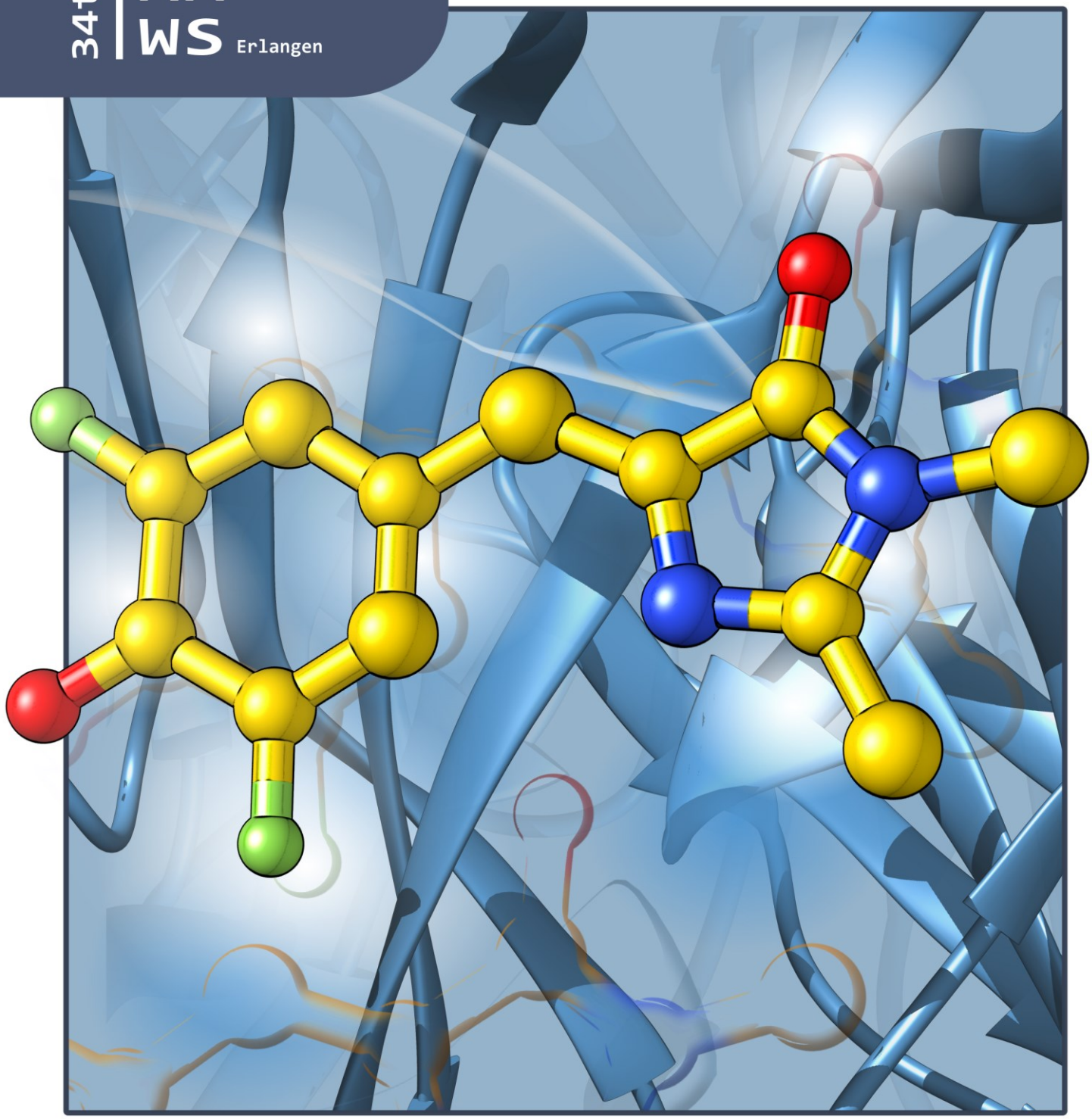
34th MMWS Erlangen

MMWS2020.MGMS-DS.DE



17/02-19/02/2020

MOLECULAR MODELLING WORKSHOP ERLANGEN

MOLECULAR MODELLING WORKSHOP 2020

Welcome to the 34th Molecular Modelling Workshop (MMWS)

This year's Workshop is the 18th held in Erlangen. The first 16 Workshops were known as the Darmstadt Molecular Modelling Workshop and, as the name suggests, took place in Darmstadt under the leadership of Jürgen Brickmann and his group. The eighth MMWS (1994) was the first to take place under the auspices of the Molecular Graphics and Modelling Society – Deutschsprachige Sektion (MGMS-DS e.V.), which has been responsible ever since.

Thus, the MMWS can look back on a long history of giving graduate students and postdocs the opportunity to present their work. It predates the annual Young Modellers' forum organized by the parent MGMS in London and the equivalent workshop run by the Association of Molecular Modellers in Australasia (AMMA) in association with the MGMS. We are proud that the MMWS has become a fixture in the molecular modelling scene in Europe and that it continues to provide students and young researchers a stage to present their work.

As in previous years, the technical conference management of the Computer-Chemie-Centrum, CCC, is supported by the Bioinformatics group headed by Heinrich Sticht. This year's scientific MMWS program was organised by Christof Jäger from the University of Nottingham.

Last year's Workshop marked the point, at which the number of Workshops held in Erlangen exceeded the number hosted in Darmstadt. This year marks the point where the technical management of the CCC will be handed over from Tim Clark to his successor Petra Imhof, who we also welcome to Erlangen during this Workshop.

This time, we are happy to welcome five plenary speakers for our MMWS. Pavel Jungwirth from the Czech Academy of Sciences (Prague), Barbara Kirchner from the University of Bonn (Germany), Charles Laughton from the University of Nottingham (United Kingdom), Birgit Strodel from the Forschungszentrum Jülich (Germany), and Petra Imhof from the University of Stavanger (Norway) and coming technical director of the CCC.

The combination of these five excellent plenary speakers is intended to cover a broad set of flavours of molecular modelling and computational chemistry and enable MMWS to remain at the forefront of the rapidly changing face of modelling in Europe and the rest of the world and to provide inspiration for young researchers.

Scientific program

Prof. Dr. Christof Jäger

Faculty of Engineering
University of Nottingham
University Park
NG7 2RD
Nottingham, UK
Tel: +44 (0)115 74 86849
christof.jaeger@nottingham.ac.uk

Technical coordination

PD Dr. Harald Lanig

Zentralinstitut für Scientific Computing
Friedrich-Alexander-Universität
Erlangen-Nürnberg (FAU)
Martensstraße 5a
91058 Erlangen, Germany
Tel: +49 (0)9131 85 20781
harald.lanig@fau.de

DEAR COLLEGUES,

The 34th Molecular Modelling Workshop (February, 17th – 19th 2020) in Erlangen provides research students and new postdoctoral scientists the perfect opportunity to present their research to the molecular modelling community. Scientists at the beginning of their academic careers are able to meet new colleagues in academia and industry.

Every year, the organisers welcome both poster and lecture contributions from all areas of molecular modelling including life sciences, physical sciences, material sciences, and the nano sciences.

The aim of the Modelling Workshop is to introduce research in progress. The workshop is the perfect venue to introduce new methods in molecular modelling that can be applied to many disciplines. The workshop is suitable for everyone, those who want to gain experience in presentation skills and those who just want to network in a friendly relaxed environment.

Contributions are welcome
from all areas of molecular modelling -
from the life sciences, computational biology,
computational chemistry to materials sciences.

Our plenary speakers this year are (in alphabetical order):

PROF. DR. PETRA IMHOF
University of Stavanger, Norway

PROF. DR. PAVEL JUNGWIRTH
Czech Academy of Sciences/Prague, Czech Republic

PROF. DR. BARBARA KIRCHNER
Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

PROF. DR. CHARLES LAUGHTON
University of Nottingham, UK

PROF. DR. BIRGIT STRODEL
Forschungszentrum Jülich, Germany

AWARDS

Traditionally, there will be two *Poster Awards* of 100 Euro each and three *Lecture Awards* for the best talks sponsored by the MGMS-DS:

1st Winner

Travel bursary to the Young Modellers Forum in the United Kingdom (travel expenses are reimbursed up to 500 Euro)

2nd Winner

up to 200 Euro travel expenses reimbursement

3rd Winner

up to 100 Euro travel expenses reimbursement

Only undergraduate and graduate research students qualify for the poster and lecture awards.

Additionally, we are happy that, for the first time, the Wiley-VCH journal "Advanced Theory and Simulations" is providing several *book token awards* to outstanding workshop contributions.

MGMS-DS E.V. ANNUAL MEETING

The general meeting of the MGMS (German Section) will be held during the workshop. We cordially invite all conference delegates to participate in the annual meeting of the society!

FEES

The conference fee amounts to 100 Euro (students: 50 Euro). This fee includes the annual membership fee for the MGMS-DS e.V.

WI-FI ACCESS

During the workshop, Wi-Fi access is possible via **eduroam** (SSID). Please have your Wi-Fi configured in advance or ask your local administrator for detailed information about your eduroam access. Links to general information about eduroam can be found on the workshop website mmws2020.mgms-ds.de

PRE-CONFERENCE WORKSHOP (SCHRÖDINGER)

After successful and positively perceived workshops in previous years, Schrödinger is again offering a pre-conference workshop. The topic of this year's workshop is "Investigating small molecule & biomacromolecule dynamics and interactions with the Schrödinger Suite". This event will take place on Monday (11 am - 1 pm).

LOCATION

The Pre-conference Workshop about the Schrödinger Suite takes place at the Computer-Chemistry-Center (CCC), Nägelsbachstr. 25 (see conference web page for details).

Conference location: All talks, coffee breaks, the poster sessions and the buffet dinner on Monday, February 17th will take place at the Institute for Biochemistry, Fahrstr. 17, 91054 Erlangen (and *not* at the old location, i.e. Institute for Organic Chemistry, Henkestr. 42).

The *Social Event "Visit at a typical Erlanger Gasthaus – Biergarten"* will take place at Gasthaus "Steinbach Bräu" (www.steinbach-braeu.de), Vierzigmannstr. 4, on Tuesday evening. Food and drinks will be available at your own expense.

Prof. Dr. Petra Imhof

Petra Imhof studied Chemistry and received her Diploma in 1998 at the Heinrich-Heine University, Düsseldorf, where she also obtained her PhD. After PostDoc stays at Mülheim a.d. Ruhr and Heidelberg, she moved to Berlin, where she became Professor for Computational Biophysics in 2013. In 2019, she moved to University of Stavanger, Norway, where she became Professor for Computational Chemistry.

Prof. Dr. Pavel Jungwirth

Pavel Jungwirth obtained his MSc in Physics in 1989 at the Charles University in Prague. He completed his PhD in Computational Chemistry at the J. Heyrovsky Institute of Physical Chemistry (Prague) in 1993 with Prof. R. Zahradník. At the moment, he is professor (External Faculty) at the Charles University in Prague, Faculty of Mathematics and Physics. Pavel Jungwirth received several national and international awads.

His research interests focus on the investigation of ions at biological interfaces and protein environment using molecular simulations. He published more than 300 review and research papers.

Prof. Dr. Barbara Kirchner

Barbara Kirchner studied chemistry in Freiburg, Mainz and Chemnitz and completed her PhD on classical molecular dynamics simulations at the University of Basel, Switzerland. Before taking up her current position in Bonn at the Mulliken Center (University of Bonn) as one of the chairs of Theoretical Chemistry, she held a chair of theoretical chemistry at the University of Leipzig. Barbara worked as a postdoctoral researcher in various institutes on ab initio molecular dynamics simulations and as a visiting scientist in Brisbane, Australia.

Barbara's work is diverse, ranging from understanding liquids and solvents, through intermolecular forces and processes in the condensed phase, to quantum chemical analysis of interesting molecules, and methodological developments. Last year she received for "a trajectory in ionic liquids" with the Ruth M. Lynden-Bell award.

Prof. Dr. Charles Laughton

Charlie Laughton is Professor of Computational Pharmaceutical Science at the University of Nottingham. Charlie got his DPhil in synthetic organic chemistry from the University of Oxford, but over the span of a couple of postdocs at first Aston University in Birmingham and then the Institute of Cancer Research in London, shifted his focus to computational chemistry, structural biology and drug design. At Nottingham his research is split between multidisciplinary drug discovery and development projects (with colleagues in Medicinal Chemistry, Chemical Engineering, and Pharmacology), and computational methods development, where he works will colleagues in applied Maths and Computer Science.

Prof. Dr. Birgit Strodel

Birgit Strodel studied chemistry at Düsseldorf (Heinrich-Heine-University) and Chapel Hill (University of North Carolina, USA). After her diploma thesis, she moved to Frankfurt, where she obtained her PhD with Prof. Dr. Gerhard Stock. She then joined the group of Prof. Dr. David J. Wales (University of Cambridge, UK) for a three-years PostDoc stay. In 2009, she returned to Germany and became head of a young investigators group at the institute of complex systems in the Forschungszentrum Jülich, where she was appointed head of the computational biochemistry group at the institute of biological information processing. Since 2018, she is also full professor at the Heinricht-Heine-University.

Birgit received the “JCP Editor’s Choice Award” in 2020 for for publishing one of the five most outstanding papers in *The Journal of Chemical Physics* in 2019.

Lectures Program

PROGRAM

Monday, February 17th 2020

11:00-13:00	**Pre-conference workshop**
11:00-14:00	**Registration**
14:00-14:10	**Welcome remarks / Agenda review**
14:10-14:30	**L01: Marcus Conrad (Erlangen, Germany)** Histamine H_2 receptor in complex with G_s protein: a comprehensive molecular dynamics study
14:30-14:50	**L02: Marko Hanževački (Nottingham, UK)** From the catalytic mechanism of the glycyl radical enzyme pyruvate formate-lyase to the dynamics of its activation
14:50-15:40	**PLENARY LECTURE I: Pavel Jungwirth** Electrons in liquid ammonia and in water: from blue electrolytes to bronze colored metallic solutions
15:40-16:10	**Coffee Break**
16:10-16:30	**L03: Krzysztof K. Bojarski (Gdańsk, Poland)** Challenges in modeling protein-glycosaminoglycan systems
16:30-16:50	**L04: Nicolas Tielker (Dortmund, Germany)** A SAMPL journey, or "There and Back Again": from 2 over 5 to 6 and back to square one
16:50-17:10	**L05: Conrad Hübler (Freiberg, Germany)** Selective complexation of carbohydrates with artificial receptors - A computational challenge
17:10-17:30	**L06: Oldamur Hollóczki (Bonn, Germany)** On the environmental effects of nanoplastics
17:30-17:50	**L07: Gianna Pohl (Merck Healthcare KGaA)** Compound optimization - What we should know about the SAR trends we don't know
18:00-19:00	**Annual Meeting of the MGMS-DS e.V.**
19:30	**Buffet – Dinner**

PROGRAM

Tuesday, February 18th 2020

09:00-09:20	**L08: Nils E. R. Zimmermann (Berkely, USA)** Potential of electrostatics-finite ion size (PfEFIS) method: towards automatic ion diffusion network analysis in solids
09:20-09:40	**L09: Patrik Melix (Dresden, Germany)** Ligand field molecular mechanics applied to MOFs
09:40-10:00	**L10: Alexandra Freidzon (Moscow, Russia)** Multireference study of charge and energy transfer in organic semiconductors
10:00-10:20	**L11: Amin Alibakhshi (Kiel, Germany)** Reproducing high accuracy QM interactions via empirical force fields
10:20-10:50	**Conference Photo & Coffee Break**
10:50-11:10	**L12: Sabahuddin Ahmad (Düsseldorf, Germany)** Understanding the substrate access mechanism in Phospholipase A from *Pseudomonas aeruginosa*
11:10-11:30	**L13: Zlatko Brkljača (Zagreb, Croatia)** Insights into mechanism of proton translocation assisted by membrane proteins
11:30-12:20	**PLENARY LECTURE II: Birgit Strodel** Modelling peptide aggregation influenced by *in vivo* conditions
12:20-13:30	**Lunch**
13:30-14:50	**POSTER SESSION**
14:50-15:10	**L14: Nadine Homeyer (Schrödinger)** Rapid exploration of synthetically tractable chemical space for hit-to-lead & lead optimization
15:10-15:30	**L15: Xiyu Chen (San Francisco, USA)** Kinematic flexibility analysis of kinase activation loops
15:30-15:50	**L16: Andrea Thorn (Würzburg, Germany)** Modelling biological data with experimental restraints
15:50-16:10	**L17: Slavica Subic (Erlangen, Germany)** Prediction of odour (PrOdour) – data exploration
16:10-16:40	**Coffee Break**

PROGRAM

Tuesday, February 18th 2020

16:40-17:00	**L18: Laura Schulz (Frankfurt, Germany)** Sampling of ligand-induced conformational changes in renin and factor VIIa
17:00-17:20	**L19: Sören von Bülow (Frankfurt, Germany)** Dynamics of crowded macromolecules from atomistic simulations
17:20-18:10	**PLENARY LECTURE III: Charles Laughton** Enhancing conformational sampling with machine learning and the cloud
19:00	**Social Event: Bierkeller (Steinbach Bräu)**

Wednesday, February 19th 2020

09:00-09:20	**L20: Akinjide Oluwajobi (Pilsen, Czech Republic)** Atomistic modelling of high entropy alloys
09:20-09:40	**L21: Özge Özkılınç (Ankara, Turkey)** Effect of chalcogendiazole groups on the conductive properties of EDOT containing polymers
09:40-10:00	**L22: Nataša Vučemilović-Alagić (Zagreb, Croatia)** Imidazolium-based ionic liquids under the magnifier of molecular simulations
10:00-10:20	**L23: Javier Luque Di Salvo (Palermo, Italy)** Bottom-up DFT-MD multi-scale modelling of hydrated bulk anion exchange membranes
10:20-10:50	**Coffee Break**
10:50-11:10	**L24: Ateeque Malani (Mumbai, India)** Adsorption of ions at solid-liquid interface: role of ion hydration structure and energies
11:10-11:30	**L25: Christian R. Wick (Erlangen, Germany)** Modelling SILP catalysis: the water-gas-shift reaction
11:30-12:20	**PLENARY LECTURE IV: Barbara Kirchner** Insights into ionic and molecular liquids and their interfaces from computational chemistry

PROGRAM

Wednesday, February 19th 2020

12:20-13:40	**Lunch**
13:40-14:00	**L26: Christina C. Roggatz (Hull, UK)** Histamine – a signalling cue influenced by fluctuating environmental conditions and climate change?
14:00-14:20	**L27: Anna Kahler (Erlangen, Germany)** Intrinsic flexibility and structural stability of proteins
14:20-14:40	**L28: Illimar Rekand (Bergen, Norway)** Druggability predictions of ribonucleic acid crystal structures
14:40-15:30	**PLENARY LECTURE V: Petra Imhof** Recognition and specificity in protein-DNA interactions – insight from molecular simulations
15:30-15:50	**Poster & Lecture awards, Closing**

Poster Session

POSTER SESSION

Tuesday, February 18th 2020 13:30-14:50

P01 **Sabahuddin Ahmad (Düsseldorf, Germany)**
Understanding the substrate access mechanism in Phospholipase A from *Pseudomonas aeruginosa*

P02 **Amin Alibakhshi (Kiel, Germany)**
Reproducing high accuracy QM interactions via empirical force fields

P03 **Felix Bänsch (Recklinghausen, Germany)**
A new approach to DPD repulsion parameter estimation

P04 **Marcel Baltruschat (Dortmund, Germany)**
Machine learning meets pK_a

P05 **Filippo Balzaretti (Bremen, Germany)**
Water contaminants on different models of rutile TiO2 - a DFT(B) / FF approach

P06 **Frank R. Beierlein (Erlangen-Nürnberg, Germany)**
Probing DNA conformation: a multi-technique approach

P07 **Iga D. Biskupek (Gdańsk, Poland)**
Studying conformational transitions of selected proteins using UNRES coarse-grained simulations with Lorentzian restraints

P08 **Manuel Deubler (Erlangen-Nürnberg, Germany)**
Disarming the glycan shield of HIV: broadly neutralizing antibody PGT122 against HIV

P09 **Lukas Eberlein (Dortmund, Germany)**
Nucleic acid building blocks: The role of tautomerism

P10 **Lennart Eisel (Dortmund, Germany)**
A QM/SQM embedded cluster RISM approach for predicting EPR parameters of protein-bound nitroxide spin probes

P11 **Enrico Gandini (Milan, Italy)**
Computational studies of antifreeze peptides

P12 **Kevin Höllring (Erlangen-Nürnberg, Germany)**
Influence of self-organization induced local clustering on self-diffusion of [CnMIm+] [NTf2-] in bulk and nanoconfined systems

P13 **Michael C. Hutter (Saarbrücken, Germany)**
Druggability prediction using gene ontology terms

P14 **Julia B. Jasper (Frankfurt, Germany)**
What can we learn from local water happiness?

Please kindly remove your posters on tuesday evening!

POSTER SESSION

Tuesday, February 18th 2020 13:30-14:50

P15 **Sascha Jung (Dortmund, Germany)**
Protonation effects of endothiapepsin-fragment complexes

P16 **Jonas Kaindl (Erlangen-Nürnberg, Germany)**
Metadynamics simulations reveal binding of PAR2 to its trypsin activated N-terminus

P17 **Jan Kaiser (Freiburg, Germany)**
Monte Carlo simulation and thermodynamic integration applied to protein charge transfer

P18 **Mattia Livraghi (Erlangen-Nürnberg, Germany)**
Highly cross-linked epoxy networks under mechanical strain

P19 **Fidele Ntie-Kang (Buea, Cameroon)**
Discovery of histone deacetylase inhibitors and HIV-1 latency reversing agents by large-scale virtual screening

P20 **Lalehan Oktay (Istanbul, Turkey)**
Targeted in silico screening of small molecule databases against breast cancer

P21 **Simon Schäfer (Erlangen-Nürnberg, Germany)**
Mind the gap - linking crystal structures and sequences without misrepresentation in antibody research

P22 **Jonas Schaub (Jena, Germany)**
Rule-based in-silico fragmentation for the analysis of natural product chemical space

P23 **Maximilian F. Schmidt (Erlangen-Nürnberg, Germany)**
Hybridization of β-adrenergic agonists and antagonists confers G protein bias

P24 **David C. Schröder (Bielefeld, Germany)**
Structure determination of peptides and peptidotriazoles by MD simulations

P25 **Laura Schulz (Frankfurt, Germany)**
Sampling of ligand-induced conformational changes in renin and factor VIIa

P26 **Slavica Subic (Erlangen-Nürnberg, Germany)**
Prediction of odour (PrOdour) – data exploration

P27 **Navista S. O. Ujiantari (Innbruck, Austria)**
Pharmacophore-based virtual screening to identify new β3-adrenergic receptor agonists

Please kindly remove your posters on tuesday evening!

POSTER SESSION

Tuesday, February 18th 2020 13:30-14:50

P28 **Emire Uyanık (Ankara, Turkey)**
Design of thiophene and thiadiazoloquinoxaline containing semiconducting polymer

P29 **Vincent Stegmaier (Berlin, Germany)**
Protonation dynamics in cytochrome c oxidase

P30 **Sören von Bülow (Frankfurt, Germany)**
Dynamics of crowded macromolecules from atomistic simulations

P31 **Marius Wenz (Berlin, Germany)**
In search of binding competent structures of the tWW domain using classical molecular dynamics simulations

P32 **Sandro Wrzalek (Berlin, Germany)**
Identification of conformational states and transitions in between from molecular dynamics trajectories and vibrational signatures

P33 **Karolina Zięba (Gdańsk, Poland)**
Simulations of the membrane proteins in the lipid bilayer with the UNRES coarse-grained force field

All abstracts are available on the conference web site:
www.mmws2020.mgms-ds.de

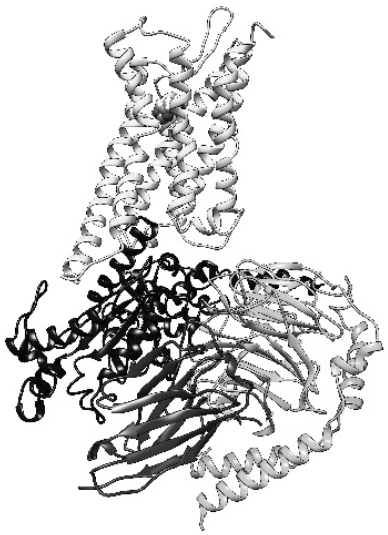
Please kindly remove your posters on tuesday evening!

Abstracts

Histamine H_2 receptor in complex with G_s protein: a comprehensive molecular dynamics study

Marcus Conrad, Christian A. Söldner, Heinrich Sticht

Bioinformatik, Institut für Biochemie, Friedrich-Alexander-Universität Erlangen-Nürnberg



The histamine H_2 receptor plays an important role in the regulation of gastric acid secretion. Therefore, it is a main drug target for the treatment of gastroesophageal reflux or peptic ulcer disease [1]. However, to date there is no crystal structure available for this receptor, which would be a prerequisite for *rational* drug design. Hence we created a model of the active histamine H_2 receptor-G_s complex based on the structure of the ternary complex of the β_2 adrenoceptor [2]. The binding mode of its natural ligand histamine was deduced using a multiple walker metadynamics protocol [3],[4]. We conducted further refinements by conventional molecular dynamics simulations and used the resulting structure to examine the interactions of the receptor with histamine and the G_s protein. To further validate the model and investigate its switchability towards the inactive ensemble we conducted several simulations including gaussian accelerated molecular dynamics [5] in which the G protein was removed. Since the main G protein interaction is mediated by the $\alpha 5$ helix of G_s, we additionally simulated a shortened version of the G_s, namely the terminal $\alpha 5$ helix to examine whether this part is sufficient to maintain the active conformation. The overall results of our study provide detailed insights into the dynamics of the H_2 receptor complex that can contribute to future drug development.

[1] Stephen Holt. "Over-the-Counter Histamine H 2-Receptor Antagonists". In: *Drugs* 47.1 (1994), pp. 1–11.

[2] Søren GF Rasmussen et al. "Crystal structure of the β-2-adrenergic receptor–Gs protein complex". In: *Nature* 477.7366 (2011), p. 549.

[3] Noureldin Saleh et al. "An efficient metadynamics-based protocol to model the binding affinity and the transition state ensemble of G-protein-coupled receptor ligands". In: *Journal of chemical information and modeling* 57.5 (2017), pp. 1210–1217.

[4] Christian A Söldner, Anselm HC Horn, and Heinrich Sticht. "Binding of histamine to the H1 receptor—A molecular dynamics study". In: *Journal of molecular modeling* 24.12 (2018), p. 346.

[5] Yinglong Miao, Victoria A Feher, and J Andrew McCammon. "Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation". In: *Journal of chemical theory and computation* 11.8 (2015), pp. 3584–3595.

