CYX 1 -8.856 -13.147 9.143 0.00 0.00
ATOM 13 N TYR 2 -10.768 -13.630 10.258 0.00 0.00
ATOM 14 H TYR 2 -11.598 -13.207 10.649 0.00 0.00
ATOM 15 CA TYR 2 -10.377 -15.043 10.507 0.00 0.00
ATOM 16 HA TYR 2 -9.396 -14.982 10.979 0.00 0.00
ATOM 17 CB TYR 2 -11.440 -15.563 11.521 0.00 0.00
ATOM 18 HB2 TYR 2 -10.873 -16.297 12.093 0.00 0.00
ATOM 19 HB3 TYR 2 -11.671 -14.828 12.292 0.00 0.00
ATOM 20 CG TYR 2 -12.840 -16.103 11.131 0.00 0.00
ATOM 21 CD1 TYR 2 -13.349 -17.224 11.870 0.00 0.00
ATOM 22 HD1 TYR 2 -12.742 -17.573 12.692 0.00 0.00
ATOM 23 CE1 TYR 2 -14.684 -17.625 11.676 0.00 0.00
ATOM 24 HE1 TYR 2 -15.032 -18.467 12.255 0.00 0.00
ATOM 25 CZ TYR 2 -15.605 -16.807 11.002 0.00 0.00
ATOM 48 HD11 ILE 3 -14.598 -15.251 4.926 0.00 0.00
ATOM 49 HD12 ILE 3 -15.356 -15.754 6.406 0.00 0.00
ATOM 50 HD13 ILE 3 -15.160 -14.055 6.121 0.00 0.00
ATOM 51 C ILE 3 -10.184 -15.610 5.732 0.00 0.00
ATOM 52 O ILE 3 -10.051 -14.390 5.693 0.00 0.00
ATOM 53 N GLN 4 -9.768 -16.389 4.735 0.00 0.00
ATOM 54 H GLN 4 -10.038 -17.342 4.767 0.00 0.00
ATOM 55 CA GLN 4 -8.938 -15.916 3.611 0.00 0.00
ATOM 56 HA GLN 4 -8.134 -15.298 4.010 0.00 0.00
ATOM 57 CB GLN 4 -8.368 -17.123 2.909 0.00 0.00
ATOM 58 HB2 GLN 4 -9.133 -17.783 2.500 0.00 0.00
ATOM 59 HB3 GLN 4 -7.714 -16.802 2.098 0.00 0.00
ATOM 60 CG GLN 4 -7.524 -17.993 3.962 0.00 0.00
ATOM 61 HG2 GLN 4 -6.750 -17.454 4.508 0.00 0.00
ATOM 62 HG3 GLN 4 -8.156 -18.337 4.781 0.00 0.00
ATOM 63 CD GLN 4 -6.763 -19.179 3.389 0.00 0.00
ATOM 64 OE1 GLN 4 -6.394 -19.329 2.200 0.00 0.00
ATOM 65 NE2 GLN 4 -6.334 -20.049 4.317 0.00 0.00
ATOM 66 HE21 GLN 4 -5.607 -20.696 4.047 0.00 0.00
ATOM 67 HE22 GLN 4 -6.652 -19.958 5.271 0.00 0.00
ATOM 68 C GLN 4 -9.700 -15.035 2.533 0.00 0.00
ATOM 69 O GLN 4 -9.093 -14.142 1.926 0.00 0.00
ATOM 70 N ASN 5 -10.984 -15.409 2.310 0.00 0.00
ATOM 71 H ASN 5 -11.388 -16.075 2.954 0.00 0.00
ATOM 72 CA ASN 5 -11.857 -14.664 1.478 0.00 0.00
ATOM 95 CD PRO 7 -16.383 -11.634 4.283 0.00 0.00
ATOM 96 HD2 PRO 7 -16.272 -11.567 5.365 0.00 0.00
ATOM 97 HD3 PRO 7 -16.287 -12.686 4.014 0.00 0.00
ATOM 98 CG PRO 7 -17.806 -11.259 3.979 0.00 0.00
ATOM 99 HG2 PRO 7 -18.434 -11.610 4.798 0.00 0.00
ATOM 100 HG3 PRO 7 -18.124 -11.662 3.017 0.00 0.00
ATOM 101 CB PRO 7 -17.654 -9.726 3.869 0.00 0.00
ATOM 102 HB2 PRO 7 -17.520 -9.355 4.884 0.00 0.00
ATOM 103 HB3 PRO 7 -18.491 -9.343 3.284 0.00 0.00
ATOM 104 CA PRO 7 -16.386 -9.574 3.018 0.00 0.00
ATOM 105 HA PRO 7 -16.657 -9.729 1.973 0.00 0.00
ATOM 106 C PRO 7 -15.767 -8.223 3.330 0.00 0.00
ATOM 107 O PRO 7 -14.914 -8.099 4.251 0.00 0.00
ATOM 108 N LEU 8 -16.192 -7.112 2.667 0.00 0.00
ATOM 109 H LEU 8 -16.950 -7.241 2.011 0.00 0.00
ATOM 110 CA LEU 8 -15.559 -5.813 2.790 0.00 0.00
ATOM 111 HA LEU 8 -14.509 -6.083 2.913 0.00 0.00
ATOM 112 CB LEU 8 -15.712 -5.009 1.489 0.00 0.00
ATOM 113 HB2 LEU 8 -16.777 -4.806 1.374 0.00 0.00
ATOM 114 HB3 LEU 8 -15.445 -5.609 0.619 0.00 0.00
ATOM 115 CG LEU 8 -15.038 -3.607 1.539 0.00 0.00
ATOM 116 HG LEU 8 -15.247 -3.046 2.450 0.00 0.00
ATOM 117 CD1 LEU 8 -13.556 -3.874 1.437 0.00 0.00
ATOM 118 HD11 LEU 8 -12.918 -2.993 1.375 0.00 0.00
ATOM 119 HD12 LEU 8 -13.252 -4.362 2.363 0.00 0.00

ATOM 26 OH TYR 2 -16.891 -17.332 10.885 0.00 0.00
ATOM 27 HH TYR 2 -17.443 -16.606 10.585 0.00 0.00
ATOM 28 CE2 TYR 2 -15.103 -15.697 10.254 0.00 0.00
ATOM 29 HE2 TYR 2 -15.773 -15.106 9.648 0.00 0.00
ATOM 30 CD2 TYR 2 -13.745 -15.308 10.411 0.00 0.00
ATOM 31 HD2 TYR 2 -13.261 -14.504 9.875 0.00 0.00
ATOM 32 C TYR 2 -10.378 -15.944 9.271 0.00 0.00
ATOM 33 O TYR 2 -10.145 -17.107 9.420 0.00 0.00
ATOM 34 N ILE 3 -10.953 -15.462 8.185 0.00 0.00
ATOM 35 H ILE 3 -11.306 -14.516 8.217 0.00 0.00
ATOM 36 CA ILE 3 -11.050 -16.270 6.816 0.00 0.00
ATOM 37 HA ILE 3 -10.687 -17.279 7.010 0.00 0.00
ATOM 38 CB ILE 3 -12.569 -16.468 6.375 0.00 0.00
ATOM 39 HB ILE 3 -12.632 -16.748 5.324 0.00 0.00
ATOM 40 CG2 ILE 3 -13.195 -17.572 7.233 0.00 0.00
ATOM 41 HG21 ILE 3 -13.036 -17.265 8.267 0.00 0.00
ATOM 42 HG22 ILE 3 -14.196 -17.873 6.923 0.00 0.00
ATOM 43 HG23 ILE 3 -12.529 -18.433 7.182 0.00 0.00
ATOM 44 CG1 ILE 3 -13.212 -15.100 6.551 0.00 0.00
ATOM 45 HG12 ILE 3 -12.549 -14.374 6.080 0.00 0.00
ATOM 46 HG13 ILE 3 -13.300 -14.796 7.594 0.00 0.00
ATOM 47 CD1 ILE 3 -14.668 -15.018 5.989 0.00 0.00
ATOM 73 HA ASN 5 -11.295 -13.884 0.964 0.00 0.00
ATOM 74 CB ASN 5 -12.421 -15.599 0.311 0.00 0.00
ATOM 75 HB2 ASN 5 -11.547 -16.135 -0.059 0.00 0.00
ATOM 76 HB3 ASN 5 -13.188 -16.304 0.632 0.00 0.00
ATOM 77 CG ASN 5 -13.021 -14.733 -0.836 0.00 0.00
ATOM 78 OD1 ASN 5 -13.454 -13.605 -0.670 0.00 0.00
ATOM 79 ND2 ASN 5 -12.989 -15.296 -1.970 0.00 0.00
ATOM 80 HD21 ASN 5 -13.220 -14.772 -2.802 0.00 0.00
ATOM 81 HD22 ASN 5 -12.642 -16.227 -2.154 0.00 0.00
ATOM 82 C ASN 5 -12.958 -13.947 2.290 0.00 0.00
ATOM 83 O ASN 5 -13.964 -14.552 2.618 0.00 0.00
ATOM 84 N CYX 6 -12.626 -12.766 2.761 0.00 0.00
ATOM 85 H CYX 6 -11.764 -12.356 2.430 0.00 0.00
ATOM 86 CA CYX 6 -13.456 -12.023 3.703 0.00 0.00
ATOM 87 HA CYX 6 -14.116 -12.759 4.161 0.00 0.00
ATOM 88 CB CYX 6 -12.578 -11.542 4.830 0.00 0.00
ATOM 89 HB2 CYX 6 -11.896 -12.345 5.111 0.00 0.00
ATOM 90 HB3 CYX 6 -11.920 -10.808 4.367 0.00 0.00
ATOM 91 SG CYX 6 -13.415 -11.064 6.338 0.00 0.00
ATOM 92 C CYX 6 -14.224 -10.939 3.050 0.00 0.00
ATOM 93 O CYX 6 -13.669 -10.195 2.218 0.00 0.00
ATOM 94 N PRO 7 -15.533 -10.698 3.467 0.00 0.00
ATOM 120 HD13 LEU 8 -13.284 -4.402 0.523 0.00 0.00
ATOM 121 CD2 LEU 8 -15.554 -2.763 0.404 0.00 0.00
ATOM 122 HD21 LEU 8 -16.605 -2.488 0.489 0.00 0.00
ATOM 123 HD22 LEU 8 -14.934 -1.868 0.343 0.00 0.00
ATOM 124 HD23 LEU 8 -15.335 -3.347 -0.489 0.00 0.00
ATOM 125 C LEU 8 -15.996 -5.002 4.045 0.00 0.00
ATOM 126 O LEU 8 -17.166 -4.767 4.276 0.00 0.00
ATOM 127 N GLY 9 -15.003 -4.432 4.790 0.00 0.00
ATOM 128 H GLY 9 -14.101 -4.768 4.486 0.00 0.00
ATOM 129 CA GLY 9 -15.187 -3.825 6.140 0.00 0.00
ATOM 130 HA2 GLY 9 -15.961 -4.365 6.685 0.00 0.00
ATOM 131 HA3 GLY 9 -14.268 -3.841 6.727 0.00 0.00
ATOM 132 C GLY 9 -15.586 -2.370 6.102 0.00 0.00
ATOM 133 O GLY 9 -15.821 -1.702 5.071 0.00 0.00
ATOM 134 N NHE 10 -15.604 -1.802 7.305 0.00 0.00
ATOM 135 HN1 NHE 10 -15.314 -2.383 8.078 0.00 0.00
ATOM 136 HN2 NHE 10 -15.943 -0.861 7.447 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 -5.589 4.480 15.967 0.00 0.00
TER 139 Cl- 11
END

OT, clinched open45pbr (OPEN)

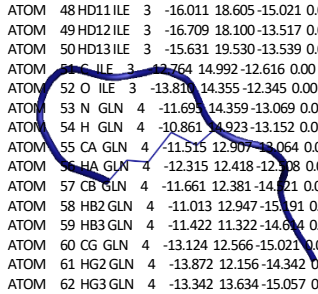
OT_MD-III_10us_T16_4



ATOM 1 N CYX 1 -8.083 -0.642 18.787 0.00 0.00
ATOM 2 H1 CYX 1 -8.766 -1.328 19.075 0.00 0.00
ATOM 3 H2 CYX 1 -8.294 -0.317 17.855 0.00 0.00
ATOM 4 H3 CYX 1 -8.114 0.199 19.347 0.00 0.00
ATOM 5 CA CYX 1 -6.664 -1.072 19.032 0.00 0.00
ATOM 6 HA CYX 1 -6.590 -1.163 20.116 0.00 0.00
ATOM 7 CB CYX 1 -6.391 -2.423 18.362 0.00 0.00
ATOM 8 HB2 CYX 1 -5.978 -2.219 17.374 0.00 0.00
ATOM 9 HB3 CYX 1 -5.630 -2.981 18.906 0.00 0.00
ATOM 10 SG CYX 1 -7.822 -3.536 18.244 0.00 0.00
ATOM 11 C CYX 1 -5.847 0.102 18.582 0.00 0.00
ATOM 12 O CYX 1 -6.089 0.702 17.589 0.00 0.00
ATOM 13 N TYR 2 -4.727 0.303 19.251 0.00 0.00
ATOM 14 H TYR 2 -4.527 -0.249 20.073 0.00 0.00
ATOM 15 CA TYR 2 -3.596 1.258 18.874 0.00 0.00
ATOM 16 HA TYR 2 -3.979 2.048 18.229 0.00 0.00
ATOM 17 CB TYR 2 -2.840 1.768 20.179 0.00 0.00
ATOM 18 HB2 TYR 2 -3.635 1.980 20.893 0.00 0.00
ATOM 19 HB3 TYR 2 -2.283 0.959 20.653 0.00 0.00
ATOM 20 CG TYR 2 -1.991 3.002 19.882 0.00 0.00
ATOM 21 CD1 TYR 2 -2.566 4.200 19.817 0.00 0.00
ATOM 22 HD1 TYR 2 -3.607 4.396 20.027 0.00 0.00
ATOM 23 CE1 TYR 2 -1.844 5.339 19.430 0.00 0.00
ATOM 24 HE1 TYR 2 -2.365 6.284 19.393 0.00 0.00
ATOM 25 CZ TYR 2 -0.433 5.269 19.139 0.00 0.00
ATOM 26 OH TYR 2 0.172 6.336 18.547 0.00 0.00
ATOM 27 HH TYR 2 1.110 6.152 18.457 0.00 0.00
ATOM 28 CE2 TYR 2 0.150 4.020 19.288 0.00 0.00
ATOM 29 HE2 TYR 2 1.158 3.903 18.919 0.00 0.00
ATOM 30 CD2 TYR 2 -0.598 2.903 19.593 0.00 0.00
ATOM 31 HD2 TYR 2 -0.224 1.910 19.388 0.00 0.00
ATOM 32 C TYR 2 -2.589 0.573 17.918 0.00 0.00
ATOM 33 O TYR 2 -1.701 1.139 17.288 0.00 0.00
ATOM 34 N ILE 3 -2.656 -0.779 17.820 0.00 0.00
ATOM 35 H ILE 3 -3.395 -1.220 18.349 0.00 0.00
ATOM 36 CA ILE 3 -1.919 -1.710 16.994 0.00 0.00
ATOM 37 HA ILE 3 -1.009 -1.176 16.719 0.00 0.00
ATOM 38 CB ILE 3 -1.553 -3.002 17.710 0.00 0.00
ATOM 39 HB ILE 3 -1.294 -3.663 16.882 0.00 0.00
ATOM 40 CG2 ILE 3 -0.243 -2.764 18.474 0.00 0.00
ATOM 41 HG21 ILE 3 -0.422 -1.884 19.093 0.00 0.00
ATOM 42 HG22 ILE 3 0.069 -3.632 19.056 0.00 0.00
ATOM 43 HG23 ILE 3 0.556 -2.595 17.752 0.00 0.00
ATOM 44 CG1 ILE 3 -2.684 -3.716 18.545 0.00 0.00
ATOM 45 HG12 ILE 3 -3.608 -3.723 17.966 0.00 0.00
ATOM 46 HG13 ILE 3 -2.759 -3.101 19.442 0.00 0.00
ATOM 47 CD1 ILE 3 -2.304 -5.178 18.870 0.00 0.00
ATOM 48 HD11 ILE 3 -1.501 -5.232 19.605 0.00 0.00
ATOM 49 HD12 ILE 3 -3.195 -5.748 19.133 0.00 0.00
ATOM 50 HD13 ILE 3 -2.002 -5.640 17.931 0.00 0.00
ATOM 51 C ILE 3 -2.794 -1.948 15.731 0.00 0.00
ATOM 52 O ILE 3 -3.989 -2.034 15.842 0.00 0.00
ATOM 53 N GLN 4 -2.045 -2.087 14.591 0.00 0.00
ATOM 54 H GLN 4 -1.041 -1.965 14.639 0.00 0.00
ATOM 55 CA GLN 4 -2.489 -2.816 13.437 0.00 0.00
ATOM 56 HA GLN 4 -3.523 -2.512 13.274 0.00 0.00
ATOM 57 CB GLN 4 -1.739 -2.336 12.204 0.00 0.00
ATOM 58 HB2 GLN 4 -2.362 -2.566 11.339 0.00 0.00
ATOM 59 HB3 GLN 4 -1.583 -1.257 12.217 0.00 0.00
ATOM 60 CG GLN 4 -0.420 -3.168 12.039 0.00 0.00
ATOM 61 HG2 GLN 4 0.149 -3.020 12.956 0.00 0.00
ATOM 62 HG3 GLN 4 -0.683 -4.226 12.024 0.00 0.00
ATOM 63 CD GLN 4 0.351 -2.716 10.777 0.00 0.00
ATOM 64 OE1 GLN 4 -0.132 -2.062 9.834 0.00 0.00
ATOM 65 NE2 GLN 4 1.585 -3.146 10.724 0.00 0.00
ATOM 66 HE21 GLN 4 2.093 -2.852 9.901 0.00 0.00
ATOM 67 HE22 GLN 4 2.058 -3.687 11.434 0.00 0.00
ATOM 68 C GLN 4 -2.569 -4.343 13.583 0.00 0.00
ATOM 69 O GLN 4 -1.811 -4.896 14.391 0.00 0.00
ATOM 70 N ASN 5 -3.416 -5.068 12.829 0.00 0.00
ATOM 71 H ASN 5 -3.846 -4.534 12.088 0.00 0.00
ATOM 72 CA ASN 5 -3.560 -6.487 12.873 0.00 0.00
ATOM 73 HA ASN 5 -4.448 -6.703 12.279 0.00 0.00
ATOM 74 CB ASN 5 -2.420 -7.208 11.991 0.00 0.00
ATOM 75 HB2 ASN 5 -2.129 -6.552 11.171 0.00 0.00
ATOM 76 HB3 ASN 5 -1.556 -7.259 12.653 0.00 0.00
ATOM 77 CG ASN 5 -2.813 -8.603 11.562 0.00 0.00
ATOM 78 OD1 ASN 5 -3.928 -9.096 11.696 0.00 0.00
ATOM 79 ND2 ASN 5 -1.881 -9.278 11.060 0.00 0.00
ATOM 80 HD21 ASN 5 -2.110 -10.222 10.784 0.00 0.00
ATOM 81 HD22 ASN 5 -0.956 -8.886 10.958 0.00 0.00
ATOM 82 C ASN 5 -3.902 -7.033 14.241 0.00 0.00
ATOM 83 O ASN 5 -3.335 -7.948 14.875 0.00 0.00
ATOM 84 N CYX 6 -4.899 -6.373 14.781 0.00 0.00
ATOM 85 H CYX 6 -5.310 -5.617 14.252 0.00 0.00
ATOM 86 CA CYX 6 -5.491 -6.660 16.071 0.00 0.00
ATOM 87 HA CYX 6 -4.685 -6.622 16.803 0.00 0.00
ATOM 88 CB CYX 6 -6.392 -5.477 16.416 0.00 0.00
ATOM 89 HB2 CYX 6 -5.785 -4.582 16.274 0.00 0.00
ATOM 90 HB3 CYX 6 -7.299 -5.404 15.816 0.00 0.00
ATOM 91 SG CYX 6 -6.943 -5.419 18.171 0.00 0.00
ATOM 92 C CYX 6 -6.299 -7.941 16.209 0.00 0.00
ATOM 93 O CYX 6 -6.956 -8.319 15.244 0.00 0.00
ATOM 94 N PRO 7 -6.087 -8.748 17.297 0.00 0.00
ATOM 95 CD PRO 7 -5.239 -8.662 18.511 0.00 0.00
ATOM 96 HD2 PRO 7 -5.470 -7.765 19.086 0.00 0.00
ATOM 97 HD3 PRO 7 -4.202 -8.551 18.195 0.00 0.00
ATOM 98 CG PRO 7 -5.439 -9.843 19.386 0.00 0.00
ATOM 99 HG2 PRO 7 -5.947 -9.531 20.298 0.00 0.00
ATOM 100 HG3 PRO 7 -4.472 -10.202 19.737 0.00 0.00
ATOM 101 CB PRO 7 -6.352 -10.761 18.648 0.00 0.00
ATOM 102 HB2 PRO 7 -7.215 -11.049 19.248 0.00 0.00
ATOM 103 HB3 PRO 7 -5.837 -11.628 18.234 0.00 0.00
ATOM 104 CA PRO 7 -6.872 -9.949 17.450 0.00 0.00
ATOM 105 HA PRO 7 -6.613 -10.406 16.495 0.00 0.00
ATOM 106 C PRO 7 -8.282 -9.536 17.554 0.00 0.00
ATOM 107 O PRO 7 -8.543 -8.532 18.206 0.00 0.00
ATOM 108 N LEU 8 -9.197 -10.325 16.984 0.00 0.00
ATOM 109 H LEU 8 -8.843 -10.987 16.309 0.00 0.00
ATOM 110 CA LEU 8 -10.627 -10.150 17.134 0.00 0.00
ATOM 111 HA LEU 8 -10.923 -9.133 17.389 0.00 0.00
ATOM 112 CB LEU 8 -11.359 -10.448 15.759 0.00 0.00
ATOM 113 HB2 LEU 8 -12.437 -10.438 15.921 0.00 0.00
ATOM 114 HB3 LEU 8 -11.091 -11.408 15.317 0.00 0.00
ATOM 115 CG LEU 8 -10.995 -9.359 14.678 0.00 0.00
ATOM 116 HG LEU 8 -9.930 -9.389 14.450 0.00 0.00
ATOM 117 CD1 LEU 8 -11.850 -9.552 13.409 0.00 0.00
ATOM 118 HD11 LEU 8 -11.384 -9.073 12.548 0.00 0.00
ATOM 119 HD12 LEU 8 -11.814 -10.605 13.130 0.00 0.00
ATOM 120 HD13 LEU 8 -12.924 -9.439 13.558 0.00 0.00
ATOM 121 CD2 LEU 8 -11.381 -8.008 15.145 0.00 0.00
ATOM 122 HD21 LEU 8 -11.460 -7.368 14.266 0.00 0.00
ATOM 123 HD22 LEU 8 -12.320 -8.009 15.700 0.00 0.00
ATOM 124 HD23 LEU 8 -10.639 -7.580 15.820 0.00 0.00
ATOM 125 C LEU 8 -11.170 -11.090 18.238 0.00 0.00
ATOM 126 O LEU 8 -10.729 -12.264 18.280 0.00 0.00
ATOM 127 N GLY 9 -12.147 -10.612 18.992 0.00 0.00
ATOM 128 H GLY 9 -12.493 -9.677 18.832 0.00 0.00
ATOM 129 CA GLY 9 -12.883 -11.417 19.926 0.00 0.00
ATOM 130 HA2 GLY 9 -12.173 -12.124 20.354 0.00 0.00
ATOM 131 HA3 GLY 9 -13.141 -10.846 20.818 0.00 0.00
ATOM 132 C GLY 9 -13.981 -12.235 19.305 0.00 0.00
ATOM 133 O GLY 9 -14.224 -12.269 18.100 0.00 0.00
ATOM 134 N NHE 10 -14.695 -12.969 20.161 0.00 0.00
ATOM 135 HN1 NHE 10 -14.504 -12.831 21.144 0.00 0.00
ATOM 136 HN2 NHE 10 -15.494 -13.458 19.782 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 -14.487 0.304 -4.128 0.00 0.00
TER 139 Cl- 11
END

OT, open (OPEN)

OT_MD-II_15us_T16_23



ATOM 1 N CYX 1 -6.682 15.205 -17.325 0.00 0.00
ATOM 2 H1 CYX 1 -6.473 15.211 -18.313 0.00 0.00
ATOM 3 H2 CYX 1 -6.240 14.362 -16.984 0.00 0.00
ATOM 4 H3 CYX 1 -6.250 16.006 -16.887 0.00 0.00
ATOM 5 CA CYX 1 -8.127 15.182 -16.973 0.00 0.00
ATOM 6 HA CYX 1 -8.773 15.319 -17.840 0.00 0.00
ATOM 7 CB CYX 1 -8.370 13.927 -16.157 0.00 0.00
ATOM 8 HB2 CYX 1 -7.576 13.804 -15.420 0.00 0.00
ATOM 9 HB3 CYX 1 -9.293 13.962 -15.578 0.00 0.00
ATOM 10 SG CYX 1 -8.413 12.426 -17.164 0.00 0.00
ATOM 11 C CYX 1 -8.367 16.295 -15.962 0.00 0.00
ATOM 12 O CYX 1 -7.669 16.353 -15.013 0.00 0.00
ATOM 13 N TYR 2 -9.596 16.851 -16.065 0.00 0.00
ATOM 14 H TYR 2 -10.150 16.601 -16.872 0.00 0.00
ATOM 15 CA TYR 2 -10.198 17.697 -15.001 0.00 0.00
ATOM 16 HA TYR 2 -9.491 17.905 -14.198 0.00 0.00
ATOM 17 CB TYR 2 -10.327 19.079 -15.587 0.00 0.00
ATOM 48 HD11 ILE 3 -16.011 18.605 -15.021 0.00 0.00
ATOM 49 HD12 ILE 3 -16.709 18.100 -13.517 0.00 0.00
ATOM 50 HD13 ILE 3 -15.631 19.530 -13.539 0.00 0.00
ATOM 51 C ILE 3 -13.764 14.992 -12.616 0.00 0.00
ATOM 52 O ILE 3 -13.810 14.355 -12.345 0.00 0.00
ATOM 53 N GLN 4 -11.695 14.359 -13.069 0.00 0.00
ATOM 54 H GLN 4 -10.861 14.923 -13.152 0.00 0.00
ATOM 55 CA GLN 4 -11.515 12.907 -13.064 0.00 0.00
ATOM 56 HA GLN 4 -12.315 12.418 -12.508 0.00 0.00
ATOM 57 CB GLN 4 -11.661 12.381 -14.521 0.00 0.00
ATOM 58 HB2 GLN 4 -11.013 12.947 -15.191 0.00 0.00
ATOM 59 HB3 GLN 4 -11.422 11.322 -14.644 0.00 0.00
ATOM 60 CG GLN 4 -13.124 12.566 -15.021 0.00 0.00
ATOM 61 HG2 GLN 4 -13.872 12.156 -14.342 0.00 0.00
ATOM 62 HG3 GLN 4 -13.342 13.634 -15.057 0.00 0.00
ATOM 63 CD GLN 4 -13.296 12.096 -16.431 0.00 0.00
ATOM 64 OE1 GLN 4 -13.358 12.867 -17.402 0.00 0.00
ATOM 96 HD2 PRO 7 -5.697 7.899 -16.509 0.00 0.00
ATOM 95 CD PRO 7 -5.359 7.267 -15.688 0.00 0.00
ATOM 97 HD3 PRO 7 -6.228 6.763 -15.264 0.00 0.00
ATOM 98 CG PRO 7 -4.279 6.313 -16.280 0.00 0.00
ATOM 99 HG2 PRO 7 -3.869 6.728 -17.200 0.00 0.00
ATOM 100 HG3 PRO 7 -4.707 5.347 -16.545 0.00 0.00
ATOM 101 CB PRO 7 -3.204 6.220 -15.224 0.00 0.00
ATOM 102 HB2 PRO 7 -2.180 6.227 -15.597 0.00 0.00
ATOM 103 HB3 PRO 7 -3.254 5.279 -14.677 0.00 0.00
ATOM 104 CA PRO 7 -3.352 7.420 -14.305 0.00 0.00
ATOM 105 HA PRO 7 -3.322 7.075 -13.272 0.00 0.00
ATOM 106 C PRO 7 -2.198 8.387 -14.507 0.00 0.00
ATOM 107 O PRO 7 -2.220 9.133 -15.467 0.00 0.00
ATOM 108 N LEU 8 -1.164 8.411 -13.649 0.00 0.00
ATOM 109 H LEU 8 -1.244 7.678 -12.959 0.00 0.00
ATOM 110 CA LEU 8 0.133 9.056 -13.703 0.00 0.00
ATOM 111 HA LEU 8 -0.085 10.045 -14.105 0.00 0.00

ATOM 18 HB2 TYR 2 -9.396 19.473 -15.994 0.00 0.00
ATOM 19 HB3 TYR 2 -10.900 18.902 -16.497 0.00 0.00
ATOM 20 CG TYR 2 -11.039 20.108 -14.639 0.00 0.00
ATOM 21 CD1 TYR 2 -12.327 20.559 -14.858 0.00 0.00
ATOM 22 HD1 TYR 2 -12.909 20.130 -15.660 0.00 0.00
ATOM 23 CE1 TYR 2 -12.958 21.452 -13.965 0.00 0.00
ATOM 24 HE1 TYR 2 -13.953 21.834 -14.138 0.00 0.00
ATOM 25 CZ TYR 2 -12.125 22.034 -12.983 0.00 0.00
ATOM 26 OH TYR 2 -12.491 23.161 -12.359 0.00 0.00
ATOM 27 HH TYR 2 -13.333 23.524 -12.642 0.00 0.00
ATOM 28 CE2 TYR 2 -10.857 21.597 -12.732 0.00 0.00
ATOM 29 HE2 TYR 2 -10.257 22.119 -12.001 0.00 0.00
ATOM 30 CD2 TYR 2 -10.305 20.650 -13.620 0.00 0.00
ATOM 31 HD2 TYR 2 -9.294 20.319 -13.438 0.00 0.00
ATOM 32 C TYR 2 -11.449 17.014 -14.449 0.00 0.00
ATOM 33 O TYR 2 -12.196 16.304 -15.131 0.00 0.00
ATOM 34 N ILE 3 -11.622 17.201 -13.101 0.00 0.00
ATOM 35 H ILE 3 -11.008 17.877 -12.670 0.00 0.00
ATOM 36 CA ILE 3 -12.740 16.533 -12.356 0.00 0.00
ATOM 37 HA ILE 3 -12.470 16.612 -11.302 0.00 0.00
ATOM 38 CB ILE 3 -14.048 17.311 -12.391 0.00 0.00
ATOM 39 HB ILE 3 -14.878 16.747 -11.965 0.00 0.00
ATOM 40 CG2 ILE 3 -13.846 18.521 -11.542 0.00 0.00
ATOM 41 HG21 ILE 3 -14.866 18.874 -11.391 0.00 0.00
ATOM 42 HG22 ILE 3 -13.408 18.225 -10.588 0.00 0.00
ATOM 43 HG23 ILE 3 -13.269 19.304 -12.035 0.00 0.00
ATOM 44 CG1 ILE 3 -14.563 17.636 -13.830 0.00 0.00
ATOM 45 HG12 ILE 3 -13.830 18.259 -14.342 0.00 0.00
ATOM 46 HG13 ILE 3 -14.790 16.751 -14.425 0.00 0.00
ATOM 47 CD1 ILE 3 -15.810 18.537 -13.952 0.00 0.00
ATOM 65 NE2 GLN 4 -13.346 10.797 -16.625 0.00 0.00
ATOM 66 HE21 GLN 4 -13.408 10.192 -15.819 0.00 0.00
ATOM 67 HE22 GLN 4 -13.461 10.458 -17.570 0.00 0.00
ATOM 68 C GLN 4 -10.127 12.501 -12.533 0.00 0.00
ATOM 69 O GLN 4 -9.125 13.056 -12.933 0.00 0.00
ATOM 70 N ASN 5 -10.159 11.403 -11.730 0.00 0.00
ATOM 71 H ASN 5 -11.061 11.063 -11.429 0.00 0.00
ATOM 72 CA ASN 5 -8.985 10.638 -11.357 0.00 0.00
ATOM 73 HA ASN 5 -8.167 11.324 -11.136 0.00 0.00
ATOM 74 CB ASN 5 -9.263 9.896 -10.112 0.00 0.00
ATOM 75 HB2 ASN 5 -10.025 10.379 -9.501 0.00 0.00
ATOM 76 HB3 ASN 5 -9.749 8.967 -10.413 0.00 0.00
ATOM 77 CG ASN 5 -8.074 9.601 -9.311 0.00 0.00
ATOM 78 OD1 ASN 5 -7.001 9.811 -9.700 0.00 0.00
ATOM 79 ND2 ASN 5 -8.182 8.953 -8.224 0.00 0.00
ATOM 80 HD21 ASN 5 -7.364 8.674 -7.702 0.00 0.00
ATOM 81 HD22 ASN 5 -9.053 8.741 -7.759 0.00 0.00
ATOM 82 C ASN 5 -8.442 9.830 -12.496 0.00 0.00
ATOM 83 O ASN 5 -9.176 8.975 -12.984 0.00 0.00
ATOM 84 N CYX 6 -7.234 10.040 -12.964 0.00 0.00
ATOM 85 H CYX 6 -6.614 10.669 -12.474 0.00 0.00
ATOM 86 CA CYX 6 -6.674 9.476 -14.207 0.00 0.00
ATOM 87 HA CYX 6 -7.223 8.612 -14.582 0.00 0.00
ATOM 88 CB CYX 6 -6.893 10.496 -15.290 0.00 0.00
ATOM 89 HB2 CYX 6 -6.241 11.354 -15.129 0.00 0.00
ATOM 90 HB3 CYX 6 -6.419 10.107 -16.191 0.00 0.00
ATOM 91 SG CYX 6 -8.610 10.995 -15.755 0.00 0.00
ATOM 92 C CYX 6 -5.225 9.004 -13.965 0.00 0.00
ATOM 93 O CYX 6 -4.603 9.659 -13.170 0.00 0.00
ATOM 94 N PRO 7 -4.719 7.955 -14.598 0.00 0.00
ATOM 112 CB LEU 8 0.818 9.134 -12.341 0.00 0.00
ATOM 113 HB2 LEU 8 0.328 9.885 -11.722 0.00 0.00
ATOM 114 HB3 LEU 8 0.633 8.141 -11.930 0.00 0.00
ATOM 115 CG LEU 8 2.314 9.352 -12.250 0.00 0.00
ATOM 116 HG LEU 8 2.779 8.546 -12.818 0.00 0.00
ATOM 117 CD1 LEU 8 2.659 10.796 -12.627 0.00 0.00
ATOM 118 HD11 LEU 8 3.696 11.026 -12.378 0.00 0.00
ATOM 119 HD12 LEU 8 2.502 10.886 -13.702 0.00 0.00
ATOM 120 HD13 LEU 8 2.151 11.504 -11.972 0.00 0.00
ATOM 121 CD2 LEU 8 2.807 9.279 -10.847 0.00 0.00
ATOM 122 HD21 LEU 8 3.895 9.222 -10.860 0.00 0.00
ATOM 123 HD22 LEU 8 2.319 10.079 -10.291 0.00 0.00
ATOM 124 HD23 LEU 8 2.522 8.400 -10.269 0.00 0.00
ATOM 125 C LEU 8 1.077 8.301 -14.690 0.00 0.00
ATOM 126 O LEU 8 1.345 7.109 -14.533 0.00 0.00
ATOM 127 N GLY 9 1.689 9.022 -15.646 0.00 0.00
ATOM 128 H GLY 9 1.462 9.999 -15.772 0.00 0.00
ATOM 129 CA GLY 9 2.665 8.442 -16.513 0.00 0.00
ATOM 130 HA2 GLY 9 2.355 7.405 -16.634 0.00 0.00
ATOM 131 HA3 GLY 9 2.786 8.884 -17.503 0.00 0.00
ATOM 132 C GLY 9 4.127 8.459 -15.984 0.00 0.00
ATOM 133 O GLY 9 4.447 9.198 -15.048 0.00 0.00
ATOM 134 N NHE 10 5.010 7.794 -16.760 0.00 0.00
ATOM 135 HN1 NHE 10 4.672 7.179 -17.486 0.00 0.00
ATOM 136 HN2 NHE 10 5.989 7.891 -16.532 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 4.874 -10.098 -16.650 0.00 0.00
TER 139 Cl- 11
END

OT, open23pbr (OPEN)