From the catalytic mechanism of the glycyl radical enzyme pyruvate formate-lyase to the dynamics of its activation

Marko Hanževački,[a,c,d] Karmen Čondić-Jurkić,[b] Radha Dilip Banhatti,[c] Anna Croft,[a] Christof Jäger,[a] Ana-Sunčana Smith[c,d] and David Smith[c]

[a]Department of Chemical and Environmental Engineering, University of Nottingham, Nottingham, United Kingdom, [b]Memorial Sloan Kettering Cancer Center, New York, United States, [c]Group for Computational Life Sciences, Division for Physical Chemistry, Ruđer Bošković Institute, Zagreb, Croatia, [d]PULS Group, Department of Physics, Interdisciplinary Center for Nanostructured Films, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Pyruvate formate-lyase (PFL) is a glycyl radical enzyme (GRE) that reversibly converts pyruvate and coenzyme A (CoA) to formate and acetyl-CoA in two half-reactions. PFL has a crucial role in the central glucose metabolism of *E. coli* and other microbes when switching to anaerobic conditions. [1] Due to their versatile functionalities, GREs have potential applications in biochemical and enzyme engineering. [2] Recently, it has been proposed that GREs also have a vital role in the numerous metabolic and biosynthetic pathways and environment due to their high abundance in the human gut microbiota.

The reactions catalyzed by GREs involve an extremely reactive and short-lived glycyl radical intermediate, which makes them very challenging for experimental studies. However, molecular modeling methods, based on classical and quantum chemical approach, provide valuable tools for the successful studies of these type of biomolecular systems.

Certain conformational changes of PFL and the entry of CoA in the active site are crucial for the second half-reaction to take place. Recently we proposed through extensive molecular dynamics (MD) simulations that the approach of CoA near the active site occurs through the newly appearing "open" state of the identified channel. [3] Further, by employing enhanced sampling methods we have shown that CoA is more likely to reside inside the active site after the first half-reaction whereby we characterized the stable and potentially reactive binding poses of CoA in the active site. [4]

Using full enzyme as a model for QM/MM calculation we explored the mechanism of both half-reactions. We confirm the progression of the first half-reaction with pyruvate in two steps without CoA occupying the active site. Further, we propose two-step mechanism for the second half-reaction with CoA. In addition, while a definite possibility of H-abstraction from CoA in the active site before the first half-reaction exists, we propose that this would cause a premature quenching of the radical and could ultimately lead to the inactivation of PFL.

In order to better understand the activation of GREs we explore the conformational dynamics of several representative GREs and the prototypical member of the radical SAM enzyme, namely pyruvate formate-lyase activating enzyme (PFL-AE). In this respect we performed preliminary multiple atomistic unrestrained MD simulations ranging to the timescales of several microseconds to uncover the connection between the fundamental conformational dynamics of GREs and the potential effects of their binding on the dynamics of the activating enzyme.

[1] L. R. F. Backman, M. A. Funk, C. D. Dawson, C. L. Drennan, *Crit. Rev. Biochem. Mol. Biol.*, **2017**, *52*, 674-695.
[2] C. M. Jäger, A. K. Croft, *Chem. Bio. Eng. Rev.*, **2018**, *5*, 143-162.
[3] M. Hanževački, K. Čondić-Jurkić, R. D. Banhatti, A-S. Smith, D. M. Smith, *Chem. Eur. J.*, **2019**, *25*, 8741-87535.
[4] M. Hanževački, R. D. Banhatti, K. Čondić-Jurkić, A-S. Smith, D. M. Smith, *J. Phys. Chem. A*, **2019**, *123*, 9345-9356.

Electrons in Liquid Ammonia and in Water: From blue electrolytes to bronze colored metallic solutions

Pavel Jungwirth

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Prague 6, Czech Republic
E-mail: pavel.jungwirth@uochb.cas.cz

The electronic structure of alkali metal – liquid ammonia mixtures at concentrations spanning from blue electrolytes to bronze colored metallic solutions is characterized by means of photoelectron spectroscopy in liquid microjets aided by ab initio molecular dynamics simulations and quantum chemical calculations. The experimental PE signature of ammoniated electrons and dielectrons with vertical detachment energy of about 2 eV is in agreement with calculations. By comparing solutions of different alkali metals at low to intermediate concentrations, we show that the vertical detachment energy is only weakly sensitive to pairing with an alkali metal cation. This is confirmed by electronic structure calculations, which also show that the transition from the ammoniated electron to dielectron regimes is connected with a minor change of vertical detachment energy.

Upon increasing the alkali metal concentration, the photoelectron peak at ~2 eV broadens asymmetrically toward higher binding energies, which primarily signifies a build-up of a metallic conduction band. The present study shows that this electrolyte–to–metal transition is a gradual process rather than an abrupt first order transition. From the molecular point of view, the transition may be understood in a simplified way as a coalescence of individual solvated electrons and dielectrons upon increasing alkali metal doping, with the metallic behavior appearing at the threshold of percolation.

Challenges in modeling protein-glycosaminoglycan systems

Krzysztof K. Bojarski, Sergey A. Samsonov

Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland
krzysztof.bojarski@ug.edu.pl

Glycosaminoglycans (GAGs), long unbranched negatively charged periodic polysaccharides are present in extracellular matrix and lysosomes where they play a vital role in diverse biological processes such as adhesion, anticoagulation or signalling cascades. Due to their predominantly electrostatic interactions with respective protein targets, GAGs can mediate biological functions of various proteins including chemokines, growth factors and collagen. It is reported that GAGs are also involved in several enzymatic reactions including cathepsins and their immature precursors – procathepsins.

Due to high cost of reagents and devices used in experiments as well as timescale needed to analyse protein-GAG systems, theoretical approaches are often used to aid *in vitro* and *in vivo* studies which allowes us to characterize studied systems at molecular level. Nevertheless, modeling protein-GAG complexes still represent substantial challenge in computational approaches. Features that make modeling GAG containing systems challenging are: *i)* extensive conformational space of GAGs in terms of their glycosidic linkages and monosaccharide rings; *ii)* GAGs highly charged nature; *iii)* GAGs preference to bind at solvent-exposed and spatially close but sequentially not necessarily successive positively charged amino acid patches made up of long and, therefore, flexible lysine or arginine residues; *iv)* the multipose binding observed in several protein-GAG complexes; *v)* highly variable sulfation pattern of GAGs known as "sulfation code" defining its structural properties, molecular recognition and functional activity; *vi)* the possibility of a protein-GAG complex formation, in which GAG orientation is 180° rotated in reference to experimental structure of the complex. In such case, the rotated GAG interacts with exactly the same aminoacid residues as in the X-ray structure, but its direction in reference to its reducing/non-reducing is opposite.

In our studies, we proposed methodology that could address these challenges. We performed long, microsecond scale Molecular Dynamics (MD) of Fibroblast Growth Factor-1-heparin (FGF-1-HP) in order to characterize the impact of glycosidic linkage conformations and ring puckers on the stability of the complex represented as the calculated binding free energies, hydrogen bonds between the ligand and receptor as well as native/non-native contacts. In the present study, we would like to explain how the orientation of a GAG on a protein surface may affect the stability of a complex. In order to do that we performed 1 μs MD simulation of FGF-1-HP complex. In our analysis, the HP of different length of the chain (dp2 and dp4 which were modeled as well as dp6 present in X-ray structure; dp stands for degree of polymerisation) was bound in parallel and antiparallel orientation on the protein surface in the binding site corresponding to X-ray structure (pdb ID: 2AXM) of FGF-1-HP dp6 complex (Figure 1).

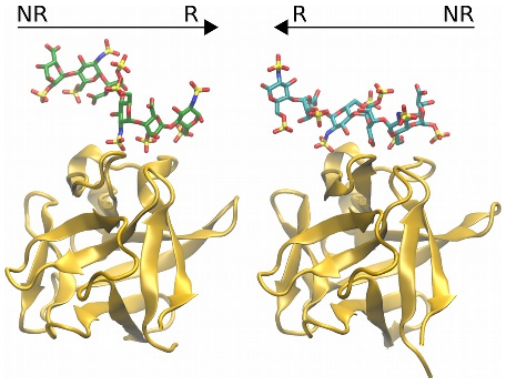


Figure 1: The structure of FGF-1 complex with HP dp6 in different orientations on the protein surface.

A SAMPL journey, or "There and Back Again": From 2 over 5 to 6 and back to square one

Nicolas Tielker,[1] Lukas Eberlein,[1] Stefan Güssregen,[2] Stefan M. Kast[1]

[1]Fakultät für Chemie und Chemische Biologie, TU Dortmund, Germany
[2]Sanofi-Aventis Deutschland GmbH, Frankfurt am Main, Germany

The SAMPL (Statistical Assessment of the Modeling of Proteins and Ligands) series of blind prediction challenges was designed to compare the performance of widely different computational methods for predicting physicochemical properties of drug-like compounds in solution. Over more than a decade our group participated in several SAMPL challenges by employing a statistical-mechanical integral equation approach to solvation thermodynamics [1,2] coupled to a quantum-level model of the solute in the form of the "embedded cluster reference interaction site model" (EC-RISM). [3-7] Such a method allows for efficient calculations of Gibbs energies of solvation of neutral and ionic molecules for a wide range of solvents and solvent mixtures. Optimized with respect to quantitative accuracy, both the electronic relaxation and the excess chemical potential governing the insertion into a solvent are taken into account for predicting the Gibbs energy of the molecule in solution, from which other physicochemical properties can be derived. One part of the latest challenge devoted to small molecule properties, SAMPL6 (2017-2019), focused on acidity constants that are difficult to predict since the compounds often contain functional groups and scaffolds that imply a multitude of tautomeric states ("microstates"). In the second part participants were asked to predict octanol-water partition coefficients, log P_{ow}, for a subset of neutral SAMPL6 compound states.

We developed and optimized a workflow to calculate all the above-mentioned properties of a given compound by determining the Gibbs energies of the most probable conformations for all tautomers of the relevant neutral and, where necessary, ionic forms by EC-RISM calculations for SAMPL6, yielding a root mean square deviation of about 1.0 pK units for acidity constants [6] and 0.5 log units for the subset of molecules for which experimental log P_{ow} were measured. [7] As we previously participated in the SAMPL2 challenge (2009) [4] for tautomer predictions and the SAMPL5 challenge (2015) [5] for cyclohexane-water distribution coefficients using workflows and methodology available back then, it is timely to revisit these earlier datasets by employing the most advanced EC-RISM-based SAMPL6 technology. The retrospective analysis reveals a number of insights and surprises in terms of expected systematic progress and the impact of error compensation on predictive capabilities.

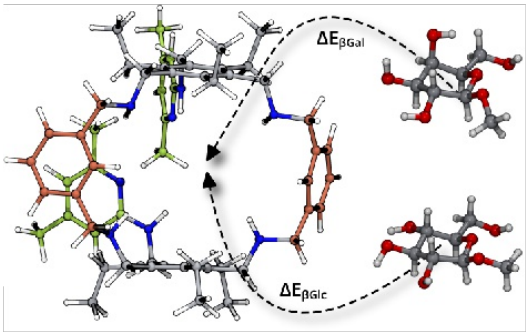
[1] E. L. Ratkova, D. S. Palmer, M. V. Fedorov, *Chem. Rev.* **2015**, *115*, 6312-6356.
[2] J. Heil, S. M. Kast, *J. Chem. Phys.* **2015**, *142*, 114107.
[3] T. Kloss, J. Heil, S. M. Kast, *J. Phys. Chem. B* **2008**, *112*, 4337-4343.
[4] S. M. Kast, J. Heil, S. Güssregen, K. F. Schmidt, *J. Comput.-Aided Mol. Des.* **2010**, *24*, 343-353.
[5] N. Tielker, D. Tomazic, J. Heil, T. Kloss, S. Ehrhart, S. Güssregen, K. F. Schmidt, S. M. Kast, *J. Comput.-Aided Molec. Des.* **2016**, *30*, 1035-1044.
[6] N. Tielker, L. Eberlein, S. Güssregen, S. M. Kast, *J. Comput.-Aided Mol. Des.* **2018**, *32*, 1151-1163.
[7] N. Tielker, D. Tomazic, L. Eberlein, S. Güssregen, S. M. Kast, *J. Comput.-Aided Mol. Des.* **2020**, https://doi.org/10.1007/s10822-020-00283-4.

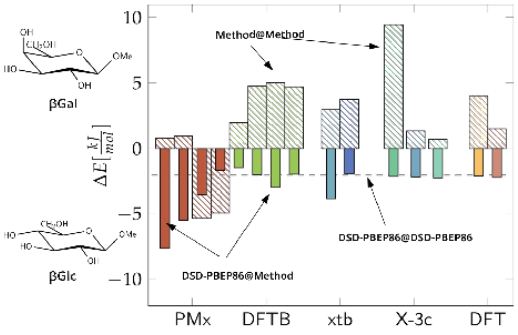
Selective complexation of carbohydrates with artificial receptors - A computational challenge

Conrad Hübler and Monika Mazik
Institut für Organische Chemie, Technische Universität Freiberg, Leipziger Straße 29, 09599 Freiberg

Carbohydrates play a key role in a wide range of biological processes and the detailed understanding of the principles of their recognition by carbohydrate-binding proteins is of particular interest for researchers. Biomimetic carbohydrate receptors [1-3] provide valuable model systems to study the underlying principles of carbohydrate-based molecular recognition events and their development is also strong motivated by the belief that such artificial carbohydrate-binding agents could be used for the detection and treatment of diseases. Although effective artificial receptors have been developed, the exact prediction of their binding strength and selectivity is still further away and it is hoped that combined theoretical and experimental studies will contribute significantly to the solution of this problem.

Our previous studies showed that receptors consisting of both a macrocyclic building block and flexible side-arms [4-6] exhibit strong selectivity towards β-D-glucoside and represent particularly interesting objects for systematic binding studies. The binding capabilities of these receptor molecules were determined by investigations in two-phase systems, such as liquid-liquid extractions of sugars from water into organic phase, and by studies in homogenous media, including ^{1}H NMR, fluorescence and microcalorimetric titrations.





We present the calculational study of the selectivity of macrocyclic carbohydrate receptors towards methyl β-D-glycosides, such as gluco- and galactopyranoside, using popular semiempirical quantum-chemical and quantum-chemical methods, combined with an analysis of the potential energy surface (PES). The PES of the glycosides was sampled using molecular dynamics and simulated annealing while the PES of the complexes of the macrocyclic receptors with the bound substrates was sampled using a simplified docking procedure. The selectivity towards the carbohydrates is then calculated at different levels of theory. Although the calculated selectivity indicates better binding of β-methylglucoside for most of these samples, the results elucidate the difficulties of modelling flexible molecules (including carbohydrates) and more studies have to be carried out.

[1] Mazik, M. *RSC Adv.* **2012**, *2*, 2630. [2] Davis, A. P. *Org. Biomol. Chem.* **2009**, *7*, 3629. [3] Mazik, M. *Chem. Soc. Rev.* **2009**, *38*, 935. [4] Lippe, J.; Mazik, M. *J. Org. Chem.* **2015**, *80*, 1427. [5] Lippe, J.; Mazik, M. *J. Org. Chem.* **2013**, *78*, 9013. [6] Amrhein, F.; Lippe, J.; Mazik, M. *Org. Biomol. Chem.* **2016**, *14*, 10648.

On the environmental effects of nanoplastics

Oldamur Hollóczki

Mulliken Center for Theoretical Chemistry, University of Bonn, Beringstr. 4+6, D-53115, Bonn, Germany

Throughout the last decade there has been an increasing scientific and public concern regarding the environmental effects of plastic wastes, especially that of smaller fragmentation products: micro- and nanoplastics. Nanoplastics, with a diameter < 100 nm, have been reported to enter the bodies and organs of animals and humans, but until very recently little was known about the extent of their actual effects. The difficulties of the corresponding research lie in the challenging detection of these small, mostly organic species in colloidal solutions of similarly sized biomolecules.

To tackle these problems, we pioneered the application of molecular modeling and theoretical chemistry for exploring nanoplastic-biomolecule interactions. We found that the presence of these particles may change the structure and dynamics of lipid bilayers, which can have severe effects on the environment and on human health as well [1]. Furthermore, we demonstrated that upon interacting with nanoplastics the secondary structure of proteins can be altered [2]. In the presence of plastic particles, β-loop-like tryptophan zippers show significant changes in the energy demand of cleaving the intramolecular hydrogen bonds that define the structure of these *de novo* proteins. Polyethylene nanoparticles apparently stabilize α-helices, which they also encompass through the rearrangement of the constituting polymer chains. Nylon, on the other hand, spontaneously changes α-helices into a more β-sheet like assembly. Since such changes are responsible for various prion diseases, and have been also related to Alzheimer's disease, these findings are particularly concerning.

Having seen these severely adverse effects it is clear that technological processes are necessary to be developed for the removal of such particles. Considering also the above demonstrated potential of molecular modeling in nanoplastic research, we aimed at designing extractants that can be used for this purpose. The first set of solvents in our focus were hydrophobic ionic liquids. Within these materials, the ions form micelles around the plastics in solution, but do not of disintegrate them into smaller particles. This feature gives ionic liquids a clear advantage over organic molecular solvents such as THF or toluene, in which the disintegration of nanoplastics produces smaller, and potentially more harmful plastic fragments. The interface of ionic liquids and water attracts the polar domain of the former species, since the adjacent water molecules interact stronger with these charged groups than with the non-polar moieties. One might imagine that this structure provides a significant barrier for the phase transfer of the plastic particle, however, as the plastic approaches the interface, the interfacial ions flip over to present their non-polar groups to the particle, allowing the extraction to occur without any activation energy.

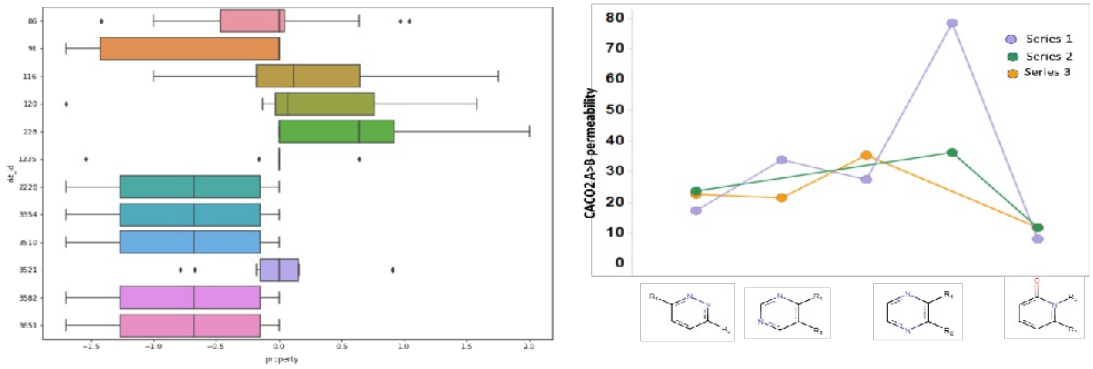
Since ionic liquids can be toxic, their use in drinking water and food treatment might be limited to a couple of compounds. For this reason, we extended the scope also to environmentally more benign liquids.

[1] O. Hollóczki, S. Gehrke, *ChemPhysChem* **2020**, *21*, 9-12.
[2] O. Hollóczki, S. Gehrke, *Sci. Rep.* **2019**, *9*, 16013.

Compound optimization - What we should know about the SAR trends we don't know

Gianna Pohl and Daniel Kuhn

Computational Chemistry & Biology, Merck Healthcare KGaA, Frankfurter Strasse 250, 64293 Darmstadt, Germany



Structure activity relationships (SAR) are important factors in drug development to understand and optimize the physiochemical and biological properties of molecules. An examination of the exchange of individual residual groups can provide information on how these groups influence the properties. Methods to investigate the influence of single residual groups are for instance Matched Molecular Pair (MMP) analysis [1] and Free-Wilson analysis [2]. To perform the necessary calculations of the analyses, there are already freely available programs, such as *mmpdb* [3] or the implementation of the Free-Wilson analysis by Pat Walters [4]. While performing MMP analysis is straightforward, these tools hardly provide any possibilities to visualize the results or to put them into context. This is where our work starts. We created several command line tools to facilitate data preparation, analysis calculations allowing for consideration of different endpoints. *Jupyter notebooks* and *voila* web applications were created to visualize the results in different ways. For example, the distributions of the influence of different transformations can be visualized as a box plot as shown in the figure above left. The results can also be exported to analysis tools such as *Spotfire* or *Jupyter notebooks* to allow analysis by broad scientific audience. These analysis tools allow that activity cliffs can be easily detected, and the influence of certain residual groups can be understood. To analyze trends between related chemical series Matched Molecular Series (MMS) analysis has been implemented. MMP and MMS analysis have been applied to ADME endpoints such as CACO2 permeability and microsomal stability. Analyzing related compound series helps to elucidate SAR trends and provides information how SAR can be transferred across series, e.g. changes in CACO2 permeability are dependent on MMS (see Figure above right). The developed methodology is applied in the optimization of ADME endpoints and examples from drug discovery projects will be shared.

[1] A. G. Dossetter, E. J. Griffen, A. G. Leach, *Drug Discovery Today, Matched molecular pair analysis in drug discovery*, **2013**, *18(15-16)*, 724-731
[2] S. M. Free, J. W. Wilson, *Journal of Medicinal Chemistry, A Mathematical Contribution to Structure-Activity Studies*, **1964**, *7(4)*, 395-399
[3] A. Dalke, J. Hert, C. Kramer, *Journal of chemical information and modeling, mmpdb: An open-source matched molecular pair platform for large multiproperty data sets*, **2018**, *58(5)*, 902-910.
[4] P. Walters, *Free Wilson Analysis*, **2018**, http://practicalcheminformatics.blogspot.com/2018/05/free-wilson-analysis.html

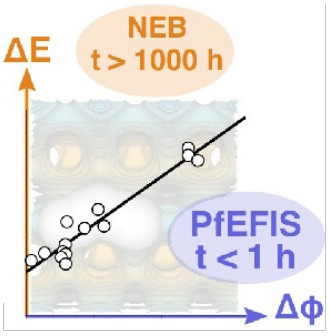
Potential of Electrostatics-Finite Ion Size (PfEFIS) method: towards automatic ion diffusion network analysis in solids

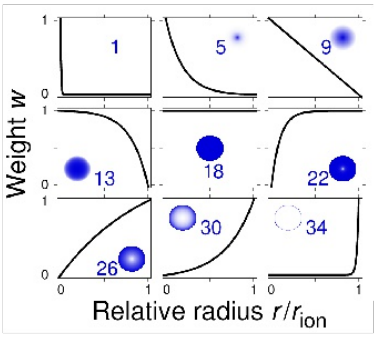
Nils E. R. Zimmermann

Lawrence Berkeley National Laboratory, Berkeley, CA, U. S. A.

We use more and more hand-held electronic devices and laptops, and, on our streets, there is a clear shift from predominantly gasoline-powered mobility towards a sizable share of electric mobility. Consequently, the demand for rechargeable battery materials constantly and swiftly increases. Finding new safe and ecological cathode materials for high-energy-density rechargeable batteries is thus an important present-day and future task for the scientific community at the interface between chemistry, physics, materials science, and engineering. Replacing lithium with magnesium as charge carrier could help addressing safety and environmental issues with current technologies. Because experimental screenings take a lot of time and are expensive, computational materials screenings based on efficient quantum mechanical calculations have become attractive alternatives. The assessment of the ionic mobility (i.e., the ion diffusion) is a bottleneck in this context due to the huge computational resources required by the conventionally employed density functional theory-nudged elastic band (DFT-NEB) calculations. For these reasons, we have developed a method, the potential of electrostatics-finite ion size (PfEFIS) barrier estimation method, which can rapidly estimate jump-diffusion barriers of ions in typical battery candidates [1]. The computational speed-up factor of approximately 10,000 (Figure left) thus permits targeting tens of thousands of materials to be screened for their ionic mobility.

The key idea of PfEFIS is to relax a chain of states in the field of the electrostatic potential that is averaged over a spherical volume using different finite size ion models (Figure right). For magnesium migrating in typical battery material candidates such as transition metal oxides, we find that the optimal model is a shell that is slightly larger than usual ion sizes found in literature [2]. This data-driven result parallels typical assumptions made in models based on Onsager's reaction field theory to quantitatively estimate electrostatic solvent effects. We believe that the method can potentially also be used to estimate proton and oxygen ion migration in solid materials, thus, facilitating computational screening studies for solid oxide fuel cells.





Work is in progress, which addresses another important challenge: automatically identifying insertion sites in the (initially empty) materials. This step would be executed prior to the PfEFIS barrier estimation in a high-throughput computational screening. Hence, reliable site identification is of particular importance in order to fully evaluate the entire diffusion network of ions in battery candidate materials in an automatic way. Preliminary results indicate that PfEFIS is also capable of tackling this issue, thus, paving the way to automatically handling ten thousands of materials in computational work flows for battery materials discovery.

[1] N. E. R. Zimmermann, D. C. Hannah, Z. Rong, M. Liu, G. Ceder, M. Haranczyk, K. A. Persson, *J. Phys. Chem. Lett.* **2018**, 9, 628–634.
[2] R. D. Shannon, *Acta Cryst.* **1976**, A32, 751–767.

Ligand Field Molecular Mechanics applied to MOFs

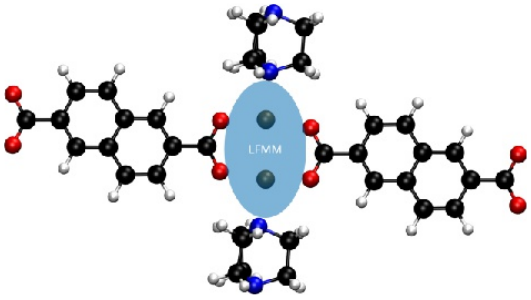
Patrick Melix, Thomas Heine

TU Dresden, Universität Leipzig

The observation of structural flexibility of certain Metal-Organic Frameworks (MOFs) triggered many theoretical investigations of these materials. The thermodynamics and kinetics of flexible MOFs can be difficult to model for various reasons: The role of the metal, conformational isomerism, dependence of the flexibility on the particle size and the role of dispersion interactions, to name only a few, are on their own already challenging for many methods. Combining these characteristics in one material challenges theoreticians to revisit their toolbox and combine tools and methods in new ways.

In our recent work [1] we apply Ligand Field Molecular Mechanics (LFMM) to perform molecular dynamics (MD) simulations of the pillared layer MOF $Ni_2(ndc)_2(dabco)$ (ndc=2,6-naphtalenedicarboxylate, dabco=1,4-diazabicyclo-[2,2,2]-octane), also known as DUT-8(Ni) [2]. This MOF shows pronounced stimuli-induced switching between a closed pore and an open pore phase upon ad- and desorption of certain gases and liquids accompanied by an enormous volume change of the unit cell [3]. The modelling of the flexible paddle wheel structure of this MOF is a very challenging task using standard density functional methods and even more so for classic force fields. By employing LFMM in combination with the general amber force field (GAFF) we modeled the loading of various guests (Nitrogen, carbon dioxide, dimethylether, methane, chlorinated methanes, methanol, ethanol and water) into the MOF to study the host-guest interactions and understand the characteristics of their adsorption inside the MOF.

LFMM is based on the calculation of the Ligand Field Stabilization Energy by use of the angular overlap model and is a well-established method that has been applied to the description of many transition metal complexes [4]. One of its distinct features is the applicability to model different spin states using only one set of parameters. Exploiting this, Paesani *et al.* introduced a hybrid Monte Carlo/MD method that allows switching of the spin state of each metal center individually during MD simulations [5].



Even though no spin state switching is experimentally observed in DUT-8(Ni) we hope that our investigation provokes others to also evaluate the applicability of LFMM and the concept of its use in the hybrid MC/MD scheme. Our goal is to extend our simulations to other flexible MOFs with dynamic spin state changes. A combination of LFMM and DFT simulations provides in our opinion a powerful toolbox for the description of many flexible MOFs, extending even to those with varying spin state.

[1] P. Melix, F. Paesani, T. Heine, *Adv. Theory Simul.*, **2019**, *2*, 1900098.
[2] N. Klein, C. Herzog, M. Sabo, I. Senkovska, J. Getzschmann, S. Paasch, M. R. Lohe, E. Brunner, S. Kaskel, *Phys. Chem. Chem. Phys.*, **2010**, *12*, 1177.
[3] V. Bon, N. Klein, I. Senkovska, A. Heerwig, J. Getzschmann, D. Wallacher, I. Zizak, M. Brzhezinskaya, U. Mueller, S. Kaskel, *Phys. Chem. Chem. Phys.*, **2015**, *17*, 17471.
[4] R. J. Deeth, A. Anastasi, C. Diedrich, K. Randell, *Coord. Chem. Rev.*, **2009**, *253*, 795.
[5] J. Cirera, V. Babin, F. Paesani, *Inorg. Chem.*, **2014**, *53*, 11020.

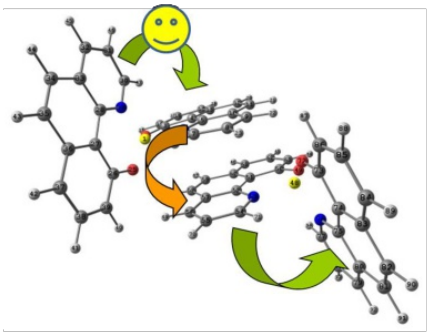
Multireference Study of Charge and Energy Transfer in Organic Semiconductors

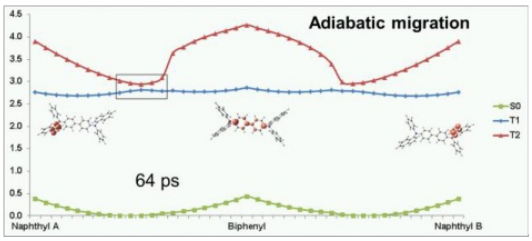
Alexandra Freidzon

[a] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)*

[b] *Photochemistry Centre of the Russian Academy of Sciences*

Transfer of charges and energy in organic semiconductor layers implies that the hopping quasiparticles (electrons, holes, or excitons) are localized within relatively small regions, such as small molecules or molecular fragments. However, band-like charge transport can be a key to high mobility in organic semiconductors. Charge and exciton migration is assisted by nuclear motion, and intra- and intermolecular vibronic interactions play an important role in the hopping process. Therefore, understanding how molecular structure and molecular packing affect charge and exciton localization and migration in organic semiconductors may help one in the design of efficient emissive and high-mobility transport layers. Multireference ab initio methods give a reliable tool for distinguishing between different localizations of charges or excitons, because they make it possible to treat equally important states on equal grounds. We propose a computational procedure based on CASSCF/XMCQDPT calculation to track the charge or exciton transfer path and to elucidate the role of vibrational modes in the hopping process. The examples of some typical emissive, electron- and hole-transporting materials are considered.





The work was financially supported by the Russian Science Foundation, grant no. 19-13-00383.

Reproducing high accuracy QM interactions via empirical force fields

Amin Alibakhshi

Department of Theoretical Chemistry, Christian-Albrechts University Kiel
(alibakhshi@pctc.uni-kiel.de)

Reliable and accurate evaluation of inter and intra-atomic interactions is one of the key steps in majority of molecular modeling approaches. Although these interactions can be precisely evaluated via high quality ab initio methods, the computational costs of such approaches specially for large systems has made applying empirical force fields still inevitable.

The empirical force fields typically fractionize the total energy of a molecule into smaller components such as bond stretching, angle and torsional bending and non-bonded interactions and evaluate each one of these energy terms via simple algebraic and empirical expressions. The conventional approaches of parametrizing such empirical force fields typically rely on transferability of force field parameters and determining those parameters such that they reproduce one or few bulk properties of the ensemble as accurate as possible. Nevertheless, remarkable inaccuracies inherent in such parameterization approaches reported in many recent studies has been the motivation of proposing next generation force fields e.g. machine learning based and ab initio derived force fields as well as more rigorous system specific parameterization strategies e.g. via the force matching technique [1].

In the present study we evaluated the performance of several conventional and more rigorous recently proposed empirical force fields in reproducing high accuracy QM interactions. As a benchmark, we carried out QM computations for 1500 configurations of boric acid and borate solvated in a small cluster of 64 water molecules. Each force field was parameterized by force matching technique to reproduce the reference QM forces. The studied configurations were extracted from molecular dynamics snapshots and selected such that they cover a wide combination of various force field variables. According to the results, none of the conventional force fields could reproduce QM forces with average unsigned relative error more accurate than 104%. Nevertheless, through a new empirical force field introduced in the present study which is a slightly modified version of a robust reactive force field recently developed in our group [2], we could achieve 8% accuracy in average unsigned relative error of evaluated forces.

1. Steffen, J. and B. Hartke, *Cheap but accurate calculation of chemical reaction rate constants from ab initio data, via system-specific, black-box force fields.* The Journal of chemical physics, 2017. **147**(16): p. 161701.
2. Hartke, B. and S. Grimme, *Reactive force fields made simple.* Physical Chemistry Chemical Physics, 2015. **17**(26): p. 16715-16718.

Understanding the substrate access mechanism in Phospholipase A from *Pseudomonas aeruginosa*

Sabahuddin Ahmad[1], Stephan Schott-Verdugo[1], Christoph Strunk[2], Filip Kovacic[2], Karl-Erich Jaeger[2], Holger Gohlke[1,3]

[1]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany and Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

[3]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich GmbH, Jülich, Germany

The Gram-negative bacterium *Pseudomonas aeruginosa* is an opportunistic pathogen and frequent cause of nosocomial infections, affecting primarily immune-compromised patients. PlaF from *P. aeruginosa* is a novel phospholipase A1 (PLA1). It is an integral inner membrane protein and has been shown to be a relevant virulence factor. A recent crystal structure of PlaF together with crosslinking and micro-scale thermophoresis experiments revealed PlaF to be in monomeric and dimeric configurations, although the protein was found to be active only in the monomeric state. Our previous free energy computations and wet-lab experiments indicated that, depending on the protein concentration, the dimer-to-monomer equilibrium is shifted to either side. The dimer-to-monomer transition is associated with a change in the orientation of the active site tunnel with respect to the membrane: A) in the dimeric form, the tunnel is *parallel* to the membrane *surface*, and the entrance resides more than 5 Å above it, which will likely hamper substrate access; B) in the monomeric form, PlaF tilts and the active site tunnel is in direct contact with the membrane, which will likely facilitate the substrate access from the membrane. We therefore hypothesize that the tilting of monomeric PlaF leads to this form being active by facilitating substrate access to the active site tunnel.

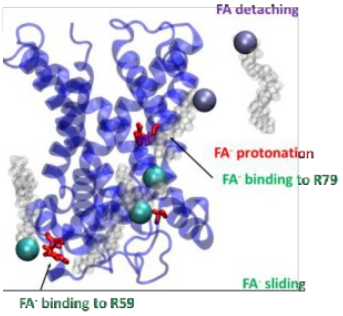
On further investigation, it was found that there are other cavities than the active site tunnel, resulting in total four access pathways (Tunnel 1-4) to/from the active site. However, which of these pathways is most favorable for substrate access remained elusive. To answer this, we performed configurational free energy computations of substrate access from the membrane to the active site. We considered different phospholipid substrates to understand this. Since a phospholipid substrate can enter by its head or either one of the two tails, we tested all three scenarios. Our computations reveal that A) Tunnel 2 is energetically most favorable for substrate access and B) the substrate entering with tail 1 first is the most favorable access mode. This access mode agrees with the fact that PlaF is a PLA1 that cleaves tail 1 of phospholipids. To validate our hypothesis about tunnel 2, we proposed tryptophan substitutions for all the four tunnels. Among proposed substitutions, the ones corresponding to tunnel 2 reduced the activity of PlaF by ~70%. Hence, it was confirmed that the tunnel 2 of Plaf is the preferred access tunnel for substrates. At present we aim to understand the egress pathway of the PlaF products. Together with our previous studies, our results can provide an explanation to the activity mechanism of PlaF at the atomistic level.

Insights into Mechanism of Proton Translocation Assisted by Membrane Proteins

Zlatko Brkljača, Sanja Škulj and Mario Vazdar

Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Zagreb, Croatia

Molecular dynamics simulations are a powerful tool for the investigation of complex systems in atomistic detail. In the last 20 years, these methods have been extensively used to study biological systems, such as DNA, proteins and model biological membranes. In this study we focus on the latter subject, particularly on the phenomenon of membrane permeation, which governs biological processes and underlying cell mechanisms. In this respect, while the neutral species can quite readily permeate across the cell bilayer, the unmediated transport of charged species is usually slow and does not occur on biologically relevant timescales. [1] The transport of such species is thus generally occurring via various supporting mechanisms which commonly involve membrane proteins, used to arbitrate and assist their passage through lipid bilayers. [2]



One of the most prominent membrane proteins located in the inner mitochondrial membrane is adenine nucleotide transporter 1 (ANT1). While its main function is transport of ADP and ATP nucleotides, ANT1 is also found to catalyze translocation of other charged species through inner mitochondrial membranes in the ATP independent manner. [3] In this work, we use classical MD simulations to obtain a detailed insight into the translocation of protons assisted by fatty acids across model bilayers with embedded ANT1 protein. In particular, we use both unbiased MD simulations and "state-of-the-art umbrella sampling along the pathway" technique to investigate this phenomenon. Our results led us to suggest a novel mechanism of proton transport assisted by fatty acid and catalyzed via ANT1 protein.

[1] S. Škulj, M. Vazdar, *Phys. Chem. Chem. Phys.*, **2019**, *21*, 10052-10060.
[2] N. P. Bethel, M. Grabe, *Proc. Natl. Acad. Sci. U. S. A.*, **2016**, *113*, 14049-14054.
[3] Y. Wang, E. Tajkhorshid, *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 9598-9603.

Modelling peptide aggregation influenced by *in vivo* conditions

Birgit Strodel

Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, 52425 Jülich, Germany

Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

One of the grand challenges of biophysical chemistry is to understand the principles that govern peptide aggregation, which is a highly complex process that is sensitive to initial conditions, operates on a huge range of length- and timescales, and has products that range from dimers to macroscopic amyloid fibrils. Aberrant aggregation is associated with more than 25 diseases, which include Alzheimer's, Parkinson's, Huntington's, and type II diabetes. Amyloid aggregation has been extensively studied *in vitro* under conditions that are far from the physiological ones. There is need to extend these investigations to *in vivo* conditions where amyloid formation is affected by a myriad of biochemical interactions. The aim of our simulation work is to elucidate the effects of various *in vivo* components and conditions, such as the presence of Cu^{2+}, oxidative stress, an acidic environment mimicking brain inflammation, but also the sheer existence of the neuronal cell membrane on the conformational dynamics and aggregation of the amyloid-β peptide (Aβ) that is strongly related to the development of Alzheimer's disease. To this end, we perform large-scale molecular dynamics simulations and develop novel simulation tools allowing us to unravel the amyloid aggregation pathways under varying external conditions. The most recent and enlightening results from our simulations will be presented in my talk.

Rapid Exploration of Synthetically Tractable Chemical Space for Hit-To-Lead & Lead Optimization

Nadine Homeyer, Thomas Steinbrecher

Schrödinger GmbH, Q7, 23, 68161 Mannheim, Germany

In the effort to find new drug molecules, vast amounts of compounds are screened *in silico* for potential drug candidates. However, despite the huge number of molecules in the compound libraries used for these computational analyses, they constitute only a very tiny fraction of the chemical space of all drug-like compounds [1, 2]. Due to the enormous size of the drug-like chemical space consisting of up to more than 10^{60} molecules [1, 3], it is highly desirable to focus *in silico* analyses on cleverly selected, structurally diverse, synthesizable drug-like compounds. Especially in the hit-to-lead and lead optimization stages of the drug design process, it would be beneficial to be able to rapidly ideate synthetically tractable, chemically diverse compounds based on the structure of a hit or lead molecule.

Schrödinger's Pathfinder module for reaction based enumeration can, based on a provided "target" molecule, easily generate data sets of sensible and synthetically accessible compounds for computational drug design campaigns. Reaction based enumeration combines rapid enumeration with the functionality of Pathfinder identifying possible routes to desired targets. It is a synthetically aware enumeration tool enabling the user to prepare large compound libraries with significant diversity for the exploration of different R-groups and cores in the later optimization phases of the drug design process.

A study in which reaction based enumeration was combined with docking, multi-parameter optimization, FEP+ binding free energy calculations and active learning [4] showed that this enumeration tool can be useful for exploring a larger chemical space than conventional Structure Activity Relationship (SAR) studies and for identifying synthetically accessible compounds with better properties and potency. Therefore, this new enumeration tool may facilitate the rapid discovery of new drug compounds from the so far unexplored chemical space.

[1] R. S. Bohacek, C. McMartin, W. C. Guida, *Medicinal research reviews*, **1996**, 16, 3-50.
[2] T. Hoffmann, M. Gastreich, *Drug Discovery Today,* **2019**, *24*, 1148-1156.
[3] P. Kirkpatrick, C. Ellis, *Nature*, **2004**, 432, 824-828.
[4] K. D. Konze, P. H. Bos, M. K. Dahlgren, K. Leswing, I. Tubert-Brohman, A. Bortolato, B. Robbason, R. Abel, S. Bhat, *J. Chem. Inf. Model*, **2019**, 59, 3782-3793.

Kinematic Flexibility Analysis of Kinase Activation Loops

Xiyu Chen[1], Sigrid Leyendecker[1], and Henry van den Bedem[2,3]

[1]*Chair of Applied Dynamics, University of Erlangen-Nuremberg*

[2]*Atomwise, Inc. 717 Market Street, San Francisco, CA 94103 USA*

[3]*Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, CA 94158 USA*

The macromolecules function relies on fast exchanges between three-dimensional conformations. We use Kino-geometric sampling (KGS) method and model the molecular framework to study functional molecular rigidity analysis and to investigate molecular transitions [1,2]. KGS represents molecules as multi-body complexes with dihedral angles as revolute degrees of freedom. In addition, non-covalent interactions (e.g. hydrogen bonds and hydrophobic interaction) are chosen as holonomic constraints, a lower-dimensional manifold in conformation space is defined by these constraints [3,4]. Protein kinases are cellular enzymes that catalyze phosphorylation reaction and they are also an important method for drug research in cancer. Inactive or activated forms of kinases are often classified into specific conformations based on their activation loops and surrounding areas [5]. The inhibitors are directed to the active or several inactive conformations, so the activation classification and the rigidity analysis of kinases may help the drug development and the general understanding of kinases conformation. This information is valuable for developing potent and selective kinase inhibitors.

In this work, KGS method is applied to analyze kinematic flexibility of macromolecular and to classify the large dataset of kinases into different cluster groups according to sequence similarity and structure similarity. The differences of degree of freedom of rigidification and the largest rigidified cluster size between the active and the inactive state kinases are researched based on the KGS rigidity analysis. It shows that the activation loop rigidified the surrounding α helix and make the rigidification of active state kinases increases.

[1] D. Budday, S. Leyendecker, and H. van den Bedem. *Journal of Chemical Information and Modeling,* **2018**, *58(10)*, 2108-2122

[2] R. Fonseca, D. Budday, and H. van den Bedem. *Journal of Computational Chemistry*, **2018**, *39*, 711-720

[3] D. Budday, R. Fonseca, S. Leyendecker, and H. van den Bedem. Proteins: *Structure, Function, and Bioinformatics,* **2017**, *85*, 1795-1807

[4] D. Budday, S. Leyendecker, and H. van den Bedem. *Journal of the Mechanics and Physics of Solids*, **2015**, *83*, 36-47.

[5] M. Vivek, L. Roland, and J. Dunbrack. *National Academy of Sciences*, **2019**, *116,* 6818-6827

Modelling biological data with experimental restraints

Thorn, Andrea

Institute of Structural Biology, Rudolf Virchow Center for Experimental Biomedicine
University of Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, Germany
E-mail: andrea.thorn@web.de

Macromolecular modelling problems often utilize experimental structures of biological macromolecules both as their basis and for validation of results. For over 50 years, X-ray crystallography has been the primary method to determine these structures. Recently, electron cryo microscopy (Cryo-EM) has also started to significantly contribute to the wealth of almost 150 000 experimentally determined structures.

However, these experimental methods have limitations: A model is strictly needed for the interpretation of experimental data [1], and has to be optimized to fit (potentially flawed) measured data. As a consequence, some biological questions - for example whether a ligand is bound - cannot be answered, some structures, such as large complexes or membrane proteins, cannot be solved at all and, worst of all, published and seemingly correct structure solutions can have flaws that might even lead to a retraction. In addition, the limitations of experimental structure solution directly affect computational modelling as a downstream method.

In this talk, I will describe these limitations and their implications in terms of both the data quantity [2] as well as the employed models. I will show how modelling can be improved (giving a previously unprecedented way to restrain the movement of bonded atoms as an example [3]) and describe how we ultimately realized that something might be fundamentally wrong with the atomic models which we employ to interpret experimental data.

[1] Holton, J.M., Classen, S., Frankel, K.A., Tainer, J.A., *FEBS*, **2014**, *281*, 4046–406
[2] Thorn, A., Parkhurst, J.M., Emsley, P., Nicholls, R., Vollmar, M., Evans, G. & Murshudov, G.N. (2017) *Acta Cryst. D,* **2017**, *73*, 729-737.
[3]Thorn, A., Dittrich, B., Sheldrick, G.M., *Acta Cryst. A*, **2012**, *68*, 448-451

Prediction of Odour (PrOdour) – Data Exploration

Slavica Subic[1], Veniamin Morgenshtern[2], Thilo Bauer[3, 4]

[1]Advanced Signal Processing and Communications Engineering, Elite Master's Study Programme, FAU Erlangen-Nürnberg
[2]Chair of Multimedia Communications and Signal Processing, Dept. Electrical, Electronic and Communication Engineering, FAU Erlangen-Nürnberg
[3]Computer Chemistry Centre (CCC), Department of Chemistry and Pharmacy, FAU Erlangen-Nürnberg
[4]Fraunhofer Institute for Process Engineering and Packaging IVV

Where exactly does the odour come from? What causes a substance to produce a certain scent?

One could approach answering this question in many ways, and one way would be to look at the language used to represent the molecules. Structural formulae are exactly that, and that is where this data exploration task starts. Looking at the language of chemistry, and individual words in its dictionary, we search for the ones most responsible for particular odours.

Starting from molecular formulae, parsing them into millions of words (features, sub-molecular structures), and looking for largest overlapping structures, we collect a dictionary of molecular features. With data collected from experiments where experts evaluated the sensory properties of substances, it was possible to build a data set containing the applicability of a list of odour descriptors to the substances used in the experiment. Using LASSO (Least Absolute Shrinkage and Selection Operator) regression regularization for feature extraction, we aim to find the features most responsible for the applicability scores by which descriptors were rated in the experiment - in other words, features most responsible for the expression of a certain odour. By grouping descriptors with clustering techniques, we can explore similarities of features among the groups and search for chemically meaningful features that define a descriptor cluster.

[1] A. Cadeddu, E. K. Wylie, J. Jurczak, M. Wampler-Doty, and B. A. Grzybowski - Organic Chemistry as a Language and the Implications of Chemical Linguistics for Structural and Retrosynthetic Analyses, *Angew. Chem.* **2014**, 8246-8250, DOI: 10.1002/ange.201403708
[2] T. Hastie, R. Tibshirani, J. Friedman – The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, Second Edition, corrected 12th printing - 13. Jan. **2017**.

Sampling of Ligand-Induced Conformational Changes in Renin and Factor VIIa

Laura Schulz,[†] Ahmad Reza Mehdipour,[†] and Gerhard Hummer[†,‡]

[†]*Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von Laue Str. 3, 60438 Frankfurt am Main, Germany*

[‡]*Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*

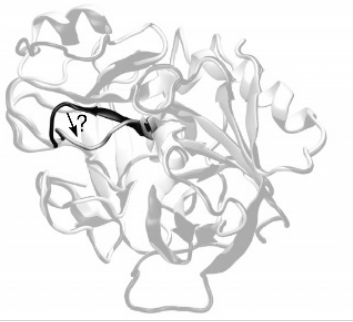
Proteins are dynamic biomacromolecules that can have diverse and nearly isoenergetic conformational states. Ligand binding can shift the equilibrium of this conformational ensemble.

We are studying protein conformational changes upon ligand binding in two different systems: Factor VIIa and Renin. In both systems we know from experiment that a specific group of ligands targeting the orthosteric site can induce conformational changes. The observed conformational changes are the distortion of a beta-sheet structure in the S1 pocket and the opening of a non-functional flap for Factor VIIa and Renin, respectively.

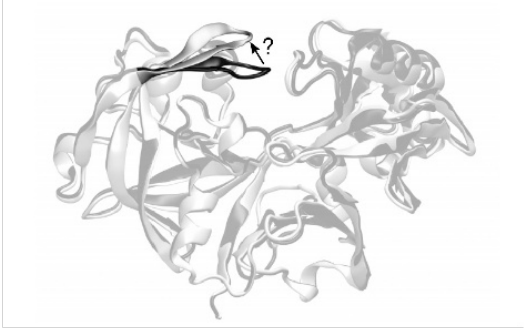
Unbiased MD simulations for Factor VIIa and for Renin (each 4 µs) with removed ligand structures led to contrasting findings. For Factor VIIa the conformational change can be captured with unbiased MD simulations, whereas the transition could not be sampled for Renin within this time scale.

For Factor VIIa starting from the initial beta sheet conformation one or both backbone hydrogen bonds were disrupted leading to the loop-like conformation. Starting with the loop-like conformation the distorted S1 pocket remained distorted.

For Renin starting from the initial closed flap conformation the flap did not open up. Using established biasing protocols for Renin such as increasing the water-protein interaction as established in the group of Gervasio [1] enhanced the protein flexibility, but did not sample the fully open flap. We are working on alternative approaches to disrupt the hydrophobic aromatic cluster holding the Renin flap in place. Starting from the open flap conformation, the open flap of Renin closed again for all eight available X-ray structures.



Factor VIIa:
stable (black) vs. distorted (grey) S1 pocket



Renin:
closed (black) vs. open (grey) flap

[1] Oleinikovas et al., J. Am. Chem. Soc. 2016

Dynamics of crowded macromolecules from atomistic simulations

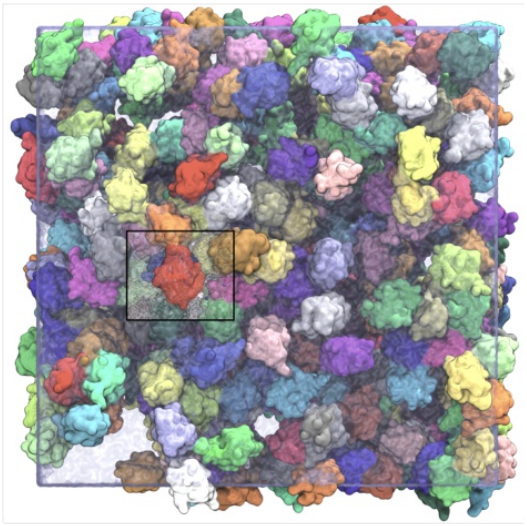
Sören v. Bülow,[†] Marc Siggel,[†] Max Linke[†] and Gerhard Hummer[†,‡]

[†]*Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue-Str. 3, 60438 Frankfurt am Main, Germany*

[‡]*Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*

The interior of the cell is densely crowded, affecting thermodynamic and kinetic properties of the macromolecules within. It has only very recently become computationally feasible to simulate crowded macromolecular systems atomistically to address detailed questions on protein stability, interactions and transport properties like diffusion.

I present a large-scale molecular dynamics simulation study of atomistically-resolved crowded protein systems containing up to 540 fully flexible proteins with 3.6 million atoms in the microsecond range. [1] We find that using the Amber99SB*-ILDN-Q protein force-field in conjunction with the TIP4P-D water model gives a surprisingly accurate picture of the dynamics of concentrated protein solutions that agrees very well with experimental results.



Snapshot of an atomistic molecular dynamics simulation of 540 fully flexible ubiquitin proteins at 200 mg/ml concentration in explicit water.

The protein species studied here form dynamic clusters between which they constantly exchange. A theoretical model, based on the Stokes-Einstein equations, nearly quantitatively links the slow-down of translational and rotational diffusion, increase in viscosity and formation of protein clusters.

Using a colloidal model of sticky hard spheres, we relate the above properties to the dissociation constant of protein-protein interaction, allowing for an estimate of the concentration-dependent slow-down of protein diffusion given only the dissociation constant.

We build on the above framework to simulate mixed systems to address research questions like the postulated solubilizing properties of ATP (acting as a "hydrotrope"), phase-separation of intrinsically disordered proteins, or interactions between RNA and transcription factors.

Reference:

[1] Sören von Bülow, Marc Siggel, Max Linke, Gerhard Hummer (2019) Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proc Natl Acad Sci* 116:9843-9852.

Enhancing conformational sampling with machine learning and the cloud

Charles Laughton

Centre for Biomolecular Sciences, University of Nottingham
University Park, Nottingham, NG7 2RD, UK

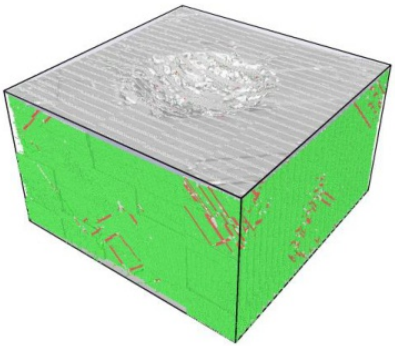
I will describe some of our recent research developing new methods for the enhanced sampling of the conformational space of biomolecules, both large and small. I will concentrate on methods that combine molecular dynamics-based sampling with machine learning based supervision in iterative, adaptive, workflows. I will also describe the cloud-based infrastructure and workflow tools we have developed to support this type of research, which we are now making generally available.