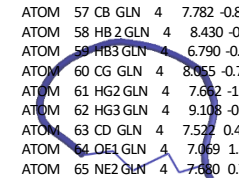
OT_MD-I_15us_T16_17

ATOM 1 N CYX 1 0.266 -11.768 8.765 0.00 0.00
ATOM 2 H1 CYX 1 1.030 -12.072 8.178 0.00 0.00
ATOM 3 H2 CYX 1 -0.207 -10.909 8.526 0.00 0.00
ATOM 4 H3 CYX 1 -0.563 -12.345 8.760 0.00 0.00
ATOM 5 CA CYX 1 0.835 -11.710 10.109 0.00 0.00
ATOM 6 HA CYX 1 0.016 -11.619 10.823 0.00 0.00
ATOM 7 CB CYX 1 1.573 -13.050 10.494 0.00 0.00
ATOM 8 HB2 CYX 1 2.502 -13.145 9.933 0.00 0.00
ATOM 9 HB3 CYX 1 1.914 -13.067 11.529 0.00 0.00
ATOM 10 SG CYX 1 0.662 -14.628 10.426 0.00 0.00
ATOM 11 C CYX 1 1.790 -10.558 10.235 0.00 0.00
ATOM 12 O CYX 1 2.943 -10.688 9.906 0.00 0.00
ATOM 13 N TYR 2 1.236 -9.445 10.755 0.00 0.00
ATOM 14 H TYR 2 0.303 -9.457 11.142 0.00 0.00
ATOM 15 CA TYR 2 1.926 -8.193 10.702 0.00 0.00
ATOM 16 HA TYR 2 2.529 -8.259 9.797 0.00 0.00
ATOM 17 CB TYR 2 0.911 -7.073 10.321 0.00 0.00
ATOM 18 HB2 TYR 2 0.769 -7.006 9.243 0.00 0.00
ATOM 19 HB3 TYR 2 -0.110 -7.379 10.548 0.00 0.00
ATOM 20 CG TYR 2 1.124 -5.650 10.900 0.00 0.00
ATOM 21 CD1 TYR 2 2.268 -4.923 10.334 0.00 0.00
ATOM 22 HD1 TYR 2 2.779 -5.379 9.498 0.00 0.00
ATOM 23 CE1 TYR 2 2.603 -3.652 10.885 0.00 0.00
ATOM 24 HE1 TYR 2 3.414 -3.049 10.504 0.00 0.00
ATOM 25 CZ TYR 2 1.941 -3.232 12.066 0.00 0.00
ATOM 26 OH TYR 2 2.236 -2.048 12.572 0.00 0.00
ATOM 27 HH TYR 2 3.019 -1.632 12.205 0.00 0.00
ATOM 28 CE2 TYR 2 0.878 -3.985 12.628 0.00 0.00
ATOM 29 HE2 TYR 2 0.377 -3.651 13.524 0.00 0.00
ATOM 30 CD2 TYR 2 0.430 -5.185 12.023 0.00 0.00
ATOM 31 HD2 TYR 2 -0.414 -5.738 12.408 0.00 0.00
ATOM 32 C TYR 2 2.813 -7.901 11.941 0.00 0.00
ATOM 33 O TYR 2 3.794 -7.234 11.761 0.00 0.00
ATOM 34 N ILE 3 2.394 -8.380 13.117 0.00 0.00
ATOM 35 H ILE 3 1.486 -8.820 13.055 0.00 0.00
ATOM 36 CA ILE 3 3.033 -8.333 14.438 0.00 0.00
ATOM 37 HA ILE 3 3.788 -7.550 14.503 0.00 0.00
ATOM 38 CB ILE 3 1.881 -8.024 15.457 0.00 0.00
ATOM 39 HB ILE 3 2.375 -7.850 16.412 0.00 0.00
ATOM 40 CG2 ILE 3 1.067 -6.689 15.151 0.00 0.00
ATOM 41 HG21 ILE 3 0.444 -6.771 14.261 0.00 0.00
ATOM 42 HG22 ILE 3 0.264 -6.611 15.884 0.00 0.00
ATOM 43 HG23 ILE 3 1.753 -5.867 15.356 0.00 0.00
ATOM 44 CG1 ILE 3 0.946 -9.242 15.615 0.00 0.00
ATOM 45 HG12 ILE 3 0.393 -9.446 14.697 0.00 0.00
ATOM 46 HG13 ILE 3 1.410 -10.199 15.853 0.00 0.00
ATOM 47 CD1 ILE 3 -0.245 -9.082 16.605 0.00 0.00
ATOM 48 HD11 ILE 3 0.192 -9.083 17.604 0.00 0.00
ATOM 49 HD12 ILE 3 -0.860 -8.205 16.404 0.00 0.00
ATOM 50 HD13 ILE 3 -1.020 -9.836 16.472 0.00 0.00
ATOM 51 C ILE 3 3.773 -9.660 14.770 0.00 0.00
ATOM 52 O ILE 3 4.657 -9.760 15.644 0.00 0.00
ATOM 53 N GLN 4 3.542 -10.743 13.969 0.00 0.00
ATOM 54 H GLN 4 2.837 -10.781 13.247 0.00 0.00
ATOM 55 CA GLN 4 4.187 -12.030 14.211 0.00 0.00
ATOM 56 HA GLN 4 5.011 -11.877 14.908 0.00 0.00
ATOM 57 CB GLN 4 3.043 -12.719 15.033 0.00 0.00
ATOM 58 HB2 GLN 4 3.527 -13.545 15.553 0.00 0.00
ATOM 59 HB3 GLN 4 2.673 -12.033 15.795 0.00 0.00
ATOM 60 CG GLN 4 1.876 -13.329 14.220 0.00 0.00
ATOM 61 HG2 GLN 4 2.212 -13.864 13.332 0.00 0.00
ATOM 62 HG3 GLN 4 1.495 -14.157 14.817 0.00 0.00
ATOM 63 CD GLN 4 0.726 -12.355 13.852 0.00 0.00
ATOM 64 OE1 GLN 4 0.903 -11.350 13.102 0.00 0.00
ATOM 65 NE2 GLN 4 -0.459 -12.598 14.347 0.00 0.00
ATOM 66 HE21 GLN 4 -1.137 -11.896 14.087 0.00 0.00
ATOM 67 HE22 GLN 4 -0.750 -13.409 14.874 0.00 0.00
ATOM 68 C GLN 4 4.779 -12.763 12.979 0.00 0.00
ATOM 69 O GLN 4 4.379 -12.550 11.830 0.00 0.00
ATOM 70 N ASN 5 5.798 -13.637 13.175 0.00 0.00
ATOM 71 H ASN 5 6.205 -13.771 14.090 0.00 0.00
ATOM 72 CA ASN 5 6.449 -14.383 12.077 0.00 0.00
ATOM 73 HA ASN 5 6.534 -13.641 11.283 0.00 0.00
ATOM 74 CB ASN 5 7.832 -14.710 12.529 0.00 0.00
ATOM 75 HB2 ASN 5 8.365 -13.806 12.826 0.00 0.00
ATOM 76 HB3 ASN 5 7.789 -15.328 13.425 0.00 0.00
ATOM 77 CG ASN 5 8.691 -15.381 11.435 0.00 0.00
ATOM 78 OD1 ASN 5 8.285 -15.612 10.285 0.00 0.00
ATOM 79 ND2 ASN 5 9.890 -15.685 11.756 0.00 0.00
ATOM 80 HD21 ASN 5 10.539 -16.090 11.097 0.00 0.00
ATOM 81 HD22 ASN 5 10.173 -15.592 12.721 0.00 0.00
ATOM 82 C ASN 5 5.622 -15.531 11.553 0.00 0.00
ATOM 83 O ASN 5 5.421 -16.525 12.334 0.00 0.00
ATOM 84 N CYX 6 5.261 -15.467 10.308 0.00 0.00
ATOM 85 H CYX 6 5.409 -14.625 9.770 0.00 0.00
ATOM 86 CA CYX 6 4.493 -16.548 9.668 0.00 0.00
ATOM 87 HA CYX 6 4.341 -17.436 10.282 0.00 0.00
ATOM 88 CB CYX 6 3.014 -16.146 9.508 0.00 0.00
ATOM 89 HB2 CYX 6 2.957 -15.208 8.955 0.00 0.00
ATOM 90 HB3 CYX 6 2.520 -16.990 9.027 0.00 0.00
ATOM 91 SG CYX 6 2.112 -16.029 10.981 0.00 0.00
ATOM 92 C CYX 6 5.142 -17.021 8.314 0.00 0.00
ATOM 93 O CYX 6 5.478 -16.119 7.561 0.00 0.00
ATOM 94 N PRO 7 5.203 -18.363 7.966 0.00 0.00
ATOM 95 CD PRO 7 4.825 -19.497 8.782 0.00 0.00
ATOM 96 HD2 PRO 7 3.739 -19.435 8.840 0.00 0.00
ATOM 97 HD3 PRO 7 5.157 -19.318 9.805 0.00 0.00
ATOM 98 CG PRO 7 5.511 -20.739 8.176 0.00 0.00
ATOM 99 HG2 PRO 7 4.884 -21.624 8.277 0.00 0.00
ATOM 100 HG3 PRO 7 6.430 -20.973 8.714 0.00 0.00
ATOM 101 CB PRO 7 5.681 -20.391 6.719 0.00 0.00
ATOM 102 HB2 PRO 7 4.901 -20.861 6.119 0.00 0.00
ATOM 103 HB3 PRO 7 6.641 -20.791 6.393 0.00 0.00
ATOM 104 CA PRO 7 5.645 -18.898 6.670 0.00 0.00
ATOM 105 HA PRO 7 6.672 -18.563 6.519 0.00 0.00
ATOM 106 C PRO 7 4.745 -18.341 5.554 0.00 0.00
ATOM 107 O PRO 7 3.642 -18.030 5.849 0.00 0.00
ATOM 108 N LEU 8 5.220 -18.299 4.362 0.00 0.00
ATOM 109 H LEU 8 6.209 -18.467 4.246 0.00 0.00
ATOM 110 CA LEU 8 4.492 -17.852 3.170 0.00 0.00
ATOM 111 HA LEU 8 3.998 -16.918 3.438 0.00 0.00
ATOM 112 CB LEU 8 5.446 -17.525 2.027 0.00 0.00
ATOM 113 HB2 LEU 8 6.146 -18.356 1.943 0.00 0.00
ATOM 114 HB3 LEU 8 4.921 -17.595 1.074 0.00 0.00
ATOM 115 CG LEU 8 6.301 -16.213 2.048 0.00 0.00
ATOM 116 HG LEU 8 6.882 -16.150 2.968 0.00 0.00
ATOM 117 CD1 LEU 8 7.223 -16.120 0.798 0.00 0.00
ATOM 118 HD11 LEU 8 6.594 -16.101 -0.092 0.00 0.00
ATOM 119 HD12 LEU 8 7.797 -15.198 0.889 0.00 0.00
ATOM 120 HD13 LEU 8 7.917 -16.961 0.789 0.00 0.00
ATOM 121 CD2 LEU 8 5.397 -14.949 2.110 0.00 0.00
ATOM 122 HD21 LEU 8 6.052 -14.107 1.885 0.00 0.00
ATOM 123 HD22 LEU 8 4.632 -15.089 1.347 0.00 0.00
ATOM 124 HD23 LEU 8 4.905 -14.877 3.080 0.00 0.00
ATOM 125 C LEU 8 3.415 -18.916 2.781 0.00 0.00
ATOM 126 O LEU 8 3.751 -20.068 2.520 0.00 0.00
ATOM 127 N GLY 9 2.133 -18.452 2.558 0.00 0.00
ATOM 128 H GLY 9 1.994 -17.454 2.634 0.00 0.00
ATOM 129 CA GLY 9 1.125 -19.317 1.895 0.00 0.00
ATOM 130 HA2 GLY 9 1.282 -20.378 2.090 0.00 0.00
ATOM 131 HA3 GLY 9 0.129 -19.156 2.306 0.00 0.00
ATOM 132 C GLY 9 1.262 -19.235 0.387 0.00 0.00
ATOM 133 O GLY 9 2.173 -18.585 -0.182 0.00 0.00
ATOM 134 N NHE 10 0.420 -20.050 -0.252 0.00 0.00
ATOM 135 HN1 NHE 10 -0.056 -20.801 0.228 0.00 0.00
ATOM 136 HN2 NHE 10 0.686 -20.161 -1.220 0.00 0.00
TER 137 NHE 10
ATOM 138 Cl- Cl- 11 0.014 14.126 -6.961 0.00 0.00
TER 139 Cl- 11
END

OT, open233-4pbr (OPEN)*

OT_MD-III_10us_T16_3

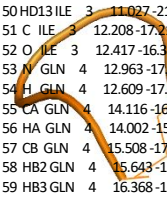
ATOM 1 N CYX 1 13.757 -0.647 -12.196 0.00 0.00
ATOM 2 H1 CYX 1 14.738 -0.885 -12.163 0.00 0.00
ATOM 3 H2 CYX 1 13.680 0.352 -12.318 0.00 0.00
ATOM 4 H3 CYX 1 13.256 -1.126 -12.931 0.00 0.00
ATOM 5 CA CYX 1 13.205 -0.860 -10.844 0.00 0.00
ATOM 6 HA CYX 1 13.704 -0.064 -10.291 0.00 0.00
ATOM 7 CB CYX 1 11.669 -0.738 -10.961 0.00 0.00
ATOM 8 HB2 CYX 1 11.281 -1.320 -11.798 0.00 0.00
ATOM 9 HB3 CYX 1 11.127 -1.118 -10.095 0.00 0.00
ATOM 48 HD11 ILE 3 12.761 -2.034 -2.252 0.00 0.00
ATOM 49 HD12 ILE 3 12.256 -0.369 -2.678 0.00 0.00
ATOM 50 HD13 ILE 3 11.071 -1.646 -2.555 0.00 0.00
ATOM 51 C ILE 3 10.399 -2.422 -6.743 0.00 0.00
ATOM 52 O ILE 3 10.639 -1.297 -7.224 0.00 0.00
ATOM 53 N GLN 4 9.194 -2.872 -6.445 0.00 0.00
ATOM 54 H GLN 4 9.156 -3.840 -6.160 0.00 0.00
ATOM 55 CA GLN 4 7.896 -2.315 -6.865 0.00 0.00
ATOM 56 HA GLN 4 7.205 -2.963 -6.326 0.00 0.00
ATOM 95 CD PRO 7 6.505 1.368 -15.184 0.00 0.00
ATOM 96 HD2 PRO 7 7.490 1.818 -15.061 0.00 0.00
ATOM 97 HD3 PRO 7 5.900 1.678 -14.332 0.00 0.00
ATOM 98 CG PRO 7 5.867 1.842 -16.488 0.00 0.00
ATOM 99 HG2 PRO 7 6.636 2.231 -17.156 0.00 0.00
ATOM 100 HG3 PRO 7 5.257 2.732 -16.330 0.00 0.00
ATOM 101 CB PRO 7 5.213 0.714 -17.173 0.00 0.00
ATOM 102 HB2 PRO 7 5.436 0.730 -18.240 0.00 0.00
ATOM 103 HB3 PRO 7 4.128 0.770 -17.082 0.00 0.00



```
ATOM  10 SG CYX  1    11.026  0.804 -11.485 0.00 0.00
ATOM  11 C  CYX  1    13.592 -2.220 -10.423 0.00 0.00
ATOM  12 O  CYX  1    13.221 -3.129 -11.044 0.00 0.00
ATOM  13 N  TYR  2    14.202 -2.420  -9.205 0.00 0.00
ATOM  14 H  TYR  2    14.419 -1.650  -8.589 0.00 0.00
ATOM  15 CA TYR  2    14.638 -3.746  -8.722 0.00 0.00
ATOM  16 HA TYR  2    14.540 -4.486  -9.515 0.00 0.00
ATOM  17 CB TYR  2    16.101 -3.827  -8.331 0.00 0.00
ATOM  18 HB2TYR  2    16.389 -4.822  -7.991 0.00 0.00
ATOM  19 HB3TYR  2    16.661 -3.486  -9.202 0.00 0.00
ATOM  20 CG TYR  2    16.430 -2.850  -7.148 0.00 0.00
ATOM  21 CD1TYR  2    16.337 -3.336  -5.839 0.00 0.00
ATOM  22 HD1TYR  2    15.964 -4.331  -5.646 0.00 0.00
ATOM  23 CE1TYR  2    16.428 -2.456  -4.739 0.00 0.00
ATOM  24 HE1TYR  2    16.206 -2.801  -3.740 0.00 0.00
ATOM  25 CZ TYR  2    16.744 -1.087  -4.946 0.00 0.00
ATOM  26 OH TYR  2    16.856 -0.266  -3.825 0.00 0.00
ATOM  27 HH TYR  2    16.628 -0.760  -3.033 0.00 0.00
ATOM  28 CE2TYR  2    16.988 -0.601  -6.287 0.00 0.00
ATOM  29 HE2TYR  2    17.230  0.432  -6.490 0.00 0.00
ATOM  30 CD2TYR  2    16.701 -1.479  -7.378 0.00 0.00
ATOM  31 HD2TYR  2    16.607 -1.044  -8.362 0.00 0.00
ATOM  32 C  TYR  2    13.631 -4.158  -7.616 0.00 0.00
ATOM  33 O  TYR  2    13.619 -5.274  -7.149 0.00 0.00
ATOM  34 N  ILE  3    12.728 -3.288  -7.134 0.00 0.00
ATOM  35 H  ILE  3    12.701 -2.375  -7.565 0.00 0.00
ATOM  36 CA ILE  3    11.521 -3.401  -6.297 0.00 0.00
ATOM  37 HA ILE  3    11.124 -4.405  -6.445 0.00 0.00
ATOM  38 CB ILE  3    11.835 -3.240  -4.775 0.00 0.00
ATOM  39 HB ILE  3    10.943 -3.436  -4.179 0.00 0.00
ATOM  40 CG2ILE  3    13.004 -4.023  -4.297 0.00 0.00
ATOM  41 HG21ILE  3   13.944 -3.762  -4.784 0.00 0.00
ATOM  42 HG22ILE  3   13.018 -4.127  -3.212 0.00 0.00
ATOM  43 HG23ILE  3   12.805 -5.044  -4.623 0.00 0.00
ATOM  44 CG1ILE  3    12.095 -1.760  -4.366 0.00 0.00
ATOM  45 HG12ILE  3   11.395 -1.094  -4.870 0.00 0.00
ATOM  46 HG13ILE  3   13.049 -1.513  -4.834 0.00 0.00
ATOM  47 CD1ILE  3    12.093 -1.430  -2.866 0.00 0.00
ATOM  57 CB GLN  4    7.782 -0.859  -6.273 0.00 0.00
ATOM  58 HB2GLN  4    8.430 -0.149  -6.787 0.00 0.00
ATOM  59 HB3GLN  4    6.790 -0.467  -6.498 0.00 0.00
ATOM  60 CG GLN  4    8.055 -0.772  -4.739 0.00 0.00
ATOM  61 HG2GLN  4    7.662 -1.616  -4.171 0.00 0.00
ATOM  62 HG3GLN  4    9.108 -0.837  -4.468 0.00 0.00
ATOM  63 CD GLN  4    7.522  0.488  -4.102 0.00 0.00
ATOM  64 OE1GLN  4    7.069  1.373  -4.791 0.00 0.00
ATOM  65 NE2GLN  4    7.630  0.776  -2.819 0.00 0.00
ATOM  66 HE21GLN  4   7.145  1.571  -2.499 0.00 0.00
ATOM  67 HE22GLN  4   8.356  0.304  -2.236 0.00 0.00
ATOM  68 C  GLN  4    7.539 -2.410  -8.378 0.00 0.00
ATOM  69 O  GLN  4    8.370 -2.457  -9.222 0.00 0.00
ATOM  70 N  ASN  5    6.228 -2.396  -8.624 0.00 0.00
ATOM  71 H  ASN  5    5.555 -2.470  -7.874 0.00 0.00
ATOM  72 CA ASN  5    5.737 -2.543  -9.981 0.00 0.00
ATOM  73 HA ASN  5    6.225 -3.387 -10.469 0.00 0.00
ATOM  74 CB ASN  5    4.260 -2.882  -9.933 0.00 0.00
ATOM  75 HB2ASN  5    4.018 -3.779  -9.362 0.00 0.00
ATOM  76 HB3ASN  5    3.647 -2.133  -9.431 0.00 0.00
ATOM  77 CG ASN  5    3.627 -3.208 -11.277 0.00 0.00
ATOM  78 OD1ASN  5    4.231 -3.964 -12.086 0.00 0.00
ATOM  79 ND2ASN  5    2.468 -2.772 -11.650 0.00 0.00
ATOM  80 HD21ASN  5   1.964 -3.120 -12.454 0.00 0.00
ATOM  81 HD22ASN  5   1.925 -2.242 -10.984 0.00 0.00
ATOM  82 C  ASN  5    6.041 -1.281 -10.817 0.00 0.00
ATOM  83 O  ASN  5    5.427 -0.203 -10.587 0.00 0.00
ATOM  84 N  CYX  6    6.725 -1.456 -11.929 0.00 0.00
ATOM  85 H  CYX  6    7.055 -2.399 -12.073 0.00 0.00
ATOM  86 CA CYX  6    7.047 -0.453 -12.895 0.00 0.00
ATOM  87 HA CYX  6    6.521  0.482 -12.703 0.00 0.00
ATOM  88 CB CYX  6    8.531 -0.212 -12.836 0.00 0.00
ATOM  89 HB2CYX  6    9.101 -1.124 -13.013 0.00 0.00
ATOM  90 HB3CYX  6    8.857  0.547 -13.547 0.00 0.00
ATOM  91 SG CYX  6    9.034  0.431 -11.216 0.00 0.00
ATOM  92 C  CYX  6    6.749 -1.001 -14.354 0.00 0.00
ATOM  93 O  CYX  6    6.807 -2.158 -14.634 0.00 0.00
ATOM  94 N  PRO  7    6.397 -0.088 -15.279 0.00 0.00
ATOM 104 CA PRO  7    5.804 -0.523 -16.567 0.00 0.00
ATOM 105 HA PRO  7    4.939 -1.171 -16.426 0.00 0.00
ATOM 106 C  PRO  7    6.802 -1.164 -17.552 0.00 0.00
ATOM 107 O  PRO  7    8.049 -1.012 -17.413 0.00 0.00
ATOM 108 N  LEU  8    6.318 -1.866 -18.584 0.00 0.00
ATOM 109 H  LEU  8    5.321 -2.016 -18.636 0.00 0.00
ATOM 110 CA LEU  8    6.997 -2.164 -19.867 0.00 0.00
ATOM 111 HA LEU  8    8.071 -2.099 -19.692 0.00 0.00
ATOM 112 CB LEU  8    6.630 -3.568 -20.292 0.00 0.00
ATOM 113 HB2LEU  8    5.573 -3.471 -20.538 0.00 0.00
ATOM 114 HB3LEU  8    6.696 -4.277 -19.467 0.00 0.00
ATOM 115 CG LEU  8    7.340 -4.112 -21.523 0.00 0.00
ATOM 116 HG LEU  8    7.342 -3.350 -22.301 0.00 0.00
ATOM 117 CD1LEU  8    8.791 -4.531 -21.331 0.00 0.00
ATOM 118 HD11LEU  8   9.437 -3.653 -21.338 0.00 0.00
ATOM 119 HD12LEU  8   8.960 -5.061 -20.394 0.00 0.00
ATOM 120 HD13LEU  8   8.994 -5.200 -22.166 0.00 0.00
ATOM 121 CD2LEU  8    6.490 -5.371 -21.977 0.00 0.00
ATOM 122 HD21LEU  8   5.475 -4.981 -22.064 0.00 0.00
ATOM 123 HD22LEU  8   6.877 -5.941 -22.821 0.00 0.00
ATOM 124 HD23LEU  8   6.469 -6.165 -21.231 0.00 0.00
ATOM 125 C  LEU  8    6.673 -1.132 -20.976 0.00 0.00
ATOM 126 O  LEU  8    5.548 -1.239 -21.524 0.00 0.00
ATOM 127 N  GLY  9    7.628 -0.247 -21.252 0.00 0.00
ATOM 128 H  GLY  9    8.532 -0.538 -20.906 0.00 0.00
ATOM 129 CA GLY  9    7.523  0.835 -22.116 0.00 0.00
ATOM 130 HA2GLY  9    8.428  0.971 -22.708 0.00 0.00
ATOM 131 HA3GLY  9    6.659  0.735 -22.773 0.00 0.00
ATOM 132 C  GLY  9    7.281  2.135 -21.465 0.00 0.00
ATOM 133 O  GLY  9    6.911  2.202 -20.217 0.00 0.00
ATOM 134 N  NHE  10   7.402  3.246 -22.190 0.00 0.00
ATOM 135 HN1NHE  10   7.742  3.253 -23.141 0.00 0.00
ATOM 136 HN2NHE  10   7.223  4.112 -21.701 0.00 0.00
TER  137    NHE  10
ATOM 138 Cl-Cl-  11   13.962  2.310  8.046 0.00 0.00
TER  139    Cl-  11
END
```

OT, intermediate saddle (OPEN/FOLDED)*

OT_MD-II_15us_T16_9

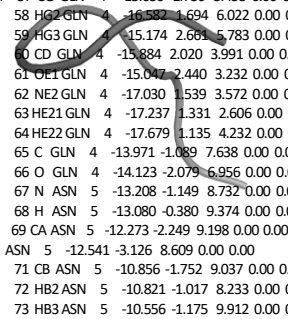


```
ATOM   1 N  CYX  1    5.069 -17.547 -11.905 0.00 0.00
ATOM   2 H1 CYX  1    4.967 -17.074 -11.018 0.00 0.00
ATOM   3 H2 CYX  1    4.905 -18.542 -11.846 0.00 0.00
ATOM   4 H3 CYX  1    4.479 -17.033 -12.543 0.00 0.00
ATOM   5 CA CYX  1    6.428 -17.425 -12.449 0.00 0.00
ATOM   6 HA CYX  1    6.526 -17.834 -13.454 0.00 0.00
ATOM   7 CB CYX  1    6.800 -15.994 -12.805 0.00 0.00
ATOM   8 HB2CYX  1    7.532 -16.076 -13.609 0.00 0.00
ATOM   9 HB3CYX  1    6.000 -15.480 -13.337 0.00 0.00
ATOM  10 SG CYX  1    7.298 -14.859 -11.422 0.00 0.00
ATOM  11 C  CYX  1    7.374 -18.174 -11.616 0.00 0.00
ATOM  12 O  CYX  1    7.042 -18.546 -10.495 0.00 0.00
ATOM  13 N  TYR  2    8.621 -18.431 -12.062 0.00 0.00
ATOM  14 H  TYR  2    8.845 -18.166 -13.011 0.00 0.00
ATOM  15 CA TYR  2    9.619 -19.068 -11.115 0.00 0.00
ATOM  16 HA TYR  2    9.250 -19.925 -10.551 0.00 0.00
ATOM  17 CB TYR  2    10.740 -19.565 -12.029 0.00 0.00
ATOM  18 HB2TYR  2    10.374 -20.229 -12.813 0.00 0.00
ATOM  19 HB3TYR  2    11.297 -18.757 -12.503 0.00 0.00
ATOM  20 CG TYR  2    11.810 -20.341 -11.261 0.00 0.00
ATOM  21 CD1TYR  2    13.143 -19.821 -11.106 0.00 0.00
ATOM  22 HD1TYR  2    13.328 -18.931 -11.689 0.00 0.00
ATOM  23 CE1TYR  2    14.044 -20.428 -10.231 0.00 0.00
ATOM  24 HE1TYR  2    15.024 -20.017 -10.042 0.00 0.00
ATOM  25 CZ TYR  2    13.708 -21.623  -9.607 0.00 0.00
ATOM  26 OH TYR  2    14.627 -22.056  -8.651 0.00 0.00
ATOM  27 HH TYR  2    14.273 -22.825  -8.197 0.00 0.00
ATOM  28 CE2TYR  2    12.522 -22.206  -9.889 0.00 0.00
ATOM  29 HE2TYR  2    12.176 -23.023  -9.273 0.00 0.00
ATOM  30 CD2TYR  2    11.519 -21.607 -10.645 0.00 0.00
ATOM  31 HD2TYR  2    10.532 -22.026 -10.776 0.00 0.00
ATOM  32 C  TYR  2    10.195 -18.150 -10.047 0.00 0.00
ATOM  33 O  TYR  2    10.058 -16.937 -10.089 0.00 0.00
ATOM  34 N  ILE  3    10.597 -18.690  -8.938 0.00 0.00
ATOM  35 H  ILE  3    10.649 -19.694  -8.844 0.00 0.00
ATOM  36 CA ILE  3    10.891 -18.042  -7.695 0.00 0.00
ATOM  37 HA ILE  3    10.037 -17.394  -7.495 0.00 0.00
ATOM  38 CB ILE  3    10.883 -19.141  -6.522 0.00 0.00
ATOM  39 HB ILE  3    11.087 -18.605  -5.595 0.00 0.00
ATOM  40 CG2ILE  3    9.475 -19.723  -6.348 0.00 0.00
ATOM  41 HG21ILE  3   9.190 -20.265  -7.249 0.00 0.00
ATOM  42 HG22ILE  3   9.442 -20.441  -5.529 0.00 0.00
ATOM  43 HG23ILE  3   8.736 -18.931  -6.227 0.00 0.00
ATOM  44 CG1ILE  3    11.862 -20.315  -6.669 0.00 0.00
ATOM  45 HG12ILE  3   11.519 -20.894  -7.526 0.00 0.00
ATOM  46 HG13ILE  3   12.842 -19.920  -6.936 0.00 0.00
ATOM  47 CD1ILE  3    12.012 -21.312  -5.540 0.00 0.00
ATOM  48 HD11ILE  3   12.686 -22.139  -5.761 0.00 0.00
ATOM  49 HD12ILE  3   12.479 -20.733  -4.743 0.00 0.00
ATOM  50 HD13ILE  3   11.027 -21.636  -5.202 0.00 0.00
ATOM  51 C  ILE  3    12.208 -17.214  -7.705 0.00 0.00
ATOM  52 O  ILE  3    12.417 -16.344  -6.846 0.00 0.00
ATOM  53 N  GLN  4    12.963 -17.309  -8.796 0.00 0.00
ATOM  54 H  GLN  4    12.609 -17.958  -9.485 0.00 0.00
ATOM  55 CA GLN  4    14.116 -16.429  -9.082 0.00 0.00
ATOM  56 HA GLN  4    14.002 -15.561  -8.432 0.00 0.00
ATOM  57 CB GLN  4    15.508 -17.159  -8.858 0.00 0.00
ATOM  58 HB2GLN  4    15.643 -17.919  -9.628 0.00 0.00
ATOM  59 HB3GLN  4    16.368 -16.495  -8.946 0.00 0.00
ATOM  60 CG GLN  4    15.612 -17.916  -7.549 0.00 0.00
ATOM  61 HG2GLN  4    14.754 -18.563  -7.368 0.00 0.00
ATOM  62 HG3GLN  4    16.547 -18.474  -7.605 0.00 0.00
ATOM  63 CD GLN  4    15.669 -17.021  -6.304 0.00 0.00
ATOM  64 OE1GLN  4    15.523 -15.797  -6.404 0.00 0.00
ATOM  65 NE2GLN  4    16.210 -17.506  -5.262 0.00 0.00
ATOM  66 HE21GLN  4   16.309 -18.495  -5.083 0.00 0.00
ATOM  67 HE22GLN  4   16.328 -16.852  -4.502 0.00 0.00
ATOM  68 C  GLN  4    14.172 -15.829 -10.487 0.00 0.00
ATOM  69 O  GLN  4    13.634 -16.499 -11.404 0.00 0.00
ATOM  70 N  ASN  5    14.793 -14.667 -10.692 0.00 0.00
ATOM  71 H  ASN  5    15.112 -14.163  -9.877 0.00 0.00
ATOM  72 CA ASN  5    14.964 -13.937 -12.005 0.00 0.00
ATOM  73 HA ASN  5    15.238 -12.936 -11.671 0.00 0.00
ATOM  74 CB ASN  5    16.168 -14.567 -12.787 0.00 0.00
ATOM  75 HB2ASN  5    16.951 -14.779 -12.059 0.00 0.00
ATOM  76 HB3ASN  5    15.874 -15.518 -13.232 0.00 0.00
ATOM  77 CG ASN  5    16.818 -13.505 -13.742 0.00 0.00
ATOM  78 OD1ASN  5    17.317 -12.523 -13.252 0.00 0.00
ATOM  79 ND2ASN  5    16.767 -13.617 -15.021 0.00 0.00
ATOM  80 HD21ASN  5   17.070 -12.846 -15.599 0.00 0.00
ATOM  81 HD22ASN  5   16.237 -14.375 -15.427 0.00 0.00
ATOM  82 C  ASN  5    13.634 -13.729 -12.852 0.00 0.00
ATOM  83 O  ASN  5    13.785 -13.595 -14.086 0.00 0.00
ATOM  84 N  CYX  6    12.454 -13.570 -12.244 0.00 0.00
ATOM  85 H  CYX  6    12.397 -13.860 -11.279 0.00 0.00
ATOM  86 CA CYX  6    11.148 -13.177 -12.791 0.00 0.00
ATOM  87 HA CYX  6    11.185 -13.318 -13.871 0.00 0.00
ATOM  88 CB CYX  6    10.087 -14.141 -12.280 0.00 0.00
ATOM  89 HB2CYX  6    10.178 -15.062 -12.856 0.00 0.00
ATOM  90 HB3CYX  6    10.298 -14.492 -11.270 0.00 0.00
ATOM  91 SG CYX  6    8.414 -13.524 -12.319 0.00 0.00
ATOM  92 C  CYX  6    10.803 -11.737 -12.428 0.00 0.00
ATOM  93 O  CYX  6    11.155 -11.346 -11.333 0.00 0.00
ATOM  94 N  PRO  7    10.156 -10.887 -13.288 0.00 0.00
ATOM  95 CD PRO  7    9.897 -11.294 -14.664 0.00 0.00
ATOM  96 HD2PRO  7    8.923 -11.775 -14.759 0.00 0.00
ATOM  97 HD3PRO  7    10.732 -11.832 -15.114 0.00 0.00
ATOM  98 CG PRO  7    9.856 -10.008 -15.459 0.00 0.00
ATOM  99 HG2PRO  7    9.129 -10.072 -16.268 0.00 0.00
ATOM 100 HG3PRO  7    10.804  -9.787 -15.949 0.00 0.00
ATOM 101 CB PRO  7    9.508  -8.931 -14.398 0.00 0.00
ATOM 102 HB2PRO  7    8.427  -8.884 -14.268 0.00 0.00
ATOM 103 HB3PRO  7    9.809  -7.939 -14.736 0.00 0.00
ATOM 104 CA PRO  7    10.131  -9.438 -13.124 0.00 0.00
ATOM 105 HA PRO  7    11.135  -9.032 -13.001 0.00 0.00
ATOM 106 C  PRO  7    9.283  -9.123 -11.898 0.00 0.00
ATOM 107 O  PRO  7    8.299  -9.745 -11.527 0.00 0.00
ATOM 108 N  LEU  8    9.724  -8.040 -11.188 0.00 0.00
ATOM 109 H  LEU  8    10.584  -7.568 -11.430 0.00 0.00
ATOM 110 CA LEU  8    9.120  -7.713  -9.869 0.00 0.00
ATOM 111 HA LEU  8    8.512  -8.542  -9.506 0.00 0.00
ATOM 112 CB LEU  8    10.233  -7.466  -8.827 0.00 0.00
ATOM 113 HB2LEU  8    10.925  -6.677  -9.119 0.00 0.00
ATOM 114 HB3LEU  8    9.652  -7.105  -7.978 0.00 0.00
ATOM 115 CG LEU  8    11.024  -8.782  -8.480 0.00 0.00
ATOM 116 HG LEU  8    11.780  -9.044  -9.220 0.00 0.00
ATOM 117 CD1LEU  8    11.680  -8.649  -7.069 0.00 0.00
ATOM 118 HD11LEU  8   10.865  -8.319  -6.425 0.00 0.00
ATOM 119 HD12LEU  8   12.236  -9.540  -6.778 0.00 0.00
ATOM 120 HD13LEU  8   12.428  -7.856  -7.056 0.00 0.00
ATOM 121 CD2LEU  8    10.092  -9.998  -8.316 0.00 0.00
ATOM 122 HD21LEU  8   10.638 -10.753  -7.750 0.00 0.00
ATOM 123 HD22LEU  8    9.271  -9.709  -7.660 0.00 0.00
ATOM 124 HD23LEU  8    9.819 -10.386  -9.297 0.00 0.00
ATOM 125 C  LEU  8    8.202  -6.431  -9.977 0.00 0.00
ATOM 126 O  LEU  8    8.540  -5.591 -10.778 0.00 0.00
ATOM 127 N  GLY  9    7.104  -6.463  -9.316 0.00 0.00
ATOM 128 H  GLY  9    6.882  -7.358  -8.904 0.00 0.00
ATOM 129 CA GLY  9    6.007  -5.590  -9.596 0.00 0.00
ATOM 130 HA2GLY  9    5.408  -5.429  -8.700 0.00 0.00
ATOM 131 HA3GLY  9    6.292  -4.599  -9.948 0.00 0.00
ATOM 132 C  GLY  9    5.009  -6.070 -10.639 0.00 0.00
ATOM 133 O  GLY  9    4.517  -7.192 -10.657 0.00 0.00
ATOM 134 N  NHE  10   4.617  -5.238 -11.624 0.00 0.00
ATOM 135 HN1NHE  10   5.022  -4.313 -11.618 0.00 0.00
ATOM 136 HN2NHE  10   4.060  -5.530 -12.415 0.00 0.00
TER  137    NHE  10
ATOM 138 Cl-Cl-  11   -9.180  -2.592  6.191 0.00 0.00
TER  139    Cl-  11
END
```

dOT, twisted saddle (FOLDED)

dOT_MD_3us_T16_4

```
ATOM   1 H1 CYE  1   -9.094  0.925 15.600 0.00 0.00
ATOM  46 HD12ILE  3  -12.981  7.149  7.916 0.00 0.00
ATOM  91 N  PRO  7   -13.680 -4.064 14.444 0.00 0.00
```



```
ATOM    2 CA CYE   1    -9.506  0.444 14.713 0.00 0.00
ATOM    3 HA CYE   1    -8.849 -0.394 14.481 0.00 0.00
ATOM    4 CB CYE   1   -10.803 -0.250 15.162 0.00 0.00
ATOM    5 HB2 CYE  1   -11.081 -0.865 14.306 0.00 0.00
ATOM    6 HB3 CYE  1   -10.588 -0.814 16.070 0.00 0.00
ATOM    7 SG CYE   1   -12.205  0.826 15.299 0.00 0.00
ATOM    8 C  CYE   1    -9.698  1.450 13.591 0.00 0.00
ATOM    9 O  CYE   1    -9.936  2.668 13.849 0.00 0.00
ATOM   10 N  TYR   2    -9.544  1.021 12.367 0.00 0.00
ATOM   11 H  TYR   2    -9.369  0.038 12.215 0.00 0.00
ATOM   12 CA TYR   2    -9.762  1.885 11.241 0.00 0.00
ATOM   13 HA TYR   2    -9.486  2.845 11.677 0.00 0.00
ATOM   14 CB TYR   2    -8.725  1.554 10.181 0.00 0.00
ATOM   15 HB2 TYR  2    -7.713  1.517 10.585 0.00 0.00
ATOM   16 HB3 TYR  2    -8.992  0.687  9.578 0.00 0.00
ATOM   17 CG TYR   2    -8.551  2.766  9.266 0.00 0.00
ATOM   18 CD1 TYR  2    -8.162  4.017  9.818 0.00 0.00
ATOM   19 HD1 TYR  2    -7.966  4.090 10.877 0.00 0.00
ATOM   20 CE1 TYR  2    -8.072  5.172  8.984 0.00 0.00
ATOM   21 HE1 TYR  2    -7.811  6.124  9.423 0.00 0.00
ATOM   22 CZ TYR   2    -8.300  5.009  7.635 0.00 0.00
ATOM   23 OH TYR   2    -8.351  6.092  6.869 0.00 0.00
ATOM   24 HH TYR   2    -8.448  6.857  7.441 0.00 0.00
ATOM   25 CE2 TYR  2    -8.628  3.740  7.130 0.00 0.00
ATOM   26 HE2 TYR  2    -8.872  3.615  6.086 0.00 0.00
ATOM   27 CD2 TYR  2    -8.725  2.611  7.946 0.00 0.00
ATOM   28 HD2 TYR  2    -8.937  1.637  7.529 0.00 0.00
ATOM   29 C  TYR   2   -11.278  1.833 10.867 0.00 0.00
ATOM   30 O  TYR   2   -12.056  0.924 11.143 0.00 0.00
ATOM   31 N  ILE   3   -11.670  2.903 10.114 0.00 0.00
ATOM   32 H  ILE   3   -10.968  3.563  9.811 0.00 0.00
ATOM   33 CA ILE   3   -12.973  2.954  9.429 0.00 0.00
ATOM   34 HA ILE   3   -13.697  2.700 10.203 0.00 0.00
ATOM   35 CB ILE   3   -13.421  4.352  8.853 0.00 0.00
ATOM   36 HB ILE   3   -14.076  4.187  7.997 0.00 0.00
ATOM   37 CG2 ILE  3   -14.202  5.141  9.993 0.00 0.00
ATOM   38 HG21 ILE 3   -14.576  6.032  9.489 0.00 0.00
ATOM   39 HG22 ILE 3   -15.072  4.539 10.257 0.00 0.00
ATOM   40 HG23 ILE 3   -13.470  5.361 10.769 0.00 0.00
ATOM   41 CG1 ILE  3   -12.215  5.127  8.309 0.00 0.00
ATOM   42 HG12 ILE 3   -11.460  5.347  9.064 0.00 0.00
ATOM   43 HG13 ILE 3   -11.631  4.450  7.686 0.00 0.00
ATOM   44 CD1 ILE  3   -12.572  6.311  7.350 0.00 0.00
ATOM   45 HD11 ILE 3   -13.386  6.051  6.674 0.00 0.00
ATOM   47 HD13 ILE 3   -11.747  6.581  6.691 0.00 0.00
ATOM   48 C  ILE   3   -13.113  1.902  8.413 0.00 0.00
ATOM   49 O  ILE   3   -12.148  1.502  7.765 0.00 0.00
ATOM   50 N  GLN   4   -14.348  1.335  8.219 0.00 0.00
ATOM   51 H  GLN   4   -15.178  1.737  8.629 0.00 0.00
ATOM   52 CA GLN   4   -14.708  0.180  7.327 0.00 0.00
ATOM   53 HA GLN   4   -15.718 -0.074  7.650 0.00 0.00
ATOM   54 CB GLN   4   -14.715  0.540  5.801 0.00 0.00
ATOM   55 HB2 GLN  4   -13.771  0.712  5.285 0.00 0.00
ATOM   56 HB3 GLN  4   -15.180 -0.319  5.317 0.00 0.00
AT OM  57 CG GLN   4   -15.650  1.736  5.458 0.00 0.00
ATOM   58 HG2 GLN  4   -16.582  1.694  6.022 0.00 0.00
ATOM   59 HG3 GLN  4   -15.174  2.661  5.783 0.00 0.00
ATOM   60 CD GLN   4   -15.884  2.020  3.991 0.00 0.00
ATOM   61 OE1 GLN  4   -15.047  2.440  3.232 0.00 0.00
ATOM   62 NE2 GLN  4   -17.030  1.539  3.572 0.00 0.00
ATOM   63 HE21 GLN 4   -17.237  1.331  2.606 0.00 0.00
ATOM   64 HE22 GLN 4   -17.679  1.135  4.232 0.00 0.00
ATOM   65 C  GLN   4   -13.971 -1.089  7.638 0.00 0.00
ATOM   66 O  GLN   4   -14.123 -2.079  6.956 0.00 0.00
ATOM   67 N  ASN   5   -13.208 -1.149  8.732 0.00 0.00
ATOM   68 H  ASN   5   -13.080 -0.380  9.374 0.00 0.00
ATOM   69 CA ASN   5  -12.273 -2.249  9.198 0.00 0.00
70 HA ASN   5   -12.541 -3.126  8.609 0.00 0.00
ATOM   71 CB ASN   5  -10.856 -1.752  9.037 0.00 0.00
ATOM   72 HB2 ASN  5   -10.821 -1.017  8.233 0.00 0.00
ATOM   73 HB3 ASN  5   -10.556 -1.175  9.912 0.00 0.00
ATOM   74 CG ASN   5   -9.840 -2.830  8.740 0.00 0.00
ATOM   75 OD1 ASN  5   -9.599 -3.096  7.544 0.00 0.00
ATOM   76 ND2 ASN  5   -9.177 -3.448  9.693 0.00 0.00
ATOM   77 HD21 ASN 5   -8.392 -4.061  9.521 0.00 0.00
ATOM   78 HD22 ASN 5   -9.390 -3.281 10.666 0.00 0.00
ATOM   79 C  ASN   5  -12.501 -2.655 10.675 0.00 0.00
ATOM   80 O  ASN   5  -11.780 -3.577 11.115 0.00 0.00
ATOM   81 N  CYX   6  -13.298 -1.921 11.481 0.00 0.00
ATOM   82 H  CYX   6  -13.831 -1.141 11.124 0.00 0.00
ATOM   83 CA CYX   6  -13.237 -2.068 12.949 0.00 0.00
ATOM   84 HA CYX   6  -12.191 -2.091 13.255 0.00 0.00
ATOM   85 CB CYX   6  -13.937 -0.792 13.488 0.00 0.00
ATOM   86 HB2 CYX  6   -13.544  0.070 12.949 0.00 0.00
ATOM   87 HB3 CYX  6   -14.995 -0.928 13.264 0.00 0.00
ATOM   88 SG CYX   6  -13.865 -0.409 15.237 0.00 0.00
ATOM   89 C  CYX   6  -14.115 -3.282 13.429 0.00 0.00
ATOM   90 O  CYX   6  -15.293 -3.375 13.040 0.00 0.00
ATOM   92 CD PRO   7  -12.348 -3.936 14.987 0.00 0.00
ATOM   93 HD2 PRO  7   -12.290 -3.206 15.794 0.00 0.00
ATOM   94 HD3 PRO  7   -11.600 -3.628 14.257 0.00 0.00
ATOM   95 CG PRO   7   -11.921 -5.305 15.500 0.00 0.00
ATOM   96 HG2 PRO  7   -11.596 -5.255 16.539 0.00 0.00
ATOM   97 HG3 PRO  7   -11.278 -5.826 14.790 0.00 0.00
ATOM   98 CB PRO   7  -13.242 -6.143 15.483 0.00 0.00
ATOM   99 HB2 PRO  7   -13.463 -6.584 16.455 0.00 0.00
ATOM  100 HB3 PRO  7   -13.082 -6.919 14.735 0.00 0.00
ATOM  101 CA PRO   7  -14.388 -5.158 15.045 0.00 0.00
ATOM  102 HA PRO   7   -14.985 -5.577 14.234 0.00 0.00
ATOM  103 C  PRO   7  -15.247 -4.763 16.163 0.00 0.00
ATOM  104 O  PRO   7  -14.780 -4.064 17.082 0.00 0.00
ATOM  105 N  LEU   8  -16.416 -5.243 16.333 0.00 0.00
ATOM  106 H  LEU   8  -16.648 -6.010 15.719 0.00 0.00
ATOM  107 CA LEU   8  -17.444 -4.802 17.381 0.00 0.00
ATOM  108 HA LEU   8   -17.323 -3.732 17.552 0.00 0.00
ATOM  109 CB LEU   8   -18.844 -5.007 16.663 0.00 0.00
ATOM  110 HB2 LEU  8   -19.487 -4.334 17.231 0.00 0.00
ATOM  111 HB3 LEU  8   -18.800 -4.617 15.646 0.00 0.00
ATOM  112 CG LEU   8   -19.521 -6.462 16.685 0.00 0.00
ATOM  113 HG LEU   8   -19.478 -6.802 17.719 0.00 0.00
ATOMATOM  114 CD1 LEU  8   -20.921 -6.521 16.066 0.00 0.00
ATOM  115 HD11 LEU  8   -21.028 -5.916 15.166 0.00 0.00
ATOM  116 HD12 LEU  8   -21.165 -7.574 15.927 0.00 0.00
ATOM  117 HD13 LEU  8   -21.695 -6.071 16.687 0.00 0.00
ATOM  118 CD2 LEU   8   -18.841 -7.494 15.769 0.00 0.00
ATOM  119 HD21 LEU  8   -18.637 -7.066 14.788 0.00 0.00
ATOM  120 HD22 LEU  8   -17.869 -7.685 16.223 0.00 0.00
ATOM  121 HD23 LEU  8   -19.360 -8.452 15.751 0.00 0.00
ATOM  122 C  LEU   8   -17.249 -5.527 18.731 0.00 0.00
ATOM  123 O  LEU   8   -16.998 -6.772 18.788 0.00 0.00
ATOM  124 N  GLY   9   -17.551 -4.892 19.813 0.00 0.00
ATOM  125 H  GLY   9   -17.858 -3.932 19.746 0.00 0.00
ATOM  126 CA GLY   9   -17.508 -5.476 21.195 0.00 0.00
ATOM  127 HA2 GLY  9   -16.703 -6.212 21.191 0.00 0.00
ATOM  128 HA3 GLY  9   -17.215 -4.628 21.814 0.00 0.00
ATOM  129 C  GLY   9   -18.847 -6.149 21.644 0.00 0.00
ATOM  130 O  GLY   9   -19.816 -6.043 20.870 0.00 0.00
ATOM  131 N  NHE  10   -18.907 -6.755 22.850 0.00 0.00
ATOM  132 HN1 NHE  10   -18.137 -6.594 23.483 0.00 0.00
ATOM  133 HN2 NHE  10   -19.770 -7.149 23.198 0.00 0.00
TER   134    NHE  10
END
```