Atomistic Modelling of High Entropy Alloys

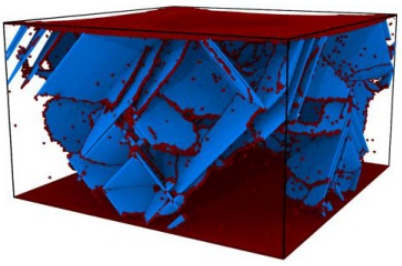
Akinjide Oluwajobi, Jan Ocenacek

New Technologies - Research Centre, University of West Bohemia, 30100 Pilsen, Czech Republic

High Entropy Alloys (HEAs) exhibit exceptional physical and mechanical properties for potential applications in many industrial sectors. To efficiently explore these multicomponent alloys, fundamental understanding of their deformation behaviour is essential.



Indentation of CoAlCuFeNi



Defect networks during indentation

As example, the Molecular Dynamics (MD) simulations of nanoindentation of CoAlCuFeNi was conducted, by using Embedded Atom Method (EAM) interatomic potential. These simulations reveal the nanoscale mechanisms of plasticity, critical to their excellent properties. The nanoidentation investigation shows dislocation evolution, cross-slip and twinning as the major mechanisms of plastic deformation.

Effect of chalcogendiazole groups on the conductive properties of EDOT containing polymers

Özge Özkılınç, Hakan Kayı*

Computational Chemical Engineering Laboratory, Chemical Engineering Department, Ankara University, Tandoğan, 06100 Ankara, Turkey

E-Mail: ozkilinc@ankara.edu.tr, *hakan.kayi@ankara.edu.tr

Tuning of the π-conjugated systems' backbone for the desired application is possible through modification of donor group as electron-rich or acceptor group as electron-deficient one. Electronic nature can be accomplished for the desired applications via changing the chalcogen atoms. In this work, impact of chalcogen atom substitution on electrochemical properties of semiconducting polymers was studied with density functional theory (DFT). A series of donor-acceptor-donor (DAD) systems with EDOT as an electron donor unit and benzoxadiazole, benzothiadiazole, benzoselenadiazole, benzotelluradiazole as electron acceptor units were investigated theoretically. Quantum chemical calculations are carried to calculate the HOMO-LUMO levels. For all conformations, B3LYP method was employed. Also, LANL2DZ and 6-31G(d) were used as basis sets. Obtained results indicate that the narrowest electronic band gap is achieved by introducing a heavy chalcogen atom, such as tellurium [1], into the donor and acceptor units of the semiconducting polymer. In the end, bandgap results are compared with the experimental findings in literature.

[1] Ozkilinc O, Kayi H, *Journal of Molecular Modeling*, **2019**, 25:167.

Acknowledgements: Authors thank Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting this study through the project number 117Z354.

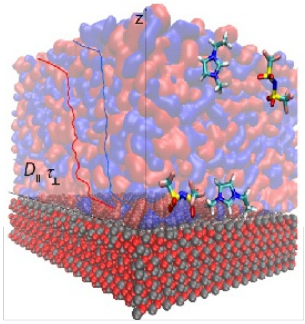
Imidazolium-based Ionic Liquids under the Magnifier of Molecular Simulations

Nataša Vučemilović-Alagić,[a,b] Kevin Höllring,[b] Christian R. Wick,[b]

David M. Smith*,[a] Ana-Sunčana Smith*,[a,b]

[a]*Group of Computational Life Sciences, Department of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia*

[b]*PULS Group, FAU Erlangen-Nürnberg, Cauerstraße 3, 91058, Erlangen, Germany*



With the aim of understanding the structural organization and dynamic properties related to Supported Ionic Liquid Phase (SILP) catalysts, we present an in-depth study of the [C_2Mim][NTf_2] ionic liquid (IL). Using classical molecular dynamics (MD) simulations, the IL is considered in three configurations, corresponding to the pure liquid (L), vacuum-liquid-vacuum (V-L-V) and solid-liquid-vacuum (S-L-V) arrangements. In the first part of the study, by combining charges derived from CHelpG, RESP-hf and RESP-b3lyp with charge scaling factors of 1.0, 0.9 and 0.8, we search for an optimum force field (ff) parameterization. This is achieved by benchmarking the static and dynamic responses of the three model systems against experimental data such as self-diffusion coefficients (L), surface tension (V-L-V) and X-ray reflectivity (S-L-V). [1] Consequently, we provide detailed atomic descriptions of the IL at the solid and vacuum interfaces, which are in very good agreement with experimental data such as low-energy ion scattering experiments. [2] We further leverage the extensive data generated to provide evaluation of lateral diffusion coefficients, in the xy plane, and residence times along the z-direction.

[1] N. Vučemilović-Alagić, R. D. Banhatti, R. Stepić, C. R. Wick, D. Berger, M. U. Gaimann, J. Harting, D. M. Smith, A. S. Smith, *J. Colloid Interface Sci.*, **2019**, *553*, 350-363.
[2] N. Vučemilović-Alagić, R. D. Banhatti, R. Stepić, C. R. Wick, D. Berger, M. U. Gaimann, J. Harting, D. M. Smith, A. S. Smith *Data in brief*, **2020**, *28*, 104794.

Bottom-up DFT-MD multi-scale modelling of hydrated bulk anion exchange membranes

J. Luque Di Salvo[1], G. De Luca[2], A. Cipollina[1], G. Micale[1]

[1]*Dipartimento di Ingegneria (DI), Università degli Studi di Palermo– viale delle Scienze Ed.6, 90128 Palermo (PA), Italy.*

[2]*Institute on Membrane Technology, ITM-CNR, Via P. Bucci 17/C, 87036 Rende (CS), Italy.*

Ion Exchange Membranes (IEMs) are key components of various separation technologies in the fields of water treatment, chemicals synthesis and recovery, and energy generation. Generally, IEMs consist in a polymer backbone functionalized with fixed charged groups. When the membrane is put in contact with an electrolytic solution, due to Donnan equilibrium and Donnan exclusion, hydrated ions showing opposite charge with respect to functional charged groups (counter-ions) are selectively adsorbed whereas ions with the same charge (co-ions) are ideally rejected, leading to ion selectivity [1]. The fixed charged groups and counter-ions in the membrane adsorb water leading to a nanophase segregation inside the bulk of the membrane where hydrophobic and hydrophilic domains co-exist (polymer backbone and adsorbed water-solution). Their relative sizes and distribution strongly depend on: membrane materials; Ion Exchange Capacity (IEC, concentration of fixed charged groups); water volume fraction; fixed charged groups and ions' hydration, which have a crucial effect on ion diffusion. Usually, experimental techniques find it difficult to unravel the key structure-kinetic properties relationships, an aspect that can be approached by *ad-hoc* quantum and atomistic models.
In this work, a bottom-up multi-scale methodology to model a series of bulk anion exchange membranes is presented, based on recent published results [2]. First, Molecular Mechanics (MM) of representative polymeric units was used to perform a conformational search of the most stable monomer conformers. Second, *ab initio* calculations in the frame of Density Functional Theory (DFT) of increasingly hydrated functionalized monomers were carried out to define the water uptake per functionalized and non-functionalized monomers, and to parametrize the partial charges later used in the Molecular Dynamics (MD) force field being important parameters controlling morphology and transport features. Third, monomer units were randomly assembled into large polymeric chains to perform MD simulations. Once the simulation boxes were equilibrated through a calibrated simulated annealing protocol, a morphology analysis was performed based on radial distribution functions, cluster analysis of hydrophilic domains as well as 2D density maps to study the percolation degree of the systems. Finally, counter-ion diffusion coefficients were determined at variating IEC and water uptake analyzing the Mean Square Displacement (MSD) curves of counter-ions and by analytical obstruction models based on the water volume fraction inside the membrane. The results refer to polysulfone-tetramethylammonium (PSU-TMA) anion exchange membranes counter-balanced with hydroxide ion [2], and of preliminary results on hydrated PSU-TMA-Cl and polystyrene-TMA-Cl membranes. The results highlight how the *ab initio* methodology for studying the structure-kinetic relationships of the anion exchange membranes is a powerful investigation tool. Moreover, thanks to the minimization of adjustable parameters, commonly adopted in IEMs modelling, it gives general outcomes and can be easily extended to any type of Ion Exchange Membrane material.

[1] T. Luo, S. Abdu, M. Wessling, Selectivity of ion exchange membranes: A review, J. Memb. Sci. 555 (2018) 429–454. https://doi:10.1016/j.memsci.2018.03.051.
[2] J. Luque Di Salvo, G. De Luca, A. Cipollina, G. Micale, Effect of ion exchange capacity and water uptake on hydroxide transport in PSU-TMA membranes: A DFT and molecular dynamics study *J. Memb. Sci.* (**2020**), https://doi.org/10.1016/j.memsci.2020.117837.

Adsorption of Ions at Solid-Liquid Interface: Role of Ion Hydration Structure and Energies

Sai Adapa and Ateeque Malani

Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, 400076, India.

The adsorption of ions on charged surface creates distribution of counter- and co-ions from surface to the bulk region, which is referred to as electrical double layer (EDL). In EDL, some of the ions are in direct association with the surface while rest of the ions are distributed away from the surface till bulk region. The recently developed super-capacitors[1,2] physically store high density charge in terms of these EDL at the electrode surface. In addition, ion adsorptions at solid-liquid interface play crucial role in remediation of heavy metal ions, water purification by ion-exchange, minerals formation and surface catalytic reactions.

The adsorption of ions at the solid-liquid interface is governed by their distribution, hydration structure and hydration energies. Here, we have performed molecular dynamics simulation studies of monovalent (Na^+ and Rb^+) and divalent (Mg^{2+} and Sr^{2+}) ions at various concentrations at mica surface. These ions have varying hydration energies and hydartrion structure, and our simulation studies cover wide ranges of concentrations. Since, divalent ions have higher hydration energies, we expected their adsorption to be higher compared to monovalent ions having lesser hydration energies. However, we observed counter-intuitive behaviour of higher monovalent adsorption at all concentrations. We have analysed the potential of mean force of these ions, their coverage and hydration structure near mica surface to explain these observed behaviors.[3]

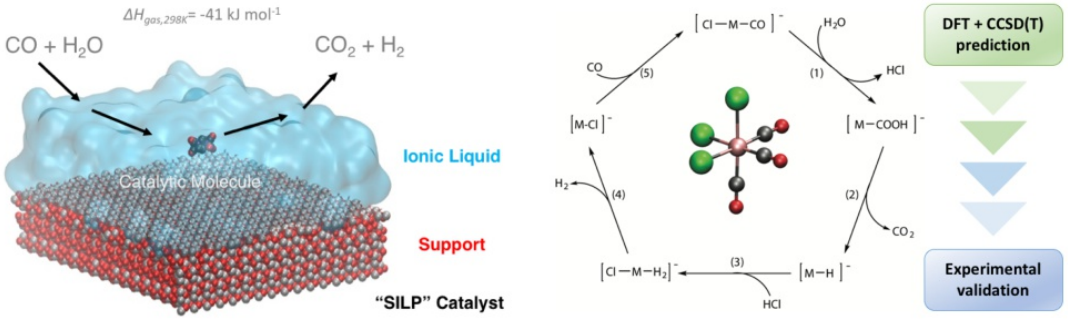
References:

[1] J. R. Miller and P. Simon, *Science*, **2008**, 321, 651-652.
[2] Miller et al., *Science*, **2010**, 329, 1637-1639.
[3] Adapa and Malani, *Scientific Reports*, **2018**, 8, 12198.

Modelling SILP Catalysis: The Water-Gas-Shift Reaction

R. Stepić,[1] C. R. Wick,[1,2] N. Vučemilović-Alagić,[1,3] D. M. Smith,[1,3] A.-S. Smith[1,3]

[1]*PULS Group, Institute of Theoretical Physics, FAU Erlangen-Nürnberg, Germany*
[2]*Zentralinstitut für Scientific Computing (ZISC), FAU Erlangen-Nürnberg, Germany*
[3]*Division of Physical Chemistry, Ruđer Bošković Institute, Zagreb, Croatia*



Supported ionic liquid phase (SILP) catalysts facilitate a highly efficient, homogenously catalyzed water-gas shift reaction (WGSR) at ultra-low temperatures between 120 °C and 150 °C. [1,2] SILP catalysts consist of an ionic liquid (IL) covering a porous support, while the actual catalytic molecule is dissolved in the IL phase. Thus, the SILP catalysts allow fine tuning of the actual catalytic reaction as well as the reaction conditions *via* variation of the support, the IL, the molecular catalysts and addition of additives.

Ru-based transition metal complexes have been identified as one of the most promising molecular catalysts in SILP conditions for the water-gas shift reaction (WGS). The active Ru-complexes have been found to exist in imidazolium chloride melts under operating conditions in a dynamic equilibrium, which is dominated by the $[Ru(CO)_3Cl_3]^-$ complex.[3] A detailed theoretical and experimental investigation showed that the reaction mechanism for this type of catalytic species includes the intermediate formation and degradation of hydrogen chloride, which effectively reduces the high barrier for the formation of the requisite dihydrogen complex. The hypothesis that the rate-limiting step involves water is supported by using D_2O in continuous catalytic WGS experiments. [4]

In addition to the elucidation of the reaction mechanism in SILP catalysis, we investigated the effect of CuCl additives with respect to the WGS reaction. Again, density functional theory (DFT) calculations and in situ diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) measurements have been used to identify the role of CuCl as CO shuttle, which leads to an enhancement of the catalytic performance. [5,6]

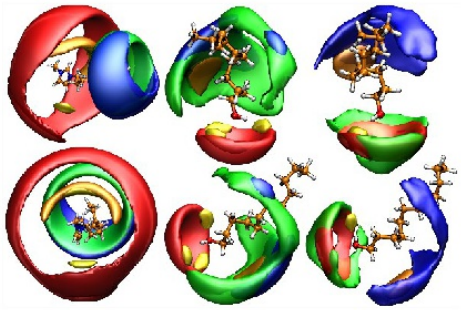
[1] C. Ratnasamy, J.P. Wagner *Cat. Rev.: Sci. Eng.*, **2009**, 51, 325-440
[2] S. Werner, N. Szesni, M. Kaiser, R. W. Fischer, M. Haumann, P. Wasserscheid, *ChemCatChem* **2010**, *2*, 1399–1402.
[3] T. Bauer, R. Stepic, P. Wolf, F. Kollhoff, W. Karawacka, C. R. Wick, M. Haumann, P. Wasserscheid, D. M. Smith, A.-S. Smith, et al., *Catal. Sci. Technol.* **2018**, *8*, 344–357.
[4] R. Stepić, C. R. Wick, V. Strobel, D. Berger, N. Vučemilović-Alagić, M. Haumann, P. Wasserscheid, A.-S. Smith, D. M. Smith, *Angew. Chem., Int. Ed. Engl.* **2019**, *58*, 741–745.
[5] P. Wolf, M. Aubermann, M. Wolf, T. Bauer, D. Blaumeiser, R. Stepic, C. R. Wick, D. M. Smith, A.-S. Smith, P. Wasserscheid, et al., *Green Chem.* **2019**, *21*, 5008–5018.
[6] D. Blaumeiser, R. Stepić, P. Wolf, C. R. Wick, M. Haumann, P. Wasserscheid, D. M. Smith, A.-S. Smith, T. Bauer, J. Libuda, *Catalysis Science & Technology* **2020**, *10*, 252–262.

Insights into ionic and molecular liquids and their interfaces from computational chemistry.

Barbara Kirchner

Mulliken Center for Theoretical Chemistry, Bonn, Germany
E-mail: kirchner@thch.uni-bonn.de

Ionic liquids (ILs) [1] are playing a role in many exciting applications as and for materials.[2] They are used in so different areas as in electrochemistry, catalysis, in coal processing, in pharmaceutical applications, to name but a few.[2] Understanding which one of all the possibilities is the right ionic liquid to use is the key to the successful outcome of the particular applications. Therefore, an understanding of the molecular level behavior, i.e., the vast molecular cosmos, [3] of the given liquid itself is desirable if the full potential of the solvent should be reached.



For example, the structure-directing or template effect [4] has been invoked several times for ILs to explain the different outcome in material synthesis when varying the IL. We showed the successful imprinting of the ILs' order in the alignment of alcohol molecules [5] in simple model systems (Fig.), which is governed by the micro-heterogeneous [6] structure of ILs. This imprinting leads to the ordering of reaction substrates in a predetermined fashion, opening new possibilities for explaining or enhancing the selectivity in chemical reactions in ILs.

In electrolyte systems, the association of the ions to form ion pairs or other, low charge aggregates is a long discussed issue,[7] since it can affect the manner and the extent of conduction through changing the number of mobile charged species in the solution. Structural diffusion in ILs [8] can be connected to electrochemical applications. Structural or Grotthuss diffusion is the extremely fast diffusion of molecular groups or protons. In particular, the proton transfer ability of solvents is a highly important feature for electrolytes and for solvents in synthesis as well. As a further example of exploring the molecular cosmos we will discuss chirality and its detection,[9] which plays an important role for the understanding of pharmaceuticals in ILs or pharmaceutical ILs. [10]

References

[1] P. Wasserscheid, P. and Welton T., eds. Ionic liquids in synthesis. John Wiley & Sons, 2008.
[2] Balducci, A. "Ionic liquids in lithium-ion batteries." Springer, Cham, 2017. 1-27; Lynden-Bell, R. M. et al,. PCCP 14.8 (2012): 2693; Welton, T. Coord. Chem. Rev. 248.21-24 (2004): 2459; Bica, K., et al. Chem. Comm. 48.41 (2012): 5013; F. Cui et al. Fuel 217 (2018) 508; Hough, W. L. et al. Bull. Chem. Soc. Jap. 80.12 (2007): 2262; F., R. Branco, et al. (2011), ChemMedChem, 6: 975
[3] Kirchner, B., et al. WIREs 5.2 (2015): 202. Schröder, C., J. Chem. Phys. 135.2 (2011): 024502; Pádua, AAH, et. al *Acc. Chem. Res.* 40.11 (2007): 1087; E. Izgorodina, Chem. Rev. 117(10), 2017; V. Ivaništšev et al. *J. Phys. Chem. C*, 2014, 118, 11, 5841; Zahn, S. and Cybik, R.; Am. J. Nano Res. App. 2014, 2, 19
[4] Cooper, E. R., et al. Nature 430.7003 (2004): 1012.
[5] Elfgen, R. et al. Acc. Chem. Res. 50.12 (2017): 2949
[6] Canongia Lopes, JNA and Pádua AAH; J. Phys. Chem. B 110.7 (2006): 3330.
[7] T. Cremer, et al., Chem. Eur. J., (2010), 16, 9018; N. Taccardi, et al., Chem. Eur. J. 2012, 18, 8288;B. Kirchner et al., J. Phys.: Condens. Matter, (2015), 27, 463002; O. Hollóczki, et al., PCCP. (2014), 16, 16880
[8] Del Pópolo, M. G. et al. J. Phys. Chem. B 110.17 (2006): 8798; B. Kirchner and A. P. Seitsonen, Inorg. Chem., (2007), 46, 2751; R. Elfgen, et al., Z. Anorg. Allg. Chem. (2017) 643, 41; J. Ingenmey et al. JCP (2018), 148:19; J. Ingenmey et al., ChemSusChem (2018), 11, 1900; D. S. Firaha et al. , J. Chem. Phys. (2016) 145, 204502, O Hollóczki, et al. Chem-Eur. J. 24 (61), 16193
[9] M. Thomas and B. Kirchner J. Phys. Chem. Lett. (2016), 7, 509
[10] Vasiloiu, M., et al Cat. today 200 (2013): 80

Histamine – a signalling cue influenced by fluctuating environmental conditions and climate change?

Christina C. Roggatz[1] & David M. Benoit[2]

[1]Energy & Environment Institute, and [2]Department of Physics and Mathematics, E.A. Milne Centre for Astrophysics & G.W. Gray Centre for Advanced Materials; University of Hull, Cottingham Road, Hull, HU6 7RX, UK

Histamine, an enzymatic decarboxylation-product of the amino acid histidine and well-known for its involvement in numerous physiological processes within the human body, also serves as an external signalling cue for invertebrates. When functioning as a settlement cue for sea urchins [1] or as a foraging cue for freshwater snails [2], histamine is exposed to a range of environmental factors, such as temperature, pH and salinity. These physical conditions are subject to large fluctuations and are projected to be influenced significantly by climate change over the course of this century. In this study, we assess the sensitivity of histamine and its essential properties to current and future environmentally relevant ranges of pH, salinity and temperature.

Employing a combination of Molecular Dynamics and static DFT geometry optimisation methods, we investigate if protonation caused by a change in pH affects the molecule's conformation, charge distribution and dipole as well as interaction with surrounding ions. Multiple MD simulations were performed with parameters representing either very low or high levels of salinity equivalent to fresh or salt water and different temperatures to simulate natural temperature ranges in coastal and riverine ecosystems. We established that changes to the protonation state of histamine are small for conditions relevant in marine ecosystems, while for freshwater systems there are two relevant protonation states that dominate alternately. The two forms differ significantly in their conformation and charge distribution as well as in their interaction with surrounding Cl^- and Na^+ ions. Charge as well as conformation are key elements in receptor-ligand interactions and significant changes to these properties, as observed here, could hint at a potential disruption of histamine's signalling function for invertebrates with consequences for essential ecosystem interactions.

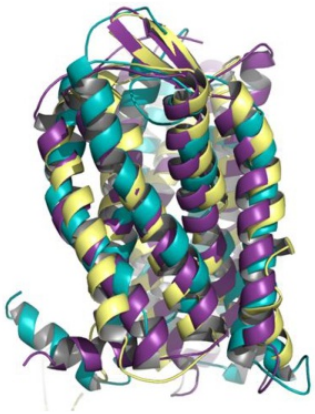
[1] R.L. Swanson *et al.*, *Biol. Bull.* **2004**, *206, 161-172*
[2] F. Lombardo, R. Maramaldo, B. Fratello, D. Sonetti, *Comp. Biochem. Physiol.*, **1992**, *101C, 2, 389 – 398*

Intrinsic Flexibility and Structural Stability of Proteins

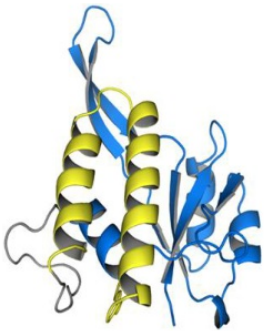
Anna Kahler

HPC@RRZE, Friedrich-Alexander-Universität Erlangen-Nürnberg

The properties of a protein are based on its fold and thus, based on the sequence of the amino acids. Molecular and structural biology have provided a wealth of information about proteins, however, the underlying dynamical processes are not yet fully understood. I will show you how I used MD (molecular dynamics) simulations to gain atomistic information on the dynamical behavior of biologically relevant proteins.



G-protein coupled receptors (GPCRs) interact with small molecules, peptides, or proteins and transmit a signal over the membrane via structural changes to activate intra-cellular pathways. GPCRs are characterized by a rather low sequence similarity and exhibit structural differences even for functionally closely related GPCRs. I propose a computational approach that relies on the generation of several independent models based on different template structures, which are subsequently refined by MD simulations. The conformational stability and the agreement with GPCR-typical structural features is then used to select a favorable model. This strategy was applied to predict the structure of the herpesviral chemokine receptor US28 by generating three independent models based on the known structures of the chemokine receptors CXCR1, CXCR4, and CCR5. Model refinement and evaluation suggested that the GPCR-typical structural features, such as a conserved water cluster or conserved non-covalent contacts, are present to a larger extent in the model based on CCR5 compared to the other models. A final model validation based on the recently published US28 crystal structure confirms that the CCR5-based model is the most accurate and exhibits 80.8 % correctly modeled residues within the transmembrane helices. [1]



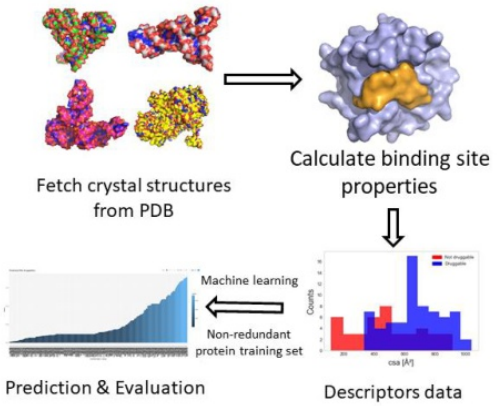
The transcription factor RfaH from Escherichia coli consists of an N-terminal domain (NTD) and a C-terminal domain (yellow alpha-helices, CTD), which tightly interact in the autoinhibited conformation of RfaH. Upon activation, the CTD is released and undergoes a large-scale $\alpha \rightarrow \beta$ structural transition. Investigation of RfaH under different environmental conditions revealed that not only high temperatures, but also a decrease in ionic strength significantly enhances CTD dynamics. None of the conditions investigated caused CTD dissociation suggesting that this process needs to be triggered by the interaction with DNA or other proteins of the transcription machinery. [2]

[1] A. Kahler, H. Sticht, AIMS Biophysics, **2016**, *3*, 211–231.
[2] A. Kahler, A.H.C. Horn & H. Sticht, Curr Biotechnol, **2015**, *4*, 26–38.

Druggability Predictions of Ribonucleic Acid Crystal Structures

Illimar Rekand, Ruth Brenk

Department of Biomedicine, University of Bergen, Norway



It is estimated that the intersection of genes which are disease-modifying and suited druggable targets constitutes about 0.02 to 0.05% of the human genome, limiting the amount of viable protein targets.[1, 2] Examples include riboswitches, which acts as 5'-UTR *cis*-regulating elements, which upon binding to small metabolites undergo a conformational change and thus switch on or off gene expression.[3-5]

A drug target needs to be relevant for a given disease and be able to be modulated by either biological or small molecules. In the latter case, the drug target needs to have a pocket that can bind drug-like ligands with high affinity, a property which is referred to as being "druggable". To identify such pockets in protein structures, we have previously derived a druggability predictor named DrugPred[6, 7], which we have in this work extended to also classify RNA binding sites. Due to the paucity of validated druggable RNA binding sites, we trained the predictor, DrugPred_RNA on protein binding sites described only by descriptors applicable for both protein and RNA structures. DrugPred_RNA, distinguishes druggable from less druggable binding sites with high accuracy. Further, known druggable RNA binding sites are classified correctly, making this a useful tool for RNA-based druggability predictions.

1. Hopkins, A. L.; Groom, C. R. *Nature Reviews Drug Discovery* **2002,** 1, 727-730.
2. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. *Nature Reviews Drug Discovery* **2006,** 5, 993-996.
3.Blount, K. F.; Breaker, R. R. *Nat Biotechnol* **2006,** 24, 1558-64.
4.Breaker, R. R. *Future Microbiology* **2009,** 4, 771-773.
5.Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013,** 152, 17-24.
6.Krasowski, A.; Muthas, D.; Sarkar, A.; Schmitt, S.; Brenk, R. *J Chem Inf Model* **2011,** 51, 2829-42.
7.Sarkar, A.; Brenk, R. *Plos One* **2015,** 10.

Recognition and specificity in protein-DNA interactions – insight from molecular simulations

Petra Imhof

Department of Chemistry, Bioscience and Environmental Engineering
University of Stavanger
4036 Stavanger, Norway

DNA processing proteins must recognise their target site with high specificity to be functional. Restriction enzymes and transcription factors discriminate between local alterations in the sequence, i.e. differences by only one base pair step, direct or inverted repeat, or subtle changes such as methylation of one base. Repair enzymes on the other hand, are able to distinguish single mismatched or damaged bases from intact, canonical DNA. Previous research, using classical molecular dynamics (MD) simulations of protein-DNA complexes at atomic level, has shown that the target site recognition often is an interplay of direct protein-DNA interactions and the indirect read-out of the DNA's shape as determined by the conformational dynamics of the DNA. Our work on the DNA repair enzyme Thymine DNA Glycosylase is an example for how specificity is controlled at various stages: by binding, conformational transition, and the chemical step. Recognition of mismatched DNA at the binding stage is achieved by the enzyme's preference for a slightly distorted conformation, not present in intact DNA. Further discrimination is achieved by the substrate base being better acommodated in the active site than non-cognate bases. As shown by QM/MM calculations, the glycosidic bond scission in the enzymatic complex is via a step-wise dissociative mechanism and largely facilitated by a proton transfer to the leaving base that is unlikely for intact cytosine bases. The chemical step can thus be understood as the last of several instances to protect intact DNA from base excision.

Water-mediated protein-DNA interactions are particularly interesting in the context of indirect (shape) read-out. Mainly the first hydration shells of the DNA and the water dynamics therein are affected by alterations in the DNA sequence or by damages such as methylation and oxidation. Their analysis further elucidates how target site recognition by the proteins may be achieved from a more distant 'search state' rather than an already, albeit loosely, bound 'interrogation state'.

Understanding the substrate access mechanism in Phospholipase A from *Pseudomonas aeruginosa*

Sabahuddin Ahmad[1], Stephan Schott-Verdugo[1], Christoph Strunk[2], Filip Kovacic[2], Karl-Erich Jaeger[2], Holger Gohlke[1,3]

[1]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany and Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

[3]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich GmbH, Jülich, Germany

The Gram-negative bacterium *Pseudomonas aeruginosa* is an opportunistic pathogen and frequent cause of nosocomial infections, affecting primarily immune-compromised patients. PlaF from *P. aeruginosa* is a novel phospholipase A1 (PLA1). It is an integral inner membrane protein and has been shown to be a relevant virulence factor. A recent crystal structure of PlaF together with crosslinking and micro-scale thermophoresis experiments revealed PlaF to be in monomeric and dimeric configurations, although the protein was found to be active only in the monomeric state. Our previous free energy computations and wet-lab experiments indicated that, depending on the protein concentration, the dimer-to-monomer equilibrium is shifted to either side. The dimer-to-monomer transition is associated with a change in the orientation of the active site tunnel with respect to the membrane: A) in the dimeric form, the tunnel is *parallel* to the membrane *surface*, and the entrance resides more than 5 Å above it, which will likely hamper substrate access; B) in the monomeric form, PlaF tilts and the active site tunnel is in direct contact with the membrane, which will likely facilitate the substrate access from the membrane. We therefore hypothesize that the tilting of monomeric PlaF leads to this form being active by facilitating substrate access to the active site tunnel.

On further investigation, it was found that there are other cavities than the active site tunnel, resulting in total four access pathways (Tunnel 1-4) to/from the active site. However, which of these pathways is most favorable for substrate access remained elusive. To answer this, we performed configurational free energy computations of substrate access from the membrane to the active site. We considered different phospholipid substrates to understand this. Since a phospholipid substrate can enter by its head or either one of the two tails, we tested all three scenarios. Our computations reveal that A) Tunnel 2 is energetically most favorable for substrate access and B) the substrate entering with tail 1 first is the most favorable access mode. This access mode agrees with the fact that PlaF is a PLA1 that cleaves tail 1 of phospholipids. To validate our hypothesis about tunnel 2, we proposed tryptophan substitutions for all the four tunnels. Among proposed substitutions, the ones corresponding to tunnel 2 reduced the activity of PlaF by ~70%. Hence, it was confirmed that the tunnel 2 of Plaf is the preferred access tunnel for substrates. At present we aim to understand the egress pathway of the PlaF products. Together with our previous studies, our results can provide an explanation to the activity mechanism of PlaF at the atomistic level.

Reproducing high accuracy QM interactions via empirical force fields

Amin Alibakhshi

Department of Theoretical Chemistry, Christian-Albrechts University Kiel
(alibakhshi@pctc.uni-kiel.de)

Reliable and accurate evaluation of inter and intra-atomic interactions is one of the key steps in majority of molecular modeling approaches. Although these interactions can be precisely evaluated via high quality ab initio methods, the computational costs of such approaches specially for large systems has made applying empirical force fields still inevitable.

The empirical force fields typically fractionize the total energy of a molecule into smaller components such as bond stretching, angle and torsional bending and non-bonded interactions and evaluate each one of these energy terms via simple algebraic and empirical expressions. The conventional approaches of parametrizing such empirical force fields typically rely on transferability of force field parameters and determining those parameters such that they reproduce one or few bulk properties of the ensemble as accurate as possible. Nevertheless, remarkable inaccuracies inherent in such parameterization approaches reported in many recent studies has been the motivation of proposing next generation force fields e.g. machine learning based and ab initio derived force fields as well as more rigorous system specific parameterization strategies e.g. via the force matching technique [1].

In the present study we evaluated the performance of several conventional and more rigorous recently proposed empirical force fields in reproducing high accuracy QM interactions. As a benchmark, we carried out QM computations for 1500 configurations of boric acid and borate solvated in a small cluster of 64 water molecules. Each force field was parameterized by force matching technique to reproduce the reference QM forces. The studied configurations were extracted from molecular dynamics snapshots and selected such that they cover a wide combination of various force field variables. According to the results, none of the conventional force fields could reproduce QM forces with average unsigned relative error more accurate than 104%. Nevertheless, through a new empirical force field introduced in the present study which is a slightly modified version of a robust reactive force field recently developed in our group [2], we could achieve 8% accuracy in average unsigned relative error of evaluated forces.

1. Steffen, J. and B. Hartke, *Cheap but accurate calculation of chemical reaction rate constants from ab initio data, via system-specific, black-box force fields.* The Journal of chemical physics, 2017. **147**(16): p. 161701.
2. Hartke, B. and S. Grimme, *Reactive force fields made simple.* Physical Chemistry Chemical Physics, 2015. **17**(26): p. 16715-16718.

A new Approach to DPD Repulsion Parameter Estimation

Felix Bänsch [1,3], Mirco Daniel [1], Harald Lanig [2], Christoph Steinbeck [3], Achim Zielesny [1]

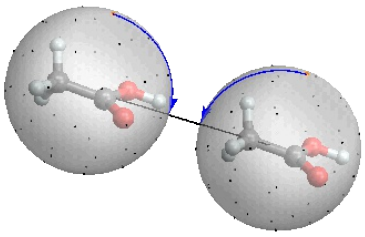
1 Institute for Bioinformatics and Cheminformatics, Westphalian University of Applied Sciences, Recklinghausen, Germany

2 Department of Chemistry and Pharmacy, Friedrich-Alexander University, Erlangen-Nürnberg, Germany

3 Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller University, Jena, Germany

Dissipative Particle Dynamics (DPD) is a mesoscopic simulation technique for complex fluids and soft matter systems. Molecular Fragment DPD is a "bottom-up" variant with particles being defined as small "fragment molecules" with a molecular weight in the order of 100 Da. Larger molecules are then partitioned into adequate smaller "fragment molecule" particles that are bonded by harmonic springs to mimic covalent connectivities and spatial 3D conformations. The conservative interaction between two DPD particles i and j is characterized by an isotropic repulsion a_{ij} [1].

This project aims at a new force field based approach for consistent a_{ij} parameter estimation to obtain a "fragment molecule" particle set especially for biomolecular simulations that contain peptides and proteins.



The a_{ij} parameters are themselves determined by a coordination number Z_{ij}, i.e. the average number of particles i around a particle j, and the differential pair interaction energy ΔE_{ij} of the particle pair so that strategies for approximation of the latter two are under investigation.

Final goal is an open protocol for automated a_{ij} parameter estimation that utilizes the Wolfram Language of the Mathematica platform [2] as a math-enabled scripting engine where all force field related molecular mechanics and dynamics calculations are to be performed with the open TINKER package [3].

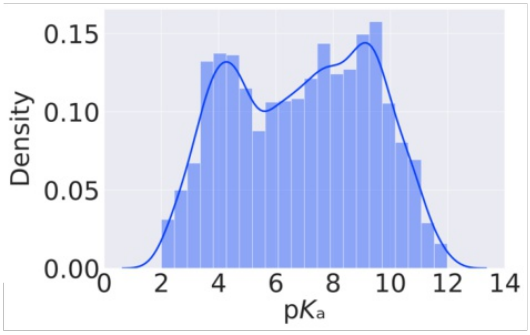
[1] R. D. Groot, P. B. Warren, J. Chem. Phys., 1997, 107, 4423
[2] Wolfram Mathematica: https://www.wolfram.com/mathematica/
[3] J. Rackers *et al.*, J. Chem. Theory Comput., 2018, 14, 5273

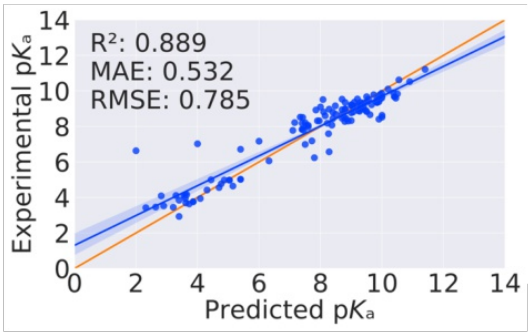
Machine learning meets pK_a

Marcel Baltruschat, Paul Czodrowski

TU Dortmund University, Faculty of Chemistry and Chemical Biology

The acid-base dissociation constant (pK_a) of a drug has a far-reaching influence on pharmacokinetics by altering the solubility, membrane permeability and protein binding affinity of the drug [1,2]. That's why high quality pK_a prediction methods are very important within the drug discovery process. However, there is no publicly available, open source and license-free pK_a prediction tool that can reach the quality of licensed programs like *MoKa* [3] or *Marvin* [4]. Our goal is to develop a new, highly accurate pK_a prediction tool based on freely accessible data and free to use for everyone.





To achieve our goal, we identified a dataset from *DataWarrior*, which contains 7913 pK_a values measured in water. In addition, we could extract 8111 values from the ChEMBL25 database. The values were then preprocessed, filtering out all molecules according to Lipinski's rule of five (one violation allowed), unwanted atoms and bad functional groups. Additionally, any salts were removed and tautomer canonization was performed with *QUACPAC* [5]. Only structures that are considered monoprotic in the pH range of 2 to 12 (determined by *Marvin* [4]) were retained. After preprocessing the structures, multiple measurements were combined and any outliers were removed. The final result was a curated monoprotic training dataset with 5994 unique structures. The distribution of the pK_a values is shown in the left figure.
This training dataset was used to test a total of seven different machine learning configurations with the basic regressors Random Forest, Support Vector Regression, Multilayer Perceptron and XGradientBoost. In addition, six different descriptor/fingerprint sets were examined for each configuration, resulting in a total number of 42 trained and evaluated machine learning models. The models were all 5-fold cross-validated and evaluated using two different external test datasets, one dataset taken from the scientific literature and another dataset provided by Novartis [6]. The predictive quality of the best-performing model as measured by the external test dataset from the literature is shown in the right figure.
In further work all commercial tools used for preprocessing will be replaced by open source applications and the prediction capabilities will be extended to multiprotic molecules.

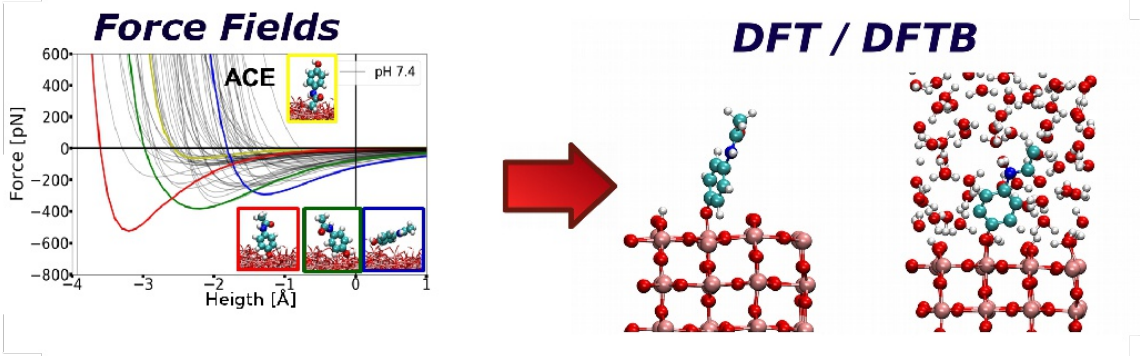
[1] Charifson, P. S., & Walters, W. P. (2014). Acidic and Basic Drugs in Medicinal Chemistry: A Perspective. *Journal of Medicinal Chemistry*
[2] Manallack, D. T. (2007). The pKa Distribution of Drugs: Application to Drug Discovery. *Perspectives in Medicinal Chemistry*
[3] MoKa, Molecular Discovery Ltd. Borehamwood, United Kingdom
[4] Marvin 20.1.0, 2020, ChemAxon Ltd, *http://www.chemaxon.com*
[5] QUACPAC 2.0.2.2: OpenEye Scientific Software, Santa Fe, NM. *http://www.eyesopen.com*
[6] Richard A. Lewis, Stephane Rodde, Novartis Pharma AG, Basel, Switzerland

Water contaminants on different models of rutile TiO2 - a DFT(B) / FF approach

F. Balzaretti, V. Gupta, M. Romero, S. Köppen

University of Bremen, Faculty of Production Engineering, Hybrid Materials Interfaces, and Bremen Center for Computational Materials Science (BCCMS) and Bremen Center for Materials and Processes (MAPEX) Bremen, Germany

The impact of TiO_2 as a fundamental material for new technologies is still nowadays worldwide recognized. Independently from its two main phases, anatase and rutile, this semiconductor proved to be able of major photocatalytical activity within visible light irradiation, which turns out to be a powerful tool for relevant engineering applications such as water splitting or environmental cleaning [1, 2]. Moreover, TiO_2 powders with negligible rate of impurities can be synthesized at very affordable prices, enabling it to reach a primary role with respect to other materials. Optimal reaction rates at basic pH conditions have been reported for several adsorbates [3, 4, 5], showing an environmental dependence. This motivates more complex simulations of small organic pollutants taking into account surface defect states, aqueous environment with varying pH conditions as well as the presence of different pollutants at different concentrations.



Here we present extensive interaction studies of nine worldwide abundant aqueous contaminants, for which force-displacement curves were evaluated at different molecules orientations through *Force Fields (FF)* methods. Among other very reactive contaminants, the Glyphosate revealed interaction forces higher by a factor of 10. Next to this prominent herbicide, the most favorable final configurations of two other pollutants belonging to the pharmaceutical and intensive breeding sectors were taken as initial structures for both *Density Functional Theory (DFT)* and *Tight Binding (DFTB)* approaches. Independently from the class of applications, all contaminants show thermodynamically stable geometries with hydroxyl and carboxylic groups towards the surface on pristine and reconstructed surfaces of rutile [6].

[1] A. Migani and L. Blancafort, *Chemical Reviews*, **2013**, 113, 3887-3948
[2] Y. Lan, Y. Lu and Z. Ren, *Nano Energy*, **2013**, 2, 1031-1045
[3] K. Dai, P. Tianyou, H. Chen et al., *Environ. Sci. Technol.*, **2008**, 42, 1505-1510
[4] X. D. Zhu, Y. K. Wang et al., *Chemosphere*, **2013**, 92, 925-932
[5] H. Gao, J. Zhou, D. Dai, Y. Qu, *Chem. Eng. Technol.*, **2009**, 32, 867-872
[6] F. Balzaretti, V. Gupta, S. Köppen, **2020**, *to be submitted*

Probing DNA Conformation: A Multi-Technique Approach

Frank R. Beierlein,[1] Jack S. Hardwick,[2] Marius M. Haugland,[2] Afaf H. El-Sagheer,[3] Denis Ptchelkine,[4] Andrew N. Lane,[5] Tom Brown,[2] Janet E. Lovett,[6] Edward A. Anderson[2]

[1] *Computer-Chemistry-Center and Interdisciplinary Center for Molecular Materials, Department of Chemistry and Pharmacy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstr. 25, 91052 Erlangen, Germany*
[2] *Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK*
[3] *Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez 43721, Egypt*
[4] *Weatherall Institute of Molecular Medicine, Department of Oncology, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DS, UK*
[5] *Center for Environmental and Systems Biochemistry and Department of Toxicology & Cancer Biology, The University of Kentucky, 789 S. Limestone St., Lexington, KY 40509 USA*
[6] *SUPA School of Physics and Astronomy and BSRC, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK*

The determination of distances between specific points in nucleic acids is essential to understanding their behaviour at the molecular level. The ability to measure distances of 2–10 nm is particularly important: deformations arising from protein binding commonly fall within this range, but the reliable measurement of such distances for a conformational ensemble remains a significant challenge. We show that MD simulations are a robust tool to interpret electron paramagnetic resonance (EPR) measurements of oligonucleotides spin-labelled with triazole-appended nitroxides at the 2' position. For two nitroxide spin-labels attached to B-form DNA duplexes, we present results from EPR spectroscopy, MD simulations, X-ray crystallography, and NMR spectroscopy. These four methods are mutually supportive and pinpoint the locations of the spin labels on the duplexes. In doing so, this work establishes 2'-alkynyl nitroxide spin-labelling as a minimally perturbing method for probing DNA conformation.

[1] J. S. Hardwick, M. M. Haugland, A. H. El-Sagheer, D. Ptchelkine, F. R. Beierlein, A. N. Lane, T. Brown, J. E. Lovett, E. A. Anderson, *2'-Alkynyl spin-labelling is a minimally perturbing tool for DNA structural analysis, Nucleic Acids Res.* **2020**, accepted.
[2] M. M. Haugland, A. H. El-Sagheer, R. J. Porter, J. Pena, T. Brown, E. A. Anderson, J. E. Lovett, 2'-*Alkynylnucleotides: A Sequence- and Spin Label-Flexible Strategy for EPR Spectroscopy in DNA, J. Am. Chem. Soc.* **2016**, *138,* 9069-9072.

Studying conformational transitions of selected proteins using UNRES coarse-grained simulations with Lorentzian restraints

Iga D. Biskupek, Cezary Czaplewski, Adam Liwo

Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, Gdańsk, Poland.
dominika.biskupek@phdstud.ug.edu.pl

Identifying the mechanism of conformational transitions of proteins is essential in the understanding of their biological functions. When investigating the conformational transitions, experimental methods provide only fragmentary information about the time evolution of the system such as the distance between donor and acceptor groups or distance distribution, therefore the use of simulation method to interpret the experimental data is necessary. An all-atom structure-based method for the study of the structural transition of proteins by utilizing Lorentzian attractive terms has been developed recently[1]. Two Lorentzian attractive interactions provide a double-well potential with a bounded energy barrier. We implemented this method to coarse-grained UNRES simulations, this enabling us to study larger systems with a smaller at lower computational expense. UNRES is a highly reduced protein model with only two interaction sites per residue[2]. Owing to this reduction, it offers ~1000-fold speed-up compared to all-atom molecular dynamics.

The aim of the study is to test the UNRES implementation of the method of modeling conformational transitions with double-well Lorentzian guiding function. For testing, we selected to fallowing two proteins with two well-defined states: the apo (PDB ID: 1LFH) and the holo (PDB ID: 1LFG) forms of lactoferrin, respectively, and the open (PDB ID: 4AKE) and the closed (PDB ID: 1AKE) forms of adenylate kinase,[3] respectively. The adenylate kinase is a phosphotransferase enzyme found in various organisms. The enzyme plays an essential role in cellular energy homeostasis by catalyzing the interconversion between ADP and ATP. Lactoferrin is an iron-binding protein present in large quantities in colostrum and in breast milk, in external secretions and in polymorphonuclear leukocytes. Lactoferrin's main function is non-immune protection[4]. Martiple conformational transitions between above-mentioned forms were observed in simulations for both proteins. The free energy landscapes were constructed and transition pathways were identified.

[1] J. Lee, K. Joo, B.R. Brooks, J. Lee, *J Chem Theory Comput*, **2015**, *11*, 3211-3224.
[2] A. Liwo, M. Baranowski, C. Czaplewski, E. Gołaś, Y. He, D. Jagieła, P. Krupa, M. Maciejczyk, M. Makowski, M.A. Mozolewska, A. Niadzvedtski, S. Ołdziej, H.A. Scheraga, A.K. Sieradzan, R. Slusarz, T. Wirecki, Y. Yin, B. Zaborowski, *J Mol Model*, **2014**, *20*, 2306.
[3] A.A. Rashin, A.H.L. Rashin, R.L. Jernigan, *Biochemistry*, **2010**, *49(27)*, 5683-5704.
[4] F. Giansanti, G. Panella, L. Leboffe, G. Antonini, *Pharmaceuticals*, **2016**, *9(4)*, 61.

Disarming the Glycan Shield of HIV: Broadly Neutralizing Antibody PGT122 against HIV

Manuel Deubler, Anselm H. C. Horn, and Heinrich Sticht

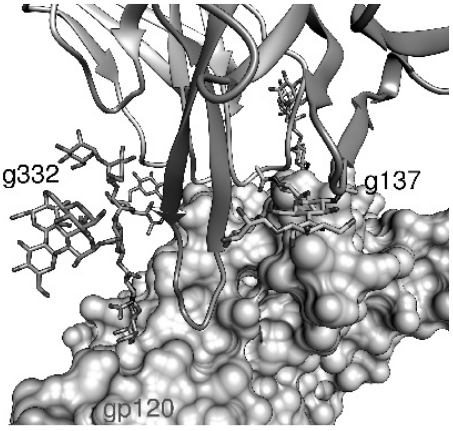
Bioinformatik, Institut für Biochemie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

The human immunodeficiency virus (HIV) establishes a latent infection causing a slow depletion of $CD4^+$ T-cells belonging to the humoral immune system. Ultimately, this results in the stage of the acquired immunodeficiency syndrome (AIDS) in which infected individuals are highly susceptible for opportunistic diseases.

The virus itself circumvents recognition by the immune system with a high mutation rate and an envelope which contains host-cell derived features. These include the numerous glycans attached to the only viral protein on the surface of HIV, gp120. The glycans form a protective layer against immune recognition called the glycan shield.

However, some patients develop antibodies that show neutralizing properties against multiple virus strains called broadly neutralizing antibodies (bnAbs). Recently, bnAbs have been investigated for immunization via passive administration and for therapeutic use. They also serve as a template for the design of mimetic peptides. One potential candidate is PGT122, a bnAb that targets the V3 region of gp120. At its binding site PGT122 interacts with several glycans through a 26 residue long complementary determining region of the heavy chain (CDRH3).

Investigation of the interactions between the antibody and the glycans on the viral surface during affinity maturation and for peptide design, all-atom MD simulations of the investigated antibody-gp120 systems were performed on the µs timescale. These revealed distinct dynamics for each of the four simulated glycans (g137, g156, g301, g332).



Only interactions to g332 in the center of the epitope remained stable throughout the simulation time whereas the other glycans displayed a high degree of conformational freedom. Replacement of the CDRH3 loop of PGT122 with its putative precursor showed less, but stable contacts to g332. This indicates subtle optimization of the precursor Ab during affinity maturation. On the other hand, interactions with g137 were lost in some simulations which underlines this glycan's distinct role in antibody recognition. The simulations of a PGT122 derived peptide showed unphysiological structural adaptations emphasizing the importance of the V_H region for H3 structural integrity.

Thus, these simulations provide insights into the dynamics of the glycan shield and stability of CDRH3 loops for novel antiviral approaches against HIV.

Nucleic acid building blocks: The role of tautomerism

Lukas Eberlein, Nicolas Tielker, Stefan M. Kast

Fakultät für Chemie und Chemische Biologie, TU Dortmund, Germany

Knowledge of the structure and thermodynamic properties of biologically active molecules, in particular their conformational, tautomer, and protonation state preferences is essential to understand their behavior in solution and mode of action in the organism. An accurate pH-, temperature- and pressure-dependent characterization of these properties in solution is of vital importance but poses a challenge to both experiment and theory even for well-studied compounds such as nucleic acid building blocks. Experimentally, rapid conformational changes and the fast proton transfer between multiple tautomeric forms make elucidating these equilibria cumbersome especially under extreme conditions, while the theoretical task is complicated due to the environmental effect on both electronic structure and solvent distributions.

Conceptually footed in the methodology employed within the SAMPL blind prediction challenges for tautomer equilibria, distribution coefficients and acidity constants, [1-3] we here combine tautomer and conformational sampling for natural and non-natural nucleic acid building blocks. Solvent effects on energetics and spectroscopic parameters in quantum-chemical calculations are considered using the "embedded cluster reference interaction site model" (EC-RISM) developed by us, which has been demonstrated to provide accurate estimates of thermodynamic quantities and spectroscopic features in solution even for high pressure solvents. [4,5] The EC-RISM workflow is refined by coupled-cluster extrapolation, which allows treatment of electron correlation effects with high accuracy, and by the incorporation of explicit solvent molecules to further improve the predictive quality for tautomer preferences and spectroscopic parameters. We obtain the contributions of all accessible states to the molecular ensemble as a function of pH, pressure and temperature, even for conformationally flexible species, [6] as well as their respective relevance for understanding experimental NMR spectra. This can not only lead to a further understanding of the implications of tautomerism for the base pairing in nucleic acids under extreme conditions and therefore for the universality of our genetic code and the extended Hachimoji code [7] in the universe, but also to the identification of variations in the purine scaffold of adenine and guanine species to develop new, tautomer-stable, building blocks for DNA-encoded libraries or synthetic biology.

[1] S. M. Kast, J. Heil, S. Güssregen, K. F. Schmidt, *J. Comput.-Aided Mol. Des.* **2010**, *24*, 343-353.
[2] N. Tielker, D. Tomazic, J. Heil, T. Kloss, S. Erhart, S. Güssregen, K. F. Schmidt, S. M. Kast, *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1035-1044.
[3] N. Tielker, L. Eberlein, S. Güssregen, S. M. Kast, *J. Comput.-Aided Mol. Des.* **2018**, *32*, 1151-1163.
[4] R. Frach, P. Kibies, S. Böttcher, T. Pongratz, S. Strohfeldt, S. Kurrmann, J. Koehler, M. Hofmann, W. Kremer, H. R. Kalbitzer, O. Reiser, D. Horinek, S. M. Kast, *Angew. Chemie Int. Ed.* **2016**, *55*, 8757-8760.
[5] T. Pongratz, P. Kibies, L. Eberlein, N. Tielker, C. Hölzl, S. Imoto, M. Beck Erlach, S. Kurrmann, P. H. Schummel, M. Hofmann, O. Reiser, R. Winter, W. Kremer, H. R. Kalbitzer, D. Marx, D. Horinek, S. M. Kast, *Biophys. Chem.* **2020**, *257*, 106258.
[6] N. Tielker, L. Eberlein, C. Chodun, S. Güssregen, S. M. Kast, *J. Mol. Model.* **2019**, *25*, 139.
[7] S. Hoshika, N. A. Leal, M.-J. Kim, M.-S. Kim, N. B. Karalkar, H.-J. Kim, A. M. Bates, N. E. Watkins Jr., H. A. SantaLucia, A. J. Meyer, S. DasGupta, J. A. Piccirilli, A. D. Ellington, J. SantaLucia Jr., M. M. Georgiadis, S. A. Benner, *Science* **2019**, *363*, 884-887.