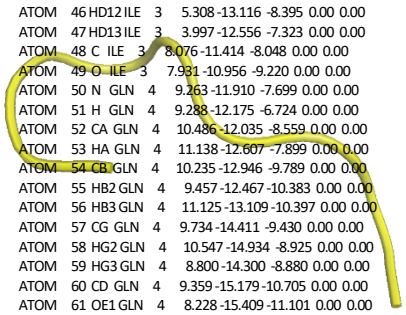
dOT, twisted saddlehelix (FOLDED)

dOT_MD_3us_T16_7

```
ATOM    1 H1 CYE   1    8.954  1.944  7.785 0.00 0.00
ATOM    2 CA CYE   1    8.356  2.609  7.162 0.00 0.00
ATOM    3 HA CYE   1    8.888  2.725  6.217 0.00 0.00
ATOM    4 CB CYE   1    7.098  1.826  6.740 0.00 0.00
ATOM    5 HB2 CYE  1    6.469  2.558  6.234 0.00 0.00
ATOM    6 HB3 CYE  1    7.288  1.030  6.021 0.00 0.00
ATOM    7 SG CYE   1    6.196  1.209  8.236 0.00 0.00
ATOM    8 C  CYE   1    8.055  3.969  7.851 0.00 0.00
ATOM    9 O  CYE   1    8.015  4.034  9.072 0.00 0.00
ATOM   10 N  TYR   2    7.764  4.981  7.027 0.00 0.00
ATOM   11 H  TYR   2    7.684  4.833  6.031 0.00 0.00
ATOM   12 CA TYR   2    7.551  6.360  7.647 0.00 0.00
ATOM   13 HA TYR   2    8.477  6.690  8.118 0.00 0.00
ATOM   14 CB TYR   2    7.140  7.369  6.542 0.00 0.00
ATOM   15 HB2 TYR  2    7.701  7.256  5.614 0.00 0.00
ATOM   16 HB3 TYR  2    6.119  7.107  6.265 0.00 0.00
ATOM   17 CG TYR   2    7.036  8.803  7.074 0.00 0.00
ATOM   18 CD1 TYR  2    8.124  9.619  7.202 0.00 0.00
ATOM   19 HD1 TYR  2    9.078  9.128  7.080 0.00 0.00
ATOM   20 CE1 TYR  2    8.057 10.951  7.704 0.00 0.00
ATOM   21 HE1 TYR  2    8.909 11.594  7.872 0.00 0.00
ATOM   22 CZ TYR   2    6.799 11.441  8.070 0.00 0.00
ATOM   23 OH TYR   2    6.795 12.661  8.681 0.00 0.00
ATOM   24 HH TYR   2    5.891 12.963  8.795 0.00 0.00
ATOM   25 CE2 TYR  2    5.691 10.692  7.969 0.00 0.00
ATOM   26 HE2 TYR  2    4.726 11.052  8.295 0.00 0.00
ATOM   27 CD2 TYR  2    5.785  9.380  7.400 0.00 0.00
ATOM   28 HD2 TYR  2    4.926  8.726  7.404 0.00 0.00
ATOM   29 C  TYR   2    6.586  6.428  8.813 0.00 0.00
ATOM   30 O  TYR   2    5.497  5.831  8.639 0.00 0.00
ATOM   31 N  ILE   3    6.965  6.935  9.972 0.00 0.00
ATOM   32 H  ILE   3    7.877  7.370  9.978 0.00 0.00
ATOM   33 CA ILE   3    6.171  6.905 11.225 0.00 0.00
ATOM   34 HA ILE   3    6.935  7.067 11.986 0.00 0.00
ATOM   35 CB ILE   3    5.269  8.122 11.459 0.00 0.00
ATOM   36 HB ILE   3    4.695  8.046 12.383 0.00 0.00
ATOM   37 CG2 ILE  3    6.167  9.336 11.609 0.00 0.00
ATOM   38 HG21 ILE 3    6.741  9.386 12.534 0.00 0.00
ATOM   39 HG22 ILE 3    6.887  9.304 10.791 0.00 0.00
ATOM   40 HG23 ILE 3    5.612 10.273 11.567 0.00 0.00
ATOM   41 CG1 ILE  3    4.141  8.159 10.459 0.00 0.00
ATOM   42 HG12 ILE 3    4.492  8.433  9.464 0.00 0.00
ATOM   43 HG13 ILE 3    3.719  7.168 10.293 0.00 0.00
ATOM   44 CD1 ILE  3    3.027  9.168 10.836 0.00 0.00
ATOM   45 HD11 ILE 3    2.476  9.374  9.919 0.00 0.00
ATOM   46 HD12 ILE 3    2.415  8.703 11.609 0.00 0.00
ATOM   47 HD13 ILE 3    3.441 10.073 11.279 0.00 0.00
ATOM   48 C  ILE   3    5.657  5.539 11.595 0.00 0.00
ATOM   49 O  ILE   3    4.642  5.451 12.288 0.00 0.00
ATOM   50 N  GLN   4    6.424  4.492 11.259 0.00 0.00
ATOM   51 H  GLN   4    7.218  4.608 10.645 0.00 0.00
ATOM   52 CA GLN   4    6.018  3.113 11.486 0.00 0.00
ATOM   53 HA GLN   4    6.656  2.497 10.853 0.00 0.00
ATOM   54 CB GLN   4    6.190  2.736 12.973 0.00 0.00
ATOM   55 HB2 GLN  4    5.441  3.257 13.569 0.00 0.00
ATOM   56 HB3 GLN  4    6.095  1.664 13.147 0.00 0.00
ATOM   57 CG GLN   4    7.494  3.229 13.672 0.00 0.00
ATOM   58 HG2 GLN  4    7.339  4.307 13.710 0.00 0.00
ATOM   59 HG3 GLN  4    7.492  2.853 14.695 0.00 0.00
ATOM   60 CD GLN   4    8.658  2.741 12.867 0.00 0.00
ATOM   61 OE1 GLN  4    9.309  3.543 12.294 0.00 0.00
ATOM   62 NE2 GLN  4    8.732  1.450 12.664 0.00 0.00
ATOM   63 HE21 GLN 4    9.465  1.130 12.047 0.00 0.00
ATOM   64 HE22 GLN 4    8.299  0.873 13.370 0.00 0.00
ATOM   65 C  GLN   4    4.605  2.743 11.017 0.00 0.00
ATOM   66 O  GLN   4    4.108  1.672 11.424 0.00 0.00
ATOM   67 N  ASN   5    3.959  3.570 10.177 0.00 0.00
ATOM   68 H  ASN   5    4.520  4.370  9.920 0.00 0.00
ATOM   69 CA ASN   5    2.514  3.512 10.025 0.00 0.00
ATOM   70 HA ASN   5    2.131  3.142 10.976 0.00 0.00
ATOM   71 CB ASN   5    1.883  4.866  9.648 0.00 0.00
ATOM   72 HB2 ASN  5    0.797  4.951  9.668 0.00 0.00
ATOM   73 HB3 ASN  5    2.233  5.527 10.440 0.00 0.00
ATOM   74 CG ASN   5    2.296  5.454  8.287 0.00 0.00
ATOM   75 OD1 ASN  5    2.911  4.775  7.495 0.00 0.00
ATOM   76 ND2 ASN  5    2.018  6.734  8.153 0.00 0.00
ATOM   77 HD21 ASN 5    2.347  7.145  7.291 0.00 0.00
ATOM   78 HD22 ASN 5    1.442  7.301  8.758 0.00 0.00
ATOM   79 C  ASN   5    2.027  2.366  9.083 0.00 0.00
ATOM   80 O  ASN   5    0.859  2.408  8.575 0.00 0.00
ATOM   81 N  CYX   6    2.822  1.443  8.703 0.00 0.00
ATOM   82 H  CYX   6    3.761  1.333  9.057 0.00 0.00
ATOM   83 CA CYX   6    2.439  0.382  7.703 0.00 0.00
ATOM   84 HA CYX   6    1.567  0.746  7.159 0.00 0.00
ATOM   85 CB CYX   6    3.663  0.233  6.808 0.00 0.00
ATOM   86 HB2 CYX  6    3.438 -0.475  6.010 0.00 0.00
ATOM   87 HB3 CYX  6    3.774  1.232  6.386 0.00 0.00
ATOM   88 SG CYX   6    5.170 -0.387  7.632 0.00 0.00
ATOM   89 C  CYX   6    2.071 -0.907  8.484 0.00 0.00
ATOM   90 O  CYX   6    2.311 -1.066  9.735 0.00 0.00
ATOM   91 N  PRO   7    1.443 -1.924  7.855 0.00 0.00
ATOM   92 CD PRO   7    1.012 -1.926  6.483 0.00 0.00
ATOM   93 HD2 PRO  7    1.724 -1.514  5.767 0.00 0.00
ATOM   94 HD3 PRO  7    0.161 -1.246  6.432 0.00 0.00
ATOM   95 CG PRO   7    0.787 -3.398  6.148 0.00 0.00
ATOM   96 HG2 PRO  7    1.699 -3.777  5.688 0.00 0.00
ATOM   97 HG3 PRO  7   -0.106 -3.479  5.528 0.00 0.00
ATOM   98 CB PRO   7    0.508 -4.104  7.449 0.00 0.00
ATOM   99 HB2 PRO  7    0.813 -5.146  7.347 0.00 0.00
ATOM  100 HB3 PRO  7   -0.562 -3.982  7.620 0.00 0.00
ATOM  101 CA PRO   7    1.331 -3.288  8.438 0.00 0.00
ATOM  102 HA PRO   7    0.763 -3.284  9.368 0.00 0.00
ATOM  103 C  PRO   7    2.723 -3.942  8.652 0.00 0.00
ATOM  104 O  PRO   7    3.577 -3.825  7.699 0.00 0.00
ATOM  105 N  LEU   8    2.930 -4.588  9.773 0.00 0.00
ATOM  106 H  LEU   8    2.123 -4.776 10.351 0.00 0.00
ATOM  107 CA LEU   8    4.219 -5.161 10.264 0.00 0.00
ATOM  108 HA LEU   8    4.024 -5.585 11.249 0.00 0.00
ATOM  109 CB LEU   8    4.686 -6.404  9.466 0.00 0.00
ATOM  110 HB2 LEU  8    4.886 -6.217  8.411 0.00 0.00
ATOM  111 HB3 LEU  8    5.684 -6.598  9.861 0.00 0.00
ATOM  112 CG LEU   8    3.968 -7.771  9.789 0.00 0.00
ATOM  113 HG LEU   8    3.871 -7.967 10.857 0.00 0.00
ATOM  114 CD1 LEU  8    2.482 -7.827  9.185 0.00 0.00
ATOM  115 HD11 LEU  8    1.865 -6.944  9.351 0.00 0.00
ATOM  116 HD12 LEU  8    2.516 -7.764  8.098 0.00 0.00
ATOM  117 HD13 LEU  8    2.033 -8.786  9.444 0.00 0.00
ATOM  118 CD2 LEU   8    4.843 -8.903  9.218 0.00 0.00
ATOM  119 HD21 LEU  8    5.018 -8.705  8.161 0.00 0.00
ATOM  120 HD22 LEU  8    5.768 -8.965  9.790 0.00 0.00
ATOM  121 HD23 LEU  8    4.484 -9.931  9.268 0.00 0.00
ATOM  122 C  LEU   8    5.371 -4.211 10.611 0.00 0.00
ATOM  123 O  LEU   8    6.535 -4.585 10.898 0.00 0.00
ATOM  124 N  GLY   9    5.110 -2.873 10.591 0.00 0.00
ATOM  125 H  GLY   9    4.143 -2.581 10.578 0.00 0.00
ATOM  126 CA GLY   9    6.157 -1.824 10.913 0.00 0.00
ATOM  127 HA2 GLY  9    5.654 -0.882 11.130 0.00 0.00
ATOM  128 HA3 GLY  9    6.701 -2.062 11.827 0.00 0.00
ATOM  129 C  GLY   9    7.275 -1.696  9.813 0.00 0.00
ATOM  130 O  GLY   9    8.229 -0.926 10.065 0.00 0.00
ATOM  131 N  NHE  10    7.316 -2.448  8.733 0.00 0.00
ATOM  132 HN1 NHE  10    6.506 -2.961  8.415 0.00 0.00
ATOM  133 HN2 NHE  10    8.145 -2.336  8.167 0.00 0.00
TER   134    NHE  10
END
```

dOT, scoop-like open23variant (OPEN/FOLDED)*

OT_MD_3us_T16_2



```
ATOM   1 H1 CYE  1   10.236 -7.670 -2.618 0.00 0.00
ATOM   2 CA CYE  1   10.603 -8.695 -2.658 0.00 0.00
ATOM   3 HA CYE  1   11.153 -8.858 -3.586 0.00 0.00
ATOM   4 CB CYE  1   11.753 -8.903 -1.707 0.00 0.00
ATOM   5 HB2 CYE  1   11.495 -8.632 -0.683 0.00 0.00
ATOM   6 HB3 CYE  1   12.140 -9.921 -1.685 0.00 0.00
ATOM   7 SG CYE  1   13.247 -7.900 -2.137 0.00 0.00
ATOM   8 C  CYE  1    9.534 -9.716 -2.358 0.00 0.00
ATOM   9 O  CYE  1    9.492 -10.361 -1.313 0.00 0.00
ATOM  10 N  TYR  2    8.466 -9.875 -3.213 0.00 0.00
ATOM  11 H  TYR  2    8.600 -9.376 -4.081 0.00 0.00
ATOM  12 CA TYR  2    7.537 -11.011 -3.105 0.00 0.00
ATOM  13 HA TYR  2    7.911 -11.886 -2.574 0.00 0.00
ATOM  14 CB TYR  2    6.172 -10.572 -2.347 0.00 0.00
ATOM  15 HB2 TYR  2    5.621 -11.365 -1.840 0.00 0.00
ATOM  16 HB3 TYR  2    6.429 -9.859 -1.564 0.00 0.00
ATOM  17 CG TYR  2    5.136 -9.848 -3.161 0.00 0.00
ATOM  18 CD1 TYR  2    4.002 -10.467 -3.813 0.00 0.00
ATOM  19 HD1 TYR  2    3.797 -11.499 -3.567 0.00 0.00
ATOM  20 CE1 TYR  2    3.159 -9.662 -4.566 0.00 0.00
ATOM  21 HE1 TYR  2    2.225 -10.003 -4.990 0.00 0.00
ATOM  22 CZ TYR  2    3.375 -8.257 -4.704 0.00 0.00
ATOM  23 OH TYR  2    2.544 -7.495 -5.448 0.00 0.00
ATOM  24 HH TYR  2    1.782 -7.972 -5.783 0.00 0.00
ATOM  25 CE2 TYR  2    4.469 -7.673 -4.125 0.00 0.00
ATOM  26 HE2 TYR  2    4.603 -6.606 -4.215 0.00 0.00
ATOM  27 CD2 TYR  2    5.430 -8.496 -3.438 0.00 0.00
ATOM  28 HD2 TYR  2    6.310 -8.005 -3.051 0.00 0.00
ATOM  29 C  TYR  2    7.095 -11.604 -4.499 0.00 0.00
ATOM  30 O  TYR  2    6.299 -12.602 -4.557 0.00 0.00
ATOM  31 N  ILE  3    7.618 -11.013 -5.576 0.00 0.00
ATOM  32 H  ILE  3    8.344 -10.332 -5.403 0.00 0.00
ATOM  33 CA ILE  3    7.067 -11.307 -6.897 0.00 0.00
ATOM  34 HA ILE  3    6.639 -12.306 -6.809 0.00 0.00
ATOM  35 CB ILE  3    5.861 -10.461 -7.278 0.00 0.00
ATOM  36 HB ILE  3    5.225 -10.525 -6.394 0.00 0.00
ATOM  37 CG2 ILE  3    6.135 -8.948 -7.348 0.00 0.00
ATOM  38 HG21 ILE  3    5.183 -8.579 -7.730 0.00 0.00
ATOM  39 HG22 ILE  3    6.349 -8.396 -6.433 0.00 0.00
ATOM  40 HG23 ILE  3    6.863 -8.796 -8.145 0.00 0.00
ATOM  41 CG1 ILE  3    4.950 -10.964 -8.432 0.00 0.00
ATOM  42 HG12 ILE  3    4.044 -10.359 -8.470 0.00 0.00
ATOM  43 HG13 ILE  3    5.481 -10.807 -9.370 0.00 0.00
ATOM  44 CD1 ILE  3    4.502 -12.393 -8.276 0.00 0.00
ATOM  45 HD11 ILE  3    3.784 -12.625 -9.063 0.00 0.00
ATOM  46 HD12 ILE  3    5.308 -13.116 -8.395 0.00 0.00
ATOM  47 HD13 ILE  3    3.997 -12.556 -7.323 0.00 0.00
ATOM  48 C  ILE  3    8.076 -11.414 -8.048 0.00 0.00
ATOM  49 O  ILE  3    7.931 -10.956 -9.220 0.00 0.00
ATOM  50 N  GLN  4    9.263 -11.910 -7.699 0.00 0.00
ATOM  51 H  GLN  4    9.288 -12.175 -6.724 0.00 0.00
ATOM  52 CA GLN  4    10.486 -12.035 -8.559 0.00 0.00
ATOM  53 HA GLN  4    11.138 -12.607 -7.899 0.00 0.00
ATOM  54 CB GLN  4    10.235 -12.946 -9.789 0.00 0.00
ATOM  55 HB2 GLN  4    9.457 -12.467 -10.383 0.00 0.00
ATOM  56 HB3 GLN  4    11.125 -13.109 -10.397 0.00 0.00
ATOM  57 CG GLN  4    9.734 -14.411 -9.430 0.00 0.00
ATOM  58 HG2 GLN  4    10.547 -14.934 -8.925 0.00 0.00
ATOM  59 HG3 GLN  4    8.800 -14.300 -8.880 0.00 0.00
ATOM  60 CD GLN  4    9.359 -15.179 -10.705 0.00 0.00
ATOM  61 OE1 GLN  4    8.228 -15.409 -11.101 0.00 0.00
ATOM  62 NE2 GLN  4    10.269 -15.781 -11.436 0.00 0.00
ATOM  63 HE21 GLN  4    9.872 -16.289 -12.214 0.00 0.00
ATOM  64 HE22 GLN  4    11.160 -15.912 -10.978 0.00 0.00
ATOM  65 C  GLN  4    11.124 -10.627 -8.869 0.00 0.00
ATOM  66 O  GLN  4    11.801 -10.396 -9.891 0.00 0.00
ATOM  67 N  ASN  5    11.061 -9.674 -7.941 0.00 0.00
ATOM  68 H  ASN  5    10.594 -10.005 -7.109 0.00 0.00
ATOM  69 CA ASN  5    11.349 -8.226 -8.131 0.00 0.00
ATOM  70 HA ASN  5    11.689 -7.962 -9.132 0.00 0.00
ATOM  71 CB ASN  5    10.013 -7.420 -7.840 0.00 0.00
ATOM  72 HB2 ASN  5    10.191 -6.353 -7.974 0.00 0.00
ATOM  73 HB3 ASN  5    9.262 -7.608 -8.608 0.00 0.00
ATOM  74 CG ASN  5    9.408 -7.567 -6.464 0.00 0.00
ATOM  75 OD1 ASN  5    9.346 -8.695 -5.967 0.00 0.00
ATOM  76 ND2 ASN  5    8.975 -6.536 -5.866 0.00 0.00
ATOM  77 HD21 ASN  5    8.865 -6.638 -4.867 0.00 0.00
ATOM  78 HD22 ASN  5    8.972 -5.613 -6.277 0.00 0.00
ATOM  79 C  ASN  5    12.512 -7.736 -7.203 0.00 0.00
ATOM  80 O  ASN  5    12.761 -6.530 -7.147 0.00 0.00
ATOM  81 N  CYX  6    13.201 -8.603 -6.561 0.00 0.00
ATOM  82 H  CYX  6    13.032 -9.585 -6.724 0.00 0.00
ATOM  83 CA CYX  6    14.123 -8.315 -5.517 0.00 0.00
ATOM  84 HA CYX  6    13.768 -7.462 -4.939 0.00 0.00
ATOM  85 CB CYX  6    14.101 -9.498 -4.497 0.00 0.00
ATOM  86 HB2 CYX  6    13.069 -9.827 -4.379 0.00 0.00
ATOM  87 HB3 CYX  6    14.792 -10.179 -4.994 0.00 0.00
ATOM  88 SG CYX  6    14.636 -9.175 -2.808 0.00 0.00
ATOM  89 C  CYX  6    15.615 -8.088 -6.023 0.00 0.00
ATOM  90 O  CYX  6    16.066 -8.887 -6.838 0.00 0.00
ATOM  91 N  PRO  7    16.423 -7.064 -5.732 0.00 0.00
ATOM  92 CD PRO  7    15.966 -6.065 -4.873 0.00 0.00
ATOM  93 HD2 PRO  7    15.934 -6.320 -3.813 0.00 0.00
ATOM  94 HD3 PRO  7    15.003 -5.656 -5.178 0.00 0.00
ATOM  95 CG PRO  7    16.958 -4.879 -5.038 0.00 0.00
ATOM  96 HG2 PRO  7    17.216 -4.445 -4.072 0.00 0.00
ATOM  97 HG3 PRO  7    16.502 -4.049 -5.578 0.00 0.00
ATOM  98 CB PRO  7    18.169 -5.410 -5.796 0.00 0.00
ATOM  99 HB2 PRO  7    19.052 -5.460 -5.160 0.00 0.00
ATOM 100 HB3 PRO  7    18.331 -4.841 -6.712 0.00 0.00
ATOM 101 CA PRO  7    17.792 -6.847 -6.192 0.00 0.00
ATOM 102 HA PRO  7    17.794 -6.904 -7.280 0.00 0.00
ATOM 103 C  PRO  7    18.763 -7.826 -5.450 0.00 0.00
ATOM 104 O  PRO  7    18.314 -8.597 -4.558 0.00 0.00
ATOM 105 N  LEU  8    20.035 -7.891 -5.883 0.00 0.00
ATOM 106 H  LEU  8    20.254 -7.246 -6.629 0.00 0.00
ATOM 107 CA LEU  8    21.089 -8.596 -5.174 0.00 0.00
ATOM 108 HA LEU  8    20.803 -9.578 -4.799 0.00 0.00
ATOM 109 CB LEU  8    22.281 -8.687 -6.121 0.00 0.00
ATOM 110 HB2 LEU  8    22.475 -7.690 -6.514 0.00 0.00
ATOM 111 HB3 LEU  8    23.168 -8.960 -5.549 0.00 0.00
ATOM 112 CG LEU  8    22.093 -9.636 -7.334 0.00 0.00
ATOM 113 HG LEU  8    21.785 -10.615 -6.967 0.00 0.00
ATOM 114 CD1 LEU  8    21.270 -9.067 -8.429 0.00 0.00
ATOM 115 HD11 LEU  8    21.496 -8.015 -8.604 0.00 0.00
ATOM 116 HD12 LEU  8    21.383 -9.737 -9.282 0.00 0.00
ATOM 117 HD13 LEU  8    20.206 -9.189 -8.229 0.00 0.00
ATOM 118 CD2 LEU  8    23.458 -9.691 -7.970 0.00 0.00
ATOM 119 HD21 LEU  8    23.600 -8.694 -8.387 0.00 0.00
ATOM 120 HD22 LEU  8    24.205 -10.000 -7.238 0.00 0.00
ATOM 121 HD23 LEU  8    23.428 -10.427 -8.774 0.00 0.00
ATOM 122 C  LEU  8    21.382 -7.841 -3.960 0.00 0.00
ATOM 123 O  LEU  8    21.219 -6.611 -3.883 0.00 0.00
ATOM 124 N  GLY  9    21.855 -8.631 -2.954 0.00 0.00
ATOM 125 H  GLY  9    21.910 -9.626 -3.119 0.00 0.00
ATOM 126 CA GLY  9    22.382 -8.208 -1.675 0.00 0.00
ATOM 127 HA2 GLY  9    22.783 -9.047 -1.106 0.00 0.00
ATOM 128 HA3 GLY  9    23.220 -7.555 -1.916 0.00 0.00
ATOM 129 C  GLY  9    21.296 -7.445 -0.881 0.00 0.00
ATOM 130 O  GLY  9    20.148 -7.329 -1.324 0.00 0.00
ATOM 131 N  NHE  10    21.562 -6.802 0.215 0.00 0.00
ATOM 132 HN1 NHE  10    22.470 -6.840 0.656 0.00 0.00
ATOM 133 HN2 NHE  10    20.754 -6.368 0.637 0.00 0.00
TER  134  NHE  10
END
```

CT, saddle (FOLDED)

CT_MD-II_5us_T16_02

```
ATOM   1 HA2 MET  1    6.430 -4.319 -21.835 0.00 0.00
ATOM   2 CA MET  1    6.055 -4.730 -20.898 0.00 0.00
ATOM   3 HA1 MET  1    6.367 -5.762 -20.738 0.00 0.00
ATOM   4 CB MET  1    6.695 -3.827 -19.879 0.00 0.00
ATOM   5 HB2 MET  1    7.656 -3.594 -20.336 0.00 0.00
ATOM   6 HB3 MET  1    6.115 -2.906 -19.820 0.00 0.00
ATOM   7 CG MET  1    6.778 -4.484 -18.504 0.00 0.00
ATOM   8 HG2 MET  1    7.006 -5.529 -18.712 0.00 0.00
ATOM   9 HG3 MET  1    7.590 -3.982 -17.978 0.00 0.00
ATOM  10 C  MET  1    4.527 -4.398 -20.954 0.00 0.00
ATOM  11 O  MET  1    4.116 -3.500 -21.668 0.00 0.00
ATOM  12 N  TYR  2    3.710 -5.183 -20.216 0.00 0.00
ATOM  13 H  TYR  2    4.211 -5.776 -19.571 0.00 0.00
ATOM  14 CA TYR  2    2.296 -5.013 -20.112 0.00 0.00
ATOM  15 HA TYR  2    2.122 -4.076 -20.640 0.00 0.00
ATOM  16 CB TYR  2    1.399 -6.146 -20.685 0.00 0.00
ATOM  17 HB2 TYR  2    0.368 -5.924 -20.409 0.00 0.00
ATOM  18 HB3 TYR  2    1.402 -6.166 -21.775 0.00 0.00
ATOM  19 CG TYR  2    1.536 -7.543 -20.103 0.00 0.00
ATOM  20 CD1 TYR  2    2.778 -8.251 -20.120 0.00 0.00
ATOM  21 HD1 TYR  2    3.719 -7.825 -20.436 0.00 0.00
ATOM  22 CE1 TYR  2    2.892 -9.570 -19.712 0.00 0.00
ATOM  23 HE1 TYR  2    3.864 -10.020 -19.572 0.00 0.00
ATOM  24 CZ TYR  2    1.748 -10.228 -19.196 0.00 0.00
ATOM  25 OS TYR  2    1.939 -11.531 -18.706 0.00 0.00
ATOM  26 CH TYR  2    0.897 -12.117 -17.868 0.00 0.00
ATOM  27 CE2 TYR  2    0.488 -9.541 -19.035 0.00 0.00
ATOM  28 HE2 TYR  2    -0.402 -9.992 -18.622 0.00 0.00
ATOM  29 CD2 TYR  2    0.409 -8.184 -19.474 0.00 0.00
ATOM  30 HD2 TYR  2    -0.591 -7.775 -19.480 0.00 0.00
ATOM  31 C  TYR  2    1.915 -4.689 -18.647 0.00 0.00
ATOM  32 O  TYR  2    2.186 -5.419 -17.696 0.00 0.00
ATOM  33 HH1 TYR  2    -0.029 -12.099 -18.442 0.00 0.00
ATOM  34 HH2 TYR  2    0.842 -11.555 -16.936 0.00 0.00
ATOM  35 HH3 TYR  2    1.119 -13.158 -17.630 0.00 0.00
ATOM  36 N  ILE  3    0.972 -3.779 -18.469 0.00 0.00
ATOM  37 H  ILE  3    0.648 -3.289 -19.291 0.00 0.00
ATOM  38 CA ILE  3    0.470 -3.146 -17.197 0.00 0.00
ATOM  39 HA ILE  3    1.425 -2.899 -16.735 0.00 0.00
ATOM  40 CB ILE  3    -0.301 -1.918 -17.634 0.00 0.00
ATOM  41 HB ILE  3    0.065 -1.507 -18.575 0.00 0.00
ATOM  42 CG2 ILE  3    -1.804 -2.274 -17.925 0.00 0.00
ATOM  48 HG13 ILE  3    0.816 -0.648 -16.261 0.00 0.00
ATOM  49 CD1 ILE  3    -0.827 0.581 -16.856 0.00 0.00
ATOM  50 HD1 ILE  3    -1.914 0.536 -16.928 0.00 0.00
ATOM  51 HD12 ILE  3    -0.566 1.302 -16.085 0.00 0.00
ATOM  52 HD13 ILE  3    -0.550 1.002 -17.823 0.00 0.00
ATOM  53 C  ILE  3    -0.264 -4.032 -16.235 0.00 0.00
ATOM  54 O  ILE  3    0.187 -3.818 -15.041 0.00 0.00
ATOM  55 N  GLN  4    -0.756 -5.283 -16.652 0.00 0.00
ATOM  56 H  GLN  4    -0.886 -5.445 -17.641 0.00 0.00
ATOM  57 CA GLN  4    -1.339 -6.274 -15.736 0.00 0.00
ATOM  58 HA GLN  4    -1.952 -5.[illegible] -15.016 0.00 0.00
ATOM  59 CB GLN  4    -2.167 -7.330 -16.525 0.00 0.00
ATOM  60 HB2 GLN  4    -1.492 -7.696 -17.298 0.00 0.00
ATOM  61 HB3 GLN  4    -2.260 -8.171 -15.837 0.00 0.00
ATOM  62 CG GLN  4    -3.460 -6.720 -17.025 0.00 0.00
ATOM  63 HG2 GLN  4    -4.162 -7.553 -17.063 0.00 0.00
ATOM  64 HG3 GLN  4    -3.785 -5.935 -16.339 0.00 0.00
ATOM  65 CD GLN  4    -3.412 -6.101 -18.404 0.00 0.00
ATOM  66 OE1 GLN  4    -2.357 -5.636 -18.855 0.00 0.00
ATOM  67 NE2 GLN  4    -4.442 -6.365 -19.194 0.00 0.00
ATOM  68 HE21 GLN  4    -4.321 -6.147 -20.173 0.00 0.00
ATOM  69 HE22 GLN  4    -5.308 -6.757 -18.853 0.00 0.00
ATOM  70 C  GLN  4    -0.234 -7.071 -14.966 0.00 0.00
ATOM  71 O  GLN  4    -0.520 -7.717 -13.964 0.00 0.00
ATOM  72 N  ASN  5    1.006 -7.090 -15.535 0.00 0.00
ATOM  73 H  ASN  5    1.161 -6.547 -16.373 0.00 0.00
ATOM  74 CA ASN  5    2.104 -7.903 -15.020 0.00 0.00
ATOM  75 HA ASN  5    1.759 -8.441 -14.137 0.00 0.00
ATOM  76 CB ASN  5    2.320 -9.036 -16.029 0.00 0.00
ATOM  77 HB2 ASN  5    1.429 -9.554 -16.382 0.00 0.00
ATOM  78 HB3 ASN  5    2.820 -8.647 -16.916 0.00 0.00
ATOM  79 CG ASN  5    3.173 -10.101 -15.448 0.00 0.00
ATOM  80 OD1 ASN  5    4.367 -10.243 -15.806 0.00 0.00
ATOM  81 ND2 ASN  5    2.588 -10.930 -14.601 0.00 0.00
ATOM  82 HD21 ASN  5    3.207 -11.621 -14.202 0.00 0.00
ATOM  83 HD22 ASN  5    1.599 -10.878 -14.399 0.00 0.00
ATOM  84 C  ASN  5    3.299 -7.057 -14.540 0.00 0.00
ATOM  85 O  ASN  5    3.718 -7.300 -13.403 0.00 0.00
ATOM  86 N  CYX  6    3.739 -6.069 -15.287 0.00 0.00
ATOM  87 H  CYX  6    3.144 -5.667 -15.996 0.00 0.00
ATOM  88 CA CYX  6    4.852 -5.203 -14.822 0.00 0.00
ATOM  89 HA CYX  6    5.275 -5.529 -13.872 0.00 0.00
ATOM  95 O  CYX  6    3.343 -3.357 -15.250 0.00 0.00
ATOM  96 N  PRO  7    5.032 -2.864 -13.882 0.00 0.00
ATOM  97 CD PRO  7    6.040 -3.071 -12.904 0.00 0.00
ATOM  98 HD2 PRO  7    6.888 -3.541 -13.403 0.00 0.00
ATOM  99 HD3 PRO  7    5.767 -3.669 -12.035 0.00 0.00
ATOM 100 CG PRO  7    6.455 -1.658 -12.397 0.00 0.00
ATOM 101 HG2 PRO  7    7.388 -1.289 -12.824 0.00 0.00
ATOM 102 HG3 PRO  7    6.513 -1.796 -11.317 0.00 0.00
ATOM 103 CB PRO  7    5.281 -0.754 -12.607 0.00 0.00
ATOM 104 HB2 PRO  7    5.619 0.257 -12.833 0.00 0.00
ATOM 105 HB3 PRO  7    4.618 -0.688 -11.744 0.00 0.00
ATOM 106 CA PRO  7    4.585 -1.391 -13.843 0.00 0.00
ATOM 107 HA PRO  7    3.500 -1.336 -13.752 0.00 0.00
ATOM 108 C  PRO  7    5.059 -0.753 -15.177 0.00 0.00
ATOM 109 O  PRO  7    6.192 -0.813 -15.564 0.00 0.00
ATOM 110 N  LEU  8    4.140 -0.130 -15.885 0.00 0.00
ATOM 111 H  LEU  8    3.160 -0.199 -15.648 0.00 0.00
ATOM 112 CA LEU  8    4.416 0.496 -17.219 0.00 0.00
ATOM 113 HA LEU  8    5.162 -0.162 -17.664 0.00 0.00
ATOM 114 CB LEU  8    3.101 0.508 -18.087 0.00 0.00
ATOM 115 HB2 LEU  8    2.515 -0.396 -17.918 0.00 0.00
ATOM 116 HB3 LEU  8    2.430 1.272 -17.697 0.00 0.00
ATOM 117 CG LEU  8    3.357 0.627 -19.578 0.00 0.00
ATOM 118 HG LEU  8    3.796 1.612 -19.740 0.00 0.00
ATOM 119 CD1 LEU  8    3.988 -0.640 -20.182 0.00 0.00
ATOM 120 HD11 LEU  8    5.041 -0.576 -19.905 0.00 0.00
ATOM 121 HD12 LEU  8    3.564 -1.573 -19.811 0.00 0.00
ATOM 122 HD13 LEU  8    3.855 -0.584 -21.263 0.00 0.00
ATOM 123 CD2 LEU  8    2.026 0.794 -20.303 0.00 0.00
ATOM 124 HD21 LEU  8    1.381 1.507 -19.789 0.00 0.00
ATOM 125 HD22 LEU  8    2.156 1.229 -21.294 0.00 0.00
ATOM 126 HD23 LEU  8    1.394 -0.089 -20.397 0.00 0.00
ATOM 127 C  LEU  8    5.037 1.836 -17.141 0.00 0.00
ATOM 128 O  LEU  8    5.642 2.363 -18.088 0.00 0.00
ATOM 129 N  GLY  9    5.227 2.455 -15.925 0.00 0.00
ATOM 130 H  GLY  9    4.713 2.009 -15.178 0.00 0.00
ATOM 131 CA GLY  9    6.026 3.619 -15.711 0.00 0.00
ATOM 132 HA2 GLY  9    6.016 3.756 -14.629 0.00 0.00
ATOM 133 HA3 GLY  9    7.065 3.373 -15.932 0.00 0.00
ATOM 134 C  GLY  9    5.576 4.904 -16.426 0.00 0.00
ATOM 135 O  GLY  9    4.597 4.994 -17.210 0.00 0.00
ATOM 136 N  NHE  10    6.114 6.051 -16.082 0.00 0.00
```