A QM/SQM embedded cluster RISM approach for predicting EPR parameters of protein-bound nitroxide spin probes

Lennart Eisel, Tim Pongratz, Stefan M. Kast

Fakultät für Chemie und Chemische Biologie, TU Dortmund, Germany

Spectroscopic methods such as electron paramagnetic resonance (EPR) have become popular tools to study structure and dynamics of biomolecular systems. First principles calculations of spectra can enhance the understanding of these processes at an atomic scale, though the size and complexity of solvated biomolecular systems such as enzymes often require a drastic simplification of the underlying physics, thus decreasing the predictive capabilities of the computational method. [1] Since their first description by Warshel and Levitt [2] multiscale methods have emerged as an effective tool to model chemical processes in a heterogeneous protein environment in a cost-effective way. For instance, a detailed quantum-mechanical (QM) treatment of a region of interest in a complex system such as a solvated protein could be coupled to a force field-based, molecular mechanics (MM) environment including an atomically resolved solvent, giving rise to the class of hybrid QM/MM models [1] that are also applicable to the prediction of spectroscopic parameters. [3] Still, further reduction of system size to include only relevant solvent molecules explicitly could further enhance computational performance as long as quantitative properties are negligibly affected. To this end, realistic implicit solvent models that retain some structural information about the environment could offer a way forward.

We here present a computational method to predict electron paramagnetic resonance (EPR) parameters, such as isotropic hyperfine constants, on the basis of the "embedded cluster reference interaction site model" (EC-RISM) [4,5] and its application to a nitroxide spin probe in free solution and attached to a protein. The methodology combines ideas from multiscale approaches with a 3D RISM-based solvation description that models the background based on distribution functions of a granular solvent model. More specifically, in addition to the statistical solvent background a spin probe can be exposed to a few explicitly QM-modeled protein residues based on the extrapolative ONIOM scheme, [6] thus facilitating mutual polarization of the spin probe and the high-level QM zone of the protein. Additionally, the description of the complete protein system by semi-empirical QM methods (QM/SQM) allows for the polarization of all remaining protein residues by the solvent environment, giving rise to a more realistic description of the system compared to conventional QM/MM approaches. First results are presented and discussed in comparison to experimental evidence.

[1] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198-1229.
[2] A. Warshel, M. Levitt, *J. Mol. Biol.* **1979**, *103* 227-249.
[3] S. Sinnecker, F. Neese, *J. Comput. Chem.* **2006**, *27*, 1463-1475.
[4] T. Kloss, J. Heil, S. M. Kast, *J. Phys. Chem. B* **2008**, *112*, 4337-4443.
[5] N. Tielker, D. Tomazic, J. Heil, T. Kloss, S. Ehrhart, S. Güssregen, K. F. Schmidt, S. M. Kast, *J. Comput.-Aided Molec. Des.* **2016**, *30*, 1035-1044.
[6] S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma, M. J. Frisch, *Theo. Chem.* **1999**, *461*, 1-21.

Computational Studies of Antifreeze Peptides

Gandini E.[1,3], Sironi M.[1,2], Pieraccini S.[1,2]

[1]Dipartimento di Chimica, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy
[2]Istituto di Scienze e Tecnologie Chimiche "G. Natta" (SCITEC-CNR), CNR, and INSTM, UdR Milano, Via Golgi 19, 20133 Milano, Italy
[3]Contact: enrico.gandini@unimi.it

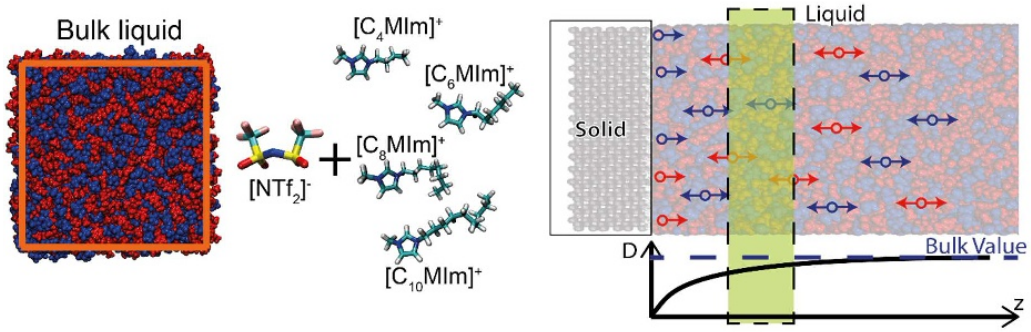
Antifreeze Proteins (AFPs) are a class of structurally diverse proteins that protect different organisms, e.g. polar fishes, from fatally freezing in icy environments. They are very attractive for potential practical applications, including food storage, cryopreservation, and anti-icing coatings. Efficient synthetic analogues of AFPs are highly desirable; three 12-residue analogues of Winter Flounder AFP (wfAFP) were successfully applied for the fabrication of anti-icing surfaces. Focus of the present work is the employment of computational techniques to elucidate the mechanism of action of such peptides.
To this aim, three kinds of simulation setups were devised, in order to observe and analyze different aspects of the antifreeze activity of the three wfAFP analogues at a molecular level. For control, a same lenght non-antifreeze peptide such as dodeca-Glycin (G12) was included in the study. Water molecules were described with TIP4P/Ice model, which is able to reproduce solvent properties near the freezing point.
Computational analysis are in good agreement with reported experimental results. The three simulation setups are able to realistically describe the properties of the antifreeze peptides at a molecular level, and to point out structural aspects of the ice-peptide interactions which will be useful for further optimization and engineering of AFP analogues.

Influence of self-organization induced local clustering on self-diffusion of [C_nMIm^+] [NTf_2^-] in bulk and nanoconfined systems

Kevin Höllring[1], Nataša Vučemilović-Alagić[2], David M. Smith[2], Ana-S. Smith[1]

[1]*PULS group, Institute of Theoretical Physics, FAU Erlangen-Nürnberg, Germany*

[2]*Division of Physical Chemistry, Ruđer Bošković Institute, Zagreb, Croatia*



The various applications of ionic liquids require comprehensive knowledge of both their static and dynamic properties. Molecular dynamics simulations can be employed in order to obtain an understanding of how these properties emerge by analyzing the behavior of individual particles within different unconfined as well as confined systems.

We extend the previous study on the properties of [C_2MIm^+] [NTf_2^-] carried out by our group [1] to additional members of the [C_nMIm^+] [NTf_2^-] - family of ionic liquids with larger cations, namely n=4, 6, 8, 10. The static and dynamic properties of these ionic liquids are analyzed in a bulk system as well as in a nano confined solid-liquid-vacuum system with a supporting sapphire crystal.

We will provide evidence that the dynamic properties of the liquids with longer cation chains are governed by the formation of local ion clusters acting as meta-particles by employing the known measure of ion pair lifetimes as well as devising the mean number of encountered closest neighbors of ions within a certain simulation timeframe as a measure for how strongly a particle is confined to a stable local environment [2].

Furthermore, we will devise a theoretical model for obtaining the absolute diffusion coefficient of particles perpendicular to an interface by determining their mean lifetime within a confined subspace without relying on mean square displacement fitting in confined subspaces [2,3].

The influence of relevant parameters of the observed ionic liquids not considered in the theoretical derivation of the estimator are studied and possible corrections to the estimator depending on rotational speed and deformation of particles examined.

[1] Vucemilovic-Alagic, N.; Banhatti, R. D.; Stepic, R.; Wick, C. R.; Berger, D.; Gaimann, M. U.; Baer, A.; Harting, J.; Smith, D. M.; Smith, A.-S. *arXiv*, **2019**, *1903.09450*
[2] Höllring, K.; Smith, A.-S. *Master thesis*, **2019**
[3] Höllring, K.; Vucemilovic-Alagic, N.; Smith, D. M.; Smith, A.-S. *in preparation*

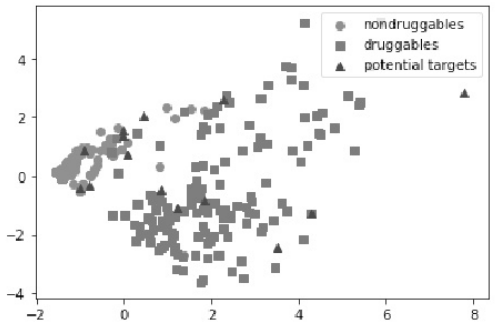
Druggability Prediction using Gene Ontology terms

Marie Detzler, Michael C. Hutter*,

** corresponding author: michael.hutter@bioinformatik.uni-saarland.de*

Center for Bioinformatics, Saarland University, Campus E2.1, 66123 Saarbrücken

Classification of potential drug targets into druggables and nondruggables has gained increased interest in the course of preclinical development and virtual screening. The considerable amount of potential targets in the human genome, as well as in pathogens has emphasized the need for corresponding *in silico* detection of valid targets. Whereas there is a large number of targets for which small molecule inhibitors have been approved as medications, [1] only few cases of targets that turned out to nondruggable were reported. Such disparate data sets where one class is highly overrepresented are a big problem for classification algorithms in general, because predictions for the underrepresented class are far less accurate. For example, simply the chance of accidentally classifying a member of the larger class correctly is much higher. Consequently, we applied several resampling techniques that comprise oversampling of the minority class and under-sampling the majority class in the framework of random forest and ada boost. [2, 3] Whereas other approaches for druggability prediction [4, 5] require 3D structures of the corresponding targets as input, we used terms form Gene Ontology, which describe the protein at hand by its associated biological processes, molecular function, and sub-cellular location. [6]



The first two principal components of the data sets indicate a separation of known druggabble (squares) and nondruggable (circles) targets, whereas the additional potential targets (triangles) are more spread among the two classes.

[1] M. Rask-Andersen, S. Masuram, H. Schiöt, *Ann. Rev. Pharmacol. Toxicol.*, **2014**, *54*, 9-26.
[2] N. Chawla, K. Bowyer, L. Hall, W. Kegelmeyer, *J. Artificial Intell. Res.*, **2002**, *16*, 321-357.
[3] D. Wilson, *IEEE Transactions Systems Man Cybernetics*, **1972**, *2-3*, 408-421.
[4] E. Fauman, B. Rai, E. Huang, *Curr. Opin. Chem. Biol.* **2011**, *15*, 463-468.
[5] A. Volkamer, D. Kuhn, F. Rippmann, M. Rarey, *Bioinformatics*, **2012**, *28*, 2074 -2075.
[6] The Gene Ontology Consortium, *Nucl. Acid. Res. D*, **2007**, *36 suppl.1*, D440-D444.

What can we learn from local water happiness?

Julia B. Jasper[1], Yannic Alber[1], Stefan Güssregen[2], Hans Matter[2], Stefan M. Kast[1]

[1]*Fakultät für Chemie und Chemische Biologie, TU Dortmund, Germany*
[2]*R&D, IDD, Structural Design and Informatics, Sanofi-Aventis Deutschland GmbH, 65926 Frankfurt am Main, Germany*

Water molecules play an essential role in protein-ligand binding as their replacement contributes to the free energy of binding ($\Delta_{bind}G$). Different approaches have been developed that aim at predicting if a water molecule can be targeted for replacement. [1-3] Prospective, quantitative measures to identify ligand chemistry to target specific thermodynamic hydration site features are highly desirable. We therefore analyze local water thermodynamics and its influence on ligand binding with novel physics-based approaches derived from 3D reference interaction site model (RISM) integral equation theory.[1,4] The following questions are addressed: Are the ligand's physicochemical properties imprinted in the *apo* water thermodynamics? Is it possible to predict by which group a water molecule should be replaced? How does a (mis)match between water "happiness" (measured by local contributions to the hydration free energy, $\Delta_{hyd}G$) and ligand properties contribute to $\Delta_{bind}G$? How can this knowledge be used for drug design?

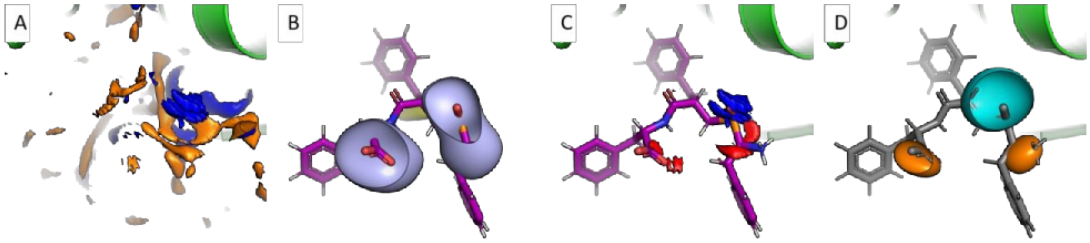


Fig. 1. Analysis for 1os0@pdb. A: "(Un)happy" (orange, blue) *apo* water regions; B: hydrophobic and hydrophilic (yellow, steel-blue) regions of the ligand; C: (mis)matching (red, blue) correlation of *apo* water and ligand properties; D: ligand regions which contribute most (un)favorably (orange, cyan) to $\Delta_{bind}G$.

3D RISM allows for the identification of localized hydration sites along with their local contribution to $\Delta_{hyd}G$. [1] Analysis of the pdbBind core set [5] shows that "unhappy" water molecules tend to be replaced by halogen atoms and aromatic carbons whereas "happy" waters are primarily replaced by polar groups. Furthermore, local contributions to *apo* ligand atom solvation and *holo* binding free energies can be estimated, which can be related to corresponding hydration site features. We can therefore quantitatively test the hypothesis that the contribution of a group to $\Delta_{bind}G$ is related to its hydrophobicity and the *apo* water "happiness", e.g. that replacement of an "unhappy" water by a hydrophobic group is favorable for binding. First results (s. Fig. 1) support this presumption.

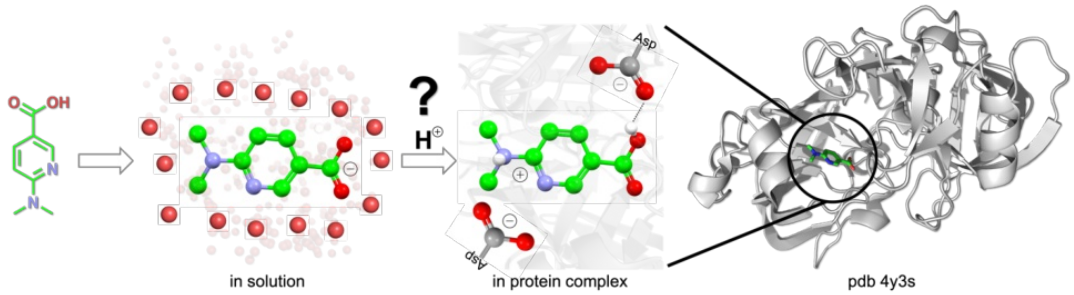
[1] S. Güssregen, H. Matter, G. Hessler, E. Lionta, J. Heil, S. M. Kast, *J. Chem. Inf. Model.* **2017**, *57*, 1652-1666.
[2] M. Miller, J. Schneider, B. K. Sathyanarayana, M. V. Toth, G. R. Marshall, L. Clawson, L. Selk, S. B. Kent, A. Wlodawer, *Science* **1989**, *246*, 1149-1152.
[3] S. B. A. de Beer, N. P. E Vermeulen, C. Oostenbrink, *Curr. Topics in Med. Chem.* **2010**, *10*, 55-66.
[4] F. Mrugalla, S. M. Kast, *J. Phys. Condens. Matter* **2016**, *28*, 344004.
[5] Y. Li, L. Han, Z. Liu, R. Wang, *J. Chem. Inf. Model.* **2014**, *54*, 1717–1736.

Protonation effects of endothiapepsin-fragment complexes

Sascha Jung, Paul Czodrowski

TU Dortmund University, Faculty of Chemistry and Chemical Biology

Endothiapepsin (EP) is an aspartic protease isolated from the ascomycete *Cryphonectria (Endothia) parasitica* and is often used as a milk clotting enzyme in cheesemaking processes. [1] Furthermore, EP often serves as a surrogate for renin (blood pressure) and β-secretase (Alzheimer's disease) in structure-based inhibitor design. [2] Due to a straightforward purification of EP with high yields, we use it as a model enzyme for the investigation of protonation changes upon protein-ligand complex formation (see figure below). Protonation changes upon binding of a ligand to a protein is often an overlooked phenomenon of molecular recognition. In fact, changes in protonation of titratable groups of either the ligand or the protein can have a serious impact on the binding geometry as well as the binding affinity. [3] In order to experimentally determine protonation changes, we perform Isothermal Titration Calorimetry (ITC) experiments with known EP-binding fragments in various buffers with differing heats of ionization. By supporting the ITC measurements with pK_a calculations (Poisson-Boltzmann calculations) we can predict the residue or ligand functional group responsible for the protonation change. [3] In a next step, we additionally want to determine the binding mode of the fragments via X-ray crystallography.



[1] N. Radeva, J. Schiebel, X. Wang, S. Krimmer, K. Fu, M. Stieler, F. Ehrmann, A.r Metz, T. Rickmeyer, M. Betz, J. Winquist, A. Park, F. Huschmann, M. Weiss, U. Mueller, A. Heine, G. Klebe , *J. Med. Chem.*, **2016**, *59(16)*, 7561–7575.
[2] M. Kuhnert, H. Köster, R. Bartholomäus, A. Park, A. Shahim, A. Heine, H. Steuber, G. Klebe, W. Diederich, *Angew. Chem. Int. Ed.*, **2015**, *54*, 2849–2853.
[3] M. Neeb, P. Czodrowski, A. Heine, L. Barandun, C. Hohn, F. Diederich, G. Klebe, *J. Med. Chem.*, **2014**, *57(13)*, 5554–5565

Metadynamics Simulations Reveal Binding of PAR2 to its Trypsin Activated N-terminus

J. Kaindl,[1,2] T. Clark,[2] P. Gmeiner[1]

[1] *Department of Chemistry and Pharmacy, Medicinal Chemistry, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nikolaus-Fiebiger-Str. 10, Erlangen, Germany*

[2] *Department of Chemistry and Pharmacy, Computer-Chemistry-Center, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstr. 25, Erlangen, Germany*

The proteinase-activated receptors (PAR) belong to the class A of G protein-coupled receptor (GPCR) and show an unusual mechanism of receptor activation. In contrast to most GPCR which are activated by diffusible ligands like small molecules, lipids, peptides or proteins, the PARs are activated by proteolytic modification. Here, serine proteases like trypsin or thrombin cleave a part the N-terminus of the receptors exposing a new terminal sequence, subsequently activating the receptor. Despite being associated with diseases like arthritis [1], asthma [2] and cancer [3], the exact pathophysiological role of PAR is not fully understood.

Up to date, four crystal structures of PARs have been resolved, three of which in complex with a small molecule antagonist [4,5] and one in complex with a PAR2-antibody [5], with two of the small molecules binding to allosteric sites of the receptor. What is not yet clear, however, is the binding of the receptor to the proteinase exposed signaling sequence of the N-terminus or other molecules with agonistic properties.

We describe an approach to deduce the stages N-terminus binding to the PAR2 applying metadynamics simulations [6,7]. With this enhanced sampling method, the N-terminus continuously explores new conformations and interactions with the receptor along a given direction. In our case from the core of the receptor to the extracellular side. A bias is introduced to shift the ligand from low energy conformation to low energy conformation. The introduced bias to shift the ligand away from a certain binding mode allows the calculation of absolute binding energies, therefore enabling the deduction of potential stages of N-terminus binding.

[1] W. Ferrell, et al., *J. Clin. Invest.*, **2003**, *111*, 35-41.
[2] N. Asokananthan, et al., *J. Immunol.*, **2002**, *168*, 3577-3585.
[3] L. Ge, et al., *J. Biol. Chem.*, **2004**, *279*, 55419-55424.
[4] C. Zhang, et al., *Nature*, **2012**, *492*, 387-392.
[5] R. K. Y. Cheng, et al., *Nature*, **2017**, *545*, 112-115.
[6] L. Sutto, et al., *Wires. Comput. Mol. Sci.*, **2012**, *2*, 771-779.
[7] N. Saleh, et al., *J. Chem. Inf. Model.*, **2017**, *57*, 1210-1217.

Monte Carlo simulation and thermodynamic integration applied to protein charge transfer

Jan Kaiser, Mike Castellano, David Gnandt, Thorsten Koslowski

Institut für Physikalische Chemie, Universität Freiburg

We introduce a combination of Monte Carlo simulation and thermodynamic integration methods to address a model problem in free energy computations, electron transfer in proteins [1]. The feasibility of this approach is tested using the ferredoxin protein from *Clostridium acidurici*. The results are compared to numerical solutions of the Poisson-Boltzmann equation and data from recent molecular dynamics simulations on charge transfer in a protein complex, the NrfHA nitrite reductase of *Desulfovibrio vulgaris* [2]. Despite the conceptual and computational simplicity of the Monte Carlo approach, the data agree well with those obtained by other methods. A link to experiments is established *via* the cytochrome subunit of the bacterial photoreaction centre of *Rhodopseudomonas viridis*.

[1] J. Kaiser, M. Castellano, D. Gnandt, T. Koslowski, J. Comp. Chem., in press, doi: 10.1002/jcc.26155

[2] G. Gnandt, S. Na, T. Koslowski, Biophys. Chem., **2018**, 241, 1-7

Highly cross-linked epoxy networks under mechanical strain

Mattia Livraghi[1], Christian Wick[1,2], David M. Smith[3] and Ana-S. Smith[1]

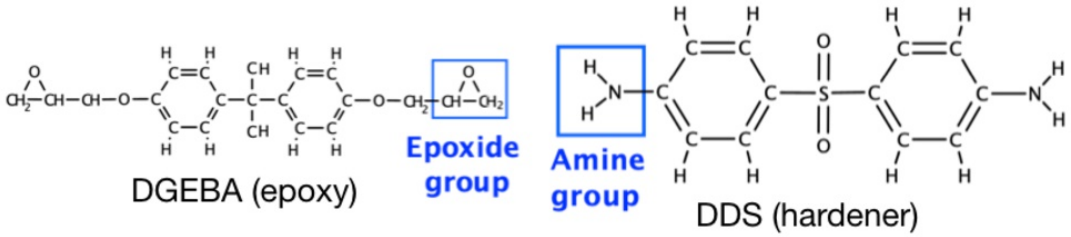
[1]PULS group, Institute of Theoretical Physics, FAU Erlangen-Nürnberg, Germany
[2]Zentralinstitut für Scientific Computing (ZISC), FAU Erlangen-Nürnberg, Germany
[3]Division of Physical Chemistry, Ruđer Bošković Institute, Zagreb, Croatia

Epoxy resins belong to a class of reactive prepolymers containing at least two epoxide groups. Thanks to their tunable properties, epoxy thermosets enjoy a breadth of applications, which shows the importance of understanding how they behave under mechanical stress. In this sense, molecular dynamics (MD) simulations can help cast light on the dynamics triggered by deformation and yield at the molecular scale.

Unfortunately, the MD description of a molecular bond, based on the harmonic potential, allows atoms to remain covalently linked at any value of bond strain, thus being unable to naturally capture bond fracture. Indeed, an accurate description of bond breakage requires quantum mechanical (QM) methods, which suffer, however, even more severe time and length scale limitations than MD does.

For this reason, combined quantum mechanics/molecular mechanics (QM/MM) approaches were developed [1], in which the bond-scission site is treated quantum mechanically, while the rest of the system is described by a much simpler and faster atomistic force-field.

This study develops a new QM/MM scheme, in which over-stretched bonds are identified by using a classical bond-length criterion, and then subject to an ONIOM [2] calculation, to decide whether or not they should break. [3,4] The ONIOM method offers the advantage of accounting for the electrostatic interaction between the QM zone and the surrounding environment, and has already been successful in modeling a number of other systems, both biological and material-based. This method will allow for the accurate modeling of highly cross-linked epoxy networks under mechanical strain.



[1] S. A. Barr, G. S. Kedziora, A. M. Ecker, J. C. Moller, R. J. Berry and Tim D. Breitzman, *J. Chem. Phys.*, **2016**, 144.
[2] T. Vreven, K. Suzie Byun, I. Komàromi, S. Dapprich, J. A. Montgomery, Jr., K. Morokuma and M. J. Frisch, *J. Chem. Theory Comput.*, **2006,** 2, 815-826.
[3] M. Livraghi, C. R. Wick, D. M. Smith, A.-S. Smith, *in preparation*
[4] M. Livraghi, *master thesis*, Integrated Life Sciences, FAU Erlangen-Nürnberg, **2019**

Discovery of histone deacetylase inhibitors and HIV-1 latency-reversing agents by large-scale virtual screening

Donya N. Divsalar[1†], Conrad V. Simoben[2,3†], Cole Schonhofer[1], Khumoekae Richard[1], Wolfgang Sippl[3], Fidele Ntie-Kang[2,3,4*], Ian Tietjen[1,5*]

[1] *Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada*

[2] *Department of Chemistry, University of Buea, Buea, Cameroon*

[3] *Institute of Pharmacy, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany*

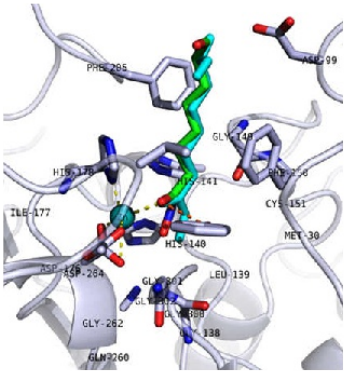
[4] *Department of Informatics and Chemistry, University of Chemistry and Technology Prague, Prague, Czech Republic*

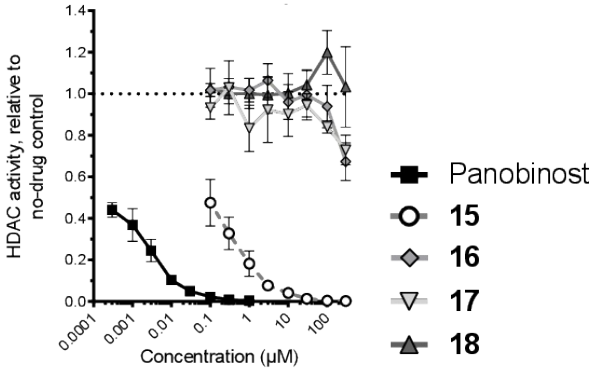
[4] *The Wistar Institute, Philadelphia, PA, USA*

[†] *These authors contributed equally*

**Corresponding author: ntiekfidele@gmail.com*

Current antiretroviral therapies used for HIV management do not target or eliminate latent viral reservoirs in humans. The experimental "shock-and-kill" therapeutic approach involves use of latency-reversal agents (LRAs) that reactivate HIV expression in reservoir-containing cells, followed by infected cell elimination through viral or host immune cytopathic effects.[1,2] Several LRAs that function as histone deacetylase (HDAC) inhibitors are reported to reverse HIV latency in cells and in clinical trials; however, none to date have consistently reduced viral reservoirs in humans, prompting a need to identify new LRAs. Toward this goal, we describe here a virtual screening (VS) approach which uses 14 reported HDAC inhibitors to probe PubChem and identifies 60 LRA candidates.





We then show that 4 screening "hits" including (*S*)-*N*-Hydroxy-4-(3-methyl-2-phenylbutanamido)benzamide (compound **15**), *N*-(4-Aminophenyl)heptanamide (16), *N*-[4-(Heptanoylamino)phenyl]heptanamide (**17**), and 4-(1,3-Dioxo-1*H*-benzo[de]isoquinolin-2(3*H*)-yl)-*N*-(2-hydroxyethyl)butanamide (**18**) inhibit HDAC activity and/or reverse HIV latency *in vitro*.[3] This proof-of-concept study demonstrates that VS-based approaches can readily identify novel HDAC inhibitors and LRAs, which in turn may help to develop new chemical leads to improve shock-and-kill-based HIV eradication efforts.

[1] R.J. Andersen, F. Ntie-Kang, I. Tietjen, *Antiviral Res.*, **2018**, *158*, 63-77.
[2] E. Abner, A. Jordan, A. *Antiviral Res.*, **2019**, *166*, 19-34.
[3] N.M. Archin, D.M. Margolis, *Curr. Opin. Infect. Dis.*, **2014**, *27*, 29-35.

Targeted In Silico Screening of Small Molecule Databases against Breast Cancer

Lalehan Oktay[a,b], Umut Yücel[b], Alpsu Olkan[b] and Serdar Durdağı[b]*

[a]Bioengineering Program, Graduate School of Natural and Applied Sciences, Bahçeşehir University, Istanbul, Turkey

[b]Computational Biology and Molecular Simulations Laboratory, Department of Biophysics, School of Medicine, Bahçeşehir University, Istanbul, Turkey

Around 75% of all breast cancers are either estrogen receptor positive and/or progesterone receptor-positive, and in these cases, estrogen is the major stimulant responsible for the progression of these types of tumors. [1] Therapy approaches for estrogen receptor and progesterone receptor-positive tumors include usage of tamoxifen - an anti-estrogen - that stops estrogen from binding to its receptors. [1] Tamoxifen functions mainly as an antagonist, but multiple data sources imply that tamoxifen also has agonist activity and that, this is the reason for its loss of effectiveness against breast cancer and, therefore, the relapse of the disease. [1] An alternative strategy to achieve estrogen deprivation is to inhibit the aromatization of androgens into estrone or estradiol through the blockage of aromatase. [2] Currently, there are two generations of aromatase inhibitors (AIs). The goal during clinical development of aromatase inhibitors has been to achieve complete estrogen suppression at doses that have no significant toxicity and no nonspecific hormonal effects. [1] Keeping this in mind, a literature check was made, and 28 molecules with known IC_{50} data for MCF-7 cell lines were retrieved from the ChEMBL chemical database and utilized in the PHASE 3D-QSAR application. From this application, a common four-sited pharmacophore hypothesis, AHHH.565, was generated and further expanded into a 3D-QSAR model. This implies that for favorable binding to a pseudo receptor, a ligand should possess a hydrogen-bond acceptor site along with 3 hydrophobic sites. The model was validated using a set of gallic acid derivatives. Using this model, screening of chemical databases (OTAVA, NCI, Peptidomimetic, SPECS, Chembridge, and Enamine) and a total of over 1.28 million molecules was conducted. Top 1000 molecules from each database were chosen according to their fitness to the model, AHHH.565, and were used in the GLIDE/SP (standard precision) mode. Top 100 resulting poses were chosen for extra precision (XP) docking using GLIDE. 10 resulting poses were chosen for Induced Fit Docking (IFD). GOLD docking program was also used for identified hits to further validate the findings. Surviving molecules were tested against 26 toxicity QSAR models and against a universal cancer therapeutic activity QSAR model through the Metacore/Metadrug of Clarivate Analytics. Surviving molecules were put through molecular dynamic (MD) simulations with the aromatase enzyme, and results were compared with known clinically used AIs to validate our screening methodology for the discovery of novel, non-toxic AIs.

[1] S. R. Johnston, M. Dowsett, *NRC*, **(2003)**, *3*, 821 – 831.

[2] A. Carmen, M. J. Carlos, *Medicinal chemistry of anticancer drugs,* **(2008)**, 54 – 90.

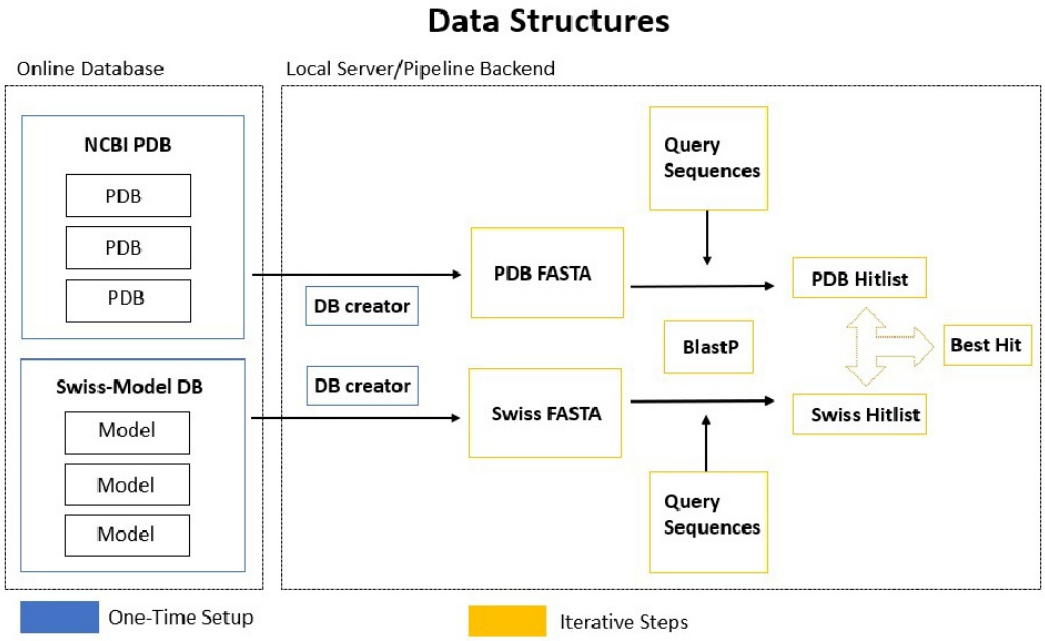
Mind the gap - linking crystal structures and sequences without misrepresentation in antibody research

Simon Schäfer[a,b], Thomas Winkler[a] and Heinrich Sticht[b]

[a]Chair of Genetics and [b]Division of Bioinformatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Protein crystal structure and model databases link structural with sequence information. Numerous crystal structures have unresolved regions (gaps) in their structural coordinates that are not represented in their corresponding sequences. This phenomenon is frequent in the C- and N-terminal regions as well as in solvent exposed loop structures [1]. In protein families with highly homologous regions and very variable loops structures, this mismatch between sequence and structure information is problematic for high throughput analyses.

In antibody structures, this problem affects the complementary determining regions. The NCBI Protein Database (PDB) entries for antibody heavy chains contain only 391 complete chains out of 586 chains in 122 PDB entries. To address the problem of possible misrepresentation for 195 chains containing gaps, we converted the PDB and the Swiss-Model Databases to a sequence data base containing only amino acids resolved in the crystal structures. The resulting database was queried with the BlastP algorithm and antibody reference sequences to find structures with high identity to the query sequences without the risk of obtaining unresolved structural elements [2]. The database utilizes the FASTA format and is compatible to sequence and structural analysis pipelines for any protein family. In addition, the database is offline deployable for closed scientific and clinical networks handling proprietary or patient data.



[1] K. Djinovic-Carugo; O. Carugo, (2015), *Intrinsically disordered proteins* 3 (1), e1095697. DOI: 10.1080/21690707.2015.1095697

[2] C. Camacho et al., (2009), *BMC bioinformatics* 10, p. 421. DOI: 10.1186/1471-2105-10-421

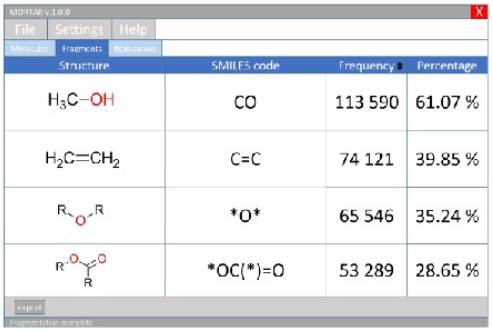
Rule-based *in-silico* Fragmentation for the Analysis of Natural Product Chemical Space

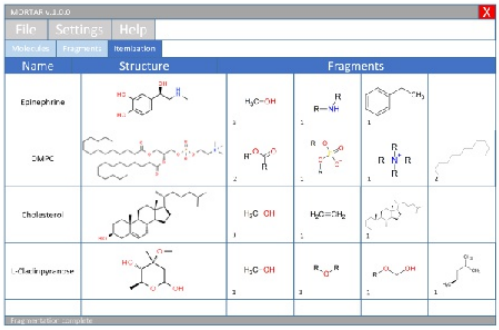
Jonas Schaub[1], Felix Bänsch[2], Achim Zielesny[2], Christoph Steinbeck[1]

[1] *Friedrich-Schiller-University, Institute for Inorganic and Analytical Chemistry, Jena, Germany*

[2] *Westphalian University of Applied Sciences, Institute for Bioinformatics and Chemoinformatics, Recklinghausen, Germany*

The *in-silico* analysis of chemical compound collections based on substructures extracted according to a defined set of rules (fragmentation) is a method used in various areas of cheminformatics. Examples of application are drug design [1], structural classification of compounds [2], and the structural investigation of functional groups [3].
This project aims at the development of an algorithm for extracting characteristic fragments from natural product structures to study their chemical space. The initial idea is to adapt the Ertl algorithm for automatic functional group identification in organic compounds [4,5], thus using functional groups as a basis for the extracted fragments.
To support the refinement of the fragmentation algorithm, a Java rich client application will be developed. The application named *MORTAR* ('MOlecule fRagmenTAtion fRamework') will provide extensive graphical functions for visualizing fragmentation results based on single compounds and entire compound collections:





The ultimate aim of this project is to improve our understanding of chemical motives in natural products in order to create better computer-assisted structure elucidation systems.

[1] G. W. Bemis, M. A. Murcko, *J. Med. Chem.*, **1996**, *39*, 2887-2893.
[2] A. Schuffenhauer, P. Ertl, S. Roggo, S. Wetzel, M. A. Koch, H. Waldmann, *J. Chem. Inf. Model.*, **2007**, *47*, 47–58.
[3] P. Ertl, T. Schuhmann, *J. Nat. Prod.*, **2019**, *82*, 1258–1263.
[4] P. Ertl, *J. Cheminform.*, **2017**, *9*, 36.
[5] S. Fritsch, S. Neumann, J. Schaub, C. Steinbeck, A. Zielesny, *J. Cheminform.*, **2019**, *11*, 37.

Hybridization of β-Adrenergic Agonists and Antagonists Confers G Protein Bias

Maximilian F. Schmidt[1,3], Markus Stanek[1], Jonas M. Kaindl[1,3], Louis-Philippe Picard[2], Harald Hübner[1], Dorothée Weikert[1], Michel Bouvier[2] and Peter Gmeiner[1]

[1]*Department of Chemistry and Pharmacy, Medicinal Chemistry, Friedrich-Alexander-University, Nikolaus-Fiebiger-Str. 10, 91058 Erlangen, Germany.*

[2]*Institute for Research in Immunology and Cancer (IRIC), Biochemistry and Molecular Medicine, University of Montréal, Québec H3C 3J7, Canada.*

[3]*Department of Chemistry and Pharmacy, Computer-Chemistry-Center, Friedrich-Alexander-University, Nägelsbachstraße 25, 91052 Erlangen, Germany.*

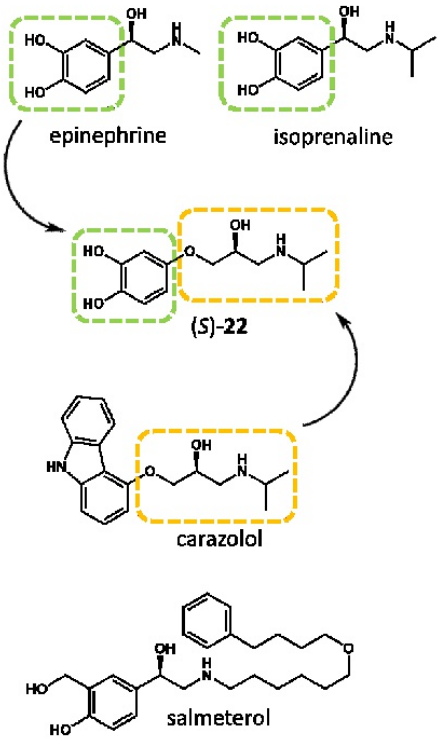
Starting from the β-adrenoceptor agonist isoprenaline and beta-blocker carvedilol, different chemotypes of agonist/antagonist hybrids were synthesized.

Investigations of ligand-mediated receptor activation revealed a predominant effect of the aromatic head group on the intrinsic activity of our ligands, whereas ligands with a carvedilol head group were devoid of agonistic activity.

Ligands composed of a catechol head group and an antagonist-like oxypropylene spacer possess significant intrinsic activity for the activation of $G\alpha_s$, while only showing weak or even no β-arrestin-2 recruitment at both β_1- and β_2-AR.

Unbiased MD simulations were performed to elucidate the binding mode of compound (*S*)-**22** in comparison to the full agonist epinephrine and the partial agonist salmeterol at the β_2-AR.

Thereby we gained insights into the origins of the functionally selective partial agonist activity for this type of catechol-beta blocker hybrid compounds.



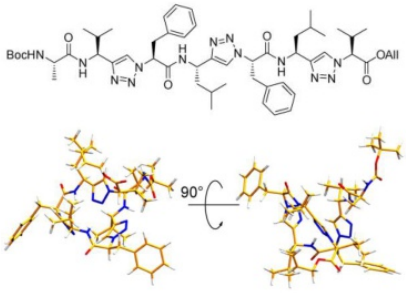
Markus Stanek, Louis-Philippe Picard, Maximilian F. Schmidt, Jonas M. Kaindl, Harald Hübner, Michel Bouvier, Dorothée Weikert, Peter Gmeiner:
Hybridization of β-Adrenergic Agonists and Antagonists Confers G Protein Bias. *Journal of Medicinal Chemistry*, **2019**, 62, 5111-5131.
DOI: 10.1021/acs.jmedchem.9b00349

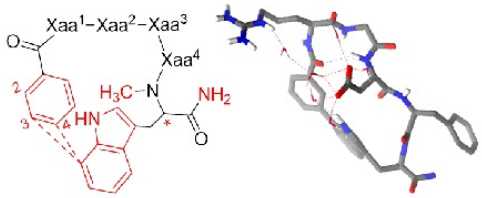
Structure Determination of Peptides and Peptidotriazoles by MD Simulations

David C. Schröder[1], Antoine Marion[2] and Norbert Sewald[1]

[1]*Organic and Bioorganic Chemistry - OCIII, Universität Bielefeld, Germany*

[2]*Department of Chemistry, Middle East Technical University, Ankara, Turkey*





	Xaa$^{1-4}$	*	connection
1	RGDF	D	7-4
2	RGD	L	7-4
3	RGDF	D	7-3

Specific protein-protein interactions are important targets in pharmaceutical research. Therefore, peptides and peptide like small molecules so called peptidomimetics are of special interest. It has been thoroughly shown that 1,4-disubstituted 1*H*-1,2,3-triazoles are potent isosteric replacements for proteolytically labile *trans* peptide bonds. Moreover, the isosteric replacement of the amide bonds can although stabilize or force secondary structure elements. We investigated an MD-based in solution conformational analysis on peptidotriazolamers, hybrid foldamers with features of peptides and triazolamers, containing alternation of amide bonds and 1,4-disubstituted 1*H*-1,2,3-triazoles with conservation of the amino acid side chains. [1] The simulations were based on the specific molecular mechanics force-field parameterization TZLff that we recently published. [2] Conformational properties of a given sequence for a homochiral (Fig. left), a heterochiral and a homochiral equivalent where every second former amino acid has been exchanged by glycine, were analyzed in DMSO as well as in Water. [1]

In a second project, we elucidated the conformation of cyclic peptides containing the RGD (Arg, Gly, Asp; Fig. right) motive being the recognition sequence for integrins. The RGD sequence naturally occurs in extracellular matrix (ECM) proteins, such as vitronectin and fibronectin, playing a crucial role in the bidirectional cell signalling. The integrin ECM interaction turned out to be a promising target in the fields of cancer treatment, radiotherapy and surface coating of implants. In our study, we found an explanation for multiple NMR data sets in DMSO as well the difference in biological activity in aqueous media. [3]

[1] D. C. Schröder et al., *Front Chem*, **2019**, *7*.
[2] A. Marion et al., *J. Chem. Inf. Model.*, **2018**, *58*, 90–110.
[3] I. Kemker, C. Schnepel, D.C. Schröder et al., *J. Med. Chem.*, **2019**, *62*, 7417-7430.

Sampling of Ligand-Induced Conformational Changes in Renin and Factor VIIa

Laura Schulz,[†] Ahmad Reza Mehdipour,[†] and Gerhard Hummer[†,‡]

[†]*Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von Laue Str. 3, 60438 Frankfurt am Main, Germany*

[‡]*Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*

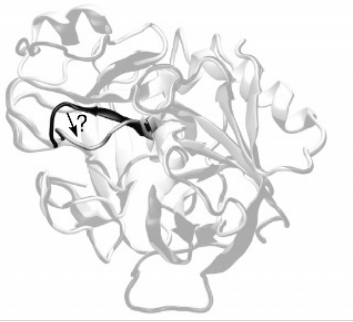
Proteins are dynamic biomacromolecules that can have diverse and nearly isoenergetic conformational states. Ligand binding can shift the equilibrium of this conformational ensemble.

We are studying protein conformational changes upon ligand binding in two different systems: Factor VIIa and Renin. In both systems we know from experiment that a specific group of ligands targeting the orthosteric site can induce conformational changes. The observed conformational changes are the distortion of a beta-sheet structure in the S1 pocket and the opening of a non-functional flap for Factor VIIa and Renin, respectively.

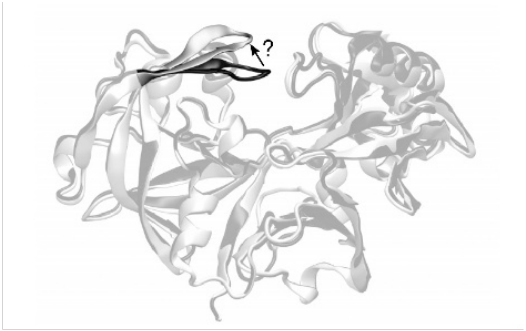
Unbiased MD simulations for Factor VIIa and for Renin (each 4 µs) with removed ligand structures led to contrasting findings. For Factor VIIa the conformational change can be captured with unbiased MD simulations, whereas the transition could not be sampled for Renin within this time scale.

For Factor VIIa starting from the initial beta sheet conformation one or both backbone hydrogen bonds were disrupted leading to the loop-like conformation. Starting with the loop-like conformation the distorted S1 pocket remained distorted.

For Renin starting from the initial closed flap conformation the flap did not open up. Using established biasing protocols for Renin such as increasing the water-protein interaction as established in the group of Gervasio [1] enhanced the protein flexibility, but did not sample the fully open flap. We are working on alternative approaches to disrupt the hydrophobic aromatic cluster holding the Renin flap in place. Starting from the open flap conformation, the open flap of Renin closed again for all eight available X-ray structures.



Factor VIIa:
stable (black) vs. distorted (grey) S1 pocket



Renin:
closed (black) vs. open (grey) flap

[1] Oleinikovas et al., J. Am. Chem. Soc. 2016

Prediction of Odour (PrOdour) – Data Exploration

Slavica Subic[1], Veniamin Morgenshtern[2], Thilo Bauer[3, 4]

[1]Advanced Signal Processing and Communications Engineering, Elite Master's Study Programme, FAU Erlangen-Nürnberg
[2]Chair of Multimedia Communications and Signal Processing, Dept. Electrical, Electronic and Communication Engineering, FAU Erlangen-Nürnberg
[3]Computer Chemistry Centre (CCC), Department of Chemistry and Pharmacy, FAU Erlangen-Nürnberg
[4]Fraunhofer Institute for Process Engineering and Packaging IVV

Where exactly does the odour come from? What causes a substance to produce a certain scent?

One could approach answering this question in many ways, and one way would be to look at the language used to represent the molecules. Structural formulae are exactly that, and that is where this data exploration task starts. Looking at the language of chemistry, and individual words in its dictionary, we search for the ones most responsible for particular odours.

Starting from molecular formulae, parsing them into millions of words (features, sub-molecular structures), and looking for largest overlapping structures, we collect a dictionary of molecular features. With data collected from experiments where experts evaluated the sensory properties of substances, it was possible to build a data set containing the applicability of a list of odour descriptors to the substances used in the experiment. Using LASSO (Least Absolute Shrinkage and Selection Operator) regression regularization for feature extraction, we aim to find the features most responsible for the applicability scores by which descriptors were rated in the experiment - in other words, features most responsible for the expression of a certain odour. By grouping descriptors with clustering techniques, we can explore similarities of features among the groups and search for chemically meaningful features that define a descriptor cluster.

[1] A. Cadeddu, E. K. Wylie, J. Jurczak, M. Wampler-Doty, and B. A. Grzybowski - Organic Chemistry as a Language and the Implications of Chemical Linguistics for Structural and Retrosynthetic Analyses, *Angew. Chem.* **2014**, 8246-8250, DOI: 10.1002/ange.201403708
[2] T. Hastie, R. Tibshirani, J. Friedman – The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, Second Edition, corrected 12th printing - 13. Jan. **2017**.

Pharmacophore-based Virtual Screening to Identify New β3-Adrenergic Receptor Agonists

Navista Sri Octa Ujiantari[a,b,c], Seungmin Ham[d], Dana Sabine Hutchinson[d], Daniela Schuster[a,c]
[a]*Institute of Pharmacy, Department of Pharmaceutical Chemistry, University of Innsbruck, Innrain 80-82/IV,Innsbruck, 6020, Austria;*
[b]*Faculty of Pharmacy, Universitas Gadjah Mada, Sekip Utara, Yogyakarta, 55281, Indonesia;*
[c]*Institute of Pharmacy, Department of Pharmaceutical and Medicinal Chemistry, Paracelsus Medical University, Salzburg, 5020, Austria;*
[d]*Drug Discovery Biology, Monash Institute of Pharmaceutical Sciences, Monash University, Victoria, 3052, Australia*

β3-Adrenergic receptors (β3-ARs), as well as β1-ARs and β2-ARs, belong to the G-protein coupled receptors (GPCRs). Activation of these receptors leads to thermogenesis and lipolysis in adipose tissues [1]. Only a few β3-AR agonists are available for clinical treatment such as mirabegron indicated for overactive bladder (OAB) [2]. Drug design and discovery using pharmacophore-based virtual screening has been applied to discover new candidate compounds [3]. Pharmacophore modeling on β3-ARs was conducted yet with no experimentally confirmation [4]. This study was aimed to generate ligand-based pharmacophore models to identify new compounds as β3-ARs agonists and validate those hits also experimentally.

Various β3-ARs agonists were collected as a dataset to build the pharmacophore model. The models were established using LigandScout v3.12 (Inte:Ligand GmbH, Vienna). The selected model was used for virtual screening against 3 commercial compound databases. To select the candidate compounds, the screening results were filtered by the physicochemical parameters derived from highly active β3-ARs agonists and a docking evaluation. The physicochemical properties were calculated with Datawarrior v4.7.2. The docking was done with GOLD v5.7.0 employing a homology model of β3-AR. To confirm the activity of the candidate compounds, the in vitro assay was performed.

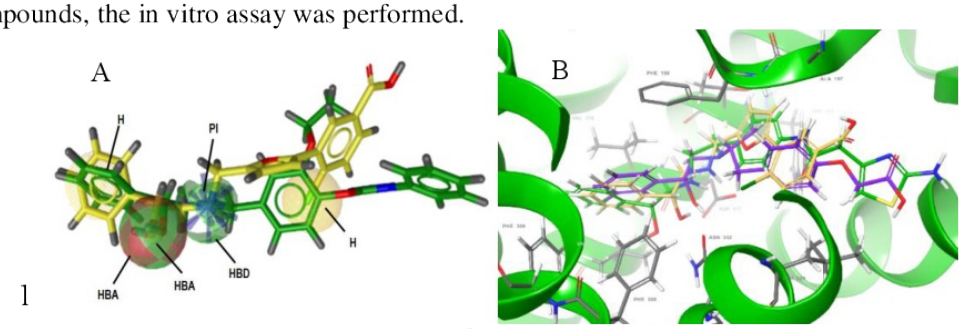


Figure 1. A: Ligand-based pharmacophore model for β3-AR consisting of both H (Hydrophobic interaction) on the hydroxy end and amine end; HBD (Hydrogen Bond Donor), HBA (Hydrogen Bond Acceptor), and PI (Positive Ionizable Area) on the center. B: Docking pose of Nav16 (yellow), BRL37344 (purple,) and mirabegron (green).

Two out of 20 tested compounds, Nav16 and Nav19 were found to be active both in CHO cell lines expressing human and mouse β3-ARs. Both compounds increased the cAMP level with EC_{50}s of 12.6 µM and 21.7 µM in CHO-hβ3AR cells and 3.43 µM and 2.14 µM in CHO-mβ3AR cells, respectively. Figure 1A depicts both compounds fitted perfectly to the pharmacophore features. Figure 1B shows the similarity of Nav16 docking pose with BRL37344 (selective β3AR agonist) and mirabegron bound to the receptor.