ATOM 43 HG21 ILE 3 -2.317 -1.484 -18.474 0.00 0.00
ATOM 44 HG22 ILE 3 -1.894 -3.208 -18.480 0.00 0.00
ATOM 45 HG23 ILE 3 -2.264 -2.426 -16.949 0.00 0.00
ATOM 46 CG1 ILE 3 -0.215 -0.838 -16.557 0.00 0.00
ATOM 47 HG12 ILE 3 -0.785 -1.161 -15.685 0.00 0.00
ATOM 90 CB CYX 6 5.890 -5.227 -15.895 0.00 0.00
ATOM 91 HB2 CYX 6 6.811 -4.792 -15.508 0.00 0.00
ATOM 92 HB3 CYX 6 5.985 -6.297 -16.081 0.00 0.00
ATOM 93 SG CYX 6 5.357 -4.362 -17.431 0.00 0.00
ATOM 94 C CYX 6 4.321 -3.730 -14.647 0.00 0.00
ATOM 137 HN1 NHE 10 6.945 6.100 -15.511 0.00 0.00
ATOM 138 HN2 NHE 10 5.612 6.881 -16.364 0.00 0.00
TER 139 NHE 10
END

CT, clinched open45pbr (OPEN)

CT_MD-I_5us_T16_8

ATOM 1 HA2 MET 1 7.269 5.503 7.705 0.00 0.00
ATOM 2 CA MET 1 7.278 5.736 6.640 0.00 0.00
ATOM 3 HA1 MET 1 6.267 5.486 6.318 0.00 0.00
ATOM 4 CB MET 1 7.321 7.254 6.391 0.00 0.00
ATOM 5 HB2 MET 1 8.194 7.600 6.944 0.00 0.00
ATOM 6 HB3 MET 1 7.463 7.415 5.322 0.00 0.00
ATOM 7 CG MET 1 6.061 7.944 6.801 0.00 0.00
ATOM 8 HG2 MET 1 5.222 7.724 6.141 0.00 0.00
ATOM 9 HG3 MET 1 5.791 7.694 7.828 0.00 0.00
ATOM 10 C MET 1 8.463 4.986 5.980 0.00 0.00
ATOM 11 O MET 1 9.603 4.956 6.522 0.00 0.00
ATOM 12 N TYR 2 8.259 4.471 4.774 0.00 0.00
ATOM 13 H TYR 2 7.466 4.924 4.342 0.00 0.00
ATOM 14 CA TYR 2 9.350 3.861 3.953 0.00 0.00
ATOM 15 HA TYR 2 9.942 3.263 4.645 0.00 0.00
ATOM 16 CB TYR 2 8.806 2.873 2.912 0.00 0.00
ATOM 17 HB2 TYR 2 9.455 2.009 2.769 0.00 0.00
ATOM 18 HB3 TYR 2 7.880 2.502 3.350 0.00 0.00
ATOM 19 CG TYR 2 8.479 3.416 1.573 0.00 0.00
ATOM 20 CD1 TYR 2 7.310 4.100 1.388 0.00 0.00
ATOM 21 HD1 TYR 2 6.564 4.241 2.156 0.00 0.00
ATOM 22 CE1 TYR 2 6.983 4.712 0.224 0.00 0.00
ATOM 23 HE1 TYR 2 6.062 5.267 0.120 0.00 0.00
ATOM 24 CZ TYR 2 7.857 4.682 -0.930 0.00 0.00
ATOM 25 OS TYR 2 7.500 5.249 -2.194 0.00 0.00
ATOM 26 CH TYR 2 6.136 5.842 -2.334 0.00 0.00
ATOM 27 CE2 TYR 2 9.038 3.929 -0.807 0.00 0.00
ATOM 28 HE2 TYR 2 9.671 3.796 -1.672 0.00 0.00
ATOM 29 CD2 TYR 2 9.356 3.299 0.442 0.00 0.00
ATOM 30 HD2 TYR 2 10.314 2.805 0.505 0.00 0.00
ATOM 31 C TYR 2 10.303 4.923 3.267 0.00 0.00
ATOM 32 O TYR 2 11.397 4.590 2.753 0.00 0.00
ATOM 33 HH1 TYR 2 5.350 5.116 -2.127 0.00 0.00
ATOM 34 HH2 TYR 2 6.116 6.136 -3.384 0.00 0.00
ATOM 35 HH3 TYR 2 6.027 6.756 -1.750 0.00 0.00
ATOM 36 N ILE 3 9.941 6.219 3.258 0.00 0.00
ATOM 37 H ILE 3 9.183 6.443 3.887 0.00 0.00
ATOM 38 CA ILE 3 10.835 7.360 2.772 0.00 0.00
ATOM 39 HA ILE 3 11.773 6.904 2.456 0.00 0.00
ATOM 40 CB ILE 3 10.237 8.024 1.498 0.00 0.00
ATOM 41 HB ILE 3 10.792 8.945 1.319 0.00 0.00
ATOM 42 CG2 ILE 3 10.453 7.245 0.170 0.00 0.00
ATOM 43 HG21 ILE 3 11.470 6.857 0.114 0.00 0.00
ATOM 44 HG22 ILE 3 9.885 6.315 0.194 0.00 0.00
ATOM 45 HG23 ILE 3 10.112 7.806 -0.700 0.00 0.00
ATOM 46 CG1 ILE 3 8.773 8.426 1.846 0.00 0.00
ATOM 47 HG12 ILE 3 8.076 7.607 2.027 0.00 0.00
ATOM 48 HG13 ILE 3 8.839 9.049 2.738 0.00 0.00
ATOM 49 CD1 ILE 3 8.274 9.337 0.769 0.00 0.00
ATOM 50 HD11 ILE 3 8.140 8.602 -0.025 0.00 0.00
ATOM 51 HD12 ILE 3 7.318 9.801 1.011 0.00 0.00
ATOM 52 HD13 ILE 3 9.063 10.066 0.588 0.00 0.00
ATOM 53 C ILE 3 11.219 8.385 3.823 0.00 0.00
ATOM 54 O ILE 3 10.575 8.395 4.889 0.00 0.00
ATOM 55 N GLN 4 12.289 9.161 3.578 0.00 0.00
ATOM 56 H GLN 4 12.707 9.039 2.666 0.00 0.00
ATOM 57 CA GLN 4 12.560 10.380 4.280 0.00 0.00
ATOM 58 HA GLN 4 12.343 10.158 5.325 0.00 0.00
ATOM 59 CB GLN 4 13.953 10.933 3.895 0.00 0.00
ATOM 60 HB2 GLN 4 14.075 10.961 2.812 0.00 0.00
ATOM 61 HB3 GLN 4 14.062 11.955 4.259 0.00 0.00
ATOM 62 CG GLN 4 15.067 10.036 4.502 0.00 0.00
ATOM 63 HG2 GLN 4 14.874 9.946 5.571 0.00 0.00
ATOM 64 HG3 GLN 4 15.116 9.060 4.018 0.00 0.00
ATOM 65 CD GLN 4 16.456 10.713 4.441 0.00 0.00
ATOM 66 OE1 GLN 4 17.333 10.299 3.695 0.00 0.00
ATOM 67 NE2 GLN 4 16.820 11.657 5.269 0.00 0.00
ATOM 68 HE21 GLN 4 17.740 12.075 5.257 0.00 0.00
ATOM 69 HE22 GLN 4 16.097 12.191 5.731 0.00 0.00
ATOM 70 C GLN 4 11.493 11.473 3.866 0.00 0.00
ATOM 71 O GLN 4 11.094 11.500 2.700 0.00 0.00
ATOM 72 N ASN 5 11.135 12.366 4.765 0.00 0.00
ATOM 73 H ASN 5 11.596 12.269 5.659 0.00 0.00
ATOM 74 CA ASN 5 10.114 13.381 4.561 0.00 0.00
ATOM 75 HA ASN 5 10.000 13.820 5.552 0.00 0.00
ATOM 76 CB ASN 5 10.669 14.440 3.618 0.00 0.00
ATOM 77 HB2 ASN 5 11.732 14.643 3.753 0.00 0.00
ATOM 78 HB3 ASN 5 10.629 14.077 2.591 0.00 0.00
ATOM 79 CG ASN 5 10.080 15.824 3.677 0.00 0.00
ATOM 80 OD1 ASN 5 9.176 16.080 4.431 0.00 0.00
ATOM 81 ND2 ASN 5 10.491 16.732 2.827 0.00 0.00
ATOM 82 HD21 ASN 5 10.099 17.658 2.923 0.00 0.00
ATOM 83 HD22 ASN 5 11.275 16.478 2.243 0.00 0.00
ATOM 84 C ASN 5 8.688 12.871 4.095 0.00 0.00
ATOM 85 O ASN 5 8.045 13.592 3.346 0.00 0.00
ATOM 86 N CYX 6 8.186 11.732 4.640 0.00 0.00
ATOM 87 H CYX 6 8.667 11.143 5.305 0.00 0.00
ATOM 88 CA CYX 6 6.862 11.326 4.314 0.00 0.00
ATOM 89 HA CYX 6 6.823 11.280 3.226 0.00 0.00
ATOM 90 CB CYX 6 6.602 9.923 4.867 0.00 0.00
ATOM 91 HB2 CYX 6 5.699 9.742 4.284 0.00 0.00
ATOM 92 HB3 CYX 6 7.326 9.173 4.549 0.00 0.00
ATOM 93 SG CYX 6 6.316 9.777 6.642 0.00 0.00
ATOM 94 C CYX 6 5.782 12.400 4.835 0.00 0.00
ATOM 95 O CYX 6 5.956 13.013 5.938 0.00 0.00
ATOM 96 N PRO 7 4.821 12.727 3.943 0.00 0.00
ATOM 97 CD PRO 7 4.626 12.305 2.551 0.00 0.00
ATOM 98 HD2 PRO 7 4.959 11.269 2.499 0.00 0.00
ATOM 99 HD3 PRO 7 5.017 12.956 1.770 0.00 0.00
ATOM 100 CG PRO 7 3.046 12.318 2.428 0.00 0.00
ATOM 101 HG2 PRO 7 2.655 11.416 2.900 0.00 0.00
ATOM 102 HG3 PRO 7 2.731 12.185 1.393 0.00 0.00
ATOM 103 CB PRO 7 2.738 13.594 3.142 0.00 0.00
ATOM 104 HB2 PRO 7 1.683 13.583 3.416 0.00 0.00
ATOM 105 HB3 PRO 7 3.058 14.489 2.608 0.00 0.00
ATOM 106 CA PRO 7 3.712 13.571 4.367 0.00 0.00
ATOM 107 HA PRO 7 4.078 14.569 4.605 0.00 0.00
ATOM 108 C PRO 7 2.908 13.126 5.585 0.00 0.00
ATOM 109 O PRO 7 2.963 11.879 5.949 0.00 0.00
ATOM 110 N LEU 8 2.218 14.062 6.323 0.00 0.00
ATOM 111 H LEU 8 1.986 14.910 5.826 0.00 0.00
ATOM 112 CA LEU 8 1.514 13.782 7.577 0.00 0.00
ATOM 113 HA LEU 8 1.831 12.818 7.974 0.00 0.00
ATOM 114 CB LEU 8 1.955 14.934 8.516 0.00 0.00
ATOM 115 HB2 LEU 8 3.016 14.706 8.614 0.00 0.00
ATOM 116 HB3 LEU 8 1.911 15.847 7.921 0.00 0.00
ATOM 117 CG LEU 8 1.318 15.246 9.896 0.00 0.00
ATOM 118 HG LEU 8 0.245 15.415 9.798 0.00 0.00
ATOM 119 CD1 LEU 8 1.625 14.083 10.878 0.00 0.00
ATOM 120 HD11 LEU 8 1.055 14.285 11.785 0.00 0.00
ATOM 121 HD12 LEU 8 1.279 13.121 10.500 0.00 0.00
ATOM 122 HD13 LEU 8 2.693 14.031 11.093 0.00 0.00
ATOM 123 CD2 LEU 8 1.931 16.560 10.419 0.00 0.00
ATOM 124 HD21 LEU 8 2.995 16.416 10.608 0.00 0.00
ATOM 125 HD22 LEU 8 1.758 17.352 9.690 0.00 0.00
ATOM 126 HD23 LEU 8 1.485 16.900 11.353 0.00 0.00
ATOM 127 C LEU 8 0.014 13.623 7.461 0.00 0.00
ATOM 128 O LEU 8 -0.699 14.351 6.729 0.00 0.00
ATOM 129 N GLY 9 -0.513 12.529 8.060 0.00 0.00
ATOM 130 H GLY 9 0.105 11.910 8.565 0.00 0.00
ATOM 131 CA GLY 9 -1.974 12.274 8.074 0.00 0.00
ATOM 132 HA2 GLY 9 -2.488 12.514 7.143 0.00 0.00
ATOM 133 HA3 GLY 9 -2.110 11.205 8.237 0.00 0.00
ATOM 134 C GLY 9 -2.758 13.071 9.130 0.00 0.00
ATOM 135 O GLY 9 -2.246 13.956 9.843 0.00 0.00
ATOM 136 N NHE 10 -4.010 12.754 9.194 0.00 0.00
ATOM 137 HN1 NHE 10 -4.405 12.227 8.429 0.00 0.00
ATOM 138 HN2 NHE 10 -4.607 13.312 9.788 0.00 0.00
TER 139 NHE 10
END

CT, open (OPEN)

CT_MD-II_5us_T16_9

ATOM 1 HA2 MET 1 -16.950 -16.076 5.912 0.00 0.00
ATOM 2 CA MET 1 -16.029 -16.413 5.436 0.00 0.00
ATOM 3 HA1 MET 1 -16.183 -17.450 5.138 0.00 0.00
ATOM 4 CB MET 1 -15.755 -15.615 4.113 0.00 0.00
ATOM 5 HB2 MET 1 -14.749 -15.935 3.840 0.00 0.00
ATOM 6 HB3 MET 1 -16.324 -16.032 3.282 0.00 0.00
ATOM 7 CG MET 1 -15.995 -14.127 4.287 0.00 0.00
ATOM 8 HG2 MET 1 -17.014 -14.017 4.657 0.00 0.00
ATOM 9 HG3 MET 1 -15.370 -13.843 5.134 0.00 0.00
ATOM 10 C MET 1 -14.955 -16.234 6.442 0.00 0.00
ATOM 11 O MET 1 -14.968 -15.282 7.266 0.00 0.00
ATOM 12 N TYR 2 -13.951 -17.128 6.388 0.00 0.00
ATOM 13 H TYR 2 -14.014 -17.914 5.757 0.00 0.00
ATOM 14 CA TYR 2 -12.998 -17.433 7.517 0.00 0.00
ATOM 15 HA TYR 2 -13.521 -17.122 8.422 0.00 0.00
ATOM 16 CB TYR 2 -12.710 -18.950 7.654 0.00 0.00
ATOM 17 HB2 TYR 2 -13.682 -19.444 7.647 0.00 0.00
ATOM 18 HB3 TYR 2 -12.206 -19.245 6.734 0.00 0.00
ATOM 19 CG TYR 2 -11.901 -19.379 8.889 0.00 0.00
ATOM 20 CD1 TYR 2 -10.780 -20.168 8.727 0.00 0.00
ATOM 21 HD1 TYR 2 -10.452 -20.500 7.753 0.00 0.00
ATOM 22 CE1 TYR 2 -10.159 -20.774 9.850 0.00 0.00
ATOM 23 HE1 TYR 2 -9.294 -21.399 9.683 0.00 0.00
ATOM 24 CZ TYR 2 -10.478 -20.412 11.201 0.00 0.00
ATOM 25 OS TYR 2 -9.600 -20.914 12.203 0.00 0.00
ATOM 26 CH TYR 2 -9.463 -20.122 13.413 0.00 0.00
ATOM 27 CE2 TYR 2 -11.525 -19.459 11.372 0.00 0.00
ATOM 28 HE2 TYR 2 -11.788 -19.198 12.386 0.00 0.00
ATOM 29 CD2 TYR 2 -12.207 -18.943 10.248 0.00 0.00
ATOM 30 HD2 TYR 2 -13.063 -18.293 10.356 0.00 0.00
ATOM 31 C TYR 2 -11.748 -16.568 7.470 0.00 0.00
ATOM 32 O TYR 2 -11.072 -16.729 6.440 0.00 0.00
ATOM 33 HH1 TYR 2 -8.715 -20.452 14.135 0.00 0.00
ATOM 34 HH2 TYR 2 -10.412 -20.101 13.948 0.00 0.00
ATOM 48 HG13 ILE 3 -7.997 -16.499 7.724 0.00 0.00
ATOM 49 CD1 ILE 3 -7.584 -17.820 9.392 0.00 0.00
ATOM 50 HD11 ILE 3 -6.807 -17.232 9.880 0.00 0.00
ATOM 51 HD12 ILE 3 -8.205 -18.336 10.124 0.00 0.00
ATOM 52 HD13 ILE 3 -7.133 -18.523 8.692 0.00 0.00
ATOM 53 C ILE 3 -9.908 -14.065 7.398 0.00 0.00
ATOM 54 O ILE 3 -8.700 -13.822 7.094 0.00 0.00
ATOM 55 N GLN 4 -10.910 -13.607 6.663 0.00 0.00
ATOM 56 H GLN 4 -11.860 -13.832 6.922 0.00 0.00
ATOM 57 CA GLN 4 -10.714 -12.804 5.473 0.00 0.00
ATOM 58 HA GLN 4 -9.712 -12.376 5.459 0.00 0.00
ATOM 59 CB GLN 4 -10.789 -13.671 4.197 0.00 0.00
ATOM 60 HB2 GLN 4 -10.634 -13.023 3.334 0.00 0.00
ATOM 61 HB3 GLN 4 -10.019 -14.440 4.132 0.00 0.00
ATOM 62 CG GLN 4 -12.049 -14.571 4.122 0.00 0.00
ATOM 63 HG2 GLN 4 -12.328 -14.947 5.107 0.00 0.00
ATOM 64 HG3 GLN 4 -12.826 -13.999 3.616 0.00 0.00
ATOM 65 CD GLN 4 -11.852 -15.750 3.133 0.00 0.00
ATOM 66 OE1 GLN 4 -12.236 -15.612 1.959 0.00 0.00
ATOM 67 NE2 GLN 4 -11.311 -16.833 3.643 0.00 0.00
ATOM 68 HE21 GLN 4 -11.132 -17.589 2.998 0.00 0.00
ATOM 69 HE22 GLN 4 -11.295 -16.868 4.652 0.00 0.00
ATOM 70 C GLN 4 -11.745 -11.730 5.401 0.00 0.00
ATOM 71 O GLN 4 -12.854 -11.844 5.871 0.00 0.00
ATOM 72 N ASN 5 -11.431 -10.664 4.674 0.00 0.00
ATOM 73 H ASN 5 -10.540 -10.616 4.202 0.00 0.00
ATOM 74 CA ASN 5 -12.377 -9.552 4.452 0.00 0.00
ATOM 75 HA ASN 5 -12.788 -9.451 5.457 0.00 0.00
ATOM 76 CB ASN 5 -11.487 -8.325 4.039 0.00 0.00
ATOM 77 HB2 ASN 5 -10.701 -8.137 4.772 0.00 0.00
ATOM 78 HB3 ASN 5 -10.966 -8.595 3.120 0.00 0.00
ATOM 79 CG ASN 5 -12.249 -6.975 3.860 0.00 0.00
ATOM 80 OD1 ASN 5 -13.415 -6.789 4.227 0.00 0.00
ATOM 81 ND2 ASN 5 -11.612 -5.924 3.423 0.00 0.00
ATOM 95 O CYX 6 -17.093 -8.510 4.036 0.00 0.00
ATOM 96 N PRO 7 -17.759 -9.027 1.916 0.00 0.00
ATOM 97 CD PRO 7 -17.801 -10.000 0.853 0.00 0.00
ATOM 98 HD2 PRO 7 -17.929 -11.022 1.210 0.00 0.00
ATOM 99 HD3 PRO 7 -16.883 -9.968 0.266 0.00 0.00
ATOM 100 CG PRO 7 -18.996 -9.534 0.055 0.00 0.00
ATOM 101 HG2 PRO 7 -19.881 -10.088 0.369 0.00 0.00
ATOM 102 HG3 PRO 7 -18.859 -9.731 -1.008 0.00 0.00
ATOM 103 CB PRO 7 -19.184 -8.056 0.387 0.00 0.00
ATOM 104 HB2 PRO 7 -20.220 -7.765 0.218 0.00 0.00
ATOM 105 HB3 PRO 7 -18.466 -7.585 -0.285 0.00 0.00
ATOM 106 CA PRO 7 -18.585 -7.885 1.782 0.00 0.00
ATOM 107 HA PRO 7 -17.961 -6.997 1.892 0.00 0.00
ATOM 108 C PRO 7 -19.754 -7.872 2.775 0.00 0.00
ATOM 109 O PRO 7 -20.365 -8.943 3.008 0.00 0.00
ATOM 110 N LEU 8 -20.091 -6.692 3.391 0.00 0.00
ATOM 111 H LEU 8 -19.626 -5.869 3.034 0.00 0.00
ATOM 112 CA LEU 8 -21.114 -6.576 4.499 0.00 0.00
ATOM 113 HA LEU 8 -21.315 -7.554 4.937 0.00 0.00
ATOM 114 CB LEU 8 -20.506 -5.697 5.560 0.00 0.00
ATOM 115 HB2 LEU 8 -20.160 -4.844 4.976 0.00 0.00
ATOM 116 HB3 LEU 8 -21.272 -5.386 6.270 0.00 0.00
ATOM 117 CG LEU 8 -19.340 -6.300 6.361 0.00 0.00
ATOM 118 HG LEU 8 -18.563 -6.488 5.621 0.00 0.00
ATOM 119 CD1 LEU 8 -18.876 -5.312 7.384 0.00 0.00
ATOM 120 HD11 LEU 8 -18.447 -4.496 6.802 0.00 0.00
ATOM 121 HD12 LEU 8 -19.730 -4.868 7.895 0.00 0.00
ATOM 122 HD13 LEU 8 -18.115 -5.626 8.098 0.00 0.00
ATOM 123 CD2 LEU 8 -19.684 -7.704 6.968 0.00 0.00
ATOM 124 HD21 LEU 8 -20.664 -7.651 7.444 0.00 0.00
ATOM 125 HD22 LEU 8 -19.697 -8.444 6.168 0.00 0.00
ATOM 126 HD23 LEU 8 -18.837 -7.989 7.591 0.00 0.00
ATOM 127 C LEU 8 -22.498 -6.117 4.026 0.00 0.00
ATOM 128 O LEU 8 -22.697 -5.126 3.303 0.00 0.00

```
ATOM  35 HH3 TYR   2    -9.182 -19.086  13.223  0.00  0.00
ATOM  36 N   ILE   3   -11.508 -15.776   8.559  0.00  0.00
ATOM  37 H   ILE   3   -12.117 -15.866   9.359  0.00  0.00
ATOM  38 CA  ILE   3   -10.261 -14.912   8.681  0.00  0.00
ATOM  39 HA  ILE   3   -10.452 -14.063   9.337  0.00  0.00
ATOM  40 CB  ILE   3    -9.163 -15.726   9.385  0.00  0.00
ATOM  41 HB  ILE   3    -8.348 -15.003   9.424  0.00  0.00
ATOM  42 CG2 ILE   3    -9.692 -15.996  10.870  0.00  0.00
ATOM  43 HG21 ILE  3    -8.818 -16.171  11.497  0.00  0.00
ATOM  44 HG22 ILE  3   -10.143 -15.061  11.201  0.00  0.00
ATOM  45 HG23 ILE  3   -10.417 -16.810  10.861  0.00  0.00
ATOM  46 CG1 ILE   3    -8.490 -16.826   8.640  0.00  0.00
ATOM  47 HG12 ILE  3    -9.267 -17.439   8.183  0.00  0.00
ATOM  82 HD21 ASN  5   -12.173  -5.096   3.282  0.00  0.00
ATOM  83 HD22 ASN  5   -10.735  -6.045   2.937  0.00  0.00
ATOM  84 C   ASN   5   -13.449  -9.851   3.341  0.00  0.00
ATOM  85 O   ASN   5   -13.127 -10.154   2.200  0.00  0.00
ATOM  86 N   CYX   6   -14.721  -9.897   3.753  0.00  0.00
ATOM  87 H   CYX   6   -14.876  -9.902   4.751  0.00  0.00
ATOM  88 CA  CYX   6   -15.855 -10.228   2.960  0.00  0.00
ATOM  89 HA  CYX   6   -15.555 -10.330   1.917  0.00  0.00
ATOM  90 CB  CYX   6   -16.368 -11.541   3.446  0.00  0.00
ATOM  91 HB2 CYX   6   -16.309 -11.653   4.529  0.00  0.00
ATOM  92 HB3 CYX   6   -17.429 -11.657   3.224  0.00  0.00
ATOM  93 SG  CYX   6   -15.624 -13.129   2.888  0.00  0.00
ATOM  94 C   CYX   6   -16.996  -9.217   3.025  0.00  0.00
ATOM 129 N   GLY   9   -23.487  -6.789   4.506  0.00  0.00
ATOM 130 H   GLY   9   -23.360  -7.669   4.985  0.00  0.00
ATOM 131 CA  GLY   9   -24.867  -6.497   4.184  0.00  0.00
ATOM 132 HA2 GLY   9   -25.075  -6.499   3.114  0.00  0.00
ATOM 133 HA3 GLY   9   -25.504  -7.298   4.559  0.00  0.00
ATOM 134 C   GLY   9   -25.429  -5.172   4.844  0.00  0.00
ATOM 135 O   GLY   9   -24.862  -4.820   5.886  0.00  0.00
ATOM 136 N   NHE  10   -26.434  -4.556   4.279  0.00  0.00
ATOM 137 HN1 NHE  10   -26.820  -4.830   3.387  0.00  0.00
ATOM 138 HN2 NHE  10   -26.791  -3.737   4.750  0.00  0.00
TER  139      NHE  10
END
```

CT, open23pbr (OPEN)

CT_MD-I_5us_T16_2

```
ATOM   1 HA2 MET   1     5.993   6.076 -11.773  0.00  0.00
ATOM   2 CA  MET   1     6.808   6.736 -11.476  0.00  0.00
ATOM   3 HA1 MET   1     6.566   7.206 -10.523  0.00  0.00
ATOM   4 CB  MET   1     8.124   5.953 -11.374  0.00  0.00
ATOM   5 HB2 MET   1     8.164   5.339 -12.274  0.00  0.00
ATOM   6 HB3 MET   1     9.015   6.580 -11.394  0.00  0.00
ATOM   7 CG  MET   1     8.166   5.109 -10.153  0.00  0.00
ATOM   8 HG2 MET   1     8.029   5.737  -9.272  0.00  0.00
ATOM   9 HG3 MET   1     7.327   4.418 -10.232  0.00  0.00
ATOM  10 C   MET   1     6.930   7.752 -12.558  0.00  0.00
ATOM  11 O   MET   1     6.423   7.593 -13.740  0.00  0.00
ATOM  12 N   TYR   2     7.523   8.923 -12.194  0.00  0.00
ATOM  13 H   TYR   2     8.001   9.003 -11.308  0.00  0.00
ATOM  14 CA  TYR   2     7.435  10.140 -12.993  0.00  0.00
ATOM  15 HA  TYR   2     7.035   9.768 -13.936  0.00  0.00
ATOM  16 CB  TYR   2     6.459  11.061 -12.385  0.00  0.00
ATOM  17 HB2 TYR   2     5.711  11.401 -13.101  0.00  0.00
ATOM  18 HB3 TYR   2     5.751  10.535 -11.743  0.00  0.00
ATOM  19 CG  TYR   2     6.953  12.302 -11.688  0.00  0.00
ATOM  20 CD1 TYR   2     7.206  13.385 -12.515  0.00  0.00
ATOM  21 HD1 TYR   2     7.017  13.332 -13.577  0.00  0.00
ATOM  22 CE1 TYR   2     7.580  14.609 -12.015  0.00  0.00
ATOM  23 HE1 TYR   2     7.583  15.390 -12.760  0.00  0.00
ATOM  24 CZ  TYR   2     7.935  14.752 -10.707  0.00  0.00
ATOM  25 OS  TYR   2     8.354  15.980 -10.165  0.00  0.00
ATOM  26 CH  TYR   2     8.453  17.087 -11.009  0.00  0.00
ATOM  27 CE2 TYR   2     7.806  13.652  -9.884  0.00  0.00
ATOM  28 HE2 TYR   2     8.236  13.681  -8.893  0.00  0.00
ATOM  29 CD2 TYR   2     7.159  12.434 -10.336  0.00  0.00
ATOM  30 HD2 TYR   2     6.972  11.616  -9.655  0.00  0.00
ATOM  31 C   TYR   2     8.771  10.798 -13.370  0.00  0.00
ATOM  32 O   TYR   2     8.898  11.306 -14.489  0.00  0.00
ATOM  33 HH1 TYR   2     8.771  17.974 -10.461  0.00  0.00
ATOM  34 HH2 TYR   2     7.509  17.315 -11.503  0.00  0.00
ATOM  35 HH3 TYR   2     9.216  17.025 -11.785  0.00  0.00
ATOM  36 N   ILE   3     9.728  10.858 -12.401  0.00  0.00
ATOM  37 H   ILE   3     9.343  10.645 -11.492  0.00  0.00
ATOM  38 CA  ILE   3    11.071  11.381 -12.522  0.00  0.00
ATOM  39 HA  ILE   3    11.087  11.534 -13.601  0.00  0.00
ATOM  40 CB  ILE   3    11.165  12.846 -11.959  0.00  0.00
ATOM  41 HB  ILE   3    10.292  13.375 -12.340  0.00  0.00
ATOM  42 CG2 ILE   3    11.231  12.861 -10.402  0.00  0.00
ATOM  43 HG21 ILE  3    10.570  12.173  -9.876  0.00  0.00
ATOM  44 HG22 ILE  3    12.226  12.530 -10.104  0.00  0.00
ATOM  45 HG23 ILE  3    11.043  13.889 -10.091  0.00  0.00
ATOM  46 CG1 ILE   3    12.413  13.642 -12.468  0.00  0.00
ATOM  47 HG12 ILE  3    13.249  13.172 -11.951  0.00  0.00
ATOM  48 HG13 ILE  3    12.529  13.316 -13.502  0.00  0.00
ATOM  49 CD1 ILE   3    12.319  15.094 -12.095  0.00  0.00
ATOM  50 HD11 ILE  3    12.422  15.310 -11.032  0.00  0.00
ATOM  51 HD12 ILE  3    13.074  15.661 -12.640  0.00  0.00
ATOM  52 HD13 ILE  3    11.302  15.392 -12.347  0.00  0.00
ATOM  53 C   ILE   3    12.142  10.396 -12.117  0.00  0.00
ATOM  54 O   ILE   3    13.314  10.781 -12.013  0.00  0.00
ATOM  55 N   GLN   4    11.765   9.180 -11.661  0.00  0.00
ATOM  56 H   GLN   4    10.766   9.043 -11.591  0.00  0.00
ATOM  57 CA  GLN   4    12.572   8.214 -10.865  0.00  0.00
ATOM  58 HA  GLN   4    13.592   8.591 -10.936  0.00  0.00
ATOM  59 CB  GLN   4    12.243   8.242  -9.379  0.00  0.00
ATOM  60 HB2 GLN   4    12.985   7.644  -8.850  0.00  0.00
ATOM  61 HB3 GLN   4    12.247   9.254  -8.974  0.00  0.00
ATOM  62 CG  GLN   4    10.928   7.693  -8.964  0.00  0.00
ATOM  63 HG2 GLN   4    10.717   6.862  -9.638  0.00  0.00
ATOM  64 HG3 GLN   4    10.951   7.247  -7.970  0.00  0.00
ATOM  65 CD  GLN   4     9.747   8.673  -9.055  0.00  0.00
ATOM  66 OE1 GLN   4     9.275   9.090 -10.056  0.00  0.00
ATOM  67 NE2 GLN   4     9.285   9.354  -8.020  0.00  0.00
ATOM  68 HE21 GLN  4     8.633  10.101  -8.212  0.00  0.00
ATOM  69 HE22 GLN  4     9.736   9.163  -7.137  0.00  0.00
ATOM  70 C   GLN   4    12.435   6.824 -11.471  0.00  0.00
ATOM  71 O   GLN   4    11.567   6.530 -12.313  0.00  0.00
ATOM  72 N   ASN   5    13.380   5.994 -11.111  0.00  0.00
ATOM  73 H   ASN   5    14.115   6.405 -10.552  0.00  0.00
ATOM  74 CA  ASN   5    13.501   4.653 -11.711  0.00  0.00
ATOM  75 HA  ASN   5    13.244   4.850 -12.752  0.00  0.00
ATOM  76 CB  ASN   5    14.939   4.132 -11.780  0.00  0.00
ATOM  77 HB2 ASN   5    15.590   4.919 -12.160  0.00  0.00
ATOM  78 HB3 ASN   5    15.256   3.953 -10.753  0.00  0.00
ATOM  79 CG  ASN   5    14.989   2.826 -12.542  0.00  0.00
ATOM  80 OD1 ASN   5    15.275   1.773 -11.962  0.00  0.00
ATOM  81 ND2 ASN   5    14.674   2.748 -13.825  0.00  0.00
ATOM  82 HD21 ASN  5    14.940   1.876 -14.259  0.00  0.00
ATOM  83 HD22 ASN  5    14.580   3.559 -14.420  0.00  0.00
ATOM  84 C   ASN   5    12.582   3.635 -11.089  0.00  0.00
ATOM  85 O   ASN   5    12.845   3.290  -9.951  0.00  0.00
ATOM  86 N   CYX   6    11.628   3.096 -11.846  0.00  0.00
ATOM  87 H   CYX   6    11.481   3.454 -12.779  0.00  0.00
ATOM  88 CA  CYX   6    10.632   2.121 -11.348  0.00  0.00
ATOM  89 HA  CYX   6    11.080   1.661 -10.468  0.00  0.00
ATOM  90 CB  CYX   6     9.371   2.889 -10.972  0.00  0.00
ATOM  91 HB2 CYX   6     8.921   3.186 -11.919  0.00  0.00
ATOM  92 HB3 CYX   6     8.776   2.146 -10.440  0.00  0.00
ATOM  93 SG  CYX   6     9.781   4.279  -9.913  0.00  0.00
ATOM  94 C   CYX   6    10.347   1.031 -12.490  0.00  0.00
ATOM  95 O   CYX   6    10.557   1.353 -13.677  0.00  0.00
ATOM  96 N   PRO   7     9.900  -0.201 -12.148  0.00  0.00
ATOM  97 CD  PRO   7     9.456  -0.582 -10.796  0.00  0.00
ATOM  98 HD2 PRO   7     8.728   0.086 -10.335  0.00  0.00
ATOM  99 HD3 PRO   7    10.400  -0.446 -10.267  0.00  0.00
ATOM 100 CG  PRO   7     9.038  -2.085 -10.758  0.00  0.00
ATOM 101 HG2 PRO   7     8.041  -2.359 -10.414  0.00  0.00
ATOM 102 HG3 PRO   7     9.697  -2.662 -10.110  0.00  0.00
ATOM 103 CB  PRO   7     9.352  -2.523 -12.207  0.00  0.00
ATOM 104 HB2 PRO   7     8.710  -3.338 -12.542  0.00  0.00
ATOM 105 HB3 PRO   7    10.372  -2.902 -12.270  0.00  0.00
ATOM 106 CA  PRO   7     9.510  -1.264 -13.062  0.00  0.00
ATOM 107 HA  PRO   7    10.317  -1.471 -13.764  0.00  0.00
ATOM 108 C   PRO   7     8.303  -0.934 -13.880  0.00  0.00
ATOM 109 O   PRO   7     7.561  -0.026 -13.540  0.00  0.00
ATOM 110 N   LEU   8     8.085  -1.644 -14.980  0.00  0.00
ATOM 111 H   LEU   8     8.668  -2.451 -15.149  0.00  0.00
ATOM 112 CA  LEU   8     6.904  -1.474 -15.865  0.00  0.00
ATOM 113 HA  LEU   8     6.380  -0.566 -15.567  0.00  0.00
ATOM 114 CB  LEU   8     7.318  -1.203 -17.285  0.00  0.00
ATOM 115 HB2 LEU   8     7.882  -2.079 -17.604  0.00  0.00
ATOM 116 HB3 LEU   8     6.464  -1.025 -17.938  0.00  0.00
ATOM 117 CG  LEU   8     8.274   0.041 -17.459  0.00  0.00
ATOM 118 HG  LEU   8     9.030   0.033 -16.674  0.00  0.00
ATOM 119 CD1 LEU   8     8.854  -0.151 -18.830  0.00  0.00
ATOM 120 HD11 LEU  8     8.071  -0.318 -19.570  0.00  0.00
ATOM 121 HD12 LEU  8     9.344   0.750 -19.197  0.00  0.00
ATOM 122 HD13 LEU  8     9.518  -1.015 -18.787  0.00  0.00
ATOM 123 CD2 LEU   8     7.500   1.350 -17.416  0.00  0.00
ATOM 124 HD21 LEU  8     8.118   2.247 -17.457  0.00  0.00
ATOM 125 HD22 LEU  8     6.879   1.427 -18.309  0.00  0.00
ATOM 126 HD23 LEU  8     6.810   1.397 -16.574  0.00  0.00
ATOM 127 C   LEU   8     5.955  -2.627 -15.716  0.00  0.00
ATOM 128 O   LEU   8     6.303  -3.833 -15.593  0.00  0.00
ATOM 129 N   GLY   9     4.668  -2.263 -15.715  0.00  0.00
ATOM 130 H   GLY   9     4.427  -1.288 -15.610  0.00  0.00
ATOM 131 CA  GLY   9     3.558  -3.222 -15.574  0.00  0.00
ATOM 132 HA2 GLY   9     3.831  -3.926 -14.788  0.00  0.00
ATOM 133 HA3 GLY   9     2.632  -2.828 -15.154  0.00  0.00
ATOM 134 C   GLY   9     3.292  -4.045 -16.824  0.00  0.00
ATOM 135 O   GLY   9     3.721  -3.702 -17.935  0.00  0.00
ATOM 136 N   NHE  10     2.475  -5.093 -16.676  0.00  0.00
ATOM 137 HN1 NHE  10     2.033  -5.320 -15.796  0.00  0.00
ATOM 138 HN2 NHE  10     2.206  -5.615 -17.498  0.00  0.00
TER  139      NHE  10
END
```

CT, intermediate saddle (OPEN/FOLDED)*

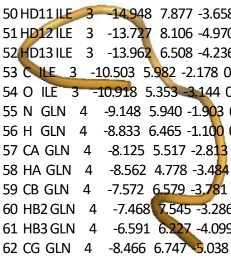
CT_MD-I_5us_T16_11

```
ATOM   1 HA2 MET   1   -16.462   2.328 -18.391  0.00  0.00
ATOM   2 CA  MET   1   -16.389   2.788 -17.406  0.00  0.00
ATOM   3 HA1 MET   1   -15.521   3.446 -17.458  0.00  0.00
ATOM   4 CB  MET   1   -17.592   3.673 -17.230  0.00  0.00
ATOM   5 HB2 MET   1   -17.589   4.480 -17.962  0.00  0.00
ATOM   6 HB3 MET   1   -18.536   3.172 -17.444  0.00  0.00
ATOM   7 CG  MET   1   -17.671   4.219 -15.875  0.00  0.00
ATOM   8 HG2 MET   1   -18.601   4.771 -15.743  0.00  0.00
ATOM   9 HG3 MET   1   -17.648   3.342 -15.228  0.00  0.00
ATOM  10 C   MET   1   -16.179   1.651 -16.440  0.00  0.00
ATOM  11 O   MET   1   -16.983   0.737 -16.355  0.00  0.00
ATOM  12 N   TYR   2   -15.059   1.719 -15.707  0.00  0.00
ATOM  13 H   TYR   2   -14.454   2.525 -15.780  0.00  0.00
ATOM  14 CA  TYR   2   -14.684   0.701 -14.749  0.00  0.00
ATOM  15 HA  TYR   2   -14.858  -0.237 -15.276  0.00  0.00
ATOM  16 CB  TYR   2   -13.183   0.892 -14.414  0.00  0.00
ATOM  17 HB2 TYR   2   -12.589   0.819 -15.325  0.00  0.00
ATOM  18 HB3 TYR   2   -13.065   1.893 -13.998  0.00  0.00
ATOM  19 CG  TYR   2   -12.633  -0.187 -13.506  0.00  0.00
ATOM  20 CD1 TYR   2   -12.695  -1.558 -13.790  0.00  0.00
ATOM  21 HD1 TYR   2   -13.049  -1.885 -14.756  0.00  0.00
ATOM  22 CE1 TYR   2   -12.168  -2.500 -12.906  0.00  0.00
ATOM  23 HE1 TYR   2   -12.195  -3.548 -13.166  0.00  0.00
ATOM  24 CZ  TYR   2   -11.550  -2.107 -11.671  0.00  0.00
ATOM  25 OS  TYR   2   -11.010  -3.063 -10.753  0.00  0.00
ATOM  26 CH  TYR   2   -11.363  -4.423 -10.850  0.00  0.00
ATOM  48 HG13 ILE  3   -14.694  -2.400 -10.479  0.00  0.00
ATOM  49 CD1 ILE   3   -15.779  -4.019  -9.837  0.00  0.00
ATOM  50 HD11 ILE  3   -14.845  -4.540  -9.628  0.00  0.00
ATOM  51 HD12 ILE  3   -16.054  -3.563  -8.886  0.00  0.00
ATOM  52 HD13 ILE  3   -16.562  -4.692 -10.189  0.00  0.00
ATOM  53 C   ILE   3   -15.757  -0.036 -10.228  0.00  0.00
ATOM  54 O   ILE   3   -16.286  -0.271  -9.116  0.00  0.00
ATOM  55 N   GLN   4   -14.736   0.796 -10.308  0.00  0.00
ATOM  56 H   GLN   4   -14.183   0.735 -11.151  0.00  0.00
ATOM  57 CA  GLN   4   -14.366   1.739  -9.294  0.00  0.00
ATOM  58 HA  GLN   4   -15.148   1.827  -8.540  0.00  0.00
ATOM  59 CB  GLN   4   -13.150   1.095  -8.560  0.00  0.00
ATOM  60 HB2 GLN   4   -13.344   0.024  -8.615  0.00  0.00
ATOM  61 HB3 GLN   4   -12.214   1.321  -9.072  0.00  0.00
ATOM  62 CG  GLN   4   -13.222   1.646  -7.130  0.00  0.00
ATOM  63 HG2 GLN   4   -13.189   2.733  -7.197  0.00  0.00
ATOM  64 HG3 GLN   4   -14.106   1.273  -6.613  0.00  0.00
ATOM  65 CD  GLN   4   -12.023   1.168  -6.344  0.00  0.00
ATOM  66 OE1 GLN   4   -11.066   0.791  -6.916  0.00  0.00
ATOM  67 NE2 GLN   4   -12.056   1.203  -5.027  0.00  0.00
ATOM  68 H E21 GLN 4   -11.224   1.019  -4.486  0.00  0.00
ATOM  69 HE22 GLN  4   -12.853   1.659  -4.606  0.00  0.00
ATOM  70 C   GLN   4   -14.026   3.097  -9.851  0.00  0.00
ATOM  71 O   GLN   4   -13.334   3.311 -10.864  0.00  0.00
ATOM  72 N   ASN   5   -14.512   4.169  -9.165  0.00  0.00
ATOM  73 H   ASN   5   -15.031   3.986  -8.318  0.00  0.00
ATOM  95 O   CYX   6   -16.632   7.799 -12.220  0.00  0.00
ATOM  96 N   PRO   7   -15.174   8.409 -13.837  0.00  0.00
ATOM  97 CD  PRO   7   -14.108   8.152 -14.806  0.00  0.00
ATOM  98 HD2 PRO   7   -14.642   7.656 -15.617  0.00  0.00
ATOM  99 HD3 PRO   7   -13.280   7.524 -14.479  0.00  0.00
ATOM 100 CG  PRO   7   -13.587   9.535 -15.333  0.00  0.00
ATOM 101 HG2 PRO   7   -13.209   9.489 -16.354  0.00  0.00
ATOM 102 HG3 PRO   7   -12.821  10.017 -14.724  0.00  0.00
ATOM 103 CB  PRO   7   -14.851  10.398 -15.244  0.00  0.00
ATOM 104 HB2 PRO   7   -15.306  10.400 -16.234  0.00  0.00
ATOM 105 HB3 PRO   7   -14.584  11.440 -15.067  0.00  0.00
ATOM 106 CA  PRO   7   -15.728   9.748 -14.050  0.00  0.00
ATOM 107 HA  PRO   7   -15.723  10.358 -13.146  0.00  0.00
ATOM 108 C   PRO   7   -17.201   9.739 -14.442  0.00  0.00
ATOM 109 O   PRO   7   -17.649   8.875 -15.247  0.00  0.00
ATOM 110 N   LEU   8   -17.998  10.644 -13.838  0.00  0.00
ATOM 111 H   LEU   8   -17.519  11.430 -13.423  0.00  0.00
ATOM 112 CA  LEU   8   -19.395  10.765 -14.203  0.00  0.00
ATOM 113 HA  LEU   8   -19.755   9.902 -14.764  0.00  0.00
ATOM 114 CB  LEU   8   -20.144  10.735 -12.886  0.00  0.00
ATOM 115 HB2 LEU   8   -19.872   9.806 -12.385  0.00  0.00
ATOM 116 HB3 LEU   8   -19.852  11.575 -12.255  0.00  0.00
ATOM 117 CG  LEU   8   -21.713  10.752 -12.884  0.00  0.00
ATOM 118 HG  LEU   8   -21.979  11.794 -13.060  0.00  0.00
ATOM 119 CD1 LEU   8   -22.307   9.807 -14.009  0.00  0.00
ATOM 120 HD11 LEU  8   -22.335  10.417 -14.913  0.00  0.00
```

```
ATOM 27 CE2 TYR 2 -11.471 -0.715 -11.438 0.00 0.00
ATOM 28 HE2 TYR 2 -11.071 -0.401 -10.485 0.00 0.00
ATOM 29 CD2 TYR 2 -11.996 0.268 -12.288 0.00 0.00
ATOM 30 HD2 TYR 2 -11.841 1.320 -12.102 0.00 0.00
ATOM 31 C TYR 2 -15.511 0.672 -13.453 0.00 0.00
ATOM 32 O TYR 2 -16.157 1.671 -13.139 0.00 0.00
ATOM 33 HH1 TYR 2 -10.868 -4.727 -11.772 0.00 0.00
ATOM 34 HH2 TYR 2 -12.439 -4.599 -10.853 0.00 0.00
ATOM 35 HH3 TYR 2 -10.927 -5.057 -10.078 0.00 0.00
ATOM 36 N ILE 3 -15.529 -0.463 -12.700 0.00 0.00
ATOM 37 H ILE 3 -14.874 -1.193 -12.941 0.00 0.00
ATOM 38 CA ILE 3 -16.374 -0.691 -11.513 0.00 0.00
ATOM 39 HA ILE 3 -17.313 -0.169 -11.697 0.00 0.00
ATOM 40 CB ILE 3 -16.748 -2.119 -11.274 0.00 0.00
ATOM 41 HB ILE 3 -17.470 -2.311 -10.480 0.00 0.00
ATOM 42 CG2 ILE 3 -17.524 -2.718 -12.453 0.00 0.00
ATOM 43 HG21 ILE 3 -18.066 -3.566 -12.036 0.00 0.00
ATOM 44 HG22 ILE 3 -18.276 -2.030 -12.839 0.00 0.00
ATOM 45 HG23 ILE 3 -16.814 -3.057 -13.207 0.00 0.00
ATOM 46 CG1 ILE 3 -15.511 -2.974 -10.917 0.00 0.00
ATOM 47 HG12 ILE 3 -15.154 -3.542 -11.776 0.00 0.00
ATOM 74 CA ASN 5 -14.067 5.586 -9.334 0.00 0.00
ATOM 75 HA ASN 5 -14.767 6.162 -8.729 0.00 0.00
ATOM 76 CB ASN 5 -12.645 5.736 -8.680 0.00 0.00
ATOM 77 HB2 ASN 5 -12.707 5.256 -7.704 0.00 0.00
ATOM 78 HB3 ASN 5 -11.859 5.193 -9.205 0.00 0.00
ATOM 79 CG ASN 5 -12.110 7.082 -8.546 0.00 0.00
ATOM 80 OD1 ASN 5 -11.204 7.517 -9.229 0.00 0.00
ATOM 81 ND2 ASN 5 -12.670 7.901 -7.711 0.00 0.00
ATOM 82 HD21 ASN 5 -12.537 8.898 -7.799 0.00 0.00
ATOM 83 HD22 ASN 5 -13.220 7.459 -6.988 0.00 0.00
ATOM 84 C ASN 5 -14.075 6.159 -10.781 0.00 0.00
ATOM 85 O ASN 5 -13.349 7.057 -11.186 0.00 0.00
ATOM 86 N CYX 6 -15.002 5.603 -11.581 0.00 0.00
ATOM 87 H CYX 6 -15.432 4.728 -11.319 0.00 0.00
ATOM 88 CA CYX 6 -15.180 6.047 -12.963 0.00 0.00
ATOM 89 HA CYX 6 -14.216 6.053 -13.473 0.00 0.00
ATOM 90 CB CYX 6 -16.166 5.084 -13.686 0.00 0.00
ATOM 91 HB2 CYX 6 -15.648 4.132 -13.566 0.00 0.00
ATOM 92 HB3 CYX 6 -17.097 5.114 -13.121 0.00 0.00
ATOM 93 SG CYX 6 -16.396 5.345 -15.449 0.00 0.00
ATOM 94 C CYX 6 -15.677 7.486 -12.984 0.00 0.00
ATOM 121 HD12 LEU 8 -21.615 8.998 -14.243 0.00 0.00
ATOM 122 HD13 LEU 8 -23.323 9.454 -13.836 0.00 0.00
ATOM 123 CD2 LEU 8 -22.229 10.316 -11.530 0.00 0.00
ATOM 124 HD21 LEU 8 -21.858 9.356 -11.172 0.00 0.00
ATOM 125 HD22 LEU 8 -21.908 11.008 -10.753 0.00 0.00
ATOM 126 HD23 LEU 8 -23.319 10.348 -11.528 0.00 0.00
ATOM 127 C LEU 8 -19.574 12.113 -14.977 0.00 0.00
ATOM 128 O LEU 8 -19.216 13.094 -14.391 0.00 0.00
ATOM 129 N GLY 9 -20.135 12.144 -16.246 0.00 0.00
ATOM 130 H GLY 9 -20.469 11.239 -16.545 0.00 0.00
ATOM 131 CA GLY 9 -20.316 13.291 -17.106 0.00 0.00
ATOM 132 HA2 GLY 9 -19.357 13.717 -17.401 0.00 0.00
ATOM 133 HA3 GLY 9 -20.750 12.925 -18.036 0.00 0.00
ATOM 134 C GLY 9 -21.277 14.354 -16.541 0.00 0.00
ATOM 135 O GLY 9 -22.101 14.069 -15.711 0.00 0.00
ATOM 136 N NHE 10 -21.325 15.546 -17.131 0.00 0.00
ATOM 137 HN1 NHE 10 -20.626 15.760 -17.828 0.00 0.00
ATOM 138 HN2 NHE 10 -22.098 16.100 -16.790 0.00 0.00
TER 139 NHE 10
END
```

CT, intermediate saddlevar (OPEN/FOLDED)*

CT_MD-II_5us_T16_6



```
ATOM 1 HA2 MET 1 -5.626 11.280 0.223 0.00 0.00
ATOM 2 CA MET 1 -6.342 11.037 -0.562 0.00 0.00
ATOM 3 HA1 MET 1 -6.578 11.973 -1.066 0.00 0.00
ATOM 4 CB MET 1 -5.840 10.032 -1.552 0.00 0.00
ATOM 5 HB2 MET 1 -6.732 9.636 -2.037 0.00 0.00
ATOM 6 HB3 MET 1 -5.210 10.439 -2.343 0.00 0.00
ATOM 7 CG MET 1 -5.047 8.941 -0.913 0.00 0.00
ATOM 8 HG2 MET 1 -4.044 9.247 -0.617 0.00 0.00
ATOM 9 HG3 MET 1 -5.590 8.508 -0.072 0.00 0.00
ATOM 10 C MET 1 -7.540 10.452 0.155 0.00 0.00
ATOM 11 O MET 1 -7.410 9.904 1.288 0.00 0.00
ATOM 12 N TYR 2 -8.659 10.568 -0.480 0.00 0.00
ATOM 13 H TYR 2 -8.581 11.034 -1.373 0.00 0.00
ATOM 14 CA TYR 2 -9.971 9.995 -0.101 0.00 0.00
ATOM 15 HA TYR 2 -9.833 9.576 0.895 0.00 0.00
ATOM 16 CB TYR 2 -10.942 11.187 -0.071 0.00 0.00
ATOM 17 HB2 TYR 2 -10.547 12.065 0.439 0.00 0.00
ATOM 18 HB3 TYR 2 -11.008 11.532 -1.103 0.00 0.00
ATOM 19 CG TYR 2 -12.387 10.940 0.373 0.00 0.00
ATOM 20 CD1 TYR 2 -13.437 10.811 -0.527 0.00 0.00
ATOM 21 HD1 TYR 2 -13.216 10.900 -1.581 0.00 0.00
ATOM 22 CE1 TYR 2 -14.714 10.491 -0.024 0.00 0.00
ATOM 23 HE1 TYR 2 -15.506 10.383 -0.750 0.00 0.00
ATOM 24 CZ TYR 2 -14.894 10.174 1.310 0.00 0.00
ATOM 25 OS TYR 2 -15.988 9.545 1.834 0.00 0.00
ATOM 26 CH TYR 2 -16.865 8.917 0.864 0.00 0.00
ATOM 27 CE2 TYR 2 -13.773 10.328 2.161 0.00 0.00
ATOM 28 HE2 TYR 2 -13.908 10.055 3.198 0.00 0.00
ATOM 29 CD2 TYR 2 -12.520 10.777 1.752 0.00 0.00
ATOM 30 HD2 TYR 2 -11.656 10.843 2.397 0.00 0.00
ATOM 31 C TYR 2 -10.320 8.851 -1.068 0.00 0.00
ATOM 32 O TYR 2 -10.009 8.981 -2.261 0.00 0.00
ATOM 33 HH1 TYR 2 -16.350 8.151 0.284 0.00 0.00
ATOM 34 HH2 TYR 2 -17.308 9.748 0.315 0.00 0.00
ATOM 35 HH3 TYR 2 -17.691 8.467 1.415 0.00 0.00
ATOM 36 N ILE 3 -11.029 7.805 -0.566 0.00 0.00
ATOM 37 H ILE 3 -11.290 7.833 0.409 0.00 0.00
ATOM 38 CA ILE 3 -11.498 6.643 -1.212 0.00 0.00
ATOM 39 HA ILE 3 -11.579 5.899 -0.418 0.00 0.00
ATOM 40 CB ILE 3 -12.890 6.810 -1.870 0.00 0.00
ATOM 41 HB ILE 3 -13.107 5.843 -2.324 0.00 0.00
ATOM 42 CG2 ILE 3 -13.929 7.111 -0.770 0.00 0.00
ATOM 43 HG21 ILE 3 -13.646 7.983 -0.179 0.00 0.00
ATOM 44 HG22 ILE 3 -14.920 7.312 -1.177 0.00 0.00
ATOM 45 HG23 ILE 3 -14.146 6.287 -0.090 0.00 0.00
ATOM 46 CG1 ILE 3 -12.928 7.932 -2.936 0.00 0.00
ATOM 47 HG12 ILE 3 -13.148 8.869 -2.426 0.00 0.00
ATOM 48 HG13 ILE 3 -11.959 7.956 -3.436 0.00 0.00
ATOM 49 CD1 ILE 3 -13.974 7.582 -4.047 0.00 0.00
ATOM 50 HD11 ILE 3 -14.948 7.877 -3.658 0.00 0.00
ATOM 51 HD12 ILE 3 -13.727 8.106 -4.970 0.00 0.00
ATOM 52 HD13 ILE 3 -13.962 6.508 -4.236 0.00 0.00
ATOM 53 C ILE 3 -10.503 5.982 -2.178 0.00 0.00
ATOM 54 O ILE 3 -10.918 5.353 -3.144 0.00 0.00
ATOM 55 N GLN 4 -9.148 5.940 -1.903 0.00 0.00
ATOM 56 H GLN 4 -8.833 6.465 -1.100 0.00 0.00
ATOM 57 CA GLN 4 -8.125 5.517 -2.813 0.00 0.00
ATOM 58 HA GLN 4 -8.562 4.778 -3.484 0.00 0.00
ATOM 59 CB GLN 4 -7.572 6.579 -3.781 0.00 0.00
ATOM 60 HB2 GLN 4 -7.468 7.545 -3.286 0.00 0.00
ATOM 61 HB3 GLN 4 -6.591 6.227 -4.099 0.00 0.00
ATOM 62 CG GLN 4 -8.466 6.747 -5.038 0.00 0.00
ATOM 63 HG2 GLN 4 -8.710 5.766 -5.445 0.00 0.00
ATOM 64 HG3 GLN 4 -9.436 7.183 -4.800 0.00 0.00
ATOM 65 CD GLN 4 -7.803 7.643 -6.131 0.00 0.00
ATOM 66 OE1 GLN 4 -7.029 8.577 -5.858 0.00 0.00
ATOM 67 NE2 GLN 4 -8.024 7.456 -7.396 0.00 0.00
ATOM 68 HE21 GLN 4 -7.609 8.079 -8.073 0.00 0.00
ATOM 69 HE22 GLN 4 -8.722 6.773 -7.654 0.00 0.00
ATOM 70 C GLN 4 -7.069 4.595 -2.243 0.00 0.00
ATOM 71 O GLN 4 -6.064 4.419 -2.898 0.00 0.00
ATOM 72 N ASN 5 -7.321 4.032 -1.055 0.00 0.00
ATOM 73 H ASN 5 -8.201 4.304 -0.639 0.00 0.00
ATOM 74 CA ASN 5 -6.476 3.133 -0.335 0.00 0.00
ATOM 75 HA ASN 5 -6.866 2.984 0.672 0.00 0.00
ATOM 76 CB ASN 5 -6.539 1.800 -1.033 0.00 0.00
ATOM 77 HB2 ASN 5 -7.586 1.557 -1.216 0.00 0.00
ATOM 78 HB3 ASN 5 -6.106 1.892 -2.028 0.00 0.00
ATOM 79 CG ASN 5 -5.771 0.688 -0.266 0.00 0.00
ATOM 80 OD1 ASN 5 -5.745 0.601 0.970 0.00 0.00
ATOM 81 ND2 ASN 5 -5.251 -0.213 -1.001 0.00 0.00
ATOM 82 HD21 ASN 5 -4.472 -0.739 -0.630 0.00 0.00
ATOM 83 HD22 ASN 5 -5.516 -0.331 -1.969 0.00 0.00
ATOM 84 C ASN 5 -5.134 3.765 -0.153 0.00 0.00
ATOM 85 O ASN 5 -4.103 3.137 -0.266 0.00 0.00
ATOM 86 N CYX 6 -5.011 5.057 0.039 0.00 0.00
ATOM 87 H CYX 6 -5.946 5.438 0.029 0.00 0.00
ATOM 88 CA CYX 6 -3.777 5.888 0.223 0.00 0.00
ATOM 89 HA CYX 6 -2.991 5.159 0.422 0.00 0.00
ATOM 90 CB CYX 6 -3.455 6.620 -1.123 0.00 0.00
ATOM 91 HB2 CYX 6 -2.681 7.372 -0.973 0.00 0.00
ATOM 92 HB3 CYX 6 -3.088 5.860 -1.813 0.00 0.00
ATOM 93 SG CYX 6 -4.894 7.502 -1.917 0.00 0.00
ATOM 94 C CYX 6 -3.898 6.838 1.499 0.00 0.00
ATOM 95 O CYX 6 -4.980 7.370 1.844 0.00 0.00
ATOM 96 N PRO 7 -2.745 7.152 2.107 0.00 0.00
ATOM 97 CD PRO 7 -1.421 6.788 1.647 0.00 0.00
ATOM 98 HD2 PRO 7 -1.376 6.802 0.558 0.00 0.00
ATOM 99 HD3 PRO 7 -1.182 5.789 2.012 0.00 0.00
ATOM 100 CG PRO 7 -0.457 7.771 2.348 0.00 0.00
ATOM 101 HG2 PRO 7 -0.176 8.623 1.728 0.00 0.00
ATOM 102 HG3 PRO 7 0.409 7.156 2.592 0.00 0.00
ATOM 103 CB PRO 7 -1.130 8.240 3.626 0.00 0.00
ATOM 104 HB2 PRO 7 -0.803 9.240 3.909 0.00 0.00
ATOM 105 HB3 PRO 7 -0.923 7.429 4.324 0.00 0.00
ATOM 106 CA PRO 7 -2.633 8.189 3.165 0.00 0.00
ATOM 107 HA PRO 7 -3.189 7.858 4.041 0.00 0.00
ATOM 108 C PRO 7 -3.205 9.559 2.695 0.00 0.00
ATOM 109 O PRO 7 -2.946 9.954 1.547 0.00 0.00
ATOM 110 N LEU 8 -3.986 10.249 3.523 0.00 0.00
ATOM 111 H LEU 8 -4.224 9.837 4.414 0.00 0.00
ATOM 112 CA LEU 8 -4.372 11.584 3.390 0.00 0.00
ATOM 113 HA LEU 8 -4.469 11.902 2.352 0.00 0.00
ATOM 114 CB LEU 8 -5.717 11.719 4.139 0.00 0.00
ATOM 115 HB2 LEU 8 -6.393 10.887 3.940 0.00 0.00
ATOM 116 HB3 LEU 8 -5.566 11.587 5.210 0.00 0.00
ATOM 117 CG LEU 8 -6.362 13.120 3.888 0.00 0.00
ATOM 118 HG LEU 8 -5.619 13.896 4.071 0.00 0.00
ATOM 119 CD1 LEU 8 -6.848 13.246 2.396 0.00 0.00
ATOM 120 HD11 LEU 8 -6.002 13.112 1.722 0.00 0.00
ATOM 121 HD12 LEU 8 -7.600 12.466 2.272 0.00 0.00
ATOM 122 HD13 LEU 8 -7.232 14.234 2.143 0.00 0.00
ATOM 123 CD2 LEU 8 -7.624 13.205 4.854 0.00 0.00
ATOM 124 HD21 LEU 8 -7.534 12.943 5.908 0.00 0.00
ATOM 125 HD22 LEU 8 -8.109 14.173 4.727 0.00 0.00
ATOM 126 HD23 LEU 8 -8.192 12.372 4.440 0.00 0.00
ATOM 127 C LEU 8 -3.327 12.478 4.061 0.00 0.00
ATOM 128 O LEU 8 -3.100 12.314 5.293 0.00 0.00
ATOM 129 N GLY 9 -2.954 13.589 3.348 0.00 0.00
ATOM 130 H GLY 9 -3.112 13.636 2.351 0.00 0.00
ATOM 131 CA GLY 9 -2.137 14.665 3.913 0.00 0.00
ATOM 132 HA2 GLY 9 -2.172 15.569 3.305 0.00 0.00
ATOM 133 HA3 GLY 9 -2.523 14.854 4.915 0.00 0.00
ATOM 134 C GLY 9 -0.682 14.303 4.079 0.00 0.00
ATOM 135 O GLY 9 -0.129 13.348 3.585 0.00 0.00
ATOM 136 N NHE 10 -0.010 15.066 4.938 0.00 0.00
ATOM 137 HN1 NHE 10 -0.393 15.954 5.229 0.00 0.00
ATOM 138 HN2 NHE 10 0.832 14.635 5.291 0.00 0.00
TER 139 NHE 10
END
```

CT, saddlevar (= folded-II) (FOLDED)*

CT_MD-II_5us_T16_5

```
ATOM 1 HA2 MET 1 11.568 -16.276 -9.059 0.00 0.00
ATOM 2 CA MET 1 11.548 -15.191 -9.163 0.00 0.00
ATOM 3 HA1 MET 1 11.655 -14.825 -8.142 0.00 0.00
ATOM 4 CB MET 1 12.766 -14.735 -10.011 0.00 0.00
ATOM 5 HB2 MET 1 13.663 -14.970 -9.437 0.00 0.00
ATOM 6 HB3 MET 1 12.850 -15.296 -10.942 0.00 0.00
ATOM 7 CG MET 1 12.721 -13.253 -10.389 0.00 0.00
ATOM 8 HG2 MET 1 13.437 -12.989 -11.167 0.00 0.00
ATOM 9 HG3 MET 1 11.774 -13.033 -10.882 0.00 0.00
ATOM 10 C MET 1 10.266 -14.804 -9.788 0.00 0.00
ATOM 11 O MET 1 9.517 -14.088 -9.151 0.00 0.00
ATOM 12 N TYR 2 9.885 -15.580 -10.829 0.00 0.00
ATOM 13 H TYR 2 10.453 -16.260 -11.314 0.00 0.00
ATOM 14 CA TYR 2 8.610 -15.446 -11.474 0.00 0.00
ATOM 15 HA TYR 2 8.512 -14.400 -11.765 0.00 0.00
ATOM 16 CB TYR 2 8.713 -16.226 -12.843 0.00 0.00
ATOM 17 HB2 TYR 2 9.600 -15.861 -13.361 0.00 0.00
ATOM 18 HB3 TYR 2 8.860 -17.298 -12.709 0.00 0.00
ATOM 48 HG13 ILE 3 4.257 -16.917 -9.589 0.00 0.00
ATOM 49 CD1 ILE 3 3.179 -17.689 -11.275 0.00 0.00
ATOM 50 HD11 ILE 3 3.297 -18.646 -10.766 0.00 0.00
ATOM 51 HD12 ILE 3 2.211 -17.327 -10.927 0.00 0.00
ATOM 52 HD13 ILE 3 3.214 -17.859 -12.351 0.00 0.00
ATOM 53 C ILE 3 5.145 -14.664 -8.448 0.00 0.00
ATOM 54 O ILE 3 4.143 -14.799 -7.823 0.00 0.00
ATOM 55 N GLN 4 6.323 -14.481 -7.911 0.00 0.00
ATOM 56 H GLN 4 7.099 -14.682 -8.526 0.00 0.00
ATOM 57 CA GLN 4 6.572 -14.370 -6.464 0.00 0.00
ATOM 58 HA GLN 4 5.618 -14.296 -5.942 0.00 0.00
ATOM 59 CB GLN 4 7.293 -15.626 -5.929 0.00 0.00
ATOM 60 HB2 GLN 4 8.347 -15.540 -6.193 0.00 0.00
ATOM 61 HB3 GLN 4 7.195 -15.591 -4.844 0.00 0.00
ATOM 62 CG GLN 4 6.698 -16.965 -6.436 0.00 0.00
ATOM 63 HG2 GLN 4 5.626 -17.000 -6.243 0.00 0.00
ATOM 64 HG3 GLN 4 6.891 -17.126 -7.497 0.00 0.00
ATOM 65 CD GLN 4 7.200 -18.185 -5.665 0.00 0.00
ATOM 95 O CYX 6 9.808 -11.516 -11.449 0.00 0.00
ATOM 96 N PRO 7 10.478 -9.414 -11.918 0.00 0.00
ATOM 97 CD PRO 7 10.712 -7.994 -11.642 0.00 0.00
ATOM 98 HD2 PRO 7 11.790 -7.840 -11.604 0.00 0.00
ATOM 99 HD3 PRO 7 10.289 -7.659 -10.695 0.00 0.00
ATOM 100 CG PRO 7 10.104 -7.219 -12.739 0.00 0.00
ATOM 101 HG2 PRO 7 10.605 -6.297 -13.033 0.00 0.00
ATOM 102 HG3 PRO 7 9.186 -6.790 -12.336 0.00 0.00
ATOM 103 CB PRO 7 9.890 -8.139 -13.888 0.00 0.00
ATOM 104 HB2 PRO 7 10.405 -7.849 -14.804 0.00 0.00
ATOM 105 HB3 PRO 7 8.827 -8.100 -14.130 0.00 0.00
ATOM 106 CA PRO 7 10.250 -9.571 -13.395 0.00 0.00
ATOM 107 HA PRO 7 9.502 -10.308 -13.674 0.00 0.00
ATOM 108 C PRO 7 11.562 -10.062 -14.084 0.00 0.00
ATOM 109 O PRO 7 12.629 -9.846 -13.502 0.00 0.00
ATOM 110 N LEU 8 11.362 -10.770 -15.190 0.00 0.00
ATOM 111 H LEU 8 10.390 -10.928 -15.412 0.00 0.00
ATOM 112 CA LEU 8 12.324 -11.496 -15.986 0.00 0.00
```

```
ATOM  19 CG TYR   2     7.463 -16.175 -13.705 0.00 0.00
ATOM  20 CD1 TYR  2     6.674 -17.316 -13.915 0.00 0.00
ATOM  21 HD1 TYR  2     6.848 -18.163 -13.269 0.00 0.00
ATOM  22 CE1 TYR  2     5.443 -17.290 -14.618 0.00 0.00
ATOM  23 HE1 TYR  2     4.801 -18.159 -14.615 0.00 0.00
ATOM  24 CZ  TYR  2     5.090 -16.123 -15.324 0.00 0.00
ATOM  25 OS  TYR  2     3.819 -16.196 -15.987 0.00 0.00
ATOM  26 CH  TYR  2     3.435 -15.041 -16.663 0.00 0.00
ATOM  27 CE2 TYR  2     5.930 -14.994 -15.245 0.00 0.00
ATOM  28 HE2 TYR  2     5.593 -14.047 -15.639 0.00 0.00
ATOM  29 CD2 TYR  2     7.075 -15.024 -14.388 0.00 0.00
ATOM  30 HD2 TYR  2     7.530 -14.073 -14.156 0.00 0.00
ATOM  31 C   TYR  2     7.374 -15.653 -10.626 0.00 0.00
ATOM  32 O   TYR  2     7.170 -16.649 -9.917 0.00 0.00
ATOM  33 HH1 TYR  2     4.284 -14.736 -17.276 0.00 0.00
ATOM  34 HH2 TYR  2     3.087 -14.343 -15.901 0.00 0.00
ATOM  35 HH3 TYR  2     2.634 -15.217 -17.380 0.00 0.00
ATOM  36 N   ILE  3     6.422 -14.633 -10.750 0.00 0.00
ATOM  37 H   ILE  3     6.538 -13.786 -11.288 0.00 0.00
ATOM  38 CA  ILE  3     5.139 -14.503 -10.016 0.00 0.00
ATOM  39 HA  ILE  3     4.949 -13.438 -10.152 0.00 0.00
ATOM  40 CB  ILE  3     4.007 -15.302 -10.655 0.00 0.00
ATOM  41 HB  ILE  3     3.109 -14.989 -10.123 0.00 0.00
ATOM  42 CG2 ILE  3     3.748 -14.742 -12.036 0.00 0.00
ATOM  43 HG21 ILE 3     4.418 -15.212 -12.756 0.00 0.00
ATOM  44 HG22 ILE 3     2.703 -15.009 -12.189 0.00 0.00
ATOM  45 HG23 ILE 3     3.927 -13.685 -12.233 0.00 0.00
ATOM  46 CG1 ILE  3     4.231 -16.764 -10.668 0.00 0.00
ATOM  47 HG12 ILE 3     5.170 -16.928 -11.196 0.00 0.00
ATOM  66 OE1 GLN  4     8.389 -18.555 -5.772 0.00 0.00
ATOM  67 NE2 GLN  4     6.439 -18.898 -4.866 0.00 0.00
ATOM  68 HE21 GLN 4     6.886 -19.640 -4.346 0.00 0.00
ATOM  69 HE22 GLN 4     5.543 -18.524 -4.589 0.00 0.00
ATOM  70 C   GLN  4     7.404 -13.092 -6.046 0.00 0.00
ATOM  71 O   GLN  4     7.101 -12.555 -5.028 0.00 0.00
ATOM  72 N   ASN  5     8.383 -12.623 -6.840 0.00 0.00
ATOM  73 H   ASN  5     8.636 -13.100 -7.694 0.00 0.00
ATOM  74 CA  ASN  5     9.439 -11.694 -6.316 0.00 0.00
ATOM  75 HA  ASN  5     8.891 -11.096 -5.588 0.00 0.00
ATOM  76 CB  ASN  5     10.770 -12.425 -5.841 0.00 0.00
ATOM  77 HB2 ASN  5     10.657 -13.066 -4.967 0.00 0.00
ATOM  78 HB3 ASN  5     11.124 -12.940 -6.734 0.00 0.00
ATOM  79 CG  ASN  5     11.957 -11.495 -5.486 0.00 0.00
ATOM  80 OD1 ASN  5     11.934 -10.310 -5.438 0.00 0.00
ATOM  81 ND2 ASN  5     13.063 -12.103 -5.260 0.00 0.00
ATOM  82 HD21 ASN 5     13.877 -11.576 -4.975 0.00 0.00
ATOM  83 HD22 ASN 5     13.104 -13.112 -5.268 0.00 0.00
ATOM  84 C   ASN  5     9.799 -10.663 -7.405 0.00 0.00
ATOM  85 O   ASN  5     9.612 -9.480 -7.279 0.00 0.00
ATOM  86 N   CYX  6     10.343 -11.139 -8.509 0.00 0.00
ATOM  87 H   CYX  6     10.459 -12.141 -8.456 0.00 0.00
ATOM  88 CA  CYX  6     10.944 -10.353 -9.619 0.00 0.00
ATOM  89 HA  CYX  6     10.846 -9.291 -9.396 0.00 0.00
ATOM  90 CB  CYX  6     12.499 -10.491 -9.633 0.00 0.00
ATOM  91 HB2 CYX  6     12.893 -10.222 -10.613 0.00 0.00
ATOM  92 HB3 CYX  6     12.842 -9.756 -8.906 0.00 0.00
ATOM  93 SG  CYX  6     13.042 -12.148 -9.050 0.00 0.00
ATOM  94 C   CYX  6     10.406 -10.467 -11.052 0.00 0.00
ATOM 113 HA LEU   8     13.196 -11.724 -15.373 0.00 0.00
ATOM 114 CB LEU   8     11.674 -12.852 -16.462 0.00 0.00
ATOM 115 HB2 LEU  8     10.877 -12.497 -17.116 0.00 0.00
ATOM 116 HB3 LEU  8     12.451 -13.395 -17.000 0.00 0.00
ATOM 117 CG LEU   8     11.129 -13.781 -15.313 0.00 0.00
ATOM 118 HG LEU   8     10.455 -13.184 -14.698 0.00 0.00
ATOM 119 CD1 LEU  8     10.423 -14.921 -16.038 0.00 0.00
ATOM 120 HD11 LEU 8     9.428 -14.549 -16.285 0.00 0.00
ATOM 121 HD12 LEU 8     10.924 -15.174 -16.972 0.00 0.00
ATOM 122 HD13 LEU 8     10.474 -15.799 -15.394 0.00 0.00
ATOM 123 CD2 LEU  8     12.141 -14.301 -14.296 0.00 0.00
ATOM 124 HD21 LEU 8     12.855 -13.537 -13.989 0.00 0.00
ATOM 125 HD22 LEU 8     11.596 -14.669 -13.427 0.00 0.00
ATOM 126 HD23 LEU 8     12.672 -15.133 -14.759 0.00 0.00
ATOM 127 C   LEU  8     12.785 -10.543 -17.105 0.00 0.00
ATOM 128 O   LEU  8     12.061 -10.192 -18.064 0.00 0.00
ATOM 129 N   GLY  9     14.105 -10.238 -17.102 0.00 0.00
ATOM 130 H   GLY  9     14.760 -10.532 -16.392 0.00 0.00
ATOM 131 CA  GLY  9     14.778 -9.588 -18.268 0.00 0.00
ATOM 132 HA2 GLY  9     14.258 -8.638 -18.391 0.00 0.00
ATOM 133 HA3 GLY  9     15.831 -9.533 -17.991 0.00 0.00
ATOM 134 C   GLY  9     14.572 -10.352 -19.609 0.00 0.00
ATOM 135 O   GLY  9     14.468 -11.555 -19.696 0.00 0.00
ATOM 136 N   NHE  10    14.424 -9.680 -20.756 0.00 0.00
ATOM 137 HN1 NHE  10    14.424 -8.674 -20.671 0.00 0.00
ATOM 138 HN2 NHE  10    14.105 -10.116 -21.610 0.00 0.00
TER  139     NHE  10
END
```

AVP, clinched open45pbr (OPEN)*

AVP_23us_T16_18

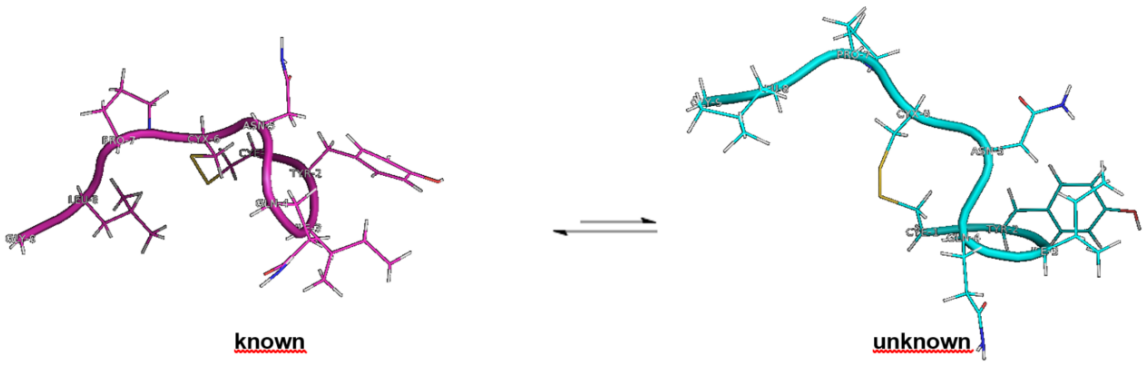
```
ATOM   1 N   CYX  1    -3.048  -7.364  19.990 0.00 0.00
ATOM   2 H1  CYX  1    -3.728  -6.660  20.239 0.00 0.00
ATOM   3 H2  CYX  1    -2.471  -7.606  20.783 0.00 0.00
ATOM   4 H3  CYX  1    -3.468  -8.226  19.675 0.00 0.00
ATOM   5 CA  CYX  1    -2.063  -6.812  19.029 0.00 0.00
ATOM   6 HA  CYX  1    -2.658  -6.365  18.233 0.00 0.00
ATOM   7 CB  CYX  1    -1.332  -5.654  19.794 0.00 0.00
ATOM   8 HB2 CYX  1    -2.049  -4.984  20.268 0.00 0.00
ATOM   9 HB3 CYX  1    -0.711  -5.999  20.621 0.00 0.00
ATOM  10 SG  CYX  1    -0.359  -4.593  18.738 0.00 0.00
ATOM  11 C   CYX  1    -1.148  -7.981  18.485 0.00 0.00
ATOM  12 O   CYX  1    -0.701  -8.868  19.249 0.00 0.00
ATOM  13 N   TYR  2    -0.872  -7.873  17.198 0.00 0.00
ATOM  14 H   TYR  2    -1.293  -7.126  16.663 0.00 0.00
ATOM  15 CA  TYR  2    -0.184  -8.915  16.362 0.00 0.00
ATOM  16 HA  TYR  2    -0.404  -9.889  16.800 0.00 0.00
ATOM  17 CB  TYR  2    -0.689  -8.927  14.830 0.00 0.00
ATOM  18 HB2 TYR  2    -0.272  -9.833  14.390 0.00 0.00
ATOM  19 HB3 TYR  2    -1.776  -9.002  14.792 0.00 0.00
ATOM  20 CG  TYR  2    -0.204  -7.779  14.084 0.00 0.00
ATOM  21 CD1 TYR  2     1.072  -7.856  13.431 0.00 0.00
ATOM  22 HD1 TYR  2     1.595  -8.751  13.733 0.00 0.00
ATOM  23 CE1 TYR  2     1.447  -6.928  12.438 0.00 0.00
ATOM  24 HE1 TYR  2     2.359  -6.973  11.862 0.00 0.00
ATOM  25 CZ  TYR  2     0.701  -5.762  12.309 0.00 0.00
ATOM  26 OH  TYR  2     1.146  -4.677  11.575 0.00 0.00
ATOM  27 HH  TYR  2     0.586  -3.905  11.468 0.00 0.00
ATOM  28 CE2 TYR  2    -0.437  -5.538  13.119 0.00 0.00
ATOM  29 HE2 TYR  2    -0.951  -4.596  13.005 0.00 0.00
ATOM  30 CD2 TYR  2    -0.970  -6.618  13.858 0.00 0.00
ATOM  31 HD2 TYR  2    -1.946  -6.524  14.312 0.00 0.00
ATOM  32 C   TYR  2     1.324  -8.837  16.489 0.00 0.00
ATOM  33 O   TYR  2     2.029  -9.782  16.089 0.00 0.00
ATOM  34 N   PHE  3     1.813  -7.716  17.046 0.00 0.00
ATOM  35 H   PHE  3     1.241  -7.006  17.479 0.00 0.00
ATOM  36 CA  PHE  3     3.235  -7.548  17.169 0.00 0.00
ATOM  37 HA  PHE  3     3.811  -8.344  16.696 0.00 0.00
ATOM  38 CB  PHE  3     3.523  -6.302  16.261 0.00 0.00
ATOM  39 HB2 PHE  3     4.610  -6.267  16.184 0.00 0.00
ATOM  40 HB3 PHE  3     3.245  -6.514  15.228 0.00 0.00
ATOM  41 CG  PHE  3     3.044  -4.923  16.670 0.00 0.00
ATOM  42 CD1 PHE  3     3.923  -4.072  17.277 0.00 0.00
ATOM  43 HD1 PHE  3     4.904  -4.457  17.514 0.00 0.00
ATOM  44 CE1 PHE  3     3.582  -2.707  17.414 0.00 0.00
ATOM  45 HE1 PHE  3     4.277  -2.065  17.935 0.00 0.00
ATOM  46 CZ  PHE  3     2.366  -2.186  16.854 0.00 0.00
ATOM  47 HZ  PHE  3     2.156  -1.134  16.977 0.00 0.00
ATOM  48 CE2 PHE  3     1.477  -3.043  16.266 0.00 0.00
ATOM  49 HE2 PHE  3     0.574  -2.619  15.851 0.00 0.00
ATOM  50 CD2 PHE  3     1.834  -4.391  16.118 0.00 0.00
ATOM  51 HD2 PHE  3     1.180  -5.062  15.581 0.00 0.00
ATOM  52 C   PHE  3     3.779  -7.214  18.561 0.00 0.00
ATOM  53 O   PHE  3     3.019  -6.851  19.433 0.00 0.00
ATOM  54 N   GLN  4     5.113  -7.426  18.817 0.00 0.00
ATOM  55 H   GLN  4     5.714  -7.621  18.030 0.00 0.00
ATOM  56 CA  GLN  4     5.722  -7.217  20.123 0.00 0.00
ATOM  57 HA  GLN  4     5.097  -7.714  20.865 0.00 0.00
ATOM  58 CB  GLN  4     7.018  -7.958  20.189 0.00 0.00
ATOM  59 HB2 GLN  4     7.684  -7.525  19.442 0.00 0.00
ATOM  60 HB3 GLN  4     7.324  -7.986  21.235 0.00 0.00
ATOM  61 CG  GLN  4     6.921  -9.482  19.899 0.00 0.00
ATOM  62 HG2 GLN  4     6.367  -9.963  20.706 0.00 0.00
ATOM  63 HG3 GLN  4     6.355  -9.639  18.981 0.00 0.00
ATOM  64 CD  GLN  4     8.268 -10.173  19.806 0.00 0.00
ATOM  65 OE1 GLN  4     8.859 -10.420  20.848 0.00 0.00
ATOM  66 NE2 GLN  4     8.752 -10.539  18.629 0.00 0.00
ATOM  67 HE21 GLN 4     9.665 -10.971  18.632 0.00 0.00
ATOM  68 HE22 GLN 4     8.168 -10.409  17.816 0.00 0.00
ATOM  69 C   GLN  4     5.812  -5.689  20.483 0.00 0.00
ATOM  70 O   GLN  4     6.044  -4.897  19.565 0.00 0.00
ATOM  71 N   ASN  5     5.946  -5.355  21.710 0.00 0.00
ATOM  72 H   ASN  5     5.921  -6.062  22.430 0.00 0.00
ATOM  73 CA  ASN  5     6.219  -3.976  22.184 0.00 0.00
ATOM  74 HA  ASN  5     6.036  -4.118  23.250 0.00 0.00
ATOM  75 CB  ASN  5     7.696  -3.746  21.880 0.00 0.00
ATOM  76 HB2 ASN  5     8.165  -4.705  22.102 0.00 0.00
ATOM  77 HB3 ASN  5     7.867  -3.617  20.811 0.00 0.00
ATOM  78 CG  ASN  5     8.308  -2.472  22.513 0.00 0.00
ATOM  79 OD1 ASN  5     7.725  -2.009  23.514 0.00 0.00
ATOM  80 ND2 ASN  5     9.255  -1.732  22.063 0.00 0.00
ATOM  81 HD21 ASN 5     9.400  -0.812  22.456 0.00 0.00
ATOM  82 HD22 ASN 5     9.759  -2.069  21.255 0.00 0.00
ATOM  83 C   ASN  5     5.256  -2.930  21.570 0.00 0.00
ATOM  84 O   ASN  5     5.696  -1.822  21.203 0.00 0.00
ATOM  85 N   CYX  6     3.969  -3.266  21.486 0.00 0.00
ATOM  86 H   CYX  6     3.605  -4.085  21.951 0.00 0.00
ATOM  87 CA  CYX  6     2.946  -2.564  20.771 0.00 0.00
ATOM  88 HA  CYX  6     3.313  -2.107  19.852 0.00 0.00
ATOM  89 CB  CYX  6     1.858  -3.633  20.246 0.00 0.00
ATOM  90 HB2 CYX  6     2.290  -4.158  19.394 0.00 0.00
ATOM  91 HB3 CYX  6     1.699  -4.273  21.114 0.00 0.00
ATOM  92 SG  CYX  6     0.326  -2.904  19.596 0.00 0.00
ATOM  93 C   CYX  6     2.377  -1.414  21.568 0.00 0.00
ATOM  94 O   CYX  6     2.083  -1.545  22.805 0.00 0.00
ATOM  95 N   PRO  7     2.123  -0.240  20.916 0.00 0.00
ATOM  96 CD  PRO  7     2.489   0.135  19.580 0.00 0.00
ATOM  97 HD2 PRO  7     1.950  -0.492  18.870 0.00 0.00
ATOM  98 HD3 PRO  7     3.523  -0.075  19.305 0.00 0.00
ATOM  99 CG  PRO  7     2.147   1.594  19.375 0.00 0.00
ATOM 100 HG2 PRO  7     1.181   1.751  18.894 0.00 0.00
ATOM 101 HG3 PRO  7     2.941   2.094  18.821 0.00 0.00
ATOM 102 CB  PRO  7     2.189   2.180  20.825 0.00 0.00
ATOM 103 HB2 PRO  7     1.558   3.057  20.965 0.00 0.00
ATOM 104 HB3 PRO  7     3.222   2.370  21.118 0.00 0.00
ATOM 105 CA  PRO  7     1.720   0.995  21.693 0.00 0.00
ATOM 106 HA  PRO  7     2.311   1.147  22.596 0.00 0.00
ATOM 107 C   PRO  7     0.243   1.048  22.055 0.00 0.00
ATOM 108 O   PRO  7    -0.554   0.609  21.242 0.00 0.00
ATOM 109 N   ARG  8    -0.070   1.805  23.133 0.00 0.00
ATOM 110 H   ARG  8     0.662   2.264  23.656 0.00 0.00
ATOM 111 CA  ARG  8    -1.502   1.954  23.652 0.00 0.00
ATOM 112 HA  ARG  8    -2.119   1.081  23.437 0.00 0.00
ATOM 113 CB  ARG  8    -1.490   2.076  25.152 0.00 0.00
ATOM 114 HB2 ARG  8    -2.510   2.131  25.532 0.00 0.00
ATOM 115 HB3 ARG  8    -1.058   1.161  25.556 0.00 0.00
ATOM 116 CG  ARG  8    -0.770   3.364  25.683 0.00 0.00
ATOM 117 HG2 ARG  8     0.129   3.602  25.114 0.00 0.00
ATOM 118 HG3 ARG  8    -1.487   4.184  25.671 0.00 0.00
ATOM 119 CD  ARG  8    -0.444   3.262  27.191 0.00 0.00
ATOM 120 HD2 ARG  8    -0.138   2.226  27.333 0.00 0.00
ATOM 121 HD3 ARG  8     0.413   3.881  27.457 0.00 0.00
ATOM 122 NE  ARG  8    -1.591   3.710  27.967 0.00 0.00
ATOM 123 HE  ARG  8    -1.854   4.665  27.772 0.00 0.00
ATOM 124 CZ  ARG  8    -2.495   3.045  28.712 0.00 0.00
ATOM 125 NH1 ARG  8    -2.369   1.787  29.147 0.00 0.00
ATOM 126 HH11 ARG 8    -1.458   1.369  29.020 0.00 0.00
ATOM 127 HH12 ARG 8    -3.140   1.246  29.511 0.00 0.00
ATOM 128 NH2 ARG  8    -3.688   3.556  29.012 0.00 0.00
ATOM 129 HH21 ARG 8    -3.967   4.483  28.726 0.00 0.00
ATOM 130 HH22 ARG 8    -4.229   3.007  29.665 0.00 0.00
ATOM 131 C   ARG  8    -2.216   3.097  22.954 0.00 0.00
ATOM 132 O   ARG  8    -1.513   3.915  22.337 0.00 0.00
ATOM 133 N   GLY  9    -3.496   2.998  22.888 0.00 0.00
ATOM 134 H   GLY  9    -3.993   2.252  23.354 0.00 0.00
ATOM 135 CA  GLY  9    -4.316   4.086  22.391 0.00 0.00
ATOM 136 HA2 GLY  9    -3.989   4.356  21.387 0.00 0.00
ATOM 137 HA3 GLY  9    -5.346   3.736  22.455 0.00 0.00
ATOM 138 C   GLY  9    -4.159   5.330  23.287 0.00 0.00
ATOM 139 O   GLY  9    -3.600   5.253  24.395 0.00 0.00
ATOM 140 N   NHE  10   -4.561   6.490  22.866 0.00 0.00
ATOM 141 HN1 NHE  10   -5.196   6.552  22.084 0.00 0.00
ATOM 142 HN2 NHE  10   -4.145   7.346  23.204 0.00 0.00
TER  143     NHE  10
ATOM 144 Cl- Cl-  11    10.024  -9.713  -3.281 0.00 0.00
TER  145     Cl-  11
ATOM 146 Cl- Cl-  12    -0.806  -6.929   1.156 0.00 0.00
TER  147     Cl-  12
END
```

A 9: Presentations and Talks

The Necessity of Long-term Molecular-Dynamics Simulations: Deamino-Oxytocin - Novel Conformational Insights

(Abstract, Poster)

Haensele E, Banting L, Clark T. The Necessity of Long-term Molecular-Dynamics Simulations: Deamino-Oxytocin - Novel Conformational Insights. (a) 26th Molecular Modeling Workshop, March 12th, 2012. Erlangen, Germany. (b) IBBS Day, May 11th, 2012. University of Portsmouth, UK. Abstract: http://mmws2012.mgms-ds.de



Extended molecular-dynamics (MD) simulations (> 1 µs) show great promise in delivering significant, practically relevant, insight into conformational processes that occur within molecular systems. If long enough, MD simulations can reveal conformational interconversions particularly in peptides and proteins. A conformational equilibrium may be unfavourable and dominated by the highly populated more stable conformation. However, the less favoured conformer is often the physiologically relevant one and may present significant difficulties for quantification by experimental techniques. Close coordination of MD analysis and experiment helps shed light on pharmacologically relevant molecular phenomena. This work is part of a series of long-term MD simulations (1) (≥ 3 µs) applied to the cyclic nonapeptides oxytocin, Arg^8-vasopressin, and deamino-oxytocin (dOT). Their moderate size and multitude of structural features presents an ideal test case to emphasise the necessity of long-term simulations and to apply diverse conformational-analysis methods (2, 3). The MD on dOT shows that (i) the results achieved with a runtime of 3 µs are in very good agreement with experimental data (4, 5) and (ii) employing DASH (2) in the analysis of these systems proves powerful and reliable in characterising conformational clusters. Furthermore, a previously undetected ring conformation of dOT was significantly populated in the simulation trajectory (390 ns/ 3000 ns, 8 transitions). This conformation indicates greater conformational flexibility of dOT *vs.* OT/ VP and thus helps explain its super-agonist properties (6).

(1) D.A. Case, T.A. Darden, T.E. Cheatham II, *et al.*, AMBER 10, 2008. (2) M.J. Ellis, University of Portsmouth, 2007, PhD, 1-227. (3) D.A. Case, T.A. Darden, T.E. Cheatham II, *et al.*, AmberTools 1.0, 2008. (4) S.P. Wood, I.J. Tickle, A.M. Treharne, *et al.*, Science, 1986, 232, 633-636. (5) J. Husain, T.L. Blundell, S. Cooper, *et al.*, Philos Trans R Soc Lond B Biol Sci, 1990, 327, 625-654. (6) B.M. Ferrier, D. Jarvis, and V. Du Vigneaud, J Biol Chem, 1965, 240, 4264-4266.

University of Portsmouth

The necessity of long-term molecular dynamics simulations

Deamino-oxytocin

Novel conformational insights

Elke Haensele, Lee Banting and Timothy Clark
Faculty of Science, School of Pharmacy and Biomedical Sciences, Center for Molecular Design, UK

Deamino-oxytocin

Deamino-oxytocin (dOT) is a synthetic analog of Oxytocin (OT) which lacks the N-terminal amino-group. OT, a hypophyseal peptide hormone, inducing *i.e.* milk ejection and uterine muscle contractions [2]. Its deamino analog dOT is biologically significantly more active (super-agonist) [3].

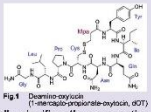


Fig.1 Deamino-oxytocin (1-ethane-β-mercaptate-oxytocin, dOT)

Structural features: A cyclic 6-residue-moiety ("ring"), a disulphide bridge between residue 1 and 6, a 3-peptide C-terminal ("tail").

Experimental studies: Numerous spectroscopic studies in the 1970s and 80s generated ambiguous results [4-7].1986 X-ray: 1XY1 ($P2_1$) and 1XY2 ($C2$), crystal structures with the same backbone conformations [8].

Conformational characteristics (1XY1/ 1XY2): 2-3-4-5 type II beta turn (ring), 6-7-8-9 type III turn (tail), intramolecular hydrogen bonds 2O5H, 5O2H, 6O9H

Methods

Molecular-dynamics simulations: Amber 10 [9]: Initial conformation dOT 1XY1; TIP4P truncated octahedron water box/ periodic boundaries [10]; non-bonded cutoff 8 Å; all-atom force-field ff99SB; Minimisation: 500 cycles steepest decent minimisation, 9500 cycles conjugated-gradient; Dynamics simulations: T 300K (Berendsen coupling [11]), constant density, PME method [12], 3000 ns production.

DASH clustering: DASH 2.10b1 [13] , a fast conformation analysis tool developed for very long MD trajectories, where cartesian methods scale poorly, analyses torsion angles from MD trajectories as time series. A DASH state is a group of similar torsion angles. Main DASH states are average torsion-angle sets which represent a conformational group or cluster *vs.* time. 5000 sets of ring torsions (Phi2/Psi2 to Phi6/Psi6) over a trajectory of 3 µs were clustered.

Conformational analysis: Ptraj/ Amber Tools 1.0 [14]: RMSD (Root Mean Square Deviation), intramolecular distances, hydrogen bond analysis, secondary structure analysis

Main ring conformations

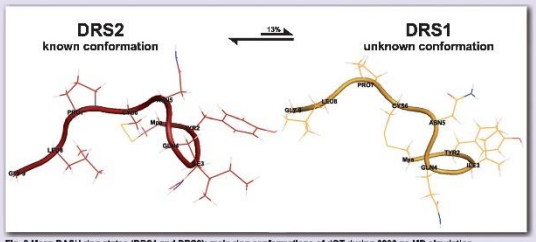


Fig. 2 Mean DASH ring states (DRS1 and DRS2): main ring conformations of dOT during 3000 ns MD simulation
(backbones: cartoon; side chains: lines; N-terminal analogue: Mpa; C-terminal: Gly9; ring: Mpa to Cys6; tail Pro7-Leu8-Gly9)

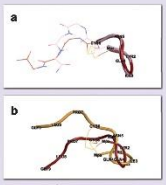


Fig. 3 Ring-fitted backbone superpositions of (a) "known" representative ring conformation DRS2 (populated with 87%/ 3000 ns) *vs.* 1XY1 (dOT X-ray structure); and (b) DRS2 *vs.* DRS1, the "unknown" representative ring conformation (populated with 13%/ 3000 ns)

Long-term MD

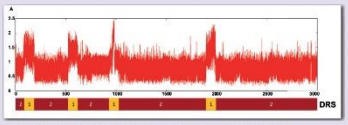


Fig. 5 RMSD (CA ring) *vs.* time [ns] and mean DASH ring states (DRS)

MD simulation

A 3000 ns free molecular-dynamics simulation at 300 K for dOT in aequous solution was carried out and clustered by DASH. Significant conformational changes of ring-RMSD values correlate precisely with the mean DASH ring states DRS2 and DRS1 presented here.

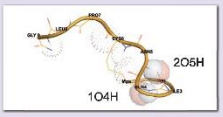
DASH clustering

DASH ring state DRS2, the main ring conformation cluster populating 87% of the 3000 ns trajectory, matches the X-ray conformation 1XY1 of dOT. DASH ring state DRS1 is a second significantly populated unexpected conformation (13%). There are eight transitions between the two states along the trajectory.

Hydrogen bonds and secondary structures

Deamino-oxytocin (dOT) sequence
1MPA-2TYR-3ILE-4GLN-5ASN-6CYS-7PRO-8LEU-9GLY

DRS2 95% **2-3-4-5 turn**



DRS1 50% **2-3-4-5 turn**
35% **3_{10}-helical turn**
28% **6-7-8-9 turn**

Fig. 4 Intramolecular hydrogen bonds
Amide protons and backbone carbonyl oxygen atoms are depicted as dotted spheres if the hydrogen bonding interactions occur with more than 60-100% (*vs.* time). Light dots indicate hydrogen bonding interactions with occupancies from 30-60% (*vs.* time).

Table 1 Population* of intramolecular hydrogen bonds and secondary structures of dOT's main conformation clusters DRS2$_{known}$ and DRS1$_{unknown}$

Hydrogen bonds	DRS2 „known"	DRS1 „unknown"
$2O5H_{i\rightarrow i+3}$	73%	66%
$2O6H_{i\rightarrow i+3}$	70%	-
$3O5H_{i\rightarrow i+3}$	29%	3%
$1O4H_{i\rightarrow i+3}$	-	62%
$4O6H_{i\rightarrow i+3}$	-	36%
$8O9H_{tail}$	6%	39%

Secondary structure	DRS2 „known"	DRS1 „unknown"
$turn_{ring}$ 2-3-4-5	95%	50%
3_{10}-$Helix_{ring}$ 1-2-3-4	-	35%
$turn_{tail}$ 6-7-8-9	6%	28%

* Hydrogen bonds were calculated for backbone amide protons and carbonyl oxygens with a distance cutoff of dNH-O = 3.0 Å, and an angle cutoff of ∠NHO = 120°. Results are given as percentage of occupancy (population) relative to occurance periods (boats) of the two main DASH states

Hydrogen bond analysis and secondary structure analysis

The ring conformation of DRS2 is characterised by a 2-3-4-5 turn stabilised by transannular hydrogen bonds 2O5H and 2O6H. The latter different to the X-ray structure of 1XY1 but like 1NPO (X-ray structure of OT) [15]. These data, among others not presented, correspond very well with experimental data.

DRS2 has not been detected by spectroscopy so far. The main transannular hydrogen bond is still 2O5H, but the 2-3-4-5 turn is weakened in favour of a 1-2-3-4 3_{10}-helical turn characterised by a 1O4H bond.

Necessity of long-term MD simulations

Without long enough simulation times it is often not possible to discover alternate conformation clusters, even less series of significant conformational transitions, but such series point to possible equilibria. The less favoured conformer of such equilibria may be the physiologically relevant one and may present technical difficulties for experimental quantification.

New conformational insight

A new conformation of dOT, with a 3_{10}-helical torsion of the ring, has been found. OT and AVP appear not to adopt this conformation. It is evident that the missing N-terminal amino group supports the formation of this new conformation due to less steric and electrostatic hindrance. The ring-moiety of dOT has a larger degree of flexibility than its natural analogues, which is a highly plausible reason for its super-agonist properties.

Co-joining of "theory and practise"

Unexpected results from either experimental or theoretical research give fresh impetus and benefit from synergy and mutual reinforcement shedding light on pharmacologically relevant molecular phenomena.

References

1. V. Du Vigneaud, C. Ressler, and S. Trippett, *J Biol Chem*, **1953**, 205, 949-957.
2. G. Gimpl and F. Fahrenholz, *Physiol Rev*, **2001**, 81, 629-683 and references therein.
3. B.M. Ferrier, D. Jarvis, and V. Du Vigneaud, *J Biol Chem*, **1955**, 240, 4264-4266.
4. L.F. Johnson, I.L. Schwartz, and R. Walter, *Proc Natl Acad Sci U S A*, **1969**, 64, 1299-1275.
5. D.W. Urry, M. Ohnishi, and R. Walter, *Proc Natl Acad Sci U S A*, **1970**, 65, 111-116.
6. A.I. Brewster, V.J. Hruby, et al., *Biochemistry*, **1973**, 12, 5294-5304.
7. J.G. Glickson, R. Rowan, et al., *Biochemistry*, **1975**, 15, 1111-1119.
8. S.P. Wood, I.J. Tickle, et al., *Science*, **1986**, 232, 633-636.
9. D.A. Case, T.A. Darden, et al., *Amber 10*, **2008**.
10. W.L. Jorgensen, J. Chandrasekhar et al., *J Chem Phys*, **1983**, 79, 926-935.
11. H.J.C. Berendsen, J.P.M. Postma, et al., *J Chem Phys*, **1984**, 81, 3684-3690.
12. T. Darden, D.M. York, and L. Pedersen, *J Chem Phys*, **1993**, 98, 10089-10092.
13. D.W. Salt, B.D. Hudson, L. Banting, et al., *J Med Chem*, **2005**, 48, 3214-3225.
14. D.A. Case, T.A. Darden, et al., *AmberTools 1.0*, **2008**.
15. J.P. Rose et al., *Nat Struct Biol*, **1996**, 3, 163-169.

DASH release 2.10b1

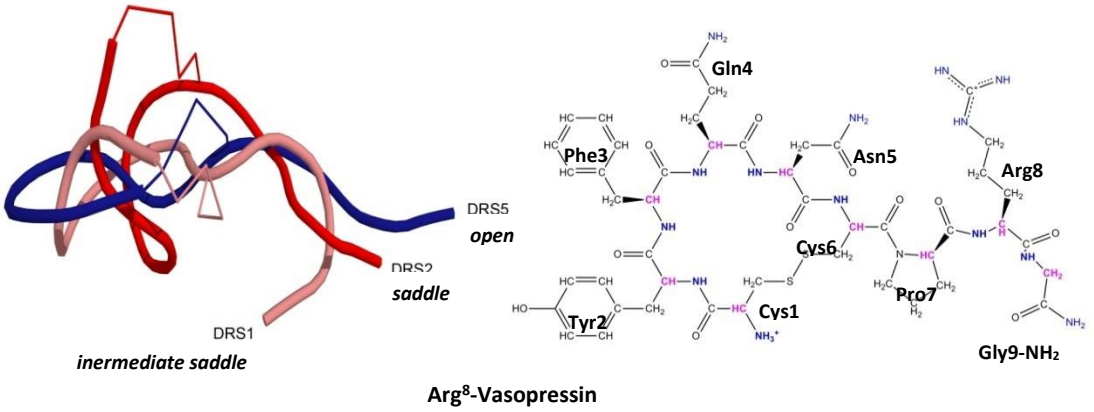
The program is released under the terms of the General Public Licens GNU and can be downloaded from the CMD (Center for Molecular Design, University of Portsmouth, UK) website: ww.port.ac.uk/research/cmd/software

Molecular-Dynamics and Umbrella-Sampling Simulations of Arg^8-Vasopressin

(Abstract, Poster)

Haensele E, Banting L, Clark T. Molecular-Dynamics and Umbrella-Sampling Simulations of Arg^8-Vasopressin. (a) 27th Molecular Modeling Workshop, Feb 25th, 2013. FAU Erlangen-Nürnberg, Germany. (b) IBBS Day, Jun 7th, 2013. University of Portsmouth, UK.

Abstract: http://mmws2013.mgms-ds.de



Arg^8-Vasopressin

Arg^8-Vasopressin (AVP) is a neurohypophyseal hormone with a wide range of endocrinological and neurological functions, *e.g.* water homeostasis, blood pressure regulation and mediation of social and sexual behaviour. Main structural characteristics are a 6-residue ring closed *via* disulphide bridging, and an α-amidated 3-residue tail. Figure: Structure and backbone conformations (blue: *open*; red: *saddle*; rose: *intermediate*; cartoon: backbones; sticks: disulphide bridges; not shown: sidechains)

A long-term (5 µs) molecular-dynamics simulation of Arg^8-vasopressin was performed in aqueous solution at 300 K. Two main conformational ring states were identified *via* DASH (1) analysis: DRS_{open}, a stretched, *open* conformation with no intramolecular hydrogen bonds in the ring; and DRS_{saddle}, a folded, *saddle*-like conformation with strong hydrogen bonding interactions between the carbonyl oxygen of the ring residue Tyr^2 and the amide protons of the ring residues Asn^5 and Cys^6. Only one transition between both main states was observed during the 5µs simulation run. In addition to these two main states, a sparsely populated DASH state, $DRS_{intermediate}$, was found with mixed conformational characteristics of the two main states. Umbrella Sampling (2, 3), post-processed with WHAM (4-6), was used to estimate the free energy profile for the conformational change from *open* to *saddle* and led to a reaction path *via* $DRS_{intermediate}$ (see video clip (7)), with barrier heights of 7.7 kcal mol^{-1} and 14.2 kcal mol^{-1} and a free energy difference between the *open* and *saddle* states of 4.0 kcal mol^{-1}.

(1) D.W. Salt, B.D. Hudson, L. Banting, *et al.*, J. Med. Chem., 2005, 48, 3214-3220. (2) G.M. Torrie and J.P. Valleau, J. Comput. Phys., 1977, 23, 187-199. (3) G.M. Torrie and J.P. Valleau, Chem. Phys. Lett., 1974, 28, 578-581. (4) S. Kumar, J.M. Rosenberg, D. Bouzida, *et al.*, J. Comput. Chem., 1992, 13, 1011-1021. (5) M. Souaille and B.T. Roux, Comput. Phys. Commun. , 2001, 135, 40-57. (6) A. Grossfield, WHAM, Version 2.0.1, 2000.

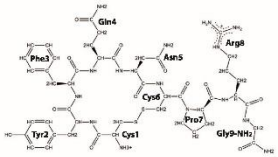
[7] https://www.youtube.com/watch?v=z0aRtSxNQ2I

Molecular Dynamics and Umbrella Sampling Simulations of 8-Arg-Vasopressin

Elke Haensele, Lee Banting and Timothy Clark
Faculty of Science, School of Pharmacy and Biomedical Sciences, Center for Molecular Design, UK

8-Arg-Vasopressin

8-Arg-Vasopressin (**AVP**) belongs to a family of highly preserved neuropeptides that exist in all animals. Structural characteristics of these hormones are a cyclic moiety of 6 residues, closed via a disulphide-bridge and a 3-residue tail with an α-amidated C-terminal. AVP, found in mammals, is responsible for a wide range of neuroendocrinological functions, inter alia pressor activity, antidiuretic effects, and mediation of social and sexual behaviour. These biological effects are developed via interaction with G-protein coupled vasopressin-receptors (for review, see *e.g.* [1]).



8-Arg-Vasopressin

Objectives and methods

Without knowledge of the intrinsic dynamic nature of receptor and ligand, even in the unbound state, an understanding of receptor activation is not possible. To study the molecular dynamics (MD) of AVP a 5 µs Amber10 (parm99SB) force-field simulation at 300 K in water (TIP4P) was carried out. The conformational space was clustered according to backbone torsion-angles via DASH [2] to identify significant conformational states. Free energy profiles were estimated for transitions between various conformations via umbrella sampling [3], post-processed with WHAM [4].

What is the bioactive conformation of 8-Arg-Vasopressin?

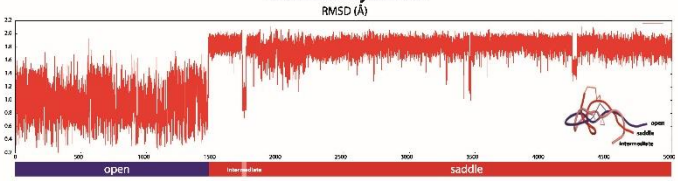
Molecular dynamics: Conformational states in solution

Open
The main conformational state during 0-1460 ns (98.3% occupancy) has a stretched, ***open*** backbone shape and **no intramolecular hydrogen bonds** inside the ring. The state corresponds to the MD starting-conformation of AVP, 1YF4, a trypsin complex [5]. 1YF4 is the only fully resolved X-ray structure of AVP to date.

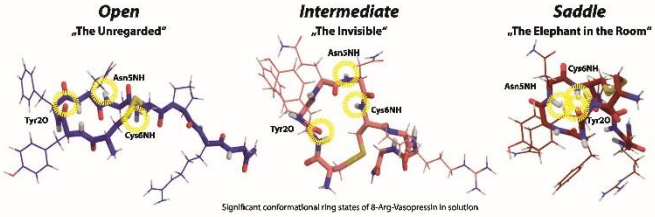
Saddle
At 1460 ns, a spontaneous conformational change takes place. The new state is characterised by a ***saddle***-like backbone shape and significant **intramolecular hydrogen bonding** interactions of Tyr2´s carbonyl-oxygen and the amide-hydrogens of Asn5 and Cys6. This *saddle* state remains until the end of the simulation (1460-5000 ns, 97.8% occupancy). The saddle conformation is, for example, known from the neurophysin complex of Lys-VP, 1JK4 [6].

Intermediate
In addition to the two conformational main states, *open* and *saddle*, a random conformation was found (0.8%, 5 µs) that shows conformational features of both main states: a ***saddle***-like backbone ring-conformation but, like the *open* conformation, with **no intramolecular hydrogen bonds**.

Molecular dynamics



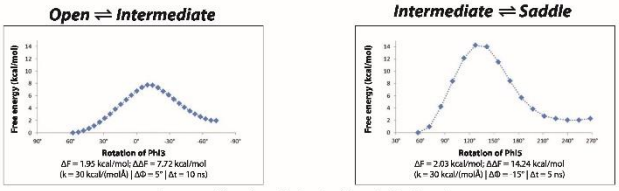
Molecular dynamics of 8-Arg-Vasopressin (runtime = 5000 ns; transition at 1460 ns)



Significant conformational ring states of 8-Arg-Vasopressin in solution

Free energy: Pathways of conformational changes

Free energy profiles were estimated via umbrella sampling. Starting with the *open* conformation, an inwards turn of the carbonyl-O of Tyr2 (via rotation of *Phi3*) forces the backbone to change to the *intermediate* conformation. The second path starts with the *intermediate* and describes the rotation of *Phi5*, which turns the amide-NH of Asn5 into the ring and leads to the formation of the ring-internal hydrogen bonds, characteristic for the *saddle* conformation.



Free energy profiles: pathways of conformational changes for 8-Arg-Vasopressin

Discussion

The dynamic analyses of AVP reveal two main backbone conformations to be considered as possible bioactive ligand conformations: *open* and *saddle*.
As the trypsin complex of AVP (*open*) is, if ever, of only little biological importance [5], in contrast to its neurophysin complex (*saddle*) [1], not much attention has been directed toward the *open* conformation (***"The Unregarded"***) up to now. NMR-studies also indicate that the *saddle* conformation (***"The Elephant..."***) is the main conformation in aqueous solution [7].
Nevertheless, the random state, *intermediate* (***"The Invisible"***), may also be a high-active conformation, even if its lifetime on the MD scale is short. This is because the population of a kinetically unfavoured ligand conformation may shift via allosteric events [8].

Video clips of 8-Arg-Vasopressin Molecular Dynamics and Umbrella Sampling on YouTube:
keyword "molecular dynamics vasopressin"
https://www.youtube.com/watch?v=ETjyKJymYeA (Molecular Dynamics)
https://www.youtube.com/watch?v=z0aRtSxNQ2I (Umbrella Sampling)

References and supplementary data

1. a) Laycock JF (2010) Perspectives on vasopressin;
b) Strand FL (1999) Neuropeptides: regulators of physiological processes pp 229-265
2. Salt DW, Hudson BD et al (2005) J Med Chem 48:3214-3220
3. a) Torrie GM, JP Valleau (1977) J Comput Phys 23:187-199;
b) Torrie GM, Valleau JP (1974) Chem Phys Lett 28:578-581
4. a) Grossfield A (2000) WHAM Version 2.0.1;
b) Kumar JM, Rosenberg JM et al (1992) J Comput Chem 13:1011-1021;
c) Souaille M, Roux BT (2001) Comput Phys Commun 135:40-57
5. Syed Ibrahim B, Pattabhi V (2005) J Mol Biol 348:1191-1198
6. Wu CK, Hu B, Rose JP et al (2001) Pro Sci 10:1869-1880
7. Sikorska, E, Rodziewicz Motowidlo S (2008) J Pept Sci 14:76-84
8. Gunasekaran K, Ma B et al (2004) Proteins 57:433-443

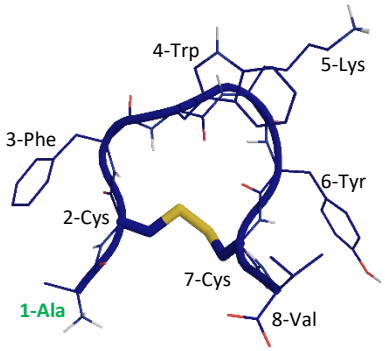
Molecular dynamics data: Amber 10; parm99SB; T=300K; p=1atm; TIP4P oct. 8 Å cutoff; periodic boundaries; SHAKE algorithm; PME method. DASH (dynamics analysis by Salt and Hudsons): torsion angle analysis method via Markov process; advantage: fast scaling of long trajectories, no predetermined number of clusters; result: time series of states = representative conformations. DASH data: 10.000 snaps of 5µs MD (= 2 snaps/ns) | Phi/Psi 2-6 (= backbone ring conformations). WHAM (weighted histogram analysis method).

Co-funding of PeReNE (The European project "Peptide Research Network of Excellence") and Interreg EU (Interreg IVA France (Channel) – England 2007-2013 programme) is gratefully acknowledged.

Urotensin-Related Peptide (URP): Long-term Molecular-Dynamics Simulation

(Abstract, Poster)

Haensele E, Banting L, Clark T. Urotensin-Related Peptide (URP): Long-term Molecular-Dynamics Simulation. (a) 28th Molecular Modeling Workshop, Mar 18th, 2014. FAU Erlangen-Nürnberg, Germany. Abstract: http://mmws2014.mgms-ds.de



Urotensin-related peptide: Ala-[Cys-Phe-Trp-Lys-Tyr-Cys]-Val
(Human-UII: Glu-Thr-Pro-Asp-[Cys-Phe-Trp-Lys-Tyr-Cys]-Val)

The hormone peptides URP (urotensin-related peptide) and U-II (urotensin II) are the natural ligands of the urotensinergic GPCR (G-protein coupled receptor) system, which plays an important role in the regulation of the cardiovascular system. Besides their physiological function, URP and U-II are also linked to pathophysiological processes such as hypertension (1). URP is an octapeptide with a six-residue ring closed by a 2Cys-7Cys-disulphide bridge, a 1-Ala N-terminal and an 8-Val C-terminal. URP differs from U-II only in the length of the N-terminal and is thus a prototype for the ring-system of these hormone peptides. Both the ring-residues Trp-Lys-Tyr and the disulphide bridge are thought to be important for receptor activation (1). Understanding the dynamic conformational properties of URP can help develop pharmacophores and direct simulations of the receptor. We describe a 5 µs molecular-dynamics simulation of URP that demonstrates the high flexibility of the peptide. *DASH* (2) analysis reveals several distinct main and transient conformational states that interchange rapidly. These states will be characterised and their properties discussed with some focus on the conformation of the disulphide bridge.

(1) D. Chatenet, T.T. Nguyen, M. Letourneau, A. Fournier, Front Endocrinol, 2012, 3, 7-13

(2) D.W. Salt, B.D. Hudson, L. Banting, M.J. Ellis, M.G. Ford, J Med Chem, 2005, 48, 3214-3220

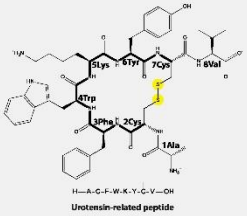
Long-term Molecular-Dynamics Simulation

Urotensin-Related Peptide (URP)

Elke Haensele, Lee Banting, David Whitley and Timothy Clark
School of Pharmacy and Biomedical Sciences, Center for Molecular Design, University of Portsmouth, UK

Structure determination

Small peptides are often very flexible, which makes an experimental structure determination difficult. Long-term molecular-dynamics simulations (MD) help sample their conformational space effectively. Here, we present a **5µs MD simulation of URP** (300K, explicit water solvation, Amber ff99SB). Representative states were determined by analysing time series of torsion angles using ***DASH*** [1].
The modular structure of the system is shown (Fig. 3) with special focus on the conformation of the ring and the disulphide bridge (Fig. 4,5).
The results are useful for pharmacophore studies, simulation of receptor-ligand complexes, and to simulate NMR spectra. Elucidation of dynamical conformational properties helps understand allosteric mechanisms of ligand/receptor interactions.



Urotensin-related peptide (URP)

URP is a cyclopeptide consisting of a 6-residue ring closed by a disulphide bridge, a 1Ala N-terminus and an 8Val C-terminus. It is a paralog of Urotensin-II (UII), a large family of G-protein coupled receptor ligands found in many species. The N-terminus of UII is highly variable in length and sequence, whereas the cyclic C-terminus is conserved for all vertebrates and thought to be responsible for receptor activation. UII and URP are vasoactive and strongly implicated in cardiovascular homeostasis. [2]
The present molecular-dynamics study investigates the conformational and dynamical properties of URP as a prototyp for the UII ring-system and is a starting point for comparative studies on human-UII.

High flexibility does not exclude structure

Molecular dynamics and representative conformational states of URP

Although highly conformationally mobile, URP exhibits distinctly modular structure (Fig. 3):

Two main ring states, **Ω-4O7H** and **Ω-open** (Fig. 2) interconvert readily.
Ω-4O7H is characterised by a **type-I 5,6-beta turn** with a 4O7H transannular hydrogen bond.
Ω-open has similar backbone shape to **Ω-4O7H**-residues 3-6, but lacks intramolecular hydrogen bonds.
The ring state **Ω-4O7H** exhibits two distinct 8Val (C-terminal tail) positions, **endo** and **exo**. These substates are equally populated and interconvert frequently. The endo 8Val enables the additional hydrogen bond 4O8H.
The ring-torsion **Psi6**, linked to a rotation of the 6Tyr7Cys-peptide bond, is the key torsion for transitions between the two main conformations **Ω-4O7H** and **Ω-open**.
In addition, the ring may be classified by 4 disulphide-bridge states (Fig. 3-5).
Besides the two main ring states **Ω-4O7H** and **Ω-open** that occupy 94% of the simulation time, two transient states, **Ω-hybrid** and **β-sheet**, were found (Fig. 6).

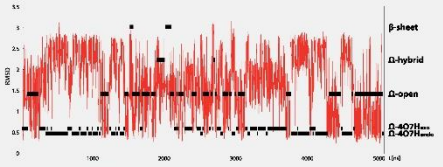
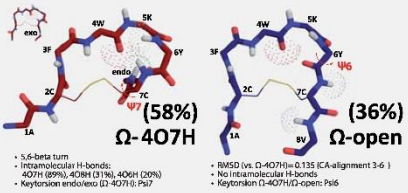


Fig. 1 Molecular dynamics of URP
RMSD (CA-backbone alignment, red) of 5µs MD simulation of URP and the *DASH* state trajectories (black) of the main backbone-states **Ω-4O7H**, **Ω-4O7Hext**, **Ω-open**, and the transient states **Ω-hybrid** and **β-sheet**. There are frequent **Ω-4O7Hendo/Ω-4O7Hext** transitions; **Ω-4O7H/Ω-open** transitions are most frequent between 2000 and 3000 ns which might correlate with the disulphide bridge conformation (cf. Fig. 4).



- 5,6-beta turn
- Intramolecular H-bonds: 4O7H (89%), 4O8H (31%), 4O6H (20%)
- Keytorsion endo/exo (Ω-4O7H): Psi7

- RMSD (vs. Ω-4O7H)= 0.135 (CA-alignment 3-6)
- No intramolecular H-bonds
- Keytorsion Ω-4O7H/Ω-open: Psi6

Fig. 2 Main representative states of URP
(small picture= Ω-4O7Hext; sticks= backbones; lines= disulphide bridges; dots= 4O, 7H, 8H)

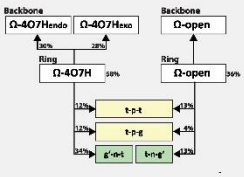


Fig. 3 Modular structering of URP
Main conformational states of URP and disulphide-bridge substates; (disulphide-bridge conformation 2χ2-2χ3-7χ2: t=trans, p=positive (+86°), n = negative (-86°), g=gauche)

Disulphide bridge: A flexibility switch?

The average disulphide torsion 2χ3 is either +86° (p) or -86° (n) with no preferred handedness. Main conformations are shown in Figure 5. The transition propensity between **Ω-open** and **Ω-4O7H** in regions with a positive 2χ3 torsion (yellow) is significantly higher. This observation gives reason to speculate about the bioactive function of the disulphide bridge: Is the SS-conformation a flexibility switch?

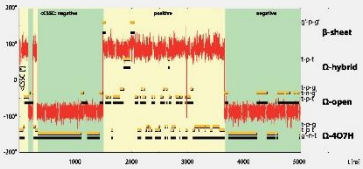


Fig. 4 Disulphide bridge torsion CSSC (red), ***DASH* ring state trajectories** (T10, black) and ***DASH* ring state trajectories including disulphide-bridge torsions** (T13ss, yellow). MD sections with negative CSSC torsion = green; sections with positive CSSC torsion = yellow.

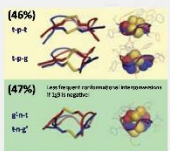
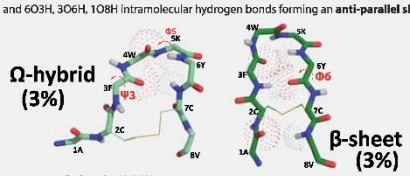


Fig. 5 Main disulphide bridge coformations of URP. (blue= Ω-open, red= Ω-4O7H, yellow= SS)

Transient states

One should not dismiss the potential importance of transient states as possible keys to bioactivity. In the 5 µs MD of URP, two interesting transient states were identified, the **Ω-hybrid** and **β-sheet** state. **Ω-hybrid** is a structural intermediate between the main Ω-shaped states and **β-sheet**. **β-sheet** is a highly structured conformer with a **4,5-turn** and 6O3H, 3O6H, 1O8H intramolecular hydrogen bonds forming an **anti-parallel sheet**.



- Conformational hybrid between Ω-states and β-sheet
- Intramolecular H-bonds: 4O6H (70%), 3O6H (39%), 3O7H (29%)

- Highly structured conformation
- Intramolecular H-bonds: 6O3H (88%), 1O8H (74%), 3O6H (48%), 4O6H (34%)
- Keytorsion Ω-4O7H/Ω-hybrid = Phi6

Fig. 6 Transient representative states of URP
(sticks= backbones; lines= disulphide bridges; dots= 1O, 3H, 3O, 4O, 5H, 6O, 8H)

References
1. Salt DW, Hudson BD, Banting L, Ellis MJ, Ford MG (2005) J Med Chem 48 (9):3214-3220
2. Vaudry H, Do Rego JC, Le Mevel JC, Chatenet D, Tostivint H, Fournier A, Tonon MC, Pelletier G, Conlon JM, Leprince J (2010) Ann NY Acad. Sci 1200:53-66
3. Chatenet D, Dubessy C, Leprince J, Boularan C, Carlier L, Ségalas-Milazzo I, Guilhaudis L, Oulyadi H, Davoust D, Scalbert E, Pfeiffer B, Renard P, Tonon M-C, Lihrmann I, Pacaud P, Vaudry H (2004) Peptides 25 (10):1819-1830

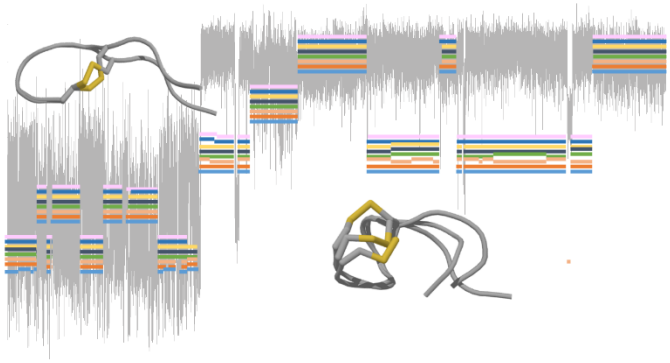
Molecular dynamics simulation
"URPnmr:" NMR structure [3] | 5 µs PMEMD | Amber ff99sb [300K | TIP4PEw/PBC | cutoff 8Å | SHAKE
***DASH* [Dynamic Analysis by Salt and Hudson]**
T14: URPnmr 5µs | Psi1, PhiPsi2-7, Phi8 | 10,000 snaps | i=20(default); T10: URPnmr 5µs | Psi2, PhiPsi3-5, Phi6 | 10,000 snaps | i=20(default); T13ss: URPnmr 5µs | Psi2, PhiPsi3-5, Phi6, 2X2, 2X3, 6X2 | 10,000 snaps | i=20(default)

Co-funding of PeReNE (The European project "Peptide Research Network of Excellence") and Interreg EU (Interreg IVA France (Channel)–England programme) is gratefully acknowledged.

DASH: Analysis of Microsecond-Scale Molecular-Dynamics Trajectories

(Abstract, Talk)

Haensele E, Whitley D, Banting L, Clark T. DASH: Analysis of Microsecond-Scale Molecular-Dynamics Trajectories (Talk). 28th Molecular Modeling Workshop, Mar 18th, 2014. FAU Erlangen-Nürnberg, Germany. Abstract: http://mmws2014.mgms-ds.de



Natural timescales for conformational changes may last milliseconds to seconds, *e.g.* protein folding. Although current molecular-dynamics simulations (MD) typically cover timescales of 10 to 100 nanoseconds, the computational power has become readily available to run simulations on a microsecond scale. However, such long simulations create the technical problem of how to analyse the increased volume of output within a reasonable time without being forced to reduce the number of considered data points drastically. This is where common clustering methods reach their limits. *DASH* (Dynamic Analysis by Salt and Hudson) (1) provides an alternative solution by using a time series of torsion angles instead of similarity matrices of Cartesian coordinates (clustering) to find representative conformations (states). Time-series analysis is very fast, making *DASH* capable of analysing considerable large datasets. The principles of *DASH* will be explained and *amberDASH*, an interface for the user-friendly application of *DASH* to AMBER trajectories, will be introduced.

The performance of *DASH* and the consistency of its results will be demonstrated using a 5-microsecond MD trajectory of Arg^8-vasopressin as an example.

DASH 1.0 Program for extracting states from molecular-dynamics simulations; distributed under the terms of the GNU General Public License; download *via* www.port.ac.uk/research/cmd/software

AmberDASH *DASH* interface for AMBER trajectories (unpublished); currently provided *via* email (please contact Dr David Whitley david.whitley@port.ac.uk)

(1) D.W. Salt, B.D. Hudson, L. Banting, M.J. Ellis, M.G. Ford, J Med Chem, 2005, 48, 3214-3220

Cyclic Peptide Hormones: Conformation, Dynamics and Pharmacophores of Urotensin and Vasopressin

(Abstract Talk)

Haensele E, Banting L, Whitley D, Read C, Cary P, Clark T, *et al.* Cyclic Peptide Hormones: Conformation, Dynamics and Pharmacophores of Urotensin and Vasopressin (Joint Lecture). Final PeReNE Meeting, Jan 15-16th, 2015. University de Le Havre, France.

Human urotensin II (h-UII), urotensin-related peptide (URP), and Arg^8-vasopressin (AVP) are natural bioactive peptides that exhibit a multitude of physiological functions such as vasoconstriction or water homeostasis. They are G-protein-coupled-receptor ligands and their common structural feature is a 6-residue ring closed by a disulphide bridge. Elucidating the conformational space of the free peptides is important in order to identify candidates for the biologically active conformation and understand the mechanisms of receptor activation and hence for drug design. Cooperatively, we study the structure and dynamics of these peptides using different methods and approaches.

1. Unrestrained, long-timescale (5µs) molecular-dynamics (MD) simulations of h-UII and URP in solution reveal two distinct major populated ring states (*Ω-shape* and *folded*) with well-defined structures but significantly different dynamics. The results agree well with experimental findings but provide extra detail not available from the experiments. Different dynamics of URP and UII indicate that a longer N-terminus may stabilise more structured ring states.
2. Replica Exchange MD simulations were applied to h-UII, URP and AVP in water and non-aqueous solvents, to extend our understanding of the conformational sampling of these peptides. The rate of convergence of the REMD conformer populations provide data regarding the peptide conformational dynamics, while the relative populations of each conformer derived from the converged simulations allow the relative free energies of each state to be estimated.
3. Receptor of Arg^8-vasopressin (called V2R) and that of urotensin II and URP (called UT) were built using the homology modelling technique. Different conformations of vasopressin, derived from the five µs MD simulations performed in Portsmouth, were docked into the binding site of the V2R model and ligand-receptor interactions were analysed. Furthermore, new data concerning the notion of biased ligands encouraged us to extend this study by considering more recent binding and pharmacological data for non-peptide, pseudo peptide and natural peptide ligands of UT. From a new set of non-peptide ligands, various pharmacophores were generated. These pharmacophores were analysed and aligned to URP and h-UII conformations resulting from long molecular dynamic simulations. To complete this work, a docking study was carried out on UT as well as a virtual screening of CERMN and French National chemical libraries.

Within a multi-allosteric view, all conformations presented may be considered as bioactive receptor-ligands and potential candidates for drug-design. Simulations of the vasopressin receptor type 2 suggest alternative binding sites.

Urotensin II and Urotensin-Related-Peptide: How to Decipher NMR-Data for Conformational Equilibria with Molecular-Dynamics Simulation and Modelling

(Abstract, Poster)

Haensele E, Mele N, Miljak M, Read CM, Whitley DC, Banting L, *et al.* Urotensin II and Urotensin-Related Peptide: How to Decipher NMR-Data for Conformational Equilibria with Molecular-Dynamics Simulation and Modelling. 13th German Peptide Symposium (DECHEMA), Mar 20-23, 2017. FAU Erlangen-Nürnberg, Germany.

The flexible peptides urotensin II (UII) and urotensin-related peptide (URP) are natural ligands of the G-protein coupled urotensin receptor, UT. They are inter alia involved in cardiovascular regulation (1). Different "single-conformations" for UII and URP have been suggested to be the reason for the different biological responses observed in some cases (2). However, these peptides cannot be described as a single-conformation in solution. We found that both UII and URP rather exist as a fast equilibrium between two main types of ring conformations, *open* and *folded*. The ratio *open*:*folded* for UII is 72:28, whereas the equilibrium for URP is shifted further towards *open* conformations with a ratio of 86:14. The conformational equilibria were characterised by combining unrestrained and enhanced molecular-dynamics simulations and simulating the NMR spectra based on the suggested equilibrium concentrations. This was achieved by comparing the experimental 1H chemical shifts with DFT-calculated chemical shifts of single conformations and conformational mixtures. The technique has already been tested for Arg^8-vasopressin (3) and is apparently able to decipher NMR data of flexible peptides for conformational equilibria.

(1) Vaudry H et al (2015) International Union of Basic and Clinical Pharmacology. XCII. Urotensin II, urotensin II-related peptide, and their receptor: from structure to function. Pharmacol Rev 67 (1):214-258. (2) Chatenet D et al (2004) Structure–activity relationships and structural conformation of a novel urotensin II-related peptide. Peptides 25 (10):1819-1830. (3) Haensele E et al (2016) Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine–Vasopressin. J Chem Inf Model 56 (9):1798-1807

University of Portsmouth

Urotensin-II and Urotensin Related Peptide

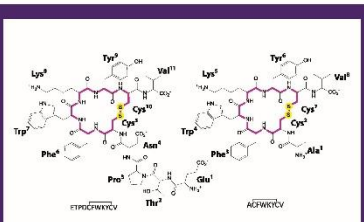
How to Decipher NMR-Data for Conformational Equilibria

with Molecular Dynamics Simulation and Modeling

Haensele E, Read C, Whitley D, Banting L, School of Pharmacy and Biomedical Sciences, and Biological Sciences, University of Portsmouth, UK; Mele N, Miljak M, Essex J, School of Chemistry,University of Southampton, UK; Delépée C, Sopkova J, Lepailleur A, Bureau R, CERMN, Université Normandie, France; Clark T, Computer-Chemie-Centrum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Introduction

Urotensin-II (UII) and urotensin related peptide (URP) are cyclic peptide hormones and natural ligands of the G-protein coupled receptor UT. They are involved in many physiological processes (e.g. cardiovascular regulation) and consequently also diseases (e.g. heart failure).[1] As structure may determine function, their conformations are of special interest. UII and URP have the same ring-sequence but differ in the N-terminal tail. The current descriptions of their conformation in solution range from unstructured [2] to distinct single-conformations [3] and the latter have even been suggested to be responsible for partially different biological functions of UII and URP.



Human Urotensin-II (UII) Urotensin Related Peptide (URP)

The Limits of NMR

NMR techniques become problematic when the time-scale for conformational equilibria in peptides is fast compared with the NMR experiment time-scale. Averaged NMR resonances appear and may be interpreted erroneously as single conformation or unstructured. UII and URP are intrinsically flexible peptides and may thus exhibit fast multiple conformational equilibria in solution.

We present a technique that allows us to decipher fast equilibria from NMR spectra using long-scale molecular dynamics (MD) simulations combined with enhanced sampling methods, NMR calculations and modeling.

Technique

1 NMR Experiments

Assignment of experimental chemical shifts (δ): δ(¹H) performs best to identify the best fitting model. δ(¹³C) can be used for control metrics.

2 MD Simulations

Identification of main conformational types with unrestrained, long (µs-scale) molecular dynamics (MD) simulations: Representative conformations are used as input for NMR calculation and enhanced sampling.

3 Enhanced Sampling

Determination of equilibrium concentrations: Unrestrained MD simulations are complemented with enhanced sampling (e.g. REMD, Metadynamics). For UII and URP, simulations converged to similar ratio of open and folded conformations.

4 DFT/NMR Calculation

Calculation of chemical shifts for each representative ("single-conformation" models) using densitiy functional theory (DFT): optimisation and NMR shielding tensor calculation (same level); δ conversion

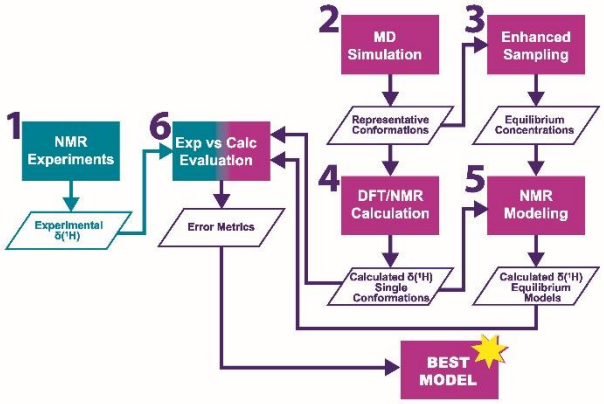
5 NMR Modeling

Modeling of equilibrium spectra via linear combination of „single-conformation" δ(¹H) based on the suggested equilibrium concentrations.

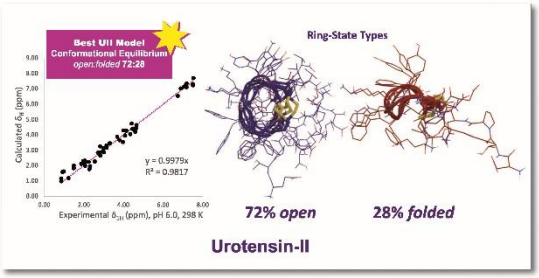
6 Evaluation

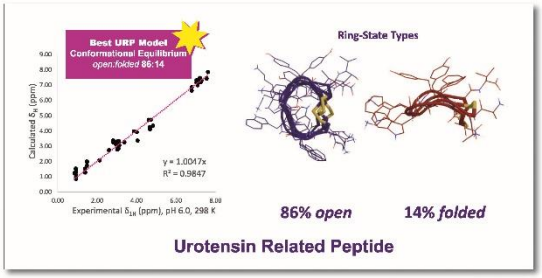
Linear regression of calculated ¹H single-conformation spectra and equilibrium spectra with the experimental spectrum. Analysis of error values (e.g. mean errors, weighted root mean square deviation, etc.). [4]

Model with smallest error values = best approximation of experimental data.



Conformational Equilibria of UII and URP





Summary

The technique presented [4] has proven to be promising for interpreting NMR spectra of flexible peptides:

The conformation of UII or URP in solution cannot be described as single conformation or unstructured. A fast conformational equilibrium of *open* and *folded* ring-type conformations provides a better interpretation of the experimental NMR data, whereby UII and URP comprise the same ring state types. Thus, partially different biological functions of UII and URP cannot be attributed to different single-conformations in solution. However, UII and URP differ in the ratio of *open* and *folded* conformations. UII shows a higher percentage of folded conformations (28%) than URP (14%).

Conformations of the *open* ring type are relatively flat with no highly populated transannular hydrogen bonds; turns are centred at either Lys-Tyr or Cys-Phe. Published single-conformations in aqueous solution belong to this type.

Conformations of the *folded* ring type are characterised by multiple turns comprising residues Phe-Trp-Lys or Trp-Lys-Tyr that are stabilised by highly populated transannular hydrogen bonds. A conformation of this type has been suggested to be the bioactive conformation in the receptor.

The results for UII and URP have been published in January 2017 [5].

Computational details

MD simulations: 55µs (12sim) | Amber10, Amber14CUDA ff99SB | TIP4PEw | trunc.oct | UII: $1Na^{+}$/ URP: $1Cl^{-}$| 300K | 1bar | 8A cutoff | PME | PBC | Shake; 3µs (2sim) | CHARMM c36b2 | TIP3P | cubic | UII: $7Na^{+}$, $6Cl^{-}$/ URP: $4Na^{+}$,$5Cl^{-}$ | NVT | 300K | 13A cutoff | PME | PCB,

REMD simulation: 32µs (6 init) | Amber ff99SB | TIP3P | cubic | UII: Na^{+}/ URP: Cl^{-} | 1bar | 10A cutoff | PME | PBC | Shake;

Clustering: DASH 2.1 | www.port.ac.uk/research/cmd/software

DFT calculation: Gaussian 09 [R3LYP/6-31G(d) opt scrf=pcm][R3LYP/6-31G(d) scrf=pcm | δ(¹H) = -0.9912σ(¹H) + 32.05

Experimental details

NMR experiments: UII (pH6) / URP (pH6) | H2O/D2O | Varian inova 600MHz | TOCSY 2D ^{1}H^{1}H | NOESY ^{1}H^{1}H | gHSQC ^{13}C^{1}H | gHSQC ^{15}N^{1}H

References: [1] For review see e.g.: Vaudry et al (2015) Pharmacol Rev 67:214; Tostivint et al (2014) J Mol Endocrinol 52:T61; [2] a) Rohr et al (2002) J Med Chem 45:1799, b) Brancaccio et al (2015) J Pept Sci 21:392; [3] a) Lescot et al (2007) J Chem Inf Model 10:3731, b) Chatenet et al (2004) Peptides 25:1619; [4] Technique explained and applied to AVP: Haensele et al (2016) J Chem Inf Model 56:1758; [5] Haensele et al (2017) J Chem Inf Model DOI: 10.1021/acs.jcim.6b00706

REFERENCES

(Note: References of Appendices "Reprint Supporting Information" of Papers 1 to 3 are not included)

1. Haensele E, Banting L, Whitley DC, Clark T. Conformation and dynamics of 8-Arg-vasopressin in solution. J Mol Model. 2014;20(11):2485(17).
2. Haensele E, Saleh N, Read CM, Banting L, Whitley DC, Clark T. Can simulations and modeling decipher NMR data for conformational equilibria? Arginine–vasopressin. J Chem Inf Model. 2016;56(9):1798-807.
3. Haensele E, Mele N, Miljak M, Read CM, Whitley DC, Banting L, *et al.* Conformation and dynamics of human urotensin II and urotensin-related peptide in aqueous solution. J Chem Inf Model. 2017;57(2):298-310.
4. Berg JM, Tymoczko JL, Stryer L, editors. Biochemie. [Korrigierter Nachdruck der 6. Auflage 2007]. Heidelberg: Spektrum Akademischer Verlag; 2010.
5. Chemie.de - Lexikon. Peptide [Internet]. Chemie.de Information Service GmbH. © 1997-2017 [cited 2016 Dec]. Available from: http://www.chemie.de/lexikon/Peptid.html.
6. Marahiel MA, Stachelhaus T, Mootz HD. Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem Rev. 1997;97(7):2651-74.
7. Lemes AC, Sala L, Ores JD, Braga ARC, Egea MB, Fernandes KF. A review of the latest advances in encrypted bioactive peptides from protein-rich waste. Int J Mol Sci. 2016;17(6):950(24).
8. Shahidi F, Zhong Y. Bioactive Peptides. J AOAC Int. 2008;91(4):914-31.
9. Loffet A. Peptides as drugs: is there a market? J Pept Sci. 2002;8(1):1-7.
10. Stevenson CL. Advances in peptide pharmaceuticals. Curr Pharm Biotechnol. 2009;10(1):122-137.
11. Vlieghe P, Lisowski V, Martinez J, Khrestchatisky M. Synthetic therapeutic peptides: science and market. Drug Discov Today. 2010;15(1-2):40-56.
12. Krieger D. Brain peptides: what, where, and why? Science. 1983;222(4627):975-85.
13. Sporn MB, Roberts AB, editors. Peptide growth factors and their receptors I. Berlin-Heidelberg: Springer-Verlag; 1990.
14. National Center for Biotechnology Information - NCBI - Medical Subject Headings - MeSH. Cobra Neurotoxin Proteins [Internet]. Rockville Pike: U. S. National Library of Medicine. 2011 [cited 2017 Feb 21]. Available from: https://www.ncbi.nlm.nih.gov/mesh/68003039.
15. Meister A, Anderson ME. Glutathione. Annu Rev Biochem. 1983;52(1):711-60.
16. Garcia MC, Puchalska P, Esteve C, Marina ML. Vegetable foods: a cheap source of proteins and peptides with antihypertensive, antioxidant, and other less occurrence bioactivities. Talanta. 2013;106:328-49.
17. Izadpanah A, Gallo RL. Antimicrobial peptides. J Am Acad Dermatol. 2005;52(3):381-90.
18. Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature. 2013;494(7436):185-94.

19. Gimpl G, Reitz J, Brauer S, Trossen C. Oxytocin receptors: ligand binding, signalling and cholesterol dependence. Prog Brain Res. 2008;170:193-204.

20. Laycock JF. Perspectives on vasopressin [Ebook]. London, UK: Imperial College Press; 2010 [cited 2012 Jul 12]. Available from: http://site.ebrary.com/lib/portsmouth/Doc?id=10422286.

21. Chatenet D, Nguyen TT, Letourneau M, Fournier A. Update on the urotensinergic system: new trends in receptor localization, activation, and drug design. Front Endocrin. 2012;3:174(13).

22. Barberis C, Mouillac B, Durroux T. Structural bases of vasopressin/oxytocin receptor function. J Endocrinol. 1998;156(2):223-9.

23. Vaudry H, Leprince J, Chatenet D, Fournier A, Lambert DG, Le Mevel JC, *et al.* International union of basic and clinical pharmacology. XCII. Urotensin II, urotensin II-related peptide, and their receptor: from structure to function. Pharmacol Rev. 2015;67(1):214-58.

24. Schwyzer R. In search of the 'bio-active conformation' - is it induced by the target cell membrane? J Mol Recognit. 1995;8(1-2):3-8.

25. Mierke DF, Giragossian C. Peptide hormone binding to G-protein-coupled receptors: structural characterization via NMR techniques. Med Res Rev. 2001;21(5):450-71.

26. Weis WI, Kobilka BK. Structural insights into G-protein-coupled receptor activation. Curr Opin Struct Biol. 2008;18(6):734-40.

27. Vanni S, Rothlisberger U. A closer look into G protein coupled receptor activation: X-ray crystallography and long-scale molecular dynamics simulations. Curr Med Chem. 2012;19(8):1135-45.

28. Saleh N, Saladino G, Gervasio FL, Haensele E, Banting L, Whitley DC, *et al.* A three-site mechanism for agonist/antagonist selective binding to vasopressin receptors. Angew Chem Int Ed Engl. 2016;55(28):8008-12.

29. Brule C, Perzo N, Joubert JE, Sainsily X, Leduc R, Castel H, *et al.* Biased signaling regulates the pleiotropic effects of the urotensin II receptor to modulate its cellular behaviors. FASEB J. 2014;28(12):5148-62.

30. Engstrom T, Barth T, Melin P, Vilhardt H. Oxytocin receptor binding and uterotonic activity of carbetocin and its metabolites following enzymatic degradation. Eur J Pharmacol. 1998;355(2-3):203-10.

31. Peters NC, Duvekot JJ. Carbetocin for the prevention of postpartum hemorrhage: a systematic review. Obstet Gynecol Surv. 2009;64(2):129-35.

32. Wood SP, Tickle IJ, Treharne AM, Pitts JE, Mascarenhas Y, Li JY, *et al.* Crystal structure analysis of deamino-oxytocin: conformational flexibility and receptor binding. Science. 1986;232(4750):633-6.

33. Fahrenholz F, Boer R, Crause P, Fritzsch G, Grzonka Z. Interactions of vasopressin agonists and antagonists with membrane receptors. Eur J Pharmacol. 1984;100(1):47-58.

34. Zingg HH. Vasopressin and oxytocin receptors. Baillieres Clin Endocrinol Metab. 1996;10(1):75-96.

35. Chini B, Mouillac B, Balestre MN, Trumpp-Kallmeyer S, Hoflack J, Hibert M, *et al.* Two aromatic residues regulate the response of the human oxytocin receptor to the partial agonist arginine vasopressin. Febs Lett. 1996;397(2-3):201-6.

36. Jard S. Vasopressin receptors. A historical survey. Adv Exp Med Biol. 1998;449:1-13.

37. Strand FL. Neuropeptides: regulators of physiological processes [Ebook]. Cambridge, Mass: The MIT Press; 1999 [cited 2012 Jul 12]. Available from: http://cognet.mit.edu/library/books/view?isbn=0262194074.

38. Chini B, Fanelli F. Molecular basis of ligand binding and receptor activation in the oxytocin and vasopressin receptor family. Exp Physiol. 2000;85:59S-66S.

39. Barlow M. Vasopressin. Emerg Med (Fremantle). 2002;14(3):304-14.

40. Boone M, Deen PM. Physiology and pathophysiology of the vasopressin-regulated renal water reabsorption. Pflugers Arch. 2008;456(6):1005-24.

41. Manning M, Misicka A, Olma A, Bankowski K, Stoev S, Chini B, *et al.* Oxytocin and vasopressin agonists and antagonists as research tools and potential therapeutics. J Neuroendocrinol. 2012;24(4):609-28.

42. Young LJ, Flanagan-Cato LM. Editorial comment: Oxytocin, vasopressin and social behavior. Horm Behav. 2012;61(3):227-9.

43. Zingg HH, Laporte SA. The oxytocin receptor. Trends Endocrinol Metab. 2003;14(5):222-7.

44. Gimpl G, Fahrenholz F. The oxytocin receptor system: structure, function, and regulation. Physiol Rev. 2001;81(2):629-83.

45. Lee HJ, Macbeth AH, Pagani JH, Young WS, 3rd. Oxytocin: the great facilitator of life. Prog Neurobiol. 2009;88(2):127-51.

46. Gimpl G, Postina R, Fahrenholz F, Reinheimer T. Binding domains of the oxytocin receptor for the selective oxytocin receptor antagonist barusiban in comparison to the agonists oxytocin and carbetocin. Eur J Pharmacol. 2005;510(1-2):9-16.

47. Chatenet D, Letourneau M, Nguyen Q, Doan N, Dupuis J, Fournier A. Discovery of new antagonists aimed at discriminating UII and URP-mediated biological activities: insight into UII and URP receptor activation. Br J Pharmacol. 2013;168(4):807-21.

48. Tostivint H, Ocampo Daza D, Bergqvist CA, Quan FB, Bougerol M, Lihrmann I, *et al.* Molecular evolution of GPCRs: somatostatin/urotensin II receptors. J Mol Endocrinol. 2014;52(3):61-86.

49. Ross B, McKendy K, Giaid A. Role of urotensin II in health and disease. Am J Physiol Regul Integr Comp Physiol. 2010;298(5):1156-72.

50. Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. Chem Biol Drug Des. 2013;81(1):136-47.

51. Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. Drug Discov Today. 2015;20(1):122-8.

52. Rezai T, Bock JE, Zhou MV, Kalyanaraman C, Lokey RS, Jacobson MP. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: Successful in silico prediction of the relative permeabilities of cyclic peptides. J Am Chem Soc. 2006;128(43):14073-80.

53. Nobelprize.org. Lists of Nobel Prizes and Laureates [Internet]. Copyright © Nobel Media AB 2017. 2017 [cited 2017 June 16]. Available from: http://www.nobelprize.org/nobel_prizes/lists/all/index.html.

54. Van 't Hoff JH. A suggestion looking to the extension into space of the structural formulas at present used in chemistry. And a note upon the relation between the optical activity and the chemical constitution of organic compounds. Archives neerlandaises des sciences exactes et naturelles. 1874;9:p. 445-54.

55. Kalsi PS. Stereochemistry: conformation and mechanism. New York, London: Wiley; 1990. p. 1

56. Eliel EL. Conformational analysis. New York, London: Interscience; 1965. p. 5.

57. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol. 1963;7(1):95-9.

58. Richardson JS. The anatomy and taxonomy of protein structure. Adv Protein Chem. 1981;34:167-339.

59. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. Ber Dtsch Chem Ges. 1894;27(3):2985-93.

60. Koshland DE. Application of a theory of enzyme specificity to protein synthesis. Proc Natl Acad Sci U S A. 1958;44(2):98-104.

61. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: a plausible model. J Mol Biol. 1965;12:88-118.

62. Koshland DE, Jr., Nemethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry. 1966;5(1):365-85.

63. Changeux JP, Edelstein SJ. Allosteric mechanisms of signal transduction. Science. 2005;308(5727):1424-8.

64. Changeux JP. Allostery and the Monod-Wyman-Changeux model after 50 years. Annu Rev Biophys. 2012;41:103-33.

65. Changeux J-P. The concept of allosteric modulation: an overview. Drug Discov Today: Technol. 2013;10(2):223-8.

66. Brunori M. Allostery turns 50: is the vintage yet attractive? Protein Sci. 2011;20(7):1097-9.

67. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. Nat Chem Biol. 2009;5(11):789-96.

68. Janin J, Sternberg MJ. Protein flexibility, not disorder, is intrinsic to molecular recognition. F1000 Biol Rep. 2013;5:2(7).

69. Viappiani C, Bettati S, Bruno S, Ronda L, Abbruzzetti S, Mozzarelli A, et al. New insights into allosteric mechanisms from trapping unstable protein conformations in silica gels. Proc Natl Acad Sci U S A. 2004;101(40):14414-9.

70. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins. 2004;57(3):433-43.

71. Fairlie DP, Abbenante G, March DR. Macrocyclic peptidomimetics - forcing peptides into bioactive conformations. Curr Med Chem. 1995;2(2):654-86.

72. Du Vigneaud V, Ressler C, Trippett S. The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. J Biol Chem. 1953;205(2):949-57.

73. Jarvis D, Du Vigneaud V. Crystalline deamino-oxytocin. Science. 1964;143(3606):545-8.

74. Ferrier BM, Jarvis D, du Vigneaud V. Deamino-oxytocin. Its isolation by partition chromatography on Sephadex and crystallization from water, and its biological activities. J Biol Chem. 1965;240(11):4264-6.

75. Low BW, Chen CC. Deamino-oxytocin and 1-gamma-mercaptobutyric acid-oxytocin: x-ray crystallographic data. Science. 1966;151(3717):1552-3.

76. Husain J, Blundell TL, Cooper S, Pitts JE, Tickle IJ, Wood SP, et al. The conformation of deamino-oxytocin: X-ray analysis of the 'dry' and 'wet' forms. Philos Trans R Soc Lond B Biol Sci. 1990;327(1243):625-54.

77. Rose JP, Wu CK, Hsiao CD, Breslow E, Wang BC. Crystal structure of the neurophysin-oxytocin complex. Nat Struct Biol. 1996;3(2):163-9.

78. Urry DW, Walter R. Proposed conformation of oxytocin in solution. Proc Natl Acad Sci U S A. 1971;68(5):956-8.

79. Koehbach J, O'Brien M, Muttenthaler M, Miazzo M, Akcan M, Elliott AG, et al. Oxytocic plant cyclotides as templates for peptide G protein-coupled receptor ligand design. Proc Natl Acad Sci U S A. 2013;110(52):21183-8.

80. Ohno A, Kawasaki N, Fukuhara K, Okuda H, Yamaguchi T. Complete NMR analysis of oxytocin in phosphate buffer. Magn Reson Chem. 2010;48(2):168-72.

81. Liwo A, Tempczyk A, Oldziej S, Shenderovich MD, Hruby VJ, Talluri S, et al. Exploration of the conformational space of oxytocin and arginine-vasopressin using the electrostatically driven Monte Carlo and molecular dynamics methods. Biopolymers. 1996;38(2):157-75.

82. Meraldi JP, Hruby VJ, Brewster AI. Relative conformational rigidity in oxytocin and (1-penicillamine)-oxytocin: a proposal for the relationship of conformational flexibility to peptide hormone agonism and antagonism. Proc Natl Acad Sci U S A. 1977;74(4):1373-7.

83. Glickson JD, Rowan R, Pitner TP, Dadok J, Bothner-By AA, Walter R. 1H nuclear magnetic resonance double resonance study of oxytocin in aqueous solution. Biochemistry. 1976;15(5):1111-9.

84. Brewster AI, Hruby VJ. 300-MHz nuclear magnetic resonance study of oxytocin aqueous solution: conformational implications. Proc Natl Acad Sci U S A. 1973;70(12):3806-9.

85. Budesinsky M, Ragnarsson U, Lankiewicz L, Grehn L, Slaninova J, Hlavacek J. Synthesis and utilization of 13C and 15N backbone-labeled proline: NMR study of synthesized oxytocin with backbone-labeled C-terminal tripeptide amide. Amino acids. 2005;29(2):151-60.

86. Bhaskaran R, Chuang LC, Yu C. Conformational properties of oxytocin in dimethyl-sulfoxide solution - NMR and restrained molecular-dynamics studies. Biopolymers. 1992;32(12):1599-608.

87. Kato T, Endo S, Fujiwara T, Nagayama K. Oxytocin solution structure changes upon protonation of the N-terminus in dimethyl sulfoxide. J Biomol NMR. 1993;3(6):653-73.

88. Brewster AI, Hruby VJ, Glasel JA, Tonelli AE. Proposed conformations of oxytocin and selected analogs in dimethyl sulfoxide as deduced from proton magnetic resonance studies. Biochemistry. 1973;12(26):5294-304.

89. Brewster AI, Hruby VJ, Spatola AF, Bovey FA. Carbon-13 nuclear magnetic resonance spectroscopy of oxytocin, related oligopeptides, and selected analogs. Biochemistry. 1973;12(8):1643-9.

90. Urry DW, Ohnishi M, Walter R. Secondary structure of the cyclic moiety of the peptide hormone oxytocin and its deamino analog. Proc Natl Acad Sci U S A. 1970;66(1):111-6.

91. Johnson LF, Schwartz IL, Walter R. Oxytocin and neurohypophyseal peptides: spectral assignment and conformational analysis by 220 MHz nuclear magnetic resonance. Proc Natl Acad Sci U S A. 1969;64(4):1269-75.

92. Lippens G, Hallenga K, Vanbelle D, Wodak SJ, Nirmala NR, Hill P, et al. Transfer nuclear overhauser effect study of the conformation of oxytocin bound to bovine neurophysin-I. Biochemistry. 1993;32(36):9423-34.

93. Ward DJ, Chen Y, Platt E, Robson B. Development and testing of protocols for computer-aided design of peptide drugs, using oxytocin. J Theor Biol. 1991;148(2):193-227.

94. Nikiforovich GV, Leonova VI, Galaktionov SG, Chipens GI. Theoretical conformational analysis of oxytocin molecule. Int J Pept Protein Res. 1979;13(4):363-73.

95. Kotelchuck D, Scheraga HA, Walter R. Conformational energy studies of oxytocin and its cyclic moiety. Proc Natl Acad Sci U S A. 1972;69(12):3629-33.

96. Liwo A, Tempczyk A, Grzonka Z. Molecular mechanics calculations on deaminooxytocin and on deamino-arginine-vasopressin and its analogs. J Comput Aided Mol Des. 1989;2(4):281-309.

97. Syed Ibrahim B, Pattabhi V. Trypsin inhibition by a peptide hormone: crystal structure of trypsin-vasopressin complex. J Mol Biol. 2005;348(5):1191-8.

98. Wu CK, Hu B, Rose JP, Liu ZJ, Nguyen TL, Zheng C, et al. Structures of an unliganded neurophysin and its vasopressin complex: implications for binding and allosteric mechanisms. Protein Sci. 2001;10(9):1869-80.

99. Schmidt JM, Ohlenschlager O, Ruterjans H, Grzonka Z, Kojro E, Pavo I, et al. Conformation of [8-arginine]vasopressin and V1 antagonists in dimethyl sulfoxide solution derived from two-dimensional NMR spectroscopy and molecular dynamics simulation. Eur J Biochem. 1991;201(2):355-71.

100. Rodziewicz-Motowidlo S, Sikorska E, Oleszczuk M, Czaplewski C. Conformational studies of vasopressin and mesotocin using NMR spectroscopy and molecular modelling methods. Part II: Studies in the SDS micelle. J Pept Sci. 2008;14(1):85-96.

101. Gryczynski I, Szmacinski H, Laczko G, Wiczk W, Johnson ML, Kusba J, et al. Conformational differences of oxytocin and vasopressin as observed by fluorescence anisotropy decays and transient effects in collisional quenching of tyrosine fluorescence. J Fluoresc. 1991;1(3):163-76.

102. Yedvabny E, Nerenberg PS, So C, Head-Gordon T. Disordered structural ensembles of vasopressin and oxytocin and their mutants. J Phys Chem B. 2014;119(3):896-905.

103. Sikorska E, Rodziewicz-Motowidlo S. Conformational studies of vasopressin and mesotocin using NMR spectroscopy and molecular modelling methods. Part I: Studies in water. J Pept Sci. 2008;14(1):76-84.

104. Walter R, Ballardin A, Schwartz IL, Gibbons WA, Wyssbrod HR. Conformational studies on arginine vasopressin and arginine vasotocin by proton magnetic resonance spectroscopy. Proc Natl Acad Sci U S A. 1974;71(11):4528-32.

105. Lubecka EA, Sikorska E, Sobolewski D, Prahl A, Slaninova J, Ciarkowski J. Arginine-, D-arginine-vasopressin, and their inverso analogues in micellar and liposomic models of cell membrane: CD, NMR, and molecular dynamics studies. Eur Biophys J. 2015;44(8):727-43.

106. Hruby VJ, Deb KK, Fox J, Bjarnason J, Tu AT. Conformational studies of peptide hormones using laser Raman and circular dichroism spectroscopy. A comparative study of oxytocin agonists and antagonists. J Biol Chem. 1978;253(17):6060-7.

107. Tu AT, Bjarnason JB, Hruby VJ. Conformation of oxytocin studied by laser Raman spectroscopy. Biochim Biophys Acta. 1978;533(2):530-3.

108. Podstawka E, Sikorska E, Proniewicz LM, Lammek B. Raman and surface-enhanced Raman spectroscopy investigation of vasopressin analogues containing 1-aminocyclohexane-1-carboxylic acid residue. Biopolymers. 2006;83(2):193-203.

109. Carotenuto A, Grieco P, Campiglia P, Novellino E, Rovero P. Unraveling the active conformation of urotensin II. J Med Chem. 2004;47(7):1652-61.

110. Pazderková M, Bednárová L, Dlouhá H, Flegel M, Lebl M, Hlaváček J, et al. Electronic and vibrational optical activity of several peptides related to neurohypophyseal hormones: Disulfide group conformation. Biopolymers. 2012;97(11):923-32.

111. Urry DW, Quadrifoglio F, Walter R, Schwartz IL. Conformational studies on neurohypophyseal hormones: the disulfide bridge of oxytocin. Proc Natl Acad Sci U S A. 1968;60(3):967-74.

112. Maxfield FR, Scheraga HA. A Raman spectroscopic investigation of the disulfide conformation in oxytocin and lysine vasopressin. Biochemistry. 1977;16(20):4443-9.

113. Tu AT, Lee J, Deb KK, Hruby VJ. Laser Raman spectroscopy and circular dichroism studies of the peptide hormones mesotocin, vasotocin, lysine vasopressin, and arginine vasopressin. Conformational analysis. J Biol Chem. 1979;254(9):3272-8.

114. Cowburn D, Live DH, Fischman AJ, Agosta WC. Side-chain conformation of oxytocin and vasopressin studies by NMR observation of isotopic isomers. J Am Chem Soc. 1983;105(25):7435-42.

115. Deslauriers R, Smith ICP, Walter R. Conformation flexibility of the neurohypophyseal hormones oxytocin and lysine-vasopressin. Carbon-13 spin-lattice relaxation study of backbone and side chains. J Am Chem Soc. 1974;96(7):2289-91.

116. FriČ I, Flegel M, Zaoral M, KodÍČEk M. Circular-dichroic spectra of vasopressin analogues and their cyclic fragments. Eur J Biochem. 1975;56(2):493-502.

117. Szmacinski H, Wiczk W, Fishman MN, Eis PS, Lakowicz JR, Johnson ML. Distance distributions from the tyrosyl to disulfide residues in the oxytocin and [Arg8]-vasopressin measured using frequency-domain fluorescence resonance energy transfer. Eur Biophys J. 1996;24(3):185-93.

118. Larive CK, Guerra L, Rabenstein DL. Cis trans conformational equilibrium across the cysteine(6) proline peptide-bond of oxytocin, arginine vasopressin, and lysine vasopressin. J Am Chem Soc. 1992;114(19):7331-7.

119. Larive CK, Rabenstein DL. Dynamics of cis trans isomerization of the cysteine(6)-proline peptide-bonds of oxytocin and arginine vasopressin in aqueous and methanol solution. J Am Chem Soc. 1993;115(7):2833-6.

120. Glasel JA, Hruby VJ, McKelvy JF, Spatola AF. Deuteron magnetic resonance studies on the microdynamical behavior of partially deuterated oxytocin with neurophysin. J Mol Biol. 1973;79(3):555-75.

121. Mori M, Sugo T, Abe M, Shimomura Y, Kurihara M, Kitada C, et al. Urotensin II is the endogenous ligand of a G-protein-coupled orphan receptor, SENR (GPR14). Biochem Biophys Res Commun. 1999;265(1):123-9.

122. Nothacker HP, Wang Z, McNeill AM, Saito Y, Merten S, O'Dowd B, et al. Identification of the natural ligand of an orphan G-protein-coupled receptor involved in the regulation of vasoconstriction. Nat Cell Biol. 1999;1(6):383-5.

123. Liu Q, Pong SS, Zeng Z, Zhang Q, Howard AD, Williams DL, Jr., et al. Identification of urotensin II as the endogenous ligand for the orphan G-protein-coupled receptor GPR14. Biochem Biophys Res Commun. 1999;266(1):174-8.

124. Sugo T, Murakami Y, Shimomura Y, Harada M, Abe M, Ishibashi Y, et al. Identification of urotensin II-related peptide as the urotensin II-immunoreactive molecule in the rat brain. Biochem Biophys Res Commun. 2003;310(3):860-8.

125. Chatenet D, Dubessy C, Leprince J, Boularan C, Carlier L, Ségalas-Milazzo I, et al. Structure–activity relationships and structural conformation of a novel urotensin II-related peptide. Peptides. 2004;25(10):1819-30.

126. Lescot E, Sopkova-de Oliveira Santos J, Dubessy C, Oulyadi H, Lesnard A, Vaudry H, et al. Definition of new pharmacophores for nonpeptide antagonists of human urotensin-II. Comparison with the 3D-structure of human urotensin-II and URP. J Chem Inf Model. 2007;47(2):602-12.

127. Brancaccio D, Merlino F, Limatola A, Yousif AM, Gomez-Monterrey I, Campiglia P, et al. An investigation into the origin of the biased agonism associated with the urotensin II receptor activation. J Pept Sci. 2015;21(5):392-9.

128. Flohr S, Kurz M, Kostenis E, Brkovich A, Fournier A, Klabunde T. Identification of nonpeptidic urotensin II receptor antagonists by virtual screening based on a pharmacophore model derived from structure−activity relationships and nuclear magnetic resonance studies on urotensin II. J Med Chem. 2002;45(9):1799-805.

129. Grieco P, Carotenuto A, Patacchini R, Maggi CA, Novellino E, Rovero P. Design, synthesis, conformational analysis, and biological studies of urotensin-II lactam analogues. Bioorg Med Chem. 2002;10(12):3731-9.

130. Manning M, Stoev S, Chini B, Durroux T, Mouillac B, Guillon G. Peptide and non-peptide agonists and antagonists for the vasopressin and oxytocin V1a, V1b, V2 and OT receptors: research tools and potential therapeutic agents. Prog Brain Res. 2008;170:473-512.

131. Wisniewski K, Alagarsamy S, Galyean R, Tariga H, Thompson D, Ly B, et al. New, potent, and selective peptidic oxytocin receptor agonists. J Med Chem. 2014;57(12):5306-17.

132. Bryant RG. The NMR timescale. J Chem Edu. 1983;60(11):933.

133. Kessler H, Griesinger C, Lautz J, Mueller A, van Gunsteren WF, Berendsen HJC. Conformational dynamics detected by nuclear magnetic resonance NOE values and J coupling constants. J Am Chem Soc. 1988;110(11):3393-6.

134. The Nobel Prize in Chemistry 2013. [Internet]. Stockholm, Sweden: The Royal Swedish Academy of Sciences. 2013 [cited 2017 Feb 10]. Available from: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/press.html.

135. Qiu L, Pabit SA, Roitberg AE, Hagen SJ. Smaller and faster: the 20-residue Trp-cage protein folds in 4 μs. J Am Chem Soc. 2002;124(44):12952-3.

136. Palmer AG. NMR characterization of the dynamics of biomacromolecules. Chem Rev. 2004;104(8):3623-40.

137. Naganathan AN, Muñoz V. Scaling of folding times with protein size. J Am Chem Soc. 2005;127(2):480-1.

138. Goldberg ME, Semisotnov GV, Friguet B, Kuwajima K, Ptitsyn OB, Sugai S. An early immunoreactive folding intermediate of the tryptophan synthease beta 2 subunit is a 'molten globule'. Febs Lett. 1990;263(1):51-6.

139. Kim PS, Baldwin RL. Intermediates in the folding reactions of small proteins. Annu Rev Biochem. 1990;59:631-60.

140. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins Struct Funct Bioinf. 2006;65(3):712-25.

141. Salt DW, Hudson BD, Banting L, Ellis MJ, Ford MG. DASH: A novel analysis method for molecular dynamics simulation data. Analysis of ligands of PPAR-Gamma. J Med Chem. 2005;48(9):3214-20.

142. Case DA, Darden TA, Cheatham II TE, Simmerling CL, Wang J, Duke RE, et al. AMBER Tools 1.0 Structure and dynamics analysis of trajectories. Version 1.5 ed. San Francisco University of California; 2008.

143. Case DA, Darden TA, Cheatham II TE, Simmerling CL, Wang J, Duke RE, et al. AMBER Tools 12. Structure and dynamics analysis of trajectories. Version 12 ed. University of California, San Francisco 2012.

144. Rapaport DC. The art of molecular dynamics simulation. 2nd ed. ed. Cambridge: Cambridge University Press; 2004.

145. Hentschke R. Molekulares Modellieren mit Kraftfeldern: Einführung in die Theorie und Praxis der Computersimulation molekularer Systeme [PDF]: Universität Wuppertal; 2004 [cited 2017 Jan 17]. Available from: www.materials.uni-wuppertal.de/skripten/skript_modellierung/molmodell.pdf, https://books.google.com.mx/books?id=fe3AtgAACAAJ.

146. Frenkel D, Smit B. Understanding molecular simulation: from algorithms to applications: Elsevier Science; 2001.

147. Haile JM. Molecular dynamics simulation: elementary methods: Wiley; 1997.

148. Newton I. Philosophiae naturalis principia mathematica. London: Tomus Primus; 1726.

149. Cino EA, Choy WY, Karttunen M. Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations. J Chem Theory Comput. 2012;8(8):2725-40.

150. Kästner J. Umbrella sampling. WIREs Comput Mol Sci. 2011;1(6):932-42.

151. Barducci A, Bonomi M, Parrinello M. Metadynamics. WIREs Comput Mol Sci. 2011;1(5):826-43.

152. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett. 1999;314(1-2):141-51.

153. Liu P, Kim B, Friesner RA, Berne BJ. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci U S A. 2005;102(39):13749-54.

154. Iida S, Nakamura H, Higo J. Enhanced conformational sampling to visualize a free-energy landscape of protein complex formation. Biochem J. 2016;473(12):1651-62.

155. Abrams C, Bussi G. Enhanced sampling in molecular dynamics using metadynamics, replica-Exchange, and temperature-acceleration. Entropy. 2014;16(1):163.

156. Sutto L, Marsili S, Gervasio FL. New advances in metadynamics. WIREs Comput Mol Sci. 2012;2(5):771-9.

157. Liwo A, Czaplewski C, Oldziej S, Scheraga HA. Computational techniques for efficient conformational sampling of proteins. Curr Opin Struct Biol. 2008;18(2):134-9.

158. Van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, et al. Biomolecular modeling: goals, problems, perspectives. Angew Chem Int Ed Engl. 2006;45(25):4064-92.

159. Okamoto Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. J Mol Graph Model. 2004;22(5):425-39.

160. Laio A, Parrinello M. Escaping free-energy minima. Proc Natl Acad Sci U S A. 2002;99(20):12562-6.

161. Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. J Comput Phys. 1977;23(2):187-99.

162. Momany FA, Mcguire RF, Burgess AW, Scheraga HA. Energy parameters in polypeptides .7. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen-bond interactions, and intrinsic torsional potentials for naturally occurring amino-acids. J Phys Chem. 1975;79(22):2361-81.

163. Allinger N. Calculation of molecular structure and energy by force-field methods. Adv Phys Org Chem. 1976;13:1-82.

164. Warshel A, Lifson S. Consistent force field calculations. II. Crystal structures, sublimation energies, molecular and lattice vibrations, molecular conformations, and enthalpies of alkanes. J Chem Phys. 1970;53(2):582-94.

165. Lifson S, Warshel A. Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. J Chem Phys. 1968;49(11):5116-29.

166. Jorgensen WL. Foundations of biomolecular modeling. Cell. 2013;155(6):1199-202.

167. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature. 1977;267:585-90.

168. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, et al. The AMBER biomolecular simulation programs. J Comput Chem. 2005;26(16):1668-88.

169. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc. 1984;106(3):765-84.

170. Weiner PK, Kollman PA. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. J Comput Chem. 1981;2(3):287-303.

171. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30(10):1545-614.

172. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem. 1983;4(2):187-217.

173. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B. 2001;105(28):6474-87.

174. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc. 1996;118(45):11225-36.

175. Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, et al. The GROMOS software for biomolecular simulation: GROMOS05. J Comput Chem. 2005;26(16):1719-51.

176. Weiner SJ, Kollman PA, Nguyen DT, Case DA. An all atom force field for simulations of proteins and nucleic acids. J Comput Chem. 1986;7(2):230-52.

177. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, et al. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J Am Chem Soc. 1995;117(19):5179-97.

178. Best RB, Buchete N-V, Hummer G. Are current molecular dynamics force fields too helical? Biophys J.95(1):L07-L9.

179. Salomon-Ferrer R, Case DA, Walker RC. An overview of the AMBER biomolecular simulation package. WIREs Comput Mol Sci. 2013;3(2):198-210.

180. Lee EH, Hsin J, Sotomayor M, Comellas G, Schulten K. Discovery through the computational microscope. Structure. 2009;17(10):1295-306.

181. Johnston JM, Filizola M. Showcasing modern molecular dynamics simulations of membrane proteins through G protein-coupled receptors. Curr Opin Struct Biol. 2011;21(4):552-8.

182. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol. 2009;19(2):120-7.

183. Perez A, Luque FJ, Orozco M. Dynamics of B-DNA on the microsecond timescale. J Am Chem Soc. 2007;129(47):14739-45.

184. Vasquez V, Sotomayor M, Cordero-Morales J, Schulten K, Perozo E. A structural mechanism for MscS gating in lipid bilayers. Science. 2008;321(5893):1210-4.

185. Jogini V, Roux B. Dynamics of the Kv1.2 voltage-gated K+ channel in a membrane environment. Biophys J. 2007;93(9):3070-82.

186. Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. BMC Biology. 2011;9(1):71.

187. Kerrigan JE. Molecular dynamics simulations in drug design. In: Kortagere S, editor. In silico models for drug discovery. Totowa, NJ: Humana Press; 2013. p. 95-113.

188. Born M, Oppenheimer R. Zur Quantentheorie der Molekeln. Ann Phys. 1927;389(20):457-84.

189. Lipparini F, Barone V. Polarizable force fields and polarizable continuum model: a fluctuating charges/PCM approach. 1. Theory and implementation. J Chem Theor Comput. 2011;7(11):3711-24.

190. Anisimov VM, Vorobyov IV, Roux B, MacKerell AD. Polarizable empirical force field for the primary and secondary alcohol series based on the classical drude model. J Chem Theor Comput. 2007;3(6):1927-46.

191. Schnieders MJ, Baker NA, Ren PY, Ponder JW. Polarizable atomic multipole solutes in a Poisson-Boltzmann continuum. J Chem Phys. 2007;126(12):124114(21).

192. Ponder JW, Case DA. Force fields for protein simulations. Adv Protein Chem. 2003;66:27-85.

193. Bernal JD, Fowler RH. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. J Chem Phys. 1933;1(8):515-48.

194. Stillinger FH, Rahman A. Improved simulation of liquid water by molecular-dynamics. J Chem Phys. 1974;60(4):1545-57.

195. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. Intermolecular Forces. Dordrecht: Reidel; 1981.

196. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983;79(2):926-35.

197. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J Chem Phys. 2004;120(20):9665-78.

198. Horn HW, Swope WC, Pitera JW. Characterization of the TIP4P-Ew water model: Vapor pressure and boiling point. J Chem Phys. 2005;123(19):194504.

199. Cerutti DS, Rice JE, Swope WC, Case DA. Derivation of fixed partial charges for amino acids accommodating a specific water model and implicit polarization. J Phys Chem B. 2013;117(8):2328-38.

200. Cerutti DS, Freddolino PL, Duke RE, Case DA. Simulations of a protein crystal with a high resolution X-ray structure: evaluation of force fields and water models. J Phys Chem B. 2010;114(40):12811-24.

201. Guillot B. A reappraisal of what we have learnt during three decades of computer simulations on water. J Mol Liq. 2002;101(1):219-60.

202. Finney JL. The water molecule and its interactions: the interaction between theory, modelling, and experiment. J Mol Liq. 2001;90(1):303-12.

203. Darden T, York D, Pedersen L. Particle Mesh Ewald - an N.log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98(12):10089-92.

204. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys. 1995;103(19):8577-93.

205. Ewald PP. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Ann Phys. 1921;369(3):253-87.

206. Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, et al. AMBER 14. A set of molecular mechanical force fields for the simulation of biomolecules and a package of molecular simulation programs. 14 ed. San Francisco University of California; 2014.

207. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc. 1990;112(16):6127-9.

208. Bashford D, Case DA. Generalized born models of macromolecular solvation effects. Annu Rev Phys Chem. 2000;51:129-52.

209. Sharp KA, Honig B. Electrostatic interactions in macromolecules - theory and applications. Annu Rev Biophys Bio. 1990;19:301-32.

210. Tomasi J, Persico M. Molecular-interactions in solution - an overview of methods based on continuous distributions of the solvent. Chem Rev. 1994;94(7):2027-94.

211. Cramer CJ, Truhlar DG. Implicit solvation models: equilibria, structure, spectra, and dynamics. Chem Rev. 1999;99(8):2161-200.

212. Zhou R. Free energy landscape of protein folding in water: explicit vs. implicit solvent. Proteins. 2003;53(2):148-61.

213. Tironi IG, Sperb R, Smith PE, Gunsteren WFv. A generalized reaction field method for molecular dynamics simulations. J Chem Phys. 1995;102(13):5451-9.

214. Barker JA, Watts RO. Monte Carlo studies of the dielectric properties of water-like models. Mol Phys. 1973;26(3):789-92.

215. Berendsen HJC, Postma JPM, Van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath J Chem Phys. 1984;81(8):3684-90.

216. Schaeffer RD, Fersht A, Daggett V. Combining experiment and simulation in protein folding: closing the gap for small model systems. Curr Opin Struct Biol. 2008;18(1):4-9.

217. Agarwal PK, Hampton S, Poznanovic J, Ramanthan A, Alam SR, Crozier PS. Performance modeling of microsecond scale biological molecular dynamics simulations on heterogeneous architectures. Concurr Comp-Pract E. 2013;25(10):1356-75.

218. Freddolino PL, Liu F, Gruebele M, Schulten K. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. Biophys J. 2008;94(10):L75-7.

219. Case DA, Darden TA, Cheatham, Simmerling CL, Wang J, Duke RE, et al. AMBER 10. A set of molecular mechanical force fields for the simulation of biomolecules and a package of molecular simulation programs. Version 10 ed. San Francisco: University of California; 2008.

220. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, DeBolt S, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun. 1995;91(1):1-41.

221. Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. J Chem Theor Comput. 2013;9(9):3878-88.

222. Goetz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theor Comput. 2012;8(5):1542-55.

223. Le Grand S, Goetz AW, Walker CR. SPFP: Speed without compromise - a mixed precision model for GPU accelerated molecular dynamics simulations. Comp Phys Comm. 2013;184:374-80.

224. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of N-alkanes. J Comput Phys. 1977;23(3):327-41.

225. Shao J, Tanner SW, Thompson N, Cheatham TE. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. J Chem Theor Comput. 2007;3(6):2312-34.

226. DASH 2.11. [Internet]. University of Portsmouth. 2015 [cited 2016 Mar 17]. Available from: www.port.ac.uk/research/cmd/software.

227. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22(12):2577-637.

228. Lewis PN, Momany FA, Scheraga HA. Chain reversals in proteins. Biochim Biophys Acta. 1973;303(2):211-29.

229. Ramachandran GN, Venkatachalam CM. Stereochemical criteria for polypeptides and proteins. IV. Standard dimensions for the cis-peptide unit and conformation of cis-polypeptides. Biopolymers. 1968;6(9):1255-62.

230. Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas. 1960;20:141-51.

231. Jolliffe I. Principal component analysis. 2nd ed. New York: Springer; 2002, 1986.

232. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and intelligent laboratory systems. 1987;2(1-3):37-52.

233. Abdi H, Williams LJ. Principal component analysis. WIREs Comput Stat. 2010;2(4):433-59.

234. SAR-Caddle. [Internet]. Kempston, UK: Cepos InSilico Ltd. 2013 [cited 2016 May 3]. Available from: http://www.ceposinsilico.de/products/sar-caddle.htm.

235. Lodewyk MW, Siebert MR, Tantillo DJ. Computational prediction of 1H and 13C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. Chem Rev. 2012;112(3):1839-62.

236. Grimme S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. J Comput Chem. 2004;25(12):1463-73.

237. Becke AD. Density-functional thermochemistry. 3. The role of exact exchange. J Chem Phys. 1993;98(7):5648-52.

238. Lee CT, Yang WT, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. Phys Rev B. 1988;37(2):785-9.

239. Koch W, Holthausen MC. A chemists´s guide to density functional theory. 2nd ed. Weinheim: Wiley-VCH; 2001.

240. Ditchfield R, Hehre WJ, Pople JA. Self-consistent molecular-orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. J Chem Phys. 1971;54(2):724-8.

241. Ramsey NF. The internal diamagnetic field correction in measurements of the proton magnetic moment. Phys Rev. 1950;77(4):567(1).

242. Ramsey NF. Magnetic shielding of nuclei in molecules. Phys Rev. 1950;78(6):699-703.

243. Kutzelnigg W, Fleischer U, van Wüllen C. Shielding calculations: IGLO method. eMagRes: John Wiley & Sons, Ltd; 2007.

244. Wolinski K, Hinton JF, Pulay P. Efficient implementation of the gauge-independent atomic orbital method for NMR chemical-shift calculations. J Am Chem Soc. 1990;112(23):8251-60.

245. Ditchfield R. Molecular-orbital theory of magnetic shielding and magnetic susceptibility. J Chem Phys. 1972;56(11):5688-91.

246. Facelli JC. Chemical shift tensors: Theory and application to molecular structural problems. Prog Nucl Mag Res Sp. 2011;58(3-4):176-201.

247. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 09, Revision C.01. Revision C.01 ed. Wallingford CT: Gaussian, Inc.; 2010.

248. Toukach FV, Ananikov VP. Recent advances in computational predictions of NMR parameters for the structure elucidation of carbohydrates: methods and limitations. Chem Soc Rev. 2013;42(21):8376-415.

249. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ. A comparison of models for calculating nuclear magnetic resonance shielding tensors. J Chem Phys. 1996;104(14):5497-509.

250. Parr RG, Yang W. Density-functional theory of atoms and molecules. New York: Oxford University Press; 1989. p. 47-70.

251. NMR magnetic shielding and molecular structure. Tossell JA, editor. Dordrecht: Springer Science+Business Media; 1993. 577 p.

252. Calculation of NMR and EPR parameters. Kaupp M, Bühl M, Malkin VG, editors. Weinheim: Wiley-VCH; 2004. 593 p.

253. Diehl P, Fluck E, Kosfeld R. NMR basic principles and progress/ NMR Grundlagen und Fortschritte: Springer Science & Business Media; 2012.

254. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. Chem Rev. 2005;105(8):2999-3093.

255. Linear regression analysis. Seber GA, Lee AJ, editors: John Wiley & Sons; 2012.

256. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis: John Wiley & Sons; 2015.

257. Regression methods. Lesson 1: simple linear regression [Internet]. Pennsylvania: The Pennsylvania State University. 2017 [cited 2017 Mar-09]. Available from: https://onlinecourses.science.psu.edu/stat501/node/250.

258. Smith SG, Goodman JM. Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: The DP4 probability. J Am Chem Soc. 2010;132(37):12946-59.

259. Du Vigneaud V, Gish DT, Katsoyannis PG. A synthetic preparation possessing biological properties associated with arginine vasopressin. J Am Chem Soc. 1954;76(18):4751-2.

260. Barberis C, Morin D, Durroux T, Mouillac B, Guillon G, Seyer R, et al. Molecular pharmacology of AVP and OT receptors and therapeutic potential. Drug News Perspect. 1999;12(5):279-92.

261. Fujiwara Y, Tanoue A, Tsujimoto G, Koshimizu TA. The roles of V1a vasopressin receptors in blood pressure homeostasis: a review of studies on V1a receptor knockout mice. Clin Exp Nephrol. 2012;16(1):30-4.

262. Pittman QJ, Bagdan B. Vasopressin involvement in central control of blood pressure. Prog Brain Res. 1992;91:69-74.

263. Pittman QJ, Wilkinson MF. Central arginine vasopressin and endogenous antipyresis. Can J Physiol Pharmacol. 1992;70(5):786-90.

264. Mogil JS, Sorge RE, LaCroix-Fralish ML, Smith SB, Fortin A, Sotocinal SG, et al. Pain sensitivity and vasopressin analgesia are mediated by a gene-sex-environment interaction. Nat Neurosci. 2011;14(12):1569-73.

265. Dunning BE, Moltz JH, Fawcett CP. Modulation of insulin and glucagon secretion from the perfused rat pancreas by the neurohypophysial hormones and by desamino-D-arginine vasopressin (DDAVP). Peptides. 1984;5(5):871-5.

266. Insel TR, O'Brien DJ, Leckman JF. Oxytocin, vasopressin, and autism: is there a connection? Biol Psychiatry. 1999;45(2):145-57.

267. Yamaguchi Y, Suzuki T, Mizoro Y, Kori H, Okada K, Chen Y, et al. Mice genetically deficient in vasopressin V1a and V1b receptors are resistant to jet lag. Science. 2013;342(6154):85-90.

268. Manning M, Chan WY, Sawyer WH. Design of cyclic and linear peptide antagonists of vasopressin and oxytocin: current status and future directions. Regulatory peptides. 1993;45(1-2):279-83.

269. Lehrich RW, Greenberg A. Hyponatremia and the use of vasopressin receptor antagonists in critically ill patients. J Intensive Care Med. 2012;27(4):207-18.

270. Verbalis JG. Disorders of body water homeostasis. Best Pract Res Clin Endocrinol Metab. 2003;17(4):471-503.

271. Verbalis JG. AVP receptor antagonists as aquaretics: review and assessment of clinical data. Cleve Clin J Med. 2006;73 Suppl 3:S24-33.

272. Czaplewski C, Kazmierkiewicz R, Ciarkowski J. Molecular modeling of the human vasopressin V2 receptor/agonist complex. J Comput Aided Mol Des. 1998;12(3):275-87.

273. Mouillac B, Chini B, Balestre MN, Elands J, Trumpp-Kallmeyer S, Hoflack J, et al. The binding site of neuropeptide vasopressin V1a receptor. Evidence for a major localization within transmembrane regions. J Biol Chem. 1995;270(43):25771-7.

274. Cui Q, Karplus M. Allostery and cooperativity revisited. Protein Sci. 2008;17(8):1295-307.

275. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983;79(2):926-35.

276. DASH 1.0. [Internet]. 2008 [cited 2012 Jul 19]. Available from: www.port.ac.uk/research/cmd/software.

277. Venkatachalam CM. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolymers. 1968;6(10):1425-36.

278. Hagler AT, Osguthorpe DJ, Dauberosguthorpe P, Hempel JC. Dynamics and conformational energetics of a peptide-hormone - vasopressin. Science. 1985;227(4692):1309-15.

279. Hruby VJ, Alobeidi F, Kazmierski W. Emerging approaches in the molecular design of receptor-selective peptide ligands - conformational, topographical and dynamic considerations. Biochem J. 1990;268(2):249-62.

280. Campbell AP, Sykes BD. The 2-dimensional transferred nuclear overhauser effect - theory and practice. Annu Rev Bioph Biom. 1993;22:99-122.

281. Zagrovic B, van Gunsteren WF. Comparing atomistic simulation data with the NMR experiment: How much can NOEs actually tell us? Proteins Struct Funct Bioinf. 2006;63(1):210-8.

282. Flaig D, Maurer M, Hanni M, Braunger K, Kick L, Thubauville M, et al. Benchmarking hydrogen and carbon NMR chemical shifts at HF, DFT, and MP2 levels. J Chem Theor Comput. 2014;10(2):572-8.

283. Zhu T, He X, Zhang JZH. Fragment density functional theory calculation of NMR chemical shifts for proteins with implicit solvation. Phys Chem Chem Phys. 2012;14(21):7837-45.

284. Hommes NJRV, Clark T. Regression formulae for ab initio and density functional calculated chemical shifts. J Mol Model. 2005;11(3):175-85.

285. Zhu T, Zhang JZH, He X. Automated fragmentation QM/MM calculation of amide proton chemical shifts in proteins with explicit solvent model. J Chem Theor Comput. 2013;9(4):2104-14.

286. Exner TE, Frank A, Onila I, Moller HM. Toward the quantum chemical calculation of NMR chemical shifts of proteins. 3. Conformational sampling and explicit solvents model. J Chem Theor Comput. 2012;8(11):4818-27.

287. Smith SG, Goodman JM. Assigning the stereochemistry of pairs of diastereoisomers using GIAO NMR shift calculation. J Org Chem. 2009;74(12):4597-607.

288. Nazarski RB, Pasternak B, Leśniak S. Three-component conformational equilibria of some flexible pyrrolidin-2-(thi)ones in solution by NMR data (δC, δH, and nJHH) and their DFT predictions: a confrontation of different approaches. Tetrahedron. 2011;67(36):6901-16.

289. Raiteri P, Laio A, Gervasio FL, Micheletti C, Parrinello M. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. J Phys Chem B. 2006;110(8):3533-9.

290. Barone G, Gomez-Paloma L, Duca D, Silvestri A, Riccio R, Bifulco G. Structure validation of natural products by quantum-mechanical GIAO calculations of 13C NMR chemical shifts. Chemistry. 2002;8(14):3233-9.

291. Barone G, Duca D, Silvestri A, Gomez-Paloma L, Riccio R, Bifulco G. Determination of the relative stereochemistry of flexible organic compounds by ab initio methods: conformational analysis and Boltzmann-averaged GIAO 13C NMR chemical shifts. Chemistry. 2002;8(14):3240-5.

292. Bern HA, Lederis K. A reference preparation for the study of active substances in the caudal neurosecretory system of teleosts. J Endocrinol. 1969;45(1):Suppl:xi-xii.

293. Ames RS, Sarau HM, Chambers JK, Willette RN, Aiyar NV, Romanic AM, et al. Human urotensin-II is a potent vasoconstrictor and agonist for the orphan receptor GPR14. Nature. 1999;401(6750):282-6.

294. Marchese A, Heiber M, Nguyen T, Heng HH, Saldivia VR, Cheng R, et al. Cloning and chromosomal mapping of three novel genes, GPR9, GPR10, and GPR14, encoding receptors related to interleukin 8, neuropeptide Y, and somatostatin receptors. Genomics. 1995;29(2):335-44.

295. Vaudry H, Do Rego JC, Le Mevel JC, Chatenet D, Tostivint H, Fournier A, et al. Urotensin II, from fish to human. Ann N Y Acad Sci. 2010;1200:53-66.

296. Tostivint H, Quan FB, Bougerol M, Kenigfest NB, Lihrmann I. Impact of gene/genome duplications on the evolution of the urotensin II and somatostatin families. Gen Comp Endocrinol. 2013;188:110-7.

297. Merlino F, Di Maro S, Munaim Yousif A, Caraglia M, Grieco P. Urotensin-II ligands: an overview from peptide to nonpeptide structures. J Amino Acids. 2013;2013:979016.

298. Zhu YC, Zhu YZ, Moore PK. The role of urotensin II in cardiovascular and renal physiology and diseases. Br J Pharmacol. 2006;148(7):884-901.

299. Labarrere P, Chatenet D, Leprince J, Marionneau C, Loirand G, Tonon MC, et al. Structure-activity relationships of human urotensin II and related analogues on rat aortic ring contraction. J Enzym Inhib Med Ch. 2003;18(2):77-88.

300. Bucharles C, Bizet P, Arthaud S, Arabo A, Leprince J, Lefranc B, et al. Concordant localization of functional urotensin II and urotensin II-related peptide binding sites in the rat brain: A typical occurrence close to the fourth ventricle. J Comp Neurol. 2014;522(11):2634-49.

301. Douglas SA, Naselsky D, Ao Z, Disa J, Herold CL, Lynch F, et al. Identification and pharmacological characterization of native, functional human urotensin-II receptors in rhabdomyosarcoma cell lines. Br J Pharmacol. 2004;142(6):921-32.

302. Jarry M, Diallo M, Lecointre C, Desrues L, Tokay T, Chatenet D, et al. The vasoactive peptides urotensin II and urotensin II-related peptide regulate astrocyte activity through common and distinct mechanisms: involvement in cell proliferation. Biochem J. 2010;428(1):113-24.

303. Ashton N. Renal and vascular actions of urotensin II. Kidney Int. 2006;70(4):624-9.

304. Stirrat A, Gallagher M, Douglas SA, Ohlstein EH, Berry C, Kirk A, et al. Potent vasodilator responses to human urotensin-II in human pulmonary and abdominal resistance arteries. Am J Physiol Heart Circ Physiol. 2001;280(2):925-8.

305. Kinney WA, Almond Jr HR, Qi J, Smith CE, Santulli RJ, de Garavilla L, et al. Structure-function analysis of urotensin II and its use in the construction of a ligand-receptor working model. Angew Chem Int Ed Engl. 2002;41(16):2940-4.

306. Foister S, Taylor LL, Feng JJ, Chen WL, Lin A, Cheng FC, et al. Design and synthesis of potent cystine-free cyclic hexapeptide agonists at the human urotensin receptor. Org Lett. 2006;8(9):1799-802.

307. McMaster D, Kobayashi Y, Rivier J, Lederis K. Characterization of the biologically and antigenically important regions of urotensin II. Proc West Pharmacol Soc. 1986;29:205-8.

308. Bandholtz S, Erdmann S, von Hacht JL, Exner S, Krause G, Kleinau G, et al. Urolinin: the first linear peptidic urotensin-II receptor agonist. J Med Chem. 2016;59(22):10100-12.

309. DASH 1.0 original release. [Internet]. 2008 [cited 2016 Nov 11]. Available from: www.port.ac.uk/research/cmd/software.

310. Bax A, Davis DG. Mlev-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. J Magn Reson. 1985;65(2):355-60.

311. Kumar A, Ernst RR, Wuthrich K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem Biophys Res Commun. 1980;95(1):1-6.

312. Kay LE, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc. 1992;114(26):10663-5.

313. Kontaxis G, Stonehouse J, Laue ED, Keeler J. The sensitivity of experiments which use gradient pulses for coherence-pathway selection. J Magn Reson Ser A. 1994;111(1):70-6.

314. Vosko SH, Wilk L, Nusair M. Accurate spin-dependent electron liquid correlation energies for local spin-density calculations - a critical analysis. Can J Phys. 1980;58(8):1200-11.

315. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields. J Phys Chem. 1994;98(45):11623-7.

316. Hehre WJ, Ditchfie.R, Pople JA. Self-consistent molecular-orbital methods .12. Further extensions of Gaussian-type basis sets for use in molecular-orbital studies of organic-molecules. J Chem Phys. 1972;56(5):2257-61.

317. Harihara.Pc, Pople JA. Influence of polarization functions on molecular-orbital hydrogenation energies. Theor Chim Acta. 1973;28(3):213-22.

318. Harihara.Pc, Pople JA. Accuracy of AHn equilibrium geometries by single determinant molecular-orbital theory. Mol Phys. 1974;27(1):209-14.

319. Gordon MS, Binkley JS, Pople JA, Pietro WJ, Hehre WJ. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. J Am Chem Soc. 1982;104(10):2797-803.

320. Cavanagh J, Fairbrother WJ, Palmer III AG, Rance M, Skelton NJ. Protein NMR Spectroscopy: Principles and Practice. 2nd ed. Oxford: Elsevier Ltd.; 2006. 996 p.

321. Grathwohl C, Wüthrich K. NMR studies of the rates of proline cis–trans isomerization in oligopeptides. Biopolymers. 1981;20(12):2623-33.

322. Vitagliano L, Berisio R, Mastrangelo A, Mazzarella L, Zagari A. Preferred proline puckerings in cis and trans peptide groups: implications for collagen stability. Protein Sci. 2001;10(12):2627-32.

323. Perczel A, Hollosi M. Turns. In: Fasman GD, editor. Circular Dichroism and the Conformational Analysis of Biomolecules. New York: Plenum Press; 1996. p. 285-380.

324. Gerthsen C, Kneser HO. Physik. 15 ed. Vogel H, editor. Heidelberg: Springer; 1986.

325. Wedler G, Freund H-J. Lehrbuch der Physikalischen Chemie. 6. vollst. überarb. u. aktualis. Auflage ed: Wiley-VCH Verlag GmbH & Co. KGaA; 2012.

326. chemie.de Lexikon. [Internet]. Berlin, Germany: CHEMIE.DE Information Service GmbH. © 1997-2017 [cited. Available from: http://www.chemie.de/lexikon/.

327. MDASH 3.0.0. [Internet]. Centre for Molecular Design, University of Portsmouth. 2016 [cited 2017 Mar 03]. Available from: www.port.ac.uk/research/cmd/software.

328. The PyMOL Molecular Graphics System, Version 1.3 educational. [Internet]. Schrödinger, LLC. 2010 [cited 2017 04]. Available from: http://pymol.org/educational/.

329. POV-Ray 3.6.2, Microsoft Windows version. [Internet]. Vision Raytracer Pty. Ltd. 2011 [cited 2017 Apr]. Available from: http://www.povray.org/.

330. Gnuplot 4.6. [Internet]. 2012 [cited 2017 Apr]. Available from: http://www.gnuplot.info/.

331. UCSF Chimera 1.6.1. [Internet]. 2012 [cited 2017 Apr]. Available from: http://www.cgl.ucsf.edu/chimera/.

332. ChemBioDraw 13. [Internet]. Alfasoft Group. 2016 [cited 2016 Mar 17]. Available from: http://www.adeptscience.co.uk/products/lab/chembiodraw/chembiodraw-ultra-suite.html.

333. Excel 2013. [Internet]. Microsoft. 2013 [cited 2016 Mar 17]. Available from: https://products.office.com/en-us/excel.

334. Adobe Photoshop CS5, 12.0.4 x64. [Internet]. Adobe Systems Incorporated. 1990-2010 [cited 2017 Apr]. Available from: http://www.adobe.com/products/photoshop.html.

335. Adobe Illustrator CS5, 15.0.2. [Internet]. Adobe Systems Incorporated. 1987-2010 [cited 2017 Apr]. Available from: http://www.adobe.com/products/photoshop.html.