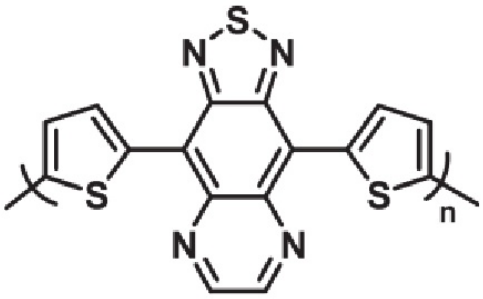
[1] S.D. Edmondson, *Elsevier*, **2017**, Vol.7.20, 714-737.
[2] J. Gras, *Drugs Today (Barc)*, **2012**, 48, 25-32.
[3] D. Schuster, C. Laggner, T.M. Steindl, A. Palusczak, R.W. Hartmann, T. Langer, *J Chem Inf Model*, **2006**, 46, 1301-11.
[4] F. Jin, C. Lu, X. Sun, W. Li, G. Liu, Y. Tang, *Mol Divers*, **2011**, 15, 817–831

Design of Thiophene and Thiadiazoloquinoxaline Containing Semiconducting Polymer

Emire Uyanık, Hakan Kayı*

Computational Chemistry Laboratory, Chemical Engineering Department,
Ankara University, Tandoğan, 06100, Ankara, Turkey
E-mail: euyanik@ankara.edu.tr, *hakan.kayi@ankara.edu.tr

Quinoxaline is a compound known as aromatic heterocyclic that has a benzene ring joined to a pyrazine group [1]. Quinoxaline itself is of scientific interest, and the quinoxaline derivatives are used in pharmaceuticals, dyes and antibiotics. Also, polymers containing quinoxaline have lower band gaps. Quinoxaline is used as electron acceptor unit (A) in semiconducting polymer systems. In this study, donor-acceptor-donor (D-A-D) type conjugated polymer based on quinoxaline acceptor was designed and electronic properties were calculated. The most stable structures of the monomer and oligomers, consisting of thiophene (D) and thiadiazoloquinoxaline (A) groups were determined. HOMO-LUMO energies of the structures were obtained by using B3LYP/6-31G (d) and B3LYP/LANL2DZ calculations without and with the inclusion of the PCM. Electronic band gap (Eg) values of all structures were obtained from the difference of HOMO and LUMO energies. Eg value of the polymer was found by extrapolation and linear fitting methodology. The results obtained from different calculation levels were compared with the experimental results.



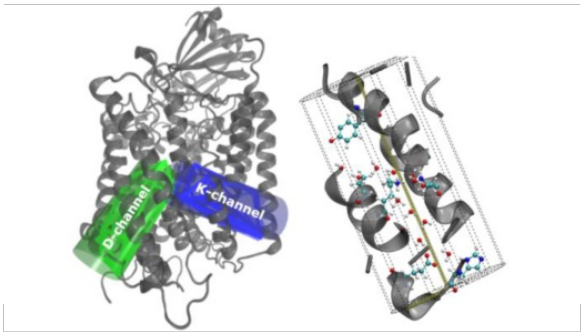
[1] S. Achelle, C. Baudequin, N. Plé, *Dyes and Pigments*, **2013**, 98, 575-600

Acknowledgements: Authors thank Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting this study through the project number 117Z354.

Protonation dynamics in Cytochrome c Oxidase

Vincent Stegmaier[a], Senta Volkenandt [b], Marcus Trautmann [b], Anne Hartmann [b], René Gorriz [b], and Petra Imhof [c]

[a] *Technische Universität Berlin, Germany;* [b] *Freie Universität Berlin, Germany;* [c] *University of Stavanger, Norway*



Cytochrome c oxidase, also called complex IV, is a membrane protein in the respiratory chain. It uses the energy gained from the reduction of molecular oxygen with protons to water to pump extra protons through the membrane. The transport of the protons to the reaction centre takes place through two so-called channels, the D- and K-channel, named after an important aspartate (D132) and lysine (K362) residue, respectively. Whereas in the D-channel crystal structures reveal a number of water molecules along which a proton transport via a Grotthuss mechanism appears to be feasible [1], the situation is more complicated in the K-channel. MD simulations reveal that the presence of a proton in the middle of the K-channel, either located at the naming residue K362 or at a water molecule in its vicinity, result in a significantly more hydrated K-channel than in cases where such an excess proton is located at the channel entrance or not present. In addition, the conformational dynamics of important residues in the K-channel also depend on the protonation state, not only of the respective residues themselves but also on the other residues in the channel. Protein residue conformation, hydration level, and width of the channel show a correlation that suggests a higher conformational flexibility of the K362 side chain in its protonated form, i.e. in the presence of more water molecules [2]. Both factors facilitate the transport of a proton through the channel, either carried by a water molecule as a hydronium ion or via the K362 residue. Simulations evaluating the probabilities of proton positions at various heights in the channel suggest an "upward" movement, that is towards the redox centre, to be more probable than a "downwards" movement in which the region around K362 can be understood as a watershed, leading to "up" or "down" movement to which K362 itself contributes in a gating manner, however, further regulated by the protonation state and thus conformation state of the other residues and consequently also the hydration level in the channel.

[1] Ghane, T., Gorriz, R. F., Wrzalek, S., Volkenandt, S., Dalatieh, F., Reidelbach, M., & Imhof, P. Hydrogen-bonded network and water dynamics in the D-channel of cytochrome c oxidase, *The Journal of membrane biology*, **2018**, *251(3)*, 299-314.
[2] Helabad, M. B., Ghane, T., Reidelbach, M., Woelke, A. L., Knapp, E. W., & Imhof, P. Protonation-state-dependent communication in cytochrome c oxidase. *Biophysical journal*, **2017**, *113(4)*, 817-828.

Dynamics of crowded macromolecules from atomistic simulations

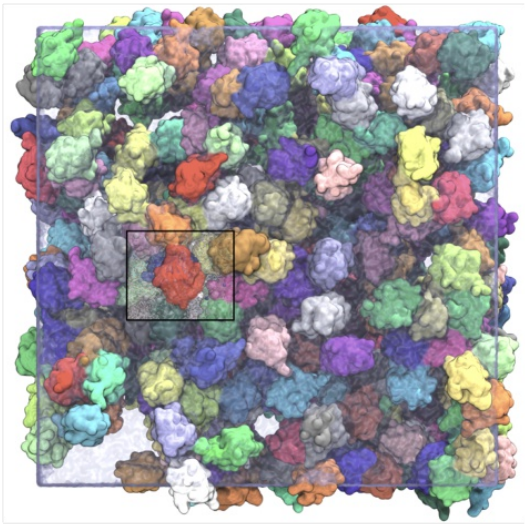
Sören v. Bülow,[†] Marc Siggel,[†] Max Linke[†] and Gerhard Hummer[†,‡]

[†]*Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue-Str. 3, 60438 Frankfurt am Main, Germany*

[‡]*Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*

The interior of the cell is densely crowded, affecting thermodynamic and kinetic properties of the macromolecules within. It has only very recently become computationally feasible to simulate crowded macromolecular systems atomistically to address detailed questions on protein stability, interactions and transport properties like diffusion.

I present a large-scale molecular dynamics simulation study of atomistically-resolved crowded protein systems containing up to 540 fully flexible proteins with 3.6 million atoms in the microsecond range. [1] We find that using the Amber99SB*-ILDN-Q protein force-field in conjunction with the TIP4P-D water model gives a surprisingly accurate picture of the dynamics of concentrated protein solutions that agrees very well with experimental results.



Snapshot of an atomistic molecular dynamics simulation of 540 fully flexible ubiquitin proteins at 200 mg/ml concentration in explicit water.

The protein species studied here form dynamic clusters between which they constantly exchange. A theoretical model, based on the Stokes-Einstein equations, nearly quantitatively links the slow-down of translational and rotational diffusion, increase in viscosity and formation of protein clusters.

Using a colloidal model of sticky hard spheres, we relate the above properties to the dissociation constant of protein-protein interaction, allowing for an estimate of the concentration-dependent slow-down of protein diffusion given only the dissociation constant.

We build on the above framework to simulate mixed systems to address research questions like the postulated solubilizing properties of ATP (acting as a "hydrotrope"), phase-separation of intrinsically disordered proteins, or interactions between RNA and transcription factors.

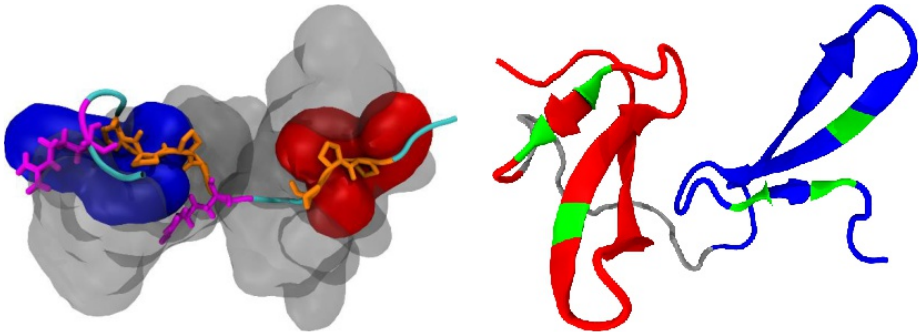
Reference:

[1] Sören von Bülow, Marc Siggel, Max Linke, Gerhard Hummer (2019) Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proc Natl Acad Sci* 116:9843-9852.

In search of binding competent structures of the tWW domain using classical molecular dynamics simulations

M. Wenz, S. Aleksić, B.G. Keller

Department of Biology, Chemistry and Pharmacy
Institute for Chemistry and Biochemistry,
Freie Universität Berlin,



WW domains, protein modules consisting of 35 to 40 amino acids, are known for mediating protein-protein interactions (PPIs) through recognition of proline-rich motifs (PRMs).[1-2] Their name is derived from two highly conserved tryptophanes (W) that have two tasks: first, stabilizing the main structure of the domain, a triple-stranded, antiparallel β −sheet, and second, forming the binding site together with serine (S) and tyrosine (Y).[3] WW domains can be found in many signalling and regulatory proteins such as the formin-binding protein 21 (FBP21).[4]

FBP21 contains two successive WW domains, forming the **tWW domain**. It is directly involved in the splicing process of the RNA as it can bind to the small nuclear ribonucleoprotein-associated protein SmB/B' that arises in four of the five major components of the spliceosome (small nuclear ribonucleoproteins).[5-6] This binding process can be traced back to the tWW domains of FBP21 recognizing a proline-rich sequence (PRS) that is formed by two PRMs of SmB/B'.[7-8]

In this work, the investigation of the apo-tWW domain with respect to the identification of possible, binding-competent structures is presented. Based on all-atom molecular dynamics simulations (MD simulations) and the density-based common-nearest neighbor algorithm [9-10], we show that it is possible to extract binding-promising metastable states out of the conformational ensemble of the tWW domain. Using the structure of a specific, literature-known PRS of SmB/B'[11], we performed docking experiments with HADDOCK protocol[12] for the most promising tWW structures. The docking results are in accordance to literature regarding the binding relevant residues.[6, 8] Ensuing MD simulations confirmed the stability of the obtained protein-protein complexes. Due to this, we can assume the tWW structures involved to be representative binding-competent structures. Running MD simulations of the protein-protein complexes, we also examined the influence of mutations within the PRS on the stability of the entire complex.

[1] R. J. Ingham, K. Colwill, C. Howard, S. Dettwiler, C. S. Lim, J. Yu, K. Hersi, J. Raaijmakers, G. Gish, G. Mbamalu, *Mol. Cell. Biol.* **2005**, *25*, 7092-7106.
[2] M. Sudol, *Prog. Biophys. Mol. Biol.* **1996**, *65*, 113-132.
[3] L. J. Ball, R. Kühne, J. Schneider-Mergener, H. Oschkinat, *Angew. Chem. Int. Ed.* **2005**, *44*, 2852-2869.
[4] M. J. Macias, M. Hyvönen, E. Baraldi, J. Schultz, M. Sudol, M. Saraste, H. Oschkinat, *Nature* **1996**, *382*, 646-649.
[5] A. K. Leung, K. Nagai, J. Li, *Nature* **2011**, *473*, 536-539.
[6] S. Klippel, M. Wieczorek, M. Schümann, E. Krause, B. Marg, T. Seidel, T. Meyer, E.-W. Knapp, C. Freund, *J. Biol. Chem.* **2011**, *286*, 38478-38487.
[7] M. T. Bedford, R. Reed, P. Leder, *Proc. Natl. Acad. Sci.* **1998**, *95*, 10602-10607.
[8] X. Huang, M. Beullens, J. Zhang, Y. Zhou, E. Nicolaescu, B. Lesage, Q. Hu, J. Wu, M. Bollen, Y. Shi, *J. Biol. Chem.* **2009**, *284*, 25375-25387.
[9] B. G. Keller, X. Daura, W. F. van Gunsteren, *J. Chem. Phys.* **2010**, *132*, 02B610.
[10] O. Lemke, B. G. Keller, *J. Chem. Phys.* **2016**, *145*, 164104.
[11] D. Kosslick, PhD thesis, Freie Universität Berlin (Berlin), **2012**.
[12] C. Dominguez, R. Boelens, A. M. Bonvin, *J. Am. Chem. Soc.* **2003**, *125*, 1731-1737.

Identification of conformational states and transitions in between from molecular dynamics trajectories and vibrational signatures

Sandro Wrzalek, Irtaza Hassan, Petra Imhof

Institute of Theoretical Physics, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

We used the Markov state modeling (MSM) technique to identify meta-stable conformations of biomolecules such as small peptides and the time scales associated with the slowest processes sampled in trajectories from classical molecular dynamics (MD) simulations. The vibrational spectra of representative conformations are calculated from the first-principle MD simulations in solution at finite temperature in the frequency range that is most sensitive to the peptide conformation, so-called amide bands (1300–1800 cm-1), see the example of the small floppy peptide Alanine-Leucine (AL)[1]. This combined approach to analyse the vibrational signature is extended to Alanine-Leucine-Alanine and Alanine-Leucine-Alanine-Leucine (ALAL). Generally, all computed spectra show two prominent bands which are assigned to the stretch vibrations of the carbonyl and carboxyl group, respectively. Variations in bandwidths and exact maxima are likely due to small fluctuations in the backbone torsion angles. The detailed vibrational spectroscopic analysis of the most probable, beta sheet-like conformations of the selected peptides (i.e., AL, ALA, ALAL) is performed with the time-frequency analysis based on the wavelet transform using the trajectories from the first principle MD simulations. Particularly, we analysed how the instantaneous frequencies of carbonyl groups (C=O) present in the peptides are affected by the local solvent environment. It is clear from the wavelet analysis of carbonyl bonds that due to change in the hydrogen bonding state or simply due to change in the local solvent environment, the state of the carbonyl bond changes which leads to change in its instantaneous frequency.

Although, MSMs work great to identify the meta-stable conformations for small molecules like peptides, they can become unreliable for molecules with more degrees of freedom, like DNA and RNA, since Markov models are, in conjunction with the demand for large data sets, highly sensitive to the preprocessing steps. Especially, commonly used coordinate projection methods like Principal Component Analysis (PCA) and Time-lagged independent component analysis (TICA) are risky: First, the state space may not be reducible to a set of only a few basis vectors leading to important information being lost (or still too many dimensions). The latter becomes difficult for clustering the data set. PCA is furthermore problematic since it rotates and reduces the space based on the variance, so it likely removes dimensions containing the slowest (least sampled=lowest variance) processes. Our approach tackles this problem by first reducing dimensions of different type (Cartesian, curvlinear, binary) within their space and combine them later with the Multi-View clustering method proposed in [2]. Our approach for polar coordinates defines two time windows with length l, separated by a constant lag time τ. Those windows move through a trajectory, while comparing their distributions, via Kullback–Leibler divergence, with each other. If the distance of two distributions exceeds a defined threshold, κ, a transition in the time frame $2l+\tau$ will be considered. A possible new state (pns) is defined as the part of the trajectory, after and before a new transition. A pns has to be distinguished from a repeated state to qualify for a new state. That distinction of states can be considered in a clustering manner: To validate if the pns is a truly new state, its distribution will be compared to that of the already known states - if the distribution is different the pns will become a new state, else the part of the trajectory will be assigned to the closest new state. Data-sensitive clustering methods like (H)DBScan can be helpful. Binary coordinates can be reduced by using auto encoders (Neural Networks) to encode those binary vectors in fewer dimensions.

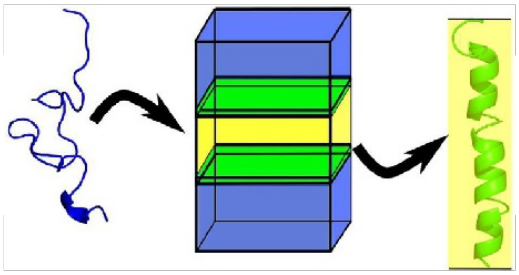
[1] I. Hassan, L. Donati, T. Stensitzki, B. Keller, K. Heyne, and P. Imhof, CPL, 2018 698:227–233.
[2] S. Kanaan-Izquierdo, PhD thesis - Universitat Politecnica de Catalunya, 2017, Multiview pattern recognition methods for data visualization, embedding and clustering

Simulations of the membrane proteins in the lipid bilayer with the UNRES coarse-grained force field

Karolina Zięba[a], Magdalena Ślusarz[a], Rafał Ślusarz[a], Adam Liwo[a,b], Cezary Czaplewski[a], Adam K. Sieradzan[a]

[a]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland

[b]School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, 130-722, Seoul, Republic of Korea



UNRES is a physics-based force field, which is a tool for a protein modeling, can be used for structure prediction and study of the protein pathway folding [1]. This model is especially dedicated to simulate proteins in water environment and, therefore, membrane proteins could not be treated. Membrane proteins are immersed in a lipid bilayer and have only partial contact with the water environment [2]. Because membrane proteins are very important for cell functioning, we recently extended the UNRES model to include the lipid bilayer as a continuous phase, with a transition zone between the lipid and water phase that corresponds to the lipid head groups. We implemented the periodic-box scheme, which is used in UNRES in our earlier work [3]. The lipid bilayer has been modeled by introducing a continuous nonpolar phase with the water-interface region of appropriate thickness. The potentials for average electrostatic and correlation interactions of the peptide groups have been rescaled to account for the reduction of the dielectric permittivity compared to the water phase and new potentials for protein side-chain–side-chain interactions inside and across the lipid phase have been introduced. The model was implemented in the UNRES package for coarse-grained simulations of proteins, and the package with the new functionality was tested for total energy conservation and thermostat behavior in microcanonical and canonical molecular dynamics simulations runs, respectively. The efficiency of the new force field was estimated by using 10 short α-helical membrane proteins with low similarity [4], which were simulated starting from extended structures. The extended UNRES force field was able to predict correctly the overall folds of the membrane proteins studied.

[1] C. Czaplewski, A. S. Karczyńska, A. K. Sieradzan, A. Liwo, *Nucl Acid Res*, **2018**, 46:W304-W309.
[2] M. Cieplak, A. Sienkiewicz, *Encyklopedia Fizyki Współczesnej*, 2004.
[3] A. K. Sieradzan, P. Krupa, H. A. Scheraga, A. Liwo, C. Czaplewski, *J Chem Theory Comput*, **2015**, 11, 817–831.
[4] W. Pulawski, M. Jamroz, M. Kolinski, A. Koliński, S. Kmiecik, *J. Chem. Inf. Model.*, **2016**, 56:2207-2215

Acknowledgements
This work was supported by the National Science Center (Poland) UMO-2017/27/B/ST4/00926. Computational resources were also provided by a) the supercomputer resources at the Informatics Center of Metropolitan Academic Network (IC MAN) in Gdańsk, and b) computational resources at Interdisciplinary Center of Mathematical and Computer Modeling in Warsaw (ICM), grant GA71-21, c) Polish Grid Infrastructure (PL-GRID) d) our 682-prcessor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk. Additional source of financing was NAWA project "International scholarship exchange of PhD candidates and academic staff", project No. POWR.03.03.00-IP.08-00-P13/18, implemented under Measure: 3.3 Internationalisation of Polish higher education, OP KED; Number of the project financing agreement PPI/PRO/2019/1/00005/U/00001.

Sponsor Pages

Boehringer Ingelheim | Value through Innovation

www.boehringer-ingelheim.com

OUR COMPANY

Boehringer Ingelheim is one of the world's 20 leading pharmaceutical companies.

Founded in 1885 in Ingelheim and family-owned to this day

~50,000 employees worldwide

FOCUS ON

- Human pharmaceuticals
- Animal health
- Biopharma-ceuticals

DIVERSITY & INCLUSION DRIVE INNOVATION

48% women*

36% female leaders*

Innovation / growth

- 14% BABY BOOMERS (1946–1964)*
- 38% GENERATION Y (1981–2000)*
- 48% GENERATION X (1965–1980)*

*Boehringer Ingelheim worldwide, 2018

DIVERSITY IS NOT ONLY ABOUT

- Minorities
- Individuals
- Compliance / ethics

DIVERSITY IS ALSO ABOUT

- Balance
- Team composition
- Innovation / growth

SUCCESSFUL BUSINESSES

HUMAN PHARMACEUTICALS

EUR **12,559 million** TOTAL NET SALES IN 2018

FOCUS ON

- Respiratory diseases
- Cardiometabolic diseases
- Central nervous system + immunology
- Oncology

ANIMAL HEALTH

EUR **3,960 million** TOTAL NET SALES IN 2018

#2 GLOBAL PLAYER

SHARE

- 56% COMPANION ANIMALS
- 43% LIVESTOCK
- 1% OTHERS

- 51% PETS
- 14% CATTLE
- 16% SWINE
- 11% POULTRY
- 5% EQUINE
- 2% VPH*
- 1% OTHERS

*Veterinary Public Health (VPH)

BIOPHARMACEUTICAL CONTRACT MANUFACTURING

Boehringer Ingelheim is a world-leading contract manufacturer of biopharmaceuticals.

We operate **1** GLOBAL NETWORK of biopharmaceutical production sites in Europe, the USA and Asia

We are a pioneer in biopharmaceuticals WITH OVER **35** YEARS of experience

EUR **734 million** TOTAL NET SALES 2018

We have manufactured MORE THAN **30** BIOPHARMACEUTICALS for global markets

SUSTAINABLE ECONOMIC & SCIENTIFIC ACTIVITY

ECONOMIC SUCCESS (2018)

EUR **17,498 million** NET SALES

19.8% RETURN ON NET SALES

AMERICAS	EUROPE	ASIA, AUSTRALIA, AFRICA (AAA)
EUR 8,088 million	EUR 5,316 million	EUR 4,094 million

RESEARCH & DEVELOPMENT

3 PILLARS

- Focus on research therapeutic areas
- Scientific platforms
- Research Beyond Borders

EUR **2,780 million** Human Pharmaceuticals expenditure in research & development in 2018

≙ around **22.1%** of Human Pharmaceuticals net sales 2018

SOCIAL RESPONSIBILITY

MAKING MORE HEALTH

ASHOKA

Making More Health is a global initiative by Boehringer Ingelheim in cooperation with Ashoka, a non-governmental organisation. Its aim is to improve healthcare for people worldwide.

MORE THAN **90** social entrepreneurs in over 30 countries received support since 2010

APPROX. **5,500** Boehringer Ingelheim employees worldwide are involved in the initiative

- Orion is OpenEye's reimagining of all the elements of CADD conducted entirely within a cloud service, in our case Amazon Web Services (AWS), delivered as either a hosted system or an in-house VPC solution
- As a 'cloud native' platform Orion completely automates and manages access to large scale AWS processing and storage
- In-cloud facilities include molecular design, 3D visualization, data analysis, results/method sharing and project organization
- All of OpenEye science is included, enabling users to construct innovative workflows with Floe, our pipelining tool
- As an open platform Orion allows for straightforward integration of third-party code (customer, academic, vendor)
- Interaction with Orion is via a simple webpage, deliverable on any internet-enabled device

Interact with Orion via web browser on any internet-enabled device.

OpenEye has built a solid reputation as a scientific leader in the field of molecular design based on two decades of delivering useful applications and programming toolkits. Our scientific approach has focussed on the power of molecular 3D structure to inform and guide, in particular via the concept of shape similarity. We have changed industry perception of what is possible with the speed, robustness and scalability of our tools and have recently built these into a ground-up, cloud-native platform, Orion. Combining unlimited computation and storage with powerful tools for data sharing, visualization and analysis in an open development platform, Orion offers unprecedented capabilities for drug discovery and optimization.

OpenEye SCIENTIFIC

www.eyesopen.com


Extraordinary
medicine requires extraordinary science.
At Roche, we are committed to both. Discovering and developing innovative medicines and diagnostic tests to help people live better, longer lives.
Roche


SANOFI
Empowering Life
Life is a health journey, with ups and downs, which can be big or small, lifelong or momentary.
We, at Sanofi, are there for those challenges, by the side of people, as a health journey partner.
Sanofi is about Empowering Life.
www.sanofi.de

Learn State-of-the-Art Chemistry Software with Schrödinger's Tools

Schrödinger's software solutions allow researchers and educators to create testable hypotheses and illuminate otherwise unattainable mechanistic insights in life and materials sciences. We train scientists worldwide to harness the latest in cutting-edge tools in computational chemistry through a wide array of training avenues.

Tailored Trainings in Science and Education
Basic and advanced training with Schrödinger scientists can be organized locally on-site or virtually, with curriculum tailored to the specific needs of each academic team or institution.

Application-Focused Workshops
Regular day-long workshops are available for both Drug Discovery and Materials solution workflows across Europe. Contact education@schrodinger.com for further information.

The Schrödinger Online Course
This self-paced online course provides practical applications of computational modeling using active learning strategies and multiple types of content. Enroll at schrodinger.com/modeling.

Web Tutorials & Videos
Access a comprehensive library of videos and step-by-step online training materials on how to use Schrödinger tools at schrodinger.com/training.

Access to Teaching Material
Interested in learning more about Schrödinger solutions education or training opportunities? Send an email to education@schrodinger.com

Request a trial:

SCHRÖDINGER

Get started at www.schrodinger.com/training

PLEASE ACKNOWLEDGE THE GENEROUS DONATIONS FROM OUR SPONSORS

(alphabetical order)

Boehringer-Ingelheim	**www.boehringer-ingelheim.de**
OpenEye	**www.eyesopen.com**
Roche	**www.roche.de**
Sanofi	**www.sanofi.com**
Schrödinger	**www.schrodinger.com**
Wiley	**www.wiley.com**

CHALLENGE YOUR MATH SKILLS

1, How many possibilities exist to obtain the number 447 as the sum of two or more successive odd positive numbers?

2, How many ways are there to obtain the number 100 as the sum of two or more consecutive integers?

IMPRINT

Publisher:	Molecular Modelling & Graphics Society - Deutschsprachige Sektion MGMS-DS e.V. Computer-Chemie-Centrum, Nägelsbachstr. 25, 91052 Erlangen, Germany
Tel.:	+49 9131 - 85 26582
Fax:	+49 9131 - 85 26565
E-Mail:	info@mgms-ds.de
WWW:	www.mgms-ds.de
V.i.S.d.P.:	PD Dr. Harald Lanig
Layout:	Dr. Anselm Horn *via* SCRIBUS (www.scribus.net)
Cover design:	Dr. Christian Wick *via* GIMP (www.gimp.org)
Cover motif:	RNA aptamer in complex with fluorophore and FAB (PDB-ID: 4kzd)
Math Challenge:	Marcus Conrad
Printed by:	DRUCKLADEN, Erlangen (www.druckladen.de)
Circulation:	120 copies

WE THANK THIS YEAR'S SPONSORS!

molecular graphics modelling society

ZISC Zentralinstitut SCIENTIFIC COMPUTING

C3 Computer Chemie Centrum

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG