Characteristics of main research directions investigated at the institute and the achievements 2010−2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

The scientific mission of IOCB is exploratory and curiosity-driven research of fundamental chemical and biological processes at the molecular level but there is also a strong record in applied research (in particular in drug development). The core scientific activities are centered in areas of chemical biology and medicinal chemistry. Specifically, considerable part of the research is focused on the identification and characterization of potential targets and biologically active compounds for therapeutic intervention in human and veterinary diseases. This includes the design, synthesis, and testing of new ligands and inhibitors, characterization of their metabolism, the mechanisms of their action, and relationships between their structure and activity. Another main research interest of the IOCB is biologically orientated chemistry including the study of the chemical nature of various natural systems and models, the use of small molecules or biomacromolecules to map and control biological processes, syntheses of artificial 'bioanalogues', etc. Progress in these research areas would be inconceivable without a strong support of the fundamental research in the areas of organic synthesis, catalysis, design of new materials, the development of modern physical chemical and analytical methods, structural biology, and computational and theoretical chemistry. The 2010-2014 IOCB research highlights in various scientific disciplines are then summarized below including a few representative figures for the major achievements.

Concerning bibliographic indicators in the 2010-2014 period it can be mentioned that the scientists at IOCB published 250-300 papers annually (in total 1448 papers). Approximately 75% of these contributions originate in the research carried out (almost fully) at IOCB: the first and/or corresponding author had IOCB affiliation. The number of citations is increasing (currently, it is at the annual rate of >10 000 citations). Major emphasis is given to publishing high-quality and innovative work rather than large number of (only) standard-quality papers. IOCB contributions were in 2010-2014 published in the following journals (number of papers): *Nature* (2), *Science* (1), *Nat. Chem.* (3), *Nat. Struct. Mol. Biol.* (3), *PNAS* (4), *J. Am. Chem. Soc.* (24), *Angew. Chemie Int. Ed.* (21), *Chem. Rev.* (4), *EMBO J.* (2), *EMBO Rep.* (1), *Chem. Comm.* (14), *Chem. Eur. J.* (33), *Nucl. Acid Res.* (11), *Cancer Res.* (1), *Org. Lett.* (6), *Nanoscale* (3), *Small* (1), *J. Phys. Chem. Lett*. (23), *J. Chem. Theor. Comput.* (46), *J. Org. Chem.* (24), *J. Med. Chem*. (20) and others which in total account for approximately 60-70 papers (*per annum*) in the journals with the impact factor greater than 5 (again, approx. ¾ of these papers reported the core research done at IOCB). Besides the growing number of publications in high-impact journals, the success of the institute is also illustrated by the two *ERC Advanced grants* awarded to IOCB scientists in 2009 and 2010.

Organic chemistry and synthesis

The research in organic chemistry covered a broad spectrum of topics including general phenomena i.e. stereochemistry and properties of extended helical aromatics as well as development of novel synthetic approaches and methodologies.

The Starý group focused on the synthesis and properties of new π-electron systems, which are attractive for enantioselective catalysis, supramolecular chemistry and molecular electronics. In particular, the attention has been paid to helically chiral aromatics (helicenes) that are enantiopure and properly functionalized. Accordingly, an easy accesses to dibenzohelicenes (*Angew. Chem. Int. Ed*. **2013**, *52*, 9970) and optically pure heterohelicenes (*Angew. Chem. Int. Ed*. **2012**, *51*, 5857; *Chem. Eur. J*. **2014**, *20*, 8477) has been developed and the helicenes were used for the formation of nanoarchitectures (*ACS Nano* **2013**, *7*, 3676).

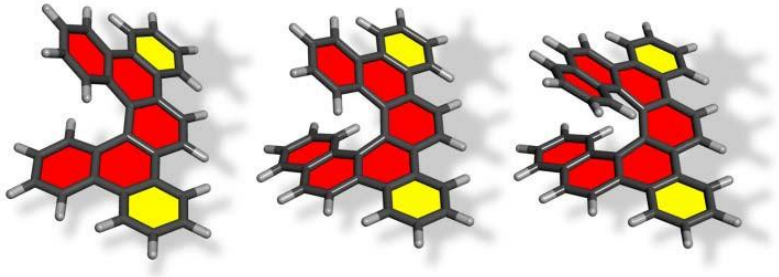


Figure 3.5.1: Novel extended dibenzohelicenes (*Angew. Chem. Int. Ed*. **2013**, *52*, 9970).

The Teplý group has established chemistry of helquats as stable chiral helical dications and studied their chiroptical and redox properties (*J. Am. Chem. Soc.* **2014**, *136*, 10826). Ability of helquats to chiroselectively aggregate Au nanoparticles has been shown (collaboration with the Hebrew University of Jerusalem: *Nano Lett.* **2012**, *12*, 5835). Also a new type of saddle-shaped saddlequats (*Chem. Sci.* **2011**, *2*, 2314) was discovered as well as a chiral dicationic compound resembling [8]circulene (*Angew. Chem. Int. Ed.* **2012**, *51*, 11972). Some of these compounds were studied as fluorescent dyes for applications in biology.

A novel stereochemical phenomenon, planamerism, has been reported by the Janeba group, where planamers are defined as rotamers of small aromatic molecules with a planar conjugated moiety that are isolable as chemical species (*Chem. Commun.* **2014**, *50*, 14892). The Jahn group discovered a new type of organic superbases with wide application potential (*Angew. Chem. Int. Ed.* **2014**, *53*, 1435). The Kraus group developed a new class of receptors based on cyclodextrin duplexes and used them for catalysis (*Chem. Commun.* **2010**, *46*, 7599) and complexation and transport of drugs, nucleotides and other biomolecules (*Chem. Eur. J.* **2012**, *18*, 12292).

Novel synthetic methodology was the major focus of the Jahn group which worked out several new methods based on electron-transfer chemistry (*Chem. Eur. J.* **2012**, *18*, 12267), on new tandem processes (*Chem. Eur. J.* **2014**, *20*, 10298), or metal catalysis (*Angew. Chem. Int. Ed.* **2014**, *53*, 9944) and applied them in the total synthesis of natural products (pheromones, steroids or isoprostanes). The Beier group focused on the chemistry of main group elements and developed new methodologies for the synthesis of several new classes of compounds and new reagents for organic synthesis. As an example, new methods for fluoroalkyl group transfer using fluorinated phosphonates were developed (*J. Org. Chem.* **2013**, *78*, 4573; *Chem. Eur.*

J. **2014**, *20*, 1453), as well as a novel access to rare sulfur pentafluorides (*Org. Lett.* **2011**, *13*, 1466; *J. Org. Chem.* **2014**, *79*, 8906). The Hocek group developed several methods of modifications of nucleobases based on C-H arylations and other C-H activations (*J. Org. Chem*. **2010**, *75*, 2302) and a new approach to the synthesis of C-nucleosides (*J. Org. Chem.* **2011**, *76*, 6619).

Medicinal chemistry and drug development

In medicinal chemistry, extensive research has been performed not only in antimetabolites of nucleic acids components (legacy of Prof. Antonín Holý) but also several new types of compounds and new areas of potential clinical applications have been studied.

In the prolific area of acyclic nucleoside phosphonates (previously approved as antiviral drugs) several new types of analogues were prepared and explored in the Janeba group as inhibitors of hypoxanthine-guanine-(xanthine) phosphoribosyltransferase, a key enzyme of the purine salvage pathway of *Plasmodium* parasites (collaboration with the University of Queensland: US/61/643,419, WO2013166545-A2; *J. Med. Chem*. **2012**, *55*, 6209; *J. Med. Chem.* **2013**, *56*, 2513; *J. Med. Chem*. **2015**, *58*, 827) with potential applications as antimalarial agents. The same group also discovered new efficient non-nucleoside reverse transcriptase inhibitors (NNRTIs) with excellent antiviral activities against HIV in collaboration with Gilead Sciences (patent applications pending) and new anti-inflammatory pyrimidine derivatives as potent inhibitors of immune-activated nitric oxide (NO) and/or prostaglandin E2 (PGE2) production with promising potential in the treatment of ulcerative colitis and/or rheumatic arthritis (Jansa P. *et al.*, US 8,883,798 B2).

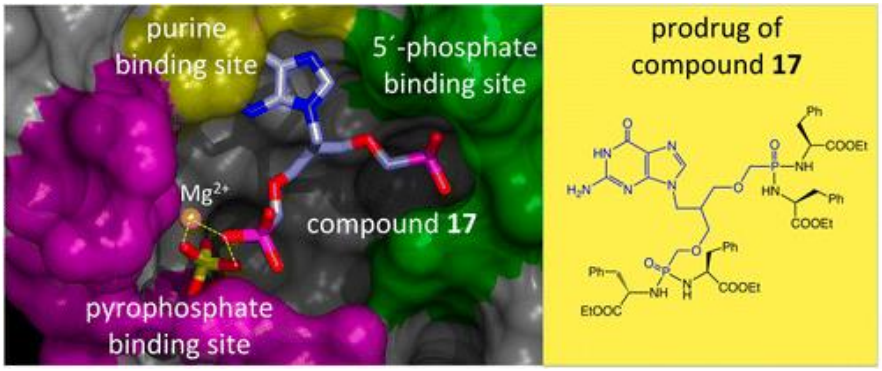


Figure 3.5.2: Crystal structure of hypoxanthine-guanine-(xanthine) phosphoribosyltransferase with a nucleotide inhibitor (*J. Med. Chem.* **2013**, *56*, 2513).

The Nencka group has developed a new methodology for one-pot construction of purine derivatives from amine precursor (*RSC Adv*. **2012**, *2*, 6970) and designed and synthesized new conformationally locked nucleoside and nucleotides as potential antiviral agents and discovered a broad group of antiviral compounds against picornaviruses, the most common human pathogens. The Krečmerová group has studied 5,6-dihydro-5-azacytosine nucleosides and their polymer conjugates for applications in epigenetic therapy and started a new programme (collaboration with the Majer group and the Johns Hopkins University) in prodrugs of 2-(phosphonomethyl)pentane-1,5-dioic acid (2-PMPA) as inhibitors of glutamate

carboxypetidase II (GCPII) targeting some inflammatory and neurodegenerative diseases (US Appl. No. 62/033926).
The Rosenberg group investigated several novel types of nucleoside phosphonates and found new inhibitors of purine and thymidine phosphorylases and cytosolic and mitochondrial 5'(3')-nucleotidases (*Org. Biomol. Chem.* **2014**, *40*, 7971). In their oligonucleotide program, they developed new stable phosphonate analogues of oligonucleotides which elicit RNase H activity and thus have potential in antisense therapy (*Nucleic Acids Res.* **2014**, *42*, 5378). The Hocek group discovered several new types of hetaryl-7-deazapurine ribonucleosides with namomolar cytostatic activities, investigated the SAR and mechanism of action (US 8,093,226 B2; *J. Med. Chem.* **2010**, *53*, 460; *J. Med. Chem.* **2011**, *54*, 5498; *J. Med. Chem.* **2014**, *57*, 1097). Several compounds from this series are now undergoing preclinical development as potential antitumor agents.
Other areas of medicinal chemistry included neuroprotective steriods where the Kudová group discovered new class of NMDA antagonists and studied the SAR (*Neuropharmacology* **2011**, *61*, 61) and insect peptide antibiotics where the Čeřovský group discovered several novel antimicrobial peptides such as lucifensin - the defensin of medicinal maggots (*Cell. Mol. Life Sci.* **2010**, *67*, *455*) and others which are studied for their applications in the treatment of osteomyelitis and some topical infectious diseases (patent applications pending).

Bioorganic chemistry and chemical biology

In bioorganic chemistry, the Hocek group developed several new methods in polymerase synthesis of base-modified nucleic acids and oligonucleotides and applied them in diagnostics for environment-sensitive fluorescent labelling (*Chem. Sci.* **2012**, *3*, 2797) or redox coding (*Chem. Eur. J.* **2013**, *19*, 12720) and in chemical biology for bioconjugations (*Angew. Chem. Int. Ed.* **2010**, *49*, 1064) and cross-linking with proteins (*Angew. Chem. Int. Ed.* **2013**, *52*, 10515) as well as for regulation of protein binding to DNA (*Angew. Chem. Int. Ed.* **2011**, *50*, 8727; *Angew. Chem. Int. Ed.* **2014**, *53*, 6734). They also discovered several new base- modified 2'-deoxyribonucleoside triphosphates which are better substrates for DNA polymerases than natural dNTPs (*Angew. Chem. Int. Ed.* **2014**, *53*, 7552) and now are pursuing the applications for metabolic labeling. The Vrábel group has started their program in developing of new bioorthogonal reactions for bioconjugations of diverse types of biomolecules.

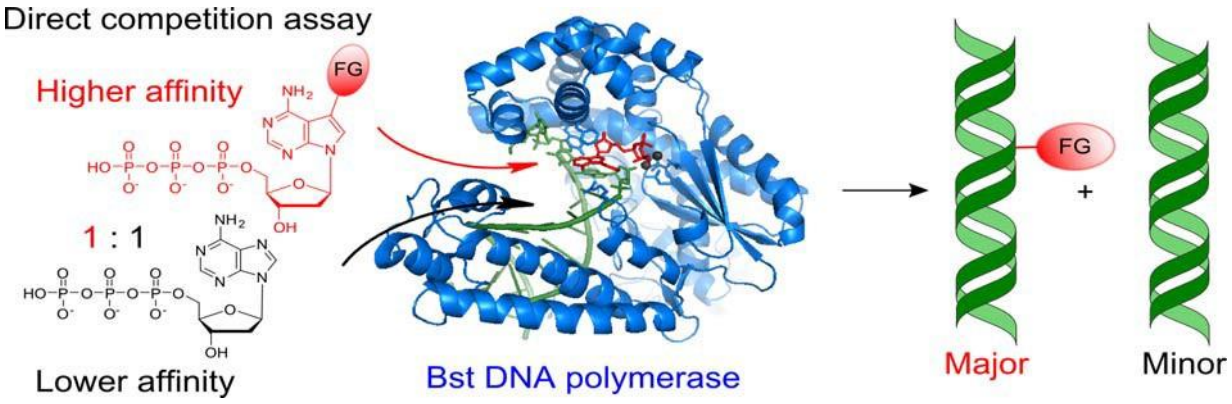


Figure 3.5.3: Unexpected preferential incorporation of 7-substituted-7-deazapurine nucleoside triphosphates by DNA polymerases in presence of the natural substrates (*Angew. Chem. Int. Ed.* **2014**, *53*, 7552).

Material Chemistry and Functionalized Molecules

Fundamental concepts from quantum chemistry are combined with synthetic organic or inorganic chemistry to develop new materials with improved or entirely new properties in the Michl research group at IOCB. The work encompasses four areas of physical organic chemistry: (i) arrays of molecular rotors where emphasis is given to the fundamental issue of artificial ferroelectricity and friction on a molecular scale, essential for anticipated practical applications (*ACS Nano* **2012**, *6*, 1901), (ii) alkylation of gold surfaces (*J. Am. Chem. Soc.* **2013**, *135*, 5669), (iii) carborane anion and radical chemistry, and (iv) singlet fission where the use of theory leads to the design of new singlet fission sensitizers (*J. Am. Chem. Soc.* **2012**, *134*, 14624). The new compounds are now being synthesized and promise to go beyond the Shockley-Queisser limit in solar cell efficiency. From a broader perspective, the first two areas relate to nanoelectronics, the third to batteries, and the fourth to photovoltaic cells.

New types of nanoparticles for use in therapeutics, imaging and diagnostics of diseases were synthesized in the Cígler group. These nanomaterials involve either bioorganic or inorganic cores and include viral-like particles, fluorescent nanodiamonds and plasmonic systems. Novel types of biocompatibilization procedures for nanoparticles were introduced which exhibited highly effective and selective targeting to cancer cells (for typical results, see *Small*, **2014**, *10*, 1106). Furthermore, ultra-bright nanodiamond fluorescence probes for background-free, near-IR cellular imaging were developed (*Nanoscale* **2013**, *5*, 3208).

Biochemistry and Molecular Biology

The research in the area of the biochemistry was mainly focused on study of molecular interactions important for pathogenesis of diseases caused by viruses, parasites, and bacteria, further on regulatory processes important for tumorous growth and on metabolic disorders (diabetes/obesity) and inflammatory diseases.

Study of interaction important for pathogenicity of different parasites, viruses, and pathogenic bacteria. The Konvalinka group studied proteases as targets for therapeutic intervention and investigated molecular mechanisms behind protease function in retroviruses. They designed and prepared specific, potent photodegradable inhibitor of HIV protease that enables to synchronize the HIV in the cell culture and trigger its maturation by light, thus opening the way to the detailed analysis of the late phase of viral replication (*J Virol.* **2014**; *88*, 13722).

The Pichová group studied interactions mediating assembly of immature and mature retroviral particles, which is critical for virus function and represents a potential target for development of new types of antiretroviral drugs. The assembly process and particle structure were analyzed by combinations of *in vivo* experiments, biochemical, NMR structural, and cryo-electron tomography and sub-tomogram averaging methods. The results open a new area for designing and testing compounds that would inhibit the CA-CA interactions critical for particle formation and stabilization (*Nature* **2012,** *487*, 385; *Retrovirology* **2014**, *11*, 94, *J. Virol.* **2014**, *88*, 14148). The Weber virology research-service group performed screening of antiviral compounds against variety of viruses (herpes simplex virus, human immunodeficiency virus 1, influenza virus, coxsackie B3 virus, hepatitis B virus and Dengue virus) and evaluated the role of drug resistance mutations in protease, reverse transcriptase and integrase on HIV-1 replicative fitness (*PLoS ONE* **2013,**

8:e65631). The Mareš group investigated cathepsins of parasitic blood flukes (schistosomes), proteolytic and anti-proteolytic systems of parasitic ticks as antigens for vaccines against ticks and tick-borne diseases. Emphasis was placed on the identification of cathepsin target - substrates, their functional and structural characterization, and on development of inhibitors. Crystallographic elucidation of 3D structure of schistosomal cathepsin SmCB1 provided a basis to design novel peptidomimetic inhibitors of SmCB1 as potential chemotherapeutics for the treatment of schistosomiasis (*J. Biol. Chem.* **2011**, *286*, 35770; *Structure* **2014**, *22*, 1786). The Pichová group studied regulation of central carbon metabolism of *Mycobacterium tuberculosis* (Mtb) during latent infection. The omics analyses of slowly growing Mtb indicated the key role of phosphoenolpyruvate carboxykinase (Pck) in the central carbon metabolism. The biochemical and structural characterization of Pck and identification of cellular interacting partner confirmed that intracellular reducing conditions and up-regulation of proteins maintaining the intracellular reduced state and antioxidant defense direct the Pck function and metabolite flow during latent Mtb infection (*J. Biol. Chem.* **2014**, *289*, 13066). The Mertlikova research-service group participated in development of novel antibacterial drugs acting as inhibitors of adenylate cyclase toxin of pathogens causing the whooping cough (*Bordetella pertusis*) and anthrax (*Bacillus anthracis*). (*Antimicrob Agents Chemother.* **2014**, *58*, 664).

Cancer research. The Řezáčová group used structural knowledge for characterization of various carborane clusters that act as active-site-directed inhibitors of human carbonic anhydrase and showed that substitution with an appropriately attached sulfamide group and other substituents leads to compounds with high selectivity toward the cancer-specific carbonic anhydrase isoenzyme IX (*Angew. Chem. Int. Ed.* **2013,** *52*, 13760). This group further studied lineage leukemia (MLL) protein, which plays a role in development of acute leukemias with poor prognosis, and its interaction with the lens epithelium–derived growth factor (LEDGF/p75), a prominent cellular cofactor for HIV integration. Using nuclear magnetic resonance they identified novel LEDGF/p75–MLL interface, which represents a new target for the development of therapeutics against MLL fusion– driven leukemic disorders (*Cancer Res.* **2014**, *74*, 5139). The Konvalinka group analyzed the molecular mechanism of the binding of substrates and inhibitors to glutamate carboxypeptidase II (GCPII), which plays important role in cancer development, and designed novel inhibitors for the specific targeting to tumors expressing GCPII (*FEBS J.* **2014,** 281; *Bioorg. Med. Chem.* **2014**, *22*, 4099).

Mechanisms of metabolic disorders. The Jiráček group designed and prepared a series of highly active insulin analogs. The X-ray structures of the analogs revealed a unique conformation different from inactive storage forms of this hormone. This conformation, called B26-turn, represents very probably the active form of insulin (*Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1966). The group also significantly contributed to the solution of the structure of a complex of insulin with insulin receptor (*Nature* **2013**, *493*, 241). These results represent a milestone in the study of insulin interaction with the receptor. Modified insulins resembling the active form of the hormone may be the starting point for the development of new more effective insulin analogs or insulin mimetics for nasal, pulmonal or oral treatment of diabetes.

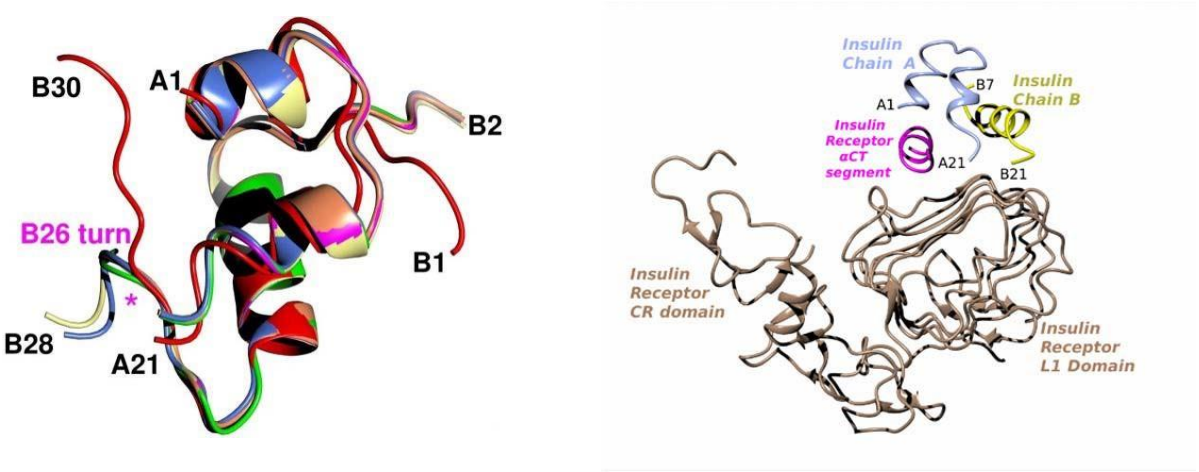


Figure 3.5.4: *Left.* Superposition of inactive form of human insulin (red) with highly active insulin analogs possessing B26-turn conformation representing the active form of the hormone. *Right.* Crystal structure of a complex of insulin with L1, CR and CT domains of insulin receptor. Insulin A chain is blue and insulin B chain is yellow.

The Maletinská targeted-research group studied development and mechanism of action of stabilized analogs of food intake regulating anorexigenic neuropeptides (prolactin-releasing peptide (PrRP), cocaine-and amphetamine-regulated transcript (CART) peptide, and orexigenic ghrelin for potential treatment of food intake disorders such as obesity or cachexia. The analogs were also investigated for potential treatment of Alzheimer's disease. (*J Pharmacol Exp Ther.* **2012** *340*, 781; PV2014-429). The Weis junior group investigated molecular and cellular mechanisms underlying neuronal human diseases caused by dysfunction of ion channels, so-called channelopathies. By combination of electrophysiological recording (patch-clamp), molecular and biochemistry techniques, and confocal imaging microscopy they described channel function regulation and their contribution to neuronal excitability, which can be applied for new relevant therapeutic avenues (*Mol Pharmacol* **2014, DOI**: 10.1124/mol.114.096008). The junior group of Stríšovsky studied the transmembrane proteins of the rhomboid family that comprise intramembrane serine proteases and related non-enzymatic members, participating in intercellular signaling, biogenesis of membrane, quality control, inflammatory and growth factor signaling, and are thus relevant for a number of chronic diseases. The X ray structure of this membrane protein and biochemical characterization contributed to elucidation of rhomboid family protein mechanism and enabled to develop small molecules interfering with the rhomboid functions (*EMBO J.* **2014**, *33*, 2408).

Chemical Ecology

Chemistry of social insects. The broad focus of IOCB in chemical sciences includes also the interest in the chemistry of natural products, namely those produced by insects in defensive or communication context. The Valterová group combined analytical chemistry, electrophysiology, biochemistry, and genetic methods to unravel the chemical identity, function and biosynthesis of semiochemicals and defensive compounds in fruit flies, bumblebees and their parasites with the ultimate goal of regulation or protection of their populations. The group elucidated the biosynthesis of sexual fatty acid pheromones in different bumblebee species by

combination of *in vitro* incubations of pheromone precursors in dissected labial glands (LG) with q RT PCR gene expression analysis. The results indicated that pheromone components are synthesized *de novo* in the LG in the bumblebee males (*ChemBioChem* **2013,** *14*, 361). The Hanus junior research group described the defensive behaviour in workers of the termite *Neocapritermes taracua*, which develop an exceptional two-component suicidal apparatus consisting of copper-containing protein crystals, stored in external pouches, and internal salivary glands. The group showed that worker's body ruptures during aggressive encounters and the crystals react in the hemolymph droplet with the salivary gland secretion to produce the defensive secretion. This exceptional defensive strategy nicely illustrates how sociality favors the evolution of extreme adaptations, not without resemblance to kamikaze actions (*Science* **2012,** *337*, 436).

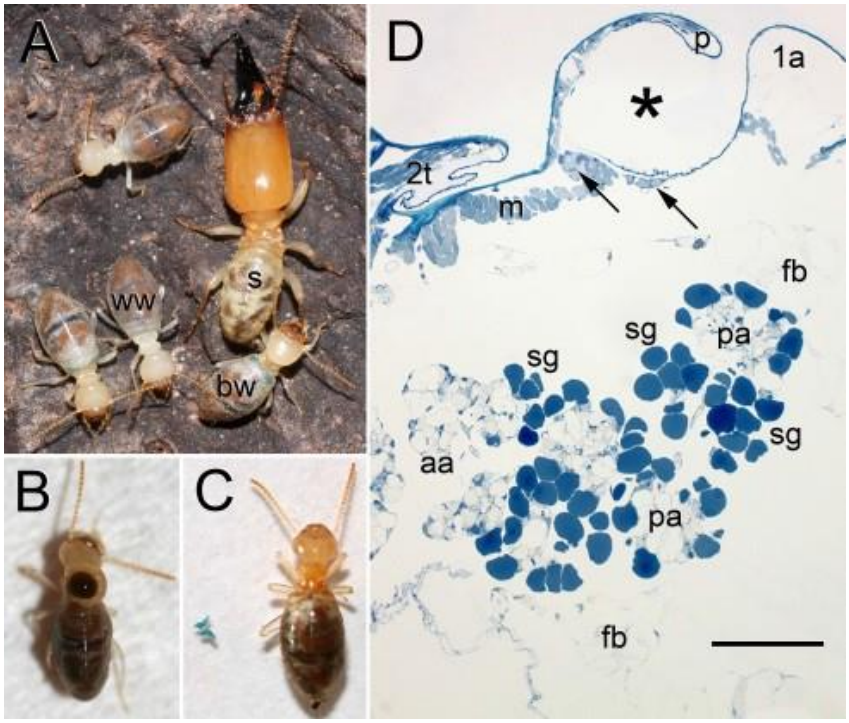


Figure 3.5.5: Anatomy of the defensive apparatus of *Neocapritermes taracua.* (A) Soldier (s) with two blue (bw) and two white workers (ww). (B) Blue worker after autothysis triggered by grasping it with tweezers. (C) Blue worker after removal of the blue crystals (placed next to it). (D) Section of anterior abdomen of a blue worker. Asteriskmarks crystal-bearing pouch (crystal dissolved); arrowsmark crystal gland cells. Scale bar, 200 μm. 1a, first abdominal segment; 2t, mesothorax; aa, anterior acini; fb, fat body; m, muscles; p, dorsal part of the crystal-bearing pouch; pa, posterior acini; sg, aggregates of secretion granules budding off posterior acini.

Computational and Theoretical Chemistry

Theoretical and computational research of noncovalent interactions ranges from accurate quantum-chemical calculations of model systems to development of new methods and their application to biochemical problems, mainly protein - ligand interactions. In the Hobza group the reference S66 database of interaction energies was developed (*J. Chem. Theory Comput.* **2011**, *7*, 2427) which became widely used tool for benchmarking computational methods. Since the non-covalent interactions

play crucial, if not decisive, role in the interactions of ligands with biomolecules, intensive computational efforts were directed towards *in silico* ligand design. As an example a theoretical and experimental study demonstrating that it is possible to modulate activity of aldose reductase inhibitors by changing the strength of the halogen bond between the inhibitor and the Thr-113 side chain of aldose reductase can be mentioned (*ACS Chem. Biol.* **2013**, *8*, 2484).
Obviously, not only the inherent properties of the small molecules or biomolecules, but also their interactions with the environment and their bulk properties are responsible for the physico-chemical behavior of real systems. To this aim, extensive molecular simulations of ions at biological aqueous interfaces, including interactions of ions with proteins and membranes with the aim of elucidating Hofmeister ion effects in protein salting out/salting in, denaturation, and enzymatic activity were performed in the Jungwirth group (see e.g. *J. Am. Chem. Soc.* **2012**, *134*, 10039). The reported studies also encompass modeling direct and indirect DNA damage and structure and dynamics of solvated electrons (e.g. *Nat. Chem.* **2014**, *6*, 697), effects of oxidation in cellular membranes, and investigating the role of ions in biological transport and signaling. The principal aim is to model at a molecular level proteins, biomembranes, and DNA, in their native aqueous environment with emphasis of ion specific effects employing state-of-the-art computational techniques spanning from classical molecular dynamics to ab initio quantum chemistry in direct contact with experimental biochemistry and molecular spectroscopy to unravel the action of ions in biological contexts.

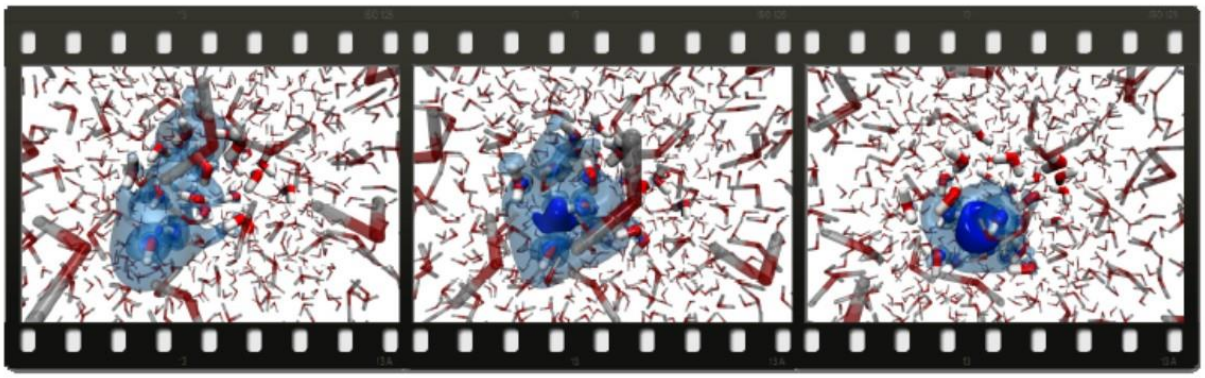


Figure 3.5.6: A series of snapshots from a computer simulation showing how the electron becomes "wet", i.e., localizes in water within a picosecond.

It is estimated that at least one third of the molecular cell machinery, mostly represented by enzymes, contain metal ions in their structure. These play either catalytic or structural role. Thus, the efforts in the Rulíšek group mostly concerned the area of theoretical bioinorganic chemistry. The high-level quantum chemical (QC) and combined quantum mechanical and molecular mechanical (QM/MM) modeling techniques were used to understand the fundamental questions, such as „Why the Nature selected particular metal ion to perform specific task?" To this aim, a protocol was developed that enables to compute at the "practical level of accuracy" the stability constants of metal ions in small complexes containing biologically relevant ligands (*Inorg. Chem.* **2013**, *52*, 10347). This may result in design of highly specific chelators in near future. At the same time, the reaction mechanisms of several intriguing metalloenzymes, such as non-heme diiron systems, were described using advanced multireference wave function techniques (such as the density matrix

renormalization group, DMRG) coupled with the QM/MM calculations (*J. Am. Chem. Soc.* **2014**, *136*, 15977). Related to these efforts were certain methodological advances in the computational electrochemistry (*J. Phys. Chem. A* **2013**, *117*, 1171).

Molecular Spectroscopy, New Instrumentation Techniques and Applications

Experimental and theoretical methods are continuously being developed for nuclear magnetic resonance (NMR), circular dichroism (CD), vibrational optical activity (VOA), and many other spectroscopic techniques. This also involved the establishment of the laboratory of Raman optical activity (ROA) and introduction of the magnetic circular dichroism (MCD) technique to IOCB. Quite naturally, the application of these methods in a number of domains including organic and medical chemistry, and biochemistry is always pursued. As an example, a simulation of VOA spectra of the entire insulin protein, and investigation of its behavior in the solution state by employing quantum-chemical computations and original tensor-transfer techniques can be mentioned (*Anal. Chem.* **2012**, *84*, 2440). The second example includes computational analysis of solvent effects in NMR spectroscopy (*J. Chem. Theory Comput.* **2010**, *6*, 288). Finally, a remarkable achievement was accomplished in 2014: the discovery of quite a new physical phenomenon, the paramagnetic optical activity (*Angew. Chem. Int. Ed.* **2014**, *53*, 9236). It involved adaptation of the spectrometer by extension with a magnetic cell and additional development of theoretical apparatus based on the theory of angular momentum to simulate and explain the spectra. The methodology can be potentially used for analyses of industrial gases and pollutants.

Furthermore, new generation capillary electrophoresis (CE) device with multidimensional detection system composed of contactless conductivity, UV-absorption and fluorescence detectors was constructed which enables separation and detection of a wide range of compounds, independently of the presence or absence of chromophores or fluorophores in their (macro)molecules. Methodology developments included the following CE methods: zone electrophoresis, isotachophoresis, isoelectric focusing, affinity electrophoresis, electrokinetic chromatography and electrochromatography. The developed device and methods were applied to the fast high-efficient separation, high-sensitive qualitative and quantitative analysis and physicochemical characterization of biomolecules (e.g. peptides, proteins, nucleosides and (oligo)nucleotides) and functional organic molecules (e.g. azahelicenes and helquats) synthesized or isolated in IOCB and potentially applicable in medicine or in molecular electronics, catalysis and nanotechnology (*Electrophoresis* **2011**, *32*, 2683; *J. Chromatogr. A* **2011**, *1218*, 4982).

Chromatography and mass spectrometry was also used in the area of lipidomics. This included development of new strategies and approaches for characterizing wax esters, including a unique method for localization of double bond position within the aliphatic chains (*Anal. Chem.* **2011**, *83*, 2978), new MALDI-MS methods for analyzing neutral lipids in vernix caseosa for disclosing gender-related differences in developing human skin and new potent matrices for MALDI-MS of lipids (*J. Am. Soc. Mass Spectrom.* **2010**, *21*, 220). Last but not least, the coupling of electrochemistry with mass spectrometry and interactions of organic compounds with inorganic cations were investigated.

Providing the aforementioned list of recent research highlights, we attempted to emphasize the multidisciplinarity of the research carried out at the IOCB, which is

considered as its prime asset concerning future scientific developments and achievements.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Václav Kašička – Electromigration methods

2.1. General characterization

The work was oriented to the research and development of instrumentation, methodology and application of high-performance capillary electromigration (HPCE or CE) methods. In the instrumentation area, a new type of HPCE apparatus with multidimensional detection system composed of electric conductivity, UV-absorption and fluorescence detectors was designed and constructed. It enables separation and detection of a wide range of compounds, independently of the presence or absence of chromophores or fluorophores in their (macro)molecules. Methodology developments included the following CE methods: zone electrophoresis (CZE), isotachophoresis (CITP), isoelectric focusing (CIEF), affinity electrophoresis (CAE), electrokinetic chromatography (CEKC) and open tubular electrochromatography (OT-CEC). The developed device and methods were applied to the fast high-efficient separation, high-sensitive qualitative and quantitative analysis and physicochemical characterization of biomolecules (e.g. amino acids, peptides, proteins, nucleosides and (oligo)nucleotides) and functional organic molecules (e.g. azahelicenes and helquats) synthesized or isolated in IOCB and/or in collaborating institutions.

2.2 Particular areas of the research and development

2.2.1. Instrumentation

A new type of a computer controlled semi-automated CE apparatus with multidimensional detection system was designed and constructed. The separation takes place in a fused silica capillary with typical inner diameter 50-75 μm, outer diameter 375 μm, total length 30-70 cm, and effective length (from injection end to the detector) 10-60 cm. Background electrolyte (BGE) and sample solutions are introduced in the capillary by the pneumatically regulated hydrodynamic system with high pressure (100-500 kPa) for filling the capillary with BGE and with low pressure (300-2000 Pa) for the sample injection. Separation voltage is provided by a high-voltage power supply with the following parameters: 0-30 kV, 0-300 μA, constant voltage or constant current mode. Hydrodynamic system and the high voltage power supply are controlled by a PC-based control unit. BGE electrode vessels and sample microvials are placed in teflon carousels and are operated manually.

A specific feature of this apparatus, which makes it different from the commercially available CE analyzers, is its multidimensional detection system composed of doubled contactless conductivity detector, UV-photometric absorption detector operating at one of the three fixed wavelengths, 206, 254 and 280 nm, and laser induced fluorescence (LIF) detector with excitation wavelength 266 nm, see Figure 1. This system allows detection of a wide range of compounds. UV-absorbing

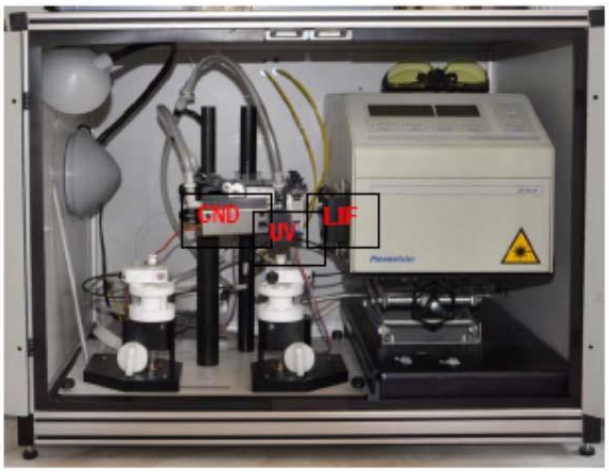


Figure 1 Photography of the home-made HPCE device with multidimensional detection system composed of contactless conductivity (CND), UV-absorption and laser induced fluorescence (LIF) detectors.

compounds are detected by a home-made UV-photometric detector, compounds with fluorophores at low UV excitation wavelengths are detected by LIF detector (Picometrics SA, Toulouse), and all compounds, including UV-transparent and non-fluorescing ones, can be detected by the universal contactless conductivity detector (Admet, Prague). LIF detector with excitation wavelength 266 nm makes possible high-sensitive detection of compounds fluorescing with the above excitation wavelength 266 nm, e.g. peptides and proteins containing aromatic amino acid residues of tryptophan and tyrosine, on the basis of their native fluorescence, i.e. without the need of their derivatization. A simultaneous data acquisition from all detectors is performed by the four channel data acquisition and evaluation chromatographic and electrophoretic station Clarity (DataApex, Prague).

2.2.2. Methodology

Methodological developments included the following HPCE methods, zone electrophoresis (CZE), isotachophoresis (CITP), isoelectric focusing (CIEF), affinity electrophoresis (CAE), electrokinetic chromatography (CEKC) and open tubular electrochromatography (OT-CEC).

2.2.1. Capillary zone electrophoresis (CZE)

The following new types of the background electrolytes (BGEs) were developed:

i) Strongly acidic or strongly alkaline BGEs for separation, analysis and characterization of analytes with low and/or high acidity (ionization) constants (pK_a), i.e. medium strong, weak and very weak acids and bases [1,2].

ii) Isoelectric or quasi-isoelectric BGEs composed of single- or multiple-component amphoteric compounds (amino acids, narrow fractions of carrier ampholytes used as separation media in isoelectric focusing) with low conductivity allowing application of high intensity electric field and thus resulting in high-efficiency and high-speed separations of peptides and proteins [3,4].

iii) Non-aqueous methanolic BGEs with well-defined conventional pH values based on using weak acid with known acidity constant (pKa) in methanol and a salt of this acid as BGE constituent applied to analysis and characterization of acid-base properties of water insoluble compounds, such as azahelicenes [5].

2.2.2. Capillary isotachophoresis (CITP)

In cooperation with the group of Prof. H. Cottet, University of Montpellier, a new method has been developed for determination of effective charge of (bio)polymers by CITP using special ITP apparatus (Villa Labeco, Spišská Nová Ves, Slovakia). The method is based on the linear dependence of the ITP zone length of the solute on its concentration and effective charge. The method was applied to determination of effective charge of dendrigraft poly-L-lysine by cationic CITP using ammonium ion as leading cation and acetate as counter ion of the leading electrolyte (pH 4.7), and acetic acid as terminating electrolyte [6]. In the next step, the method was generalized for its usage on ordinary, commercially available CE instruments [7]. This method is applicable to a broad range of strong or weak polyelectrolytes. Determined effective charges of analyzed polymers were in a good agreement with those obtained by capillary electrophoresis with indirect UV-photometric detection.

2.2.3. Capillary isoelectric focusing (CIEF)

Methodology improvement of CIEF developed also in cooperation with the group of Prof. H. Cottet, consists in flattening the pH gradient generated by commercial carrier ampholytes Servalyte pH 6-8 or Biolyte pH 6-8 by the addition of narrow pH fraction of pH 7.2 prepared by preparative fractionation of Servalyte pH 4-9 carrier ampholytes. This way increased resolving power allowed separation of glycated hemoglobin A1c (HbA1c) from non-glycated HbA0 form and made possible CIEF determination of HbA1c in blood [8].

2.2.4. Capillary affinity electrophoresis (CAE)

In the CAE, the following methodological improvements were achieved:

i) Acidic BGEs with stereoselective additives, randomly sulfated α-, β- and γ-cyclodextrins (CDs) were developed and applied to separation of enantiomers of helquats (N-heteroaromatic dicationic helicene-like species) – a new class of functional organic molecules [9−13].

ii) Non-aqueous and hydro-organic BGEs using methanol (MeOH) or MeOH/water mixtures as solvents were developed and applied to investigation of noncovalent interactions of water insoluble or sparingly soluble compounds, e.g. benzo-18-crown-6-ether and hexaarylbenzene-based receptor with small cations [14,15].

iii) Alkaline BGEs with native and derivatized α-, β- and γ-CDs were developed and applied to chiral analysis acyclic nucleoside phosphonates, important class of compounds with antiviral, antiparasitic and antineoplastic activities [16,17].

iv) A new mode of CAE, partial filling CAE (PF-CAE) has been developed and applied to investigation of noncovalent interactions between double stranded DNA oligonucleotide (Dickerson dodecamer) and classical DNA intercalator ligand (ethidium bromide) or oligophenylene derivatives-based potential new type of DNA ligands [18].

v) CAE methodology for investigation of non-covalent interactions of (bio)molecules and for determination of binding constants K_b of (bio)molecular complexes was made more precise. Prior to K_b calculation, the measured effective electrophoretic mobilities were first corrected to reference temperature 25°C and constant ionic strength [19].

2.2.5. Capillary electrokinetic chromatography (CEKC) and open-tubular capillary electrochromatography (OT-CEC)

In these two modes, new types of pseudostationary and stationary phases have been developed in cooperation with the groups of Prof. V. Král (Institute of Chemical Technology, Prague) and Prof. I. Mikšík (Institute of Physiology ASCR, Prague) and tested for separation of various types of compounds.

i) Gold nanoparticles (GNPs) added as a free pseudostationary phase in the mobile phase in the CEKC or physically adsorbed to the fused silica capillary wall as stationary phase in OT-CEC were employed for separation of peptides and polyaromatic hydrocarbons (BGE) [20] and for CE-UV profiling of trypsin digested native and glycated proteins [21].

ii) Cyclodextrin modified GNPs covalently attached or physically adsorbed to the fused silica capillary wall were applied as stationary phase in OT-CEC separation of polyaromatic hydrocarbons [22].

2.3. Applications

In this section, the most important applications of HPCE methods to separation, analysis and physico-chemical characterization of biomolecules and functional organic molecules are described.

2.3.1. Analysis

The most important results were achieved in the area of chiral analysis.

Chiral analysis of helquats

A high-efficient CE method has been developed for chiral analysis of a new class of functional organic molecules, helical *N*-heteroaromatic dications – helquats, synthesized in the group of F. Teplý [23]. Enantiomers of 24 helquats comprising 5, 6 and 7 fused rings participating in the helical backbone were separated in acidic 22/35 mM sodium/phosphate BGE, pH 2.4, using randomly sulfated α-, β- a γ-CDs as chiral selectors, see Figure 2. Low picomole amounts of helquat enantiomers in nanoliter

injected sample volumes were separated in a short time 5-10 min [9]. The method was extensively employed in the research of helquats, e.g. in the their chiral purity control and racemization monitoring [10,11,24], and especially in the discovery of their new related structures, [6]saddlequat [12] and [8]circulenoid [13]. The method was applicable also for chiral analysis of helicenes-like compounds based on imidazolium motif [25].

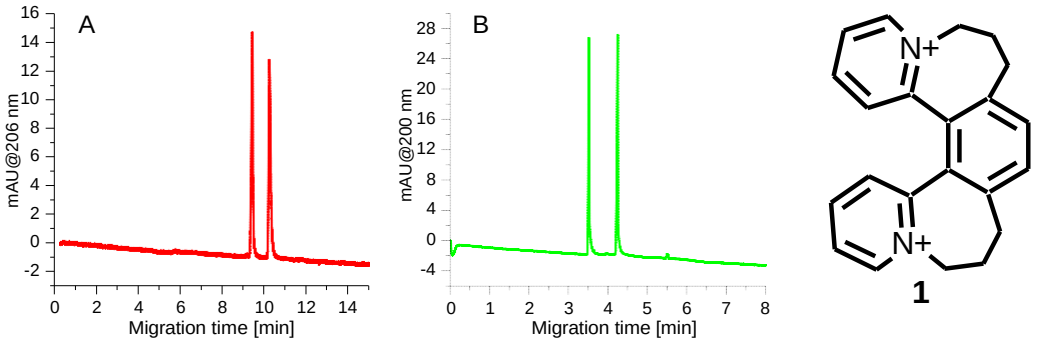


Figure 2 CE separation of *P* and *M* enantiomers of pentahelquat **1** by CE in 22/35 mM sodium/phosphate BGE, pH 2.4, using 4 mM sulfated β-CD (A) and 4 mM sulfated γ-CD (B) as chiral selectors. Separation was performed in the CE analyzer P/ACE MDQ (Beckman Coulter, Fullerton) in the fused silica capillary (id/od 50/375 μm, total/effective length 60/71 cm (A) or 30/40 cm (B) at separation voltage -25 kV (A) or -16 kV (B), at 20˚C. Adapted from ref. [9].

Chiral analysis of acyclic nucleoside phosphonates

A new fast CE method has been developed for chiral analysis of acyclic nucleoside phosphonates (ANPs) – an important class of compounds synthesized in our Institute and exhibiting antiviral, antiparasitic and antineoplastic acivities [26,27]. Complete (base line) separations of several ANPS were achieved within a short time (5–15 min) in sodium tetraborate BGE, pH 10.0, using native and derivatized CDs as chiral selectors. With UV-absorption detection at 206 nm, concentration detection limits of the analyzed ANPs were in the submicromolar range, few femtomoles of ANPs in 2-5 nanoliters were injected for one analysis. The method is applied for the chiral purity control of new preparations of ANPs prior to their delivery for biological tests in the pharmaceutical company Gilead Sciences, Inc., partner of IOCB in the Gilead Sciences & IOCB Research Center [16,17].

Qualitative and quantitative analysis of amino acids, peptides and proteins

CZE and MEKC methods in various BGEs within a broad pH scale (pH 2.1-10) and with anionic or cationic detergent based pseudostationary phases were applied to qualitative and quantitative analysis of various preparations of amino acids and peptides isolated or (bio)synthesized in the group of V. Čeřovský, e.g. thirteen mono-*N*-acyl derivatives of 2,6-diaminopimelic acid – new potential inhibitors of the enzyme *N*-succinyl-L,L-diaminopimelic acid desuccinylase [28], melectin antimicrobial peptides [29] and double chain peptide fragment of human IgG1 hinge region [30]. The purity degree of these amino acids and peptides was determined as the ratio of the corrected (migration time normalized) peak area of amino acid or peptide itself to the sum of corrected areas of all peaks present on electrophoregrams.

In cooperation with the group of Prof. A. Cifuentes, Institute of Food Science and Research, CSIC, Madrid, water soluble proteins from *Bacillus thuringiensis* (Bt)-transgenic (Aristis-Bt) and two native non-transgenic (Aristis and Coventry) maize varieties and tryptic digests of these proteins were analyzed by CZE in classical and isoelectric acidic BGEs (formic, acetic and iminodiacetic acids, pH 2.1-2.6) [3]. Among the tested BGEs, the best resolution of extracted proteins and their tryptic peptide fragments was achieved in the isoelectric BGE, 200 mM iminodiacetic acid, pH 2.26. Some significant relative qualitative and quantitative differences in CZE-UV profiling of the above proteins and their tryptic digests were found, which can be potentially used to differentiate transgenic Aristis Bt and non-transgenic Aristis varieties or two native non-transgenic varieties, Aristis and Coventry.

Recent developments and applications of HPCE methods to separation, analysis and characterization of peptides including peptide drugs were summarized in four invited review articles [31–34].

2.3.2. Physico-chemical characterization

Effective mobility, effective charge and isoelectric point

CE methods have been widely employed in the physicochemical characterization of biomolecules as well as functional organic molecules. In most of the above applications, the analyzed compounds were characterized by their effective electrophoretic mobilities, which reflect their charge to size (M_r) ratio. In the case of GnRH peptides and their analogs and fragments, from the dependence of electrophoretic mobilities of these peptides on pH, their isoelectric points were determined [4]. Newly developed CITP method has been applied to determination of effective charge of linear and dendrigraft poly-L-lysine [6], linear poly-L-glutamic acid and some other (bio)polymers [7].

Acidity (ionization) constant (pK_a) and ionic mobilities

Non-aqueous CZE in methanolic BGEs within a wide conventional pH range (pH_{MeOH} = 1.6-13.3) was employed to determine pK_a of ionogenic groups of azahelicenes – unique three-dimensional aromatic systems synthesized in the group of I. Starý [35]. They consist of *ortho*-fused benzene/pyridine units and exhibit helical chirality. The pK_a values of pyridinium groups of the mono- and dibasic azahelicenes and ionic mobilities of mono- and dicationic forms were determined from the dependence of their effective electrophoretic mobility on pH by a nonlinear regression analysis [5].

Thermodynamic pK_a of seven zwitterionic heterocyclic very weak bases, designed as chiral Lewis base catalysts for enantioselective reactions, were determined by CZE from the pH dependence of their effective electrophoretic mobility in strongly acidic aqueous BGEs (pH 0.85 – 2.80). Prior to pK_a calculation by non-linear regression analysis, the CZE measured effective mobilities were corrected to reference temperature, 25 °C, and constant ionic strength, 25 mM. Thermodynamic values of these bases were found to be rather low, in the range 0.04-0.32. In addition to pK_a, also actual ionic mobilities and hydrodynamic radii of the cationic forms of the bases were estimated [1].

Thermodynamic pK_a of ionogenic groups and actual ionic mobilities of polyprotic peptides – synthetic human and salmon gonadotropin-releasing hormones (GnRHs) and their derivatives and fragments were determined from their CZE separations in series of BGEs in a wide pH range 2-12 [2]. First, the mixed acidity constants, $pK_{a,mix}$,

of ionogenic groups, and actual ionic mobilities, m_i, of GnRH peptides were determined by nonlinear regression analysis of pH dependence of their effective electrophoretic mobilities. Second, the p$K_{a,mix}$ values were recalculated to thermodynamic pK_as using the Debye-Hückel theory.

Binding (association, stability) constant (K_b) of (bio)molecular complexes

Capillary affinity electrophoresis (CAE) was applied to investigation of non-covalent interactions of (bio)molecular ligands or receptors with small cations, alkali metal ions and ammonium ion. Strength of the interaction was quantified by the binding (association, stability) constant of the formed complexes. The binding constant was determined from the dependence of effective mobility of the ligand on the concentration of the cation in the BGE, see Figure 3. Similarly as in the case of pK_a determination, prior to K_b calculation by non-linear regression analysis, the effective mobilities were first corrected to the reference temperature 25°C and constant ionic strength, 10 mM or 25 mM. Using this procedure, the binding constants of the complexes of benzo-18-crown-6 ether and hexaarylbenzene-based receptor with alkali metal ions and ammonium ion in methanol were determined [14,15,19,36,37].

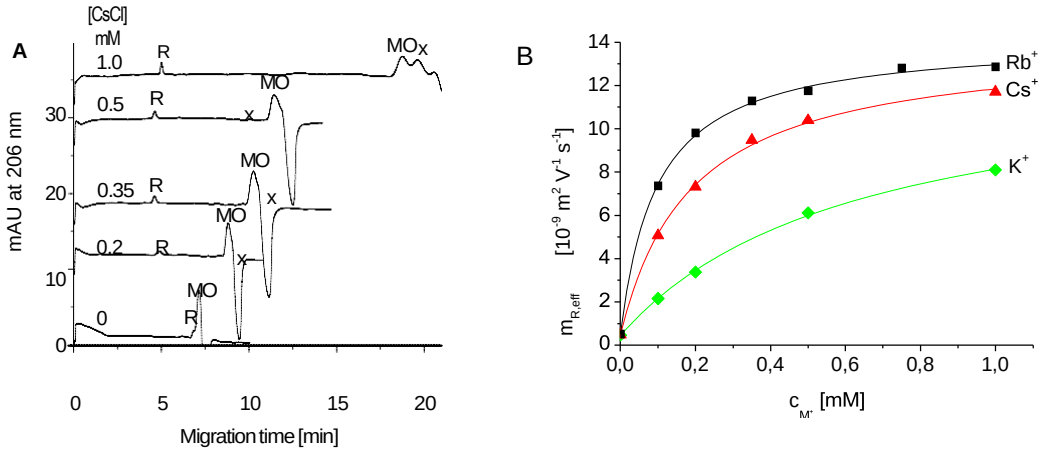


Figure 3 A) Electropherograms of hexaarylbenzene-based receptor (R) at variable concentration of Cs^+ ions in methanolic BGE 0.025 M Tris, 0.050 M chloroacetic acid, pH_{MeOH} 7.8; home-made CE analyzer with fused silica capillary, total/effective length 305/195 mm, id/od 50/375 μm, separation voltage 12 kV, 22°C, MO = mesityloxide (EOF marker). B) Dependence of effective mobility of receptor R, $m_{R,eff}$, on increasing Cs^+ ion concentration in BGE c_{M+}. Adapted from [36] and [37].

This methodology allowed quantitative evaluation of both weak and medium strong interactions the above ligands. In addition to K_b determination, in cooperation with theoretical chemist Dr. P. Toman from the Institute of Macromolecular Chemistry of the Czech Academy of Sciences, the detailed structural features of the above complexes, such as alignment of both complex constituents and interatomic distances within the complex, were determined by density functional theory.

Binding constants K_b of complexes of double stranded DNA oligonucleotide (Dickerson dodecamer) with classical DNA intercalator (ethidium bromide (EtBr)) or oligophenylene derivatives-based potential new type of DNA ligands were determined by partial filling CAE method [18]. K_b was calculated from the dependence of migration time changes of DNA oligomer (applied as analyte) on the length of ligand zones introduced beforehand as plugs of various lengths in hydroxypropylcellulose coated fused silica capillary.

1. Concluding remarks

To summarize, we believe that the above description of our work in the period 2010-2014 demonstrates that significant results were achieved in the research and development of instrumentation, methodology and applications of HPCE methods. A novel special HPCE apparatus with multidimensional detection system was constructed, which broadened the detectability, increased sensitivity and enlarged applicability of the CE methods to qualitative and quantitative analysis of a wide range of compounds, independently of the presence or absence of chromophores or fluorophores in their (macro)molecules. New methodology in several HPCE modes was developed and applied to high-efficient separation, high-sensitive qualitative and quantitative analysis and determination of important physicochemical characteristics of relevant biologically and pharmaceutically active compounds as well as of new types of functional organic molecules.

2. References

1. Ehala, S., Grishina, A. A., Sheshenev, A. E., Lyapkalo, I. M., Kasicka, V. (2010) Determination of acid-base dissociation constants of very weak zwitterionic heterocyclic bases by capillary zone electrophoresis. *J. Chromatogr. A*, 1217 (51), 8048-8053.
2. Solinova, V., Kasicka, V. (2013) Determination of acidity constants and ionic mobilities of polyprotic peptide hormones by CZE. *Electrophoresis*, 34 (18), 2655-2665.
3. Sazelova, P., Kasicka, V., Leon, C., Ibanez, E., Cifuentes, A. (2012) Capillary electrophoretic profiling of tryptic digests of water soluble proteins from Bacillus thuringiensis-transgenic and non-transgenic maize species. *Food Chem.*, 134 (3), 1607-1615.
4. Solinova, V., Poitevin, M., Koval, D., Busnel, J. M., Peltre, G., Kasicka, V. (2012) Capillary electrophoresis in classical and carrier ampholytes-based background electrolytes applied to separation and characterization of gonadotropin-releasing hormones. *J. Chromatogr. A*, 1267, 231-238.
5. Napagoda, M., Rulisek, L., Jancarik, A., Klivar, J., Samal, M., Stara, I. G., Stary, I., Solinova, V., Kasicka, V., Svatos, A. (2013) Azahelicene Superbases as MAILD Matrices for Acidic Analytes. *ChemPlusChem*, 78 (9), 937-942.
6. Ibrahim, A., Koval, D., Kasicka, V., Faye, C., Cottet, H. (2013) Effective Charge Determination of Dendrigraft Poly-L-lysine by Capillary Isotachophoresis. *Macromolecules*, 46 (2), 533-540.
7. Chamieh, J., Koval, D., Besson, A., Kasicka, V., Cottet, H. (2014) Generalized polymer effective charge measurement by capillary isotachophoresis. *J. Chromatogr. A*, 1370, 255-262.
8. Koval, D., Kasicka, V., Cottet, H. (2011) Analysis of glycated hemoglobin A1c by capillary electrophoresis and capillary isoelectric focusing. *Anal. Biochem.*, 413 (1), 8-15.
9. Koval, D., Severa, L., Adriaenssens, L., Vavra, J., Teply, F., Kasicka, V. (2011) Chiral analysis of helquats by capillary electrophoresis: Resolution of helical N-heteroaromatic dications using randomly sulfated cyclodextrins. *Electrophoresis*, 32 (19), 2683-2692.
10. Severa, L., Koval, D., Novotna, P., Oncak, M., Sazelova, P., Saman, D., Slavicek, P., Urbanova, M., Kasicka, V., Teply, F. (2010) Resolution of a configurationally stable [5]helquat: enantiocomposition analysis of a helicene congener by capillary electrophoresis. *New J. Chem.*, 34 (6), 1063-1067.

11. Vavra, J., Severa, L., Svec, P., Cisarova, I., Koval, D., Sazelova, P., Kasicka, V., Teply, F. (2012) Preferential Crystallization of a Helicene-Viologen Hybrid - An Efficient Method to Resolve [5]Helquat Enantiomers on a 20 g Scale. *Eur. J. Org. Chem.*, (3), 489-499.
12. Adriaenssens, L., Severa, L., Koval, D., Cisarova, I., Belmonte, M. M., Escudero-Adan, E. C., Novotna, P., Sazelova, P., Vavra, J., Pohl, R., Saman, D., Urbanova, M., Kasicka, V., Teply, F. (2011) [6]Saddlequat: a [6]helquat captured on its racemization pathway. *Chem. Sci.* 2 (12), 2314-2320.
13. Severa, L., Oncak, M., Koval, D., Pohl, R., Saman, D., Cisarova, I., Reyes-Gutierrez, P. E., Sazelova, P., Kasicka, V., Teply, F., Slavicek, P. (2012) A Chiral Dicationic [8]Circulenoid: Photochemical Origin and Facile Thermal Conversion into a Helicene Congener. *Angew. Chem. Int. Ed.*, 51 (48), 11972-11976.
14. Ehala, S., Marklik, E., Toman, P., Kasicka, V. (2010) ACE applied to the quantitative characterization of benzo-18-crown-6-ether binding with alkali metal ions in a methanol-water solvent system. *Electrophoresis*, 31 (4), 702-708.
15. Ehala, S., Toman, P., Makrlik, E., Rathore, R., Kasicka, V. (2011) Affinity capillary electrophoresis and density functional theory applied to binding constant determination and structure elucidation of hexaarylbenzene-based receptor complex with ammonium cation. *J. Chromatogr. A*, 1218 (30), 4982-4987.
16. Solinova, V., Kaiser, M. M., Lukac, M., Janeba, Z., Kasicka, V. (2014) Enantiopurity analysis of new types of acyclic nucleoside phosphonates by capillary electrophoresis with cyclodextrins as chiral selectors. *J. Sep. Sci.*, 37 (3), 295-303.
17. Krecmerova, M., Jansa, P., Dracinsky, M., Sazelova, P., Kasicka, V., Neyts, J., Auwerx, J., Kiss, E., Goris, N., Stepan, G., Janeba, Z. (2013) 9-[2-(R)-(Phosphonomethoxy)propyl]-2,6-diaminopurine, (R)-PMPDAP, and its prodrugs: Optimized preparation, including identification of by-products formed, and antiviral evaluation in vitro. *Bioorg. & Med. Chem.*, 21 (5), 1199-1208.
18. Ruzicka, M., Cizkova, M., Jirasek, M., Teply, F., Koval, D., Kasicka, V. (2014) Study of deoxyribonucleic acid-ligand interactions by partial filling affinity capillary electrophoresis. *J. Chromatogr. A*, 1349, 116-121.
19. Ehala, S., Toman, P., Rathore, R., Makrlik, E., Kasicka, V. (2011) Affinity capillary electrophoresis and quantum mechanical calculations applied to the investigation of hexaarylbenzene-based receptor binding with lithium ion. *J. Sep. Sci.*, 34 (18), 2433-2440.
20. Rezanka, P., Ehala, S., Koktan, J., Sykora, D., Zvatora, P., Vosmanska, M., Kral, V., Miksik, I., Cerovsky, V., Kasicka, V. (2012) Application of bare gold nanoparticles in open-tubular CEC separations of polyaromatic hydrocarbons and peptides. *J. Sep. Sci.*, 35 (1), 73-78.
21. Miksik, I., Lacinova, K., Zmatlikova, Z., Sedlakova, P., Kral, V., Sykora, D., Rezanka, P., Kasicka, V. (2012) Open-tubular capillary electrochromatography with bare gold nanoparticles-based stationary phase applied to separation of trypsin digested native and glycated proteins. *J. Sep. Sci.*, 35 (8), 994-1002.
22. Rezanka, P., Navratilova, K., Zvatora, P., Sykora, D., Matejka, P., Miksik, I., Kasicka, V., Kral, V. (2011) Cyclodextrin modified gold nanoparticles-based open-tubular capillary electrochromatographic separations of polyaromatic hydrocarbons. *J. Nanopart. Res.*, 13 (11), 5947-5957.

23. Adriaenssens, L., Severa, L., Salova, T., Cisarova, I., Pohl, R., Saman, D., Rocha, S. V., Finney, N. S., Pospisil, L., Slavicek, P., Teply, F. (2009) Helquats: A Facile, Modular, Scalable Route to Novel Helical Dications. *Chem. Eur. J.* 15 (5), 1072-1076.
24. Severa, L., Jirasek, M., Svec, P., Teply, F., Revesz, A., Schroder, D., Koval, D., Kasicka, V., Cisarova, I., Saman, D. (2012) Counterion-Induced Inversion of Conformer Stability of a [5]Helquat Dication. *ChemPlusChem*, 77 (8), 624-635.
25. Cizkova, M., Saman, D., Koval, D., Kasicka, V., Klepetarova, B., Cisarova, I., Teply, F. (2014) Modular Synthesis of Helicene-Like Compounds Based on the Imidazolium Motif. *Eur. J. Org. Chem.*, (26), 5681-5685.
26. Holy, A. (2006) Antiviral acyclic nucleoside phosphonates structure activity studies. *Antiviral Research*, 71 (2-3), 248-253.
27. Kaiser, M. M., Jansa, P., Dracinsky, M., Janeba, Z. (2012) A novel type of acyclic nucleoside phosphonates derived from 2-(phosphonomethoxy)propanoic acid. *Tetrahedron*, 68 (21), 4003-4012.
28. Hlavacek, J., Vitovcova, M., Sazelova, P., Picha, J., Vanek, V., Budesinsky, M., Jiracek, J., Gillner, D. M., Holz, R. C., Miksik, I., Kasicka, V. (2014) Mono-N-acyl-2,6-diaminopimelic acid derivatives: Analysis by electromigration and spectroscopic methods and examination of enzyme inhibitory activity. *Anal. Biochem.*, 467, 4-13.
29. Niederhafner, P., Bednarova, L., Budesinsky, M., Safarik, M., Ehala, S., Jezek, J., Borovickova, L., Fucik, V., Cerovsky, V., Slaninova, J. (2010) Melectin MAPs: the influence of dendrimerization on antimicrobial and hemolytic activity. *Amino Acids*, 39 (5), 1553-1561.
30. Niederhafner, P., Gut, V., Jezek, J., Budesinsky, M., Kasicka, V., Wunsch, E., Hlavacek, J. (2010) Synthetic study on cystinyl peptides using solution and solid phase methodology: human IgG1 hinge region. *Amino Acids*, 39 (3), 641-650.
31. Kasicka, V. (2010) Recent advances in CE and CEC of peptides (2007-2009). *Electrophoresis*, 31 (1), 122-146.
32. Kasicka, V. (2012) Recent developments in CE and CEC of peptides (2009-2011). *Electrophoresis*, 33 (1), 48-73.
33. Kasicka, V. (2014) Recent developments in capillary and microchip electroseparations of peptides (2011–2013). *Electrophoresis*, 35 (1), 69-95.
34. Stepanova, S., Kasicka, V. (2014) Determination of impurities and counterions of pharmaceuticals by capillary electromigration methods. *J. Sep. Sci.*, 37 (15), 2039-2055.
35. Misek, J., Teply, F., Stara, I. G., Tichy, M., Saman, D., Cisarova, I., Vojtisek, P., Stary, I. (2008) A straightforward route to helically chiral N-heteroaromatic compounds: Practical synthesis of racemic 1,14-diaza[5]helicene and optically pure 1-and 2-aza[6]helicenes. *Angew. Chem. Int. Ed.*, 47 (17), 3188-3191.
36. Ehala, S., Toman, P., Rathore, R., Makrlik, E., Kasicka, V. (2011) Affinity capillary electrophoresis and density functional theory employed for the characterization of hexaarylbenzene-based receptor complexation with alkali metal ions. *Electrophoresis*, 32 (9), 981-987.
37. Ehala, S., Toman, P., Makrlik, E., Rathore, R., Kasicka, V. (2012) Combined Theoretical and Experimental Study of the Complexation of a Hexaarylbenzene-Based Receptor with the Potassium Cation. *J. Sol. Chem.*, 41 (10), 1812-1824.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Jiří Jiráček - Chemistry and Biology of Insulin and Insulin-like Growth Factors

During the period 01/2010 – 12/2011, **Jiří Jiráček's group** consisted of two scientifically independent teams: the **team of Jiří Jiráček** and the **team of Lenka Maletínská**. Starting from 01/2012 Lenka Maletínská was awarded a targeted reserch group leader in IOCB. Therefore, Maletínská´'s team scientific activities and achievements are reported here separately and only for the period 01/2010 – 12/2011.

At the beginning of 2010, **J. Jiráček's team** started with three main research topics:

(i) Inhibitors for metalloenzymes - design and synthesis to study the functions of enzymes,
(ii) Proteomics - to investigate the mechanisms of breast cancer and insect immunity, and
(iii) Insulin analogues - for diabetes and insulin receptor studies.

A few years ago, we realized that our group deals with too many different scientific topics and that our effort could be much more effective, particularly with respect to the quality of publications, if we focused on one specific project. Despite valuable scientific results obtained in research on metalloenzyme inhibitors and proteomics, we realized that the most effective exploitation of our experience and skills would be through performing exclusive work on the chemistry and biology of insulin and insulin-related hormones. Therefore, in 2010, we decided to focus, in the future, exclusively on **the development of novel insulin and insulin-like growth factors-I/II (IGF-I/II) analogues/mimetics in the context of their receptors' (IR/IGF-1R) structure–function relationships.** The projects dealing with proteomics were closed with the end of 2010 and projects dealing with enzyme inhibitors (and respective grants) were closed with the end of 2012 (but some results from these project were published later).

Ad (i) Inhibitors for metalloenzymes.

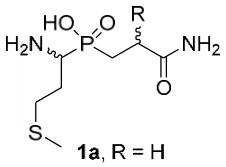
Until the end of 2012, Jiráček's group has been involved in the design, synthesis and evaluation of transition-state mimicking inhibitors for metalloenzymes. Potent and selective inhibitors may be very useful for studies of enzyme functions *in vivo*, especially if an animal model with deletion of the respective enzyme is not available.

We designed and synthesized series of phosphinate pseudopeptides (compounds **1**-**8**, Ref. 1) as potential *inhibitors for methionine aminopeptidases (MetAPs) and leucine aminopeptidases (LeuAPs)*. Inhibitors for MetAPs, in particular, could find applications as potential antibiotics or in the treatment of cancers. The compounds were tested for their inhibitory activities in the collaborating laboratories of Richard C. Holz from Loyola University in Chicago and Marcin Drag and Artur Mucha from the Technical University in Wroclaw, Poland. Some of the compounds specifically inhibited MetAPs (Ref. 2) or LeuAP (Ref. 3) and provided interesting information about the structural requirements of the active sites of the respective enzymes.

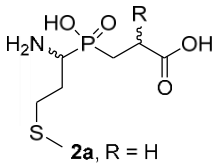
In a collaboration with Jan Hlaváček from IOCB and Richard C. Holz from Loyola University in Chicago, we also participated in the synthesis of *inhibitors for bacterial N-succinyl-L,L-diaminopimelic acid desuccinylase (DapE)* (compounds of the general formula **9**, Ref. 4) *and N^{α}-acetyl-L-ornithine deacetylase (ArgE)* (compounds of the general formula **10,** Ref. 5). These compounds may also find applications as potential antibiotics.

Since 2002, in a collaboration with Timothy A. Garrow from the University of Illinois at Urbana-Champaign, we have also been developing *inhibitors for betaine-homocysteine S-methyltransferase (BHMT)*, and, since 2010, we have been developing inhibitors for a similar enzyme, *BHMT-2*. Both enzymes are Zn-metallotransferases involved in the mammalian liver

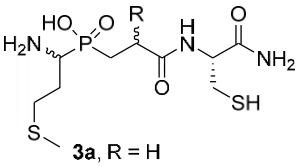
and kidney metabolism of sulfur amino acids. BHMT uses betaine and BHMT-2 *S*-methylmethionine as a methyl donor to produce methionine. Specific inhibitors may be very important tools for the studies of physiological functions of the enzymes *in vivo*.



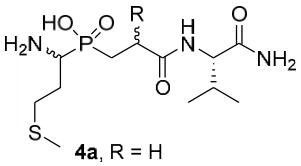
1a, R = H
1b, R = CH_3



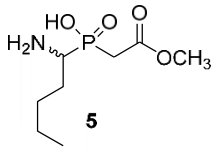
2a, R = H
2b, R = CH_3



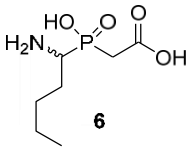
3a, R = H
3b, R = CH_3



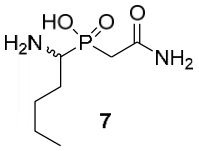
4a, R = H
4b, R = CH_3



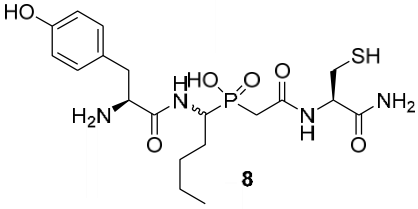
5



6

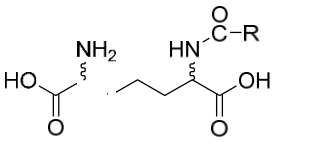


7

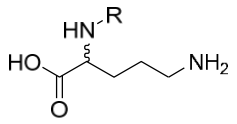


8

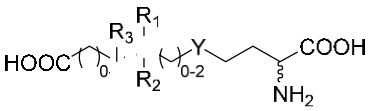
Previously, we published the synthesis of a series of BHMT inhibitors that mimic the transition state of BHMT substrates, homocysteine and methionine. The inhibitory properties of new compounds (general structure **11**) yielded interesting information about the mechanism of BHMT catalysis and the specificity of the active site. Some of the inhibitors (e.g. **12**) inhibited BHMT in the nanomolar range. Selected compounds were used in U.S. for *studies of BHMT physiological function in rats* (Ref. 6), which confirmed the important role of BHMT in sulfur metabolism *in vivo*.



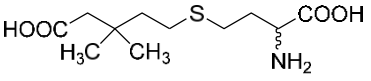
9, R = 15 different substituents



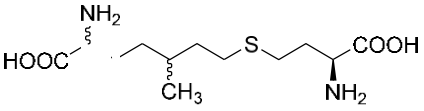
10, R = 18 different substituents



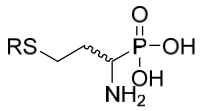
11, 20 different compounds
X is S, N, CH_2
Y is S, Se, O, CH_2
R_{1-3} is H, CH_3, n/a



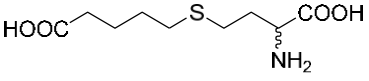
12



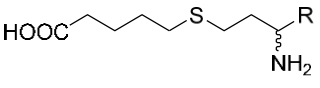
13



14a, R = Ph
14b, R = *tert*-butyl
14c, R = Bn



15



16, R = 11 different substituents

Homocysteine and methyl-group metabolic pathways are interconnected with the metabolism of numerous biological compounds, the maintenance of redox status and epigenetic regulations. The consequences of disruptions of the processes are difficult to predict. In collaboration with Apigenex s.r.o. (Prague, CZ), we employed the *2-DE and MS methodologies to study the changes of hepatocyte proteome together with changes in concentrations of key homocysteine-related metabolites after BHMT inhibition and*

homocysteine treatment. First, we used the cancer derived cell line HepG2 as a "training" model system (Ref. 7), and then we have performed a complex analysis of freshly isolated human hepatocytes in primary culture (Ref. 8, IOCB awarded publication in 2013 for biochemistry). The intra-cellular concentrations of the metabolites (homocysteine, cysteine, methionine, *S*-adenosylmethionine, *S*-adenosylhomocysteine, dimethylglycine, cystathionine, betaine and choline) changed significantly upon treatment. The changes differed between the BHMT-inhibited and homocysteine-treated cells. We also found treatment-dependent changes in intra-cellular as well as secreted proteomes of human hepatocytes. The proteomic data were confirmed with Western blotting (Figure 1). This data will contribute to the understanding of the processes behind hyperhomocysteinemia and facilitate understanding of BHMT physiological functions.

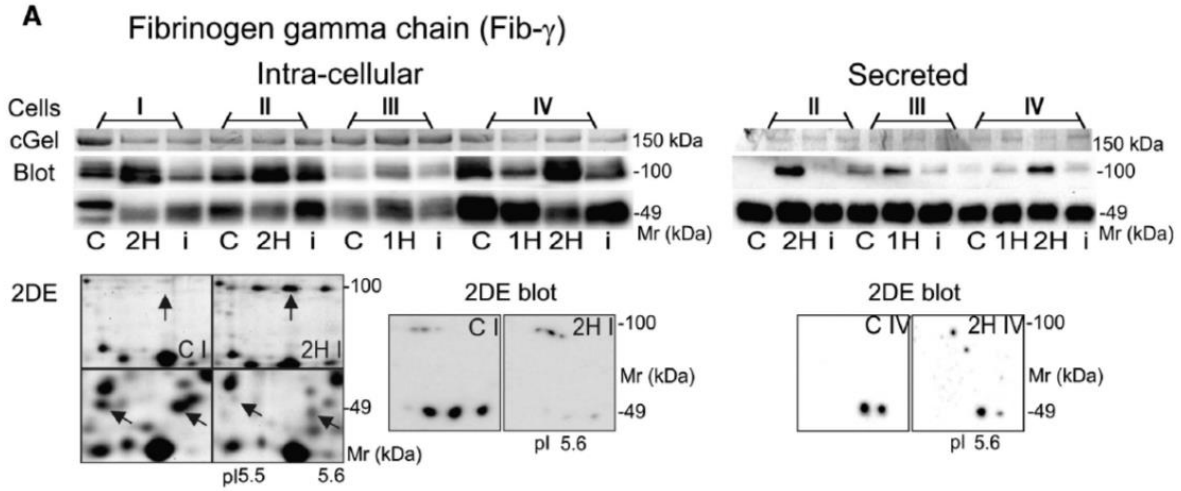


Figure 1. An illustrative example of proteomic analysis of fibrinogen gamma content in Hcy-treated and control primary hepatocytes (from Ref. 8).

We also developed the first specific *inhibitor for BHMT-2*. The work with BHMT-2 was difficult due to the very low stability of the enzyme. However, we found that BHMT-2 is stabilized by co-purification with BHMT. This finding evokes questions about the natural co-oligomerization of both enzymes in the cell. Stabilization of BHMT-2 has allowed us conduct inhibition studies with the enzyme. We designed and prepared a first series of BHMT-2 inhibitors, and we found that compound **13** is a potent BHMT-2 inhibitor (K_i of approximately 77 nM). Moreover, compound **13** only weakly inhibits BHMT (IC_{50} of approximately 77 μM) (Ref. 9, IOCB awarded publication in 2012 for Biochemistry).

We have also developed an efficient methodology for the synthesis of *S-substituted derivatives of phosphonohomocysteine* (compounds of the general formula **14**) as precursors for the synthesis of *new BHMT inhibitors* (Ref. 10, invited and dedicated article).

In the next study (Ref. 11), we systematically investigated the tolerance of BHMT for modifications at the "homocysteine" part of the previously reported potent inhibitor **15**. In the new compounds of the general formula **16**, which are S-alkylated homocysteine derivatives, we replaced the carboxylic group in the "homocysteine" part of inhibitor **15** with different chemical moieties. This study helped to completely *define the structural requirements of the active site of BHMT* and revealed the remarkable selectivity of the enzyme for homocysteine.

Finally, in collaboration with Pavel Jungwirth's group (IOCB) and with two U.S. laboratories, we discovered and thoroughly studied *BHMT activation by potassium ions*. In brief, theoretical chemists predicted the binding site of K^+ in BHMT and this was confirmed

by crystal structure of the enzyme and kinetic experiments. This work (Ref. 12) is a nice example of an effective collaboration of theoretical and experimental scientists.

Ad (ii) Proteomics.

Until the end of 2010, we were involved in proteomics studies of breast cancer cells and insect immunity. In the *breast cancer proteomics* project, we compared the phenotypes of breast cell lines EM-G3, HCC1937, MCF7 and MDA-MB-231 using 2D-electrophoresis (2-DE). Transformed phenotypes are common to cell lines derived from various cancers. Proteome profiling is a valuable tool that may reveal uncharacteristic cell phenotypes in transformed cells. Changes in expression of glutathione *S*-transferases (GSTs) and other proteins that interact with glutathione (GSH) in model cell lines may be of particular interest. Our results confirmed the applicability of GSH-targeted affinity chromatography to proteome profiling and allowed us to characterize the phenotypes of four breast cancer cell lines (Ref. 13). In the *insect immunity proteomics* project, we performed a 2D-electrophoretic comparison of the proteomes of hemocytes and fat bodies of infected and intact *S. bullata* larvae (Ref. 14).

Ad (iii) Insulin analogues.

We believe that, over the past several years, we have achieved extensive expertise in the field of design, synthesis and evaluation of *insulin analogues* for diabetes and insulin receptor studies, and we have obtained several exciting scientific results, which are very promising for the current and future development of this scientific topic by our team.

Insulin is a key protein hormone that regulates blood glucose levels, and it thus has widespread impact on lipid and protein metabolism. Despite numerous years of scientific effort, the structure of the insulin–insulin receptor (IR) complex is still only partly understood. Insulin analogues may be useful for studies of the interaction of the hormone with IR. Moreover, modified insulins with a faster onset of action, as well as analogues with sustained action along with a flatter time–action, profile are required for better control of glycemia in diabetic patients.

To this point, we have designed, prepared and characterized about 100 insulin analogues. At first, we focused on *insulin analogues modified at the C-terminus of the B-chain*, which is one of the key binding epitopes in insulin molecule. We prepared and tested binding affinities of a series of *insulin analogues modified at position B26*. Some of these analogues, especially B26-*N*-methylated or with D-amino acids at B26, are among the most potent insulins ever reported, being 4-fold stronger in binding to IR than human insulin. During the past five years, we have established an excellent collaboration with the structural biology group of Andrzej M. Brzozowski from the University of York in the U.K., with a focus on the *elucidation of an active structure of insulin* that it must adopt upon binding to IR. The description of crystal structures of the above-mentioned B26-modified analogues was certainly a breakthrough in our insulin research (Ref. 15, selected as a significant IOCB result in 2010). Here, we reported highly potent insulin analogues with previously undescribed structures at the C- (B26-turn in Figure 2) and at N-termini (newly observed I-state, B2 in Figure 2) of insulin's B-chain. We postulated that the structures of these analogues might reflect insulin's active conformation upon binding to IR.

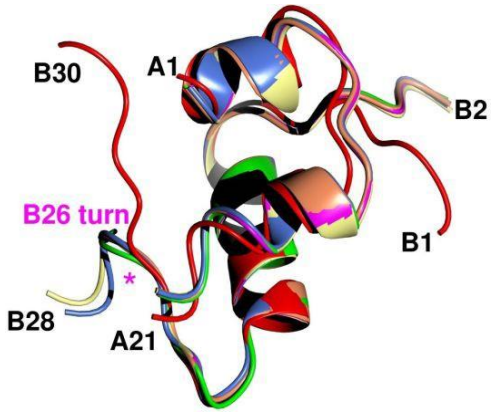


Figure 2. An overlay of the general fold of human insulin (red) and several analogues *N*-methylated at

B26. The B26-turn is marked by an asterisk, and the new intermediate conformation (I-state) of the *N*-terminus is labeled as B2. From Ref. 15.

Monomeric behavior of insulin is a necessary condition for the formation of its active form. Therefore, using isothermal microcalorimetry and X-ray crystallography, we investigated insulin analogues with selective *N*-methylations of peptide bond amides at positions B24, B25 or B26 to delineate their structural and functional contribution to the dimer interface (Ref. 16). All *N*-methylated analogues showed impaired binding affinities to IR, which suggests a direct IR-interacting role for the respective amide hydrogens. Moreover, selective *N*-methylations resulted in reduced dimerization abilities. This effect may be attributed to the loss of *inter*- and *intra*-molecular hydrogen bonding in insulin dimer.

In collaboration with the University of Catanzaro (Italy) and Thomas Jefferson University (Philadelphia) laboratories, we participated in the study of the *mechanism of enhanced mitogenic properties of insulin-like growth factor II* (IGF-II) mediated by isoform A of IR (Ref. 17). The results suggest that the lower affinity of IGF-II for the IR-A promotes lower IR-A phosphorylation and activation of early downstream effectors, vis-à-vis insulin, but may protect IR-A and IRS-1 from down-regulation, thereby evoking sustained and robust mitogenic stimuli. Similar results were also obtained in experiments using our analogue, *N*MeTyrB26-insulin, an insulin analogue with IR-A binding affinity similar to that of IGF-II.

Next, we further studied *the extent of the detachment of the C-terminus of the B-chain of insulin during binding to IR* (Ref. 18, IOCB awarded publication in 2013 for Interdisciplinary research and selected as a significant IOCB result in 2013). We focused on the Phe at B24, which is crucial for binding to IR. We found that the structural integrity of the B24 position of insulin is essential for its activity; i.e. insulin binding pocked for PheB24 side chain must be filled in by an aromatic side chain. In other words; the data suggest limited, B25-B30 only, unfolding of the C terminus of the B-chain upon insulin activation and further support our theory of the B26-turn.

In December 2010, we established a collaboration with Michael C. Lawrence (WEHI, Parkville, Australia), who is associated with a world-leading laboratory in structural biology of IR and IGF-1R. Our collaboration (WEHI/IOCB/York collaboration agreement from December 2010) yielded the long-awaited first structures of insulin-IR complexes including a complex of IR with our D-ProB26-DTIA insulin (Ref. 19, selected as a significant IOCB result in 2013). However, this undoubtedly breakthrough result in insulin/IR research provided only a partial information about binding of both molecules because the interaction involves only Site 1 of IR (L1 and α-CT domains) and the information about interaction with Site 2 of IR (FnIII domains) is missing (Figure 3). Moreover, both flexible parts of insulin, the C- and N-termini of the B-chain are not visible in the complex. Hence, further studies were needed for elucidation of insulin active (IR-bound) conformation.

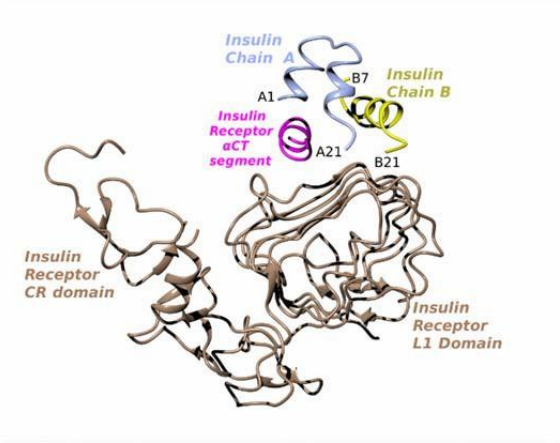


Figure 3. Structure of insulin bound to L1 and α-CT domains of IR (adapted from Ref. 19).

In the next study (Ref. 20), we prepared a series of insulin analogs modified at B26 by natural amino acids. Remarkably, AsnB26- insulin was highly active and displayed significant selectivity for the metabolic B isoform of IR. MD simulations indicated that Asn

side chain should be repulsed from insulin B26 binding pocket. Accordingly, the crystal structure of the analogue showed a bended conformation at positions B25-B27, which was fully compatible with IR receptor complex overpassing the α-CT peptide and IR glycosylation site (Figure 4 left). Moreover, simultaneously with the publishing of our article, M.C. Lawrence published (PNAS 111, E3395, 2014) the refinement of the crystal structure of the native insulin B-chain C-terminus. Comparison of both structures (Figure 5 right) reveals high similarity of both conformations and validated our findings with B24- (PheB24 structural integrity) and B26-modified (B26-turn) analogues.

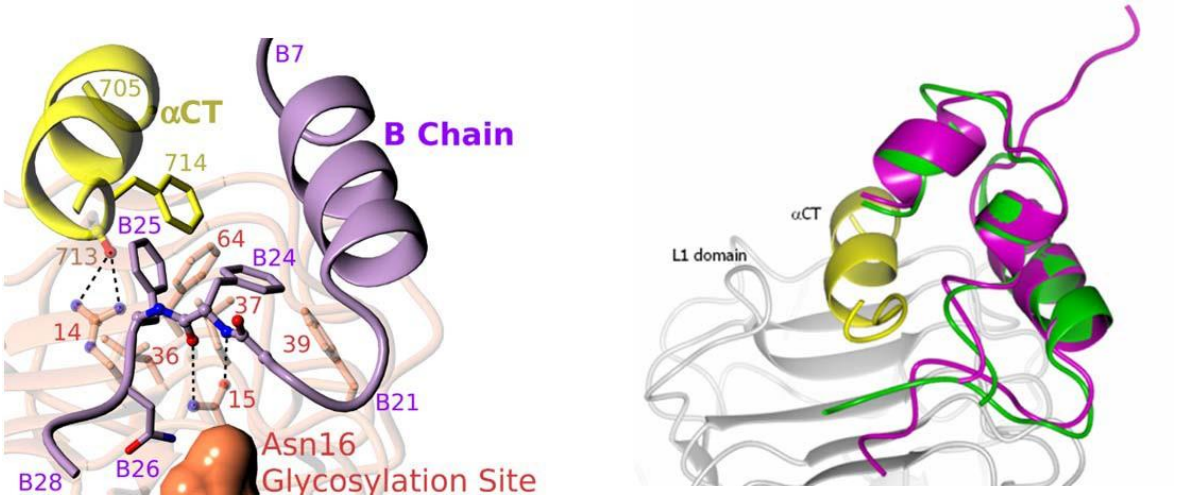


Figure 4. Left. Superposition of AsnB26-insulin B-chain (violet) to the IR complex (from Ref 19). **Right.** An overlay of AsnB26-insulin (violet) with IR-complexed human insulin (green, adapted from PNAS 111, E3395, 2014).

In the next study (Ref. 21), we focused on the elucidation of *the role of the T/R/I transition* (Figure 5) of the *N*-terminus of the B-chain of insulin in binding to the IR. The goal of the project was to prepare insulin analogues with stabilized T or R conformations at the *N*-terminus of the B-chain and to determine the biological potency of the respective analogues. Via total chemical synthesis, we prepared several insulin analogues with achiral Aib (α-aminoisobutyric acid exhibiting a strong helical propensity) at position B3, B5 or B8 to structurally lock the R-state. For the stabilization of the T-conformation we prepared analogues with D-Pro and N-Me-Ala (turn-inducing amino acids) at B8. The results of this study suggest that the R-state is incompatible with high insulin activity. On the other hand T- or I-states could represent IR-bound conformation of the hormone but under the condition of high flexibility of this part of insulin which is naturally provided by flexible GlyB8.

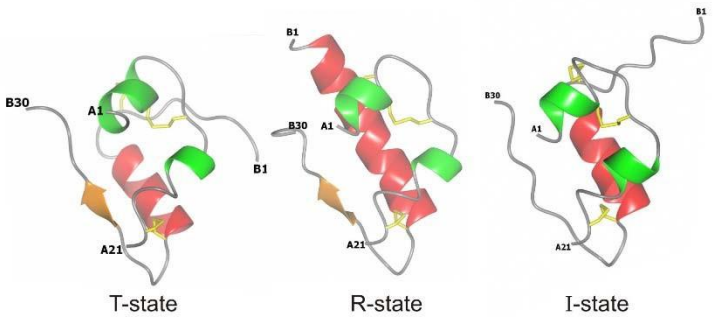


Figure 5. Schematic representation of the 3-D-fold of the insulin monomer. The A-chain is shown in green, and the B-chain is shown in red. Structures of insulin in the T-state (left), R-state (center) and the I-state (right) highlight the different conformations of residues B1-B6 from the *N*-terminus of the B-chain.

In the next study (Ref. 22), we chemically synthesized [GlnB22]-insulin, which is a naturally occurring insulin mutant causing diabetes in MODY (diabetes variant) patients. We characterized its biological and structural properties. The chemical synthesis of this insulin analogue revealed that its folding ability is severely impaired. In vitro and in vivo tests showed that its binding affinity and biological activity are reduced (both approximately 20% that of human insulin). Comparison of the NMR structure of [GlnB22]-insulin with the structure of native human insulin revealed that the most significant structural effect of the mutation is distortion of the B20-B23 β-turn. The partially disordered [GlnB22]-insulin structure appears to be one reason for the reduced binding potency of this mutant and may also be responsible for its low folding efficiency in vivo.

After our discovery that insulin analogues with β-turn structures at B25-B27 (B26-turn, Figures 2 and 4) may be highly active in binding to IR, we decided to synthesize a series of *insulins with covalently stabilized cyclic structures at the C-terminus of the B-chain*. The original strategy and aim of this research was to probe whether insulin-active-like conformation (seen in its IR complex and highly-active analogues), i.e. the bend of the B23-B30 chain around B25-B27 sites, can be mimicked and stabilized by a covalent bridge performed within this region of the hormone. We have chosen Huisgen Cu^{1+}-catalyzed azide-alkyne cycloaddition to form a triazole-bridge between the selected residues in the B-chain. The main steric parameters were: (i) positions of alkyne and azido precursors within B23-B30 chain, (ii) the length of the crosslinked protein main chain, (iii) the chirality or the precursors, and (iv) the position of the triazole ring (i.e. up- or down-stream of the main chain). Till now, this work yielded 14 novel, systematically crosslinked insulin analogues, which were structurally and functionally characterized (three crystal and two NMR structures). Most interestingly, one of the analogues, insulin with triazole bridge between B26 and B29 sites, was highly active in binding to both isoforms of IR but with a significant preference for the IR-B (217% for IR-A and 550% for IR-B). Our results indicate the potential of chemistry-driven modulation of insulin function, shedding also light on the functional importance of hormone's B-chain C-terminus for its IR-B specificity. These unpublished and confidential results will be published in 2015. We have already submitted a Utility model, which was awarded in 2013 (U1-24985) and submitted a Czech patent application (PV 2014-450).

In 2013, we started our first attempts for the *chemical and recombinant production of IGF-I and IGF-II analogues*. To this end, with have already successfully prepared by total chemical synthesis two-chain IGF-I with a cleavage between residues Arg36 and Arg37 but also a single-chain IGF-1 with a triazole bridge between Arg36 and Arg37. In collaboration with the IOCB Structural Biology group (Václav Veverka) we have also successfully developed the recombinant production of these hormones and we have already prepared and characterized first IGF-I and -II analogues Theses topics will be described in detail in Future plans.

In 2014, we have started a project on the combinatorial synthesis of non-peptide insulin/IGF mimetics. This project will be described in detail in Future plans.

We have also optimized the synthesis of chiral shift reagents Sm-(R)-pdta and Sm-(S)-pdta useful for NMR determination of absolute configurations of water soluble compounds as amino acids (Ref. 23 and awarded Czech patent B6-304433). Finally, Václav Vaněk from our group also participated in an elegant study of Pavel Jungwirth's group (IOCB) showing that Coulomb explosion occurs during the early stages of the reaction of alkali metals with water (Ref. 24).

On the next pages, we will describe the scientific achievements of **Lenka Maletínská's team** during the period 2010-2011 when they belonged to Jiráček's group.

In 2010-2011, L. Maletínská was the coordinator of two projects, and B. Železná was the coordinator of one project of the Grant Agency of the Czech Republic). The projects dealt with food-intake-regulating peptides: anorexigenic **cocaine- and amphetamine-regulating transcript (CART) peptide** (project 303/10/1368), anorexigenic **prolactin-releasing peptide (PrRP)** (project 303/10/1368), both of which are neuropeptides that are produced and act in the brain, and the stomach-born centrally acting orexigenic hormone **ghrelin** (303/09/0744).

Food-intake-regulating peptides not only affect hunger or satiety but also energy homeostasis; they also affect metabolic disturbances such as leptin and insulin resistance, which results in so-called metabolic syndrome. **Metabolic syndrome** includes metabolic disturbances that lead to obesity and type 2 diabetes; it is currently a major public health problem associated with morbidity and mortality and continues to increase worldwide. Recently discovered anorexigenic neuropeptides involved in food intake regulation, such as PrRP and CART peptides, as well as antagonists of orexigenic ghrelin represent new possibilities in the development of future **anti-obesity agents**. We studied the previously mentioned peptides both ***in vitro*** in physiologically relevant cell lines that express the respective peptide receptors and ***in vivo*** in mouse models of obesity and leptin and insulin resistance. Our research has a strongly multidisciplinary character: it involves peptide chemistry, biochemistry, physiology and pharmacology. Our cooperation with the Faculty of Medicine, Charles University and our international collaborations allow us to study a variety of aspects of peptides involved in food intake and energy metabolism and to contribute to obesity research.

The relevance of **CART peptides** among central anorexigenic neuropeptides results from the fact that CART is one of the most abundant transcripts in the brain. Despite enormous effort, no receptors for CART peptides have been isolated and cloned to date. For CART peptide, we reported the synergistic and long-lasting effect of intracerebroventricularly administered peptide CART(61-102) and intraperitoneally administered satiety signal cholecystokinin octapeptide (CCK-8) on food intake and Fos immunoreactivity in hypothalamus and brainstem. We found that CART peptide and CCK-8 have a synergistic anorexigenic effect and additively activate their common food-intake-regulating targets: paraventricular and dorsomedial nuclei of the hypothalamus and nucleus tractus solitarii of the brainstem. The synergistic effect of CART(61-102) and CCK-8 *via* CCK-A receptor on the regulation of food intake points to their integrated action in the central nervous system (Ref. 25).

The magnitude of the anorexigenic effect of **PrRP** is clearly demonstrated by the fact that both PrRP and PrRP receptor knock-out mice are hyperphagic and obese. We compared full-length PrRP31 and its shorter fragments PrRP20 and PrRP13 by specific binding, MAPK/ERK and CREB phosphorylation, and prolactin release in three rodent tumor pituitary cell lines: GH3, AtT20 and RC-4B/C cells; we also investigated the anorexigenic effect in mice after their intracerebroventricular administration to find the minimal active core of the molecule (Ref. 26). Based on the first study, we synthesized more stable PrRP20 analogs with the C-terminal Phe replaced with non-coded amino acids. We found that the *in vitro* biological effects of the analogs were comparable with those of PrRP20 and that the anorexigenic effects in fasted mice were even enhanced and longer-lasting (Ref. 27). In particular, [PheNO$_2^{31}$]PrRP20, [1-Nal^{31}]PrRP20, [2-Nal^{31}]PrRP20 and [Tyr^{31}]PrRP20 showed not only in vitro effects comparable or greater than those of PrRP20, but also a very significant and long-lasting anorexigenic effect after central administration in fasted mice.

Ghrelin is the only orexigenic peptide of gut origin that acts in the brain. Therefore, its agonists are promising agents for treatment of cachexia in chronically or terminally ill patients. We found that peripherally administered ghrelin and its various peptide agonists

significantly activated neurons in food-intake-responsible areas, such as arcuate and paraventricular nuclei of the hypothalamus, the nucleus of the solitary tract, and area postrema (Ref. 28).

Ghrelin is 28-peptide and the only hormone whose hydroxyl group of its serine residue in position 3 is acylated by *n*-octanoic acid, which is necessary for its biological activity. Because of the lability of the ester bond in Ser-O-octanoyl in position 3 of ghrelin, only 10-20% of ghrelin in blood is octanoylated. To stabilize the ghrelin molecule, we designed ghrelin analogs with an amide bond instead of an ester bond between Ser^3 and octanoic acid and with non-coded amino acids in the peptide chain. The synthesized analogs were ghrelin agonists with affinity to the receptor and with orexigenic effects similar to ghrelin but with a substantially higher stability (Ref. 29, IOCB awarded publication in 2012 for Biochemistry).

In the other part of the ghrelin-related project, we followed the effects of ghrelin antagonists and inverse agonists on short- and long-term food intake and energy utilization in **ovariectomized (OVX) mice with DIO** and compared them with those under estrogen supplementation**.** Ghrelin antagonists and inverse agonists seem to be promising compounds in anti-obesity treatment; we therefore hypothesized that they could act beneficially at DIO of females combined with estrogen deficiency and could be promising for the treatment of metabolic syndrome in women after menopause. We showed that peptidic ghrelin antagonist [D-Lys^3]GHRP-6 attenuated body weight and insulin resistance in OVX mice with DIO (Ref. 30), a model of common menopausal obesity in women, where estrogen deficiency and a high-fat diet are conditions of central resistance to anorexigenic hormone leptin (Ref. 31) and augmented ghrelin activity.

References

1. Pícha, J., Buděšínský, M., Fiedler, P., Šanda, M., Jiráček, J.* (2010) Synthesis of alpha-carboxyphosphinopeptides derived from norleucine. *Amino Acids*, **39**, 1265-1280.
2. Poreba, M., Gajda, A., Pícha, J., Jiráček, J., Marschner, A., Klein, C. D., Salvesen, G. S., Drag, M.* (2012) S1 pocket fingerprints of human and bacterial methionine aminopeptidases determined using fluorogenic libraries of substrates and phosphorus based inhibitors. *Biochimie*, **94**, 704-710.
3. Pícha, J., Liboska, R., Buděšínský, M., Jiráček, J., Pawelczak, M., Mucha, A.* (2011) Unusual activity pattern of leucine aminopeptidase inhibitors based on phosphorus containing derivatives of methionine and norleucine. *J. Enzym. Inhib. Med. Chem.*, **26**, 155-161.
4. Vaněk, V., Pícha, J., Buděšínský, M., Šanda, M., Jiráček, J., Holz, R. C., Hlaváček, J.* (2010) Synthesis of N-succinyl-L,L-diaminopimelic acid mimetics via selective protection. *Protein Pept. Lett.*, **17**, 405-409.
5. Hlaváček, J.*, Pícha, J., Vaněk, V., Jiráček, J., Slaninová, J., Fučík, V., Buděšínský, M., Gilner, D., Holz, R. C. (2010) Inhibitors of N (alpha)-acetyl-L-ornithine deacetylase: synthesis, characterization and analysis of their inhibitory potency. *Amino Acids*, **38**, 1155-1164.
6. Straková, J., Williams, K. T., Gupta, S., Schalinske, K. L., Kruger, W. D., Rozen, R., Jiráček, J., Li, L., Garrow, T. A.* (2010) Dietary intake of S-(alpha-carboxybutyl)-DL-homocysteine induces hyperhomocysteinemia in rats. *Nutr. Res.*, **30**, 492-500.
7. Kořínek, M., Šístek, V., Mládková, J., Mikeš, P., Jiráček, J., Selicharová, I.* (2012) Quantification of homocysteine-related metabolites and the role of betaine–homocysteine S-methyltransferase in HepG2 cells. *Biomed. Chromatogr.*, **27**, 111-121.
8. Selicharová, I.*, Kořínek, M., Demianová, Z., Chrudinová, M., Mládková, J., Jiráček, J. (2013) Effects of hyperhomocysteinemia and betaine-homocysteine S-methyltransferase inhibition on hepatocyte metabolites and the proteome. *Biochim. Biophys. Acta*, **1834**, 1596-1606.
9. Mládková, J., Vaněk, V., Buděšínský, M., Elbert, T., Demianová, Z., Garrow, T.A., Jiráček, J.* (2012) Double-headed sulfur-linked amino acids as first inhibitors for betaine-homocysteine S-methyltransferase 2. *J. Med. Chem.*, **55**, 6822-6831.

10. Pícha, J., Buděšínský, M., Fiedler, P., Jiráček, J.* (2012) The synthesis of phosphonic acids derived from homocysteine via transesterification reactions (PK-6943DP). *Arkivoc*, 2012, **(iv)**, 80-99. *Dedicated and invited article.*
11. Pícha, J., Vaněk, V., Buděšínský, M., Mládková, J., Garrow, T.A., Jiráček, J.* (2013) The development of a new class of inhibitors for betaine-homocysteine S-methyltransferase. *Eur. J. Med. Chem.*, **65**, 256-275.
12. Mládková, J., Hladílková, J., Diamond, C.E., Tryon, K., Yamada, K., Garrow, T.A., Jungwirth, P., Koutmos, M., Jiráček, J.* (2014) Specific potassium ion interactions facilitate homocysteine binding to betaine-homocysteine S-methyltransferase. *Proteins*, **82**, 2552-2564.
13. Mládková, J., Šanda, M., Matoušková, E., Selicharová, I.* (2010) Phenotyping breast cancer cell lines EM-G3, HCC1937, MCF7 and MDA-MB-231 using 2-D electrophoresis and affinity chromatography for glutathione-binding proteins. *BMC Cancer*, **10**, 449, 1-10.
14. Mášová, A., Šanda, M., Jiráček, J., Selicharová, I.* (2010) Changes in the proteomes of the hemocytes and fat bodies of the flesh fly Sarcophaga bullata larvae after infection by Escherichia coli. *Proteome Sci.*, **8**, 1-11.
15. Jiráček, J., Žáková, L., Antolíková, E., Watson, C. J., Turkenburg, J. P., Dodson, G. G., Brzozowski, A. M.* (2010) Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 1966-1970.
16. Antolíková, E., Žáková, L., Turkenburg, J. P., Watson, C. J., Hančlová, I., Šanda, M., Cooper, A., Kraus, T., Brzozowski, A. M.*, Jiráček, J.* (2011) Non-equivalent Role of Inter- and Intramolecular Hydrogen Bonds in the Insulin Dimer Interface. *J. Biol. Chem.*, **286**, 36968-36977.
17. Morcavallo, A., Genua, M., Palummo, A., Kletvíková, E., Jiráček, J., Brzozowski, A. M., Iozzo, R. V., Belfiore, A.*, Morrione, A.* (2012) Insulin and insulin-like growth factor II differentially regulate endocytic sorting and stability of the insulin receptor isoform A. *J. Biol. Chem.*, **287**, 11422-11436.
18. Žáková, L., Kletvíková, E., Veverka, V., Lepšík, M., Watson, C.J., Turkenburg, J.P., Jiráček, J.*, Brzozowski, A.M.* (2013) Structural integrity of the B24 site in human insulin is important for hormone functionality. *J. Biol. Chem.*, **288**, 10230-10240.
19. Menting, J.G., Whittaker, J., Margetts, M.B., Whittaker, L.J., Kong, G.K.-W., Smith1, B.J., Watson, C.J., Žáková, L., Kletvíková, E., Jiráček, J., Chan, S.J., Steiner, D.F., Dodson, G.G.,Brzozowski, A.M., Weiss, M.A., Ward, C.W., Lawrence, M.C. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature*, **493**, 241-245.
20. Žáková, L., Kletvíková, E., Lepšík, M., Collinsová, M., Watson, C.J., Turkenburg, J.P., Jiráček, J., Brzozowski, A.M.* (2014) Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex. *Acta Crystallogr. D*, **70**, 2765- 2774.
21. Kosinová, K., Veverka, V., Novotná, P., Collinsová, M., Urbanová, M., Moody, N.R., Turkenburg, J.P., Jiráček, J., Brzozowski, A.M., Žáková, L.* (2014) Insight into the structural and biological relevance of the T/R transition of the N‑terminus of the B‑chain in human insulin. *Biochemistry*, **53**, 3392-3402.
22. Křížková, K., Veverka, V., Maletínská, L., Hexnerová, R., Brzozowski, A.M., Jiráček, J., Žáková, L.* (2014) Structural and functional study of the GlnB22-insulin mutant responsible for Maturity-Onset Diabetes of the Young. PloS One, **9** (11), 1-16, e112883.
23. Hrubá, L., Budešínsky, M., Pícha, J., Jiráček, J., Vaněk, V.* (2013) Simplified syntheses of the water-soluble chiral shift reagents Sm-(R)-pdta and Sm-(S)-pdta. *Tett. Letts.*, **54**, 6296-6297.
24. Mason, P.E., Uhlig, F., Vaněk, V., Buttersack, T., Bauerecker, S., Jungwirth, P.* (2015) Coulomb explosion during early stages of the reaction of alkali metals with water. *Nature Chem.*, **7**, 250-254.
25. Pirnik, Z., Maixnerová, J., Matyšková, R., Koutová, D., Železná, B., Maletínská, L., Kiss, A.* (2010) Effect of anorexinergic peptides, cholecystokinin (CCK) and cocaine and amphetamine regulated transcript (CART) peptide, on the activity of neurons in hypothalamic structures of C57Bl/6 mice involved in the food intake regulation. *Peptides*, **31**, 139-144.

26. Maixnerová, J., Špolcová, A., Pýchová, M., Blechová, M., Elbert, T., Řezáčová, M., Železná, B., Maletínská, L.* (2011) Characterization of prolactin-releasing peptide: Binding, signaling and hormone secretion in rodent pituitary cell lines endogenously expressing its receptor. *Peptides*, **32**, 811-817.
27. Maletínská, L.*, Špolcová, A., Maixnerová, J., Blechová, M., Železná, B. (2011) Biological properties of prolactin-releasing peptide analogs with a modified aromatic ring of a C-terminal phenylalanine amide. *Peptides*, **32**, 1887-1892.
28. Pirnik, Z., Bundziková, J., Holubová, M., Pýchová, M., Fehrentz, J. A., Martinez, J., Železná, B., Maletínská, L., Kiss, A.* (2011) Ghrelin agonists impact on Fos protein expression in brain areas related to food intake regulation in male C57BL/6 mice. *Neurochem. Int.*, **59**, 889-895.
29. Maletínská, L.*, Pýchová, M., Holubová, M., Blechová, M., Demianová, Z., Elbert, T., Železná, B. (2012) Characterization of new stable ghrelin analogs with prolonged orexigenic potency. *J. Pharmacol. Exp. Ther.*, **340**, 781-786.
30. Maletínská, L.*, Matyšková, R., Maixnerová, J., Sýkora, D., Pýchová, M., Špolcová, A., Blechová, M., Drápalová, J., Lacinová, Z., Haluzík, M., Železná, B. (2011) The Peptidic GHS-R antagonist [D-Lys3]GHRP-6 markedly improves adiposity and related metabolic abnormalities in a mouse model of postmenopausal obesity. *Mol. Cell. Endocrinol.*, **343**, 55-62.
31. Matyšková, R., Železná, B., Maixnerová, J., Koutová, D., Haluzík, M., Maletínská, L.* (2010) Estradiol Supplementation Helps Overcome Central Leptin Resistance of Ovariectomized Mice on a High Fat Diet. *Horm. Metab. Res.*, **42**, 182-186.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Jan Konvalinka - Proteases of Human Pathogens

The group of Jan Konvalinka aims to identify, characterize and exploit enzymes, especially proteases, as therapeutic targets. At present we work on one well established target (HIV protease), one recently identified one (glutamate carboxypeptidase II, also known as the prostate specific membrane antigen) and an emerging therapeutic target (human and mouse serine racemase). In all cases, we established appropriate recombinant expression systems, developed purification procedures and activity assays for these enzymes, approached crystallization and structure determination, identified small molecule inhibitors, prepared specific monoclonal and/or polyclonal antibodies, mapped the protein product expression in human and mammalian tissues by immunochemical methods, performed structure-activity study using site-directed mutagenesis and characterized polymorphisms of the naturally occurring forms of the enzymes *in vivo*. The current research team consists of four postdocs, six PhD students, ten undergraduate students and three staff scientists and technicians.

As documented by numerous joint grant projects and publications, we closely collaborate also with several other teams in the Institute. The most important include Pavel Majer's group (Medicinal Chemistry: joint project on HIV assembly inhibition), Pavel Hobza, Lubos Rulisek and Pavel Jungwirth's groups (computational theoretic chemistry), Pavlina Rezacova's group (structural biology: X-ray structure determination of enzyme-inhibitor complexes), and several others. The international collaborations involve Brain Research Institute, Johns Hopkins, Baltimore (prof. Barbara Slusher), Department of Virology, University of Heidelberg (prof. Hans-Georg Kraeusslich) and National Cancer Institute, Laboratory of Mocromolecular Crystallography (prof. Alex Wlodawer).

The group has expertise in recombinant protein expression in *E.coli*, insect and mammalian cells, protein purification, assay development, enzyme kinetics, inhibitor design and synthesis, and protein crystallisation and structure determination by X-ray crystallography. Recently we introduced fluorescent microscopy, protein NMR (in collaboration), microcalorimetry and real-time PCR techniques to characterize individual spliced forms of the studied proteins. In collaboration we also perform knock-out experiments and xenografts of human cancer in mice

In the reported period, the group worked mostly on two different, yet complementary models:

i/ the processing and maturation of HIV (viral protease, its inhibition and development of viral resistance, identification of novel PR inhibitors, design and testing of small molecule protein-protein interaction inhibitors to block viral assembly);

ii/ expression, structure and **function of glutamate carboxypeptidase II from human brain and prostate carcinoma** (expression, domain mapping, assay development, substrate specificity, X-ray structure and mutagenesis of a neuropeptidase and prostate specific membrane antigen, specific targeting to GCPII-expressing cells).

Analysis of the polyprotein processing and maturation of HIV

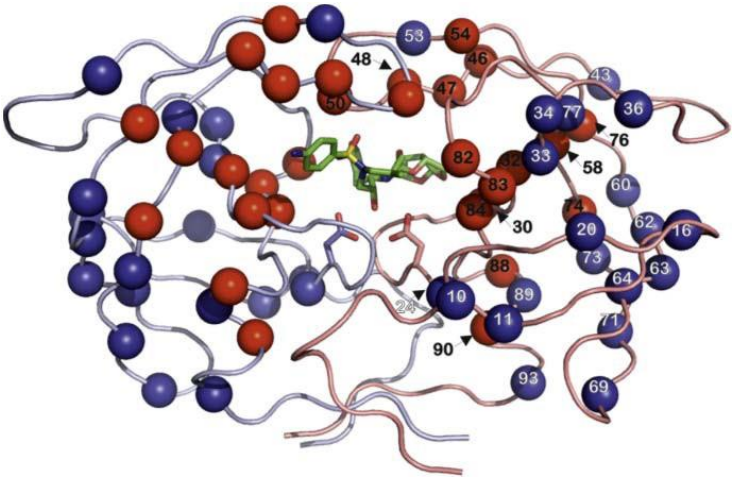


Fig. 1*: Sites of known primary and secondary resistance-associated mutations in HIV-1PR. The model depicts mutations associated with resistance to clinically used PI. The inhibitor (green) bound to the active site is represented as stick models. Mutated residues are indicated with their Cα atoms (spheres) and colored red and blue for major (primary) and minor (secondary) mutations, respectively (adopted from Konvalinka et al., 2015).*

The life cycle of HIV offers number of "Achilles's heels" of the pathogen that might serve as useful therapeutic targets. Our lab traditionally focuses on the viral maturation step, when endogenous viral protease cleaves the viral polyproteins into structural proteins and enzymes of the mature virus. More recently, we started a project to design and analyze low molecular ligands of the HIV capsid proteins that prevent HIV core assembly.

Detailed characterization of HIV strains selected under the pressure of protease inhibitors: mutations in the enzyme and in the viral substrate

The major problem that limits the therapeutic efficiency of protease inhibitors (PIs) is drug resistance caused by extensive mutations in PR. During the reported period, we concentrated on the detailed description of events leading to antiviral resistance on molecular level. We cloned, expressed, characterized and crystallized a series of HIV PR species isolated from patients on prolonged therapy by PI darunavir (DRV), and characterized the amino acid mutations in the PR region that lead to resistance development (Kožíšek et al., 2012, Kožíšek et al., 2014). We have characterized number of PR samples amplified from HIV-positive patients from San Diego under prolonged treatment by DRV. One PR species that we identified, contained as many as 21 mutations in the PR region, the highest number of mutations described so far in the literature, and exhibited pronounced DRV resistance. The substitutions critical for the development of DRV resistance were identified. We performed X-ray structure analysis of five DRV-resistant PR complexes and explained the resistance mechanism for those specific mutants on structural level. The significantly lower affinity and higher *Ki* value of the resistant variants toward DRV was caused by the loss of two hydrogen bonds between inhibitor and PR main-chain and one water-mediated contact in the inhibitor binding pocket due to a substantial shift of the aminophenyl moiety of the inhibitor (Kožíšek et al., 2014).

Apart from amino acid exchange, insertions have been described as common mechanism of antiviral resistance development, especially for the HIV reverse transcriptase inhibitors. However, very few reports on the selection of insertions in the HIV PR have been published. It is still unclear whether or not these insertions impact PI resistance and/or viral replication capacity. In collaboration with the chemists at Gilead Sciences, Foster Ciry, USA and

laboratory of Monique Nijhuis from Utrecht we have identified amino acid insertions at positions 33 and 35 of PR of an HIV-1 infected patients who had undergone prolonged treatment with PIs and characterized the contribution of these insertions to viral resistance. We prepared the corresponding mutated, recombinant PR variants with or without insertions at positions 33 and 35 and characterized them in terms of enzyme kinetics and crystal structures. We also analyzed the relative inhibitory activity of a series of clinically used and investigational inhibitors of HIV PR and established that a novel compound, inhibitor GS 8374 developed at Gilead Sciences does not lose its inhibitory potency upon the mutations. This finding has been explained by X-ray structure analysis of the corresponding complexes.

Besides primary and secondary mutations in the protease itself, amino acid changes in the cleavage sites of viral Gag polyprotein are also introduced, leading to improved ability of the mutated enzyme to bind and cleave its substrate. The viral resistance could be explained in part by mutations close to the NC/SP2 cleavage site which enhanced processing of the NC/p6 intermediate and also affected overall Gag processing. We characterized HIV protease from a child patient, identified by our collaborators from Wurzburg, Germany. The PR of this patient accumulated up to 22 mutations. Biochemical analysis *in vitro* showed that the patient-derived PR is highly resistant to most of the currently used PIs and exhibits also a very poor catalytic activity. Determination of the crystal structure of the PR revealed prominent changes in the flap elbow region and S1/S1' active site subsites. The introduction of patient derived Gag and Pol sequences into the proviral genom containing patient's virus protease rescued viral infectivity near to wild-type levels. This indicates a very efficient co-evolution of enzyme and substrate maintaining high viral loads in vivo under constant drug pressure. This case of mutual interplay between an enzyme and substrate under selection pressure represents unique model for molecular analysis of evidence and will be analyzed in the lab further (Kožíšek et al., 2012).

Induced maturation of HIV virions by triggering of polyprotein processing by light

Even after many years of research, the regulation, activation, and kinetics of HIV polyprotein processing and virion maturation are still not completely understood. One of the reasons is that the virologists deal with a pool of virus particles in various stages of their life cycle. The possibility to induce HIV polyprotein processing in a synchronized manner with an external stimulus would be an important step towards addressing these questions. Therefore, we designed, synthesized, and tested a potent, specific, photodegradable HIV PR inhibitor. The compound, an analogue of the tight-binding HIV PR inhibitor ritonavir (RTV), shows subnanomolar inhibitory activity *in vitro*. It contains a photolabile 7-diethylamino-4-(hydroxymethyl)coumarin moiety connected *via* a carbonate linker to an RTV fragment. This fragment is released from the molecule upon irradiation with 405-nm light, reducing the inhibitory potential by four orders of magnitude (Figure 2). We determined the structure of the PR-inhibitor complex, analyzed its photolytic products, and showed that the enzymatic activity of inhibited HIV PR could be fully restored upon inhibitor photolysis. The photocleavage products display a moderate or no inhibitory with K_i value of the larger fragment of 3.2 ± 0.3 μM (Fig. 2**)**. The photocleaved coumarine group does not show any inhibition activity at concentration of 10 μM.

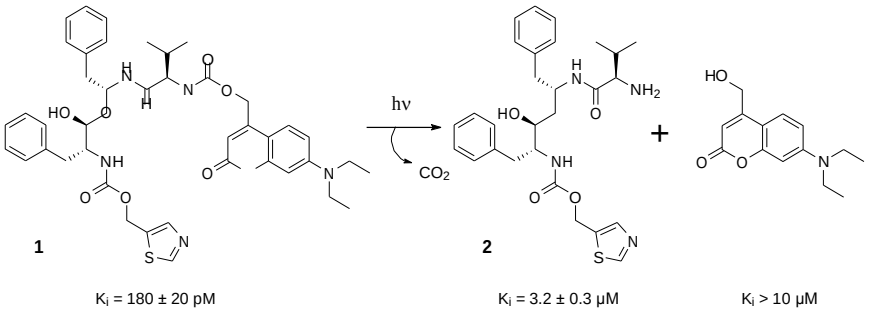


Figure 2. *The photodegradable inhibitor of HIV PR and the products of its photolysis. Inhibition values for HIV PR are shown for each compound (Schimer et al., 2015).*

Using this inhibitor and a novel irradiation setup that we designed we were able to almost fully restore enzymatic activity with both purified HIV PR *in vitro* and isolated HIV virions leading to complete induced polyprotein processing in immature viral particles. We thus demonstrated that polyprotein processing of isolated immature HIV preparations containing the inhibitor could be triggered by light (Schimer et al., 2015). An analogous approach may be used for photocaging of other biologically relevant proteases and for analysis of their function *in situ.*

In collaboration with prof. Hans-Georg Krausslich from the University of Heidelberg and John Briggs from EMBL we now characterize the ultrastructure of the mature HIV virions by cryoelectron microscopy (Schimer et al., manuscript in preparations).

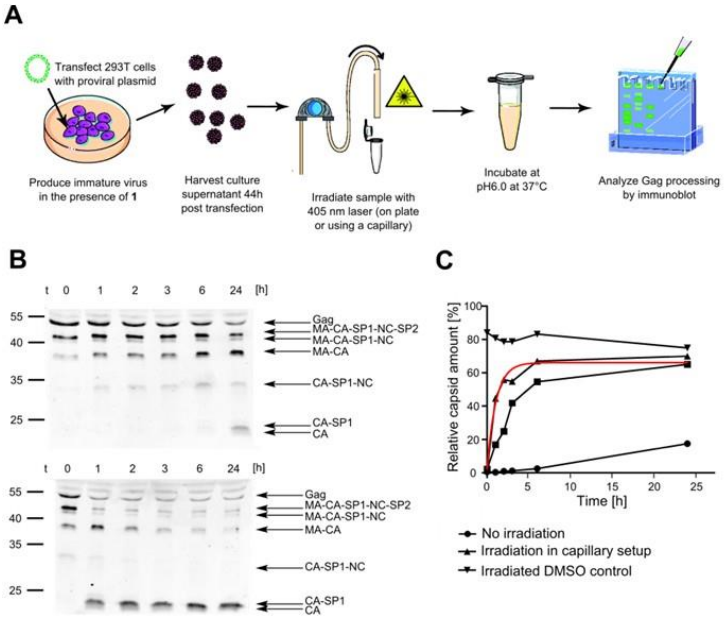


Figure 3. A) *Schematic illustration of the irradiation experiment to trigger HIV maturation using photoinactivation of photolabile inhibitor (PDI). HEK293T cells were transfected with a proviral plasmid and grown in the presence of 2 µM PDI. The tissue culture supernatant was then harvested and either left untreated or subjected to irradiation. Subsequently, samples were incubated for various lengths of time and analyzed by Western blot. B) The figure shows samples incubated for the indicated times without prior irradiation (top panel) or following irradiation (bottom panel) and analysed by Western blot. C) Quantitative analysis of the processing products. Anti-CA reactive bands from the immunoblots shown in B were quantified (adopted from Schimer et al., 2015).*

Structural and enzymological characterization of glutamate carboxypeptidase II and its homologs

Glutamate carboxypeptidase II (GCPII; EC 3.4.17.21) is a Zn^{2+}-dependent metalloprotease of the M28 peptidase family. It is also known as prostate specific membrane antigen (PSMA) or folate hydrolase I (FOLH1). In the brain, the enzyme cleaves the peptide neurotransmitter *N*-acetyl-L-aspartyl-L-glutamate (NAAG) into *N*-acetyl-L-aspartate and free glutamate, a potent and potentially neurotoxic neurotransmitter. The same enzyme also possesses a folate hydrolase activity, cleaving gamma-linked glutamates from folyl-poly-gamma-glutamates and facilitating thus the absorption of dietary folates in the small intestine. While the physiological function of GCPII in the prostate is not known, its expression levels are elevated in metastatic prostate carcinoma. Furthermore, the enzyme is also expressed in the cancer-associated neovasculature of most solid tumors.

In the nervous system, GCPII inhibition leads to an increase in NAAG concentration and has been shown to be neuroprotective in animal models of ischemic or traumatic brain injury, neuropathic pain or amyotrophic lateral sclerosis. In tumors, GCPII represents a very promising target for specific diagnostics, imaging and targeted delivery of anticancer drugs.

The contribution of our laboratory to the analysis of various aspects of GCPII biochemistry dates back to 2002 when we were able to develop the first efficient expression system for GCPII in insect (Schneider's S2) cells (Bařinka et al., 2002). We were able to purify the extracellular part of GCPII, determine its substrate specificity, kinetic parameters, the relationship between the activity and N-glycosylation and, importantly, determine its 3D structure by X-ray structure analysis (Mesters et al., 2006).

In the reviewed period, we built on these preliminary results. We introduced a novel, efficient expression-purification system for GCPII and its homologs based on *in vivo* biotinylation by Bir ligase (Tykvart et al., 2012). In the collaboration with the laboratory of Jacek Lubkowski at NCI Frederick, we studied the 3D-structure of the protein, mapped the structural requirements of its S1 and S1' subsites and analyzed its interactions with substrates and inhibitors by X-ray structure analysis. The results of this endeavor have been published in a series of joint publications with Barinka and Lubkowski labs and are schematically summarized in figure below.

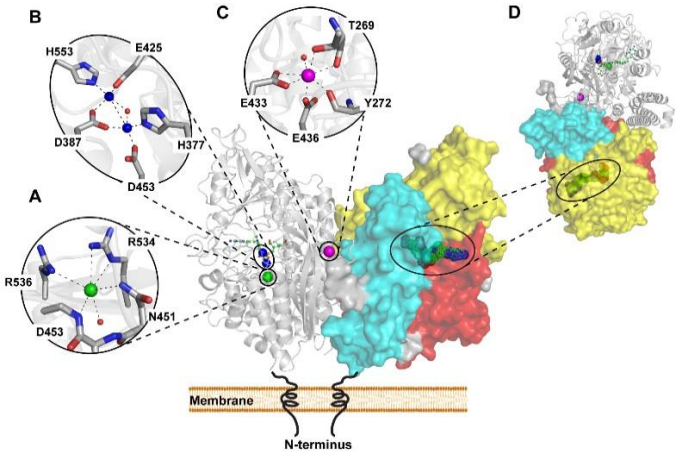


Fig. 4*: Overview of the important structural features of GCPII as revealed by X-ray structure analyses in the Konvalinka, Barinka and Lubkowski labs. One monomer of GCPII shown in semitransparent surface representation with individual domains of the extracellular part colored red (protease domain), yellow (apical domain), and cyan (C-terminal). N-linked sugar moieties are colored gray, and the entrance into the S1 funnel is circled. The second monomer is in cartoon representation (grey), with inorganic ions shown as blue, green, and*

*magenta spheres for Zn^{2+}, Cl^{-} and Ca^{2+}, respectively. Close-ups of the chloride anion surroundings (**A**); the active-site zincs (**B**); and the calcium loop (**C**). **Panel D**: GCPII homodimer (top view). The S1 funnel and the S1' pocket are delineated by the black line, with an inhibitor (PDB code 3BI1) shown in ball-and-stick representation (figure adopted from the review by Hlouchová et al., 2012).*

The availability of pure recombinant enzyme enabled us to prepare several very specific and sensitive monoclonal antibodies. We compared all available mABs against GCPII and its homologs, identified their epitopes and quantified their binding properties to the antigen (Tykvart et al., 2014). We were also able to clone, express and characterize homologous GCPIII, a protein with unknown function that might compensate for the lack of GCPII activity in knock-out mouse. We determined the 3D-structure of this homolog (Hlouchová et al., 2009). Using novel, highly sensitive method for the detection of GCPII in complex biological matrices, we quantified the amount of GCPII in human serum (Navratil, V., et al., 2014).

In addition to its well-characterized role in the central nervous system, GCPII acts as a folate hydrolase in the small intestine, participating in the absorption of dietary polyglutamylated folates ($FolGlu_n$), which are the provitamin form of folic acid (also known as vitamin B_9). Despite years of research, nothing has been known about the processing of polyglutamylated folates by GCPII at the structural or enzymological level. We prepared a series of X-ray structures of complexes between a catalytically inactive GCPII mutant (Glu424Ala) and a panel of naturally occurring polyglutamylated folates, performed inhibitor binding and enzyme kinetic analyses with polyglutamylated folates as substrates. As a result, the crystallographic data reveal considerable details about the binding mode of polyglutamylated folates to GCPII, especially the engagement of the so called arene binding site in recognizing the folic acid moiety. (Navrátil, M., et al., 2014).

Structural and biochemical characterization of a novel aminopeptidase from human intestine homologous to GCPII

N-acetylated alpha-linked acidic dipeptidase-like protein (NAALADase L), encoded by the *NAALADL1* gene, is a close homolog of glutamate carboxypeptidase II (GCPII), a metallopeptidase that has been intensively studied as a target for imaging and therapy of solid malignancies and neuropathologies. However, neither the physiological functions nor structural features of NAALADase L are known at present. We performed a thorough characterization of the protein product of the human *NAALADL1* gene, including heterologous overexpression and purification, structural and biochemical characterization, and analysis of its expression profile.

By solving the NAALADase L X-ray structure, we provided the first experimental evidence that it is a zinc-dependent metallopeptidase with a catalytic mechanism similar to that of GCPII, yet distinct substrate specificity. A proteome-based assay revealed that the *NAALADL1* gene product possesses previously unrecognized aminopeptidase activity, but no carboxy- or endopeptidase activity. These findings were corroborated by site-directed mutagenesis and identification of bestatin as a potent inhibitor of the enzyme. Analysis of *NAALADL1* gene expression at both the mRNA and protein levels revealed the small intestine as the major site of protein. Our data imply that the *NAALADL1* gene product's primary physiological function is associated with the final stages of protein/peptide digestion and absorption in the human digestive system. Based on these results, we suggest a new name for this enzyme: human ileal aminopeptidase (HILAP; Tykvart et al., 2015).

Rational design of urea-based glutamate carboxypeptidase II (GCPII) inhibitors as tools for specific drug targeting, delivery and imaging

Low-molecular-weight inhibitors of GCPII could be used as specific ligands for imaging and drug delivery to prostate cancer cells. However, they must be modified with a linker to enable connection of the ligand with an imaging molecule, anticancer drug, and/or nanocarrier. We performed a structure-activity relationship (SAR) study of GCPII inhibitors with linkers suitable for imaging and drug delivery. Structure-assisted inhibitor design and targeting of a specific GCPII exosite resulted in a 7-fold improvement in K_i value compared to the parent structure. X-ray structural analysis of the inhibitor series led to the identification of several inhibitor binding modes. We also optimized the length of the inhibitor linker for effective attachment to a biotin-binding molecule and showed that the optimized inhibitor could be used to target nanoparticles to cells expressing GCPII (Tykvart et al., *Bioorg. Med. Chem.* **2014**, *22*, 4099-4108).

Development of the DNA-Inhibitor Antibody Assay (XXX) for the quantitative determination of HIV PR, aspartic proteases and carbonic anhydrase IX

We developed new, sensitive and specific analytical tool for the determination of small quantities of GCPII protein molecules in complex biological samples. To extend this finding and evaluate its potential use for general application, we prepared specific DNA-labeled ligands for several more biologically relevant enzymes (proteases of pathogens or cancer markers), including a specific inhibitor of HIV (a ritonavir derivative attached to oligonucleotide probe), pepstatine derivative for the detection of aspartic proteases, specific inhibitor of carbonic anhydrase IX, an emerging cancer marker (Figure 6.3) and an inhibitor of neuraminidase from influenza.

All the compounds are specific and potent inhibitors of their corresponding targets and can be used for the quantification of the corresponding targets with sensitivity ranging from one hundred molecules of the target (GCPII) to hundred pikograms (CA IX) (Navrátil, V. et al., patent application and publication in preparations).

XXX

"IBodies": synthetic antibody analogs for detection, imaging and isolation of proteins and specific drug delivery

We developed synthetic polymers based on *N*-(2-hydroxypropyl)methacrylamide(HPMA) copolymer decorated by a specific, low-molecular weight inhibitors of glutamate carboxypeptidase II (GCPII), a fluorescent label for visualisation, biotin residue as an affinity tag and/or anticancer drug (doxorubicin, DOX) that is bound to the copolymer via degradable spacer. These polymers recognize and visualize GCPII on the surface of GCPII-positive cells and could be used for immunohistochemistry, in fluorescent-aided cell sorting (FACS) analyses; in a sandwich ELISA and due to their selectivity and specificity, also for immunoprecipitation via streptavidin-coated beads. Finally, they internalize into GCPII-

expressing cells upon binding to the target and could be thus used for targeted drug delivery to these cells.

These polymer conjugates are effective, specific and sensitive replacement of monoclonal antibodies raised against GCPII showing sensitivity and selectivity comparable to the best antibodies described so far. They are cheap, easily available in bulk quantities, very stable, non-toxic and biocompatible.

In order to analyze whether this concept, developed for GCPII, could be used for other medicinally relevant targets, we prepared a series of HPMA copolymers conjugated with low molecular inhibitors of HIV protease (ritonavir), general aspartic protease inhibitor (pepstatine), and specific inhibitors of neuraminidase and carbonic anhydrase IX. These polymer conjugates show analogous behavior as those decorated by GCPII-inhibitor: tightly bind to their target enzymes and inhibit their activity (in some cases showing synergistic effect of multiple binding), could be used for effective "immunoprecipitation" of HIV PR, cathepsin D and CA IX and for specific imaging of neuraminidase and CA IX-expressing cells. We work on the application of these conjugates for the drug delivery and imaging of target protein-expressing cells in animal models.

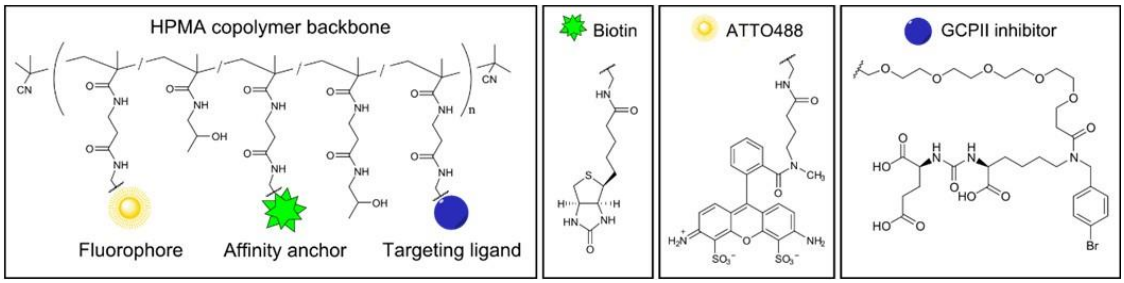


Fig. 6*: Schematic representation of "iBodies" conjugates and examples of targeting ligands.*

Selected papers

Grantz Sasková, K., Kozísek, M., Stray, K., de Jong, D., Rezácová, P., Brynda, J., van Maarseveen, N.M., Nijhuis, M., Cihlár, T., Konvalinka J. (2014): GS-8374, a Prototype Phosphonate-Containing Inhibitor of HIV-1 Protease, Effectively Inhibits Protease Mutants with Amino Acid Insertions. *J. Virol*. 88, 3586-3590

Fun, A., van Maarseveen, N.M., Pokorná, J., Maas, R.E.M., Schipper, P.J., Konvalinka, J. and Nijhuis, M. (2011) HIV-1 protease inhibitor mutations affect the development of HIV-1 resistance to the maturation inhibitor bevirimat. *Retrovirology* 8, (70)

Hlouchová, K., Navrátil, V., Tykvart, J., Šácha, P. & Konvalinka, J. (2012) GCPII variants, paralogs and orthologs. *Curr. Med. Chem*. 19, 1316-1322

Kožíšek, M., Grantz Šašková, K., Jacobs, G.B., Schuch, A., Buchholz, B., Müller, V., Kräusslich, H.-G., Řezáčová, P., Konvalinka, J. and Bodem, J. (2012) Mutations in HIV-1 gag and pol Compensate for the Loss of Viral Fitness Caused by a Highly Mutated Protease. *Antimicrob. Agents Chemother*. 56, 4320-4330

Kožíšek, M., Lepšík, M., Grantz Šašková, K., Brynda, J., Konvalinka, J. and Řezáčová, P. (2014). Thermodynamic and structural analysis of HIV protease resistance to darunavir – analysis of heavily mutated patient-derived HIV-1 proteases. *FEBS J*. 281, 1834–1847.

Knedlík T., Navrátil, V., Vik, V., Pacík, D., Sácha, P., Konvalinka, J. (2014). Detection and quantitation of glutamate carboxypeptidase II in human blood. *Prostate* 74, 768-780.

Mattei, S., Anders, M., Konvalinka, J., Kraeusslich, H.-G., Briggs, J.A.G and Mueller, B. (2014) Induced maturation of human immunodeficiency virus. *J. Virol*. 88, 13722-13731.

Navrátil, M., Ptáček, J., Šácha, P., Starková, J., Bařinka, C., Lubkowski, J., Konvalinka, J. (2014). Structural and Biochemical Characterization of Folyl-poly-γ-L-glutamate

Hydrolysing Activity of the Human Glutamate Carboxypeptidase II. *FEBS J.* 281, 3228-3242.

Plechanovová, A., Byun, Y., Alquicer, G., Škultétyová,L., Mlčochová, P., Němcová, A., Kim, H.-Y., Navrátil, M., Mease, R., Lubkowski,J., Pomper, M., Konvalinka, J., Rulíšek, L. and Bařinka, C. (2011). Novel substrate-based inhibitors of glutamate carboxypeptidase II with enhanced lipophilicity J. Med. Chem. 54, 7535-46

Sedlák, F., Šácha, P., Blechová, M., Březinová, A., Šafařík, M., Šebestík, J. and Konvalinka, J. (2013) Glutamate carboxypeptidase II does not process amyloid-β peptide. *FASEB J.* 27, 2626-2632

Schimer, J., Cígler, P., Veselý, J., Grantz Šasková, K., Lepšík, M., Brynda, J., Řezáčová, P., Kožíšek, M., Císařová, I., Oberwinkler, H., Kraeusslich, H.-G & Konvalinka, J. (2012) Structure-aided design of novel inhibitors of HIV protease based on a benzodiazepine scaffold. *J. Med. Chem*., 55, 10130−10135

Schimer, J., Konvalinka, J. (2014) Unorthodox Inhibitors of HIV Protease: Looking Beyond Active-Site-Directed Peptidomimetics. *Curr Pharm Des.* 20, 3389-3397

Schimer, J., Pávová, M., Anders, M., Pachl, P., Šácha, P., Cígler, P., Weber, J., Majer, P., Řezáčová, P., Kräusslich, H.-G., Müller, B. and Konvalinka, J. (2015) Triggering HIV polyprotein processing by light using rapid photodegradation of a tight-binding protease inhibitor. *Nature Commun*. 6, 6461, doi:10.1038/ncomms7461

Tykvart, J.; Šácha, P.; Bařinka, C.; Knedlík, T.; Starková, J.; Lubkowski, J. & Konvalinka, J. (2012) Efficient and versatile one-step affinity purification of *in vivo* biotinylated proteins: Expression, characterization and structure analysis of recombinant human Glutamate Carboxypeptidase II. *Protein Exp. Purif.* 82, 106-115.

Tykvart, J., Schimer, J., Bařinková, J., Pachl, P., Poštová-Slavětínská, L., Majer, P., Konvalinka, J. and Šácha, P. (2014) Rational design of urea-based glutamate carboxypeptidase II (GCPII) inhibitors as versatile tools for specific drug targeting and delivery. *Bioorg. Med. Chem*. 22, 4099-4108

Tykvart, J., Navrátil, V., Sedlák, F., Corey, E., Colombatti, M., Fracasso, G., Koukolík, F., Bařinka, C., Šácha, P., Konvalinka, J. (2014) Comparative analysis of monoclonal antibodies against prostate-specific membrane antigen (PSMA). *The Prostate* 74, 1674-1690.

Tykvart,J., Bařinka, C., Svoboda, M., Navrátil, V., Souček, R., Hradílek, M., Šácha, P., Lubkowski, J. and Konvalinka, J. (2015) Structural and biochemical characterization of human N-acetylated alpha-linked acidic dipeptidase-like (NAALADase L) protein, a novel aminopeptidase from human intestine. *J. Biol. Chem., doi:10.1074/jbc.M114.62814*

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Pavlína Řezáčová - Structural Biology

2.1 Introduction

During the evaluated period, the team of P. Maloy Řezáčová existed as a research-service with the research to service ratio about 70:30 percent. In June 2013, following the evaluation, the team was transformed into research team.

The main interests of the team are structural studies of various proteins of biological or medicinal interest. We use the knowledge of three-dimensional structure in understanding the protein function and in some projects also in modulating its function by design of specific inhibitors.

In addition to own research projects, the team also participates by its structural biology expertise on other projects. Relatively high involvement in collaborative projects is due to the fact that our expertise is quite unique within our country and we are often approached by potential collaborators with interesting structural biology projects. The team is open to collaboration, because running several own and collaborative projects in parallel increases the chance to obtain positive results - especially in protein crystallography, where crystallization is very often a bottleneck (ifnot an obstacle) to successful conclusion of the project. Also a balance between high-risk projects and safe projects ensures sustainable scientific output of the team, which is essential (not only) for successful grant applications.

2.2 Structure-based design of inhibitors of human carbonic anhydrase IX

In this project we focused on design of novel and original inhibitors targeting therapeutically relevant isoenzyme of human carbonic anhydrase (CA). Project was supported by grant from technology Agency (since 2012).

Human carbonic anhydrases play important roles in various physiological and pathological processes (e.g. tumorigenicity, obesity, and epilepsy) and thus many CA isoenzymes represent established diagnostic and therapeutic targets. In particular, there is a significant interest in the development of selective inhibitors targeting carbonic anhydrase isoform IX (CAIX). CAIX is a tumor associated transmembrane isoenzyme, the overexpression of which is induced by hypoxia. CA IX is used as a tumor marker and as a prognostic factor for several human cancers and thus represents a valuable target for antitumor therapy. Design of a novel generation of selective inhibitors is the current challenge in the development of new therapeutic agents able to inhibit specific isoenzymes. The conventional carbonic anhydrase inhibitors contain a sulfonamide, sulfamate or sulfamide group connected to organic moiety usually composed of a 5- or 6-membered aromatic ring or conjugated ring system. Novelty of our approach lies in utilization of substituted boron clusters compounds as active-site inhibitors of CA isoenzymes.

The idea to use boron clusters and the first compound we have design originated from the analysis of X-ray structures of CAII in complex with organic sulfonamide-based inhibitors we have determined previously [1]. Our analyses revealed that structurally related inhibitors can bind by two distinct binding modes and engage two opposite sides of the active site. Following this analysis, we hypothesized that the

binding space within the enzyme active site cavity can be effectively filled with inhibitor with three-dimensional scaffold and hydrophobic nature and we have designed molecule containing sulfamide group connected to carborane cluster optimally filling active site cavity of CAII. Our compounds inhibited the enzyme activity with Ki value in submicromolar range and showed increased selectivity toward the tumor-associated isoform CAIX over widespread CAII. Crystal structure of this compound in complex with CAII confirmed the binding into the enzyme active site in the predicted pose and revealed key interactions responsible for binding of the inhibitors and enzyme inhibition and provided information that can be applied to the structure-based design of specific inhibitors [2, 3].

2.3 Structural studies of bacterial transcription factors

This project focused on structural studies of selected catabolic repressors from *Bacillus subtilis*, was supported by grant for bilateral collaboration between the Team and Prof. Otwinowski from UT Southwestern Medical Center at Dallas (until 2012).

In many bacteria, the expression of the genes involved in the metabolism of secondary carbon sources is typically repressed in the presence of the most favorable nutrient. This mechanism of regulation, designated carbon catabolite repression, requires participation of numerous specific repressors and activators controlling the level of transcription of catabolic genes. Transcriptional repressors involved in catabolic repression mainly belong to one of the two large families of prokaryotic transcriptional regulators, GntR and DeoR, respectively. Members of these families share an N-terminal helix-turn-helix DNA-binding domain andstructurally dissimilar C-terminal domains that are involved in effector binding and oligomerization.

To understand the modulation of transcription repressors and mechanisms of carbon catabolite repression in *B. subtilis* on a molecular level, we selected several functionally characterized repressors as targets for structural characterization. We determined crystal structures of three of them alone or in complexes with small molecular ligands. We combined X-ray crystallographic studies with complementary techniques, such as chemical cross-linking, isothermal titration calorimetry, differential scanning fluorimetry and electromobility shift assay to characterize effector binding and modulation of the oligomeric states of these proteins.

Crystal structure of AraR (L-arabinose operon repressor) identified dimer organization which is distinctive from other members of the family [4]. Crystal structure of DeoR (deoxyribonucleoside operon repressor) was obtained through sophisticated optimization of crystallization conditions [5] and showed covalent binding of effector molecule to the protein [6] .

These studies provided the first structural information on transcription factors involved in catabolite repression in B. subtilis, setting the stage for understanding this phenomenon in this organism and other gram-positive bacteria as well as the mechanisms of gene regulation in general.

2.4. Structural studies of human transcription factors, binding partners of LEDGF/p75

Project focused on structural studies of human transcription factors JPO2 and PogZ, binding partners of LEDGF/p75 was supported by grant from the FP7 of the European Union coordinated by Zeger Debyser from KU Leuven, Belgium (until 2012).

Lens epithelium-derived growth factor p75 (LEDGF/p75) is a prominent cellular cofactor for human immunodeficiency virus (HIV) integration and is crucial for HIV replication. HIV integrase interacts with the C-terminal part of LEDGF/p75, region designated integrase-binding domain (IBD, amino acids residues 347 - 429). While

the role of LEDGF/p75 in HIV integration is well characterized, very little is known about its physiological function. As a transcriptional co-activator, LEDGF/p75 is implicated not only in HIV replication, but also in human cancers and autoimmunity disorders. The LEDGF/p75 was shown to interact through its IBD with several cellular proteins and recent evidence implies that LEDGF/p75 is a general adaptor protein tethering various factors to chromatin.
In this project, we characterized two LEDGF/p75 physiological binding partners JPO2, pogo transposable element (pogZ) and mixed lineage leukemia (MLL) protein. The aim of our study was to obtain structural information on the LEDGF/p75 interaction with its physiological binding partners. Such structural information is essential for understanding the LEDGF/p75 biological role and might help in design of inhibitors selectively blocking interaction with HIV integrase while not interfering with the LEDGF/p75 biological function. NMR methods recently allowed for an identification of the JPO2 and PogZ interacting region of IBD. Interestingly, both proteins share similar interaction mechanism with HIV-1 integrase (publication submitted). The structure determination of the complexes is now pursued using NMR spectroscopy. Preliminary results suggested that only a relatively small region of JPO2 undergoes conformational re-arrangement upon binding to IBD. For MLL we identified novel LEDGF/p75–MLL interface which represents a newtarget for the development of therapeutics against MLL fusion–driven leukemic disorders [7].

2.5. Fragment based drug discovery

The aim of the project is a knowledge-based design of active compounds targeting several biomedicinally important enzymes using a fragment-based approach together with biomolecular NMR spectroscopy and modern methods of structural biology, biochemistry, medicinal and computational chemistry. The project has started in April 2012 and is supported by a 5-year programme NAVRAT (Ministry of Education). The biomolecular NMR techniques are utilized within the projects for structural characterization of interactions between targeted proteins and small molecule inhibitors, which are being developed either through a ‘traditional’ medicinal chemistry elaborations of known pharmacophores or from small fragments identified in the NMR-based screening campaigns planned for the next 31/2 years of the project.
We have successfully screened our in-house library of small fragments using NMR spectroscopy based techniques against three targets: UEV domain from Tsg101, which is a subunit of ESCRT-I complex responsible for the budding of HIV-1 viral particles; IBD domain of LEDGF/p75 that is implicated in the interaction with MLL in MLL-dependent leukemia; and kinase domain of Aurora-A, whose aberrant expression plays an important role in a broad range of malignancies. The screen yielded a substantial number of promising hits for all three proteins that will be after careful validation further developed using the structural information obtained by either X-ray crystallography or NMR spectroscopy.

2.5 References

(Authors and/or co-authors of the evaluated team are in bold; corresponding author from the evaluated team is marked by asterisk)

1. Mader, P**., Brynda, J**., Gitto, R., Agnello, S., **Pachl, P**., Supuran, C. T., Chimirri, A. & **Rezacova, P.*** (2011). Structural Basis for the Interaction Between Carbonic Anhydrase and 1,2,3,4-tetrahydroisoquinolin-2-ylsulfonamides. *Journal of Medicinal Chemistry* **54**, 2522-2526.

2. **Brynda, J.,** Mader, P., Sicha, V., Fabry, M., **Poncova, K**., Bakardiev, M., Gruner, B., Cigler, P. & **Rezacova, P.*** (2013). Carborane-Based Carbonic Anhydrase Inhibitors. *Angewandte Chemie-International Edition* **52**, 13760-13763.
3. Pecina, A., Lepsik, M., Rezac, J., **Brynda, J.,** Mader, P**., Rezacova**, P., Hobza, P. & Fanfrlik, J. (2013). QM/MM Calculations Reveal the Different Nature of the Interaction of Two Carborane-Based Sulfamide Inhibitors of Human Carbonic Anhydrase II. *Journal of Physical Chemistry B* **117**, 16096- 16104.
4. **Prochazkova, K., Cermakova, K., Pachl, P., Sieglova,** I., Fabry, M., Otwinowski, Z. & **Rezacova, P.*** (2012). Structure of the effector-binding domain of the arabinose repressor AraR from Bacillus subtilis. *Acta Crystallogr D Biol Crystallogr* **68**, 176-85.
5. **Pisackova, J. Procházkova, K**., Fabry M., **Rezacova P**.* (2013) Crystallization of the Effector-Binding Domain of Repressor DeoR from Bacillus subtilis. *Cryst. Growth Des.* **13**, 844–848.**Skerlová J,** Fábry M, Hubálek M, Otwinowski Z, **Rezáčová P.*** (2014) Structure of the effector-binding domain of deoxyribonucleoside regulator DeoR from Bacillus subtilis. *FEBS J.* **281**:4280-92.
6. Cermáková, K; **Tesina, P**; Demeulemeester, J; El Ashkar, S; Méreau, H; Schwaller, J; **Rezáčová, P; Veverka, V***; De Rijck, J. (2014). Validation and Structural Characterization of the LEDGF/p75-MLL Interface as a New Target for the Treatment of MLL-Dependent Leukemia. *Cancer Res.* **74**, 5139-5151.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Michael Mareš - Cathepsin Proteases in Pathology

Introduction to the project Cathepsin Proteases in Pathology

The research of the team is focused on cathepsin proteases and proteolytic systems controlled by these enzymes. Emphasis is placed on cathepsins involved in pathological processes. Cathepsin research is one of the most progressive areas in proteolysis due to the critical involvement of cathepsins in human pathologies, including cancer, arthritis, osteoporosis, and neurodegenerative diseases. Furthermore, cathepsin proteases of parasites have emerged as therapeutic targets against several important parasitic diseases.

The research program covers the identification of cathepsin targets, their functional and structural characterization, and the development of cathepsin inhibitors as a tool for novel intervention strategies in biomedicine. The research integrates protein biochemistry, functional proteomics, enzymology, protein production, 3D structure analysis and rational ligand design. Our program extends itsreach by systematically building on collaboration with teams focused on biology and compound synthesis to broaden the impact of potential applications in biomedicine. For this interdisciplinary research, we were able to secure an extensive grant support, including collaborative and international grants.

Research areas and major results

This chapter summarizes main research activities in four branches of our program during the evaluation period.

(1) Cathepsins of parasitic helminths as targets for inhibitor drugs

Schistosomiasis (bilharzia) is a chronic infectious disease caused by trematode blood flukes of the genus *Schistosoma*. It is a global health problem as these helminth parasites infect over 200 million people. *Schistosoma mansoni* is the most widespread schistosoma species. Treatment and control of schistosomiasis now relies on just one drug, and because of the growing concern over resistance, there is pressure to identify new anti-schistosomal chemotherapeutics.

Adult schistosomes live in the cardiovascular system, and host blood proteins are a primary source of nutrients. In the schistosome gut, a network of cathepsin proteases performs the digestion of hemoglobin and other blood proteins. *S. mansoni* cathepsin B1 (SmCB1) is a critical enzyme of this network and has been validated in a murine model of schistosomiasis as a molecular target for therapy. Inhibition of SmCB1 represents an attractive option for anti-schistosomal drug development;

however, rational design of inhibitors has been hampered by a lack of structural information for the enzyme.

Our schistosoma program is focused primarily on the regulation of SmCB1 activity and also on survey of other potential drug targets among *S. mansoni* proteases. The project is carried out in collaboration with the Center for Discovery and Innovation in Parasitic Diseases, University of California San Francisco (USA) involved in the biological part and *in vivo* screening.

(A) Inhibition of SmCB1: We determined the first crystallographic structure of SmCB1 in complex with three covalent peptidomimetic inhibitors [Jilkova *et al*., J. Biol. Chem. 2011]. Based on the binding mode (Fig. 1), interactions of the inhibitors with the subsites of the active-site cleft were evaluated by quantum chemical calculations. These data and inhibition profiling with a panel of peptidomimetic derivatives identified key binding interactions and provided insight into the specificity of SmCB1 inhibition. Furthermore, functional profiling of SmCB1 using synthetic peptides and the natural substrate hemoglobin revealed a dual exo- and endopeptidase activity of SmCB1, which was explained by the conformational flexibility of the active site.

Critically, we directly demonstrated that inhibition of SmCB1 by various covalent peptidomimetic inhibitors correlates with their anti-schistosomal potency tested on parasites in culture [Jilkova *et al*., J. Biol. Chem. 2011]. Also, we developed a novel scoring function for the prediction of the inhibitor effectiveness using quantum chemical calculations as an efficient tool for the *in silico* design of SmCB1 inhibitors [Fanfrlik *et al*., J. Phys. Chem. B 2013]. The above results provided a starting point from which the rational design of specific SmCB1 inhibitors is continued; our currentwork is focused on peptidomimetic vinyl sulfones and nitriles as the most promising scaffolds.

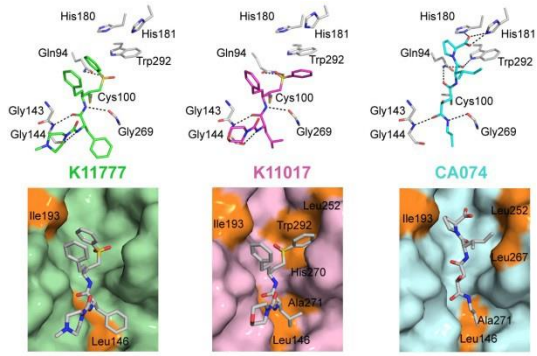


Fig. 1: Interactions between the SmCB1 active-site residues and reactive peptidomimetic inhibitors - the epoxide CA074 (1.3 Å resolution, PDB code 3QSD) and the vinyl sulfones K11017 (1.8 Å, 3S3Q) and K11777 (2.5 Å, 3S3R) [Jilkova *et al*., J. Biol. Chem. 2011].

Further research was on the biosynthesis of the active mature form of SmCB1. It is generated from the inactive precursor (zymogen) in which the N-terminal propeptide blocks the active site. Using biochemical and structural analyses, we described a unique processing mechanism for the zymogen activation by which the propeptide is proteolytically removed and its regulation by glycosaminoglycans (GAGs) [Jilkova *et al*., Structure 2014]. For this purpose, crystal structures of three molecular forms of SmCB1 along the activation pathway were determined (Fig. 2). We demonstrated that GAGs are essential for the autocatalytic processing of SmCB1, as they interact with a specific heparin-binding domain in the propeptide. An alternative processing route is mediated by an *S. mansoni* asparaginyl endopeptidase which is downregulated by GAGs, indicating that GAGs act as a molecular switch between both activation mechanisms.

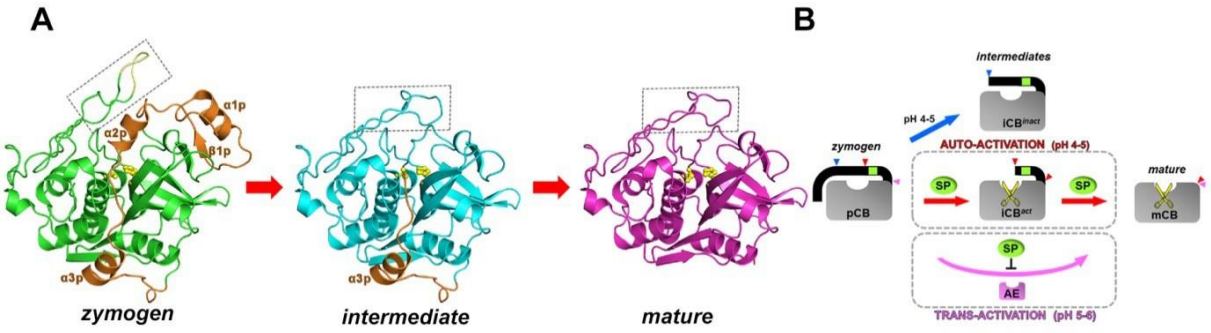


Fig. 2: (A) Spatial structure of three processing forms of SmCB1 (1.3-1.9 Å resolution, PDB codes 4I04, 4I05, 4I07): during the activation process, the propeptide (*orange*) is proteolytically removed from the inactive zymogen of SmCB1 to open the active site (*yellow*) of the active mature enzyme. (B) A scheme of the activation pathway of SmCB1 with two alternative routes [Jilkova *et al.*, Structure 2014].

In the next step, we employed the 3D structure of the SmCB1 zymogen and synthetic propeptide-derived fragments to analyze structure-inhibition relationships of how the SmCB1 propeptide interacts with the enzyme core [Horn *et al.*, ACS Chem. Biol. 2011]. The critical region was identified within the propeptide that governs the inhibition of the active site. Using this region as a template, we designed a new inhibitor scaffold mimicking the propeptide structure (Fig. 3). The developed inhibitors were selective for SmCB1 over human cathepsins and represent promising lead compounds for a new class of inhibitor drugs.

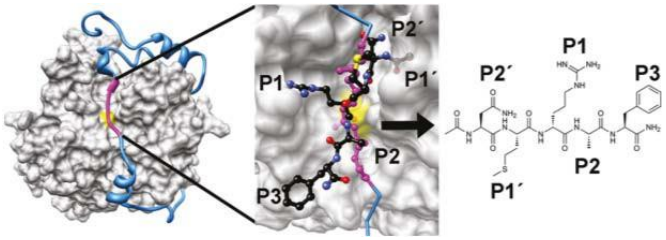


Fig. 3: The development of inhibitors of SmCB1 mimicking the structure of the SmCB1 propeptide [Horn et al., ACS Chem. Biol. 2011]. The 3D structure of the SmCB1 zymogen with the propeptide (*blue ribbon*) is on the left. The inhibitory region of the propeptide binding to the active site (*magenta*) was used as a template to design highly selective inhibitors of SmCB1.

(B) Other schistosomal proteases: This part of our schistosoma program is focused on the discovery of new potential drug targets among *S. mansoni* proteases, their functional characterization and inhibition regulation. First, the initial screen of digestive proteases by RNAi in schistosomes [Stefanic *et al.*, PLoS Negl. Trop. Dis. 2010] followed by chemical genomics analysis with selective protease inhibitors identified new potential targets, namely *S. mansoni* cathepsins C and D; projects on detailed analysis of these enzymes were initiated. Second, we identified a new group of *S. mansoni* serine proteases from the family S1 with trypsin- and chymotrypsin-like activities and described them using genomic, transcriptomic, phylogenetic and functional proteomic approaches [Horn *et al.*, PLoS Negl. Trop. Dis. 2014]. Third, we investigated for the first time *S. mansoni* prolyl oligopeptidase (named SmPOP) from the serine peptidase family S9 [Fajtova *et al.*, PLoS Negl. Trop. Dis., *in press*]. We

showed that this enzyme is localized in the tegument (surface) of schistosomes where it cleaves host vasoregulatory hormones, and thus may importantly contribute to the survival of the parasite. SmPOP substrate specificity was determined, and based on these data, we designed potent nanomolar inhibitors of SmPOP that were effective against cultured schistosomes. The study provided evidence that SmPOP is important for parasite survival and is a potential target for the development of anti-schistosomal therapeutics.

(2) Proteolytic systems of parasitic ticks as vaccination targets

Ticks are important ectoparasites that transmit a wide range of infectious agents including viruses, bacteria and protozoa causing diseases in humans and domestic animals. The ticks of the genus *Ixodes* are important vectors for Lyme disease (borreliosis) and tick-borne encephalitis in Europe and North America. In addition to the threat to public health, a number of tick species negatively impact on livestock productivity. The tick gut and saliva represent the critical interface for vector-pathogen-host interactions and the transmission of the pathogens. Therefore, these compartments are increasingly recognized as a source of potential antigens for the development of “anti-tick” vaccines that protect against the ticks and the tick-borne diseases.

This project is focused on two important groups of proteins that are involved in proteolytic interactions of ticks with the host, namely protease inhibitors from the tick saliva and proteases from the tick gut. In a collaborative arrangement, our team is responsible for structural and biochemical studies, while the biological component is conducted at the Institute of Parasitology ASCR (Czech Rep.).

Salivary protease inhibitors: The saliva of ticks is a cocktail of potent pharmacologically active components able to disarm the host haemostatic system and alter the inflammatory and immune responses. Protease inhibitors are important protein constituents of the tick saliva controlling the physiology at the feeding site, however, little is known about the inhibitors of host cathepsins. XXX

The salivary cystatin OmC2 from the tick *Ornithodoros moubata* was found to be an immunomodulatory protein that suppresses the host’s adaptive immune response and is important for survival of ticks in infestation experiments [Salat *et al*., Biochem. J. 2010]. Furthermore, we demonstrated that OmC2 is an effective broad-specificity inhibitor of host cysteine cathepsins that are physiologically relevant to the function of several types of immune cells. The results suggested that OmC2 mimics specific host-derived cystatins to interfere with their *in vivo* functions in controlling cathepsin-mediated proteolysis. The crystal structure of OmC2 was determined and used to explain its inhibitory specificity (Fig. 4).

The salivary serpin IRS-2 from the tick *Ixodes ricinus* functions as a unique inhibitor of host acute inflammation and platelet aggregation; these are key processes in defense to a skin injury. We demonstrated that IRS-2 primarily inhibits cathepsin G and chymase, and also affects thrombin activity. The mechanism of IRS-2 action is predicted to involve the modulation of neutrophils, mast cells and platelets at the early steps of their activation. The inhibitory specificity of IRS-2 was explained using the crystal structure of the protease-cleaved form of IRS-2 (Fig. 4); this is the first report on structure of a serpin from a parasitic organism [Chmelar *et al.*, Blood 2011; Kovarova *et al.*, Acta Crystallogr. Sect. F 2010].

XXX

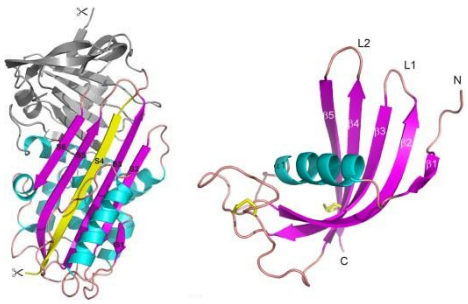


Fig. 4: Crystal structures of two protease inhibitors from the tick saliva investigated in this project. *Left*: serpin IRS-2 solved at a resolution of 1.8 Å (PDB code 3NDA) [Chmelar *et al.*, Blood 2011]. *Right*: cystatin OmC2 solved at a resolution of 2.45 Å (3L0R) [Salat *et al*., Biochem. J. 2010].

(A) Digestive proteases: Host blood is a rich source of proteins, and hemoglobin is a primary nutritive component for the ticks. It is digested intracellularly in the acidic lysosomal vesicles of gut cells; however, this critical process has been poorly understood at the molecular level. To fill this gap, we performed the first systematic mapping of the digestive hemoglobinolytic pathway in ticks using *I. ricinus* as a model species. Functional proteomic approaches (including imaging with activity-based probes and activity profiling) were applied to deconvolute the hemoglobinolytic cascade and a suite of gut proteases that operate in an ordered pathway to complete the hydrolysis of hemoglobin (Fig. 5). The work included cDNA sequencing of the component proteases and their dynamic profiling upon blood feeding. The results ofthis comprehensive research were summarized in [Sojka *et al*., Trends Parasitol. 2013, Franta *et al*., Parasites & Vectors 2010] and provided a catalog of novel vaccination antigens to target the tick gut.

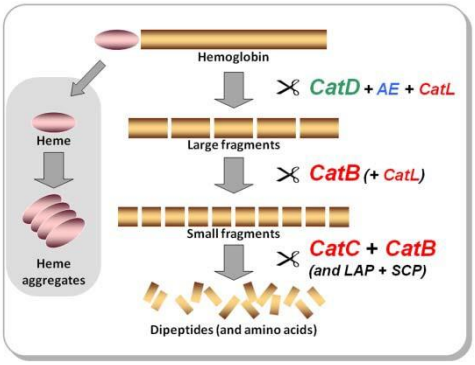


Fig. 5: The hemoglobinolytic pathway in the tick *I. ricinus* [Sojka *et al*., Trends Parasitol. 2013]. The endopeptidases, cathepsin D (CatD) supported by cathepsin L (CatL) and legumain (AE), are responsible for primary cleavage of hemoglobin. The production of secondary small fragments is dominated by the endopeptidase activity of cathepsin B (CatB). Exopeptidase action is through the carboxydipeptidase activity of CatB and the aminodipeptidase activity of cathepsin C (CatC).

In the next step, we focused on individual digestive cathepsins from *I. ricinus* to understand in detail their physiological role, biochemical properties and structure-activity relationships. The first papers in this series were devoted to two cathepsins that initiate the hemoglobinolytic pathway, namely the isoenzyme IrCL1 (a cathepsin L-type cysteine protease) [Franta *et al*., Int. J. Parasitol. 2011] and the isoenzyme IrCD1 (a cathepsin D-type aspartic protease) [Sojka *et al*., J. Biol. Chem. 2012]. In addition, we completed a large structural study on IrCD1 with four crystal structures that allowed us to describe the conversion of the inactive zymogen to the mature active IrCD1 [Hanova *et al*., manuscript in preparation]. Also, we discovered here a new mechanism for inhibition of aspartic proteases through binding a synthetic ligand to an exosite distinct from the active site cleft (Fig. 6). This provided a basis for the rational design of a novel class of exosite inhibitors of cathepsins D and other medically relevant aspartic proteases.

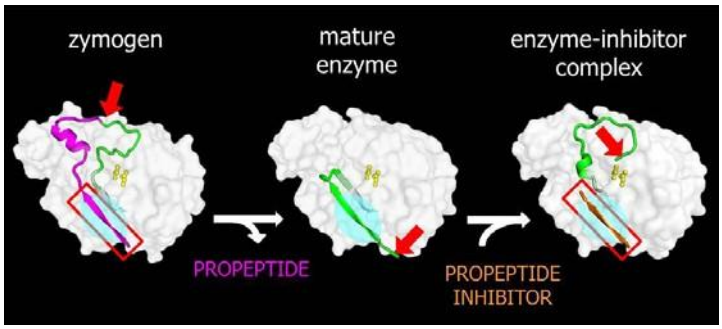


Fig. 6: Three crystal structures of IrCD1: the zymogen is activated by the removal of the propeptide (*magenta*), and the mature enzyme is inhibited by an exosite inhibitor derived from the propeptide (*orange*) [Hanova *et al.*, manuscript in preparation].

(3) Natural regulation of aspartic cathepsins in pathology

Aspartic proteases are critically involved in a number of pathologies and represent therapeutic targets, however, knowledge of their regulation during pathophysiological processes is very limited. The project is focused on the identification of novel molecular mechanisms of the regulation of aspartic protease activity that are employed by natural inhibitors. The discovered structural motifs for inhibition will be evaluated as a scaffold for the development of e.g. anticancer drugs. The research concentrates on two representative aspartic cathepsins: human cathepsin D and an insect cathepsin D.

For the project, we developed selective activity-based probes as a tool for the detection of active cathepsins D (and other cathepsin proteases) in complex proteomes [Nussbaumerova *et al.*, ChemBioChem 2009, Mertens *et al.*, ChemBioChem 2014].

(A) Human cathepsin D (hCD) is associated with tumor development and progression through its complex proliferative and proteolytic action, and also it is an independent marker of poor prognosis in breast cancer [for review, see Fusek *et al.*, Handbook of Proteolytic Enzymes 2013]. No endogenous inhibitors of hCD are known. We discovered that the enzymatic activity of hCD is specifically modulated by bioactive sphingolipids in a complex manner [Zebrakovska *et al.*, Biochim. Biophys. Acta 2011]. Certain sphingosines and ceramides act as inhibitory regulators of hCD, while phosphosphingolipids were identified as activators. Thus, sphingolipids and phosphosphingolipids, known to be antagonistic in their cell-signaling functions, displayed opposite modulation of hCD. Based on these results, we proposed that sphingolipid-based inhibitors of hCD are templates for anticancer drug design. The project now continues with the crystallization of hCD complexes with sphingolipids to explain their binding mode.

(B) Insect cathepsin D: Colorado potato beetle, *Leptinotarsa decemlineata*, is a major pest of potato plants. We performed a functional profiling of proteolytic enzymes in the digestive tract of *L. decemlineata*. The analysis identified a cathepsin D-type aspartic protease named LdCD that is critical for initiating protein digestion. The recombinant LdCD was biochemically characterized and demonstrated to be specifically inhibited by PDI ("potato cathepsin D inhibitor"). PDI is a Kunitz-type protein expressed in potato leaves upon wounding. This indicated that LdCD is a natural target for PDI acting as an antifeedant in plant defense against insect herbivory. Recently, we determined crystal structures of both LdCD and PDI (Fig. 7) [Srp *et al.*, manuscript in preparation]. This work now continues with the crystallization of the LdCD-PDI complex aimed at the determination of a new mechanism of inhibition of aspartic peptidases.

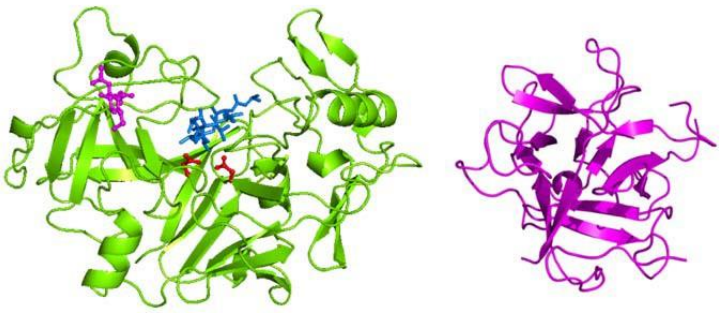


Fig. 7: Crystal structure of LdCD, a digestive cathepsin D from Colorado potato beetle (*left*), and crystal structure of its inhibitor PDI produced in potato leaves (*right*). [Srp *et al*., manuscript in preparation]

(4) Proteases and allergens of arthropod pests

Arthropod pests, such as mites or cockroaches, are a major source of allergens and contributor to the rising incidence of allergic diseases, including bronchial asthma and atopic dermatitis. Proteases produced by household arthropods in their fecal pellets appear to be potent allergens boosted by their proteolytic activity and represent a therapeutic target for the treatment of allergic hypersensitivity. Our research was focused on digestive systems of several important arthropod pests to identify and characterize new component proteases as potential allergens. The topics of this project were as follows:

(A) Digestive proteases of allergenic mites: Using activity-based and MS-based proteomic approaches, we performed the first detailed comparative study of the protease repertoire (degradome) in allergenic mites [Pytelkova *et al*., manuscript in preparation]. XXX Our survey of the digestive system of *A. siro* led to the discovery of a new allergen with α-amylase activity named Aca s 4, which was characterized biochemically and immunologically in detail [Pytelkova *et al*., BMC Biochem. 2012].

(B) Wood-feeding termites are considered major economic arthropod pests worldwide. Cellulose, the major component of termite diet, is digested with the help of a symbiotic microflora. Proteolytic digestion in termites, however, has not been studied so far; we performed its first detailed analysis in the model species *Reticulitermes santonensis* and *Coptotermes formosanus* [Pytelkova *et al*., manuscript in preparation]. XXX This project was extended by the analysis of the labial salivary glands involved in a unique defensive mechanism of the termite *Neocapritermes taracua* [Sobotnik *et al*., Science 2012, Sillam-Dusses *et al*., PLoS One 2012].

References (publications of the team in 2010-2014):

Chmelar, J., Oliveira, C.J., Rezacova, P., Francischetti, I.M., Kovarova, Z., Pejler, G., Kopacek, P., Ribeiro, J.M., Mares, M., Kopecky, J. & Kotsyfakis, M. (2011). A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood,* 117, 736-744.

Fajtova, P., Stefanic, S., Hradilek, M., Dvorak, J., Vondrasek, J., Jilkova, A., McKerrow, J.H., Caffrey, C.R., Mares, M., & Horn, M. (2015). Prolyl oligopeptidase from the blood fluke *Schistosoma mansoni*: from functional analysis to anti-schistosomal inhibitors. *PLoS Neglected Tropical Diseases,* in press.

Fanfrlik, J., Brahmkshatriya, P.S., Rezac, J., Jilkova A., Horn, M., Mares, M., Hobza, P. & Lepsik, M. (2013). Quantum mechanics-based scoring rationalizes the irreversible inactivation of parasitic *Schistosoma mansoni* cysteine peptidase by vinyl sulfone inhibitors. *Journal of*

Physical Chemistry B, 117, 14973-14982.

Franta, Z., Frantova, H., Konvickova, J., Horn, M., Sojka, D., Mares, M. & Kopacek, P. (2010). Dynamics of digestive proteolytic system during blood feeding of the hard tick *Ixodes ricinus*. *Parasites & Vectors,* 3, 119.

Franta, Z., Sojka, D., Frantova, H., Dvorak, J., Horn, M., Srba, J., Talacko, P., Mares, M., Schneider, E., Craik, C. S., McKerrow, J. H., Caffrey, C.R. & Kopacek, P. (2011). IrCL1-The haemoglobinolytic cathepsin L of the hard tick, *Ixodes ricinus*. *International Journal of Parasitology,* 41, 1253-1262.

Fusek, M., Mares, M. & Vetvicka, V. (2013). Cathepsin, D. In: *Handbook of Proteolytic Enzymes, 3rd Edition (Rawlings ND, Salvesen GS, eds.)*, Academic Press, Oxford, pp.54–63.

Horn, M., Jilkova, A., Vondrasek, J., Maresova, L., Caffrey, C. R. & Mares, M. (2011). Mapping the pro-peptide of the *Schistosoma mansoni* cathepsin B1 drug target: modulation of inhibition by heparin and design of mimetic inhibitors. *ACS Chemical Biology,* 6, 609-617.

Horn, M., Fajtova, P., Rojo Arreola, L., Ulrychova, L., Bartosova-Sojkova, P., Franta, Z., Protasio, A.V., Opavsky, D., Vondrasek, J., McKerrow, J.H., Mares, M., Caffrey, C.R. & Dvorak, J. (2014). Trypsin- and chymotrypsin-like serine proteases in *Schistosoma mansoni* - 'the undiscovered country'. *PLoS Neglected Tropical Diseases,* 8, e2766.

Jilkova, A., Rezacova, P., Lepsik, M., Horn, M., Vachova, J., Fanfrlik, J., Brynda, J., McKerrow, J.H., Caffrey, C.R. & Mares, M. (2011). Structural basis for inhibition of cathepsin B drug target from the human blood fluke, *Schistosoma mansoni*. *Journal of Biological Chemistry,* 286, 35770-35781.

Jilkova, A., Horn, M., Rezacova, P., Maresova,, L., Fajtova, P., Brynda, J., Vondrasek ,J., McKerrow, J.H., Caffrey, C.R. & Mares, M. (2014). Activation route of the *Schistosoma mansoni* cathepsin B1 drug target: structural map with a glycosaminoglycan switch. *Structure,* 22, 1786-1798.

Kovarova, Z., Chmelar, J., Sanda, M., Brynda, J., Mares, M. & Rezacova, P. (2010). Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications,* 66, 1453-1457.

Mertens, M.D., Schmitz, J., Horn, M., Furtmann, N., Bajorath, J., Mares, M. & Gutschow, M. (2014). A coumarin-labeled vinyl sulfone as tripeptidomimetic activity-based probe for cysteine cathepsins. *ChemBioChem,* 15, 955-959.

Nussbaumerova, M., Srp, J., Masa, M., Hradilek, M., Sanda, M., Reinis, M., Horn, M. & Mares, M. (2010). Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome. *ChemBioChem,* 11, 1538-1541.

Pytelkova, J., Lepsik, M., Sanda, M., Talacko, P., Maresova, L., & Mares, M. (2012). Enzymatic activity and immunoreactivity of Aca s 4, an alpha-amylase allergen from the storage mite *Acarus siro*. *BMC Biochemistry,* 13, 3.

Salat, J., Paesen, G.C., Rezacova, P., Kotsyfakis, M., Kovarova, Z., Sanda, M., Majtan, J., Grunclova, L., Horka, H., Andersen, J.F., Brynda, J., Horn, M., Nunn, M.A., Kopacek, P., Kopecky, J. & Mares, M. (2010). Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*. *Biochemical Journal,* 429, 103-112.

Sillam-Dusses, D., Krasulova, J., Vrkoslav, V., Pytelkova, J., Cvacka, J., Kutalova, K., Bourguignon, T., Miura, T. & Sobotnik, J. (2012). Comparative study of the labial gland secretion in termites (*Isoptera*). *PLoS One,* 7, e46431.

Sobotnik, J., Bourguignon, T., Hanus, R., Demianova, Z., Pytelkova, J., Mares, M., Foltynova, P., Preisler, J., Cvacka, J., Krasulova, J., & Roisin, Y. (2012). Explosive backpacks in old termite workers. *Science,* 337, 436.

Sojka, D., Franta, Z., Horn, M., Caffrey, C.R., Mares, M. & Kopacek, P. (2013). New insights into the machinery of blood digestion by ticks. *Trends in Parasitology,* 29, 276-285.

Sojka, D., Franta, Z., Frantova, H., Bartosova, P., Horn, M., Vachova, J., O'Donoghue, A.J., Eroy-Reveles, A.A., Craik, C.S., Knudsen, G.M., Caffrey, C.R., McKerrow, J.H., Mares, M. & Kopacek, P. (2012). Characterization of Gut-associated Cathepsin D Hemoglobinase from Tick *Ixodes ricinus* (IrCD1). *Journal of Biological Chemistry,* 287, 21152-21163.

Stefanic, S., Dvorak, J., Horn, M., Braschi, S., Sojka, D., Ruelas, D.S., Suzuki, B., Lim, K.C., Hopkins, S.D., McKerrow, J.H. & Caffrey, C.R. (2010). RNA interference in *Schistosoma mansoni* schistosomula: selectivity, sensitivity and operation for larger-scale screening. *PLoS Neglected Tropical Diseases,* 4, e850.

Zebrakovska, I., Masa, M., Srp, J., Horn, M., Vavrova, K. & Mares, M. (2011). Complex modulation of peptidolytic activity of cathepsin D by sphingolipids. *Biochimica et Biophysica Acta – Molecular and Cell Biology of Lipids,* 1811, 1097-1104.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Iva Pichová - Microbial Proteins

The research in the Group of microbial proteins is primarily focused on investigation of key steps critical for infectivity of major human pathogens – retroviruses, *Mycobacterium tuberculosis,* and *Candida* spp. and on their interaction with host cells and cellular machineries. The research is interdisciplinary, combines experiments with viruses in host cells, molecular biology experiments such as gene silencing, qRT PCR analyses of gene expression, next generation sequencing, protein engineering, *in vitro* characterization of proteins, structural studies with computation analyses. This interdisciplinary approach is based on collaboration with number of researchers from different fields. The group has a close collaboration with several laboratories from IOCB (P. Řezáčová, M. Hájek, M. Hocek, U. Jahn, V. Veverka, J. Fanfrlík, I. Valterová, L. Bednárová), from local Universities (T. Ruml, R. Hrabal - Institute of Chemical Technology in Prague; P. Hamal, M. Hajduch - Faculty of Medicine, Olomouc; V. Sklenář - Masaryk University Brno), Academic Institute (H. Sychrová - Institute of Physiology, ASCR, Prague) as well as with foreign researchers (E. Hunter - Emory University, R. Tuma, Institute of Biotechnology - University of HelsinkiFinland, H. Vogel - Max Planc Institute for Chemical Ecology, Jena, U. Sauer - ETH Zurich,
M. Jaskolski - A. Mickiewicz University in Poznan, J. Briggs - EMBL Heidelberg; D. Laubitz - Institute of Biochemistry and Biophysics, Polish Academy of Sciences; D. Baker - University of Washington, Seattle,USA). The group published 51 papers, two chapters in Handbook of proteolytic enzymes, 1 patent that was licenced by Generi Biotech company during 2010-2014. The research was supported during this period by 4 grants from Czech science foundation (GACR), 2 centres and 1 NPU project from Ministry of Education, one FP7 collaborative project in programme Healts, and 1 project from Czech technology agency (TACR). The group is a member of IOCB-Gilead research centre.

Retroviruses

One of our major goals in this period was to contribute to understanding of **retroviral assembly.** Retroviruses such as Human immunodeficiency virus type 1 (HIV-1) are of great medical importance. Retroviral assembly proceeds in two stages. The multi-domain Gag polyproteins oligomerize into a hexameric lattice, and together with interacting genomic RNA, the immature viral particle is formed either at the plasma membrane (e.g., HIV) or at a distinct location within the cytoplasm of the host cells (e.g., M-PMV and mouse mammary tumor virus (MMTV). During second stage, Gag is cleaved by the viral protease, leading to internal rearrangement of the virus into the mature, infectious form. Assembly is critical for virus function and also represents a potential target for inhibition and development of new types of antiretroviral drugs. We have used HIV-1, M-PMV, MMTV, and murine leukemia virus (strain XMRV) as models for our experiments.

Contribution of M-PMV Gag-structural motifs to assembly of immature particles

The hexameric lattice of an immature retroviral particle consists of Gag polyprotein, which is the precursor of all viral structural proteins. Lentiviral and alpharetroviral Gag

contains a peptide sequence called the spacer peptide (SP), which is localized between the capsid (CA) and nucleocapsid (NC) proteins. SP plays a critical role in intermolecular interactions during the assembly of immature particles of several retroviruses. In contrast, M-PMV does not contain any distinct SP sequence, and the CA- NC connecting region is not organized into a clear rod-like structure. Nevertheless, the CA-NC junction comprises a sequence critical for assembly of immature M-PMV particles.
To find the mechanism working in M-PMV, we prepared a series of CA-NC mutants and examined formation of virus-like particles (VLPs) *in vitro*. We also cloned a proviral construct containing mutated Gag domains and observed virus assembly in 293T cells. We identified the motif of basic residues within the N-terminus of NC that mediates non-specific interaction with nucleic acids, and we found that this motif is necessary for proper virus assembly. An additional critical motif for M-PMV assembly that we identified is a helical segment at the junction between CA and NC, which mimics SP1. The results were published in **Journal of Virology** (Bohmova et al., 2010). The following paper investigated this region in more details. We provided biochemical data confirming the critical role of M-PMV SP-like domain in immature particle assembly, release, processing and infectivity. Using the 8 Å cryo-electron microscopy density maps of immature M-PMV particles (Bharat et al., 2012), we prepared computational models of the SP-like domain, which indicated the structural features required for M-PMV immature particle assembly. This work was published in **Journal of Virology** (Strohalmová-Bohmová et al., 2014).

8Å resolution structure of immature particles
Molecular architecture of mature retroviruses has been significantly advanced by both crystallographic and electron microscopy studies of reconstituted assemblies of viral proteins. These have provided a detailed structural view of the arrangement of the domains of capsid protein (CA) within the mature capsid core in HIV-1 and RSV. In contrast to mature core, there was a lack of detailed information about the interactions in immature particles, mainly due to availability of suitable model and quality of samples for structural studies. We have succeeded in preparation of M-PMV tubular structures from fusion protein capsid-necleocapsid (CANC) faithfully mimicking the assemblies of immature particles. In collaboration with Dr. Briggs group (EMBL, Heidelberg) we solved and published an 8 Å resolution structure of these immature M-PMV tubular particles obtained by combining cryo-electron microscopy and tomography approaches. This structure showed that the transition of an immature retrovirus into its mature infectious form involves dramatic rotations and translations of CA domains and that the CA interactions stabilizing the immature and mature viruses are almost completely distinct. This finding opens a new area for designing and testing compounds that would inhibit the CA-CA interactions critical for immature particle formation and stabilization. The results were published in **Nature** (Bharat et al., 2012). The measurements at cryo-EM and calculations were made in Germany. This work would never be possible to perform without system that we developed previously for *in vitro* assembly of immature retroviral particles using fusion protein CANC, enabling to prepare homogenous population of immature particles

Structural motives influencing assembly of mature viral particles
Maturation of the immature particle takes place during or shortly after budding through the host cell membrane. The viral protease is autocatalytically activated and cleaves Gag into the individual structural proteins matrix (MA), capsid (CA), and nucleocapsid (NC). Proteolytic cleavage initiates changes in the CA structure, enabling subsequent formation of a mature hexameric lattice that forms the viral core. Because the mature CA shell protects the viral genomic RNA, proper rearrangement of CA allowing correct assembly of CA proteins is crucial for the viral life cycle. Retroviral capsid proteins consist of two helical domains, the

N-terminal domain (NTD) and C-terminal domain (CTD), connected by a short flexible segment. Despite low sequence homology among CAs from different retroviruses, the structural arrangement of both CA domains is highly conserved. The NTD consists of an N-terminal β-hairpin followed by six or seven α-helices, and the CTD consists of four α-helices. Upon maturation, the N-terminus of retroviral CA folds to form a β-hairpin, which is primarily stabilized by a salt bridge between the N-terminal proline and a highly conserved aspartate residue. We studied the necessity of additional β-hairpin stabilization in mature CA for formation of M-PMV mature particle both, at the structural level using NMR, and also at the cellular level by analysis of pulse-chase experiments of mutated M-PMV proviral vectors in host cells. The results confirmed that in addition to critical interaction between the N-terminal Pro with Asp 57, the interactions of residues in the vicinity of helix 5 with R14 from the β-hairpin loop are essential for infectivity of the virus. This work was published in **Retrovirology** (Obr et al., 2014).

“

Unique domain in Mouse mammary tumor virus essential for assembly of immature particles

Mouse mammary tumor virus (MMTV) is a member of the betaretrovirus genus and is associated with both mammary adenocarcinomas and T cell lymphomas in mice. While the assembly and budding of other retroviruses have been examined intensively, similar morphogenic analysis of MMTV has not been performed, partly due to difficulties encountered with the construction of molecular clones. MMTV expression in mammalian cells requires stimulation by glucocorticoid hormones, and the resulting expression level of Gag-related polyproteins in host cells is low. We overcame this by engineering of proviralMMTV clone by introduction of foreign promoters and LTR sequences. The MMTV Gag polyprotein contains the non-canonical domains pp21, p3, p8, and “n” of unknown function(s). These domains have unique sequences with low similarity to other retroviral Gag proteins. We found that the region spanning the p8 and n domains are critical for shape determination and assembly. The observed phenotype of MMTV Gag mutants in 293T cells was verified in an RBA rat mammary cell line. These results show that the unique non-canonical domains significantly contribute to MMTV assembly and suggest that actin might be involved in Gag transport and assembly. We published these results in **Journal of Virology** (Zabranský et al., 2010).

The unique assembly of the gammaretroviral immature particles

Using an *in vitro* assembly system of capsid-nucleocapsid protein (CANC) we studied the formation of virus-like particles (VLP) of a gammaretrovirus, the xenotropic murine leukemia virus (MLV)-related virus (XMRV). We showed that, unlike other retroviruses, MLV CA and CANC do not assemble tubular particles characteristic of mature assembly. The prevention of β-hairpin formation by the deletion of either the N-terminal proline or 10 initial amino acids enabled the assembly of ΔProCANC or Δ10CANC into immature-like spherical particles. Detailed three-dimensional (3D) structural analysis of these particles using cryo-electron microscopy combined with cryo-electron tomography revealed that in contrast to other retroviruses (HIV, RSV, M-PMV), the outermost layer in the MLV particles does not form an ordered hexameric lattice below a disordered N-terminal CA layer. The results were published in **Journal of Virology** (Hadravová et al., 2012).

Development of high-throughput method for monitoring the assembly

We developed a rapid and simple suitable for testing libraries of compounds inhibiting assembly process. The method is based on using fluorescently labeled oligonucleotide (ON) that binds to the NC moiety to initiate the *in vitro* assembly of HIV-1 CANC particles that protect the ON. The non-associated ON is degraded by nuclease and emitted

fluorescence is then recorded. Inhibitors that block the assembly of helical tubes render the ON accessible to nuclease degradation and the level of inhibition may be calculated based on the released fluorescence. The strength of this method resides in its simplicity and suitability both for analyzing of inhibitors of CA dimerization and compounds interfering with nucleic acid binding to the NC. This assay was **patented** (Rumlová et al., 2012, PV 2010-718) and licenced by company Generi Biotech, CR. Currently is available on the market.

Role of the G-patch domain in the retroviral life cycle

M-PMV, like some other betaretroviruses and some class II-related endogenous retroviruses, encodes a glycine rich G-patch domain (GPD), which is localized at the 3´ end of the *pro* gene upstream of the *pro-pol* ribosomal frameshift sequence. During maturation of the Gag-Pro polyprotein, the GPD transiently remains a C-terminal part of the protease, from which it is autocatalytically cleaved during maturation. GPDs are typically involved in mRNA splicing and DNA repair in eukaryotic organisms. We investigated the role of GPD in various steps of the retroviral life cycle. We analyzed the possibility that GPD might play a role in splicing of M-PMV *env* mRNA, which is used for translation of Env polyprotein precursor Pr86. We also examined whether GPD might mediate transport of full-length genomic RNA or its incorporation into viral particles. We found that GPD is not necessary for these processes. We also showed that GPD is not incorporated *in trans* in virions; however, we confirmed that GPD is a part of mature reverse transcriptase (RT) in released M-PMV virions and contributes to RT activity and subsequently to virus infectivity. The results were published in **Journal of Virology** (Křížová et al., 2012).

Viral – cell proteins interactions

By using yeast two-hybrid screen, we identified breast cancer-associated protein (BCA3) as a novel binding partner of M-PMV and HIV-1 proteases. The interaction was confirmed by co-immunoprecipitation and immunocolocalization of M-PMV and HIV-1 PRs and BCA3. We found that full-length, but not C-terminally truncated BCA3, was incorporated into M-PMV virions. We ruled out the potential role of the G-patch domain, a glycine-rich domain located at the C-terminus of M-PMV PR, in BCA3 interaction and virion incorporation. Expression of BCA3 does not affect M-PMV particle release and proteolytic processing; however, it slightly increases M-PMV infectivity. This work was published in **Journal of General Virology** (Rumlová et al., 2014).

HIV-1 protease-induced apoptosis

Apoptosis is one of the presumptive causes of $CD4^+$ T cell depletion during HIV infection and progression to AIDS. However, the precise role of HIV-1 in this process remains unexplained. HIV-1 protease (PR) has been suggested as a possible factor, but a direct link between HIV-1 PR enzymatic activity and apoptosis has not been established. We found that expression of active HIV-1 PR induces death in HeLa and HEK-293 cells via the mitochondrial apoptotic pathway. This conclusion is based on *in vivo* observations of the direct localization of HIV-1 PR in mitochondria, a key player in triggering apoptosis. Moreover, we observed an HIV-1 PR concentration-dependent decrease in mitochondrial membrane potential and the role of HIV-1 PR in activation of caspase 9, PARP cleavage and DNA fragmentation. In addition, *in vitro* data demonstrated that HIV-1 PR mediates cleavage of mitochondrial proteins Tom22, VDAC and ANT, leading to release of AIF and Hsp60 proteins. We also found that BCA3 accelerates p53 transcriptional activity on the *bax* promoter, thus elevating the cellular level of pro-apoptotic Bax protein. This work was published in **Retrovirology** (Rumlová et al., 2014)

Solution of the high resolution structure of M-PMV 13PR monomer by unprecedented way

Retroviral proteases (PR) have been investigated for years in this group, results were published in 16 international journals and our data indicated unusual folding of Mason Pfizer Monkey (MPMV) virus protease in a form of monomer. We prepared a series of protease mutants and asked Polish crystallographers and American programmers for collaboration on solution of MPMV PR structure. The crystallographers could crystallized PR in several crystal forms however, solution of the structure resisted to attempts, which utilized all available programs and existing crystallographic models of retropepsins (full dimers and individual subunits). The mr-rosetta algorithm, which has an outstanding record of success with difficult structures also failed to produce a solution using the existing models. This daunting protein-folding problem was therefore presented as a challenge to players of internet game Foldit, which is based on Rosetta algorithm and was developed in Baker's laboratory. This game represents one of the first examples of "crowd" science. The model of M-PMV protease monomer was in 2010 the first real-world model for game players, who generated over one million models in less than three weeks, starting from the NMR coordinates from our previous measurements. One of these solutions, when submitted to MR calculations in mr-rosetta did produce a plausible crystal structure that could be easily refined to an R factor of 0.169 with excellent geometry. The winning Foldit players were co-authors of our publication in **Nature Structure Biology** in 2011 (Khatib et al., 2011). The structure was later published in **Acta Crystallographica D** (Gilski et al., 2011). This paper was announced by different world media, including Czech newspapers, it resulted in radio and TV dialogs in the USA, Poland and Czech Republic, it served as an inspiration for book story. The paper is highly cited (91 citations) and has contributed significantly to development of networked science and to interest of general public in protein folding.

***Pathogenic yeasts of genus* Candida**

Candida spp. are opportunistic pathogens that can cause diseases ranging from superficial mycoses to disseminated and often fatal infections. The processes important for virulence of *C. albicans*, the member of the *Candida* genus most commonly found in humans, include adherence to host tissues, yeast-hypha morphogenesis, secretion of hydrolases (in particular aspartic proteases and phospholipases), and ability to grow on diverse nitrogen and carbon sources under variety of conditions. Secreted aspartic proteases (Saps) facilitate the breakdown of host tissues and thus mediate penetration of pathogen into the host and contribute to utilization of host macromolecules as a source of nutrients. Pathogenic *Candida* spp. contain more *SAP* genes than non-pathogenic species. *SAP* gene and Sap protein regulation remain poorly understood. Our long-term objective in this project is to describe the extracellular proteolysis of *Candida* spp. in details.

Evidence for the presence of proteolytically active secreted aspartic proteinase 1 of *Candida parapsilosis* in the cell wall

We studied positioning of Sapp1p in the upper layer of the cell wall during the secretion. Using biotinylation of cell wall proteins followed by extraction and mass-spec analysis of labeled proteinase molecules we showed that Sapp1p secretion from the yeast cells is not a random process. In addition, we performed proteinase activity assay using the whole *C. parapsilosis* cells and revealed the proteinase is active and therefore fully folded even prior to secretion. The non-biotinylated lysine residues were located in the N-terminal part of the Sapp1p molecule, showing that in the large population of proteinase molecules the C-terminal domain emerges to the extracellular space first. The pathogenic yeast might use this cell-associated proteinase activity to enhance degradation of appropriate substrates. The ability of

yeast cells to retain active Sap molecules in the cell wall before their secretion into extracellular space might provide a great benefit for pathogenic *Candida spp.*, mostly during biofilm formation, but also during adhesion and colonization of host tissues. The results were published in **Protein Science** (Vinterová et al., 2011)

The crystal structure of protease Sapp1p from *Candida parapsilosis* in complex with ritonavir

Oropharyngeal candidiasis (OPC) is one of the common diseases related to HIV/AIDS. Marked reduction of OPC has been observed in the AIDS patients who receive treatment with protease inhibitors. It has been hypothesized that HIV PR inhibitors are active also against aspartic proteases secreted by *Candida* species (Sap), and can thus alleviate opportunistic candidiasis. Nevertheless, K_i values that were obtained for HIV PR inhibitors and Saps from several pathogenic *Candida* species were within micromolar range or higher. To contribute to discussion, whether HIV protease inhibitors can act against opportunistic mycoses, we crystallized protease Sapp1p from *C. parapsilosis* with ritonavir and determined the structure of this complex at 2.4 Å resolution. Since ritonavir inhibits Sapp1p with K_i - 1.9 μM, it was rather difficult to obtain crystals of reasonable quality and stability, and to achieve a better resolution. Despite these problems, the structural analysis showed that ritonavir binds to the active site of Sapp1p. The central hydroxyl group of ritonavir interacts with catalytic aspartates in a similar way as pepstatin A, a classical inhibitor of aspartic proteases. However, the interactions with individual substrate binding pockets substantially differ. Ritonavir displays low shape complementarity with individual substrate binding subsites. Moreover, non-polar interactions between ritonavir and Sapp1p are quite limited. The structure of Sapp1p-RTV complex shed light on the concept multi-target effect of ritonavir. This work was published in **Journal of Enzyme Inhibition and Medicinal Chemistry** (Dostal et al., 2012)

Investigation of Vacuolar Degradation of proteins in *Candida albicans*

Vacuoles play an important role in the physiology of pathogenic Candida spp. However, information on *Candida albicans* vacuolar enzymes, their properties, and regulation is scarce. Expression of the genes APR1 and CPY1 encoding vacuolar aspartic protease and serine carboxypeptidase, respectively, was analyzed using a clinical isolate of *C. albicans.* Analysis of APR1 and CPY1 expression under nitrogen-limited conditions revealed that the genes were regulated on both the transcriptional and translational levels and detectable amounts of Apr1p were synthesized only when *C. albicans* was grown in nitrogen-limited media. We found that the *apr1Δ* strain was not able to form hyphae in the medium, which contained prolin as sole nitrogen source. Filamentation of *cpy1Δ* strain was not diminished. Apr1p therefore seems to be important for morphological transition under nitrogen limited conditions. The results were published in **Canadian Journal of Microbiology** (Bauerová et al., 2012) and **Folia Microbiologica** (Bauerova et al., 2014)

Mycobacterium tuberculosis

Tuberculosis (TB) is now the biggest bacterial killer of human population. In recent years the TB epidemic is accelerated by the emergence of multi- and extensively-drug resistant *Myobacterium tuberculosis* (Mtb) strains. Moreover, HIV activates latent Mtb, accelerates TB progression, and increases Mtb transmission rates, critical not only for diagnosed TB sufferers, but also for the more than two billion individuals with undiagnosed and/or asymptotic latent Mtb forms and their uninfected contacts. The biology of *Mtb* that persists in infected people is poorly understood. *Mtb* reprograms metabolism during latent infection in response to the host environment. Current knowledge of *Mtb* suggests that adaptation of the bacteria to the host

environment is a defining feature of *Mtb* pathogenicity. Recent findings suggest that simultaneous catabolism of lipids and glucose and regulation of carbon metabolism represent significant determinants of the ability of *Mtb* to persist in host. The long term goal in this project is to investigate the regulation of metabolic pathways in Mtb. This topic initiated within the European project SysteMTb of the 7FP in collaboration with several laboratories where omics (transcriptomics, proteomics, and metabolomics) experiments identified the critical nodes and enzymes that shoud contribute to reprogramming the metabolism during latent Mtb infection. The role of Pichova group was to perform functional and structural studies of selected metabolic enzymes. In addition to this focus, we started collaboration within IOCB with organic chemists on developing of inhibitors of metabolic Mtb enzymes.

Regulation of the phosphoenolpyruvate-pyruvate-oxalacetate (PEPO) node in *M. tuberculosis*

The main pathways of central metabolism (glycolysis, gluconeogenesis, and the tricarboxylic acid cycle) are inter-connected *via* the PEPO node. The reactions within this node are catalyzed by a set of enzymes that can be regulated under different conditions by various factors, including enzyme activities and specificities, substrate availability, and product inhibition. This flexible node can switch the carbon flux distribution within the central metabolism. Phosphoenolpyruvate carboxykinase (Pck, EC 4.1.1.32) is the enzyme at the center of PEPO node and metabolic rearrangements during latent Mtb infection. Although Pck is associated with gluconeogenesis under standard growth conditions, the enzyme can catalyze the reverse reaction, supporting the synthesis of oxaloacetate (OAA) by an anaplerotic CO_2-fixing reaction under conditions leading to slowed or stopped bacterial replication. Transcriptomics analyses indicated high up-regulation of Pck in slowly growing Mtb and metabolomics analyses reported intracellular metabolite concentrations strongly favoring anaplerotic directions. The mechanism of this switch however, was not known. To study the mechanisms that regulate the switch between Pck's two functions, we cloned this enzyme and systematically investigated factors influencing the enzyme's gluconeogenic and anaplerotic reaction kinetics. We demonstrate that a reducing environment, as found under hypoxia-triggered non-replicating conditions, accelerates the reaction in the anaplerotic direction. Furthermore, we identified cellular proteins that interact with Pck. The interaction between Pck and the reduced form of mycobacterial thioredoxin, gene expression of which is increased under hypoxic conditions, also increased the enzyme's anaplerotic activity. We thus propose that a reducing environment and the protein-protein interaction with thioredoxin in particular modulate the Pck's activity in "dormant" *M. tuberculosis* and thus enable the enzyme's anaplerotic function under fermentative growth conditions. Our results show that MTb Pck operating in the anaplerotic direction can fully profit from increased intracellular reducing conditions and from up-regulation of proteins maintaining the intracellular reduced state and antioxidant defense. The data were published in **Journal of Biological Chemistry** (Machova et al., 2014). The PCK complexes were identified in the Polish laboratory, the metabolic flow was confirmed by MS experiments from ETH Zurrich.

Insight into the mechanism of action of 7-substituted-7-deazapurine nucleosides on human (*h* ADK) **and mycobacterial adenosine kinases (Mtb ADK) by X ray, NMR, enzyme kinetics, and Semiempirical quantum mechanical (SQM) based analyses**

Enzymes involved in the biosynthesis of purine nucleotides are potential targets for development of new types of compounds active against Mtb. In most organisms, purine nucleotides are formed from purine bases and phosphoribosyl-1-pyrophosphate or by phosphorylation of nucleosides in the purine salvage pathway, or they are synthesized *de novo* in a multi-step sequence. Mtb expresses enzymes from both pathways; however, the

interdependence and regulation of these processes remains unclear. Of the variety of possible target enzymes from the Mtb purine salvage pathway, adenosine kinase (ADK) is considered as a promising target for drug development. ADK catalyzes phosphorylation of adenosine to adenosine monophosphate (AMP) in the phosphoryl transfer reaction using adenosine 5′-triphosphate (ATP) as a substrate and releasing adenosine 5′-diphosphate (ADP). Mtb ADK this shares low structural similarity with human ADK. In collaboration with M. Hocek group we have designed and tested series of 7-deazaadenine ribonucleosides bearing small and bulky substituents in different positions. Several compounds were specific inhibitors of Mtb enzyme, were not cytotoxic, were active against Mtb strain My331/88 and drug-resistant strain Praha 131 *in vitro* (MIC = 2 - 4 μM) with good therapeutic index. These compounds are promising lead structures for further drug development. The results were published in **CheMedChem**, **J.Med. Chem.**, and **Agents. Med. Chem. Commun**. (Spáčilová et al., 2010, Bourderioux et al., 2011, Perlíková et al., 2013).

Testing of large series of these compounds with human and mycobacterial kinases revealed great difference in phosphorylation of modified nucleosides by these enzymes.While no synthesized analog was phosphorylated by Mtb ADK, several compounds served as good substrates for *h*ADK. This may indicate differences in the structures of these enzymes and/or different binding modes of nucleoside analogs into the adenosine and ATP binding sites. To gain insight into the mechanism of action of 7-substituted-7-deazapurine nucleosides on *h*ADK and Mtb ADK, we solved the X-ray structures of complexes of Mtb and *h*ADKs with inhibitor, used 1D ^{1}H STD NMR experiments to establish the mode of competition of 7- hetaryl-7-deazapurine derivatives with adenosine and ATPγS for their respective binding sites, and applied semiempirical quantum mechanical (SQM) based analysis for calculation of binging energies in Mtb and human ADK. The results showed that following binding of the 7- (het)aryl-7-deazaadenine ribonucleosides into the adenosine site, Mtb ADK adopts a unique semi-open conformation of the lid domain and *h*ADK the close conformation. Enzyme kinetics and 1D ^{1}H STD NMR analysis of the inhibitors' competition with adenosine and ATPγS showed that 7-(het)aryl-7-deazaadenine ribonucleosides bind preferentially into the adenosine site in *h*ADK and are readily accommodated in both, to the adenosine and ATP sites in Mtb ADK. Quantum mechanical analysis indicated that these compounds have higher affinity for the ATP site of Mtb ADK in the open conformation but movement of the lid domain, increases binding affinity of these inhibitors in the adenosine site. Both the
preferential binding of 7-(het)aryl-7-deazaadenine ribonucleosides in the Mtb ADK ATP binding site and the unique semi-open conformation of Mtb ADK during inhibitor binding in the adenosine site explain the observation that 7-(het)aryl-7-deazaadenine ribonucleosides are not phosphorylated by Mtb ADK. Similar mechanisms may also occur for deazaadenine ribonucleoside derivatives modified at other positions. Results were published in **J.Med. Chem**. (Snášel et al. 2014). In this project we coordinated and planned experiments, cloned enzymes, crystallized proteins, end performed kinec and inhibition studies.

Papers related to this report:

Bohmová K, Hadravová R, Stokrova J, Tuma R, Ruml T, Pichová I, Rumlová M. The effect of dimerizing domains and basic residues on in vitro and in vivo assembly of Mason-Pfizer monkey virus and Human immunodeficiency virus. J Virol. **84**, 1977-88, (2010).

Zabransky A, Hoboth P, Hadravová R, Stokrova J, Sakalian M, Pichová I. The noncanonical Gag domains p8 and n are critical for assembly and release of mouse mammary tumor virus. J Virol. **84**, 11555 – 11559, (2010).

Spácilová P, Naus P, Pohl R, Votruba I, Snásel J, Zábranská H, Pichová I, Ameral R, Birkus G, Cihlár T, Hocek M. CycloSal-phosphate pronucleotides of cytostatic 6-(Het)aryl-7-deazapurine

ribonucleosides: Synthesis, cytostatic activity, and inhibition of adenosine kinases. ChemMedChem. **5**, 1386-96, (2010).

Khatib F, Dimaio F, Foldit Contenders Group, Foldit Void Crushers Group, Cooper S, Kazmierczyk M, Gilski M, Krzywda S, Zabranska H, Pichova I, Thompson J,Popović Z, Jaskolski M, Baker D. Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat Struct Mol Biol. **18**, 1175-1177, (2011).

Gilski M, Kazmierczyk M, Krzywda S, Zabranska H, Cooper S, Popovic Z, Khatib F, DiMaio F, Thompson J, Baker D, Pichova I, Jaskolski M. High-resolution structure of a retroviral protease folded as a monomer. Acta Cryst. **67**, 907-914, (2011).

Vinterová Z, Sanda M, Dostál J, Hrušková-Heidingsfeldová O, Pichová I. Evidence for the presence of proteolytically active secreted aspartic proteinase 1 of Candida parapsilosis in the cell wall. Protein Sci. **12**, 2004-12 (2011).

Bourderioux A, Naus P, Perlíková P, Pohl R, Pichová I, Votruba I, Dzubák P, Konecný P, Hajdúch M, Stray KM, Wang T, Ray AS, Feng JY, Birkus G, Cihlar T, Hocek M. Synthesis and Significant Cytostatic Activity of 7-Hetaryl-7-deazaadenosines. J Med Chem. **54**, 5498-507, (2011). Dostál J, Brynda J, Hrušková-Heidingsfeldová O, PachlP, Pichová I, Řezáčová P. The crystal structure of protease Sapp1p from *Candida parapsilosis* in complex with the HIV protease inhibitor ritonavir. J. Enz. Inhib. Med. Chem. **27**(1), 160-165, (2012).

Hadravová R, de Marco A, Ulbrich P, Stokrová J, Dolezal M, Pichová I, Ruml T, Briggs JA, Rumlová M. *In vitro* assembly of virus-like particles of a Gammaretrovirus, the Murine Leukemia Virus (XMRV). J Virol. **86**(3), 1297-1306, (2012).Křížová I, Hadravová R, Štokrová J, Günterová J, Doležal M, Ruml T, Rumlová M, Pichová I. The G-patch domain of Mason-Pfizer monkey virus is a part of reverse transcriptase. J Virol. **86**(4), 1988-1998, (2012).

Bharat TAM, Davey NE, Ulbrich P, Riches JD, de Marco A, Rumlova M, Sachse C, Ruml T, Briggs JAG. **Structure of the immature retroviral capsid at 8 A resolution by cryo-electron microscopy**. Nature **487**, 385-389, (2012).

Bauerová V, Pichova I, Hrušková-Heidingsfeldová O. Nitrogen source and growth stage of Candida albicans influence expression level of vacuolar aspartic protease Apr1p and carboxypeptidase Cpy1p. Can. J. Microbiol. **58**, 678-681, (2012).

Strohalmová-Bohmová K, Spiwok V, Lepšík M, Hadravová R, Křížová I, Ulbrich P, Pichová I, Bednárová L, Ruml T, Rumlová M. Role of Mason-Pfizer monkey virus CA-NC spacer peptide-like domain in assembly of immature particles. J Virol. **88**(24), 14148-60, (2014).

Obr M, Hadravová R, DoleŽal M, KříŽová I, Papoušková V, Zídek L, Hrabal R, Ruml T, Rumlová M. Stabilization of the β-hairpin in Mason-Pfizer monkey virus capsid protein- a critical step for infectivity. Retrovirology **11**, 94, (2014).

Rumlová M, Křížová I, Hadravová R, Doležal M, Strohalmová K, Keprová A, Pichová I, Ruml T. Breast cancer-associated protein--a novel binding partner of Mason-Pfizer monkey virus protease. J Gen Virol. **95**, 1383-9, (2014).

Rumlová M, Křížová I, Keprová A, Hadravová R, Doležal M, Strohalmová K, Pichová I, Hájek M, Ruml T. HIV-1 protease-induced apoptosis. Retrovirology **11**, 37, (2014).

Bauerová V, Hájek M, Pichová I, Hrušková-Heidingsfeldová O. Intracellular aspartic proteinase Apr1p of Candida albicans is required for morphological transition under nitrogen-limited conditions but not for macrophage killing. Folia Microbiol (Praha). **59**(6), 485-93, (2014).

Spácilová P, Naus P, Pohl R, Votruba I, Snásel J, Zábranská H, Pichová I, Ameral R, Birkus G, Cihlár T, Hocek M. CycloSal-phosphate pronucleotides of cytostatic 6-(Het)aryl-7-deazapurine ribonucleosides: Synthesis, cytostatic activity, and inhibition of adenosine kinases. 2010 ChemMedChem. 5, 1386 - 1396 (2010).

Bourderioux A, Naus P, Perlíková P, Pohl R, Pichová I, Votruba I, Dzubák P, Konecný P, Hajdúch M, Stray KM, Wang T, Ray AS, Feng JY, Birkus G, Cihlar T, Hocek M. Synthesis and Significant Cytostatic Activity of 7-Hetaryl-7-deazaadenosines. **J Med Chem**. 54, 5498-507, (2011).

Perlíková, P., Konečný, P., Nauš, P., Snášel, J., Votruba, I., Džubák P., Pichová, I., Hajdúch, M., Hocek M. 6-Alkyl-, 6-Aryl- and 6-Hetaryl-7-deazapurine Ribonucleosides as Inhibitors of Human or MTB Adenosine Kinase and Potential Antimycobacterial Agents. Med. Chem. Commun. **4**, 1497-1500, (2013).

Snášel J, Nauš P, Dostál J, Hnízda A, Fanfrlík J, Brynda J, Bourderioux A, Dušek M, Dvořáková H, Stolaříková J, Zábranská H, Pohl R, Konečný P, Džubák P, Votruba I, Hajdúch M, Rezáčová P,

Veverka V, Hocek M, Pichová I. Structural basis for inhibition of mycobacterial and human adenosine kinase by 7-substituted 7-(Het)aryl-7-deazaadenine ribonucleosides. J Med Chem. **57**(20), 8268-79, (2014).

Machová I, Snášel J, Zimmermann M, Laibitz D, Plocinski P, Oehlmann W, Dostál J, Sauer U, Pichová I. *Mycobacterium tuberculosis* phosphoenolpyruvate carboxykinase is regulated by redox mechanisms and interaction with thioredoxin. J. Biol. Chem. **289**(19), 13066-78, (2014).

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Pavel Hobza - Quantum Chemical Calculations on Model Complexes

2.1. Benchmark calculations in model noncovalent complexes and method development

Our major achievement in this field was the introduction of the S66 data set,[1.1, 1.2] a set of accurate CCSD(T)/CBS interaction energies in variety of noncovalent complexes of organic molecules. It became the de facto standard for benchmarking and parameterization of methods targeted at description of noncovalent interactions. The extensions of this data set cover also large number of nonequilibrium geometries of the 66 model complexes with total of more than thousand points calculated at CCSD(T)/CBS level. To extend the coverage of the benchmark data, we developed additional data sets of halogenated compounds[1.3] and large noncovalent complexes.[1.4] These allowed us to test the transferability of methods to other elements and larger systems.

The new benchmark data enabled us to parameterize accurate correlated QM methods. The methods we have developed yield the best results at the given level of computational complexity. The spin-component scaled CCSD parameterized for noncovalent interactions[1.5] (SCS-MI-CCSD) yields results very close to CCSD(T) at substantially reduced cost. We continued our work on scaled MP3 approach, introducing the MP2.X method[1.6] and showing that MP2.5 yield excellent results in wide range of systems.[1.7, 1.8]

The progress in computation allowed us to explore noncovalent interactions at unprecedentedly high level, from CCSDT and CCSDT(Q) to full configuration interaction.[1.9, 1.10] We have also explored the basis set dependence of interaction energies computed at this level.[1.10] Using this knowledge, we developed the A24 dataset of accurate interaction energies[1.11] (all-electron, accurate CSCD(T)/CBS with CCSDT(Q) and relativistic corrections) that enabled assessing the accuracy of the common benchmark calculations. These accurate methods were also compared to accurate experimental data available for HF dimer.[1.12]

To bring accurate description of noncovalent interactions to very large systems, we have developed a series of corrections for semiempirical QM methods. We have improved the hydrogen-bonding correction we have introduced earlier[1.13, 1.14] and updated the dispersion correction[1.14] to reflect the latest development in DFT-D methodology. Additionally, we were the first to enable accurate description of halogen bonding in semiempirical QM methods by the means of specific correction.[1.15]

References

2.1.1 Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. J. Chem. Theory Comput. **2011**, 7 (8), 2427–2438.

2.1.2 Řezáč, J.; Riley, K. E.; Hobza, P. Extensions of the S66 Data Set: More Accurate Interaction Energies and Angular-Displaced Nonequilibrium Geometries. J. Chem. Theory Comput. **2011**, 7 (11), 3466–3470.

2.1.3 Mintz, B. J.; Parks, J. M. Benchmark Interaction Energies for Biologically Relevant Noncovalent Complexes Containing Divalent Sulfur. J. Phys. Chem. A **2012**, 116 (3), 1086–1092.

2.1.4 Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. J. Chem. Theory Comput. **2013**, 9 (8), 3364–3374.

2.1.5 Pitoňák, M.; Řezáč, J.; Hobza, P. Spin-Component Scaled Coupled-Clusters Singles and Doubles Optimized towards Calculation of Noncovalent Interactions. Phys. Chem. Chem. Phys. **2010**, 12 (33), 9611.

2.1.6 Riley, K. E.; Řezáč, J.; Hobza, P. MP2.X: A Generalized MP2.5 Method That Produces Improved Binding Energies with Smaller Basis Sets. Phys Chem Chem Phys **2011**, 13, 21121– 21125.

2.1.7 Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of

Halogenated Molecules. J. Chem. Theory Comput. **2012**, 8 (11), 4285–4292.
2.1.8 Sedlák, R.; Riley, K. E.; Řezáč, J.; Pitoňák, M.; Hobza, P. MP2.5 and MP2.X: Approaching CCSD(T) Quality Description of Noncovalent Interaction at the Cost of a Single CCSD Iteration. ChemPhysChem **2013**, 14 (4), 698–707.
2.1.9 Řezáč, J.; Šimová, L.; Hobza, P. CCSD[T] Describes Noncovalent Interactions Better than the CCSD(T), CCSD(TQ), and CCSDT Methods. J. Chem. Theory Comput. **2012**.
2.1.10 Šimová, L.; Řezáč, J.; Hobza, P. Convergence of the Interaction Energies in Noncovalent Complexes in the Coupled-Cluster Methods Up to Full Configuration Interaction. J. Chem. Theory Comput. **2013**, 9 (8), 3420–3428.
2.1.11 Řezáč, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the "Gold Standard," CCSD(T) at the Complete Basis Set Limit? J. Chem. Theory Comput. **2013**, 9 (5), 2151–2155.
2.1.12 Řezáč, J.; Hobza, P. Ab Initio Quantum Mechanical Description of Noncovalent Interactions at Its Limits: Approaching the Experimental Dissociation Energy of the HF Dimer. J. Chem. Theory Comput. **2014**.
2.1.13 Korth, M.; Pitoňák, M.; Řezáč, J.; Hobza, P. A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods. J. Chem. Theory Comput. **2010**, 6 (1), 344–352.
2.1.14 Řezáč, J.; Hobza, P. Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. J Chem Theory Comput **2012**, 8 (1), 141–151.
2.1.15 Řezáč, J.; Hobza, P. A Halogen-Bonding Correction for the Semiempirical PM6 Method. Chem. Phys. Lett. **2011**, 506 (4-6), 286–289.

1.1. Development of SQM scoring and its applications

We introduced a fast and reliable rescoring scheme for docked complexes based on a semiempirical quantum mechanical (QM) PM6-DH2 method [2.1]. The method utilizes a PM6-based Hamiltonian with corrections for dispersion energy and hydrogen bonds. The total score is constructed as the sum of the PM6-DH2 interaction enthalpy, the empirical force field (AMBER) interaction entropy, and the sum of the deformation (PM6-DH2, SMD) and the desolvation (SMD) energies of the ligand. The SMD method was chosen because it outperformed other implicit solvation methods [2.2]. The main advantage of the procedure is the fact that we do not add any empirical parameter for either an individual component of the total score or an individual protein-ligand complex. This rescoring method is applied to a very challenging system, namely, the HIV-1 protease with a set of ligands. The difficulty of the system is in the assignment of the protonation of the active site aspartates [2.3]. As opposed to the conventional DOCK procedure, the PM6-DH2 rescoring based on all of the terms distinguishes between binders and nonbinders and provides a reliable correlation of the theoretical and experimental binding free energies. Such a dramatic improvement, resulting from the PM6-DH2 rescoring of all the complexes, provides a valuable yet inexpensive tool for rational drug discovery and de novo ligand design. This QM-based scoring function [2.4] has been modified and extended to treat covalent binding of inhibitory ligands [2.5]. The enhancements are (i) the description of the covalent bond breakage and formation using hybrid QM/semiempirical QM (QM/SQM) restrained optimizations and (ii) the addition of the new $\Delta G_{cov}'$ term to the noncovalent score, describing the "free" energy difference between the covalent and noncovalent complexes. This enhanced QM-based scoring function is applied to a series of 20 vinyl sulfone-based inhibitory compounds inactivating the cysteine peptidase cathepsin B1 of the Schistosoma mansoni parasite (SmCB1). The available X-ray structure of the SmCB1 in complex with a potent vinyl sulfone inhibitor K11017 is used as a template to build the other covalently bound complexes and to model the derived noncovalent complexes. We present the correlation of the covalent score and its constituents with the experimental binding data. Four outliers are identified. They contain bulky R1′ substituents structurally divergent from the template, which might induce larger protein rearrangements than could be accurately modeled. In summary, we propose a new computational approach and an optimal protocol for the rapid evaluation and prospective design of covalent inhibitors with a conserved binding mode.
The above-mentioned QM-based scoring function has been applied to a variety of difficult protein-ligand systems, involving halogen bond (CK2 [2.6], aldose reductase [2.7]), covalent binding (Schistosoma mansoni Cathepsin B1 [2.5]) or structural water molecules (serine racemase [2.8]). Apart from this, two studies on CDK2 were carried out ([2.9], [2.10]).

References

1.1.1 Fanfrlík, J.; Bronowska, A.K.; Řezáč, J.; Přenosil,O.; Konvalinka, J.; Hobza, P. A Reliable Docking/Scoring Scheme Based on the Semiempirical Quantum Mechanical PM6-DH2 Method Accurately Covering Dispersion and H-Bonding: HIV-1 Protease with 22 Ligands. J. Phys. Chem. B **2010**, 114, 12666-12678.

1.1.2 Kolář, M.; Fanfrlík, J.; Lepšík, M.; Forti, F.; Luque, F. J.; Hobza, P. Assessing the Accuracy and Performance of Implicit Solvent Models for Drug Molecules: Conformational Ensemble Approaches. J. Phys. Chem. B **2013**, 117, 5950–5962.

1.1.3 Pecina, A.; Přenosil, O.; Fanfrlík, J.; Řezáč, J.; Granatier, J.;Hobza, P.; Lepšík, M. On the Reliability of the Corrected Semiempirical Quantum Chemical Method (PM6-DH2) for Assigning the Protonation States in HIV-1 Protease/Inhibitor Complexes, Collect. Czech. Chem. Commun. **2011**, 76, 457-479.

1.1.4 Lepšík, M., Řezáč, J., Kolář, M., Pecina, A., Hobza P., and Fanfrlík, J. The Semiempirical Quantum Mechanical Scoring Function for In-Silico Drug Design. ChemPlusChem. **2013**, 78, 921-931.

1.1.5 Fanfrlík, J.; Brahmkshatriya, P.S.; Řezáč, J.; Jílková, A.; Horn, M.; Mareš, M.; Hobza, P.; Lepšík, M. Quantum Mechanics-Based Scoring Rationalizes Irreversible Inactivation of Parasitic Schistosoma mansoni Cysteine Peptidase by Vinyl Sulfone Inhibitors. J. Phys. Chem. B, **2013**, 117, 14973-14982.

1.1.6 Dobeš, P.; Řezáč, J.; Fanfrlík, J.; Otyepka, M.; Hobza, P. Semiempirical Quantum Mechanical Method PM6-DH2X Describes the Geometry and Energetics of CK2-Inhibitor Complexes Involving Halogen Bonds Well, While the Empirical Potential Fails. J. Phys. Chem. B **2011**, 115, 8581–8589.

1.1.7 Fanfrlík, J.; Kolář, M.; Kamlar, M.; Hurný, D.; Ruiz, F.X.; Cousido-Siah, A.; Mitschler, A.; Řezáč, J.; Munusamy, E.; Lepšík, M.; Matějíček, P.; Veselý, J.; Podjarny, A.; Hobza, P. The modulation of aldose reductase inhibition by halogen bond tuning. ACS Chem. Biol., **2013**, 8, 2484-2492.

1.1.8 Vorlova, B.; Nachtigallova, D.; Jiraskova-Vanickova, J.; Ajani, H.; Jansa, P.; Rezac, J.; Fanfrlik, J.; Otyepka, M.; Hobza, P.; Konvalinka, J.; Lepsik, M. Malonate-based inhibitors of mammalian serine racemase: Kinetic characterization and structure-based computational study. Eur. J. Med. Chem. **2015**, 89, 189-97.

1.1.9 Dobeš, P.; Fanfrlík, J.; Řezáč, J.; Otyepka, M.; Hobza, P. Transferable scoring function based on semiempirical quantum mechanical PM6-DH2 method: CDK2 with 15 structurally diverse inhibitors. J. Comput. Aided Mol. Des., **2011**, 25, 223

1.1.10 Brahmkshatriya, P. S.; Dobeš, P.; Fanfrlík, J.; Řezáč, J.; Paruch, K.; Bronowska, A. K.; Lepšík, M.; Hobza, P. Quantum Mechanical Scoring: Structural and Energetic Insights into Cyclin-dependent Kinase 2 Inhibition by Pyrazolo[1,5-a]pyrimidines. Curr. Comput. Aided Drug Des. **2013**, 9, 118–129.

1.2. Polyhedral Boron Hydrides

The noncovalent interactions of heteroboranes with aromatic systems have only recently been acknowledged as a source of stabilization in supramolecular complexes.[3.1] We have studied the physical basis of these interactions has been studied in several model complexes using advanced computational methods.[3.2] The highly accurate CCSD(T)/complete basis set (CBS) value of the interaction energy for the model diborane···benzene complex in a stacking geometry exhibiting a B2H···π hydrogen bond was calculated to be −4.0 kcal·mol−1. The DFT-SAPT/CBS approach, which is shown to reproduce the CCSD(T)/CBS data reliably asserted that the major stabilizing component was dispersion, followed by electrostatics. Furthermore, the effect of the benzene heteroatom- and exosubstitutions was studied and found to be small. Next, when aromatic molecules were changed to cyclic aliphatic ones, van der Waals complexes stabilized by the dispersion term only were formed. As the last step, interactions of two larger icosahedral borane cages with benzene were explored. The complex of the monoanionic CB11H12− exhibited two minima: the first stacked above the plane of the benzene ring with a C−H···π hydrogen bond and the second planar, in which the carborane cage bound to benzene via five B−H···H−C dihydrogen bonds. The DFT-SAPT/CBS calculations revealed that both of these binding motifs were stabilized by dispersion followed by electrostatic terms, with the planar complex being 1.4 kcal·mol−1 more stable than the stacked one. The dianionic B12H122− interacted with benzene only in the planar geometry, similarly as smaller anions do. The large stabilization energy of 11.0 kcal·mol−1 was composed of dominant attractive dispersion and slightly

smaller electrostatic and induction terms. In summary, the borane/carborane···aromatic interaction is varied both in the complex geometries and in the stabilizing energy components. The detailed insight derived from high-level quantum chemical computations can help us understand such important processes as host−guest complexation or carborane···biomolecule interactions.[3.2]
In order to study carborane...biomolecule interactions, we analyzed the crystal structures of two novel carborane-sulfamide inhibitors in the complex with human carbonic anhydrase II (hCAII) using QM/MM calculations.[3.3] Even though both complexes possess the strongly interacting sulfamide···zinc ion motif, the calculations have revealed the different nature of binding of the carborane parts of the inhibitors. The neutral closo-carborane cage was bound to hCAII mainly via dispersion interactions and formed only very weak dihydrogen bonds. On the contrary, the monoanionic nido cage interacted with the protein mainly via electrostatic interactions. It formed short and strong dihydrogen bonds (stabilization of up to 4.2 kcal/mol; H···H distances of 1.7 Å) with the polar hydrogen of protein NH2 groups. This type of binding is unique among all of the classical organic and inorganic inhibitors of hCAII. Virtual glycine scanning allowed us to identify the amino-acid side chains, which made important contributions to ligand-binding energies. In summary, using QM/MM calculations, we have provided a detailed understanding of the differences between the interactions of two carborane sulfamides, identified the amino acids of hCAII with which they interact, and thus paved the way for the computer-aided rational design of selective boron-cluster-containing hCAII inhibitors.[3.3]
The chalcogen bond is a nonclassical σ-hole-based noncovalent interaction with emerging applications in medicinal chemistry and material science. It is found in organic compounds, including 2D aromatics, but has so far never been observed in 3D aromatic inorganic boron hydrides. Thiaboranes, harboring a sulfur heteroatom in the icosahedral cage, are candidates for the formation of chalcogen bonds. The phenyl-substituted thiaborane were synthesized and crystalized. X-ray crystal structure showed that they form sulfur-σ-hole … π type chalcogen bonds.[3.4] Quantum chemical analysis revealed that these interactions are considerably stronger than both in their organic counterparts and in the known halogen bond.[3.4] The reason is the existence of a highly positive σ-hole on the positively charged sulfur atom. This discovery expands the possibilities of applying substituted boron clusters in crystal engineering and drug design.

References

1.2.1 Řezáčová, P., Cígler, P., Matějíček, P., *Lepšík*, *M.*, Pokorná, J., Gruner, B., Konvalinka, J. (2011) Chapter I.3, Medicinal Application of Carboranes: Inhibition of HIV Protease. /In/: Boron Science- New Technologies and Applications/. /N. S. Hosmane (ed). New York, ISBN-978-1-4398266-3-8: CRC Press, pp. 41-70.

1.2.2 Sedlák R., Fanfrlik J., Hnyk D., Hobza P., Lepšík M. Interactions of Boranes and Carboranes with Aromatic Systems: CCSD(T) Complete Basis Set Calculations and DFT-SAPT Analysis of Energy Components. J. Phys. Chem. A **2010**, 114, 11304-11311.

1.2.3 Pecina, A.; Řezáč, J.; Brynda, J.; Mader, P.; Řezáčová, P.; Lepšík, M.; Hobza, P.; Fanfrlík, J. QM/MM Calculations Reveal Different Nature of Interaction of Two Carborane-based Sulfamide Inhibitors of Human Carbonic Anhydrase II. J. Phys. Chem. B **2013**, 117, 16096-16104.

1.2.4 Fanfrlik, J.; Přáda, A.; Padělková, Z.; Pecina, A.; Macháček, J.; Lepšik, M.; Holub, J.; Růžička, A.; Hnyk, D.; Hobza P. Dominant Role of Chalcogen Bonding in 2D and 3D Aromatics Revealed by Crystal Packing and Quantum Chemical Calculations. Angew. Chem. Int. Ed., **2014**, 53, 10139–10142.

1.3. Halogen Bonding

The structures and intermolecular interactions in the halogen bonded complexes of various anaesthetics with formaldehyde were studied by MP2 and CCSD(T) methods. In all complexes the C-halogen bond is shortened, in comparison with free compound, and an increase of the respective stretching frequency (blue shift) is observed. The resulting contraction of the C-halogen bond length was explained in the terms of Pauli repulsion. [4.1]
The desdcription of halogen bonding by standard molecular mechanics (MM) has been poor, owing to the lack of the so-called sigma hole localizeed at the halogen. This was improved by by modelling the sigma hole by a massless point charge attached to tye halogen. The position and charge of this particle was determined on the basis of CCSD(T) calculations. [4.2] The modified MM was successfully used in molecular docking to reproduce protein - ligand experimental geometries. [4.3]
O-H...X and O-H...O hydrogen bonds as well as C-X..X dihalogen and C-X...O halogen bonds have been investigated in halomethanol dimers. Halogen and dihalogen bonds, being of comparable strength, are weaker than both hzdrpogen bonds, but are still significant. [4.4]

Origin of the sigma-hole was investigated using the SAPT perturbation technique. The most important charcteristics of halogen bonds are their mixed dispersion - electrostatic nature, their high degree of directionalitz and their tunability. [4.5]
Halogen bonds in crystal TTF derivatives as well as in iodine molecule complexes were investigated by advanced quantum mechanical methods. [4.6], [4.7] Stabilisation energies of halogen bonds investigated were large and exceed 20 kcal/mol. The dominant contribution originates in dispersion, charge-transfer and electrostatic energies.
To contribute to the understanding of noncovalent binding of halogenated molecules with biological activity, electrostatic potential (ESP) maps of more than 2500 compounds were thoroughly analysed. [4.8] A peculiar region of positive ESP, called the sigma-hole, is a concept of central importance for halogen bonding. The results are in fair agreement with crystallographic surveys of small molecules as well as biomolecular complexes and may help to improve the intuition of chemists when dealing with halogenated compounds.
The sigma-holes of halogen atoms on various aromatic scaffolds were described in terms of their size and magnitude. Both the size and magnitude of the sigma-hole increase when passing from chlorinated to iodinated analogues. Also, the sigma-hole properties were studied upon chemical substitution of the aromatic ring as well as in the aromatic ring. The strength of the halogen bond between halogenbenzenes and Ar atoms and HF molecules increases while its directionality decreases when passing from chlorine to iodine. The decrease of the directionality of the halogen bond is larger for a HF-containing complex and is caused by electrostatic and exchange-repulsion energies. These findings are especially valuable for protein-halogenated ligand-binding studies, applied in the realm of rational drug development and lead optimisation. [4.9]
In order to study halogen bonding in biomolecular complexes, we designed series of aldose reductase (AR) inhibitors,[2.7] which was derived from a known AR binder, which had previously been shown to form a halogen bond between its Br atom and the O atom of the Thr-113 side chain of AR. In the series, the strength of the halogen bond was modulated by two factors, namely Br–I substitution and the fluorination of the aromatic ring in several positions. The role of the single halogen bond in AR–ligand binding was elucidated by advanced binding free energy calculations involving the semiempirical quantum chemical Hamiltonian. The results were complemented with ultrahigh-resolution X-ray crystallography and IC50 measurements. All of the AR inhibitors studied were shown by X-ray crystallography to bind in an identical manner. Further, it was demonstrated that it was possible to decrease the IC50 value by about 1 order of magnitude by tuning the strength of the halogen bond by a monoatomic substitution. The calculations revealed **that** the protein–ligand interaction energy increased upon the substitution of I for Br or upon the addition of electron- withdrawing F atoms to the ring. However, the effect on the binding affinity was found to be more complex due to the change of the solvation/desolvation properties within the ligand series. We have shown that it is possible to modulate the strength of a halogen bond in a protein–ligand complex as was designed based on the previous studies of low-molecular-weight complexes.[2.7]

References

1.3.1 Zierkiewicz, W.; Wieczorek, R.; Hobza, P. Michalska, D.; Halogen bonded complexes between volatile anaesthetics (chloroform, halothane, enflurane, isoflurane) and formaldehyde: a theoretical study. PCCP **2011**, 13, 5105.

1.3.2 Kolar, M.; Hobza, P. On Extension of the Current Biomolecular Empirical Force Field for the Description of Halogen Bonds. J. Chem. Theory Comput. **2012**, 8. 1325.

1.3.3 Kolar, M.; Hobza, P.; Bronowska A.K. Plugging the explicit sigma-holes in molecular docking. Chem. Commun. **2013**, 49, 981 - 983.

1.3.4 Riley, K. E.; Rezac, J.; Hobza, P. Competition between halogen, dihalogen and hydrogen bonds in bromo- and iodomethanol dimers. J. Mol. Model. **2013**, 19, 2879-2883.

1.3.5 Riley, K. E.; Hobza, P. The relative roles of electrostatics and dispersion in the stabilization of halogen bonds. Phys. Chem. Chem. Phys. 2013, 15, 17742-17751.

1.3.6 Deepa, P.; Sedlak, R.; Hobza, P. Times Cited: On the origin of the substantial stabilisation of the electron-donor 1,3-dithiole-2-thione-4-carboxyclic acid ... I2 and DABCO ...I2 complexes. Phys. Chem. Chem. Phys. **2014**, 16, 6679-6686.

1.3.7 Deepa, P.; Pandiyan, B. V.; Kolandaivel, P.;Hobza P. Halogen bonds in crystal TTF derivatives: an ab initio quantum mechanical study. Phys. Chem. Chem. Phys. **2014**, 16,2038 - 2047.

1.3.8 Kolar, M.H.; Carloni, P.; Hobza, P. Statistical analysis of sigma-holes: a novel complementary view on halogen bonding. Phys. Chem. Chem. Phys. **2014**, 16, 19111-19114.

1.3.9 Kolar, M.; Hostas, J.; Hobza, P. The strength and directionality of a halogen bond are co-determined by the magnitude and size of the sigma-hole. Phys. Chem. Chem. Phys. **2014**, 16, 9987-9996.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the ASCR, v. v. i.

Scientific team	Pavel Jungwirth - Computational Chemistry

1. Overview

The theoretical and computational investigations of the group have focused in the period of 2010-2014 on the influence of the solvent (in particular water and dissolved ions, but also osmolytes and electrons) on the properties of biomolecules, in close contact to experiments performed either within the group or by collaborators in the Institute or elsewhere. Hydration of ions and specific interactions of ions with biomolecules play a key role in many natural and biotechnological processes. It is our aim to obtain fundamental molecular understanding of these interactions, which would allow for a rational design of practical applications such as controlling protein association, enzymatic activity, membrane structure, etc. In particular, we have put a lot of effort to the understanding of the lyotropic or Hofmeister series which orders cations and anions according to their ability to salt-out proteins.

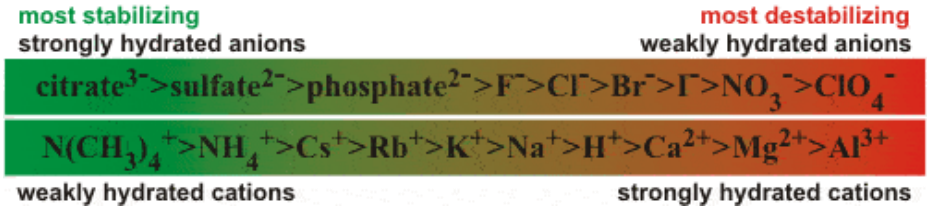


Figure 1: The lyotropic (Hofmeister) series of common cations and anions (adopted from www.lsbu.ac.uk/water/Hofmeister.html).

Since its formulation in 1888, this series has been invoked in a whole plethora of effects, ranging from salting-out, salting-in, and crystallization of proteins, over enzyme activities and swelling of tissues, to salt solubilities, ion exchange, surface tension of electrolytes, and bubble coalescence.[1,2] Although it has been clear that the Hofmeister series was intimately connected with ion hydration in homogeneous and heterogeneous environments and with ion pairing, its molecular origin has not been fully understood.[3] This situation could have been summarized as follows: Biochemists and physical chemists used the Hofmeister series as a mantra to put a label on ion specific behavior in various environments, rather than to reach a molecular level understanding and, consequently, an ability to predict a particular biophysical or biochemical effect of a given salt ion.

Dissolving salt in water is the usual way how to get charged particles into the solution. Another option is photoionization which leads to a creation of an electron and a radical cation. In the biological context, ionizing radiation is involved in direct and indirect DNA damage, which is utilized, e.g., in cancer therapy.[4] Within the study of direct DNA damage, we have performed an extended computational study, complemented by liquid microjet photoelectron spectroscopy, of ionization potentials of aqueous nucleic acid components, i.e., bases, nucleosides, and nucleotides. For indirect damage, the most important process is photoionization of water, which we have followed investigating both the nascent water radical cation and the forming solvated electron. For the former, we have unraveled the reaction mechanism leading to the formation of hydroxyl radical as a key DNA oxidizing agent. For the latter, we have provided the formation mechanism and a new structural model for the solvated electron.

In our recent work, we have thus primarily focused on specific ionic interactions with peptides and proteins in aqueous solutions and with lipid membranes, as well as on processes

leading to direct or indirect radiative DNA damage. Based on carefully designed and experimentally verified molecular dynamics simulations we were able to formulating general rules for interactions of Hofmeister ions with solvated proteins and membranes, opening also a way to go "beyond Hofmeister",[5] i.e., to future systematic rationalization of effects of ions on biological functions. (*For sake of saving space, from here on we restrict references to our publications from the evaluated period 2010-2014.*)

2. Ion-protein interactions

Our studies, as well as work of other research groups, have shown that interactions of (particularly monovalent) ions with protein surfaces in water have a rather local character which can be to the first approximation described in terms of ion pairing with charged side chain groups at protein surface and segregation at the backbone surface patches of the protein, with the rest of the protein surface playing only a minor role.[5] Due to this locality a reductionist approach can be successfully applied[6] and the basic physics of interactions of monovalent ions with proteins can be learned from ion interactions with small peptides, individual amino acids or even relevant functional groups (represented, e.g., by carboxylate anions and ammonium or guanidinium cations). Therefore, in addition to simulating whole proteins in solution, we also modeled the energetics and dynamics of ion-pairing between salt ions and model peptides, amino acids, as well as the ions representing the charged side chains. In terms of energetics we aimed at relative ordering salt ions according to the strength of their interaction with the oppositely charged amino acids. This also allowed us to critically test the empirical and qualitative Law of matching water affinities stating that small (large) cations or cationic groups tend to pair with small (large) anions or anionic groups.[5]

2.1 Ion pairing in aqueous solutions

The first thing we needed to understand at the molecular level was the specificity of ion pairing in water, e.g., why does sodium pair better than potassium with carboxylate anions? The answer to this question is beyond purely academic interest since it may hold one of the clues to why the intracellular environment is rich in potassium but poor in sodium. We have, therefore, collected a whole set of ion-pairing data, based on molecular dynamics simulations and electronic structure calculations,[7] which fit well with experiment and also (in most cases) with the Law of matching water affinities.[5] This data eventually allowed us to unravel a quantitative model, rationalizing the Law of matching water affinities in terms of unfavorable effects of overlapping hydration shells in pairs formed by ions of unequal size (or charge density).[8] Recently, we also looked at divalent ions like magnesium (with Heather Allen at Ohio State and Doug Tobias at UC Irvine),[9] calcium,[10] or sulfate,[11] which are more difficult to describe within classical force fields.

In the course of our investigations of ion-pairing we also stumbled across an unexpected phenomenon of formation of a contact ion pair between two ions of the same charge, namely guanidiniums. As previously suggested, guanidinium cations indeed have a tendency to form "Coulomb-defying" contact pairs in water, the energetics of which we have quantified in terms of quadrupole, dispersion, and cavitation interactions.[12,13]

2.2 Ions at non-polar aqueous interfaces

On average, about 40% of protein surface is non-polar. Therefore, we also have to pay attention to ion behavior at these hydrophobic patches. We have shown that, despite textbook knowledge that ions are generally repelled from hydrophobic surfaces, some ions (particularly large and soft one) exhibit appreciable affinities for these regions.[14,15,16] The fact that anions like iodide or thiocyanate can indeed appear at aqueous surfaces has been confirmed experimentally and is now accepted in the community. Also, certain cations such as guanidinium can accommodate at the water surface at an orientation parallel to it.[17] At the same time, there has been a major controversy concerning the surface behavior of the two inherent water ions – hydronium and hydroxide, with most physical scientist agreeing that the former is weakly attracted and the latter weakly repelled from the surface, while most colloid scientists claim that hydroxide is an extremely strong surfactant. In collaboration with spectroscopists, we have dedicated a lot of attention to defending the former view and even suggested possible explanations for the negative surface charge observed in electrophoretic and titration measurements.[16,18] In the context of ion-protein interactions, our recent studies showed that unlike extended hydrophobic interfaces non-polar amino acid side chains at the protein surface do not exhibit any appreciable affinity for ions.[19] In addition, the strength of interaction of ions with polar (uncharged) side chains is comparable with that with the aqueous solvent, therefore, there is practically no preferential binding of ions to these groups.[5]

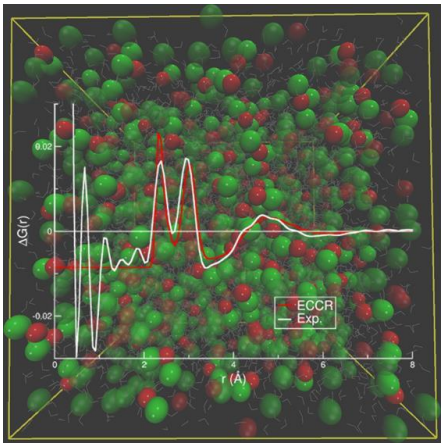


Figure 2: Structure of aqueous calcium chloride solution – perfect match between molecular simulations and neutron scattering (from Ref. 10).

2.3 Interactions between ions and amino acids or peptides

As a crucial step on the way toward rationalizing the Hofmeister series, we evaluated using molecular dynamics simulations the relative affinities of Hofmeister cations and anions for the protein backbone and charged side chains, which were then confirmed experimentally by our collaborator Paul Cremer at Pennsylvania State University. We showed that cations exhibit normal Hofmeister ordering both at the side chains of glutamate and aspartate and at the protein backbone with the latter interactions being, however, very weak.[20,21] In contrast, simulations and experiments show that anions follow a reversed Hofmeister ordering at charged side chains of lysine and arginine, while following normal ordering at the protein backbone.[19,22-24] We also investigated in detail the counter-intuitive like-charge pairing between arginine side chains and between guanidinium and arginine, which was subsequently confirmed by electrophoretic measurements.[25]

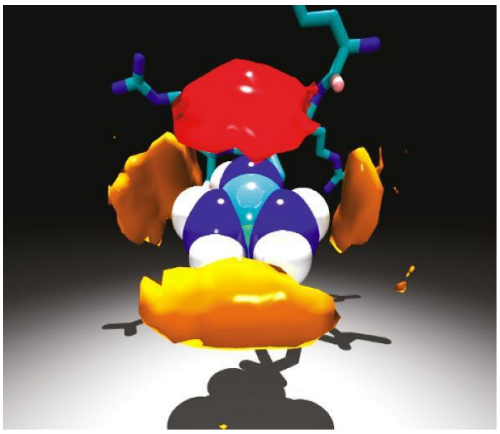


Figure 3: Pairing of arginine with guanidinum (red) and chloride (yellow). From Ref. 25.

Ion-protein interactions and salt effects on enzymatic activity

At this stage, the "holy grail" has been for us to understand how ions interact with proteins influencing their solubility (salting out), i.e, to provide a molecular interpretation of the Hofmeister series. Since investigations suggested that reductionism to a large extent works,[5] we took what we learnt above and generalized it to proteins. In a nutshell, the Hofmeister series can be thus rationalized as follows: i) the action of ions is primarily at the backbone and charged amino acid side chains exposed at the surface of an aqueous protein, ii) anions follow normal Hofmeister series at the backbone but reversed series at the side chains, interacting strongly with both, iii) cations follow only the normal series and their interaction with the backbone is weak, iv) as a result, anions are more interesting and their action can even lead for Hofmeister reversal (such as is the case for lysozyme), while cations are less important and interesting from the point of view of the Hofmeister series.[5]

As the next logical step, we started investigating effects of ions on protein/peptide stability. First, we used molecular dynamics to rationalize pH dependent interfacial behavior of a beta-amyloid observed by non-linear spectroscopy (experiments by Richard Saykally at University of California Berkeley).[26] In a subsequent study we unraveled by molecular simulations in combination with CD spectroscopy (performed by Chris Dempsey at University of Bristol) how specifically ions unfold (denature) small model peptides with different secondary motifs – a triptophane zipper and polyalanine.[27] We showed that while the hydrophobic core of a tryptophane zipper is destabilized both by "hydrophobic" tetrapropyl ammonium and "hydrophilic" guanidinium, only the latter is able to unwind alpha-helical structures. Moreover, we demonstrated that strongly pairing counterions (such as sulfate in the case of guanidinium) can significantly decrease the action of a protein-destabilizing ion. These systems allowed us to add to the detailed understanding of the denaturing effects of guanidinium and urea on the smallest self-folding protein, the tryptophane cage (with

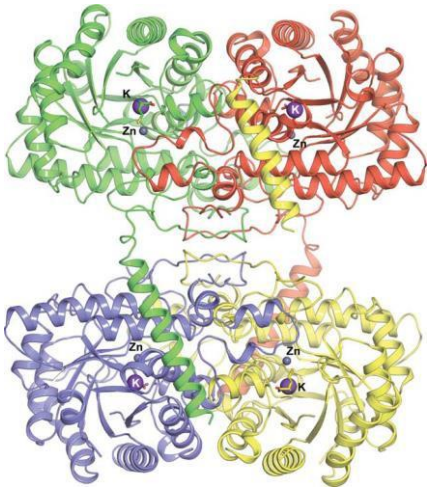


Figure 4: Cationic binding sites in the BHMT tetramer.[29]

calorimetry performed in the group of Jan Konvalinka, CD spectroscopy in the group of Lucie Beranova, bioinformatics by Jiří Vondrášek (all at the Institute), and NMR done by Gary Thompson at University of Leeds,).[28] Our results led us to the observation that while guanidinium and urea interact differently with the protein, the unfolding pathways are very similar to each other.

These studies gave us confidence to start looking also at the effect of ions on enzymatic activity. As a case study, we investigated by molecular dynamics the unusually strong positive effect potassium, compared to other alkali cations, on the activity of the betaine-homocysteine S-methyltransferase (BHMT) (the corresponding activity measurements performed in the group of Jiří Jiráček at the Institute and structural studies were done by Tim Garrow at Univ. of Illinois and Markos Koutmos in Bethesda).[29] Based on our results we were able to rationalize the enhanced enzymatic activity in the presence of potassium by its ability to stabilize the substrate in the active site in a specific binding arrangement.

3. Modeling of cellular membranes

Molecular dynamics simulations have opened the possibility to approach the description and control of the membrane biophysical state on an atomistic level. In particular, we have been focusing on the role of ions and oxidized phospholipids on the biophysical membrane parameters like lipid mobility, packing density, and degree of hydration. These parameters influence key properties of the membranes, such as the bilayer integrity and pore formation. Simulations have yielded a comprehensive picture of how alkali cations and halide anions, cationic cell penetrating peptides, as well as selected oxidized phospholipids with high physiological relevance, interact with the bilayer and control the properties of the membrane.

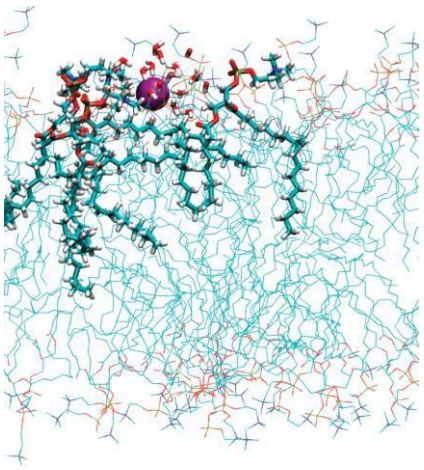


Figure 5*: A simulation snapshot of iodide dsorbed at a phospholipid membrane surface.*[30]

3.1 Interactions of ions with phospholipid bilayers

The structure of the lipid bilayer interior is relatively easy to determine experimentally, but the properties of the lipid-aqueous salt solution interface are harder to measure. This is an unfortunate situation, since many processes, ranging from protein folding and adsorption to the generation of second messengers from lipid precursors, take place in this region. It is also the headgroup part which is responsible for interactions with ions. The region is characterized by large gradients of all relevant physicochemical parameters, providing a complex environment, which can be modified by both alterations of lipid bilayer composition and/or properties of the aqueous phase. This picture is further complicated by the dynamic character of the membrane, the components of which are free to undergo a variety of motions due to thermal agitation as well as macro-scale membrane movements that affect the bilayer and adjacent water layers. We have established a direct combination of our molecular dynamics simulations with fluorescence spectroscopy measurements (performed in the group of Martin Hof at the Heyrovsky Institute) as a powerful tool for exploring the interactions of solution ions with the lipid bilayer.[30-32] On the computational side, we have first developed a reliable methodology and then applied it to realistic models of cellular membranes We have established the phosphate and carbonyl headgroup regions as binding sites for small cations and showed that sodium interacts more strongly with the headgroups than potassium. Ion specificity also concerns anions and we showed that soft ions like iodide can penetrate deep into the lipid bilayer.[30] The results of our simulations, confirmed by fluorescence spectroscopy measurements, allowed us to formulate basic rules for interactions of alkali cations and halide anions with phospholipid bilayers.[30,31] On top of it, we started investigations of interactions of short arginine-rich cationic peptides

with lipid bilayers with the future goal of elucidating the detailed mechanism allowing them to penetrate into cells.[33]

3.2 Effects of lipid oxidation on cellular membranes

Oxidized phospholipids are now known to be involved in several pathological conditions, such as atherosclerosis, inflammation, cancer, type 2 diabetes, and Alzheimer's disease. However, a coherent overall view of the causalities and mechanisms is lacking, mainly because of insufficient understanding of the cellular as well as molecular mechanisms. Within a Eurocores consortium, which represents an integrated interdisciplinary approach of European research laboratories, we have characterised the influence of lipid oxidation on structural parameters in model and living cell membranes by molecular simulations performed in our group in direct contact with fluorescence spectroscopy measurements (performed in Martin Hof's group). At the molecular level, oxidized phospholipids are truly unique. Firstly, oxidative modification introduces polar moieties to the hydrophobic parts of lipids, which changes dramatically their biophysical properties and, as a consequence, the properties of membranes containing them. Secondly, some of the oxidized phospholipids become chemically reactive, capable of covalent linkage to other biomolecules. The resulting structures are complex both chemically and from the point of view of their biophysical and biological properties.

We have used molecular dynamics to characterize both local and global changes in the phospholipid bilayers, as well as self-supported monolayers, caused by varying degrees of oxidation. For physiologically realistic oxidation levels of ~10% the changes are primarily local, leading to reorientation of the oxidized chains from the hydrophobic membrane interior toward the headgroup region and even into the aqueous phase, which has also implications on the lipid hydration and lateral mobility.[34-36] At higher oxidation levels, the global integrity of the bilayer becomes compromised and water pores across the membrane can form.[37] Most importantly, we have found (in collaboration with the experimental group of Paavo Kinnunen at Aalto University) that lipid oxidation reduces the barrier for transfer of phosphatidyl serine from the inner to the outer leaflet of the membrane (the so called lipid scrambling), which is a crucial step in the early stages of apoptosis.[38]

4. Direct and indirect DNA damage by ionizing radiation

For the study of direct radiation damage, as well as redox processes in DNA it is essential to know the ionization potentials of DNA and its components. These values have been known, both from theory and experiment for gas phase and microhydrated species, while the biologically relevant processes take place in aqueous solutions. With the advent of the liquid microjet photoelectron spectroscopy, it became possible to obtain the bulk liquid values experimentally. In parallel, we have established vertical ionization potentials of aqueous components of DNA by means of ab initio calculations employing a non-equilibrium polarizable continuum model. Turns out, however, that indirect processes triggered by ionization within the environment of DNA (water in particular) are more important for radiation damage in biology than the direct processes, simply because there is dramatically more water than DNA in our bodies. We have been, therefore, also investigating the basic physics and chemistry of radiation-induced damage in water. Photoionization of water leads to the formation of a cationic hole and an excess electron. Provided the ionizing radiation is well above the ionization threshold, the electron is "kicked" far enough from the cationic hole such that the two species can be treated separately. Within this spirit, we have followed by ab initio molecular dynamics simulations the ultrafast dynamics of both the cationic hole and excess electron created in water upon photoionization.

4.1 Ionization potentials of DNA components in water

In collaboration with the computational group of Petr Slavíček (Institute of Chemical Technology, Prague) and the experimental groups of Bernd Winter (BESSY Berlin) and Stephen Bradforth (University of Southern California) we have developed a computationally

robust and simple scheme for evaluation of vertical ionization potentials of neutral and ionic solutes in water.[39,40] This approach is based on a non-equilibrium polarizable continuum model of the solvent around the bare solute or, if necessary, a solute, which is first microsolvated by a small number of explicit water molecules. The ground state electronic structures before and after ionization are described using second order Moller-Plesset perturbation theory and the higher states after ionization are obtained via the time-dependent density functional theory. While our main focus has been on DNA components, the approach is quite general and it is possible to address also other solutes including amino acid functional groups and system interesting from the point of view of molecular electronics. Recently, we have evaluated photoelectron spectra of all aqueous nucleic acid bases, nucleosides, and nucleotides, in a very good agreement with photoelectron spectroscopy (wherever available).[39,40] We have also established the ionization energy of the most easily ionizable base (guanine) in the context of a short piece of DNA in water and showed that its value is only marginally influenced by the DNA context.[41] The main conclusion of these study is that the effect of the solvent is dramatic and cannot be recovered by merely adding one or several water molecules to the gas phase species. Namely, while in the gas phase the ionization energy of a base is strongly influenced by the ribose and by the phosphate, this influence all but disappears in bulk water. This observation points to the remarkable (and hitherto underappreciated) ability of the aqueous solvent to screen the electronic effects of the chemical environment of the nucleic acid base.

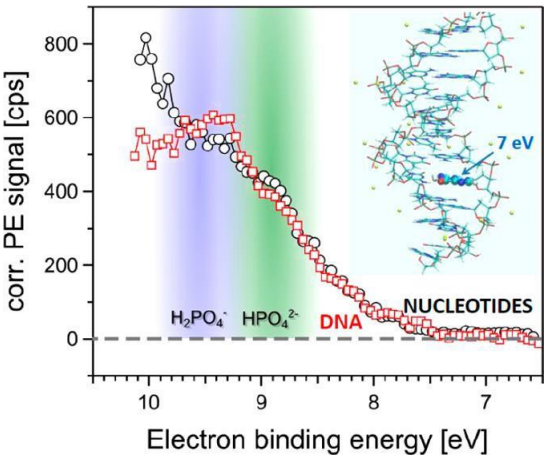


Figure 6: Almost identical ionization energies of a piece of DNA and a mixture of nucleotides in water (from Ref. 41).

4.2 Reactivity of the cationic hole in water

Photoionization in water by ionizing radiation can be formally written as $H_2O + h\nu \rightarrow H_2O^+ + e^-$. But what actually is the H_2O^+ species right after photoionization? Is it localized, partially delocalized, or a completely delocalized? And, if it is initially (at least partially) delocalized, then how long does it take to localize the cationic hole on a single water molecule? This is important to know, since only a localized hole reacts with a neighboring water molecule to form a hydronium cation and an OH radical, the latter being the key species involved in indirect radiative damage to DNA. By ab initio molecular dynamics simulations performed in our group, in tandem with ultrafast UV spectroscopy (performed in the group of Stephen Bradforth at University of Southern California), we were able to answer quantitatively all these questions.[42] We found that the radical cation in water is "born" upon photoionization as partially delocalized, however, it only takes few tens of femtoseconds for it to localize. Once it is localized on a single water molecule it almost immediately (within several femtoseconds) forms the OH radical. Spin localization thus turns out to be the rate limiting step of this key process in radiation chemistry of water. It is interesting to note that in the view of the extremely fast reactivity of the cationic hole in water, the OH radical (which is one of the most reactive species in chemistry) looks almost as a stable species. The extreme reactivity of the cationic hole in water is due to the fact that it is energetically unfavorable to localize charge and spin on the

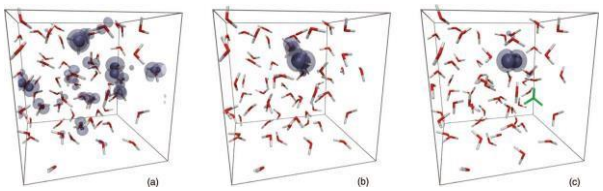


Figure 7: The spin localization and reactive steps following photoionization in bulk water (from Ref. 42).

same water molecule in the bulk liquid. Our work has thus unraveled the molecular details connected with the formation of the principal radical involved in the indirect DNA damage.

4.3 Structure, dynamics, and reactivity of the hydrated electron

Although low-energy electrons in water are only of secondary importance for DNA damage in biological tissues (not because they do not react with DNA but because under physiological conditions they react on a microsecond timescale before they can reach DNA), a lot of attention has been devoted to their characterization. This is partly due to the fact that hydrated electron is a very peculiar solvated charged species (a "fluffy" anion without a positive nucleus to keep it in shape) and partly due to the fact that its reactivity can lead to formation of hydrogen, which is a very unwanted species in any aqueous nuclear waste. In our work, we have performed ab initio molecular dynamics simulations of an excess electron in large water clusters, which allowed us to address the corresponding structural, dynamical, and reactivity issues. Our first studies (in collaboration with Burkhard Schmidt at Freie Universität Berlin) focused on the formation of a solvated electron by attachment of an electron to neutral water at ambient conditions.[43] We have shown, that at liquid conditions the electron always localizes to a well bound species within one or two picoseconds. We have also demonstrated,[44] that the more weakly bound isomers observed in cluster experiments can only exist at cryogenic conditions, which are pertinent to these measurements, but are not relevant at ambient conditions. Most recently, we moved with simulations directly to the aqueous bulk, which allowed us to characterize the structure[45-47] and the ultrafast dynamics (with terahertz experiments performed by Peter Hamm at University of Zurich)[48] of the excess electron in water, resolving existing controversies about its character ("cavity" vs. "diffuse" electron) both in the aqueous bulk and at the water-vapor intefrace. Using ab initio molecular dynamics we have shown that the hydrated electron has a complex structure with a dominant cavity, but also with important diffuse and radical anion contributions.

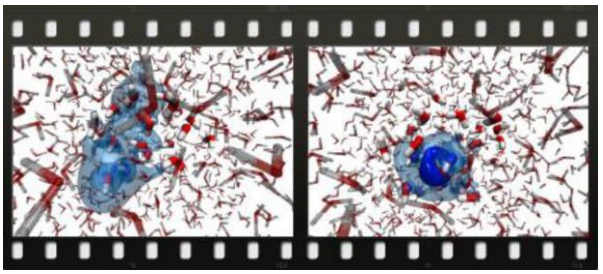


Figure 8: Two snapshots from a localization trajectory of an excess electron in water by ab initio molecular dynamics (from Ref. 48).

Ab initio molecular dynamics simulations in principle allow following the quenching reactions of hydrated electrons. We were able to model (in collaboration with Joost VandeVondele at ETH Zurich) the most basic and fundamental reaction of the hydrated electron, namely that with a proton, leading to the formation of a hydrogen atom. In the gas phase, this is the simplest imaginable textbook chemical process, leading to a release of 13.6 eV. In water, however, things are much more complicated, since water is not only a polar solvent which stabilizes the charged reactants but it is even directly chemically involved by binding to the proton. As a result, the reaction in water is only mildly exothermic and has a small barrier. Consequently, as earlier flash photolysis studies showed, the electron-proton reaction in water is about an order of magnitude slower than a diffusion limited process and, moreover, proceeds as proton transfer rather than electron transfer. Our simulations have provided the detailed molecular mechanism for this reaction, in which the proton induces an asymmetry in the shape of the solvated electron, which helps to overcome the desolvation barrier. The proton then hops along a water chain to the center of the electron spin density eventually forming a neutral (and, therefore, poorly solvated) hydrogen atom.[49,50] Understanding the molecular details of this reaction is crucial, e.g., for developing efficient strategies to quench it (and thus prevent potential hydrogen gas explosions) during nuclear waste storage.

5. Summary

Within the evaluation period we have completed rebuilding the group around biophysical/ biochemical topics involving ions, proteins, and membranes. In this way, we managed not to lose the previously acquired know-how about ions in aqueous solutions, but instead creatively applied it in new contexts of ion-protein interactions and dynamics of water around biologically relevant ions. At the same time, we have extended the scope of research to unravelling basic processes involving photoionization of water and formation of solvated electrons, as well as photoionization of aqueous DNA and its components in the context of radiation damage to DNA during radiation cancer therapy. As a result of the refocus on ion-protein interactions, we have recently succeeded to provide a molecular-level rationalization underlying the 125 years old concept of the Hofmeister series of ions. These findings at the same time pave the way "beyond Hofmeister" toward unraveling molecular mechanisms behind biological functions of the same ions.[5]

One of the greatest joys on this scientific journey has been building and maintaining a vigorously active, scientifically thrilling, and personally pleasant group of about 15-20 members including students (both post- and pre-graduate level), postdocs, and senior scientists. In order to be able to perform cutting edge research we strive to maintain within the group a broad expertise focused around molecular dynamics simulation techniques, but at the same time extending on one side to electronic structure calculations and, on the other side, to coarse graining approaches all the way to linear electric circuits. Such a broad research scope is achievable by having a relatively large and very multidisciplinary group - we hire students and postdocs from chemistry, physics, biology, as well as applied mathematics.

We also purposefully aims at blurring the boundary between simulations and experiment. The way to achieve this is twofold: group members are encouraged to "get their hands dirty" and perform at least some of the experiments themselves. In some cases, such as the neutron scattering experiments on structures of salt solutions we are actually able to perform most of the experimental work ourselves (with qualified local help at the ILL in Grenoble by Dr. Henry Fisher). In most situations we, however, rely to a large extent on our experimental collaborators. Nevertheless, group members very often actively participate on the measurements, e.g., in the fluorescence spectroscopy labs of Prof. Martin Hof and Dr. Josef Lazar, and the biochemistry lab of Dr. Jiří Jiráček (all at the Academy of Sciences in Prague), or the spectroscopy and optical labs of Prof. Stephen Bradforth (USC Los Angeles), Dr. Bernd Winter (BESSY Berlin), and Prof. Sigurd Bauerecker (University of Braunschweig). Only rarely do we fully "outsource" the experiment, as in the case of the spectroscopy and microfluidics lab of Prof. Paul Cremer (Penn State) or ultrafast spectroscopy lab of Prof. Peter Hamm (University of Zurich).

References

(1) Hofmeister, F. *Arch. Exp. Pathol. Pharmakol. (Leipzig)* **1888**, *24*, 247.

(2) Kunz, W.; Lo Nostro, P.; Ninham, B. W. *Current Opinion Colloid & Interface Sci.* **2004**, *9*, 1.

(3) Zhang, Y. J.; Furyk, S.; Bergbreiter, D. E.; Cremer, P. S. *J. Am. Chem. Soc.* **2005**, *127*, 14505.

(4) von Sonntag, C. *Free-Radical-Induced DNA Damage and Its Repair: A Chemical Perspective*; Springer: Berlin, 2006.

(5) Cremer, P. S.; Jungwirth, P.: *Nature Chem.* **2014**, *6*, 261.

(6) Lund, M.; Heyda, J.; Jungwirth, P.. in: Specific ion effects, ed. Kunz, W.; World Scientific, 2010, 217.

(7) Pluharova, E.; Marsalek, O.; Schmidt, B.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2013**, *4*, 4177.

(8) Lund, M.; Jagoda-Cwiklik, B.; Woodward, C. E.; Vacha, R.; Jungwirth, P. *J. Phys. Chem. Lett.* **2010**, *1*, 300.

(9) Casillas-Ituarte, N. N.; Callahan, K. M.; Tang, C. Y.; Chen, X. K.; Roeselova, M.; Tobias, D. J.; Allen, H. C. *Proc. Nat. Acad. Sci. USA* **2010**, *107*, 6616.
(10) Kohagen, M.; Lepsik, M.; Jungwirth, P.: *J. Phys. Chem. B* **2014**, *118*, 7902.
(11) Pegado, L.; Marsalek, O.; Jungwirth, P.; Wernersson, E.: *Phys. Chem. Chem. Phys.* **2012**, *14*, 10248.
(12) Vazdar, M.; Vymetal, J.; Heyda, J.; Vondrasek, J.; Jungwirth, P. *J. Phys. Chem. A* **2011**, *115*, 11193.
(13) Vazdar, M.; Uhlig, F.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2012**, *3*, 2021.
(14) Ottosson, N.; Heyda, J.; Wernersson, E.; Pokapanich, W.; Svensson, S.; Winter, B.; Ohrwall, G.; Jungwirth, P.; Bjorneholm, O.: *Phys. Chem. Chem. Phys.* **2010**, *12*, 10693.
(15) Vazdar, M.; Pluharova, E.; Mason, P. E.; Vacha, R.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2012**, *3*, 2087.
(16) Vacha, R.; Uhlig, F.; Jungwirth, P.: *Adv. Chem. Phys.* **2014**, *155*, 69.
(17) Wernersson, E.; Heyda, J.; Vazdar, M.; Lund, M.; Mason, P. E.; Jungwirth, P. *J. Phys. Chem. B* **2011**, *115*, 12521.
(18) Vacha, R.; Rick, S. W.; Jungwirth, P.; de Beer, A. G. F.; de Aguiar, H. B.; Samson, J. S.; Roke, S. *J. Am. Chem. Soc.* **2011**, *133*, 10204.
(19) Rembert, K.; Paterova, J.; Heyda, J.; Hilty, C.; Jungwirth, P.; Cremer, P. S.: *J. Am. Chem. Soc.* **2012**, *134*, 10039.
(20) Heyda, J.; Vincent, J. C.; Tobias, D. J.; Dzubiella, J.; Jungwirth, P. *J. Phys. Chem. B* **2010**, *114*, 1213.
(21) Pluharova, E.; Baer, M. D.; Mundy, C. J.; Schmidt, B.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2014**, *5*, 2235.
(22) Mason, P. E.; Heyda, J.; Fischer, H. E.; Jungwirth, P. *J. Phys. Chem. B* **2010**, *114*, 13853.
(23) Paterova, J.; Rembert, K.; Heyda, J.; Kurra, Y.; Okur, H. I.; Liu, W. R.; Hilty, C.; Cremer, P. S.; Jungwirth, P.: *J. Phys. Chem. B* **2013**, *117*, 8150.
(24) Hladilkova, J.; Heyda, J.; Rembert, K. B.; Okur, H. I.; Kurra, Y.;; Liu, W. R.; Hilty, C.; Cremer, P. S.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2013**, *4*, 4069.
(25) Kubickova, A.; Krizek, T.; Coufal, P.; Wernersson, E.; Heyda, J.; Jungwirth, P. *J. Phys. Chem. Lett.* **2011**, *2*, 1387.
(26) Miller, A. E.; Petersen, P. B.; Hollars, C. W.; Saykally, R. J.; Heyda, J.; Jungwirth, P. *J. Phys. Chem. A* **2011**, *115*, 5873.
(27) Dempsey, C. E.; Mason, P. E.; Jungwirth, P. *J. Am. Chem. Soc.* **2011**, *133*, 7300.
(28) Heyda, J.; Kozisek, M.; Bednarova, L.; Thompson, G.; Konvalinka, J.; Vondrasek, J.; Jungwirth, P. *J. Phys. Chem. B* **2011**, *115*, 8910.
(29) Mladkova, J.; Hladilkova, J.; Diamond, C. E.; Tryon, K.; Yamada, K.; Garrow, T. A.; Jungwirth, P.; Koutmos, M.; Jiracek, J.: *Proteins: Structure, Function, Bioinformatics* **2014**, *82*, 2552.
(30) Vacha, R.; Jurkiewicz, P.; Petrov, M.; Berkowitz, M. L.; Bockmann, R. A.; Barucha-Kraszewska, J.; Hof, M.; Jungwirth, P. *J. Phys. Chem. B* **2010**, *114*, 9504.
(31) Jurkiewicz, P.; Cwiklik, L.; Vojtiskova, A.; Jungwirth, P.; Hof., M.; *Biochim. Biophys. Acta – Biomembranes* **2012**, *1818*, 609.
(32) Pokorna, S.; Jurkiewicz, P.; Vazdar, N.; Cwiklik, L.; Jungwirth, P.; Hof, M.: *J. Chem. Phys.* **2014**, *141*, 22D512.
(33) Vazdar, M.; Wernersson, E.; Khabiri, M.; Cwiklik, L.; Jurkiewicz, P.; Hof, M.; Kolusheva, S.; Jelinek, R.; Jungwirth, P.: *J. Phys. Chem. B* **2013**, *117*, 11530.
(34) Beranova, L.; Cwiklik, L.; Jurkiewicz, P.; Hof, M.; Jungwirth, P. *Langmuir* **2010**, *26*, 6140.
(35) Jurkiewicz, P.; Olzynska, A.; Cwiklik, L.; Conte, E.; Jungwirth, P.; Megli, F. M.; Hof., M.: *Biochim. Biophys. Acta – Biomembranes* **2012**, *1818*, 2388.

(36) Vazdar, M.; Jurkiewicz, P.; Hof, M.; Jungwirth, P.; Cwiklik, L.: *J. Phys. Chem. B* **2012**, *116*, 6411.
(37) Lis, M.; Wizert, A.; Przybylo, M.; Langner, M.; Swiatek, J.; Jungwirth, P.; Cwiklik, L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17555.
(38) Volinsky, R.; Cwiklik, L.; Jurkiewicz, P.; Hof, M.; Jungwirth, P.; Kinnunen, P. K. J. *Biophys. J.* **2011**, *101*, 1376.
(39) Pluharova, E.; Jungwirth, P.; Bradforth, S. E.; Slavicek, P. *J. Phys. Chem. B* **2011**, *115*, 1294.
(40) Pluharova, E.; Oncak, M.; Seidel, R.; Schroeder, C.; Schroeder, W.; Winter, B.; Bradforth, S. E.; Jungwirth, P.; Slavicek, P.: *J. Phys. Chem. B* **2012**, *116*, 13254.
(41) Pluharova, E.; Schroeder, C.; Seidel, R.; Bradforth, S. E.; Winter, B.; Faubel, M.; Slavicek, P.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2013**, *4*, 3766.
(42) Marsalek, O.; Elles, C. G.; Pieniazek, P. A.; Pluharova, E.; VandeVondele, J.; Bradforth, S. E.; Jungwirth, P. *J. Chem. Phys.* **2011**, *135*.
(43) Marsalek, O.; Uhlig, F.; Frigato, T.; Schmidt, B.; Jungwirth, P. *Phys. Rev. Lett.* **2010**, *105*, 043002.
(44) Marsalek, O.; Uhlig, F.; P., J. *J. Phys. Chem. C* **2010**, *114*, 20489.
(45) Uhlig, F.; Marsalek, O.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2012**, *3*, 3071.
(46) Marsalek, O.; Uhlig, F.; VandeVondele, J.; Jungwirth, P.: *Acc. Chem. Res.* **2012**, *45*, 23.
(47) Uhlig, F.; Marsalek, O.; Jungwirth, P.: *J. Phys. Chem. Lett.* **2013**, *4*, 338.
(48) Savolainen, J.; Uhlig, F.; Ahmed, S.; Hamm, P.; Jungwirth, P.: *Nature Chem.* **2014**, *6*, 697.
(49) Marsalek, O.; Frigato, T.; VandeVondele, J.; Bradforth, S. E.; Schmidt, B.; Schutte, C.; Jungwirth, P. *J. Phys. Chem. B* **2010**, *114*, 915.
(50) Uhlig, F.; Jungwirth, P.: *Zeit. Phys. Chemie* **2013**, *227*, 1583.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Lubomír Rulíšek - Theoretical Bioinorganic Chemistry

The core of the research carried out in the Rulíšek's group in the title period (2010-2014) falls into the area of **theoretical bioinorganic chemistry**. It included theoretical studies of reaction mechanisms of selected metalloproteins (mostly containing polynuclear metal sites in the active site): trinuclear copper site found in multi-copper oxidases (MCOs), dinuclear non-heme iron Δ^9-desaturase (Δ^9D), dinuclear zinc glutamate carboxypeptidase II (GCPII), and mononuclear manganese super-oxide dismutase (MnSOD). All of these systems represent challenging systems for quantum chemical methodology either due to the complexity of the electronic structure of the active site (MCOs, MnSOD, Δ^9D) or a conformational complexity of the substrate in presumably conceptually simpler dizinc GCPII. It is only a tight interplay between theory and experiment (X-ray, spectroscopic, kinetic, and thermodynamic data) that enables the formulation of consensus reaction mechanisms for these complicated systems (as documented in more details below). The general guidelines for this tight correlation between theory and experiment in bioinorganic chemistry were outlined in an invited perspective.[1]

[1] Rokob, T. A.; Srnec, M.; Rulíšek, L.: Theoretical Calculations of Physico-Chemical and Spectroscopic Properties of Bioinorganic Systems: Current Limits and Perspectives. *Dalton Trans.* **2012**, *41*, 5754-5768. *Invited Perspective.*

Related to theoretical bioinorganic chemistry is our research with the long-term vision of ***in silico* design of smaller artificial metalloenzymes** (catalytic metallopeptides). Currently, it is at the stage of the design of shorter peptide sequences selectively binding selected metal ions. Condition *sine qua non* for successful accomplishment of this project is the ability to quantitatively calculate the **complexion (free) energy changes associated with the binding of metal ions in biomolecules**.
Methodological issues originating from our calculations of thermodynamic, kinetic, and spectroscopic properties of these systems gave rise to **modest contributions to quantum chemical** and combined quantum mechanical/ molecular mechanical (**QM/MM**) **theory**. Last but not least, the accumulated experience in computational chemistry enabled us to participate in many projects related to **organic reactivity** (homogeneous catalysis), spectroscopy, and interactions of smaller ligands with biomolecules and finally, we made modest impact in the area of **computational electrochemistry** which encompasses theoretical calculations of reduction potentials (and pK_a values), ranging from small systems to metalloproteins.
These efforts resulted in approximately 40 scientific papers published by the scientists in the Rulíšek group over the period, mostly in top-tier specialized journals with several of them in the top-level chemical journals, such as *J. Am. Chem. Soc.* (2), *Angew. Chemie Int. Ed.* (1) and *Chem. Eur. J.* (6). The 22 contributions (with the first or corresponding authorship from the group) can be considered as the key contributions of the group and the most important findings reported therein will be mentioned in more details below.

2.1. Reaction Mechanisms of Metalloproteins.

2.1.1 Multi-copper oxidases (MCOs). The MCOs couple four one-electron oxidations of substrates at the mononuclear type 1 copper (Cu-T1) site with the four-electron reduction of dioxygen at the trinuclear copper cluster (TNC). The TNC consists of three copper ions arranged in a unique triangular fashion (Figure 1). In its oxidised form and in some experimentally observed intermediates (the peroxy and native intermediates), this leads to a magnetic coupling of the unpaired electrons of the three copper ions, resulting in unusual spectroscopic features (the so-called spin frustration). By correlating experimental and theoretical data, an unambiguous mapping between the structural, energetic and spectroscopic properties of the various intermediates in the MCO reaction cycle can be established. Our studies involved quantum mechanics (QM; density-functional theory and multi-reference self-

consistent field – e.g. CASSCF/CASPT2) calculations which resulted in the first application of multi-reference methods (CASSCF, CASPT2) for calculations of molecular *g*-tensors of bioinorganic polynuclear systems, such as the TNC (in collaboration with the groups of Prof. Frank Neese and Prof. Kristine Pierloot).[2] Correlation of the calculated *g*-values with the electronic paramagnetic resonance (EPR) experimental data provided a unique electronic structure/spectroscopy correlation which is, in turn, independent and convincing proof of the correct structural assignment of the MCO intermediates. Next, combined QM and molecular mechanics (QM/MM) modelling, ranging from standard QM/MM optimisations[3] to the QM/MM free-energy perturbations (QTCP) to address the O_2 reactivity in the TNC and phenomena such as the Cu-T1 → TNC electron transfer (reorganization energies) were carried out.[4,5] In the collaboration with the groups of Prof. Ulf Ryde (Lund University, Sweden) and Prof. Edward Solomon (Stanford University, U. S. A.) we were able to provide a unique set of consistent theoretical data (mostly by correlating the calculated data to spectroscopic results) that led to the consensus reaction mechanism of the dioxygen cleavage in the MCOs including the re-reduction of

---

[2] Vancoillie, S.; Chalupský, J.; Ryde, U.; Solomon, E. I.; Pierloot, K.; Neese, F.; Rulíšek, L.: Multireference *Ab Initio* Calculations of g tensors for Trinuclear Copper Clusters in Multicopper Oxidases. *J. Phys. Chem. B* **2010**, *114,* 7692-7702.
[3] Srnec, M.; Ryde, U.; Rulíšek, L.: Reductive Cleavage of the O–O Bond in Multicopper Oxidases: QM/MM and QM Study. *Faraday Discuss.* **2011**, *148*, 41-53.
[4] Shleev, S.; Andoralov, V.; Falk, M.; Reimann, C. T.; Ruzgas, T.; Srnec, M.; Ryde, U.; Rulíšek, L.: On the Possibility of Uphill Intramolecular Electron Transfer in Multicopper Oxidases: Electrochemical and Quantum Chemical Study of Bilirubin Oxidase. *Electroanalysis* **2012**, *24*, 1524-1540.
[5] Hu, L.-H.; Farrokhnia, M.; Heimdal, J.; Shleev, S.; Rulíšek, L.; Ryde, U.: Reorganisation Energy for Internal Electron Transfer in Multicopper Oxidases. *J. Phys. Chem. B* **2011**, *115*, 13111-13126.

the TNC after the dioxygen cleavage. Our achievements and the work of others concerning theoretical calculations of MCOs are summarized in the recent comprehensive review.[6]

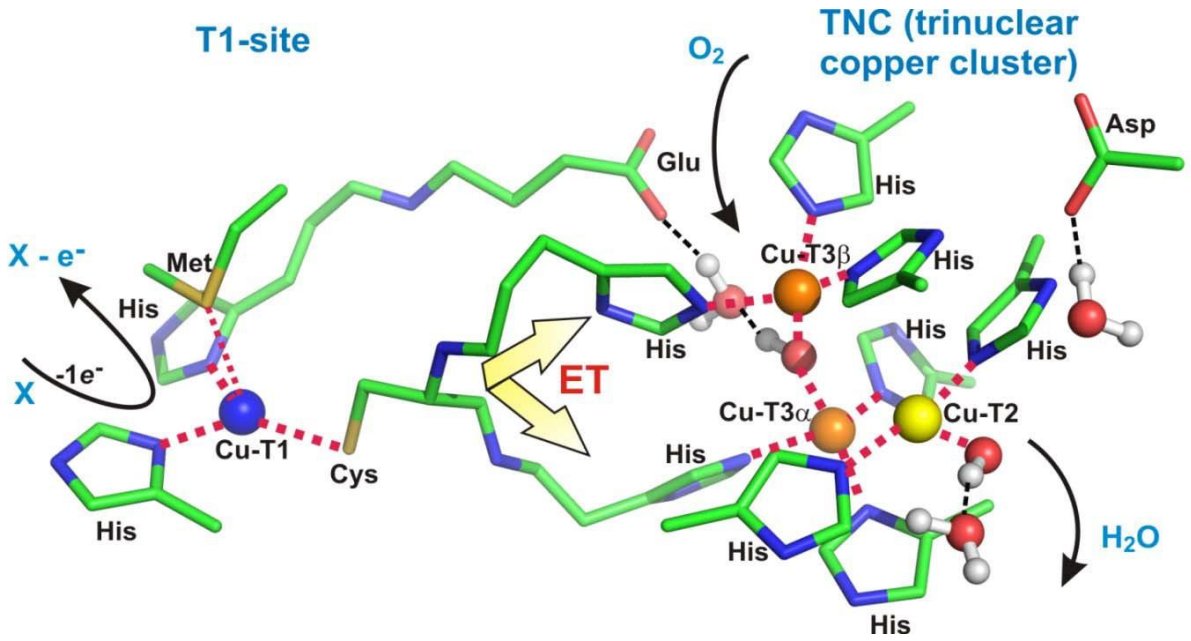


Figure 1: The general architecture of the trinuclear copper cluster site (the site of the four-electron $O_2 \rightarrow H_2O$ reduction) and of the Cu-T1 site (the site of one-electron oxidations of organic substrates or metal ions).

2.1.2. Non-heme diiron Δ^9 Desaturase. In the recent study[7] we published the computational data (complemented by a small amount of experimental data from the Solomon's lab –MCD spectra) on the most challenging system studied insofar in our laboratory - Δ^9 desaturase (Δ^9D). Δ^9D is a non-heme diiron enzyme capable of inserting a double bond into an alkyl chain by double hydrogen atom

abstraction using molecular O_2 as cofactor. The constitution of its active site by itself represents a formidable task for contemporary computational methods whereas the ultimate goal is to decipher the reaction mechanisms for this type of enzymes. It can be expected that along the reaction coordinate many reactive intermediates containing iron-oxo units are to be encountered, as well as radicals ensuing from the homolytic cleavage of the C-H bonds, and many other 'non-trivial' species. While the application of DFT methodology might be probably the only practical solution to the problem, its accuracy has not yet been fully grasped. It is difficult, if not impossible, to benchmark DFT methods even on smaller systems (as is normally the standard procedure), as the CASPT2 (RASPT2) or MRCI calculations would necessitate more active orbitals to be included in the calculations than is within the reach of current computational power (the so-called DMRG, mentioned below, might be the only alternative among wave function methods). Still, the correlation of several calculated (DFT) spectroscopic parameters (absorption, CD, vibrational, and Mössbauer) for multiple structural models of the experimentally defined peroxodifferic intermediate (**P**) with the experimental values enabled us to lock in the structures in the initial steps of the catalytic cycle. We suggested that protonation of the peroxide moiety, possibly preceded by water binding in the Fe_A coordination sphere, could be responsible for the conversion of the **P** intermediate in Δ^9D into a form capable of hydrogen

---

[6] Rulíšek, L.; Ryde, U.: Theoretical Studies of the Active-Site Structure, Spectroscopic, and Thermodynamic Properties, and Reaction Mechanism of Multicopper Oxidases. *Coord. Chem. Rev.* **2013**, *257*, 445-458. *Invited Review.*

[7] Srnec, M.; Rokob, T. A.; Schwartz, J. K.; Kwak, Y.; Rulíšek, L.*; Solomon, E. I.*: Structural and Spectroscopic Properties of the Peroxodiferric Intermediate of *Ricinus Communis* Soluble ^{9} Desaturase. *Inorg. Chem.* **2012**, *51,* 2806-2820.

abstraction. Last but not least, the results were compared with recent findings on the related ribonucleotide reductase and toluene/methane monooxygenase enzymes.

An emerging method in the domain of multireference wave function techniques is the density matrix renormalization group which scales only polynomially and not exponentially with respect to the size of the active space. Its application in the bioinorganic chemistry is a matter of debate since the number of studies is rather limited. Notably, there was no study reported on usage of this advanced technology for the bioinorganic reactivity. Therefore, the results of density matrix renormalization group complete active space self-consistent field (DMRG-CASSCF) and second-order perturbation theory (DMRG-CASPT2) calculations were presented on various structural alternatives for the O—O and first C—H activating step of the catalytic cycle the Δ^9D.[8] The reaction step – abstraction of the first hydrogen atom - studied therein is presumably associated with the highest activation barrier along the full pathway; therefore, its quantitative assessment is of key importance to the understanding of the catalysis. The DMRG approach allows unprecedentedly large active spaces for the explicit correlation of electrons in the large part of the chemically important valence space, which is apparently *conditio sine qua non* for obtaining well-converged reaction energetics. The derived reaction mechanism involves protonation of the previously characterized 1,2-μ peroxy $Fe^{III}Fe^{III}$ (**P**) intermediate to a 1,1-μ hydroperoxy species, which abstracts an H atom from the C_{10} site of the substrate. An Fe^{IV}-oxo unit is generated concomitantly, supposedly capable of the second H atom abstraction from C_9. In addition, several popular DFT functionals were compared to the computed DMRG-CASPT2 data. Notably, many of these show a preference for heterolytic C–H cleavage**,** erroneously predicting substrate hydroxylation. This study shows that, despite its limitations, DMRG-CASPT2 is a significant methodological advancement toward the accurate computational treatment of complex bioinorganic systems, such as those with the highly open-shell diiron active sites.

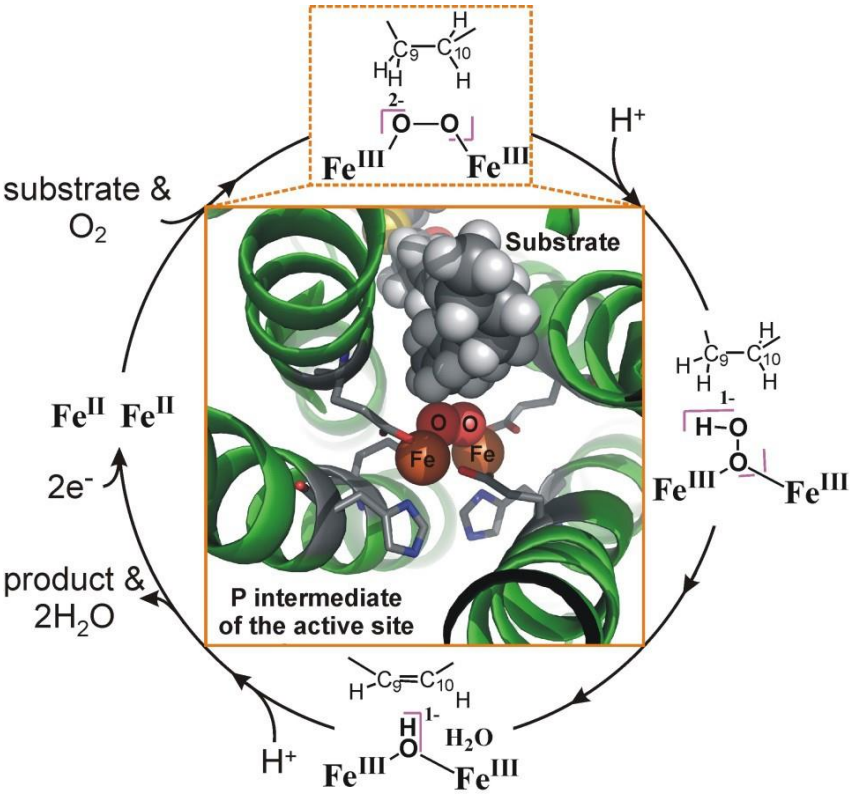


Figure 2: The postulated catalytic cycle of the Δ^9 desaturase. The formal oxidation states of key intermediates are depicted together with the structure of the active site of Δ^9

[8] Chalupský, J.; Rokob, T. A.; Kurashige, Y.; Yanai, T.; Solomon, E. I.; Rulíšek, L.; Srnec, M.: Reactivity of the Binuclear Non-Heme Iron Active Site of Δ^9 Desaturase Studied by Large-Scale Multireference *Ab Initio* Calculations. *J. Am. Chem. Soc.* **2014**, *136*, 15977-15991.

2.1.3. Glutamate carboxypeptidase II (GCPII). GCPII is an enzyme, which plays an important role in the regulation of levels of one of the neurotransmitters – *N*-Ac-Asp-Glu (NAAG) – in synapses. In collaboration with Jan Konvalinka's (IOCB) and Cyril Bařinka's groups, we recently proposed a detailed and consistent picture of the reaction mechanism of this highly interesting enzyme at the atomic level. This lead us to general considerations related to the theoretical aspects of the peptide bond hydrolysis, which included comparison of the „non-catalyzed" hydrolysis with the same reaction catalyzed by mono- and dizinc model systems, and also included calibration and benchmarking of various cheaper and practical quantum chemical methods (such as DFT) against the reference CCSD(T) data.[9] Last but not least, we provided a rationale for the interaction of novel substrate-based inhibitors of GCPII with enhanced lipophilicity.[10]

2.2. Theoretical Aspects of Metal Ion Complexation in Biomolecules and Design of Specific Metal Binding Sites.

As part of our ongoing efforts to design new catalytic metallopeptides and selective metal chelators, we addressed computationally the physico-chemical aspects of binding of metal ions in the biomolecular scaffolds. Specifically, we tried to contribute to the discussion related to the metal ion selectivity. Experimentally (and in reality) this quantity is described by the differences in the stability constant of a given ion in the particular metal-binding site (complex). In theory, one needs to use thermodynamic cycle to compute the stability constant by adding gas-phase complexation energies/enthalpies to solvation/desolvation energies of the species involved. We have attempted to calibrate quantum chemical methods for the calculations of the former and devise a robust computational protocol that would yield the satisfactorily accurate stability constants. The results of these studies are mentioned in more detail below.

2.2.1. On the Accuracy of the Calculated Free Energies Associated with Metal Ion Complexation. As mentioned above, to address fundamental questions in bioinorganic chemistry, such as metal ion selectivity, accurate computational protocols for both the gas-phase association of metal-ligand complexes and solvation/desolvation energies of the species involved are needed. We attempted to critically evaluate the performance of the *ab initio* and DFT electronic structure methods available and recent solvation models in calculations of the energetics associated with metal ion complexation. On the example of five model complexes ($[M^{II}(CH_3S)(H_2O)]^+$, $[M^{II}(H_2O)_2(H_2S)(NH_3)]^{2+}$, $[M^{II}(CH_3S)(NH_3)(H_2O)(CH_3COO)]$, $[M^{II}(H_2O)_3(SH)(CH_3COO)(Im)]$, $[M^{II}(H_2S)(H_2O)(CH_3COO)(PhOH)(Im)]^+$ in typical coordination geometries) and four metal ions (Fe^{2+}, Cu^{2+}, Zn^{2+}, and Cd^{2+}; representing open- and closed-shell and the first- and second-row transition metal elements), we provided reference values for the gas-phase complexation energies, as presumably obtained using the CCSD(T)/aug-cc-pVTZ method, and compared it with cheaper methods, such as DFT and RI-MP2, that can be used for large-scale calculations.[11] The computational data highlighted several intricacies in the theoretical predictions of the experimental stability constants: the covalent character of some metal-ligand bonds (e.g. Cu(II)-thiolate) causing larger errors in the gas-phase complexation energies, inaccuracies in the treatment of solvation of the charged species and difficulties in the definition of the reference state for Jahn-Teller unstable systems (e.g., $[Cu(H_2O)_6]^{2+}$).

2.2.2. Predicting Stability Constants of Metal Ion Complexes from the First Principles. As mentioned above, the most important experimental quantity describing the thermodynamics of metal ion binding with various (in)organic ligands, or biomolecules is the stability constant of the complex (β). In principle, it can be calculated as the free energy change associated with the metal ion complexation, i.e. its uptake from the solution under standard conditions. Since this process is associated with interactions of charged species, large values of interaction and solvation energies are in general involved. Using the standard thermodynamic cycle (*in vacuo* complexation, and solvation/desolvation of the reference state and of the resulting complexes), one usually subtracts values of several hundreds of kcal.mol^{-1} to obtain final results in the orders of units or tens kcal.mol^{-1}. Employing the density functional theory (DFT) and Møller-Plesset second order perturbation theory (MP2) calculations together with the conductor-like screening model for realistic solvation (COSMO-RS method) the stability constants of selected complexes – $[M(NH_3)_4]^{2+}$, $[M(NH_3)_4(H_2O)_2]^{2+}$, $[M(Imi)(H_2O)_5]^{2+}$, $[M(H_2O)_3(His)]^+$, $[M(H_2O)_4(Cys)]$, $[M(H_2O)_3(Cys)]$, $[M(CH_3COO)(H_2O)_3]^+$, $[M(CH_3COO)(H_2O)_5]^+$, $[M(SCH_2COO)_2]^{2-}$ – with eight divalent metal ions (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) were calculated.[12] Using the currently available computational protocols it was shown that it is possible to achieve the *relative* accuracy of 2-4 kcal.mol^{-1} (1-3 orders of magnitude in β) after subtracting the ligand-dependent and metal-dependent systematic shifts. In summary, the protocol represents an exciting (and complementary) alternative to evaluate the metal ion selectivity from the first principles and for the entirely general metal-binding site.

---

[9] Navrátil, V.; Klusák, V.; Rulíšek, L.: Theoretical Aspects of Hydrolysis of Peptide Bonds by Zinc Metalloenzymes. *Chem. Eur. J.* **2013**, *19*, 16634-16645.

[10] Plechanovová, A.; Byun, Y; Alquicer, G.; Škultétyová, Ľ; Mlčochová, P.; Němcová, A.; Kim, H.-J.; Navrátil, M.; Mease, R.; Lubkowski, J.; Pomper, M.; Konvalinka, J.; Rulíšek, L.; Bařinka, C.: Novel Substrate-Based Inhibitors of Human Glutamate Carboxypeptidase II with Enhanced Lipophilicity. *J. Med. Chem.* 2011, *54*, 7535-7546.

[11] Gutten, O.; Beššeová, I.; Rulíšek, L.: Interaction of Metal Ions with Biomolecular Ligands: How Accurate Are Calculated Free Energies Associated with Metal Ion Complexation? *J. Phys. Chem. A* **2011**, *115*, 11394-11402. (Correction: *J. Phys. Chem. A* **2012**, *116*, 8407-8407.

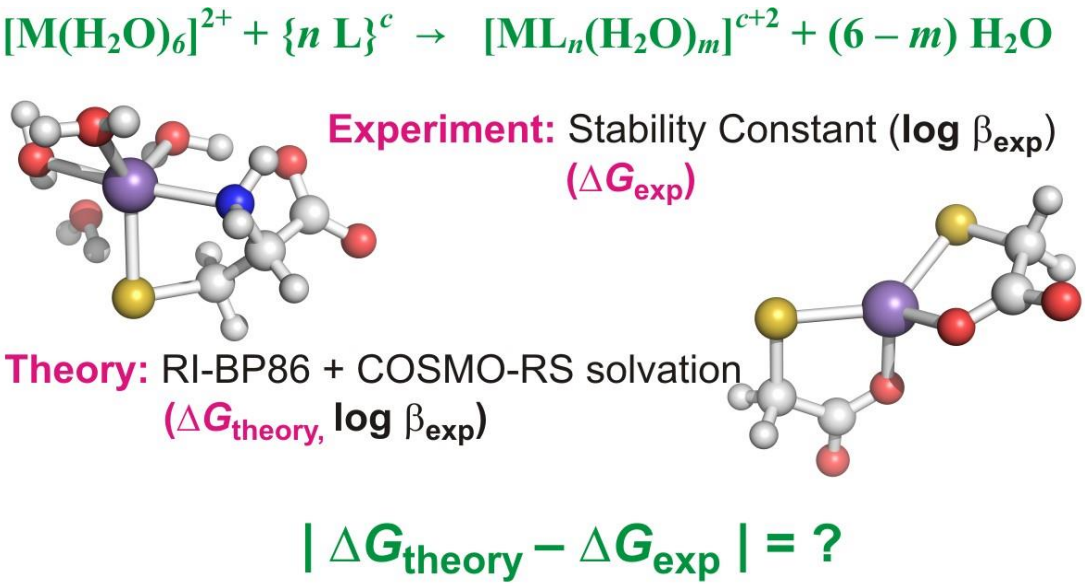


Figure 3: Predicting stability constants of transition metal-ion complexes in solution represents a challenge for which contemporary computational chemistry may not yet be able to provide final answers. However, a careful use of modern computational methods in combination with advanced solvation models, such as COSMO-RS, can at least provide quantitative estimates of *relative* affinities for a series of metal-ions and a number of insights into this intriguing topic.

2.3. Computational Electrochemistry: From Small Molecules to Metalloproteins.

Recent progress in quantum chemical methodology together with the development of reliable solvation models have resulted in the situation where many physicochemical properties of molecular systems of increasing complexity can be predicted with quantitative or near-quantitative accuracy. One such property of a prime interest is the reduction potential. Its value represents an inherent propensity of the particular system to accept or deliver one or several of its electrons. While several experimental methods are (often routinely) used to determine this value, theoretical calculations may complement the data by providing a unique and experimentally unattainable structure/energy(E^0) mapping. If sufficiently accurate, such a mapping implies that reduction potentials can be used as the key descriptors in our understanding to elementary electron-transfer steps in electrochemistry, molecular electronics and biological redox processes to give just several examples of the practical importance. In

[12] Gutten, O.; Rulíšek, L.: Predicting the Stability Constants of Metal-Ion Complexes from First Principles. *Inorg. Chem.* **2013**, *52*, 10347-10355.

our group, some emphasis was given to efforts to design robust and sufficiently accurate protocols for calculations of reduction potentials. Various classes of molecules were studied, ranging from small organic molecules to metalloproteins.

2.3.1. The Reduction Pathways of 2,4,6-Trinitrotoluene: An Electrochemical and Theoretical Study. The reduction pathways of trinitrotoluene were studied using electrochemical and computational methods. The electrochemical reduction of three nitro groups in 2,4,6-trinitrotoluene is characterized by three major reduction peaks in cyclic voltammograms at the peak potentials of - 0.310, -0.463 and -0.629 V vs. normal hydrogen electrode (NHE). The second and third peaks coincide with the two peaks observed for the 2-amino-4,6-dinitrotoluene (at the potentials of -0.475 and -0.627 V vs. NHE), whereas the two peaks in the 4-amino-2,6-dinitrotoluene voltammograms appear at -0.537 and -0.623 V and deviate more significantly from the corresponding two peaks in 2,4,6-trinitrotoluen. It suggests that the first NO_2 group reduced in the overall process is the one in *ortho* position with respect to CH_3 group. Analogously, the 2,6-diamino-4-nitrotoluene exhibits a reduction peak at -0.629 V, almost identical to the third and second reduction peaks of 2,4,6- trinitrotoluene and 2-amino-4,6-dinitrotoluene, respectively. Since the other isomer, 2,4-diamino-6- nitrotoluene, exhibits a reduction peak at -0.712 V, we conclude that the second reduction occurs also in the *ortho* position with respect to the methyl group. These observations were corroborated by quantum chemical calculations, which yielded reduction potentials in a good agreement with the experimental values. Thus, studying in detail all of the possible protonation and redox states in the reduction of the first nitro group and the key steps in the reduction of the second and third nitro groups, we have obtained a comprehensive and detailed picture of the mechanism of the full $18e^-/18H^+$ reduction of TNT.[13]

2.3.2. Electrochemistry of Acceptor-Substituted Ferrocenium Salts as New Strong Single-Electron Oxidants. A series of mono- and 1,1'-diheteroatom-substituted ferrocene derivatives as well as acylated ferrocenes and their ferrocenium salts was prepared in the group of Dr. Ullrich Jahn. The redox potentials of the synthesized ferrocenes were determined by cyclic voltammetry and it was observed that all new ferrocenium salts have stronger oxidizing properties compared to standard ferrocenium hexafluorophosphate. Quantum chemical calculations of the reduction potentials of the substituted ferrocenium ions (25 species in total) were carried out using a standard thermodynamic cycle involving the gas-phase energetics and solvation energies of the contributing species. A remarkable agreement between theory and experiment was found; the mean average deviation amounted to only 0.030 V and the maximum deviation to 0.1 V (on the fairly broad range of E^0_{exp} = 0.46–1.19 V). This enabled the analysis of various physical contributions to the computed reduction potentials of these ferrocene derivatives providing insight into their electronic structure and physicochemical properties. To our best knowledge, this was the best agreement between experimental and calculated reduction potentials reported to date.[14]

2.3.3. Accuracy of Calculated Reduction Potentials of Selected Group 8 (Fe, Ru and Os) Octahedral Complexes. The theoretical calculations of reduction potentials for the $[M(H_2O)_6]^{2+/3+}$, $[M(NH_3)_6]^{2+/3+}$, $[M(en)_3]^{2+/3+}$, $[M(bipy)_3]^{2+/3+}$, $[M(CN)_6]^{4-/3-}$ and $[MCl_6]^{4-/3-}$ systems (M = Fe, Os, Ru) were carried out. The DFT(PBE)/def2-TZVP//DFT(PBE)/def2-SVP quantum chemical method was employed to obtain presumably accurate ionization energies whereas the conductor-like screening model for real solvents (COSMO-RS) was selected as the most suitable method for calculations of solvation energies of the oxidized and reduced forms of the studied species. It has been shown that COSMO-RS may overcome problems related to directionality of hydrogen bonds in the second solvation sphere that previously lead to errors of ~1V for $[Ru(H_2O)_6]^{2+}$ complex employing PCM-like models. Thus, most of the values for (2+) → (3+) oxidations are now within 0.1 – 0.2 V from the experimental data, once the anticipated spin-orbit coupling effects in Os complexes (downshifting the calculated reduction potentials by ~0.3 V) are taken into account. The robustness of the DFT(PBE)/COSMO-RS computational protocol is further verified by showing that reduction potentials

[13] Chua, C. K.; Pumera, M.; Rulíšek, L.: Reduction Pathways of 2,4,6-Trinitrotoluene: An Electrochemical and Theoretical Study. *J. Phys. Chem. C* **2012**, *116,* 4243-4251.

[14] Khobragade, D. A.; Mahamulkar, S. G.; Pospíšil, L.; Císařová, I.; Rulíšek, L.; Jahn, U.: Acceptor-Substituted Ferrocenium Salts as Strong, Single-Electron Oxidants: Synthesis, Electrochemistry, Theoretical Investigations, and Initial Synthetic Application. *Chem. Eur. J.* **2012**, *18,* 12267-12277.

obtained for selected 2+/3+ redox pairs with and without the inclusion of explicit second-sphere water molecules are almost identical.[15]

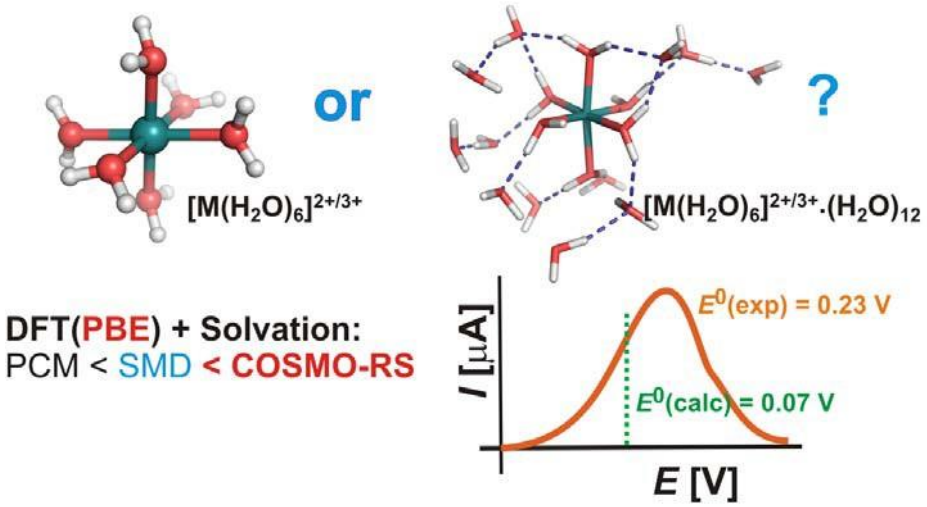


Figure 4: The schematic representation of the potential caveats involved in the calculations of reduction potentials of (formally) charged transition metal complexes with the anticipated directionality of hydrogen bonds in the second solvation sphere, thus necessitating the use of explicit solvation models. These problems are overcome by the COSMO-RS protocol.

2.3.4. Reduction potentials and acidity constants of Mn superoxide dismutase calculated by QM/MM free-energy methods. Two theoretical methods to estimate reduction potentials and acidity constants in Mn superoxide dismutase (MnSOD) were used, viz. combined quantum mechanical and molecular mechanics (QM/MM) thermodynamic cycle perturbation (QTCP) and the QM/MM-PBSA approach, in which the QM/MM energies are combined with continuum solvation energies calculated by solving the Poisson–Boltzmann equation (PB) or by the generalised Born approach (GB) and non-polar solvation energies calculated from the solvent-exposed surface area. We showed that using the QTCP method, we can obtain accurate and precise estimates of the proton-coupled reduction potential for MnSOD, ~0.30 V, which compares favourably with experimental estimates of 0.26–0.40 V. However, the calculated potentials depend strongly on the DFT functional used: the B3LYP functional gives 0.6 V more positive potentials than the PBE functional. The QM/MM-PBSA seems in general somwhat inferior to the QTCP approach but it may still be useful to estimate relative shifts.[16]

2.4. Theory and Method Development.

2.4.1. Curvature Correction for Microiterative Optimizations with QM/MM Electronic Embedding. Our contribution to the method development has been represented by the improvement of the QM/MM minimization procedure. We suggested a computationally cheap second-order correction that employs an estimated Hessian from the Davidon-Fletcher-Powell method to tackle the problems caused by the too small curvature of the QM/MM potential energy surface. Test calculations on four metalloenzymatic systems (~100 QM atoms, ~2000 relaxed MM atoms, ~20000 atoms in total) show that our approach efficiently restores the convergence in cases where gradient correction leads to oscillations and therefore, improves the robustness of the QM/MM minimization.[17]

---

[15] Rulíšek, L.: On the Accuracy of Calculated Reduction Potentials of Selected Group 8 (Fe, Ru and Os) Octahedral Complexes. *J. Phys. Chem. C* **2013**, *117*, 16871-16877.

[16] Heimdal, J.; Kaukonen, M.; Srnec, M.; Rulíšek, L.; Ryde, U.: Reduction Potentials and Acidity Constants of Mn Superoxide Dismutase Calculated by QM/MM Free-Energy Methods. *ChemPhysChem* **2011**, *12*, 3337-3347.

[17] Rokob, T. A.; Rulíšek, L.: Curvature Correction for Microiterative Optimizations with QM/MM Electronic Embedding. *J. Comput. Chem.* **2012**, *33,* 1197-1206.

2.5. Organic Reactivity

An active collaboration with experimental groups at IOCB resulted in several studies of practical interest that can be grouped as **theoretical organic reactivity**. These problems often boil down to finding and characterizing the transition-state structures (TS) along the reaction coordinate (and correlating them with the experimental reaction rates). Besides yielding fundamental understanding of the studied reactions, it enables us also to predict stereoselectivities or regioselectivities in organic reactions. Thus, we studied reaction mechanism of asymmetric allylation of aldehydes with trichloroallylsilanes catalyzed by Quinox and Methox, a chiral isoquinoline *N*-oxides as a prominent example of stereoselective organocatalysis,[18] thiocyanation of *closo*-dodecaborate $B_{12}H_{12}^{2-}$,[19] copper-mediated C−S bond formation in the intramolecular disproportionation of imine disulfides,[20] and an elegant explanation of divergent pathways and competitive mechanisms observed experimentally in metathesis reactions between 3-arylprop-2-ynyl esters and aldehydes.[21]

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

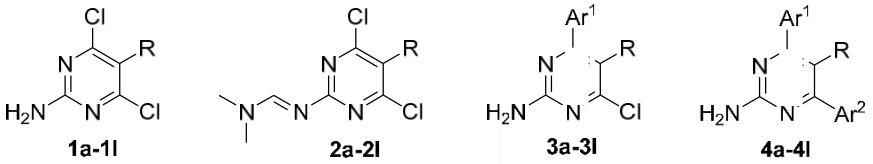
Scientific team	Combined Junior groups of medicinal chemistry

Main research topics of the group of Zlatko Janeba:

Development of potent non-nucleoside reverse transcriptase inhibitors. In 2013 and beginning of 2014, part of the team of Dr. Janeba cooperated with Gilead Sciences, Inc. (Foster City, California, USA) on development of novel types of non-nucleoside reverse transcriptase inhibitors (NNRTIs). Some 450 new compounds were prepared at IOCB, while the biological testing was performed in Gilead Sciences. Submission of patent applications is in progress. The study is commercial-in-confidence and for this reason, to date, we have not been able to publish the very promising results obtained.

Polysubstituted pyrimidines with anti-inflammatory properties. The pyrimidine moiety represents an important structural motif commonly found either in number of natural products or in synthetic derivatives that exhibit important biological effects. We have recently reported a series of 5-substituted 2-amino-4,6-dichloropyrimidines **1** (Figure 1) as potent inhibitors of immune-activated nitric oxide (NO) production (Jansa P. et al., *Med. Chem. Res.* **2014**, *23*,

4482). The most effective was 5-fluoro-2-amino-4,6-dichloropyrimidine with an IC_{50} of 2 μM (higher activity than the most potent reference compound 1,400 W), while the IC_{50}s of other derivatives were within the range of 9–36 μM (Jansa P. et al., *Med. Chem. Res.* **2014**, *23*, 4482). Later it was shown that also 5-substituted 4,6-dichloro-2-[(*N*,*N*-dimethylamino)methyleneamino]pyrimidines **2** (Figure 1) strongly inhibited NO production, with the IC_{50}s <5 μM in most cases (Jansa P. et al., *Med. Chem. Res.* **2015**, doi:10.1007/s00044-014-1285-5; Jansa P. et al., US 8,883,798 B2).



a = H, **b** = methyl, **c** = ethyl, **d** = propyl, **e** = isopropyl, **f** = propargyl, **g** = allyl, **h** = butyl, **i** = sec. butyl, **j** = phenyl, **k** = benzyl, **l** = fluoro; Ar^1 and Ar^2 = (het)aryl

Figure 1. General structures of polysubstituted pyrimidines **1-4** with anti-inflammatory properties.

Subsequently, a large series of pyrimidine derivatives bearing one (in C-4 position) or two (in C-4 and C-6 positions) (het)aryl substituents, compounds **3** and **4** (Figure 1), respectively, were synthesized which were shown to reduce simultaneously the production of NO and prostaglandin E2 (PGE2) (Jansa P. et al., US 8,883,798 B2). The compounds are not cytotoxic and they did not exhibit any negative effect on the viability of cells at concentrations decreasing the production of these factors by up to 50%. These compounds have a great potential in the treatment of inflammatory and cancer diseases. Pharmacokinetic and pharmacodynamic data are collected for the selected candidate, compound WQE-134, and its synthesis is being optimized for the bulk preparations. *In vivo* experiments on animal models of ulcerative colitis and rheumatoid arthritis are in progress. The exact mechanism of action of these analogues remains to be elucidated.

The team of Dr. Janeba is responsible for design, synthesis and characterization (structural, purity) of the compounds. The biological evaluations are done by collaborators (Dr. Hájek from the IOCB, Dr. Zídek from the Institute of Experimental Medicine).

Polysubstituted pyrimidines with strong intramolecular hydrogen bonds. Planamerism. The ability to form strong hydrogen bonds represents one of the most significant properties of biologically important pyrimidines. Intermolecular hydrogen bonding is crucial for encoding the genetic information (e.g. Watson-Crick base pairs), but intramolecular hydrogen bonding occurs within nucleic acids as well (e.g. in single-stranded transfer RNAs) and allows to fold them into a variety of specific three-dimensional structures. Nevertheless, the formation of intramolecular hydrogen bond (IMHB) has a great impact on the molecular structure and physicochemical properties of small molecules as well. High stability of IMHB has often been observed in cases where a six-membered ring can be formed and the linker atoms are sp^2-hybridized (amides, enols of β-diketones and β-enaminones, heteroaromatic rings etc.). Such structures are planar and the high stability of IMHB has been proposed to be due to stabilization by resonance (resonance-assisted hydrogen bond, RAHB).

Substituted 5-nitrosopyrimidines are compounds of special interest. It has been previously speculated that polysubstituted 5-nitrosopyrimidines with IMHB are mimics of fused heterobicyclic moieties such as purines.

Recently, a series of polysubstituted 5-nitrosopyrimidines has been prepared in our lab (Figure 2) and strong IMHBs between the oxygen atom of the 5-nitroso group and hydrogen atom of the neighbouring amino groups (namely in positions C-4 and/or C-6 of the pyrimidine moiety) have been reported (Procházková E. et al., *J. Org. Chem.* **2013**, *78*, 10121).

Furthermore, various ratios of two possible rotamers **7** and **8** (Figure 2) in equilibrium were observed depending on the character of the substituent attached to the amino groups in C-4 and C-6 positions.

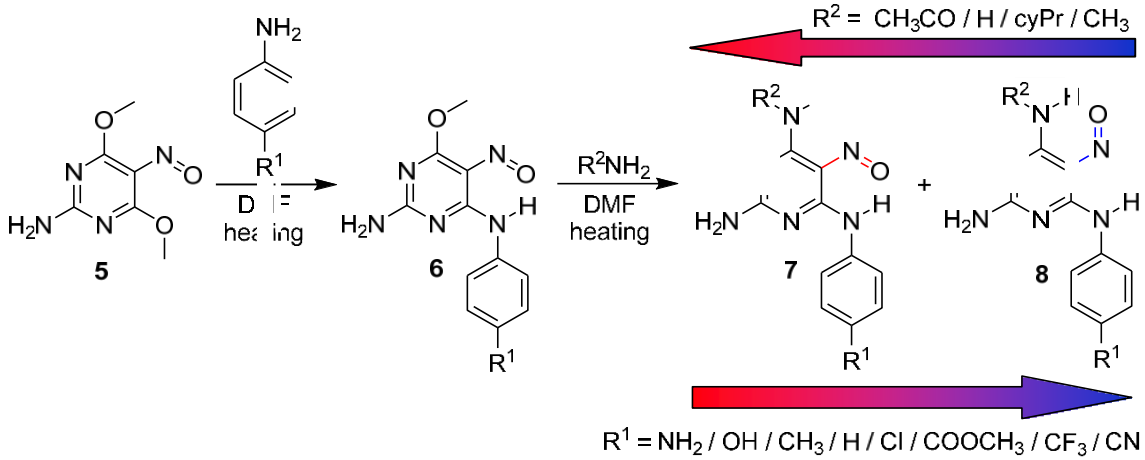


Figure 2. General synthesis of polysubstituted 5-nitrosopyrimidines and structures of the two possible rotamers **7** and **8**.

Even more recently, we have shown that in several cases two possible rotamers of 2,4,6-triamino-5-nitrosopyrimidines (e.g. isomers **9** and **10**, Figure 3) can be isolated and characterized as chemical entities (Čechová L. et al., *Chem. Commun.* **2014**, *50*, 14892). We were quite astonished by the stability of the single rotamers as they could be separated by means of the standard silica gel chromatography at room temperature. Although separation of single conformers of large molecules (e.g. proteins) has been reported before, to the best of our knowledge, no such a precedent is known for small (hetero)aromatic molecules. Unlike atropisomers (Figure 3), whose separation is achieved through steric hindrance (and display chirality), the isolation of our planar rotamers is achieved through IMHBs (and are achiral). To make a distinction from relatively common atropisomerism, we suggested a term 'planamerism' and defined 'planamers'as small aromatic molecule rotamers with a planar conjugated moiety that are isolable as chemical species (Figure 3).

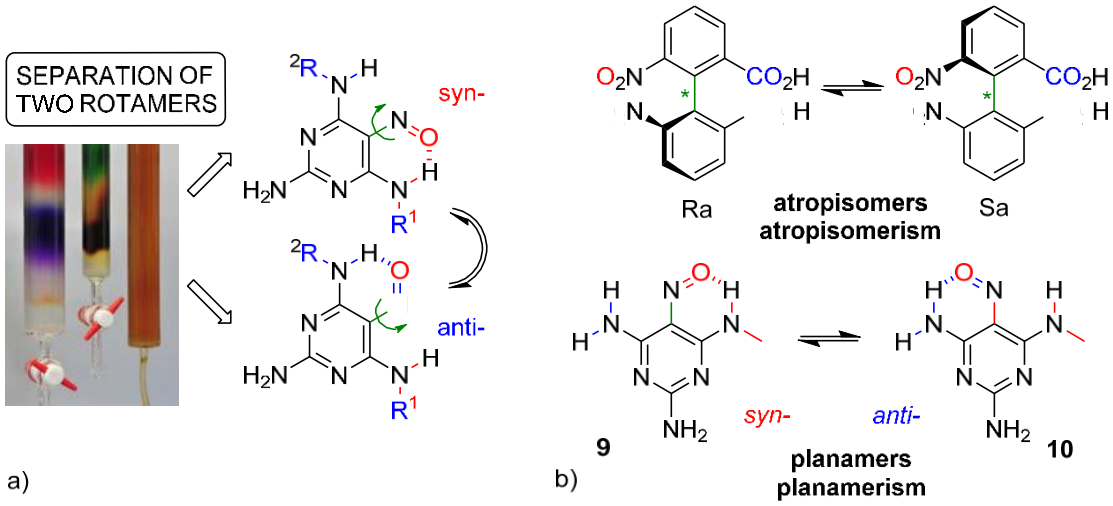


Figure 3. a) An unique isolation of pairs of planamers as chemical species differing only in nitroso group orientation. b) A comparison of atropisomerism and planamerism.

The team of Dr. Janeba is responsible for design, synthesis and characterization (structural, purity) of the compounds. The DFT calculations and NMR experiments are done by collaborators (Dr. Dračínský from the IOCB).

Acyclic nucleoside phosphonates as inhibitors of 6-oxopurine phosphoribosyl-transferases and potential antimalarial agents. Hypoxanthine-guanine-(xanthine) phosphoribosyltransferase (HG(X)PRTase) catalyses the formation of 6-oxopurine nucleoside monophosphates and is the key enzyme of the **purine salvage pathway**. While mammals are able to synthesize purine derivatives *de novo* as well as by salvage of preformed purine bases, malarial parasites of the genus *Plasmodium* depend completely on the transport of preformed bases from the host cell. Recently, we have discovered (in collaboration with the group of Dr. Luke Guddat, University of Queensland, Australia) inhibitors of PRTases based on the **acyclic nucleoside phosphonate** (**ANP**) scaffold (de Jersey J. et al., *Curr. Top. Med. Chem.* **2011**, *11*, 2085). ANPs were synthesized by Prof. A. Holý at IOCB as antiviral agents that target DNA polymerases and reverse transcriptases. Since that initial discovery several ANPs have progressed into clinical use as potent antivirals. The flexibility of the acyclic chain enables the compounds to adopt a conformation suitable for interaction with the active site of the enzymes.

We have successfully improved the selectivity and inhibitory activity of this first generation of ANP-based inhibitors of PRTases. The effect of acyclic chain length as well as the influence of various substituents on ability and selectivity of inhibitors were studied (Keough D. T. et al., *Mol. Biochem. Parasitol.* **2010**, *173*, 165; Česnek M. et al., *Bioorg. Med. Chem.* **2012**, *20*, 1076; Krečmerová et al., *Bioorg. Med. Chem.* **2012**, *20*, 1076 and Baszczyňski O. et al., *Eur. J. Med. Chem.* **2013**, *67*, 81). Based on our knowledge of the crystal structures of the human enzyme in complex with several ANPs, we have been able to design new compounds that can precisely fill the active site of PRTase. Our goal was design of compounds, where a single inhibitor molecule can occupy all three key binding sites (i.e. purine base, 5´-phosphate group and pyrophosphate, Figure 4A) in the active site of the enzyme. We designed a series of bisphosphonates containing a second phosphonate group attached to the ANP scaffold (Keough, D. T. et al., *J. Med. Chem.* **2013**, *56*, 2513; Figure 4A and 4B). Later, by positioning of the nitrogen atom at the branching point in acyclic moiety, a series of aza-ANPs was synthesized (Hocková D. et al., *J. Med. Chem.* **2012**, *55*, 6209 and Keough D. T. et al., *J. Med. Chem.* **2015**, *58*, 827; Figure 4D). Patent on these structurally novel inhibitors (US/61/643,419, WO2013166545-A2) was filled. In agreement with the literature, we have observed that the preparation of suitable prodrugs of the parent ANP-based inhibitors can significantly improve activity in cell based antimalarial assays. While the penetration of free phosphonic acids to the cells was very limited, phosphoramidate prodrugs (Figure 4C) were effective in the erythrocytes infected by *P. falciparum*. The above extensive SAR-study provides insights for the design of potential antimalarial drug leads with entirely novel mechanism of action.

The team of Dr. Janeba is responsible for design, synthesis and characterization of the enzyme inhibitors (more than 500 compounds). The biological evaluations are done by collaborators (group of Dr. Luke W. Guddat, University of Queensland, Australia and Dr. Michael D. Edstein Australian Army Malaria Institute, Enoggera).

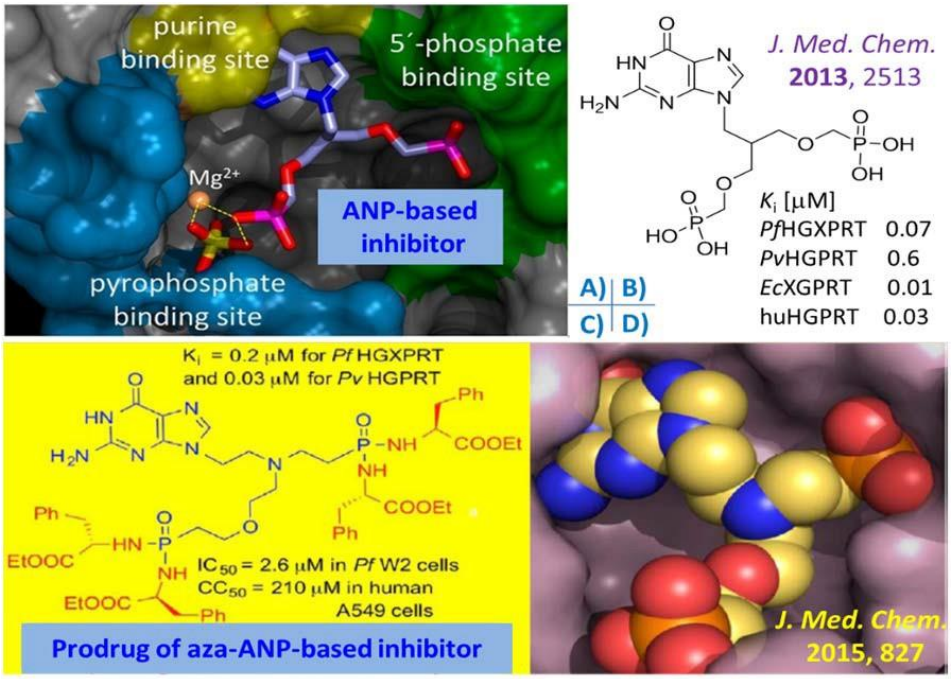


Figure 4. Examples of ANP-based PRTase inhibitors, crystal structures and phosphoramidate prodrug.

Acyclic nucleoside phosphonates as inhibitors of bacterial adenylate cyclases. Adenylate cyclase toxin (ACT) and edema factor (EF) are the key virulence factors of *Bordetella pertussis* and *Bacillus anthracis*, respectively, that facilitate their invasion into the mammalian body. A series of novel 9-[2-(phosphonomethoxy)ethyl]adenine (PMEA) amidate prodrugs was prepared as 9-[2-(phosphonomethoxy)ethyl]adenine diphosphate (PMEApp), the active metabolite of the antiviral drug bis(POM)PMEA (adefovir dipivoxil), has been shown to inhibit ACT. Although the amidate prodrugs did not inhibit ACT as effectively as bis(POM)PMEA, they were significantly less cytotoxic (Šmídková M. et al., *Antimicrob. Agents Chemother*. **2014**, *58*, 664). Moreover, they all reduced apoptotic effects of ACT and prevented an ACT-induced elevation of intracellular $[Ca^{2+}]i$. The amidate prodrugs were less susceptible to degradation in Caco-2 cells compared to bis(POM)PMEA, while they exerted good transepithelial permeability. As a consequence, a large amount of intact amidate prodrug is expected to be available to target macrophages *in vivo*. This feature makes nontoxic amidate prodrugs attractive candidates for further investigation as novel antimicrobial agents.

The team of Dr. Janeba is responsible for design, synthesis and characterization (structural, purity) of the compounds. Biological experiments are done by collaborators (Dr. Mertlíková-Kaiserová from the IOCB, Dr. Oto Pavliš from Biological Defence Department in Těchonín).

Main research topics of the group of Marcela Krečmerová:

1. 5-Azacytosine and 5,6-dihydro-5-azacytosine nucleosides with epigenetic (hypomethylating) activity

Anticancer agents 2'-deoxy-5-azacytidine (decitabine, Dacogen®) and 5-azacytidine (Vidaza®) are compounds working on epigenetic principle based on inhibition of DNA methyl transferases. A serious disadvantage of 5-azacytosine drugs is their instability in aqueous solutions. The drugs are applied in a form of cold infusions which is extremely unpleasant for

patients. Therefore, we aimed to increase stability of 5-azacytosine nucleosides by appropriate modifications (not decreasing hypomethylating activity) and in parallel, also to find new hypomethylating agents with sufficient stability.

We managed to discover epigenetic activity of stable 5,6-dihydro-5-azacytosine derivatives:

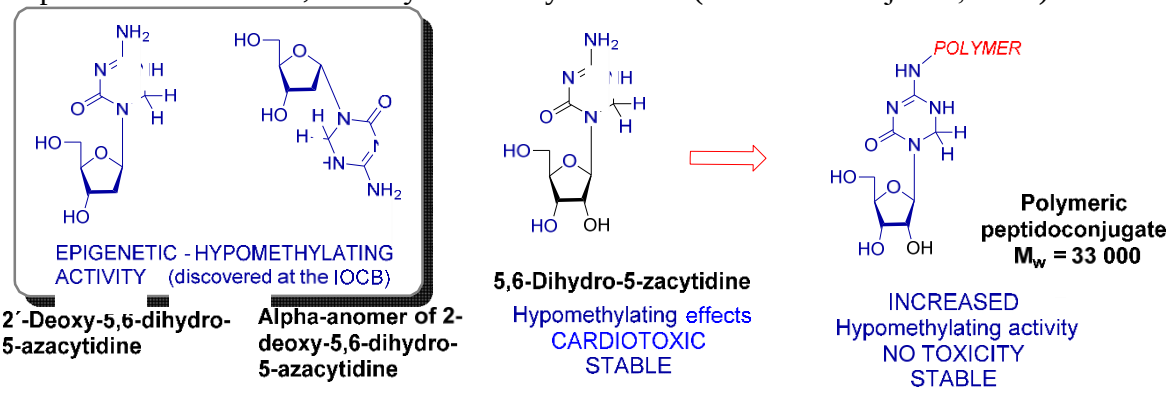
- **5,6-dihydro-2'-deoxy-5-azacytidine**
- **alpha anomer of 5,6-dihydro-2'-deoxy-5-azacytidine – so far considered biologically inactive.**

This discovery was made in collaboration with the biochemical department of the IOCB).

We focused also to another epigenetic compound, 5,6-dihydro-5-azacytidine. It has been already studied in clinical trials but it was found cardiotoxic. We managed to solve this problem as follows:

Transformation of 5,6-dihydro-5-azacytidine to a polymeric peptidoconjugate lead to a significant increase of the hypomethylating capability in comparison with the parent 5,6-dihydro-5-azacytidine. This polymer was completely nontoxic and stable.

All compounds are currently investigated *in vitro* and *in vivo* in the Laboratory of Experimental Medicine, Palacký University Olomouc (Dr. Marian Hajdúch, Head).



Synthetic methodology for appropriate reactive derivatives of 5,6-dihydro-5-azacytidine for their attachment to a polymer carrier was worked out. These methods mostly use formation of hydrazon by reaction with a hydrazide group of polymer. Besides, we also managed to form polymeric conjugate of 5-azacytidine (Vidaza). In both cases, the bonds to polymeric carrier are hydrolytically labile and enable release of the active substance. Clinical application of the original drug (Vidaza) can be thus substantially improved by control release from the polymeric conjugate.

2. 5-Azacytosine acyclic nucleoside phosphonates

1-(*S*)-[3-Hydroxy-2-(phosphonomethoxy)propyl]-5-azacytosine (HPMP-5-azaC), its cyclic form and ester prodrugs were identified as a new class of potent anti-DNA virus agents with lower toxicity and better activity compared to cidofovir (antiviral selectivity index was found 2- to 16-fold higher). Antiviral investigation was performed in collaboration with Rega Institute KU Leuven and resulted in the joint patent EP 1966226 B1. A special effort was paid to development of diverse types of prodrugs and other modifications to increase bioavailability and/or influence instability of the 5-azacytosine ring. (*S*)-HPMP-5-azaC and its prodrugs were also found the most efficient and the most selective compounds against polyomavirus infections and potential candidates against BK virus-associated nephropathies.

3. Development of prodrugs

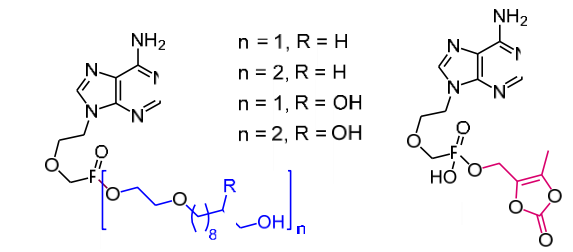
3.1. Prodrugs of acyclic nucleoside phosphonates

a) Development of new **functionalized ANP ester prodrugs** of with (di)hydroxylated (decyloxy)ethyl moieties, mono or diesters with undecyl-derived moiety containing a (masked) semipolar diol system and hexaethylene glycol based prodrugs of adefovir and cidofovir as more polar analogs of usual alkoxyalkyl prodrugs and dioxolenone esters, e.g. (5-

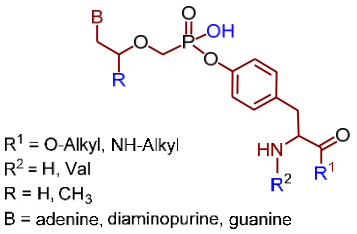
methyl-2-oxo-1,3-dioxolen-4-yl)methyl esters. Very promising antiviral data (HIV, herpes viruses) were found especially in the dioxolenone derivatives of PMEA (adefovir) (Fig. 2).

b) **Amino acid and peptidomimetic prodrugs** – a collaborative project with the USC, Los Angeles (Professor C.E. McKenna). Very good pharmacokinetic properties were found in case of the Val-Ser dipeptide promoiety where a valine carboxyl function was further esterified, e.g. Val-Ser-COOiPr cHPMPC. After the pilot syntheses with cidofovir and (*S*)-HPMPDAP (as a part of our anti-pox virus program), we continued with syntheses of tyrosine based prodrugs of anti-cancer compounds PMEG, i.e. 9-[2-(phosphonomethoxy)ethyl]guanine and its N^6-cyclopropyl-2,6-diaminopurine counterpart. Compounds had very good cytostatic activity including drug resistant cell lines, low toxicity for normal healthy cells and good bioavailability.

c) **Poxvirus program** – a collaborative program with Rega Institute and Gilead Sciences aimed to develop drug candidates against poxviruses as a potential bioterrorist weapon (*Variola*). (*S*)-HPMPDAP was selected as the best candidate. To increase its bioavailability for oral application, a large series of ester prodrugs was synthesized (POM, POC, trifluoroethyl, alkoxyalkyl and alkyl salicylyl esters and phosphoramidates). POM-HPMPDAP was finally selected as the best candidate for further development.



(Di)hydroxylated (decyloxy)ethyl esters and dioxolenone prodrug of adefovir (PMEA)



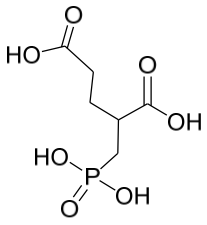
Tyrosine prodrugs of PME and PMP derivatives

3.2. Prodrugs of 2-PMPA. Inhibitors of glutamate carboxypeptidase II (GCPII)

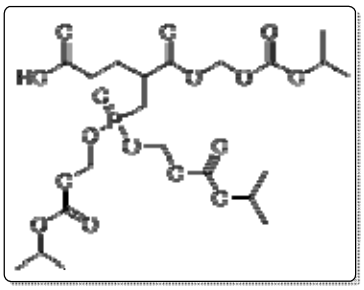
In 2012, the group was invited to start collaboration with the Johns Hopkins University, Baltimore, MD, (Professor Barbara Slusher) together with the team of Dr. Pavel Majer (IOCB) on systematic development of oral prodrugs of 2-PMPA - a compound inhibiting GCPII. GCPII is the enzyme implicated in the liberation of glutamate in brain synapsis formed by hydrolysis of a neuropeptide N-acetylaspartyl glutamate. Glutamate is an activator of ionotropic glutamate receptors whose overactivation has been implicated in diverse neurodegenerative disorders, e.g. stroke, ALS, chronic pain, diabetic neuropathy, etc. Inhibition of GCPII is thus a promising approach for the treatment of all disorders associated with glutamate excitotoxicity.

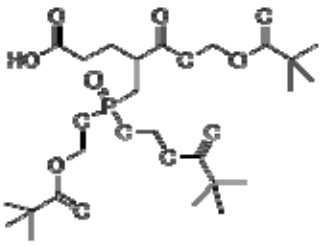
One of the most effective GCPII inhibitor is 2-(phosphonomethyl)pentane-1,5-dioic acid (2-PMPA) but its highly polar character with a low bioavailability limits its potential use as a drug. Synthesis of a large series of structurally diverse types of prodrugs resulted in finding that simple alkyl esters of the acids such as methyl, ethyl, and propyl were not successful, due to excess stability of these moieties on the carboxylates. Pivaloyloxymethyl (POM) and isopropyloxycarbonyloxymethyl (POC) on phosphonate groups demonstrated the right combination of lability *in vitro* and provided the highest levels of prodrug derived species when dosed orally. Even a compound with a free γ carboxylate demonstrated good bioavailability. However, none of these compounds released 2-PMPA *in vivo* to any appreciable extent. The presently disclosed subject matter shows that POM and POC on the bis(phosphonate) and a free alpha carboxylate (i.e. trisPOM, trisPOC derivatives) were ideal for enhancing the permeability (approximately 20 fold), as well as release of the parent

compound upon oral dosing. A similar great effectivity was also found in a corresponding dioxolenone derivative.



2-PMPA


2-PMPA-Tris-POC


2-PMPA-Tris-POM

***In vivo* results of oral 2-PMPA-Tris-POC for IBD (Inflammatory bowel diseases):**

- The compound enhances 2-PMPA exposures in mice and dogs
- decreased colitis severity in DSS treated mice (DSS = Dextran sulfate sodium induced colitis)
- prevented the shortening of the colon in DSS-colitis and decreased colitis severity

Future clinical investigations of the selected drug candidate(s) is highly probable.
The joint patent application between the IOCB and JHU was filled (US Appl. No. 62/033926, August 06, 2014).

4. Other projects

- Within the scope of the malaria project a series of purine 2-hydroxy-3-(phosphonomethoxy)-propyl ("iso-HPMP") derivatives) and a series of prodrugs of selected ANPs (PMEG, PEEG, PEEHx, HPMPG) were synthesized as inhibitors of *Plasmodium* enzyme HG(X)PRT.
- Enzymatic reactions in synthesis of prodrugs of antivirally active nucleoside and nucleotide analogues were developed as topics of Ph.D. and bachelor´s thesis of two students (lipase catalyzed esterifications by vinyl esters, enzymatic galactosylations).
- New inhibitors of human thymidine phosphorylase in a series of 5-aryl-6-(phosphonomethoxy)-uracils and 8-aza-7,9-dideazaxanthine derivatives were developed.
- Synthetic methodology leading to geminal difluorinated ANPs was worked out, starting from the commercially accessible 2-benzyloxyacetaldehyde using difluoromethylphosphonate as a fluorination reagent. The multi-step process leads to the reactive CF_2 -containing synthon which can be subsequently coupled with diverse nucleobases.

Main research topics of the group of Radim Nencka

Conformationally locked nucleosides and nucleotides

Our research was focused on the synthesis and antiviral evaluation of novel conformationally locked nucleoside and nucleotide analogues with sugar moiety substituted with the bicyclo[2.2.1]heptane (norbornane) scaffold. The compounds prepared in this series are closely related to the clinically used carbocyclic nucleosides Abacavir and Entecavir. Our aim was to comprehensively investigate the potency of conformationally constrained derivatives in antiviral treatment. Within the course of our work, we decided to prepare a number of derivatives locked in various conformations including North, South or East (Figure 1.).
The results of our studies were quite surprising. Although we targeted primarily on antiviral activity, the compounds from North and East series were mostly devoid of significant antiviral effect, except for anti-enterovirus activity of 6-chloropurine derivatives (see below). On the other hand, the adenosine derivative from the North series exerted interesting selective

inhibition of PI4K II< < (Dejmek et al. *Bioorg. Med. Chem.* 2015, 184). In the series of nucleoside derivatives locked in South conformation we have identified compounds active against Feline herpes virus an important pathogen of cats. Since the active compounds do not have the hydroxymethyl moiety, which bears triphosphate necessary for incorporation of nucleotide into polynucleotide chain and allows chain termination, the mechanism of action of these compounds seems to be different and yet unknown (Dejmek et al. *Bioorg. Med. Chem.* 2014, 2974). Thus, all these series afforded interesting hits, which we would like to use as a starting point for our further research.

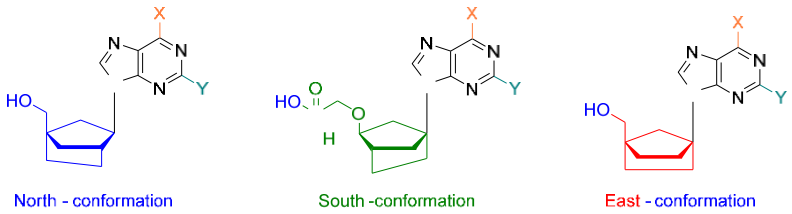


Figure 1.

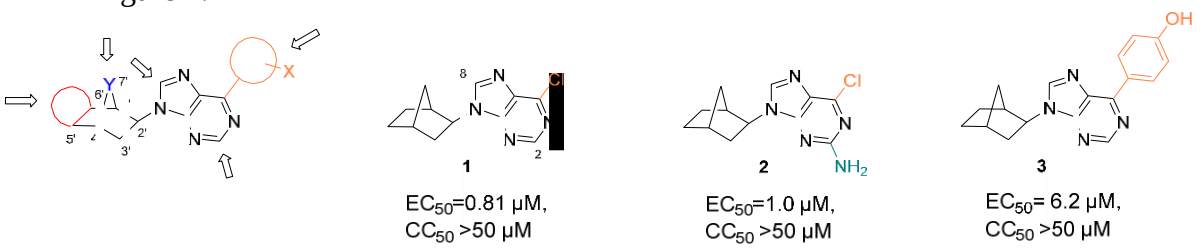
9-Substituted purines as agents against Coxsackieviruses

Coxsackieviruses (*Picornavirideae*) are important human pathogens. Among Enteroviruses that are suspected to cause 10-15 millions symptomatic infections each year in the USA only, Coxsackieviruses belong to the most commonly identified. Besides banal flu-like infections, they can also cause outbreaks of hand, foot and mouth disease (HFMD) and in some cases also life-threatening conditions including myocarditis and aseptic meningitis (usually caused by one of the B serotypes).

Our project focused on the synthesis of potential inhibitors of Coxsackievirus replication was based on observation that several 6-chloropurine nucleosides possess significant antiviral activity against CVB3 and CVB4. Our further investigation proved, however, that this activity of 6-chloropurine derivatives is not necessarily connected to the nucleoside-like skeleton and the sugar moiety can be substantially modified without any loss of antiviral potency. We used 9-norbornylpurine derivative **1** (Šála et al., *Bioorg. Med. Chem.* 2010, 18, 4374) as an ideal model structure for our further studies because of its excellent chemical stability and significant antiviral activity, although a number of other variously 9-substituted-6-chloropurine derivatives exert a significant anti-CVB3 potency (proved on a series of more than 30 compounds with bicyclic and polycyclic substituents, Dejmek et al. *Arch. Pharm. Chem. Life Sci.* 2014, 487)).

Since the mechanism of action was unknown, we had to systematically investigate the structure-activity relationship. Our preliminary results showed that there were a number of modifications of the skeleton, which could lead to enhanced antiviral activity (Figure 2). Firstly, we have made an extensive study of almost 40 purine base modifications revealing that a simple 6-chloropurine derivative **1** is the most active compound of the series, although the introduction of small groups into the position C-2 (e.g., **2**) and C-8 (OH or OCH_3) could lead to the desired drop of hydrophobicity without a significant loss of activity (Šála et al., *Bioorg. Med. Chem. Lett.*, 2011, 21, 4271).

Figure 2.



EC_{50}=0.81 μM, CC_{50} >50 μM

EC_{50}=1.0 μM, CC_{50} >50 μM

EC_{50}= 6.2 μM, CC_{50} >50 μM

Furthermore, this investigation suggested that the chlorine atom at the C-6 position can be substituted with a specific aromatic ring with the retention of anti-CVB3 activity. Therefore, we introduced a plethora of aromatic and heteroaromatic rings into this position and obtained a series of 35 derivatives. All the compounds were prepared by a Suzuki or Negishi cross coupling reaction using the derivative **1** as a precursor. However, only a couple of compounds exerted anti-CVB3 activity on a micromolar level (for instance compound **3**). On the other hand, several derivatives exerted significant activity against Chikungunya virus, an insect-borne virus with similar symptoms as Dengue fever (unpublished results). Second, the positions C-5' and C-6' obviously offer a wide space for further modifications as the derivative with an annulated benzene ring in these positions (**4**) presents the most active compound we have ever obtained. Therefore, we decided to modify these positions by the introduction of various substituents or by annulation of heterocyclic rings other than benzene (Figure 3). We prepared a large series of derivatives bearing various hydrophilic and hydrophobic side chains attached at these two positions (**5**) either by ether or ester bond (Šála et al., *Bioorg. Med. Chem. Lett.* 2012, 22, 1963). Furthermore, we studied synthetic approaches towards thiophenyl derivatives such as **6** and **7** (Hřebabecký et al., *Tetrahedron* 2012, 68, 3195). Although a number of derivatives from these two series exerted significant antiviral effect against CVB3, none of them showed activity on a submicromolar level. Several thiophenyl and tetrahydrothiophenyl analogues exerted significant cytostatic activity on CCRF-CEM and/or HL60 leukemia cell lines.

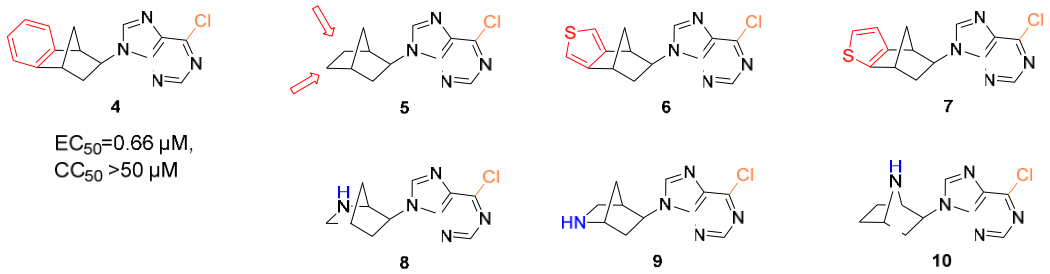


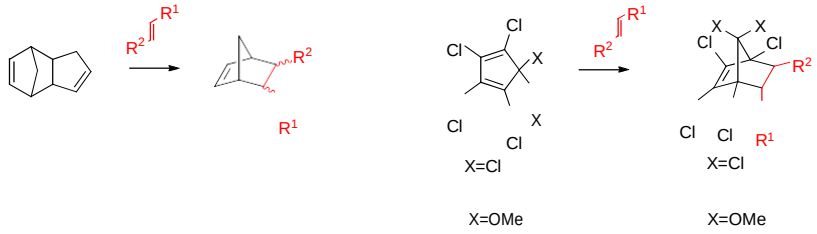
Figure 3.

Next, the effect of the heteroatom incorporation at the positions C-5', C-6' or C-7' was investigated. We explored various synthetic pathways towards derivatives containing a nitrogen atom in all mentioned position and analogues with oxygen inserted at C-6'. The obtained azanorbornyl derivatives **8**, **9** and **10** were not only evaluated for antiviral activity but, as analogues of non-opioid anesthetic epibatidine, we explored their potential as ligands for nicotine acetylcholine receptors (nAChRs) as well. It is noteworthy that, unlike epibatidine, the compounds act as antagonists of muscular nAChRs rather than agonists of neuronal type of such receptors (Hřebabecký et al., *Tetrahedron* 2012, 68, 1286). Although the exact mechanism of action of all these compounds is not clear yet, we have performed an extensive study (in collaboration with group of Dr. Mertlíková-Kaiserová) aiming on transport mechanism and metabolism of these compounds in order to elucidate this process (Plačková et al. *J. Enzyme Inhib. Med. Chem.*, 2015, 57).

Microwave assisted solvent-free synthesis of norbornene analogues

We introduced a novel technique for a Diels-Alder reaction of widely used cyclopentadiene derivatives with various dienophiles. This method excels not only in reaction times and yields

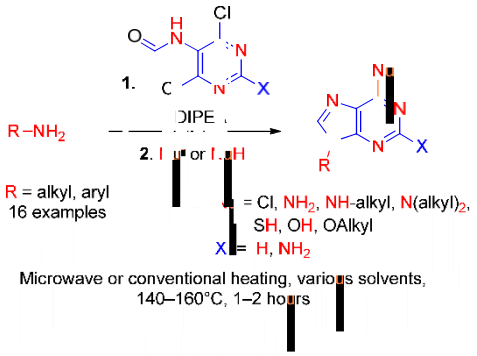
but also in purity of the crude products and simplicity of their isolation due to the use of solvent-free conditions and microwave irradiation (Scheme 1).
Furthermore, the reactions of dicyclopentadiene may be carried out at temperatures lower than the temperature required for cyclopentadiene preparation, which allows the Diels-Alder reactions even with low-boiling dienophiles under only low pressure. This method proved equally useful for multi-gram syntheses of the desired compounds (Dejmek et al., *Synthesis* 2011, 4077). We utilized this methodology in a number of cases for the preparation of suitable precursors for both above mentioned projects.



Scheme 1. Solvent-free Diels Alder reaction under microwave irradiation.

One-pot build-up procedure for the synthesis of variously substituted purine derivatives

We developed a significantly simplified approach to 9-substituted 6-chloropurine and 2-amino-6-chloropurine derivatives, which could be further converted to other, differently 6-substituted purines in one pot (Scheme 2). This is the first such technique allowing the preparation of 9-substituted purine derivatives from amine precursors in a single step, which makes it the method of choice for such transformations. We proved the usefulness of this methodology on the synthesis of a commercially successful antiviral drug Abacavir and showed its potential on a number of structurally diverse substrates (Dejmek et al. *RSC Advances* 2012, 6970).



Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Josef Michl - Organic Synthesis

Josef Michl is in charge of two research group of comparable size, one at the IOCB in Prague, Czech Republic, and the other at the University of Colorado in Boulder in the USA. He splits his time evenly between the two locations. In the period 2010 - 2014, the activities of the Prague team concentrated on four main areas:

(i) Preparation of Arrays of Molecular Rotors (effort split roughly evenly between Prague and Boulder)

(ii) Alkylation of Gold Surfaces (effort pursued exclusively in Prague)

(iii) Carborane Chemistry (about 90% in Prague)

(iv) Singlet Fission (about 50% in Prague)

Some additional projects are pursued only in Boulder (electronic structure and photophysics of oligosilanes, polymerization catalysis by lithium cations) and are of no interest for the present report. The total number of publications in the years 2010 - 2014 is 60 and three patents, but 20 of these publications report work done primarily or exclusively in Boulder. Several additional publications by the Prague team members have appeared without the name of Michl as coauthor.

In Prague, much of the work on singlet fission and occasionally some work on the other projects is performed collaboratively with Zdenek Havlas (IOCB), with whom Michl currently has a joint GACR grant on singlet fission. The work on alkylation of gold surfaces benefits from a collaboration with Ivan Stibor (TU Liberec), with whom Michl also has a joint GACR grant, and from collaboration with Zdenek Bastl (UFCH), who measures XPS on samples prepared in the Michl group. Studies of the redox behavior of carborane anions are done in collaboration with Jiří Ludvík (UFCH). Some electrochemistry is done jointly with Lubomír Pospíšil (IOCB and UFCH).

(i) Preparation of Arrays of Molecular Rotors

This project examines the feasibility of an unprecedented concept, preparation of regular two-dimensional flat arrays of artificial surface-mounted dipolar molecular rotors and control of their motion by the application of an outside electric field. Such arrays would represent an entirely new type of electroactive material with numerous possible scientific and technological applications in nanoelectronics and analog electronics in general.

We proposed and successfully demonstrated two complementary ways of accomplishing the goal: (i) Preparation and study of perforated surface rotor (PSR) arrays and (ii) synthesis and examination of self-assembled monolayer rotor (SMR) arrays. In a series of studies of gradually increasing sophistication, we have produced PSR arrays in which dipolar rotors are inserted into the surface of trigonally arranged channels in hexagonal tris-(*o*-phenylenedioxy)cyclotriphosphazene (TPP), and AMR arrays in which dipolar rotors are located above one or between two layers of triangularly interlocked triptycenes acting as a scaffold. In both cases, we achieved very low rotational barriers by a suitable structural choice for the rotating unit. We found that often but not always there is enough inherent flexibility in both types of dipolar rotor arrays that the unfavorable electrostatic interactions associated with the trigonal arrangement of the rotor dipoles introduce defects and cause a

distortion of the lattice from hexagonal to lower symmetry. As a result, the ground state is not ferroelectric, which would be the most interesting outcome and would be of considerable importance for fundamental physical studies of two-dimensional collective behavior and for future possible practical use. At present, we are looking for ways to minimize the number of defects and achieve ferroelectricity.

Theoretical and computational work on molecular rotors has guided much of the effort. Specifically, the magnitude and nature of friction on the molecular scale was elucidated. This is very important for any future efforts to fabricate practical devices containing molecular rotors.

The work on PSR arrays led to an unexpected discovery that appears worth further pursuit. We noted that the insertion of electron acceptors into the channels of hexagonal TPP produces materials suitable for charge transfer, either spontaneous if the acceptor is strong (quinone) or photoinduced if it is weak (tetracene). In these materials, which can be prepared in the form of very thin layers with TPP channels running perpendicular to the surface, negative polarons collect in the channels and positive polarons collect in TPP layers. Both are expected to have anisotropic mobility and conductivity. These materials need to be investigated in more detail, as they hold promise as absorbers of sunlight that permit exciton and charge transport management. Their oxidation-reduction properties could be used to catalyze chemical reactions and for interfacing to other molecules and catalysts. These properties could be tailored and the detailed architecture of these inclusion compounds could be controlled from molecular scale to the mesoscale.

(ii) Alkylation of Gold Surfaces

We discovered a new way of attaching alkyl residues to gold surfaces by direct C-Au bonds in solution under ambient conditions using main group element organometallics. Organomercurials work best but are unattractive for practical applications because of their poisonous nature. They have the advantage that elemental mercury, which is codeposited with the alkyls, can be easily removed either by heating or by electrochemical oxidation, yielding a gold surface coated only with alkyls. The much less poisonous organostannanes work very well, too, and it now appears that certain derivatives of germanium, a non-poisonous element, can also be used. This discovery has been protected by a patent and could have an immediate technological impact in nanoelectronics and other areas of nanoscience.

The scope of the alkylation with stannanes has been examined in considerable detail. A methyl and an ethyl group are never transferred from the tin atom to the gold surface, whereas longer alkyl groups are. In most but not all cases, tin oxide is also deposited on the gold. The reaction appears to proceed by a complicated mechanism, since the rate of growth of surface coverage is proportional to the square root of the bulk stannane concentration. The resulting monolayers are disordered and do not block the gold electrode very much. This suggests that they might be useful for the coating of gold nanoparticles that is sufficient to prevent their aggregation yet insufficient to block the access of reagents to the catalytic gold surface, and this possibility is currently under investigation.

(iii) Carborane Chemistry

The primary objective of this project is the development of the chemistry of highly oxidized states of matter. Specifically, we work with substituted monocarba-*closo*-dodecaborane anions and stable radicals with the ultimate goal of producing the world's strongest reversible neutral oxidants suitable for applications in electrochemistry (high- voltage batteries) and for the preparation of hitherto inaccessible highly oxidized species. The experimental work is closely coupled to quantum chemical computations. We have built a vacuum line for work with elemental fluorine and use it to introduce trifluoromethyl groups and fluorine substituents on the monocarba-*closo*-dodecaborane cage. The stable radical

$HCB_{11}F_5(CF_3)_6$ has remarkable oxidizing power (calculated redox potential 4 V above ferrocene/ferricenium!) and is reversibly interconverted with the anion $HCB_{11}F_5(CF_3)_6$. This is one of the least nucleophilic and least oxidizable anions known, suitable to serve as a stable counterion for extremely aggressive cations.

The practical utilization of the new powerful oxidants requires access to solvents that are inert under strongly oxidizing conditions and capable of dissolving both the radicals and the salts of the anions; the search for such solvents represents another important aspect of the project.

(iv) Singlet Fission

Increasing the efficiency of photovoltaic solar cells while reducing their cost is some of the most pressing tasks of the day. Singlet fission is a photophysical process that promises to allow the theoretical efficiency of single-junction solar cells to surpass the Shockley-Queisser limit of 1/3 by a significant factor of about 1.4. In this process, a singlet excited state of a chromophore shares some of its energy with the ground state of a neighboring chromophore, and both chromophores end up in their excited triplet states. If each of the triplets then produces an electron-hole pair, one produces two electrons and two holes from a single photon of sufficient energy, thus doubling the current, albeit at half the voltage. If one also harvests those photons whose energy is insufficient for singlet fission in an underlying layer containing an ordinary sensitizer, the overall theoretical efficiency of the cell reaches nearly 1/2.

The reason why no singlet fission solar cells are on the market today is the tiny number of known materials that perform singlet fission efficiently. Most are highly sensitive to oxygen in the presence of light and none of them appear to be practical for inexpensive solar cells. Our efforts are directed toward the development of simple structural design principles that would permit the synthesis of sensitizers that perform singlet fission efficiently. Once a large number of such materials are known, it will be possible to optimize their other essential properties, such as redox potentials, absorption spectra, light fastness, etc. We started with the principles of quantum chemistry and used them to identify two classes of chromophores likely to have the necessary relations between the energies of low- lying singlet and triplet states for singlet fission to be mildly exothermic and thus fast and competitive with all other processes that depopulate the initially reached singlet excited state. We have proposed a series of specific structures suitable for the purpose and are now synthesizing them. They are very different from the few structures that are already known to perform singlet fission efficiently and promise to open new areas to exploration. Many, for example certain derivatives of indigo or aminoquinones, appear to be very sturdy and practically useful.

The next task is to position the selected chromophores next to each other in a way that promotes interactions conducive to fast singlet fission. In some crystalline materials this happens spontaneously as dictated by the crystal structures, and we were fortunate that this was the case in the first material ever designed with singlet fission in mind, 1,3-diphenylisobenzofuran. In others, suitable positioning has to be forced by crystal engineering or by the synthesis of dimers or higher oligomers. First, however, one needs to know which mutual positions of the two chromophores are optimal. Using the principles of quantum chemistry, we have developed a simple model that provides the answer and are currently testing it experimentally.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Petr Beier - Organic Chemistry of Fluorine, Phosphorus, Sulfur and Silicon

Our research is focused on synthetic organic chemistry of main group elements such as fluorine, sulfur, phosphorus, and silicon. We develop methodologies for the synthesis of novel compound with interesting physicochemical properties. The use of modern analytical and instrumentation methods helps us to unravel reaction mechanisms and discover new reactivity. Our research lies at the interface of organic chemistry, physical chemistry and biochemistry.
During the years of 2010-2014 the research was conductive in several areas: 1. The chemistry of fluorinated phosphonates, 2. Sulfur-based reagents for tetrafluoroethyl and tetrafluoroethylene group transfer, 3. Synthesis and reactivity of (pentafluorosulfanyl)benzenes, 4. Investigation of properties of perfluorocarbon solvents, and 5. Synthesis and modification of plant growth regulators.

1. The chemistry of fluorinated phosphonates

α-Fluorinated phosphonates are, in terms of their function and bioactivity, important non-hydrolyzable mimics of natural phosphates. α-Fluorinated phosphonates are often utilized as enzyme inhibitors and metabolic probes. We have focused on using fluorinated phosphonates in synthetic chemistry as radical or carbanionic synthons for the transfer of fluorinated groups**.**
It was found that diethyl trifluoromethylphosphonate affects trifluoromethylation of non-enolizable ketones to trifluoromethyl-containing alcohols. With a suitable nucleophilic initiator it also reacted with aldehydes to afford phosphonates of trifluoromethylated alcohols or with disulfides or diselenides to afford trifluoromethythio and trifluoromethylseleno compounds, respectively (*Tetrahedron Lett.* **2010**, *51*, 252).
We have discovered a new process for the synthesis of (*E*)-α-fluorovinylphosphonates from α,α-difluorophosphonates in a highly stereoselective manner (*Synlett* **2011**, 331).
Currently, several drugs for the treatment of various bone diseases are based on the hydroxymethylenebisphosphonate structure. The biological activity is almost completely retained upon substitution of the hydroxyl for fluorine atom. We have shown a new synthesis of α-alkyl-α-fluoromethylenebisphosphonates based on alkylation of easily accessible fluoromethylenebisphosphonate (*Org. Biomol. Chem.* **2011**, *9*, 4035). This work was done in collaboration with the group of Prof. S. Prakash, USC, Los Angeles, CA, USA, both teams contributed equally in this project. Fluoromethylenebisphosphonate was also found to undergo conjugate addition to Michael acceptors leading to a further chemical diversity of the resulting biologically relevant α-alkyl-α-fluoromethylenebisphosphonates (*J. Fluorine Chem.* **2011**, *132*, 363).
Lithium salt of diethyl difluoromethylphosphonate was found to undergo 1,4-addition

to α,β-unsaturated ketones and other Michael acceptors. Factors influencing 1,2- versus 1,4-addition were studies (*J. Fluorine Chem.* **2012**, *137*, 34). The same lithium salt reacted with imines to provide important N-substituted α,α-difluoro-β-aminophosphonates (*J. Fluorine Chem.* **2012**, *141*, 76).
Another important fluorinated phosphonate – diethyl 1-fluoro-1-phenylsulfonylmethanephosphonate – underwent conjugate addition to Michael acceptors and α,β-unsaturated compounds. In case of α,β-enones a new type of phosphonate to phosphate rearrangement was discovered (*J. Org. Chem.* **2013**, *78*, 4573).
Previously unknown diethyl fluoronitromethylphosphonate was synthetized, it's pKa of the α-hydrogen atom and the pKa's values of related compounds were investigated computationally in collaboration with the group of Dr. L. Rulíšek. The new reagent was found to undergo a new type of decomposition pathway induced by base or fluoride ions. More importantly, the reagent proved to be very useful in nucleophilic additions to carbonyl compounds, conjugated additions, and alkylations affording structurally diverse new fluoronitroolefins and fluoronitrophosphonates (*Chem. Eur. J.* **2014**, *20*, 1453).
The chemistry of fluorinated phosphonates was reviewed by us (*Chem. Listy* **2014**, *108*, 923).

2. Sulfur-based reagents for tetrafluoroethyl and tetrafluoroethylene group transfer

We have developed several sulfur-based reagents for radical or nucleophilic CF_2CF_2 group transfer starting from 1,2-dibromotetrafluoroethane. The research in this area was inspired by the fact that while there are number of methods for the introduction of perfluoroalkyl, difluoromethyl and $(CF_2)_n$ where $n > 2$ groups, the methods for the introduction of CF_2CF_2 or CF_2CF_2H groups are very scarce and suffer from low yields and low substrate scope. Our reagents are easily accessible and allow efficient transformation to these functional groups on diverse substrates. For example, $PhSCF_2CF_2Br$ acted as a double radical synthon in additions to alkenes (*J. Fluorine Chem.* **2013**, *156*, 307), $PhSCF_2CF_2SiMe_3$ acted as a tandem radical-anion synthon in addition to carbonyl compounds (*Eur. J. Org. Chem.* **2011**, 4528) and N-substituted cyclic imines (*Synlett* **2012**, 1187). We have received a grant from the Czech Science Foundation supporting this research (2011-2013, P207/11/0421).

3. Synthesis and reactivity of (pentafluorosulfanyl)benzenes)

Only a few years ago, compounds with the pentafluorosulfanyl group (SF_5) have been considered as chemical curiosities despite their great stability and very interesting physical properties. This was largely due to the fact that even basic building blocks with this group were not available. In 2010 we started working in this area with the aim of improving availability of these compounds and showing synthetic access to new classes of compounds with this functional group in order to speed up the development of applications in agrochemistry, drug development and material science. During several years, we have succeeded in fulfilling this goal to a large extent. Currently, about half of all commercially available compounds with the SF_5 group come from our laboratory and we have established a business partnership with the Spirochem AG company. A big part of the know-how and intellectual property was protected by our Institute (two patents) and we are currently in the process of transferring the technology to another company for further development and scale up. Concerning the science, most of the work is based on derivatization of commercial nitro(pentafluorosulfanyl)benzenes through nucleophilic aromatic

substitution for the nitro group (*Org. Lett.* **2011**, *13*, 1466), nucleophilic aromatic substitution for hydrogen – vicarious (*Tetrahedron Lett.* **2011**, *52*, 4392; *J. Org. Chem.* **2011**, *76*, 4781; *Eur. J. Org. Chem.* **2012**, 2123), oxidative (*J. Fluorine Chem.* **2012**, *143*, 130), or other (*Beilstein J. Org. Chem.* **2013**, *9,* 411), and further chemical modifications toward the synthesis of alkenyl-SF_5-benzenes (*Beilstein J. Org. Chem.* **2012**, *8*, 1185), and pharmaceutically important indoles and oxindoles (*Synlett* **2013**, 855). A grant from the Czech Science Foundation supported this research (2012-2014, P207/12/0072).

In collaboration with the group of Prof. G.-V. Röschenthaler, Bremen, Germany (DFG starting collaboration grant) we have looked at the two step synthesis of (pentafluorosulfanyl)benzenes and found an improved conditions for safe, atom economic, and efficient fluorination method of arylsulfur chlorotetrafluorides avoiding anhydrous hydrogen fluoride (*J. Fluorine Chem.* **2014**, *157*, 79). In this work, our group contributed with majority of the work including conceiving the idea, performing most of the experimental work and writing the research article.

In collaboration with Dr. Cormac Murphy, Dublin, Ireland, we looked at microbial transformations of some compounds with the SF_5 group (*Environ. Sci. Pollut. Res.* **2014**, *21*, 753). Here we have contributed only by the supplying the substrates for the metabolic study and discussion of the results.

A conceptually novel access to aliphatic SF_5 compounds by selective oxidation of some SF_5 aromatics was disclosed by us recently (*J. Org. Chem.* **2014**, *79*, 8906). Interestingly, the observed aliphatic products obtained using rather harsh chemical oxidations resemble products of aerobic metabolic oxidation of aromatics.

4. Investigation of properties of perfluorocarbon solvents

Perfluorocarbons have intriguing properties including unusual miscibility behaviour, hydrophobic and oleophobic character, high density, and low surface tension. As a continuation of our earlier report of surprising miscibility behaviour of perfluorocarbons with ethers (*J. Fluorine Chem.* **2008**, *129*, 397) we have, in collaboration with Prof. K. Řehák, ICT, Prague measured liquid-liquid equilibria of some perfluorocarbons in the ternary system with water (or methanol) and hexafluoroisopropyl alcohol (*J. Chem. Eng. Data* **2014**, *59*, 3510). Interesting results were obtained which have implication in fluorous biphase separation techniques using these solvents. Our group contributed in conceiving the idea, supplying the material and partially writing the article.

5. Synthesis and modification of plant growth regulators

Karrikins (a group of butenolides first isolated from the smoke of burning vegetation) are known as incredibly potent germination promotors. The most active among them is 3-methyl-2*H*-furo[2,3-*c*]pyran-2-one (KAR1), with activity at 10^{-9} M concentration. The smoke also contains the 3,4,5-trimethylfuran-2(5*H*)-one (TMB), which inhibits the germination of seeds and reduced the effect of Karrikins. We have joined this multidisciplinary project involving chemists, plant physiologists and molecular biologists in 2011. This complex research is focused on the study of mode of action of the above mentioned compounds, their effect to germination of seeds and growth of young plants. The study is done in collaboration with the group of Prof. Van Staden, University KwaZulu-Natal, Pietermaritzburg, South Africa, Dr. V. Soós and Prof. E. Balázs, Agriculture Research Institute, Martonvásár, Hungary and Dr. Katherine S. Downes, Department of Enviroment and Agriculture, Curtin University, Australia. Our group is focused mainly on the synthesis of naturally occurring smoke-

derived plant growth regulators in order to support teams carrying out molecular biology and plant physiology research. Our own research is aimed at the development of novel efficient synthetic protocols toward synthesis of these regulators, preparation of synthetic analogues of TMB including its molecular probes and metabolic intermediates.

Together with the group of Prof. Balász, we studied molecular aspects of the antagonistic interaction of KAR1 and TMB. It has been found that KAR1 and TMB are not direct competitors and despite their structural similarity they do not bind to the same receptor (*New Phytologist* **2012**, *196*, 1060). Our group has contributed to this work by the synthesis of KAR1, TMB, and ^{13}C labeled TMB and we also participated in the preparation of the manuscript.

With Dr. Downes we have focused on the study of effect of smoke-derived germination regulators on the seeds of various plant species (*Annals of Botany* **2013**, *111*, 489, and *Austr. J. Bot.* **2014**, *62*, 347). For this project, our group synthesized KAR1, TMB and also glyceronitrile, another germination stimulant present in the smoke. We also participated on the preparation of manuscripts.

We are collaborating with the group of Prof. Van Staden on the research towards evaluation of germination inhibitory activity of synthetic analogues of TMB through structure-activity relationship study (*J. Plant Physiol.* **2013**, *170*, 1235). Our group contributed in conceiving the idea, we have developed the synthetic protocol, synthesized analogues and prepared major part of the article. Together, we also reported X-ray structures of several Karrikins, including their synthetic sulfur analogue 3-methyl-2H-thiopyrano[3,4-*b*]furan-2-one (*South Afr. J. Bot.* **2013**, *88*, 107and *South Afr. J. Bot.* **2014**, *91*, 53). Our contribution was the synthesis of karrikins, growing single-crystal and preparation of manuscript. We have demonstrated that TMB is non-toxic (*Mutat. Res.* **2015**, *778*, 1) and therefore it represents an ecological alternative to commonly used herbicides. Our group contributed to this work only by the supplying the smoke-derived germination regulators.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Ivo Starý – Helikální aromáty, funkční π–elektronové systémy

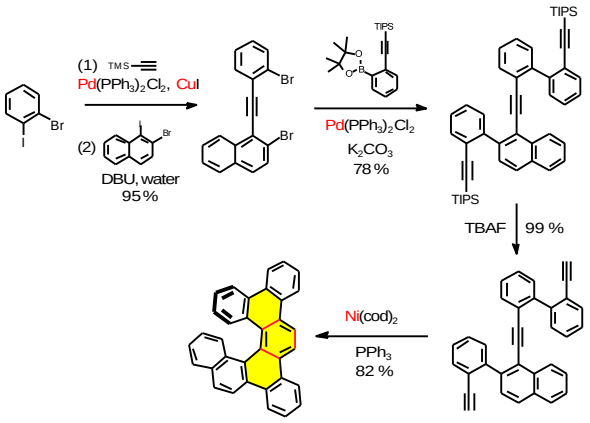
In the period of 2010-2014, our research activities were focused mainly on the design, synthesis, property studies and exploitation of advanced π-electron systems using both experimental and theoretical approaches. Chirality aspects were the golden thread of our endeavour. We strived for fulfilling a vision of multidisciplinary research stemming from our expertise in modern organic synthesis and tackling challenging problems in chemistry and physics provided sophisticated molecules could play a central role in their solutions. There are numerous reasons why to deal with the advanced π-electron systems but one of them is in particular attractive: The structural and functional diversity as well as complexity of man-made π-electron molecules (including aromatics) is enormous overwhelming in that natural resources. Along with the further development in chemistry of helically chiral aromatics, which represents the flag ship of our research, we used the past period also to identify new and attractive niches in the research. As required methodologies and instrumentation went beyond our expertise or in-house available equipment, respectively, we initiated the interdisciplinary collaboration with physicists.

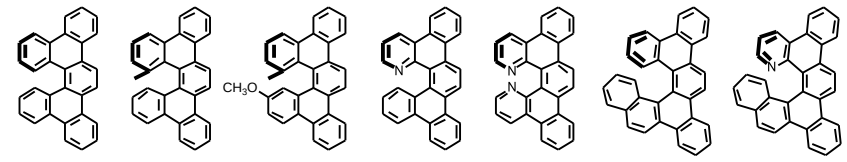
Between years 2010 and 2014 we focused mainly (not exclusively) on the following tasks, which we report on herein:

1 The development of synthetic methodologies for an easy access to functionalised helical (hetero)aromatics in a racemic and/or optically pure form

1.1 The development of a methodology for the rapid synthesis of dibenzohelicenes and their functionalized derivatives

Visions of applying unique helically chiral π-electron systems in diverse areas of chemistry and physics have fuelled the intensifying research aiming at the development of a practical and general methodology for the preparation of functionalised helicenes and their heteroanalogues. Despite the fact, that significant progress has recently been achieved in this regard, there is plenty of room to substantially shorten, simplify and generalise the synthesis of helicenes. If such an endeavour meets with success, then an on-demand preparation of tailor-made helicenes would not be ultimately a limiting step for their wider exploitation. We have developed a general approach to dibenzo[5]-, dibenzo[6]- and dibenzo[7]helicenes as well as their functionalised derivatives.[1] These helically chiral aromatics can usually be synthesised within four to six operations in overall yields ranging from 24% to 67% by employing a short sequence of reliable processes such as Sonogashira coupling, Suzuki-Miyaura coupling, desilylation and [2+2+2] alkyne cycloisomerisation. There are only scattered examples of dibenzohelicenes described in the literature and this study on their preparation and properties is so far the most comprehensive. Dibenzohelicenes have the advantage over the parent helicenes in the simplicity of their nonphotochemical preparation and, therefore, they have a potential to mimic or even substitute parent helicenes in the envisaged applications. Moreover, we have demonstrated one of the highest enantiocontrol in the asymmetric synthesis of helicenes by transition metal-catalysed [2+2+2] alkyne cycloisomerisation and manifested a straightforward approach to the optically pure paradigmatic dibenzo[6]helicene.





98 % (4 steps, 63 %) | 90 % (4 steps, 33 %) | 76 % (5 steps, 51 %) | 73 % (4 steps, 50 %) | 88 % (4 steps, 40 %) | 82 % (5 steps, 63 %) | 71 % (5 steps, 55 %)

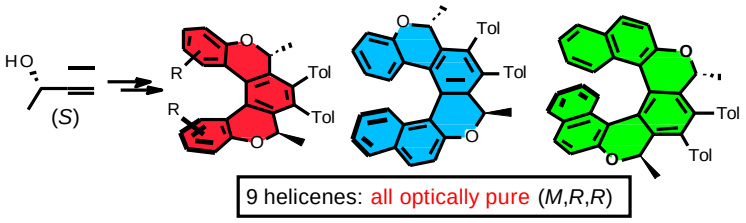
The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. A.J., J.R. and K.C. performed the synthesis and characterisation. J.V.C. and J.V. performed the quantum chemical calculations. Collaborators: R.P., L.B. and P.F. carried out the NMR, ECD and IR spectroscopic analyses. I.C. performed the single-crystal analyses.

---

[1] Jančařík, A.; Rybáček, J.; Cocq, K.; Vacek Chocholoušová, J.; Vacek, J.; Pohl, R.; Bednárová, L.; Fiedler, P.; Císařová, I.; Stará, I. G.; Starý, I. Rapid Access to Dibenzohelicenes and Their Functionalized Derivatives. *Angew. Chem. Int. Ed.* **2013**, *52*, 9970–9975.

1.2 The development of a general approach to optically pure [5]-, [6]-, and [7]heterohelicenes

The lack of a general methodology for the effective synthesis of nonracemic helicenes and their analogs has been a major hurdle that has limited a wider exploitation of these helically chiral aromatic systems in enantioselective catalysis, molecular recognition, self-assembly, surface science, chiral materials, and other branches of science. Ideally, a practical asymmetric synthesis should be independent of both the length of the helical backbone and the presence of functional groups. Since the pioneering studies by Martin *et al*. and Katz *et al*., who successfully used diastereoselective photodehydrocyclization of stilbene-type precursors, various concepts of the asymmetric synthesis of helicenes have been explored but no general protocol for obtaining optically pure helicenes or their analogs with a wide structural diversity has yet been reported. We have developed a general methodology for the preparation of optically pure [5]-, [6]-, and [7]heterohelicenes for the first time.[2] It is based on a Co^{I}- or Ni^{0}-catalyzed diastereoselective [2+2+2] cycloisomerization of centrally chiral triynes to deliver helicenes comprising two 2*H*- pyran rings. The major advantage of this methodology is that the stereochemical outcome (*dr* uniformly 100:0) depends neither on the helicene length nor the functional group(s) present. Moreover, the synthesized 2*H*-pyran [5]heterohelicenes exist as single helices even at higher temperature (in contrast to the parent [5]helicene that racemizes at room temperature) and both enantiomers of but-3-yn-2-ol (a key chiral building block) are commercially available. Finally, the helicity of the products can be easily predicted computationally. The 2*H*-pyran-modified helicenes might widely be applied to, *e.g*., enantioselective catalysis.



The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. J.Ž., A.J., A.A. and M.Š. performed the synthesis and characterisation. J.V.C. and J.V. performed the quantum chemical calculations. Collaborators: R.P. and D.Š. carried out the NMR spectroscopic analyses. I.C. performed the single-crystal analyses.

1.3 A new reaction discovered: Cyclisation of one alkyne and two nitriles to pyridazines

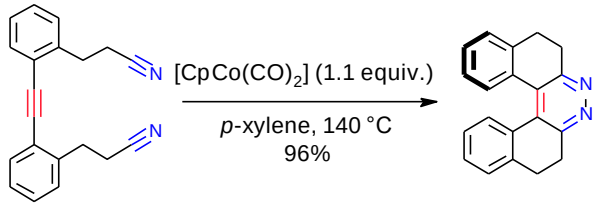
Transition--metal-catalysed cyclisation of alkynes to form benzene derivatives has attained a considerable attention due to its exceptional synthetic usefulness and versatility. The value of this methodology has been further underscored by the introduction of the cyclisation of nitriles with two alkynes thus enabling an effective *de novo* synthesis of pyridine derivatives. However, the complementary cyclisation of one alkyne and two nitrile units to afford pyridazine derivatives has not yet been described until the recent report on this new reaction by our group[3] and independent

---

[2] Žádný, J.; Jančařík, A.; Andronova, A.; Šámal, M.; Vacek Chocholoušová, J.; Vacek, J.; Pohl, R.; Šaman, D.; Císařová, I.; Stará, I. G.; Starý, I. A General Approach to Optically Pure [5]-, [6]-, and [7]Heterohelicenes. *Angew. Chem. Int. Ed*. **2012**, *51*, 5857–5861.

[3] Chercheja, S.; Klívar, J.; Jančařík, A.; Rybáček, J.; Salzl, S.; Tarábek, L.; Pospíšil, L.; Vacek Chocholoušová, J.; Vacek, J.; Pohl, R.; Císařová, I.; Starý, I.; Stará, I. G. The Use of Cobalt-Mediated

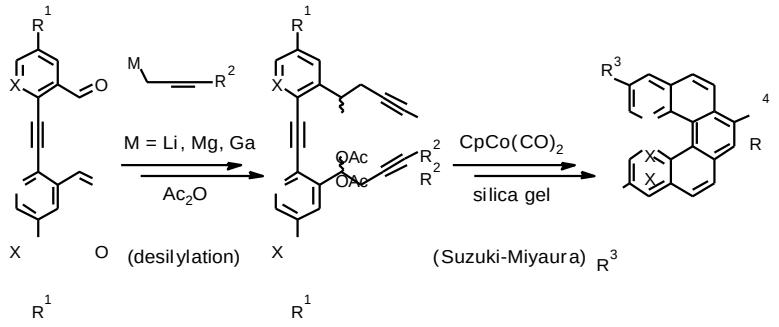
observation by Snyder *et al*.,[4] both published in 2014. We have found that the formation of pyridazines (in our case embedded in helicene scaffolds) by cobalt-mediated intramolecular cycloisomerisation of ynedinitriles is feasible and proceeds in good to high yields. The construction of the pyridazine nucleus from one alkyne and two nitrile units is proposed to follow either a conventional organometallic mechanism or to be triggered by a single-electron transfer from a Co^{II} species. We applied this synthetic methodology to the preparation of a series of helical pyridazines including [5]-, [6]- and [7]helicene derivatives. We showed by DFT calculations that cyclisation of a representative ynedinitrile to the corresponding helicene pyridazine is an exergonic reaction although being less downhill in energy than that of the analogous triyne or diynenitrile. We believe that this new cyclisation reaction might develop into a useful tool for the preparation of complex pyridazines, the importance of which has already been noticed.



The team: I.S. and I.G.S. planned the project. I.S., I.G.S. and S.C. designed the compounds. S.C., J.K., A.J., J.R. and S.S. performed the synthesis and characterisation. J.V.C. and J.V. performed the quantum chemical calculations. I.S. and I.G.S. supervised the experiments and wrote the manuscript. Collaborators: J.T. and R.P. carried out the EPR and NMR spectroscopic analyses. L.P. performed the electrochemical analyses. I.C. performed the single-crystal analyses.

1.4 A versatile synthesis of functionalized pentahelicenes

Although remarkable progress in the preparation of helicenes has recently been accomplished, the modular synthesis of their functionalized derivatives or heteroanalogues remains a challenging task. A general synthetic methodology for the nonphotochemical preparation of functionalized (hetero)helicenes has been developed.[5] It employs the sequence of a double propargyl organometallics (Li, Mg, Ga/In) addition to a tolan-2,2'-dialdehyde-type intermediate, a cobalt-catalyzed [2+2+2] cycloisomerization of a triyne intermediate, and a double silica gel-assisted acetic acid elimination to receive pentahelicene, 1,14-diazapentahelicene, 3,12-dichloro-, 3,12-dichloro-7-trimethylsilyl-, and 3,12-di-*t*-butylpentahelicene. 3,12-Dichloropentahelicene undergoes a Suzuki-Miyaura coupling with aryl boronic acids (or ester) under palladium catalysis to afford 3,12-diarylpentahelicenes representing the first



X = CH or N; R^1 = H, Cl or *t*-Bu; R^2 = H, TIPS or TMS; R^3 = H, Cl, *t*-Bu or aryl; R^4 = H or TMS

Cycloisomerisation of Ynedinitriles in the Synthesis of Pyridazinohelicenes. *Chem. Eur. J*. **2014**, *20*, 8477–8482.

[4] Snyder, J. K.; Cai, C.; Audet, M. A. *N*,*N*-Bond Formation in Intramolecular Cobalt-Catalyzed [2+2+2] Cyclizations of Alkynyl-Linked Bisnitriles, and the Preparation of Annulated Pyridazines. *Heterocycles* **2014**, *88*, 179–186.

[5] Songis, O.; Míšek, J.; Schmid, M. B.; Kollárovič, A.; Stará, I. G.; Šaman, D.; Císařová, I.; Starý, I. A Versatile Synthesis of Functionalized Pentahelicenes. *J. Org. Chem.* **2010**, *75*, 6889–6899.

successful reaction of this type in helicene chemistry. Thus, we have demonstrated a practical methodology for the preparation of various functionalized pentahelicenes, which can serve as a paradigm for the synthesis of higher or differently functionalized helicenes.

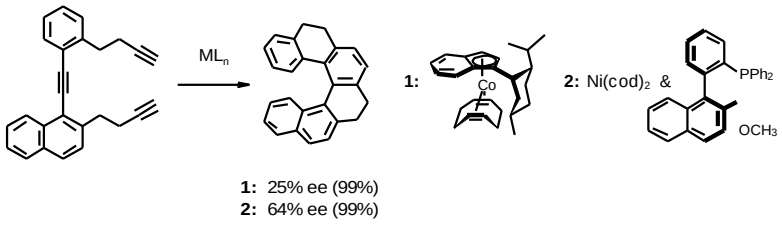
The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. O.S., J.M., M.S. and A.K. performed the synthesis and characterisation. Collaborators: D.Š. carried out the NMR spectroscopic analyses. I.C. performed the single-crystal analyses.

1.5 Chiral Co^{I} and Ni^{0} complexes in the synthesis of nonracemic helicenes through the enantioselective [2+2+2] cyclotrimerisation of alkynes

The enantioselective intramolecular [2+2+2] cycloisomerisation of triynes under catalysis by chiral transition-metal complexes (Co^{I}, Ni^{0}) in order to receive nonracemic helicene derivatives has been explored.[6] The use of the chiral neomenthylindene Co^{I} complex led to a moderate 25% ee of tetrahydro[6]helicene, which was the first example of such a reaction catalysed by the chiral Co^{I} complex. The alternative Ni^{0} catalysis employing privileged axially chiral monophosphines such as (-)-(*aS*)-NAPHEP led to tetrahydro[6]helicene with 64% ee, which is among the highest enantiomeric excesses so far observed for this Ni^{0}-catalysed reaction. It is proposed that stereodiscrimination takes place in an early stage of the catalytic cycle at the point of the formation of the diastereomeric bisalkyne complexes from the chiral metal catalyst and achiral triyne.

1.5 Chiral Co^{I} and Ni^{0} complexes in the synthesis of nonracemic helicenes through the enantioselective [2+2+2] cyclotrimerisation of alkynes

The enantioselective intramolecular [2+2+2] cycloisomerisation of triynes under catalysis by chiral transition-metal complexes (Co^{I}, Ni^{0}) in order to receive nonracemic helicene derivatives has been explored.[7] The use of the chiral neomenthylindene Co^{I} complex led to a moderate 25% ee of tetrahydro[6]helicene, which was the first example of such a reaction catalysed by the chiral Co^{I} complex. The alternative Ni^{0} catalysis employing privileged axially chiral monophosphines such as (-)-(*aS*)-NAPHEP led to tetrahydro[6]helicene with 64% ee, which is among the highest enantiomeric excesses so far observed for this Ni^{0}-catalysed reaction. It is proposed that stereodiscrimination takes place in an early stage of the catalytic cycle at the point of the formation of the diastereomeric bisalkyne complexes from the chiral metal catalyst and achiral triyne.



The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. A.A. performed the synthesis and characterisation. Collaborators: H.H. and M.H. designed the compounds and supervised the experiments. C.F. performed the synthesis and characterisation.

---

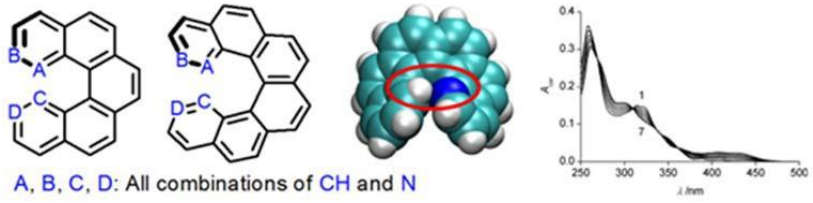
[6] Heller, B.; Hapke, M.; Fischer, C.; Andronova, A.; Starý, I.; Stará, I. G. Chiral cobalt^{I} and nickel^{0} Complexes in the Synthesis of Nonracemic Helicenes through the Enantioselective [2+2+2] Cyclotrimerisation of Alkynes. *J. Organomet. Chem.* **2013**, *723*, 98–102.

[7] Heller, B.; Hapke, M.; Fischer, C.; Andronova, A.; Starý, I.; Stará, I. G. Chiral cobalt^{I} and nickel^{0} Complexes in the Synthesis of Nonracemic Helicenes through the Enantioselective [2+2+2] Cyclotrimerisation of Alkynes. *J. Organomet. Chem.* **2013**, *723*, 98–102.

2 The experimental and theoretical studies on chemical and physicochemical properties of helical aromatics and their applications

2.1 A study on physico-chemical properties of pyridohelicenes

We have provided a detailed and comprehensive description of the physico-chemical properties of a series of pyrido[5]- and pyrido[6]helicenes having the nitrogen atom(s) in various positions.[8] Using both computational (DFT calculations) and experimental methods (kinetics measurements, UV-Vis titrations, DC and AC polarography, cyclic voltammetry and EPR spectroscopy), the (*P*)/(*M*) inversion barriers, protonation constants and redox behaviour of these helically chiral objects have been determined. Such knowledge, which is so-far scattered or missing in the literature, can be useful for planning practical applications of pyridohelicenes to, *e.g*., catalysis, organocatalysis and their asymmetric versions. The following conclusions can be drawn from the undertaken investigations: (a) the inversion barriers of pyridohelicenes and their protonated forms can be treated computationally with a good accuracy (usually within a 1 kcal/mol difference with respect to the experimental values) using the DFT (B3LYP/cc-pVTZ) method; (b) the presence of the nitrogen atom(s) in the innermost position(s) of pyridohelicenes (1 or 1,14 at pyrido[5]helicenes, 1 or 1,16 at pyrido[6]helicenes) significantly influences the free energy of activation for inversion depending upon their protonation state; (c) protonation constants of monoprotic pyridohelicenes in solution are comparable to that of pyridine within a single log unit, while the basicity of diprotic pyridohelicenes with the nitrogen atoms in the innermost positions is slightly higher than that of pyridine (about *ca*. two log units for 1,14-diazahelicene); (d) protonation constants of monoprotic pyridohelicenes can be calculated theoretically with a good accuracy (within a factor of 10 with respect to the experimental values), while calculated basicities of single-protonated diprotic pyridohelicenes with the nitrogen atoms in the innermost positions are overestimated by about two log units; (e) monoprotic pyridohelicenes are redox silent within a potential window of *ca*. –1.8 V to 1.1 V exhibiting otherwise two reduction peaks and a single oxidation one.

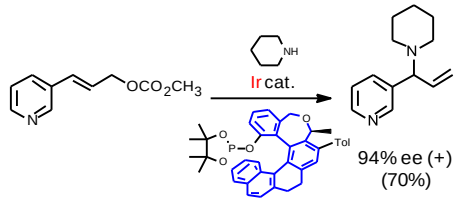


<u>The team:</u> I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. A.A., J.M., O.S. and M.Š. performed the synthesis. J.V.C. and J.V. performed the quantum chemical calculations. A.A. carried out the UV-Vis spectroscopic analyses. <u>Collaborators:</u> M.M. planned the project, supervised the experiments and co-wrote the manuscript. M.B. carried out the UV-Vis spectroscopic analyses. L.P. performed the electrochemical analyses.

[8] Vacek Chocholoušová, J.; Vacek, J.; Andronova, A.; Míšek, J.; Songis, O.; Šámal, M.; Stará, I. G.; Meyer, M.; Bourdillon, M.; Pospíšil, L.; Starý, I. On the Physicochemical Properties of Pyridohelicenes. *Chem. Eur. J.* **2014**, *20*, 877–893.

2.2 Helicene-based phosphite ligands in asymmetric transition-metal catalysis: Exploring Rh-catalysed hydroformylation and Ir-catalysed allylic amination

Starting from optically pure [6]helicene-like alcohol, four helical phosphites were prepared using corresponding chlorophosphites.[9] These ligands containing parent or substituted 1,3,2-dioxaphospholan-2-yl or dibenzo[*d*,*f*][1,3,2]dioxaphosphepin-6-yl moieties were applied to asymmetric hydroformylation of terminal alkenes catalysed by $Rh(acac)(CO)_2$ and asymmetric allylic amination of cinnamyl-type carbonates catalysed by $[Ir(cod)Cl]_2$. The helical phosphite containing the dibenzo[*d*,*f*][1,3,2]dioxaphosphepin-6-yl group was most successful in an asymmetric hydroformylation of styrene, leading to moderate enantiomeric excesses (up to 32% *ee*), high regioselectivity in favour of the branched product and mostly high conversion, whereas the helical ligand containing the 4,4,5,5-tetramethyl-1,3,2-dioxaphospholan-2-yl fragment was most effective in an asymmetric allylic amination, exhibiting high enantioselectivity (up to 94% *ee*), excellent regioselectivity in favour of the branched products and good reactivity. This study represents the first use of helicene-like ligands in asymmetric reactions such as hydroformylation and allylic amination, and the promising results indicate the potential of the helicene moieties as chiral inductors.



(+)-(*P*,*S*)

The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. Z.K. and P.S. performed the synthesis and characterisation. Collaborators: P.E. planned the project and supervised the experiments. B.P.B. and S.C. performed the synthesis and characterisation. D.Š. carried out the NMR spectroscopic analyses.

3 Self-assembly of helical aromatics on solid surfaces and their on-surface reactivity

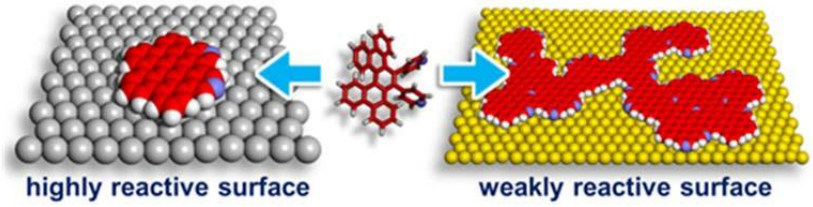
3.1 Diffusion-controlled on-surface (cyclo)dehydrogenation of heteroaromatics: Formation of *N*-doped nanohelicenes, nanographenes, nanodomes and 2D polyaromatic networks

Surface-assisted cyclodehydrogenation or dehydrogenative polymerisation of polycyclic (hetero)aromatic hydrocarbons is one of the most important strategies for bottom-up assembly of new nanostructures from their constituent molecular building blocks. Although diverse compounds have been formed in recent years using this methodology, a limited knowledge on the molecular machinery operating at the nanoscale has so far disallowed to control the reaction outcome. We show that the strength of the polyaromatics-substrate interaction rules the competitive reaction pathways (cyclodehydrogenation versus dehydrogenative polymerisation).[10] Then, starting from the same molecular precursor and controlling its diffusion by the nature of the supporting surface, temperature-triggered dehydrogenation takes place to

[9] Krausová, Z.; Sehnal, P.; Bondzic, B. P.; Chercheja, S.; Eilbracht, P.; Stará, I. G.; Šaman, D.; Starý, I. Helicene-Based Phosphite Ligands in Asymmetric Transition-Metal Catalysis: Exploring Rh-Catalyzed Hydroformylation and Ir-Catalyzed Allylic Amination. *Eur. J. Org. Chem.* **2011**, 3849–3857.

[10] (a) Pinardi, A. L.; Otero-Irurueta, G.; Palacio, I.; Martinez, J. I.; Sanchez-Sanchez, C.; Tello, M.; Rogero, C.; Cossaro, A.; Preobrajenski, A.; Gómez-Lor, B.; Jančařík, A.; Stará, I. G.; Starý, I.; Lopez, M. F.; Méndez, J.; Martin-Gago, J. A. Tailored Formation of N-Doped Nanoarchitectures by Diffusion-Controlled on-Surface (Cyclo)Dehydrogenation of Heteroaromatics. *ACS Nano* **2013**, *7*, 3676–3684; (b) Pinardi, A. L.; Martinez, J. I.; Jančařík, A.; Stará, I. G.; Starý, I.; López, M. F.; Méndez, J.; Martín-Gago, J. Á. Sequential Formation of N-Doped Nanohelicene, Nanographene and Nanodome by Surface-Assisted Chemical (cyclo)dehydrogenation of Heteroaromatics. *Chem. Commun.* **2014**, *50*, 1555–1557.

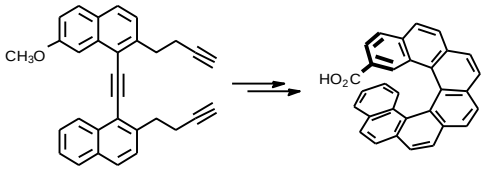
provide molecular, oligomeric or polymeric structures of variable dimensionality. By controlling the diffusion of *N*-heteroatomatic precursor, the on-surface dehydrogenation can lead to monomolecular diazahexabenzocoronene (*N*-doped nanographene), nanodome , *N*-doped oligomeric/polymeric networks or to carbon rich monolayers (membranes or graphene). Governing the on-surface dehydrogenation is a step forward towards the tailored fabrication of molecular 2D nanoarchitectures distinct to graphene and exhibiting new properties of a fundamental and technological interest.

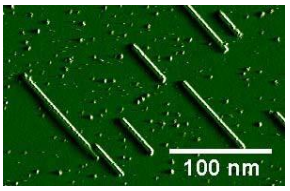


The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and co-wrote the manuscript. A.J. performed the synthesis and characterisation. Collaborators: J.A.M.-G. planned the project, supervised the experiments and co-wrote the manuscript. A.L.P., G.O.-I., I.P., J.I.M., C.S.-S., M.T., C.R., A.C., A.P., B.G.-L., M.F.L. and J.M. performed the surface analyses and quantum chemical calculations.

3.2 The synthesis of racemic and enantiopure functionalised helically chiral aromatics and their self-assembly into nanowire-like aggregates

For bottom-up fabrication of molecular structures, the promising concept of self-assembly has attracted great attention. However, only little progress has been made so far on insulating substrates, although future molecular electronic applications will require non-conducting substrates. Chirality can have a decisive influence on molecular structure formation upon self-assembly on surfaces. We developed a straightforward synthesis of [7]helicene-2-carboxylic acid.[11] The key operation in building the helical skeleton was Co^{I}-mediated [2+2+2] cycloisomerisation of aromatic triyne. The carboxylic moiety was introduced by the Pd^{II}-catalysed methoxycarbonylation reaction being performed with the corresponding helicene triflate. The racemic [7]helicene-2-carboxylic acid was resolved into enantiomers by semipreparative HPLC on a chiral column and their helicity was assigned by CD spectra correlation. Racemic [7]helicene-2-carboxylic acid was deposited on calcite (10-14) to undergo self-assembly into nanowire-like aggregates as demonstrated by noncontact atomic force microscopy (NC-AFM). This study showed that selecting properly functionalised molecules enabled the self-assembly of molecular wire-like structures even on insulating surfaces, where high molecular mobility has so far hampered the self-assembly of tailor-made molecular





[11] (a) Rahe, P.; Nimmrich, M.; Greuling, A.; Schütte, J.; Stará, I. G.; Rybáček, J.; Huerta-Angeles, G.; Starý, I.; Rohlfing, M.; Kühnle, A. Toward Molecular Nanowires Self-Assembled on an Insulating Substrate: Heptahelicene-2-Carboxylic Acid on Calcite (1014). *J. Phys. Chem. C* **2010**, *114*, 1547–1552; (b) Rybáček, J.; Huerta-Angeles, G.; Kollárovič, A.; Stará, I. G.; Starý, I.; Rahe, P.; Nimmrich, M.; Kühnle, A. Racemic and Optically Pure Heptahelicene-2-Carboxylic Acid: Its Synthesis and Self-Assembly into Nanowire-Like Aggregates. *Eur. J. Org. Chem.* **2011**, 853–860; (c) Hauke, C. M.; Rahe, P.; Nimmrich, M.; Schütte, J.; Kittelmann, M.; Stará, I. G.; Starý, I.; Rybáček, J.; Kühnle, A. Molecular Self-Assembly of Enantiopure Heptahelicene-2-Carboxylic Acid on Calcite (1014). *J. Phys. Chem. C* **2012**, *116*, 4637–4641.

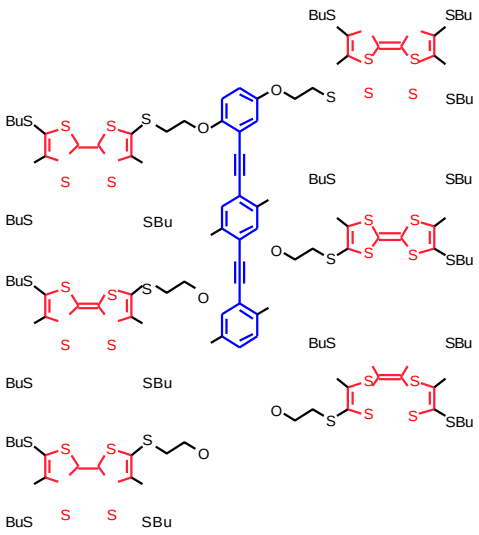
structures. Obviously, the formation of these nanowire-like aggregates was governed by a balanced interplay of the π-π interactions between large aromatic systems and hydrogen bonding between the carboxylic groups themselves or the electrostatic interaction between the carboxylate moieties and the calcite substrate. In sharp contrast to the racemate, the (*M*)-enantiomer forms molecular islands. Our results elucidate the details governing the molecular self-assembly, indicating that heterochiral recognition is responsible for the formation of the uni-directional double row structures formed by the racemate.

The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and co-wrote the manuscripts. J.R., A.A., G.H.-A. and A.K. performed the synthesis and characterisation. Collaborators: A.K. planned the project, supervised the experiments and co-wrote the manuscript. C.M.H., P.R., M.N., J.S., A.G, M.R. and M.K. performed the surface analyses and quantum chemical calculations.

4 The synthesis and properties of other π-electron systems

4.1 The synthesis and physicochemical properties of tetrathiafulvalene-oligo(*p*-phenyleneethynylene) conjugates (TTF-*p*OPE)

Within a programme focused on the development of new zipper-type multiple noncovalent interactions, we have turned our attention to the short monodisperse oligo(*p*-phenyleneethynylene)s *p*OPE rods equipped laterally with TTF units.[12] The model compounds have been synthesised from functionalised aromatic building blocks using the Sonogashira cross-coupling methodology. The unusual redox properties of these TTF-*p*OPE conjugates have been observed by employing electrochemical methods such as cyclic voltammetry and exhaustive electrolysis. We have found that one half of the TTF units in the *p*OPE monomer (having two TTFs), dimer (having four TTFs) and trimer (having six TTFs) are electrochemically silent during the first-step oxidation at 0.49 V. We propose the intramolecular formation of persistent mixed-valence complexes from the TTF and $TTF^{+\bullet}$ units comprised in an equal ratio in a molecule. Such mixed-valence dyads (single or multiple in the partially oxidised molecules) exhibit an unusual stability towards oxidation until the potential of the second oxidation at 0.84 V is achieved. This suggests that below this potential the oxidation of the respective mix-valence complexes is extremely slow or, more probably, that the first-step oxidation wave is split to such an extent (*ca* 350 mV) that its higher-potential part overlaps with the second-step oxidation wave.



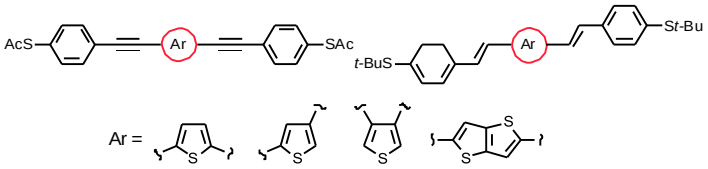
The team: I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. Š.L. and M.B. performed the synthesis and characterisation. J.V.C. and J.V. performed the quantum chemical calculations. Collaborators: V.K., L.P., M.H. and J.F. performed the electrochemical analyses. R.P. carried out the NMR spectroscopic analyses. L.P. co-wrote the manuscript.

---

[12] Lipnická, Š.; Bělohradský, M.; Kolivoška, V.; Pospíšil, L.; Hromadová, M.; Pohl, R.; Chocholoušová, J. V.; Vacek, J.; Fiedler, J.; Stará, I. G.; et al. Tetrathiafulvalene–Oligo(para-Phenyleneethynylene) Conjugates: Formation of Multiple Mixed-Valence Complexes upon Electrochemical Oxidation. *Chem. Eur. J.* **2013**, *19*, 6108–6121.

4.2 The synthesis of π-electron molecular wires with a thiophene or thieno[3,2-*b*]thiophene core unit and sulphur alligator clips

We prepared a series of the oligo(*p*-phenyleneethynylene)- and oligo(*p*-phenylenevinylene)-type molecular wires with the thiophene or thieno[3,2-*b*]thiophene core unit and end-capped with sulphur anchoring groups (AcS-, *t*-BuS-) using the Sonogashira coupling or Horner-Wadsworth-Emmons (*E*)-olefination methodology.[13] These linear/bent, conjugated/cross-conjugated systems have been characterised by UV-VIS spectroscopy and, accordingly, optical HOMO-LUMO gaps.



The experiments to estimate their molecular conductivities by using the mechanically controllable break-junction method are currently under way.

<u>The team:</u> I.S. and I.G.S. planned the project, designed the compounds, supervised the experiments and wrote the manuscript. A.S. and V.D. performed the synthesis and characterisation. J.V.C. and J.V. performed the quantum chemical calculations. <u>Collaborators:</u> J.S. planned the project and supervised the experiments.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Petr Bouř - Biomolecular Spectroscopy

Summary

The Biomolecular Spectroscopy research group at IOCB was established in 2007. Within 2010-2014 it comprised experienced researchers, postdocs, master and PhD students. The research program is oriented towards combined studies of molecules by spectroscopic and theoretical method. We are convinced that we need to understand the structure and physical principles governing the world of bimolecules, in order to change or exploit their function.

The senior researchers currently include Petr Bouř (quantum chemistry, spectroscopy), Valery Andrushchenko (optical spectroscopy), Jakub Kaminský (computational models), Vladimír Sychrovský (DNA models, NMR), Michal Straka (fullerene and heavy metal chemistry, relativity), and Jaroslav Šebestík (organic synthesis, spectroscopy).

We **develop experimental and theoretical method** with an emphasis on **detailed interpretation of spectroscopic data** (nuclear magnetic resonance, NMR, circular dichroism, CD, vibrational optical activity, VOA, etc.), **multi-level computations**, **development of new experimental** (e.g. vibrational optical activity, magnetic circular dichroism) and **theoretical** (e.g. tensor transfer techniques) **procedures**. We also promote **applications** of these methods in a number of domains including **organic and medical chemistry, and biochemistry**.

Collaboration projects outside the Institute involved Charles University (Prague, Prof. V. Baumruk), Palacký University (Olomouc, Dr. Josef Kapitán), University of Chemical Technology (Prague, Prof. Marie Urbanová, Doc. V. Setnička), Masaryk University (Brno, Dr. P. Kulhánek, Prof. R. Marek, Prof. V. Sklenář), University of Opole (Prof. Teobald Kupka), University of Wyoming (USA, Prof. J. Kubelka), University of Illinois at Chicago (USA, Prof. T. A. Keiderling), University of Manchester (UK, Prof. E. Blanch), University of Helsinki (Finland, Profs. P. Pyykkö, D. Sundholm, M. Patzschke and S. Taubert), University of Oulu (Profs. J. Vaara and P. Lantto), Tohoku University (Japan, Prof. Y. Tanaka), Kwansei Gakuin University (Japan, Prof. S. Yamamoto), Friedrich-Schiller University (Germany, Prof. W. Pohle), University of Calgary (Canada, Prof. H. Wieser), Institute for Low Temperature Physics and Engineering (Ukraine, Dr. V. Sorokin), and University of Tromsø (Norway, Prof. K. Ruud, Dr. K. Hopmann).

Grant Project

Apart of institute financing the research was primarily financed by the Grant Agency (P208-11-0105, Expanding the Optical Activity Method to the Realm of Biomolecules; P208/10/0559, Theoretical and spectroscopic studies of nucleic acid self-assemblies; 13-03978S, Magnetic circular dichroism as an analytical tool for

fullerenes and carbon nanostructures; 15-09072S, Development of theoretical and spectroscopic tools for studies of amyloid fibrils; P208/10/P356, Development of precise molecular mechanic force fields for vibrational spectroscopy) Ministry of Education (LH11033, Development of spectroscopic methods for structural molecular studies) and the Academy of Sciences (M200551205, Application of chiral spectroscopy for studies of macromolecular complexes, and a bilateral grant between Czech Academy of Sciences and National Academy of Sciences of Ukraine "Tuning the structure of DNA/RNA adsorbed on carbon nanotubes or graphene oxide by small ligands") of the Czech Republic. We also obtained several fellowships (Marie Curie reintegration grant– to Michal Straka, IOCB postdoc program – Shigeki Yamamoto, Tao Wu, Daniele Padula, Wichterle prize and Fulbright– Jakub Kaminský) and Oulu University Research Council Travel Grants (Michal Straka).

Overview of main results

(Selected results are approximately sorted by a prevalent theme:

Raman Spectroscopy
Raman optical activity (ROA)
Theoretical methods
Magnetic Circular Dichroism (MCD)
Vibrational circular dichroism
Nuclear magnetic resonance (NMR)
Spectroscopy and Chemistry of Nucleic Acids

Raman Spectroscopy

We explore the Raman spectroscopy as a modern non-invasive technique that is currently finding many new applications in analytical chemistry, security (detection of explosives), medical imaging (discrimination of healthy and malignant tissues), etc. The development is propelled by new technologies, e.g. available lasers, surface-enhancement techniques, and optical elements; at the same time the information contents that can be obtained from the spectra is significantly increased by advanced computer analyses and computational methods. Our main contribution consists of developing and testing ab initio quantum chemical methods that can be used for this purpose.

Kapitán, J.; Dračínský, M.; Kaminský, J.; Benda, L.; Bouř, P. *J. Phys. Chem. B* **2010,** *114,* 3574–3582. "**Theoretical Modeling of Magnesium Ion Imprints in the Raman Scattering of Water**"

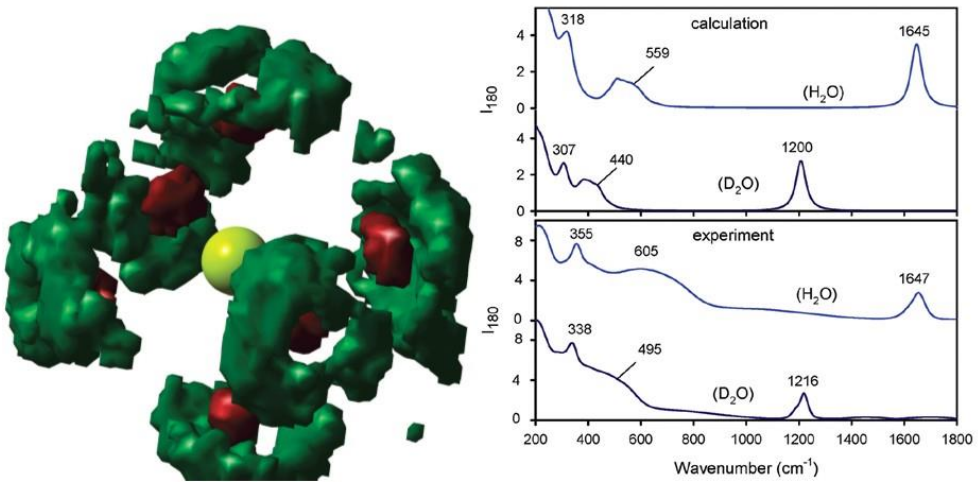


Figure 1. Calculated hydration sphere of the magnesium ion and its calculated and experimental Raman spectra in H2O and D2O solutions.

This is shown as a typical result, also indicating where we would like to go in the future in this field; Raman bands of the hydrated magnesium and other ions lie in the low-frequency region, are rather broad, and difficult to model and analyze. Also very interesting low-frequency motions of proteins and other biomolecules contribute to this part of the spectra. Yet modern Raman spectrometers provide means how to measure reliably to very low wavenumbers, because of a new generation of narrow filters; also new techniques appeared providing specifically this part of molecular spectra, such as Fourier-transform based laser pulse techniques. Theoretically the description is challenging because of temperature-excited anharmonic ill-defined vibrational modes accompanying interaction of molecules and ions with the environment. In the cited work we experimentally characterized a series of ions in this region and found a model that enables us to understand the spectrum, even in a quantitative way (**Figure 1**).

Similar task, developing a theory that can be used for analytical chemistry, was pursued in Dračínský, M; Procházková, E.; Kessler, J.; Šebestík, J.; Matějka, P.; Bouř, P. *J. Phys. Chem. B*, **2013**, *117 (24)*, 7297–7307. "**Resolution of Organic Polymorphic Crystals by Raman Spectroscopy**." Here, different theoretical challenges were encountered, consisting in modeling of periodic crystal structures; the applications were more straightforward because of the importance of discrimination between various polymorphic forms of drugs in the pharmaceutical industry.

Other works dedicated to Raman spectroscopy was Šebestík, J; Šafařík, M.; Bouř, P. *Inorg. Chem.* **2012**, *51 (8)*, 4473–4481. "**Ferric Complexes of 3-Hydroxy-4-pyridinones Characterized by Density Functional Theory and Raman and UV–vis Spectroscopies**." and two studies (where the organic-synthetic part was perfomed in Univesity of Tohoku, Japan) was Benda, L.; Straka, M.; Sychrovský, V.; Bouř, P.; Tanaka, Y. *J. Phys. Chem. A.* **2012**, 116, 8313-8320. "**Detection of Mercury−TpT Dinucleotide Binding by Raman Spectra: A Computational Study**." and Uchiyama, T.; Miura, T.; Takeuchi, H.; Dairaku, T.; Komuro, T.; Kawamura, T.; Kondo, Y.; Benda, L.; Sychrovský, V.; Bouř, P.; Okamoto, I.; Ono, A.; Tanaka, Y. *Nucleic Acids Res.* **2012**, *40*, 5766-5774. "**Raman spectroscopic detection of the T-HgII-T base pair and the ionic characteristics of mercury**."

Raman optical activity (ROA)

Raman optical activity – detection of a tiny difference in scattering of right- and left-circularly polarized light –is the youngest and most dynamic chiral spectroscopic technique that became lately available also on a commercial basis. ROA methodology combines higher resolution of vibrational spectra with the higher conformational sensitivity of the chiral techniques, and is especially suitable for biologically relevant molecules as it enables measurement in the natural aqueous environment. Chiroptical spectroscopy is a core of our research; we profit from a long tradition of high-quality research on this field in the Institute and Czech Republic. We are developing ROA methodology in collaboration with the Charles University (Prof. V. Baumruk) and Palacký University (Dr. Josef Kapitán). In 2010 we established an independent ROA laboratory and within 2010-2014 we already achieved several unique results, such as discovery of gas ROA, first observation of paramagnetic ROA, and spectroscopic determination of valinomycin and insulin conformation in solution.

Yamamoto, S.; Straka, M.; Watarai, H.; Bouř, P. *Phys. Chem. Chem. Phys.* **2010,** *12*, 11021-11032, "**Formation and structure of the potassium complex of valinomycin in solution studied by Raman optical activity spectroscopy**."

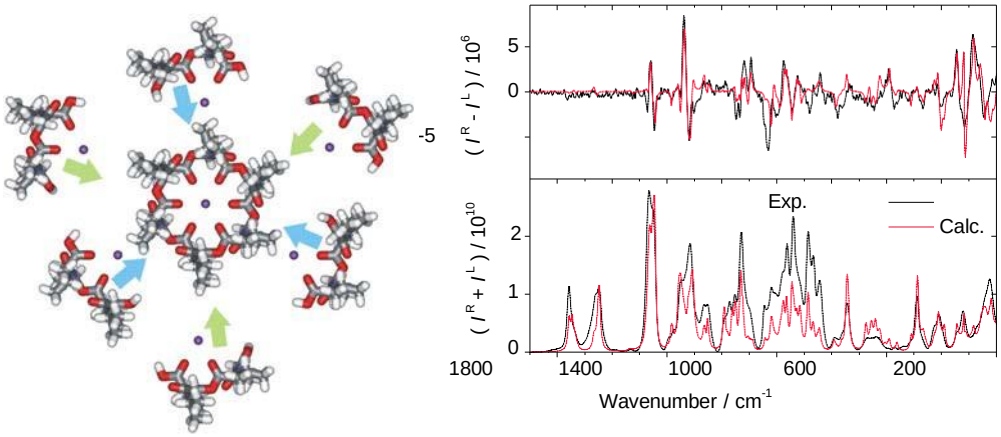


Figure 2. For interpretation of the valinomycin peptide spectra, a fragment tensor transfer technique was used, as developed at IOCB.

In this study, we showed how ROA could contribute to understanding molecular structure and flexibility of a biologically relevant compound, antibiotic and ionophore valinomycin. To understand spectra of this molecule we stretched the possibilities of the ROA experiment and theory. For example, among over 6000 molecular conformers three most important ones could be identified in solution, one of them not identifiable by NMR. The ROA computations were based on advanced tensor transfer technique that we developed as well (**Figure 2**). A new valynomycin conformation was discovered, previously not visible to NMR, because of molecular symmetry.
In Yamamoto, S.; Watarai, H.; Bouř, P. *ChemPhysChem* **2011**, *12*, 1509-1518. **"Monitoring the Backbone Conformation of Valinomycin by Raman Optical Activity**." this methodology was extended to non-symmetric valinomycin forms.

Šebestík, J.; Bouř, P. *J. Phys. Chem. Lett.,* **2011**, *2*, 498-502. "**Raman Optical Activity of Methyloxirane Gas and Liquid**."

Because of the lower sensitivity, gas ROA has not been known so far. However, we could measure it with a special experimental setup that prevented geometry change of the cell compartment between baseline and sample acquisition. This broadens the ROA application span, and it is important for benchmarking of the computations. For example, we could show the effect of the rotational broadening and anharmonic interactions in the spectra. Immediately after the measurement, the results were used for first coupled cluster calculation of ROA by K. Ruud and D. Crawford (ChemPhysChem 2011, 12, 3442-3448).

Yamamoto, S.; Kaminský, J; Bouř, P. *Anal. Chem.* **2012**, *84*, 2440-2451. "**Structure and Vibrational Motion of Insulin from Raman Optical Activity Spectra**."

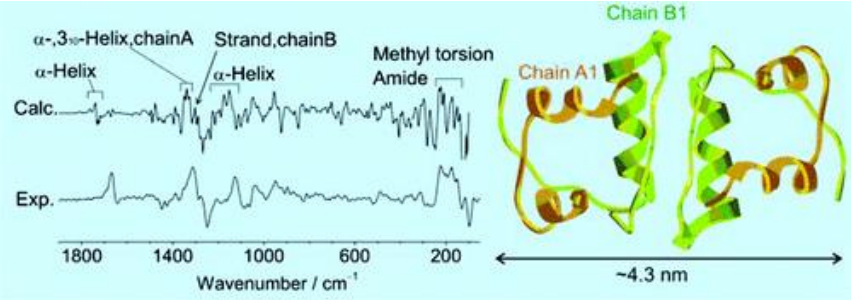


Figure 3. Analysis of insulin ROA spectrum based on quantum-chemical computations provided information about its structural features in solution.

Insulin was the largest molecule for which vibrational ROA spectrum was ever simulated by quantum-chemical methodology. The simulation was possible due to the tensor-transfer algorithm for a long time developed in our group. Then the spectrum revealed structural features that could be linked to specific parts of insulin molecule (**Figure 3**). In detail, the relation of the ROA signal to peptide conformation and relative conformer energies could be studied on simpler molecules, in Parchaňský, V.; Kapitán, J.; Kaminský, J.; Šebestík, J.; Bouř, P. *J. Phys. Chem. Lett.* **2013**, *4*, 2763-2768. „**Ramachandran Plot for Alanine Dipeptide as Determined from Raman Optical Activity**.“ and Yamamoto, S.; Furukawa, T.; Bouř, P.; Ozaki, Y. *J. Phys. Chem. A.* **2014**, 118 (20), 3655-3662. “**Solvated States of Poly-l-alanine α-Helix Explored by Raman Optical Activity.**”

Šebestík, J.; Bouř, P. *Angew. Chem. Int. Ed.* **2014**, 53 (35), 9236-9239. “**Observation of Paramagnetic Raman Optical Activity of Nitrogen Dioxide**.”

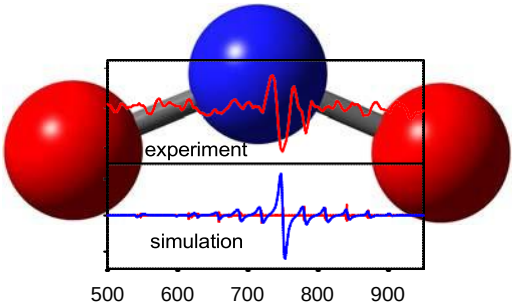


Figure 3. Experimental paramagnetic Raman optical activity of NO_2 was verified through comparison with a simulation.

Chiroptical phenomena have been attracting people’s attention since the time of Luis Pasteur (1848). We were then happy that we could report a new flavor of the magneto-optic spectroscopy, paramagnetic Raman optical activity of gases (**Figure 3**). This was considered difficult, but we succeeded because of a new magnetic cell, and a theoretical analysis. The results suggest that the technique can bring about unique information about molecular properties, and may be even usable for determination of molecular properties (even individual polarizability components) and for a characterization of industrial gases.
The interpretation of the rotationally-resolved spectral features required to develop and implement the relevant theory, which was also a topic of the previous study, Wang, B.; Bouř, P.; Keiderling, T. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9586-9593.

"**Rotationally Resolved Magnetic Vibrational Circular Dichroism of the Paramagnetic Molecule NO**."

As another possible application of ROA and other chiroptical techniques, we investigated its use for organometalic complexes (Wu, T.; Hudecová, J.; Bouř, P. *Chem. Eur. J.* **2015**, 21 (15), 5807-5813. "**Comparison of Electronic and Vibrational Optical Activity of a Europium^{III} Complex**", Wu, T.; You, X. Z.; Bouř, P. *Coord. Chem. Rev.* **2015**, *284*, 1-18. "**Applications of chiroptical spectroscopy to coordination compounds.**", Yamamoto, S.; Bouř, P. *J. Comput. Chem.* **2013,** *34*,2152-2158. „**Transition Polarizability Model of Induced Resonance Raman Optical Activity**." and Yamamoto, S.; Bouř, P. *Angew. Chem. Int. Ed.* **2012**, *51(44)*, 11058-11061. " **Detection of Molecular Chirality by Induced Resonance Raman Optical Activity in Europium Complexes**.")
[

Theoretical methods

Several theoretical algorithms and methods were developed and implemented, aimed at interpretation of molecular spectra, but also usable in general for modeling, predicting and understanding molecular behavior.

Parchaňský, V.; Bouř, P. *J. Chem. Phys.* **2010,** *133,* 044117. "**Transferability of anharmonic force fields in simulations of molecular vibrations.**" (Nominated for a prize of the Academy of Sciences)

The harmonic limit becomes a bottleneck for spectral interpretations as the precision of the computational methods increases. In this study the harmonic limit is partially overcome even for large molecules. It shows that molecular anharmonic force field constants can also be transferred in a similar way as the harmonic ones. This can make computations even on very large molecules much more precise, with a significant saving of computer time.

The modeling of the anharmonic interaction was also pursued in Hudecová, J.; Profant, V.; Novotná, P.; Baumruk, V.; Urbanová, M.; Bouř, P. *J. Chem. Theory Comput.* **2013**, *9*, 3096−3108. "**CH Stretching Region: Computational Modeling of Vibrational Optical Activity**." where we could significantly improved the harmonic results, using original algorithm based on the limited vibrational configuration interaction and two-step selection of harmonic oscillator excited states.

A detailed analysis of the transferability of molecular properties (force field, intensity tensors) and its usability for various vibrational spectroscopies were presented in Yamamoto, S.; Li, X.; Ruud, K.; Bouř, P. *J. Chem. Theory Comput.* **2012,** *8(3)*, 977-985. „**Transferability of Various Molecular Property Tensors in Vibrational Spectroscopy.**" and Andrushchenko, V.; Bouř, P. *Chirality*, **2010**, *22*, E96-E114. "**Applications of the Cartesian Coordinate Tensor Transfer Technique in the Simulations of Vibrational Circular Dichroism Spectra of Oligonucleotides**."

General molecular dynamics procedures were tested in Kessler, J.; Jakubek, M.; Dolenský, B.; Bouř, P. *J. Comput. Chem.* **2012**, 33, 2310-2317, "**Binding Energies of Five Molecular Pincers Calculated by Explicit and Implicit Solvent Models**", Kessler, J.; Bouř, P. *J. Comput. Chem.* **2014**, *35*, 1552-1559. "**Molecular Dynamics with Helical Periodic Boundary Conditions.**", and Kessler, J.;

Dračínský, M.; Bouř, P. *J. Comput. Chem.* **2013**, 34, 366-371. "**Parallel Variable Selection of Molecular Dynamics Clusters as a Tool for Calculation of Spectroscopic Properties**". The last two studies present algorithms that can be used to make molecular dynamics studies including ab initio molecular dynamics faster and more efficient. In Gauger, D. R.; Andrushchenko, V. V.; Bour, P.; Pohle, W., *Anal. Bioanal. Chem.* **2010,** 398, (2), 1109-1123. "**A spectroscopic method to estimate the binding potency of amphiphile assemblies**" carefully chosen molecular dynamics procedures were used to rationalize experimental spectroscopic results on amphiphilic lipid assemblies where X-ray techniques were not applicable.We also aim to new computationally efficient quantum-mechanical procedures, which was successful in Štěpánek, P.; Bouř, P. *J. Comput. Chem.* **2013**, *34 (18),* 1531-1539. "**Computation of Magnetic Circular Dichroism by Sum Over States Summations**." were the simulation time could be reduced by an order if compared to a classical methodology. Likewise, a semiempirical model of the surface enhanced Raman scattering could be used to interpret ours and other results, as summarized in Novák, V.; Šebestík, J.; Bouř, P. *J. Chem. Theory Comput.* **2012***, 8*, 1714-1720. "**Theoretical Modeling of the Surface-Enhanced Raman Optical Activity**."

Magnetic Circular Dichroism (MCD)

Štěpánek, P.; Andrushchenko, V.; Ruud, K.; Bouř, P. *J. Phys. Chem. A* **2012,** *116 (1)*, 778–783. **"MCD Porphyrin Protonation Studied by Magnetic Circular Dichroism.**"

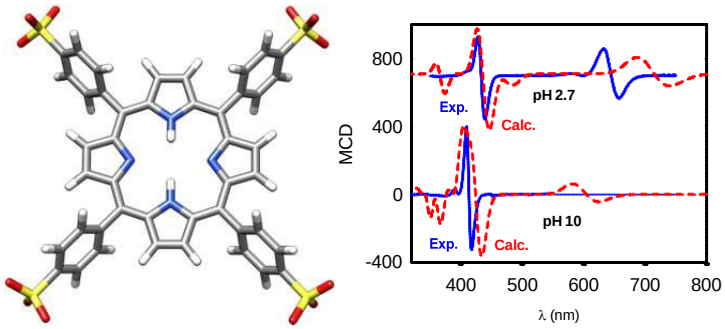


Figure 4. Monitoring of porphyrin protonation by MCD.

MCD spectroscopy measures absorption difference of the left- and right-circularly polarized light under presence of a static magnetic field. It goes through a revival period because of the possibility to interpret and understand the spectra from the first principles.
We introduced this technique to IOCB because we believe in its potential to reveal precious information about molecular structure and electronic properties. In the study given as an example in **Figure 4** we could show that MCD is able to monitor a simple chemical reaction, and the computation is able to explain the experiment.

Meanwhile, we found other MCD applications, such as insight into the properties of organometallic complexes (Andrushchenko, V.; Padula, D.; Zhivotova, E.; Yamamoto, S., Bouř, P. *Chirality* **2014**, *26*, 655-662. "**Magnetic Circular Dichroism of Porphyrin Lanthanide M^{3+} Complexes**."), characterization of carbon-based new materials (Štěpánek, P.; Straka, M.; Andrushchenko, V.; Bouř, P. *J. Chem. Phys.* **2013**, 138, 151103. "**Fullerene resolution by the magnetic circular dichroism**.") and monitoring of biomolecular properties in solution (Štěpánek, P.; Bouř, P. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20639-20649. "**Multi-scale modeling of electronic spectra of three aromatic amino acids: importance of conformational averaging and explicit solute–solvent interactions**.")

With Oulu University (Prof. J. Vaara, Finland) we investigated the potential of recently suggested novel magneto-optic spectroscopy, nuclear-spin induced circular dichroism (NSCD). Unlike MCD, it provides atomic site specific resolution. One can imagine that NMR signal is detected by means of optical spectroscopy. Our theoretical calculations confirmed the locality of this technique (Straka, M.; Štěpánek,P.; Coriani, S.; Vaara, J. *Chem. Commun.* **2014**, *50*, 15228." **Nuclear spin circular dichroism in fullerenes: a computational study**."

Vibrational circular dichroism (VCD)

VCD, difference in absorption of left- and right-circularly polarized infrared light, provides information on molecular vibrational properties, and is thus a natural counterpart of ROA. We concentrate on development of the theoretical foundations of the technique, in collaboration with experimental laboratories of Prof. M Urbanová, University of Chemical technology, Prague, Prof. T. A. Keiderling, University of Illinois at Chicago, and Prof. H. Wieser, Uviversity of Calgary. Using experiemntal results from the last laboratory, for example, we used combined MD/DFT approach coupled with tensor transfer method for unambiguous assignment of $(dG)_8$ solution conformation to a guanine quadruplex (G-quadruplex) and established characteristic IR and VCD spectral features for this conformation. The obtained results can be used for convenient and reliable detection of G-quadruplex structures in solutions (Andrushchenko, V.; Tsankov, D.; Krasteva, M.; Wieser, H.; Bour, P., *J. Am. Chem. Soc.* **2011,** 133, (38), 15055-15064. **"Spectroscopic Detection of DNA Quadruplexes by Vibrational Circular Dichroism.")**

Lately, VCD was found useful to monitor formation of protein plugs and aggregates in relation to the neurodegenerative diseases (Alzheimer, Parkinson, Huntington). We developed a computational model based on a periodic crystal approximation, and could explain several features in VCD fibrillar spectra, related to their structure and kind of side chains (Kessler, J.; Keiderling, T. A.; Bouř, P. *J. Phys. Chem. B* **2014**, *118*, 6937-6945. „**Arrangement of Fibril Side Chains Studied by**

Molecular Dynamics and Simulated Infrared and Vibrational Circular Dichroism Spectra.")

Quite a new VCD application, for Magnetic Coordination Complexes, was explored in Wu, T.; Zhang, X. P.; You, X. Z., Li, Y. Z., Bouř, P. *ChemPlusChem* **2014**, 79, 698-707. "**Chirality Transfer in Magnetic Coordination Complexes Monitored by Vibrational and Electronic Circular Dichroism**."

The VCD topic is challenging also theoretically. Although the basic quantum-chemical methodology was developed in 1985, new challenges (solvent, dynamics) are encountered even in small-molecule studies, which were addressed in Hudecová, J.; Horníček, J.; Buděšínský, M.; Šebestík, J.; Šafařík, M.; Zhang, G.; Keiderling, T. A.; Bouř, P. *ChemPhysChem* **2012**, *13(11)*, 2748–2760. "**Three Types of Induced Tryptophan Optical Activity Compared in Model Dipeptides: Theory and Experiment**." and Li, X.; Hopmann, K. H.; Hudecová, J.; Stensen, W.; Novotná, J.; Urbanová, M.; Svendsen, J. S.; Bouř, P.; Ruud. K. *J. Phys. Chem. A* **2012***, 116*, 2554-2563. "**Absolute Configuration of a Cyclic Dipeptide Reflected in Vibrational Optical Activity: Ab Initio and Experimental Investigation**."

Nuclear magnetic resonance (NMR)

Theoretical method can also significantly enhance interpretations of NMR experiments. We completed several methodological studies to support NMR applicability in studies of molecular behavior: Dračínský, M.; Bouř, P. *J. Comput. Chem.* **2012**, *33*, 1080-1089. "**Vibrational Averaging of the Chemical Shift in Crystalline α-Glycine**.", Li, X.; Hopmann, K. H.; Hudecová, J.; Isaksson, J.; Novotná, J.; Stensen, W.; Andrushchenko, V.; Urbanová, M.; Svendsen, J. S.; Bouř, P.; Ruud, K. *J. Phys. Chem. A,* **2013,** *117 (8)*, 1721−1736. "**Determination of Absolute Configuration and Conformation of a Cyclic Dipeptide by NMR and Chiral Spectroscopic Methods**." (synthetic and some computational part in collaboration with University of Tromsø, Norway), and Dračínský, M.; Jansa, P.; Bouř, P. *Chem. Eur. J.* **2012,** *18,* 981-986. "**Computational and Experimental Evidence of Through-Space NMR Spectroscopic J-Coupling of Hydrogen Atoms**."

In a very detailed and already highly-cited Dračínský, M.; Bouř, P. *J. Chem. Theory Comput.* **2010,** *6,* 288–299. "**Computational Analysis of Solvent Effects in NMR Spectroscopy.**" we addressed the problematics of modelling chemical shift and spin-spin coupling constants in solution.

NMR is the prevalent technique used for analytic identification of fullerenes and new compounds and materials based on them. We therefore studied the effect of **solvent and dynamics on the ^{13}C NMR** chemical shifts in fullerenes. We could determine the contribution of surprisingly large dynamical and solvent effects in Kaminský, J.; Buděšínský, M.; Taubert, S.; Bouř, P.; Straka, M. *Phys. Chem. Chem. Phys.* **2013**, 15, 9223-9230. "**Fullerene C70 characterization by 13C NMR and the importance of the solvent and dynamics in spectral simulations**."

Xeon atom in fullerenes sensitively reacts on environment and can be used as an imaging probe in medicine. To understand this effect, we performed simulations of ^{129}Xe NMR parameters at experimental conditions in confined Xe atom systems and novel Xe molecules. We have developed a piecewise approximation that includes relativistic effects via Breit-Pauli perturbation theory while simulating effects of

dynamics and explicit dynamical solvent, including the relativistic corrections. Our three studies opened the way to quantum-chemical studies of ^{129}Xe NMR parameters simulated including experimental conditions of Xe atom enclosed in different guest-host systems, such as liquids, clathrates, or molecular cages. The simulations provide understanding of experimental results on the microscopic structure and were published in three separate papers, e.g. Standara, S.; Kulhánek, P.; Marek, R.; Straka, M. *J. Comput. Chem.* **2013**, *129*, 1890-1898. „**Xe NMR Chemical Shift in Xe@C_{60} Calculated at Experimental Conditions: Essential Role of the Relativity, Dynamics, and Explicit Solvent.***"* (**Figure 5**, left).

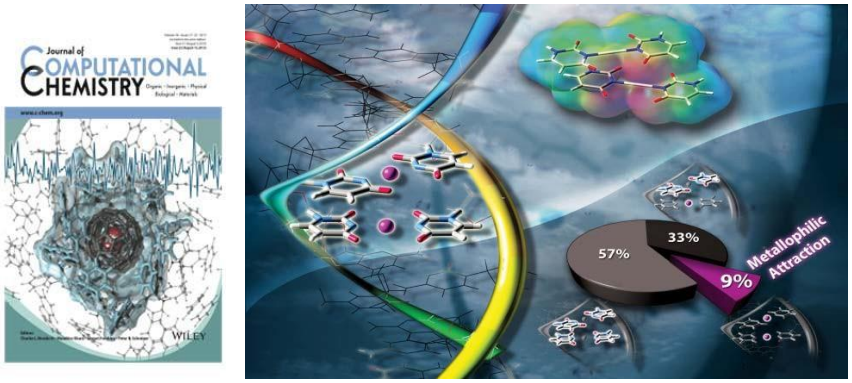


Figure 5. Journal cover pictures of our studies: left, Xe@C_{60} dissolved in benzene, right, metallophilic interactions discovered in complexes of nucleic acid components and heavy metals.

.

Fullerenes, but also biomolecules, such as nucleic acids, form complexes with heavy metals, understanding of which requires special computational procedures including relativistic efects. In two studies, we could explain the so called HALA effect (heavy-atom-light-atom) in chemical shifts of light atoms in Lantto, P.; Riedel, S.; Standara, S.; Vaara, J.; Straka, M. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10944. „**Exploring new ^{129}Xe chemical shift ranges in HXeY compounds: hydrogen more relativistic than xenon.**" and Vícha, J.; Straka, M.; Munzarová, M. L.; Marek, R. *J. Chem. Theor. Comp.* **2014**, *10*, 1489. „**Mechanism of Spin–Orbit Effects on the Ligand NMR Chemical Shift in Transition-Metal Complexes: Linking NMR to EPR.**"

Spectroscopy and Chemistry of Nucleic Acids

We study nucleic acids (NA) to understand the biological role, but also as potential new materials for nano and electrotechnologies. Using the experiments of Prof. Tahala (Tohuku University, Japan) we could computationally for the first time explain the role of metallophilic interactions in metallated DNA for conformational stabilization (**Figure 5**, right, Benda, L.; Straka, M.; Tanaka, Y.; Sychrovský, V. *Phys. Chem. Chem. Phys.* **2011**, *13*, 100. „ **On the role of mercury in the non-covalent stabilisation of consecutive U–Hg^{II}–U metal-mediated nucleic acid base pairs: metallophilic attraction enters the world of nucleic acids.")**, and several studies followed (Benda, L.; Vokáčová, Z. S.; Straka, M.; Sychrovský, V. *J. Phys. Chem. B* **2012**, *116*, 3823. Benda, L.; Straka, M.; Sychrovský, V.; Bouř, P.; Tanaka, Y. *J. Phys. Chem. A* **2012**, *116*, 8313, Šebera, J.; Burda, J.; Straka, M.; Ono, A.; Kojima, C.; Tanaka, Y.; Sychrovský, V. *Chem. Eur. J.* **2013**, *19*, 9884.).

On natural and metallated NA we thus interpreted spectroscopic parameters in

terms of structure, solvation effects, metal coordination, and molecular flexibility. E.g. the covalent character of Hg-N bond between mercury and thymine in T-Hg-T “metallic” base pair was discovered by Raman spectroscopy. The **structure** containing tandem of consecutive T-Hg-T base pairs **was determined** and the chemical reaction describing formation of T-Hg-T base pair was proposed and calculated in full correspondence with experiments (Prof. Tanaka, Japan). The **metallophilic attraction** between the mercury atoms in two consecutive T-Hg-T base pairs was responsible for ca 9 % of their total stabilization, i.e. it is one of **principle stabilizing factors**. Theoretical modeling of **charge transport properties** of Hg-DNA unveiled structural and electronic details of the effect caused by mercury and measured by fluorescence spectroscopy (in collaborating laboratory of Doc. Vala, TU Brno, and Doc. Kratochvílová, FZU AV ČR).

We also casted new light on the DNA **base-excision mechanism**. In humans, it is directed by the 8-oxoguanine glycosylase 1 DNA repair protein (hOGG1). We found that the excision is dependent on the pyramidalization of the glycosidic nitrogen, larger for **damaged nucleobase** (e.g. by a radiation). Specific solvation of nucleosides in DNA G-quadruplex enforces the pyramidalization. Also the distinct N-pyramidalization of damaged guanosine (8-oxoG) enforced within hOGG1 catalytic pocket makes the excision easier. The predicted activation energy and reaction path agree with available experimental data.

Research Report of the team in the period 2010–2014

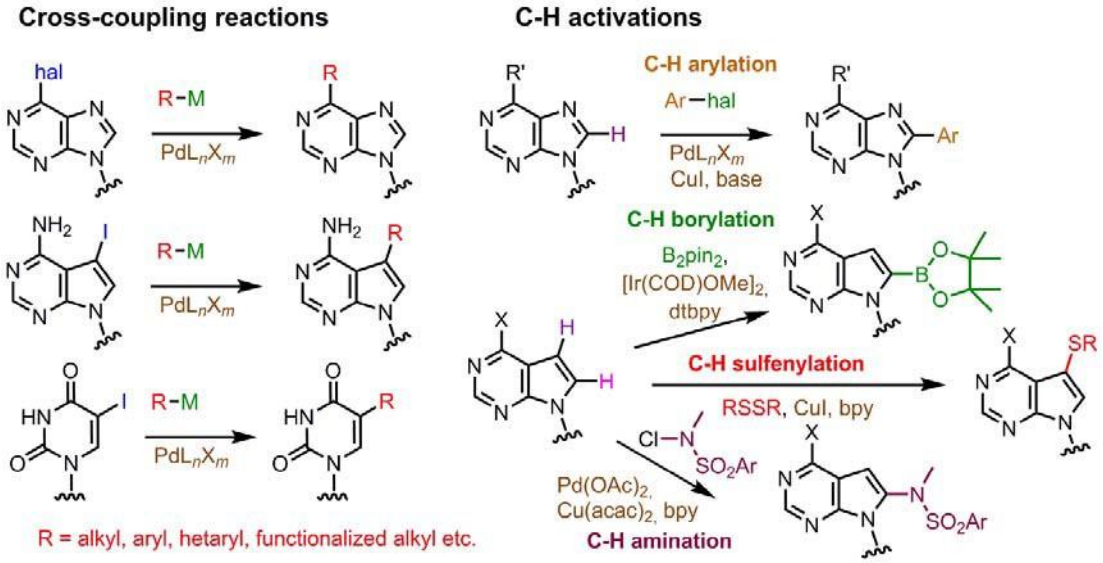
Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Michal Hocek - Bioorganic and Medicinal Chemistry of Nucleic Acids

Major focus of the Hocek group was on the synthetic, bioorganic and medicinal chemistry of base-modified nucleosides, nucleotides and nucleic acids aimed at applications in drug development, chemical biology and diagnostics. The Hocek group has published 75 papers and filed 3 patent applications during 2010-2014 (2 patents were granted in the US, EU, JP etc.).

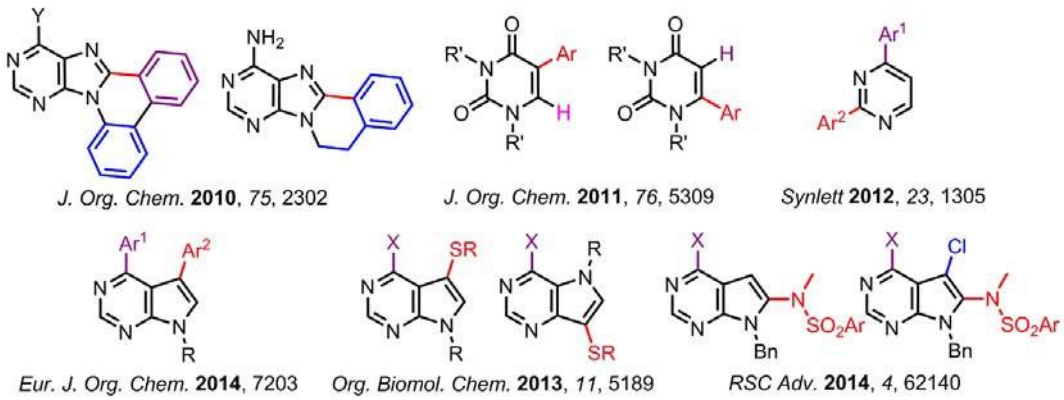
1. Development of synthetic methodology for modified nucleobases, nucleosides and nucleotides

In basic *synthetic methodology*, we have been developing novel cross-coupling and C-H activation reactions of purines, deazapurines and pyrimidines. In the cross-coupling reactions, we have finished the development of reactions of halogenated nucleobases with functionalized organometallics in order to introduce functionalized C-substituents and their follow-up functional group transformations.[1] The most important development was the direct aqueous-phase Suzuki and Sonogashira cross-coupling reactions of unprotected halogenated nucleosides, nucleotides and nucleoside triphosphates which enabled a one-step modification of unprotected nucleos(t)ides for medicinal chemistry and an easy access to substrates for enzymatic synthesis of modified nucleic acids (vide infra).



Major efforts were focused on the development of novel C-H activation reactions on purines, deazapurines and pyrimidines, as alternative or complementary reactions to cross-couplings. Previously, we have succeeded in the development of the first Pd-catalyzed Cu-mediated direct C-H arylation of purine bases and nucleosides, as well as their combinations with cross-coupling reactions and N-arylations for the synthesis of a large series of highly substituted derivatives and their intramolecular versions for the synthesis of fused purines. More recently, we successfully applied intramolecular C-H arylations for the synthesis of several novel fused purine heterocycles.[2] In pyrimidines, we have developed regioselective C-H arylations either at position 5 or 6 depending on the presence or absence of Cu-additives[3] and C-H trifluoromethtylation.[4] Moreover, we focuced on other C-H activations of deazapurines and succeeded in development of Ir-catalyzed C-H borylations, Cu-catalyzed C-H sulfenylations[5] and Pd/Cu-catalyzed C-H aminations[6] and applied these reactions for the synthesis of libraries of di- and trisubstituted deazapurine derivatives for biological activity testing. Several papers have also been published on development of orthogonal cross-couplings and C-H activations for the synthesis of disubstituted pyrimidines[7] and deazapurines.[8]

Examples of modified nucleobases/heterocycles by chemoselective cross-coupling and/or C-H activations:



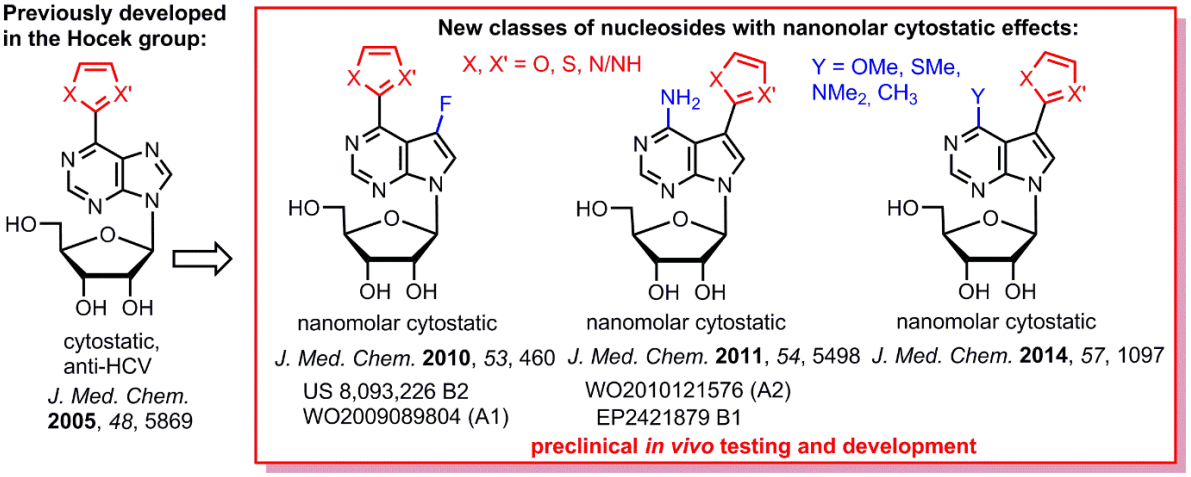
J. Org. Chem. **2010**, *75*, 2302 — *J. Org. Chem.* **2011**, *76*, 5309 — *Synlett* **2012**, *23*, 1305

Eur. J. Org. Chem. **2014**, 7203 — *Org. Biomol. Chem.* **2013**, *11*, 5189 — *RSC Adv.* **2014**, *4*, 62140

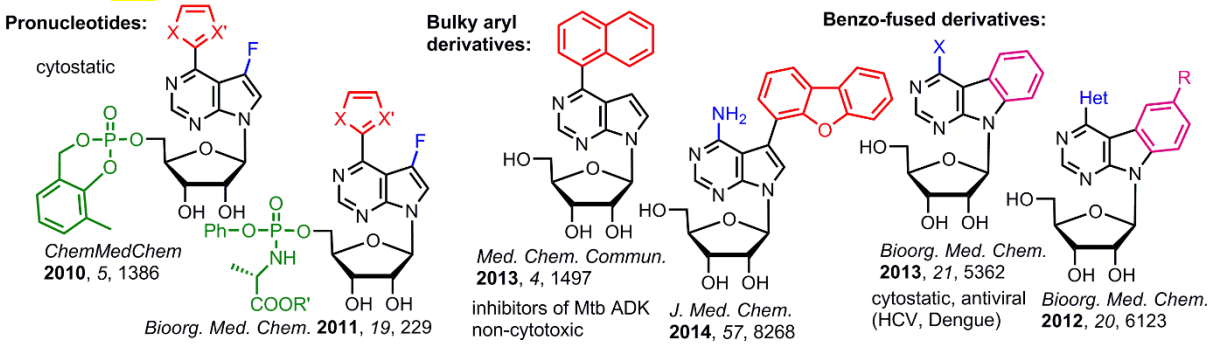
2. Medicinal chemistry of modified nucleosides

In medicinal chemistry, we have designed and prepared a library of ca. 1500 (3000 if counting previously prepared compounds) of new modified nucleobases, base-modified and sugar-modified nucleosides and nucleotides that have been tested for antiviral (HCV, HIV, HBV, Dengue, RSV – collaboration with Gilead Sciences and Novartis) and cytostatic (panel of cancer and leukemia cell lines – collaboration with Gilead and IMTM Palacký University) activity, as well as for inhibition of selected target enzymes (adenosine kinase, MAP kinase, purine nucleoside phosphorylase etc. – internal collaboration with Dr. Mertlíková-Kaiserová and Dr. Pichová), leading to the identification of several new biologically active lead structures. In all those collaborative projects, the collaborating companies or research groups were responsible for biological testing and profiling of compounds designed and synthesized in the Hocek group. Purine ribonucleosides bearing diverse functionalized *C*-substituents at the position 6 were found to possess submicromolar cytotoxic and anti-HCV effects but, due to low selectivity, their further development was discontinued. The most important findings

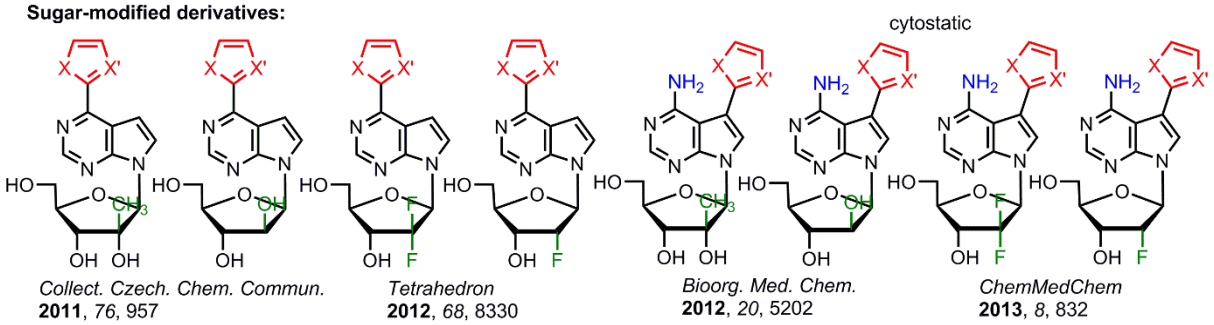
were the discovery of three novel classes of nanomolar cytostatics: 6-hetaryl-7-deazapurine ribonucleosides,[9] 7-hetaryl-7-deazaadenosines[10] and, most recently, 6-substituted 7-hetaryl-7-deazapurine ribonucleosides.[11] Due to the very high and selective activities against a broad panel of cancer and leukemia cell lines (they are non-toxic to non-proliferating cell lines), they were patented and studied in detail. The mechanism of action of 6-hetaryl-7-deazapurine involves the phosphorylation and inhibition of RNA polymerases, whereas the 7-hetaryl-7-deazaadenosines exert a complex mechanism involving incorporation of modified nucleotides both to DNA (causing DNA damage) and RNA (causing translation inhibition). The most active examples were selected from each group of nucleosides which are undergoing preclinical *in vivo* testing and pharmakokinetic studies and the preliminary results are very promising.



In order to increase the cellular uptake and improve the pharmakokinetic properties, we have prepared two types of phosphate prodrugs (cycloSal[12] and phosphoramidates[13]) but none of them improved the activity. Also we found that deazapurine nucleosides bearing bulky aryl-substituents at position 6 or 7 are potent and selective inhibitors of mycobacterial adenosine kinase (ADK).[14,15] Benzo-fused analogues of deazapurine nucleosides were prepared to reveal interesting antiviral and cytostatic activities.[16,17]



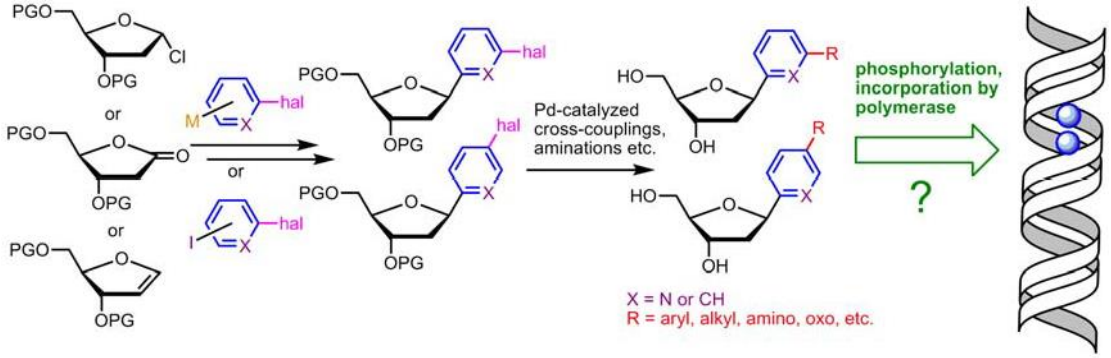
We have prepared diverse sugar-modified derivatives in order to investigate the SAR and possibly to achieve some selectivity towards HCV RNA-polymerase but most of them were less active than the parent ribonucleosides.[18-21]



3. Synthesis of modified *C*-nucleosides for chemical biology

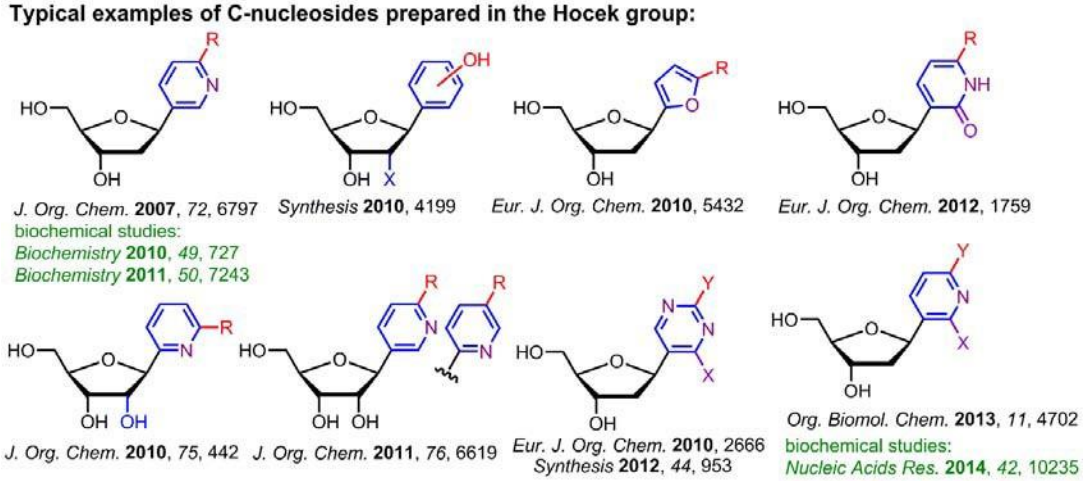
C-nucleosides are a very important class of modified nucleic acids components that have wide range of applications in medicinal chemistry and chemical biology. However, their synthesis is difficult and

most current methods suffer from low yields and selectivity. Therefore, we have developed a general and modular methodology of synthesis of diverse aryl and hetaryl *C*-nucleosides. Our approach is based on the preparation of halo-(het)aryl *C*-nucleoside intermediates and on follow-up Pd-catalyzed cross-coupling or nucleophilic substitution modifications (alkylation, arylation, amination, carbonylation, hydroxylation etc.). A large library of diverse aryl- and hetaryl C-nucleosides have been synthesized in the Hocek group.[22-30]



Other two modular methodologies based on allylic substitution[31] or indole formation[32] was developed in collaboration with Prof. Kocovsky or Prof. Kotora (minor contribution of the Hocek group). Most of the *C*-nucleosides were tested for biological activities and selected examples are used in collaborative projects in chemical biology (mechanistic studies of polymerases with R. D. Kuchta[33,34] and extension of the genetic alphabet with F. E. Romesberg[35] – these works were mostly done at the collaborating groups with minor contribution of the Hocek group).

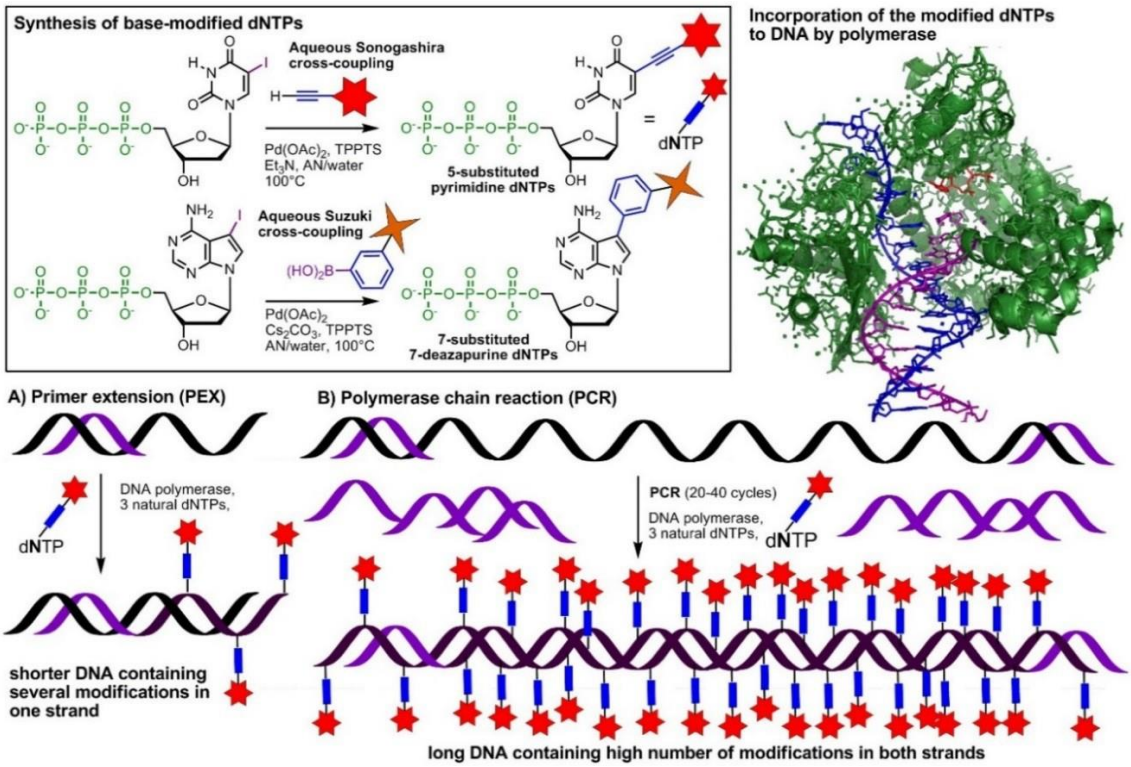
Typical examples of C-nucleosides prepared in the Hocek group:



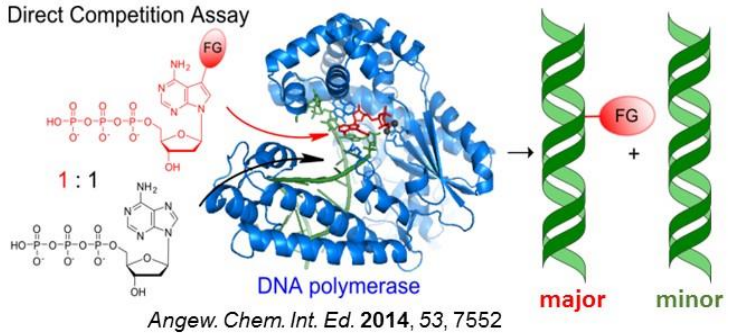
4. Bioorganic chemistry/chemical biology of base-modified nucleic acids

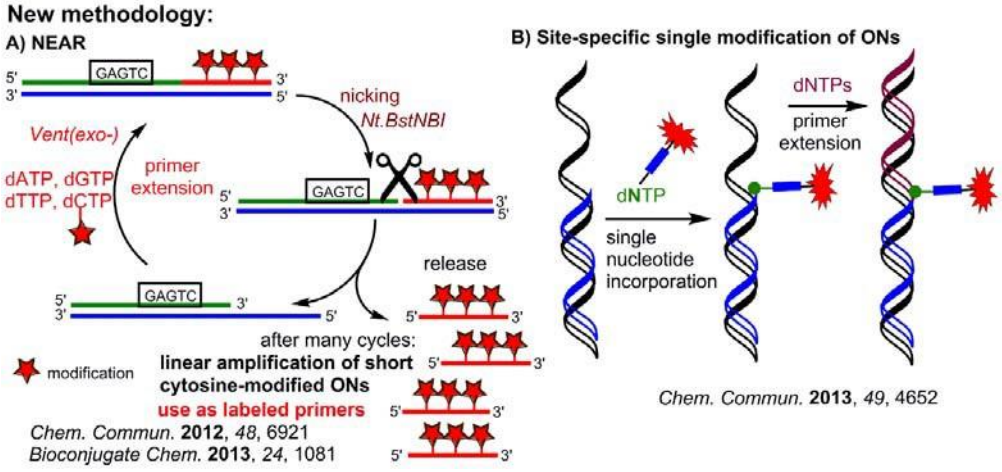
In bioorganic chemistry, we have developed a general and efficient two-step synthesis of base-modified nucleic acids based on cross-coupling modifications of halogenated nucleoside triphosphates followed by polymerase incorporation to DNA or RNA (for reviews see refs[36,37]). The first step is the direct aqueous-phase cross-coupling reaction (Suzuki or Sonogashira) of unprotected halogenated dNTPs (or NTPs) with substituted arylboronic acid or substituted terminal acetylene. The second step is the incorporation of the (d)NTPs to nucleic acids by polymerases. Most of our work dealt with incorporation of dNTPs to DNA (although successful experiments with RNA synthesis have also been performed). We found that 5-substituted pyrimidine and 7-substituted 7-deazapurine dNTPs are good substrates for at least some DNA polymerases (i.e. Pwo, Vent(exo-), KOD XL, DeepVent, DyNAzyme etc.) and most of them can be efficiently incorporated to DNA even if bulky substituents are attached to them. For the synthesis of shorter DNA sequences containing one or several modifications in one strand, we use primer extension (PEX) experiment that was also adapted for the isolation of modified single-strand oligonucleotides (ssONs) by using biotinylated template and magnetoseparation on streptavidine-coated magnetic beads. On the other hand, polymerase chain reaction (PCR) could be used for the construction of large dsDNA sequences containing high number of modifications in both strands. Very recently, terminal transferase was used for 3'-end labeling of ONs and nicking-enzyme amplification reaction for the synthesis of short modified ssONs (collaboration with the group of M.

Fojta from IPB, where the Hocek group did the synthesis and part of biochemistry, whereas Fojta group did part of biochemistry and all bioanalytical applications). [38]

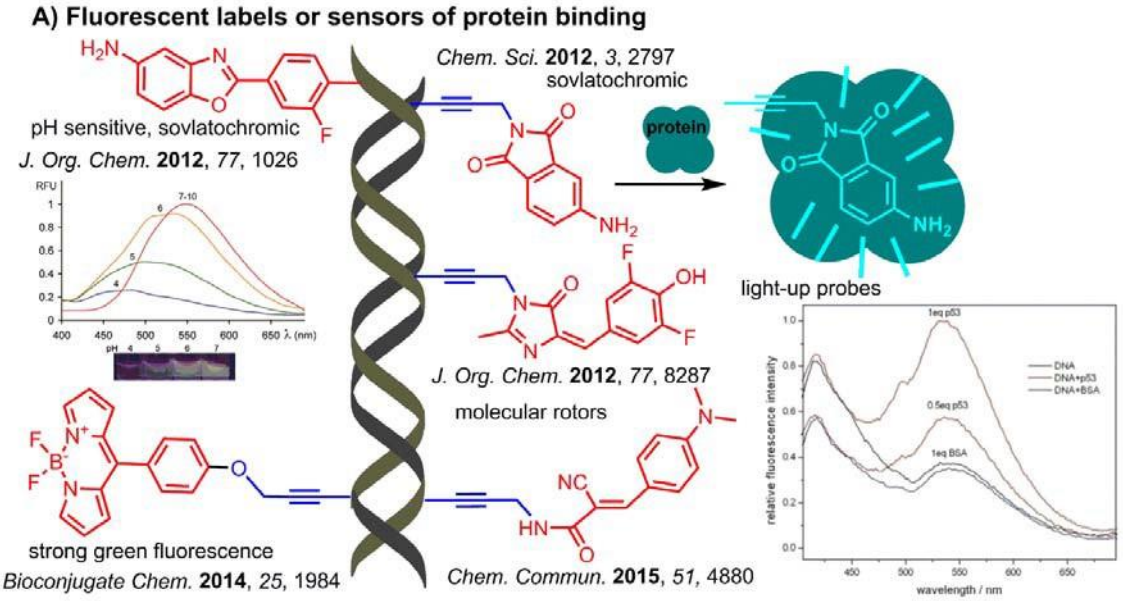


Our recent study of competitive incorporations of modified dNTPs in the presence of their natural counterparts showed a very surprising result that some 7-alkynyl- and 7-aryl-7-deazaadenine dNTPs are better substrates for the polymerases than natural dATP. [39] This finding is now applied in in vivo biosynthesis of modified DNA probes.

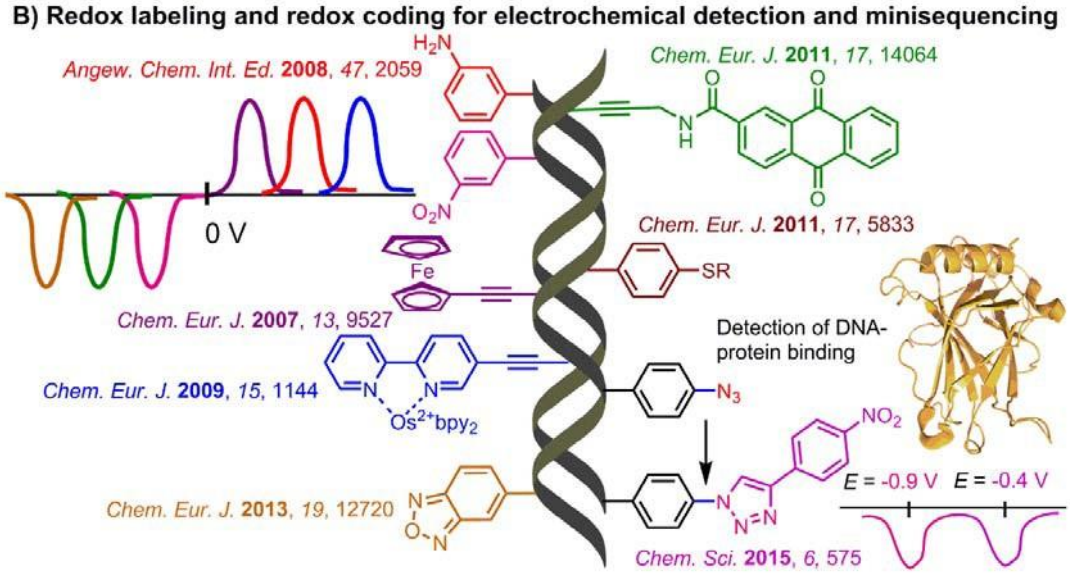




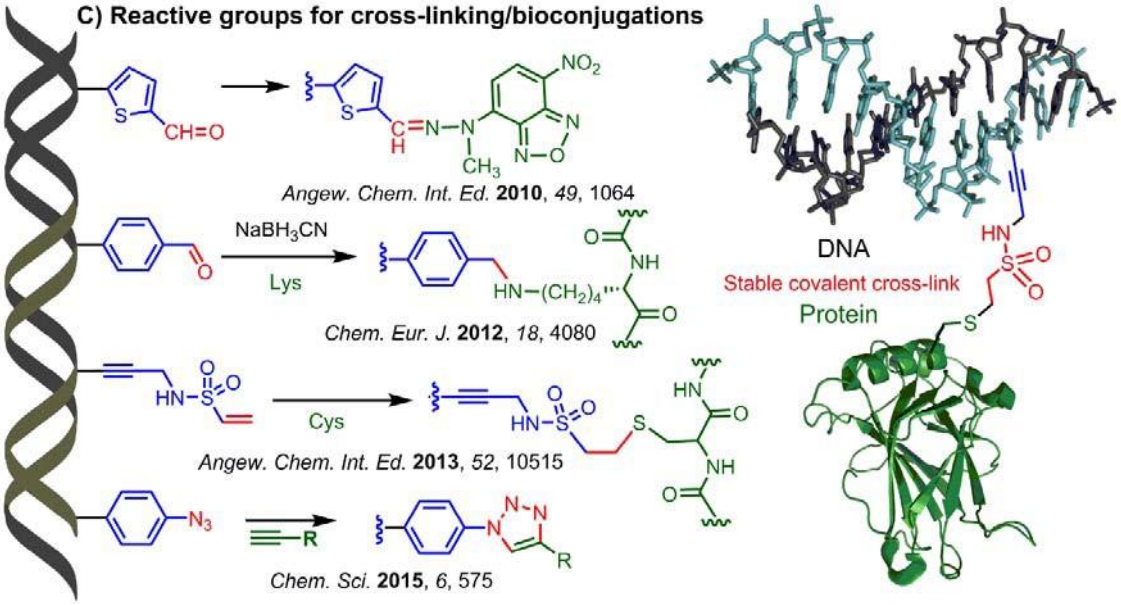
Our methodology efforts continued by development of Nicking Enzyme Amplification Reaction for synthesis of short modified ONs[40,41] and development of site specific single modification of ON probes (collaboration with the Fojta group who studied electrochemistry of the labelled ONs).[42]



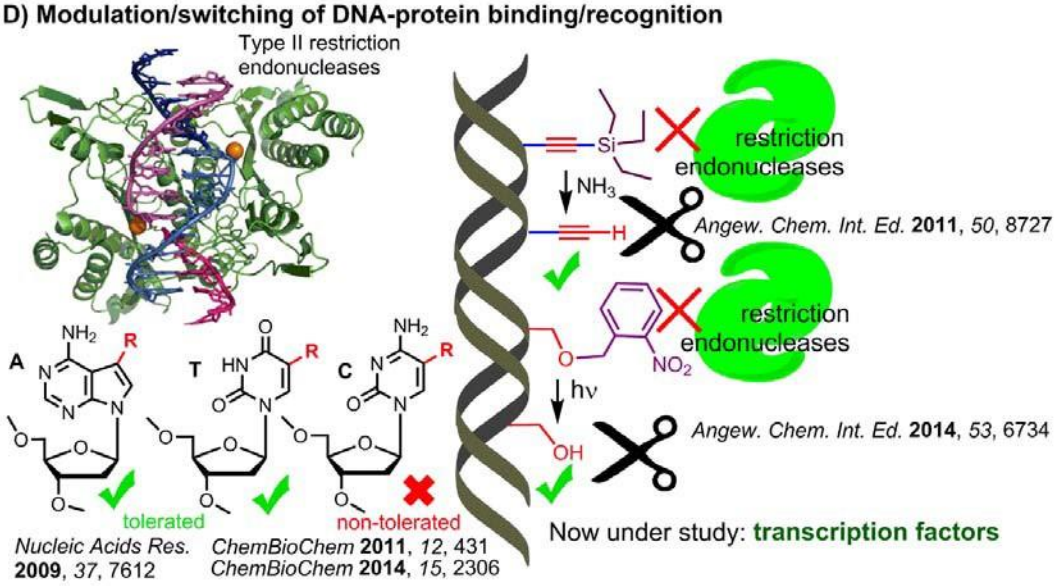
The areas of application of modified nucleic acids include bioanalysis and chemical biology. In diagnostics we focused mainly on fluorescent and redox labelling. We have developed several types of solvatochromic fluorescent labels (aminobenzoxazole[43] or aminophthalimide[44]) and incorporated these to DNA and ON probes for studies of their hybridization, SNP typing, pH microenvironment and/or protein binding. Molecular rotors (GFP-like[45] or Knovenagel adducts[46]) were also used for sending of protein –DNA interactions as light-up probes. Moreover, Bodipy-linked nucleotides were used for fluorescent labelling of DNA.[47]



Labeling of DNA or ON probes by redox-active groups that can be oxidized or reduced on electrodes was one of the major fields of application (collaboration with Prof. Fojta – in all these papers the synthesis and biochemistry was done in the Hocek group whereas the electrochemistry and analytical applications in the Fojta group). We have designed and prepared a number of dNTPs bearing diverse redox labels (nitro- and aminophenyl, ferrocene, Os(bpy)$_3$, anthraquinone,[48] sulfides,[49] methoxyphenol,[50] hydrazones,[51] acrylates,[52] benzofurazan[53] or azidophenyl[54]) and successfully incorporated them to ssONs. These redox-labeled ON probes were studied by electrochemistry and applied in bioanalysis for electrochemical detection (hybridization, SNP typing, minisequencing etc.). Multi-potential redox coding for DNA bases[53] and the use in study of DNA-protein interactions[54] are currently under development.



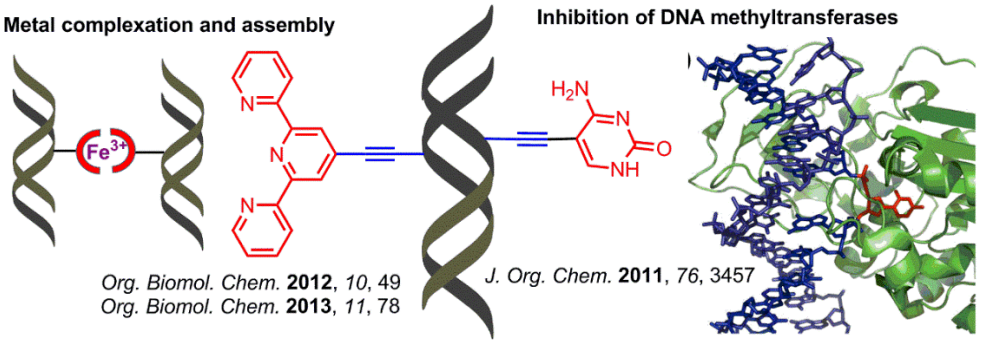
Other areas of applications are in chemical biology. Incorporation of reactive groups to DNA was developed for bioconjugations and cross-linking with proteins. A series of steroid-DNA conjugates was prepared through PEX or PCR.[55] Aldehyde-modified nucleotides were incorporated to DNA and used for staining by hydrazone-formation[56] or for bioconjugations with peptides by reductive amination.[57] Michael acceptor (vinylsulfonamide) was also incorporated to DNA and used for cross-linking with Cys-containing peptides and proteins.[58] Azidophenyl-modification has been developed for bioconjugations through Cu-catalyzed alkyne-azide click reaction.[54]



Our systematic study of cleavage of base-modified DNA by restriction endonucleases revealed surprisingly high tolerance of some REs to modifications at A or T (but not C) within the recognition sequences.[59,60] This eventually led to the development of transient switchable protection of DNA against the cleavage by the restriction endonucleases by the introduction of (triethylsilyl)ethynyl group at 7-deazaA (cleavage switched off) which can be deprotected by ammonia and the resulting acetylene-modified DNA is then cleaved by REs.[61] More recently, we have developed a major-groove photocaging where the cleavage by RE enzymes can be activated by UV irradiation.[62] Also we used the major-groove protection for a novel approach for gene cloning and expression.[63] The chemical modulation (protection and triggering) of a specific binding of a protein to DNA may open a new field of "artificial chemical epigenetics".

Other applications include incorporation of bi- and terpyridine ligands for metal complexations and assembly.[64,65] Further we designed double-headed nucleotides containing another nucleobase in major groove to mimic flipped-out nucleobase by DNA methyltransferase and the modified DNA was found to inhibit DNA methylation.[66]

E) Other applications



Some other collaborative works with minor contribution from our lag (the synthesis of the studied compounds) in NMR spectroscopy[67-69] (collaboration with prof. Marek from MU Brno), synthesis of C-glycosides (Kotora group),[70,71] electrochemistry[72-74] (Fojta group) and DNA cross-linking[75] (Madder group from Ghent) were published during the evaluated period. One methodology paper was published in Current Protocols in Chemical Biology.[76]

References [members of the Hocek group underlined, corresponding author(s) marked with *]:

1. Hasník, Z.; Pohl, R.; Hocek, M.* "Synthesis of (Purin-6-yl)methylphosphonate Bases and Nucleosides" *Tetrahedron Lett*. **2010**, *51*, 2464-2466.
2. Čerňa, I.; Pohl, R.; Klepetářová, B.; Hocek, M.* "Intramolecular Direct C-H Arylation Approach to Fused Purines. Synthesis of Purino[8,9-f]phenanthridines and 5,6-Dihydropurino[8,9-a]isoquinolines" *J. Org. Chem*. **2010**, *75*, 2302-2308.
3. Čerňová, M.; Čerňa, I.; Pohl, R.; Hocek, M.* "Regioselective Direct C-H Arylations of Protected Uracils. Synthesis of 5- and 6-Aryluracil Bases" *J. Org. Chem.* **2011**, *76*, 5309-5319.
4. Čerňová, M.; Pohl, R.; Klepetářová, B.; Hocek, M.* "C-H Trifluoromethylations of 1,3-Dimethyluracil and Reactivity of the Products in C-H Arylations" *Heterocycles* **2014**, *89*, 1159-1171.
5. Klečka, M.; Pohl, R.; Čejka, J.; Hocek, M.* "Direct C-H sulfenylation of purines and deazapurines" *Org. Biomol. Chem.* **2013**, *11*, 5189-5193.
6. Sabat, N.; Klečka, M.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "Direct C-H amination and C-H chloroamination of 7-deazapurines" *RSC Adv.* **2014**, *4*, 62140-62143.
7. Čerňová, M.; Pohl, R.; Klepetářová, B.; Hocek, M.* "A general regioselective synthesis of 2,4-diarylpyrimidines from 2-thiouracil through two orthogonal cross-coupling reactions" *Synlett* **2012**, *23*, 1305-1308.
8. Krömer, M.; Klečka, M.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "Chemoselective synthesis of 4,5-diarylpyrrolo[2,3-*d*]pyrimidines (6,7-diaryl-7-deazapurines) by consecutive Suzuki and Liebeskind-Srogl cross-couplings" *Eur. J. Org. Chem*. **2014**, 7203-7210.
9. Nauš, P.; Pohl, R.; Votruba, I.; Džubák, P.; Hajdúch, M.; Ameral, R.; Birkuš, G.; Wang, T.; Ray, A. S.; Mackman, R.; Cihlar, T.; Hocek, M.* "6-(Het)aryl-7-Deazapurine Ribonucleosides as Novel Potent Cytostatic Agents" *J. Med. Chem.* **2010**, *53*, 460-470.
10. Bourderioux, A.; Nauš, P.; Perlíková, P.; Pohl, R.; Pichová, I.; Votruba, I.; Džubák, P.; Konečný, P.; Hajdúch, M.; Stray, K. M.; Wang, T.; Ray, A. S.; Feng, J. Y.; Birkus, G.; Cihlar, T.; Hocek, M.* "Synthesis and significant cytostatic activity of 7-hetaryl-7-deazaadenosines" *J. Med. Chem.* **2011**, *54*, 5498-5507.
11. Nauš, P.; Caletková, O.; Konečný, P.; Džubák, P.; Bogdanová, K.; Kolář, M.; Vrbková, J.; Slavětínská, L.; Tloušťová, E.; Perlíková, P.; Hajdúch M.; Hocek, M.* "Synthesis, cytostatic, antimicrobial and anti-HCV activity of 6-substituted 7-(het)aryl-7-deazapurine ribonucleosides" *J. Med. Chem.* **2014**, *57*, 1097-1110.
12. Spáčilová, P.; Nauš, P.; Pohl, R.; Votruba, I.; Snášel, J.; Zábranská, H.; Pichová, I.; Ameral, R.; Birkuš, G.; Cihlář, T.; Hocek, M.* "CycloSal-Phosphate Pronucleotides of Cytostatic 6-(Het)aryl-7-Deazapurine Ribonucleosides: Synthesis, Cytostatic Activity, and Inhibition of Adenosine Kinases" *ChemMedChem* **2010**, *5*, 1386–1396.
13. Perlíková, P.; Pohl, R.; Votruba, I.; Shih, R.; Birkuš, G.; Cihlář, T.; Hocek, M.* "Phosphoramidate pronucleotides of cytostatic 6-aryl-7-deazapurine ribonucleosides" *Bioorg. Med. Chem.* **2011**, *19*, 229-242.
14. Perlíková, P.; Konečný, P.; Nauš, P.; Snášel, J.; Votruba, I.; Džubák, P.; Pichová, I.; Hajduch, M.; Hocek, M.* "6-Alkyl-, 6-Aryl- or 6-Hetaryl-7-deazapurine Ribonucleosides as Inhibitors of Human or MTB Adenosine Kinase and Potential Antimycobacterial Agents" *Med. Chem. Commun.* **2013**, *4*, 1497-1500.
15. Snášel, J.; Nauš, P.; Dostál, J.; Hnízda, A.; Fanfrlík, J.; Brynda, J.; Bourderioux, A.; Dušek, M.; Dvořáková, H.; Stolaříková, J.; Zábranská, H.; Pohl, R.; Konečný, P.; Džubák, P.; Votruba, I.; Hajdúch, M.; Řezáčová, P.; Veverka, V.; Hocek, M.*; Pichová, I.* "Structural basis for inhibition of mycobacterial and human adenosine kinase by 7-substituted 7-(het)aryl-7-deazaadenine ribonucleosides" *J. Med. Chem.* **2014**, *57*, 8268-8279.

16. Tichý, M.; Pohl, R.; Xu, H.Y.; Chen, Y.-L.; Yokokawa, F.; Shi, P.-Y.; Hocek, M.* "Synthesis and antiviral activity of 4,6-disubstituted pyrimido[4,5-*b*]indole ribonucleosides" *Bioorg. Med. Chem.* **2012**, *20*, 6123-6133.
17. Tichý, M.; Pohl, R.; Tloušťová, E.; Weber, J.; Bahador, G.; Lee, Y.-J.; Hocek, M.* "Synthesis and Biological Activity of Benzo-Fused 7-Deazaadenosine Analogues. 5- and 6-Substituted 4-Amino- or 4-Alkylpyrimido[4,5-b]indole Ribonucleosides" *Bioorg. Med. Chem.* **2013**, *21*, 5362-5372.
18. Nauš, P.; Perlíková, P.; Pohl, R.; Hocek, M.* "Sugar-modified derivatives of cytostatic 6-(het)aryl-7-deazapurine nucleosides: 2'-C-methylribonucleosides, arabinonucleosides and 2'-deoxy-2'-fluoroarabinonucleosides" *Collect. Czech. Chem. Commun*. **2011**, *76*, 957-988.
19. Perlíková, P.; Jornet Martinez, N.; Slavětínská, L.; Hocek, M.* "Synthesis of 2'-Deoxy-2'-fluororibo- and 2'-Deoxy-2',2'-difluororibonucleosides Derived from 6-(Het)aryl-7-deazapurines" *Tetrahedron* **2012**, *68*, 8300-8310.
20. Nauš, P.; Perlíková, P.; Bourderioux, A.; Pohl, R.; Slavětínská, L.; Votruba, I.; Bahador, G.; Birkuš, G.; Cihlář, T.; Hocek, M.* "Sugar-modified derivatives of cytostatic 7-(het)aryl-7-deazaadenosines: 2'-C-methylribonucleosides, 2'-deoxy-2'-fluoroarabinonucleosides, arabinonucleosides and 2'-deoxyribonucleosides" *Bioorg. Med. Chem.* **2012**, *20*, 5202-5214.
21. Perlíková, P.; Eberlin, L.; Ménová, P.; Raindlová, V.; Slavětínská, L.; Tloušťová, E.; Bahador, G.; Lee, Y.-J.; Hocek, M.* "Synthesis, Cytostatic and Antiviral activity of 2'-Deoxy-2',2'-difluororibo- and 2'-Deoxy-2'-fluororibonucleosides Derived from 7-(Het)aryl-7-deazaadenines" *ChemMedChem*, **2013**, *8*, 832-846.
22. Štefko, M.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "A General and Efficient Synthesis of Pyridin-2-yl C-Ribonucleosides Bearing Diverse Alkyl, Aryl, Amino and Carbamoyl Groups in Position 6" *J. Org. Chem.* **2010**, *75*, 442-449.
23. Kubelka, T.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "Synthesis of 2,4-Disubstituted Pyrimidin-5-yl C-2'-Deoxyribonucleosides by Sequential Regioselective Reactions of 2,4-Dichloropyrimidine Nucleosides" *Eur. J. Org. Chem.* **2010**, 2666-2669.
24. Bárta, J.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "Modular Synthesis of 5-Substituted Furan-2-yl *C*-2'-Deoxyribonucleosides and Biaryl Covalent Base-pair Analogues" *Eur. J. Org. Chem.* **2010**, 5432-5443.
25. Štefko, M.; Hocek, M.* "Synthesis of Phenol and Pyridone C-ribo-and 2'-deoxyribonucleosidesby Pd-catalyzed Hydroxylations of Haloaryl C-nucleosides" *Synthesis* **2010**, 4199–4206.
26. Štefko, M.; Slavětínská, L.; Klepetářová, B.; Hocek, M.* "General and Modular Synthesis of Isomeric 5-Substituted Pyridine-2-yl and 6-Substituted Pyridine-3-yl C-Ribonucleosides Bearing Diverse Alkyl, Aryl, Hetaryl, Amino, Carbamoyl and Hydroxy Groups" *J. Org. Chem.* **2011**, *76*, 6619-6635.
27. Kubelka, T.; Slavětínská, L.; Hocek, M.* "A General Regioselective Approach to 2,4-Disubstituted Pyrimidine-5-yl C-2'-Deoxyribonucleosides" *Synthesis* **2012**, *44*, 953-965.
28. Chapuis, H.; Joubert, N.; Kubelka, T.; Pohl, R.; Hocek, M.* "Synthesis of 6-substituted 2(1H)-pyridon-3-yl C-2'-deoxyribonucleosides" *Eur. J. Org. Chem.* **2012**, 1759-1767.
29. Kubelka, T.; Slavětínská, L.; Hocek, M.* "Synthesis of Substituted Benzyl Homo-*C*-Ribonucleosides and -Nucleotides as Carba-Analogues of Phosphoribosylanthranilate" *Eur. J. Org. Chem.* **2012**, 4969-4981.
30. Kubelka, T.; Slavětínská, L.; Eigner, V.; Hocek, M.* "Synthesis of 2,6-Disubstituted Pyridin-3-yl C-2'-Deoxyribonucleosides through Chemoselective Transformations of Bromo-chloropyridine C-Nucleosides" *Org. Biomol. Chem.* **2013**, *11*, 4702-4718.
31. Štambaský, J.; Kapras, V.; Štefko, M.; Kysilka, O.; Hocek, M.; Malkov, A. V.; Kočovský, P.* "A Modular Approach to Aryl-C-ribonucleosides via the Allylic Substitution and Ring-Closing Metathesis Sequence. A Stereocontrolled Synthesis of All Four alpha-/beta- and D-/L-C-Nucleoside Stereoisomers" *J. Org. Chem.* **2011**, *76*, 7781-7803.
32. Nečas, D.; Hidasová, D.; Hocek, M.; Kotora, M.* "Modular Synthesis of 1-α- and 1-β-(Indol-2-yl)-2'-deoxyribose C-Nucleosides" *Org. Biomol. Chem.* **2011**, *9*, 5934-5937.
33. Lund, T.; Cavanaugh, N.; Joubert, N.; Urban, M.; Patro, J.; Hocek, M.; Kuchta, R.D.* "B Family DNA Polymerases Asymmetrically Recognize Pyrimidines and Purines" *Biochemistry* **2011**, *50*, 7243-7250.
34. Urban, M.; Joubert, N.; Purse, B.; Hocek, M.; Kuchta, R.* "Mechanisms by which Human DNA Primase Chooses to Polymerize a NTP" *Biochemistry* **2010**, *49*, 727-735.
35. Dhami, K.; Malyshev, D. A.; Ordoukhanian, P.; Kubelka, T.; Hocek, M.; Romesberg, F. E.* "Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of the genetic alphabet" *Nucleic Acids Res*. **2014**, *42*, 10235–10244.
36. Hocek, M.*; Fojta, M.* "Nucleobase Modification as Redox DNA Labelling for Electrochemical Detection" *Chem. Soc. Rev.* **2011**, *40*, 5802-5814.
37. Hocek, M.* "Synthesis of base-modified 2'-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified DNA for applications in bioanalysis and chemical biology" *J. Org. Chem.* **2014**, *79*, 9914-9921.
38. Horáková, P.; Macíčková-Cahová, H.; Pivoňková, H.; Špaček, J.; Havran, L.; Hocek, M.*; Fojta, M.* "Tail-Labelling of DNA Probes Using Modified Deoxynucleotide Triphosphates and Terminal Deoxynucleotidyl Tranferase. Application in Electrochemical DNA Hybridization and Protein-DNA Binding Assays" *Org. Biomol. Chem.* **2011**, *9*, 1366-1371.
39. Kielkowski, P.; Fanfrlík, J.; Hocek, M.* "7-Aryl-7-deazaadenine 2'-Deoxyribonucleoside Triphosphates (dNTPs): Better Substrates for DNA polymerases than dATP in Competitive Incorporations" *Angew. Chem. Int. Ed.* **2014**, *53*, 7552-7555.

40. Ménová, P.; Hocek, M.* "Preparation of Short Cytosine-Modified Oligonucleotides by Nicking Enzyme Amplification Reaction" *Chem. Commun.* **2012**, *48*, 6921-6923.
41. Ménová, P.; Raindlová, V.; Hocek, M.* "The Scope and Limitation of the Nicking Enzyme Amplification Reaction for the Synthesis of Base-Modified Oligonucleotides and Primers for PCR" *Bioconjugate Chem.* **2013**, ***24*, 1081-1093**.
42. Ménová, P.; Cahová, H.; Plucnara, M.; Havran, L.; Fojta, M.; Hocek, M.* "Polymerase Synthesis of Oligonucleotides Containing a Single Chemically Modified Nucleobase for Site-Specific Redox Labelling" *Chem. Commun.*, **2013**, *49*, 4652-4654.
43. Riedl, J.; Pohl, R.; Rulíšek, L.; Hocek, M.* "Synthesis and Photophysical Properties of Biaryl-Substituted Nucleos(t)ides. Polymerase Synthesis of DNA Probes Bearing Solvatochromic and pH Sensitive Dual Fluorescent and ^{19}F NMR Labels" *J. Org. Chem.* **2012**, *77*, 1026-1044.
44. Riedl, J.; Pohl, R.; Ernsting, N. P.; Orsag, P.; Fojta, M.; Hocek M.* "Labelling of nucleosides and oligonucleotides by solvatochromic 4-aminophthalimide fluorophore for studying DNA-protein interactions" *Chem. Sci.* **2012**, *3*, 2797-2806.
45. Riedl, J.; Ménová, P.; Pohl, R.; Orság, P.; Fojta, M.; Hocek, M.* "GFP-like fluorophores as DNA labels for studying DNA-protein interactions" *J. Org. Chem.* **2012**, *77*, 8287-8293.
46. Dziuba, D.; Pohl, R.; Hocek, M.* "Polymerase synthesis of DNA labelled with benzylidene cyanoacetamide-based fluorescent molecular rotors: fluorescent light-up probes for DNA-binding proteins" *Chem. Commun.* **2015**, ***51***, 4880-4882.
47. Dziuba, D.; Pohl, R.; Hocek, M.* "Bodipy-labelled nucleoside triphosphates for polymerase synthesis of fluorescent DNA" *Bioconjugate Chem*. **2014**, *25*, 1984-1995.
48. Balintová, J.; Pohl, R.; Horáková, P.; Vidláková, P.; Havran, L.; Fojta, M.*; Hocek, M.* "Anthraquinone as a redox label for DNA. Synthesis, enzymatic incorporation and electrochemistry of anthraquinone-modified of nucleosides, nucleotides and DNA" *Chem. Eur. J.* **2011**, *17*, 14063-14073.
49. Macíčková-Cahová, H.; Pohl, R.; Horáková, P.; Havran, L.; Špaček, J.; Fojta, M.*; Hocek, M.* "Alkylsulfanylphenyl derivatives of cytosine and 7-deazaadenine nucleosides, nucleotides and nucleoside triphosphates. Synthesis, polymerase incorporation to DNA and electrochemical study" *Chem. Eur. J.* **2011**, *17*, 5833-5841.
50. Simonova, A.; Balintová, J.; Pohl, R.; Havran, L.; Fojta, M.*; Hocek, M.* "Methoxyphenol and dihydrobenzofuran as new oxidizable labels for electrochemical detection of DNA" *ChemPlusChem* **2014**, *79*, 1703-1712.
51. Raindlová, V.; Pohl, R.; Klepetářová, B.; Havran, L.; Šimková, E.; Horáková, P.; Pivoňková, H.; Fojta, M.*; Hocek, M.* "Synthesis of hydrazone-modified nucleotides and their polymerase incorporation to DNA for redox labelling" *ChemPlusChem* **2012**, *77*, 652-662.
52. Dadová, J.; Vidláková, P.; Pohl, R.; Havran, L.; Fojta, M.; Hocek, M.* "Aqueous Heck cross-coupling preparation of acrylate-modified nucleotides and nucleoside triphosphates for polymerase synthesis of acrylate-labeled DNA" *J. Org. Chem.* **2013**, *78*, 9627-9637.
53. Balintová, J.; Plucnara, M.; Vidláková, P.; Pohl, R.; Havran, L.; Fojta, M.*; Hocek, M.* "Benzofurazane as a New Redox Label for Electrochemical detection of DNA. Towards Multipotential Redox Coding of DNA Bases" *Chem. Eur. J.* **2013**, *19*, 12720-12731.
54. Balintová, J.; Špaček, J.; Pohl, R.; Brázdová, M.; Havran, L.; Fojta, M.*; Hocek, M.* "Azidophenyl as a click-transformable redox label of DNA suitable for electrochemical detection of DNA-protein interactions" *Chem. Sci.* **2015**, *6*, 575-587.
55. Ikonen, S.; Macíčková-Cahová, H.; Pohl, R.; Šanda, M.; Hocek, M.* "Synthesis of nucleoside and nucleotide conjugates of bile acids and polymerase construction of bile acid-functionalized DNA" *Org. Biomol. Chem.* **2010**, *8*, 1194-1201.
56. Raindlová, V.; Pohl, R.; Šanda, M.; Hocek, M.* "Direct polymerase synthesis of reactive aldehyde-functionalized DNA and its conjugation and staining with hydrazines" *Angew. Chem. Int. Ed.* **2010**, *49*, 1064-1066.
57. Raindlová, V.; Pohl, R.; Hocek, M.* "Synthesis of aldehyde-linked nucleotides and DNA and their bioconjugations with lysine and peptides through reductive amination" *Chem. Eur. J.* **2012**, *18*, 4080-4087.
58. Dadová, J.; Orság, P.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M.* "Vinylsulfonamide and acrylamide modification of DNA for cross-linking with proteins" *Angew. Chem. Int. Ed.* **2013**, *52*, 10515-10518.
59. Macíčková-Cahová, H.; Pohl, R.; Hocek, M.* "Cleavage of functionalized DNA containing 5-modified pyrimidines by type II restriction endonucleases" *ChemBioChem* **2011**, *12*, 431-438.
60. Mačková, M.; Pohl, R.; Hocek, M.* "Polymerase synthesis of DNA bearing vinyl groups in major groove and their cleavage by restriction endonucleases" *ChemBioChem* **2014**, *15*, 2306-2312.
61. Kielkowski, P.; Macíčková-Cahová, H.; Pohl, R.; Hocek, M.* "Transient and switchable (triethylsilyl)ethynyl protection of DNA against cleavage by restriction endonucleases" *Angew. Chem. Int. Ed.* **2011**, *50*, 8727-8730.
62. Vaníková, Z.; Hocek, M.* "Polymerase Synthesis of Photocaged DNA Resistant against Cleavage by Restriction Endonucleases" *Angew. Chem. Int. Ed.* **2014**, *53*, 6734-6737.
63. Kielkowski, P.; Brock, N. L.; Dickschat. J. S.*; Hocek, M.* "Nucleobase Protection Strategy for Gene Cloning and Expression" *ChemBioChem*, **2013**, *14*, 801-804.

64. Kalachova, L.; Pohl, R.; Hocek, M.* "Synthesis of nucleoside mono- and triphosphates bearing oligopyridine ligands, their incorporation into DNA and complexation with transition metals" *Org. Biomol. Chem.* **2012**, *10*, 49-55.
65. Kalachova, L.; Pohl, R.; Bednárová, L; Fanfrlík, J.; Hocek, M.* "Synthesis of nucleosides and dNTPs bearing oligopyridine ligands linked through octadiyne tether, their incorporation into DNA and complexation with transition metal cations" *Org. Biomol. Chem.* **2013**, *11*, 78-89.
66. Kielkowski, P.; Pohl, R.; Hocek, M.* "Synthesis of acetylene linked double-nucleobase nucleos(t)ide building blocks and polymerase construction of DNA containing cytosines in the major groove" *J. Org. Chem.* **2011**, *76*, 3457-3462.
67. Standara, S.; Maliňáková, K.; Marek, R.; Marek, J.; Hocek, M.; Vaara, J; Straka, M.* "Understanding the NMR Chemical Shifts for 6-Halopurines: Role of Structure, Solvent and Relativistic Effects" *Phys. Chem. Chem. Phys*. **2010**, *12*, 5126-5139.
68. Marek, R.*; Křístková, A.; Maliňáková, K.; Toušek, J.; Marek, J.; Hocek, M.; Malkina, O. L.; Malkin, V. G. "Interpretation of Indirect Nuclear Spin-spin Couplings in Isomers of Adenine: Novel Approach to Analyze Coupling Electron Deformation Density using Localized Molecular Orbitals" *J. Phys. Chem*. **2010**, *114*, 6689-6700.
69. Standara, S.; Maliňáková, K.; Straka, M. Zacharová, Z.; Hocek, M.; Marek, J.; Marek, R.* "Interpretation of Substituent Effects on ^{13}C and ^{15}N NMR chemical shifts in 6-Substituted Purines" *Phys. Chem. Chem. Phys.* **2011**, *13*, 15854-15864.
70. Bobula, T.; Hocek, M.; Kotora, M.* "Sonogashira reactions of α- and β-1-ethynyl-2-deoxyribosides. Synthesis of acetylene-extended C-nucleosides" *Tetrahedron* **2010**, *66*, 530-536.
71. Opekar, S.; Turek, P.; Pohl, R.; Klepetářová, B.; Votruba, I.; Hocek, M.; Kotora, M.* "[2+2+2]-Cocyclotrimerization of 6-Alkynyl-7-benzylpurines with α,ω-Diynes" *Heterocycles* **2010**, *82*, 895-907.
72. Fojta, M.*; Havran, L.; Pivoňková, H.; Horáková, P.; Hocek, M. "Redox Labels and Indicators Based on Transition Metals and Organic Electroactive Moieties for Electrochemical Nucleic Acids Sensing" *Curr. Org. Chem*. **2011**, 15, 2936-2949.
73. Daňhel, A.; Raindlová, V.; Havran, L.; Pivoňková, H.; Hocek, M.; Fojta, M.* "Electrochemical Behaviour of 2,4-Dinitrophenylhydrazi(o)ne as Multi-Redox Centre DNA Label at Mercury Meniscus Modified Silver Solid Amalgam Electrode" *Electrochim. Acta* **2014**, *126*, 122-131.
74. Daňhel, A.; Raindlová, V.; Havran, L.; Barek, J.; Hocek, M.; Fojta, M.* "Voltammetric Study of dsDNA Modified by Multi-redox Label Based on *N*-methyl-4-hydrazino-7-nitrobenzofurazan" *Electrochim. Acta* **2014**, *129*, 348-357.
75. Stevens, K.; Claeys, D.; Catak, S.; Figaroli, S.; Hocek, M.; Tromp, J. M.; Schürch, S.; Van Speybroeck, V.; Madder, A.* "Furan-oxidation triggered inducible DNA cross-linking: acyclic versus cyclic furan containing building blocks. On the benefit of restoring the cyclic sugar backbone" *Chem. Eur. J.* **2011**, *17*, 6940-6953.
76. Macíčková-Cahová, H.; Vrábel, M.; Hocek, M. "Cross-coupling Modification of Nucleoside Triphosphates, PEX and PCR Construction of Base-Modified DNA" *Curr. Protoc. Chem. Biol.* **2010**, *2*, 1-14.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Irena Valterová - Chemical Ecology of Social Insects

Research Team of Infochemicals groups together specialists in various fields of chemistry and biology with a common interest in chemical ecology and lipidomics of arthropods, namely in two groups of social insects of prime ecological and economic importance, bumblebees and termites. Our diversified approaches combine various methods of analytical and organic chemistry and biochemistry with complementary biological techniques, such as anatomy, behavioural analysis, and electrophysiology. The aspects of molecular biology underlying our topics are studied in collaboration with the IOCB team of I. Pichová.

The goal of the team is to obtain knowledge on chemistry, biochemistry and biological significance of arthropod semiochemicals and defensive compounds, including the biochemistry of lipids and its relation to the production of semiochemicals. Beside the general focus on basic research, we are aiming our projects at particular aspects of life history of pest species with a future prospect of their control. In this respect, we extend our field of interest to economically important solitary insects, such as pest species of Lepidoptera (leaf miners, bumblebee wax moths), Diptera (fruit flies), Coleoptera (bark beetles), other insect orders.

In the period of evaluation, the team has published scientific papers covering following research areas: i) bumblebee pheromones, their biosynthesis and related exocrine glands, ii) termite pheromones and defensive chemicals and related anatomic adaptations, iii) semiochemicals of (pest) insects (pheromones, kairomones, defensive substances), iv) insect lipids and lipidomics, and v) enzymes hypothetically participating in pheromone formation and lipid mobilization. The team has been using, optimizing, and developing methods of extraction and chemical analysis of both volatile and non-volatile substances (GC/MS, GC×GC/MS, HRGC/MS, LC/MS, MALDI), electrophysiological techniques (electroantennography, GC/EAD), behavioural bioassays (species-specific designs), methods of enzyme purification and characterization (electrophoresis, chromatography, radioactivity assays including radiolabelled substrates).

Most significant achievements

1. Bumblebee pheromones

Summary*. In bumblebees, males attract conspecific females by marking places with a chemical signal secreted by their cephalic labial gland. The composition of the male marking/sex pheromones is strictly species-specific and it is used (together with other traits) as a tool for chemotaxonomical a phylogenetical considerations. A combined method for species delimitation in bumblebees has been developed.*

For species determination, the chemical approach is used in cases where morphological traits are not reliable or difficult to use. Our studies use molecular, chemical, and morphological criteria to establish the taxonomic status. This combined approach was used for two cuckoo-bumblebee taxa, *Bombus (Psithyrus) barbutellus* and *Bombus (Psithyrus) maxillosus*. In the past, these two sympatric taxa were discriminated by morphological criteria (wing darkness and hair length). We developed a combined molecular dataset from one nuclear gene (EF-1α) and one mitochondrial gene (COI) spanning 1623 bp and a chemical dataset of sexual marking pheromones (73 compounds). The data were subjected statistical analysis to test divergences between the two species. The resulting phylogenetic trees show no consistent divergence between the two taxa. Thus, we conclude that *B. maxillosus* is regarded as a synonym of *B. barbutellus* (published in *Syst. Entomol.* 2011).

An opposite case was observed in *Bombus terrestris*. This species is widely distributed all over Europe and exhibits a great colour variation. This trait led to defining a number of subspecies in *B. terrestris* in the past. We studied both molecular genetic data and chemical composition of the labial gland secretion in four subspecies (*B. terrestris terrestris, B. t. lusitanicus, B. t. sassaricus,* and *B. t. dalmatinus*). While the genetic data didn't lead to distinguishing of subspecies, the labial gland secretions (male sex pheromone) showed differences statistically significant for separate subspecies. Differences were proved by preferences of females for males of the same subspecies in behavioural experiments. Thus, some of the *B. terrestris* subspecies have a potential to be considered good species (published in *Chemoecology* 2011).

The combined approach was also used for comparison of mainland *versus* insular (Corsican) populations of two species, *B. terrestris* and *B. lucorum*. Our results show a morphological differentiation (color pattern) and genetic divergence of all Corsican endemics. Likewise, chemical analyses of male pheromone indicate that four Corsican endemics are differentiated from the mainland. Phylogenetic analyses, DNA sequence-based species automatic delimitation approach, and comparative chemical studies show that two Corsican taxa can be considered as endemic species while four other are considered as subspecies. We advocate the importance to conserve the endemic subspecies as well as endemic species of the Corsican bumblebee fauna (published in *PLoS One* 2013). The combined approach also helps to understand the evolution of bumblebee taxa in Europe and the influence of climatic oscillations on divergence of reproductive traits as shown on the example of *B. lapidarius* (published in *BMC Evol. Biol.* 2013).

2. Pheromone biosynthesis in male bumblebees

***Summary**. Experiments were done and different approaches were used to decipher steps in the biosynthetic pathways leading to the formations of pheromonal components in B. terrestris, B. lucorum, and B. lapidarius. Main aliphatic components are ethyl dodecanoate in B. terrestris, ethyl tetradec-9-enoate in B. lucorum, and hexadec-9-en-1-ol B. lapidarius. Two pathways may be involved in the biosynthesis - de novo formation in the gland vs transport and modification of lipidic precursors. ^{2}H, ^{13}C, and ^{14}C-labelled precursors applied both in vivo and in vitro indicate that both pathways are possible. In collaboration with the biochemical team, we identified and characterized desaturases and lipases that may participate in the pheromone biosynthesis, quantified the expression of genes encoding these enzymes in bumblebees of different age. The quantification was based on reference genes that were selected and validated for this purpose.*

In two model bumblebee species, *Bombus terrestris* and *B. lucorum,* we have done a series of experiments *in vitro.* Dissected labial glands were incubated with radioactive [1,2-^{14}C]acetate to see whether the pheromonal components can be formed *de novo*. The labelled substrate was incorporated into several types of compounds such as terpenic alcohols, fatty acids, esters, and hydrocarbons (Fig. 1). A comparative incubation of the [1,2- ^{14}C]acetate with fat bodies led to the formation of fatty acids (FAs), triacylglycerols (TAGs), and hydrocarbons. To support the results from the *in vitro* incubations, qPCR experiments with the gene coding the fatty acid synthase (FAS) were performed. This technique however required data normalization using reference genes. Therefore, we assessed the stability of 8 reference genes in the labial gland and fat body of the bumblebees *B. terrestris* and *B. lucorum* of different ages. To date, no reference genes have been identified for these species. Our data showed that arginine kinase (AK) and phospholipase A2 (PLA2) were the most stable genes in both tissues of *B. terrestris*. In *B. lucorum,* the most stable genes were the elongation factor 1a (EEF1A) and PLA2 (published in *Anal. Biochem.* 2010).

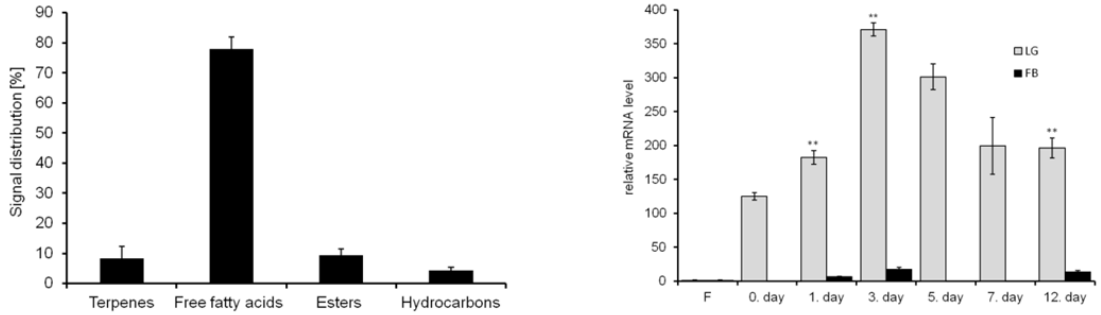


Fig. 1. Left: Distribution of the radioactivity in different types of compounds in the labial gland extract. Right: FAS transcript levels in different aged labial glands (LG) and fat bodies (FB) of *B. terrestris*.

Based on the reference genes, we have done qPCR experiments. The relative quantification of the FAS transcription level showed that the mRNA abundance of FAS gene depends on the *B. terrestris* male's age. A comparison of the FAS mRNA gene transcription level between fat bodies and labial glands of *B. terrestris* and *B. lucorum* showed that a high biosynthetic activity undergoes in the labial glands of both species. These experiments confirm that pheromonal components can be synthesised *de novo* in the labial gland. Another series of *in vitro* incubations was done with dissected labial glands and deuterium- labelled FAs of different chain length. Ethyl esters of labelled FAs were formed with a preference towards shorter chains (C12, C14). These results revealed that esterification of fatty acids proceeds in the labial gland (published in *ChemBioChem* 2013).

In the *in vivo* incubation experiments, *B. lucorum* and *B. lapidarius* were used as model species. Both species produce only aliphatic compounds in the labial gland secretions: ethyl esters of FAs (dominated by ethyl tetradec-9-enoate) in *B. lucorum* and unbranched primary alcohols (dominated by hexadec-9-enol) in *B. lapidarius*. A series of deuterium-labelled FAs of different chain length (C12-C18) were applied to the males' abdomens. The labelled acids were metabolised to pheromonal components, both main and minor ones. Decarboxylation, reduction, esterification, and desaturation of applied labelled FAs was observed (Fig. 2). The carbon chain of precursors underwent elongation but no chain shortening. Besides, the labelled acids were incorporated in TAGs in the fat body. These results led us to the hypothesis that the pheromone may be formed by a transformation of lipidic precursor. However, a greater portion of the applied precursors was metabolised *in vivo* to pheromonal components than to lipids. Considering results from *in vitro* and *in vivo* experiments, they have refuted neither the biosynthesis of pheromones from common lipids nor *de novo* synthesis of pheromones in the labial gland as an exclusive pathway. It is likely that both pathways contribute to the pheromone formation depending on the condition, physiological state of the insect, or availability of the substrates (published in *ChemPlusChem* 2015).

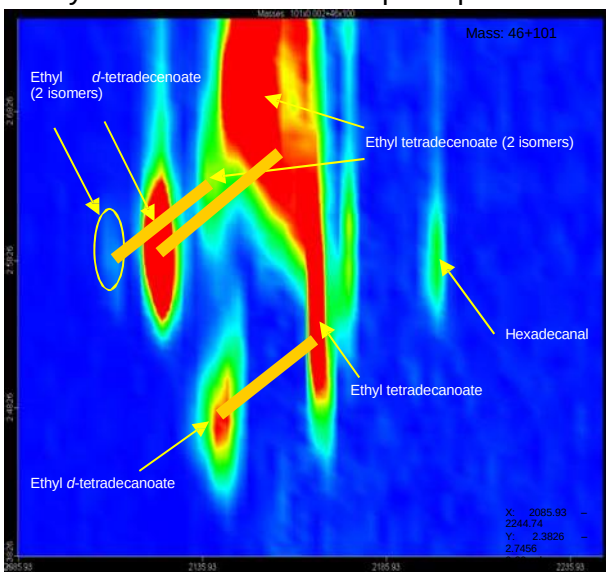


Fig. 2. Extended part of a GCxGC recording of *B. lucorum* labial gland extract (*in vivo* application of *d*-myristic acid).

Since the incubation experiments proved the activity of several enzymes in the labial gland, we focused on detailed studies of two types, desaturases and lipases. We cloned the gene encoding the Δ9 desaturase from cDNA prepared from the total RNA of the pheromone gland and fat bodies of *B. lucorum* male. The functional expression of BlucNPVE desaturase in *Saccharomyces cerevisiae* and the GC-MS analyses revealed its preference for palmitic acid (C18). This suggests that recombinant Δ9 desaturase is involved in the desaturation of metabolic FAs stored in triacylglyceroles (TAGs) because oleic acid is the most abundant fatty acid bound in TAGs in *B. lucorum*. However, its role in the desaturation of myristic acid in the labial gland cannot be excluded (published in *Insect Biochem. Mol. Biol.* 2013).

Another enzyme participating hypothetically in the biosynthesis is a lipase. If the "lipidic precursor hypothesis" was true, there should be a selective lipase associated with the labial gland breaking the diacylglycerols transported by the haemolymph. Therefore, we isolated lipases from the labial gland and fat body and studied their properties. We studied the lipase activities in tissues of *B. terrestris* males (freshly emerged to 16-day-old specimens). Lipolytic enzymes were detected in crude tissue extracts using *p*-nitrophenyl laurate. In the labial gland, we found a high lipase activity in 3-day-old males. Lipase activity in the fat bodies was highly variable. Optimum pH of the lipolytic enzymes was 8.3, the temperature optimum was found at 50 °C. According to SDS-PAGE and zymography, the molecular weight of lipolytic enzymes from labial glands and fat bodies was 67 kDa. The methodology used for studies of bumblebee lipases and their properties was based on our previous knowledge obtained with lipases from other organisms. A unique neutral lipase (BT-1) was isolated from the labial gland of *B. terrestris* males, purified and characterised (published in *PLoS One* 2013).

3. Bumblebee lipids and lipid analyses, lipidomics

***Summary**. Following the above mentioned hypothesis on formation of pheromones from lipidic precursors, similarities in the structure of pheromonal components and fatty acids in*

triacylglycerols were studied in 11 species. In some of them, striking similarities in the chain length and double bond positions were found. Changes of the lipid composition during lifespan were followed. The lipidomic studies resulted also in several methodological papers (development of analytical method, comparison of different MS techniques, fragmentation studies, computer-based evaluation of analytical data).

Another indication of lipids as precursors of male pheromones brought our study in species *B. ruderatus, B. bohemicus,* and *B. campestris*. Striking structural similarities between fatty acids (FAs) bound in TAGs and components of the male marking pheromone were found in these species (and earlier in *B. pratorum* and other bumblebee species). Higher proportion of specific FAs in pool lipids speak for the hypothesis that FAs may serve as precursors in the pheromone biosynthesis (published in *Molecules* 2014).

The expected pheromone/lipid similarities were found in several studied species, but not all of them met our expectations. Based on our knowledge on the dynamics of the pheromone production in *B. terrestris* and *B. lucorum*, the age-dependent changes (males 0–30 days old) in the TAG composition in fat bodies were studied. The total amount of TAGs in *B. lucorum* was about 2.7 times higher than that in *B. terrestris*. The highest content of TAGs was found in males 1-3 days old. In both species, the qualitative composition of FAs in TAGs was similar, but the mean relative abundance differed (Fig. 3). A statistical evaluation of the dynamic changes in the TAG composition revealed that in *B. terrestris* different age classes were well separated while in *B. lucorum* the TAG did not change substantially during the male's life. The TAG analyses provide more precise information on the differences between the classes studied than the FA composition alone (used however in most papers on insect lipids). The difference in the dynamics of lipid production between the two species agreed with our previous results on the production of sex pheromone (published in *Lipids* 2011).

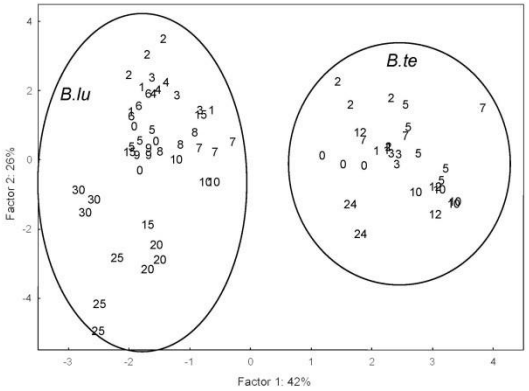


Fig. 3. Separation of species based on FA composition. Males of different age.

Beside TAGs, the wax esters (WEs) were studied with the aim to find the best method for analysis and identification of WEs in natural samples. WEs are often found in insect glands and their identification is sometimes not easy due to occurrence of isomers and coelutions. An HPLC/APCI-MS method was developed for the WEs analysis (published in *J. Chromatogr. A* 2010). A method of localization of double bonds in carbon chains was developed utilizing fragmentation of acetonitrile-related adducts in APCI-MS (published in *Anal. Chem*. 2011). The interpretation of EI mass spectra of straight-chain and methyl-branched saturated and unsaturated wax esters was studied in a series of 154 standards (published in *J. Lipid. Res*. 2012). The most important fragments indicative of the structure of the acid and alcohol chains were identified and summarized for WEs with various number of double bonds in the chains (Fig. 4). Most WEs provide acylium ions allowing structural characterization of the acid part, whereas the alcohol part gives corresponding alkyl radical cations.

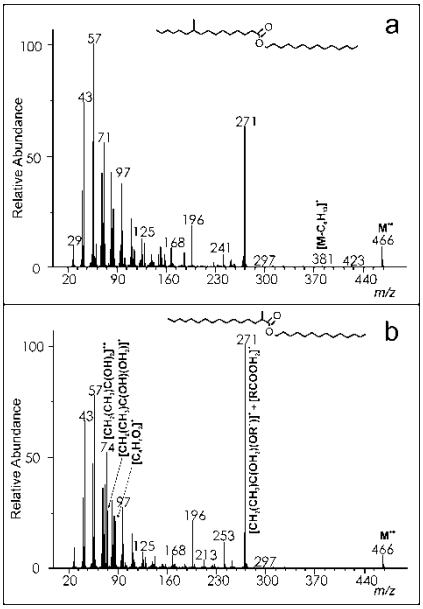


Fig. 4a,b. Mass spectra of two WE positional isomers.

The ion abundances change with the length and unsaturation of the aliphatic chains. We created a large

mass spectral database that can be used as a reference for the analysis of the GC/EI-MS data.

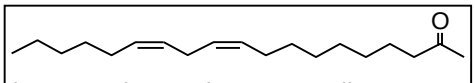
Our expertise in lipidomics was also used for analyses of lipids in other natural sources. Thus, mobilisation of lipids using adipokinetic hormone was studied in bugs (published in *Comp. Biochem. Physiol. B* 2010, *Eur. J. Entomol.* 2010, *J. Insect. Physiol*. 2012 and 2014). Besides, our methods were used for analyses of components of human cerumen that was a complex mixture of hydrocarbons, squalene, wax esters and cholesterol esters, TAGs, free FAs, free fatty alcohols, monoacylglycerols, free cholesterol, free sterols, and free hydroxy acids (published in *Lipids* 2011). Besides, a cost effective, robust and reliable coupled separation techniques for the identification and quantification of phospholipids in complex biological matrices has been developed (published in *J. Sep. Sci.* 2014).

4. Chemistry of pheromones and defensive compounds in termites

***Summary**. A set of studies on chemical ecology, anatomy of exocrine organs and caste system was performed on poorly known species from phylogenetically important lineages (genera Glossotermes,Psammotermes and Termitogeton), including the identification of several pheromones and defensive compounds. Among them we identified a new trail-following pheromone with unexpected structure in comparison with C12 unsaturated alcohols occurring in all other advanced termite families. Among other important projects belong the first description of fertility related proteins in kings and queens in the genus Prorhinotermes and a unique defensive strategy in workers of Neocapritermes taracua. Based on our experience in the field of termite chemical defence, we were invited to publish a review Chemical warfare in termites for J. Insect Physiol.* (2010).

Chemical ecology of termites is studied using a complex set of methods ranging from a detailed description of exocrine organs, through chemical identification of their products (pheromones and defensive chemicals) to the evaluation of behavioural and physiological role of particular compounds by means by bioassays and electrophysiology.

In the frame of research project entitled 'Biology, chemical ecology, and phylogeny of critical termite genera from families Rhinotermitidae and Serritermitidae' we studied the structure and function of defensive glands in soldiers (published in *Biol. J. Linn. Soc*. 2010) and identified a structurally new trail-following pheromone, (10*Z*,13*Z*)-nonadeca-10,13-dien-2-one, in serritermitid *Glossotermes oculatus* (published in *Chem. Senses* 2012).



(10*Z* ,13*Z*)-nonadeca-10,13-dien-2-one

Within the same project we described the developmental pathways (published in *PloS One* 2012) and identified trail-following and sex pheromones (published in *J. Chem. Ecol.* 2011) as well as the composition of the defensive secretion in rhinotermitid *Psammotermes hybostoma* (published in *J. Chem. Ecol.* 2012). Furthermore, the chemical composition of the complex multi-component alarm pheromone in rhinotermitid *Termitogeton planus* was identified (published in *J. Chem. Ecol.* 2014).

Our previous studies in the genus *Prorhinotermes* were supplemented by other papers published in *J. Insect Physiol*. 2010 and *ChemBioChem* 2014. It is worth noticing that we succeeded in discovering a set of compounds, characteristic to each sex of *Prorhinotermes* reproductives, kings and queens, and correlated with their fertility. These compounds are peptides and proteins; partial sequencing of some of these compounds suggests that they might belong to odorant binding proteins, involved in the transmission and release of volatile pheromone components (published in *Proc. R. Soc. B* 2010).

Beside the main projects, we were studying defensive adaptations in another few genera of tropical termites. Workers of *Neocapritermes taracua* possess a pair of dark blue, elongated dorsal spots at the thorax–abdomen junction. These "blue workers" burst when seized and emit this way a drop of fluid, which in a few seconds becomes sticky as the blue colour fades out. The blue spots are a pair of hard and fragile clove-like structures ("blue crystals") enclosed within pouches formed by posterior outgrowths of the metanotum over the first abdominal segment. The blue defensive substance consists predominantly of a 76

kDa protein ("blue crystal protein"), which contains 9 ± 2 ng of copper per crystal. Based on a partial sequencing, we tentatively assigned the protein to the hemocyanin/phenoloxidase family. Such a defence strategy and chemistry is unique within termites, but also within insects in general. The main findings on *Neocapritermes* defensive mechanism were published in *Science* 2012.

We also collaborate on identification of two new pheromones in basal termite *Hodotermopsis sjoestedti* (published in *J. Insect. Physiol.* 2011), we used CHs profiles together with DNA barcoding for termite taxonomy (published in *Mol. Phylogent. Evol.* 2013) and studied mutual use of trail-following pheromones by a termite host and its inquiline (published in *PLoS One* 2014).

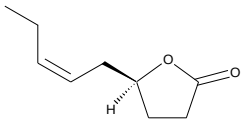
5. Chemical ecology of other insects

***Summary**. Pheromones, defence substances, and food attractants were studied in several insect species belonging to families Lepidoptera (bumble bee wax moth Aphomia sociella, sugarcane-borer Diatraea flavipennella), Coleoptera (bark beetles), Heteroptera (sting bugs), and Diptera (fruit flies). The most interesting results are discussed below in detail.*

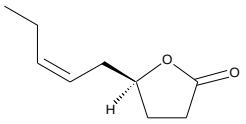
The bumblebee wax moth, Aphomia sociella (Lepidoptera: Pyralidae: Galleriinae) is a parasitic moth of causing great damage on bumblebee colonies. Male wing gland secretes male sex pheromone consisting of hexan-1-ol, 2-phenylethanol, [(*R*),(*Z*)]-nona-2,6-dien-4-olide (**1**), [(*S*),(*Z*)]-nona-6-en-4-olide (**2**), mellein (**3**), phytone (6,10,14-trimethylpentadecan-2-one, **4**), and a mixture of C_{18} fatty acids.

Beside the pheromone, males produce ultrasonic signals to excite approaching females. Observations of *A. sociella* mating behaviour and recordings of male acoustic signals confirmed that males initiate the mating process. During calling behaviour (stationary wing fanning and pheromone release), males disperse pheromone from their wing glands. When female approaches, males cease calling and begin to produce ultrasonic songs as part of the courtship behaviour. Replaying of recorded courting songs to virgin females proved that male ultrasonic signals stimulate females to accept mating. Greenhouse experiments with isolated pheromone glands confirmed that the male sex pheromone mediates long-range female attraction (published in *PLoS One* 2011).

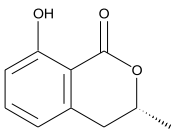
The recognition of a female and the start of ultrasonic signalling is based on the female courtship pheromone. Female extracts contained three antennally active compounds: hexan-1-ol, phytone (**4**), and 6,10,14-trimethylpentadecan-2-ol (**5**). In laboratory bioassays, alcohol **5** and, at higher doses, ketone **4** initiated male courtship behaviour associated with ultrasonic production. Hexan-1-ol and ketone **4** enhanced the activity of alcohol **5**. Thus, the moths use a complex multimodal mating strategy that is unique within the subfamily Galleriinae. This fascinating “love” story is published in *J. Nat. Prod.* 2009*, PLoS One* 2011*,* and *J. Chem. Ecol* 2012*.*



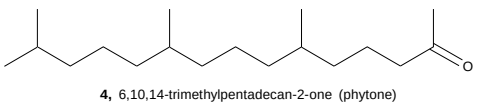
1, (4*R*)-(2*Z*,6*Z*)-nona-2,6-dien-4-olide



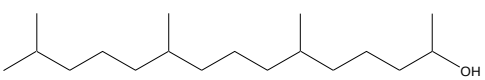
2, (4*S*)-(6*Z*)-non-6-en-4-olide



3, (*R*)-(–)-mellein



4, 6,10,14-trimethylpentadecan-2-one (phytone)



5, 6,10,14-trimethylpentadecan-2-ol

Defensive secretion of European stink-bug, Graphosoma lineatum
contains 57 compounds consisting mainly of aliphatic aldehydes. (*E*)-4-oxohex-2-enal and (*E*)-dec-2-enal were found to be the major components in the adult bug secretions followed by lower amounts of n-alkenal (C5–C12), n-alkenyl acetate (C5–C11), n-alkane (C11–C17) homologs, dienals and other compounds. The identification of (*E*)-4-oxohex-2-enal is of particular interest, since it was previously identified incorrectly. Our results demonstrate a power of our approach (GC×GC/TOF-MS) in yielding artefact-free secretion profiles (published in *J. Chromatogr. B* 2012).

The sugarcane-borer Diatraea flavipennella
is a moth that causes extensive damage to plants and, consequently, significant economic losses. Two antennally active compounds, (Z)- hexadec-9-enal (Z9-16:Ald) and (Z)-hexadec-11-enal (Z11-16:Ald) were identified in the pheromonal gland extract in the proportions 25:75. Behavioural bioassays demonstrated that a mixture of Z9–16:Ald and Z11–16:Ald in the ratio 25:75 induced a response in *D. flavipennella* males similar to living virgin females. Thus, Z9–16:Ald and Z11–16:Ald are considered to form the female sex pheromone of *D. flavipennella* (published in *J. Appl. Entomol.* 2012).

Comparative analysis of chemical communication in fruit flies
Fruit flies are among the worst plant pests in agriculture, causing major losses in fruit and vegetable production, and are major targets of insecticide use in nearly all tropical, subtropical and some temperate countries worldwide. They are also of major economic importance because they result in the establishment of quarantines and the disruption of international agricultural trade. The Sterile Insect Technique is increasingly applied operationally against the Mediterranean fruit fly; however, there are some important constraints to its use against other major fruit fly pests. To determine, whether chemical communication differs between different populations of target species, we analysed the composition of male sex pheromones and cuticular hydrocarbons (*Anastrepha fraterculus*, *Ceratitis* FAR complex) using GCxGC/TOFMS and compared them by means of multivariate statistical analyses. Data obtained so far show significant inter-population differences in chemical signalling suggesting cryptic speciation. Our data are currently compared with morphogenetic data collected by Brazilian co-workers and will be used to resolve the taxonomic relationships and to estimate the gene flow within these complexes (published in *Bull. Entomol. Res.* 2014*, Florida Entomol,* 2013, *J. Chem. Ecol.* 2012, and *J. Agric. Food Chem.* 2012).

Research Report of the team in the period 2010–2014

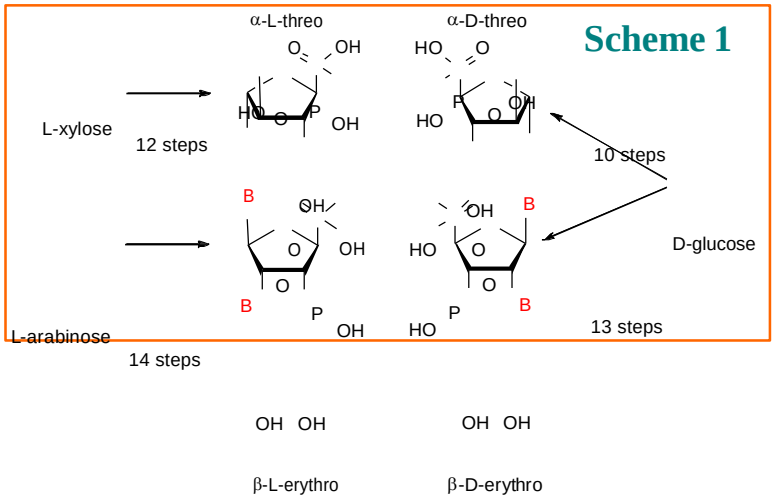
Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Ivan Rosenberg - Phosphonate Nucleotides and Oligonucleotides

Scientific program of the Group of Nucleotides and Oligonucleotides has been directed toward basic research in the area of synthesis of novel **(A)** ***Nucleoside Phosphonic Acids*** *(NPAs)* as potential antimetabolites and **(B)** ***Phosphonate Oligonucleotide Analogs*** *(POAs)* as antisense (or other biologically relevant) compounds capable of interfering with gene expression.

(A) Nucleoside Phosphonic Acids. Specifically, the interest has been focused on the synthesis of structurally diverse NPAs as *small molecules* based on nucleosides containing both classic (furanose) and modified (pyrrolidine, prolinol, piperidine, and some other heterocyclic) rings.

A-1.*Tetrofuranose nucleoside phosphonates.*[i] In frame of our systematic research of nucleoside phosphonic acids, new isoelectronic, non-isosteric phosphonate analogues of pyrimidine nucleoside 5′-phosphates possessing the phosphorus moiety directly attached to



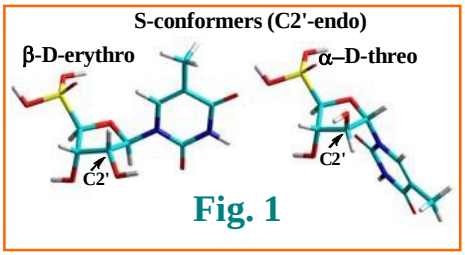
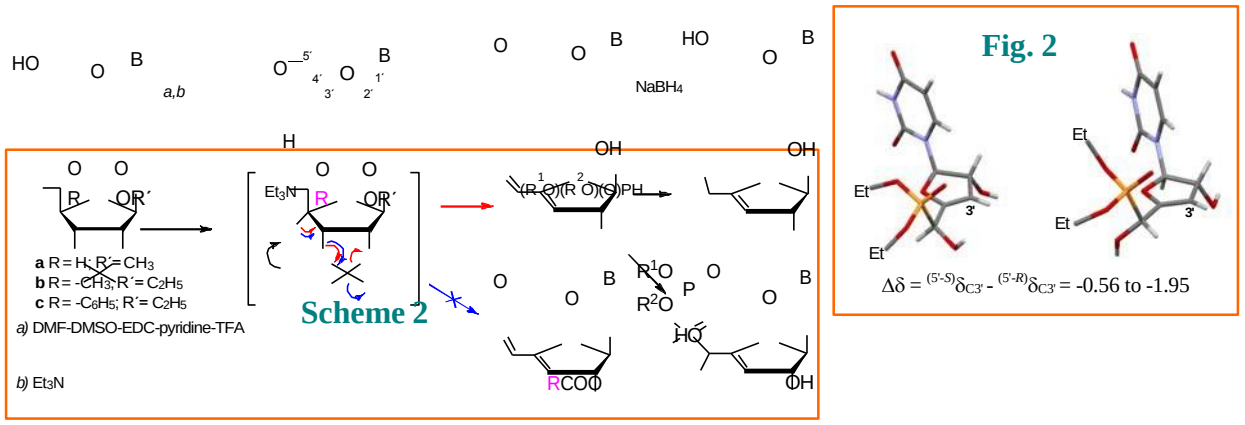


Fig. 1

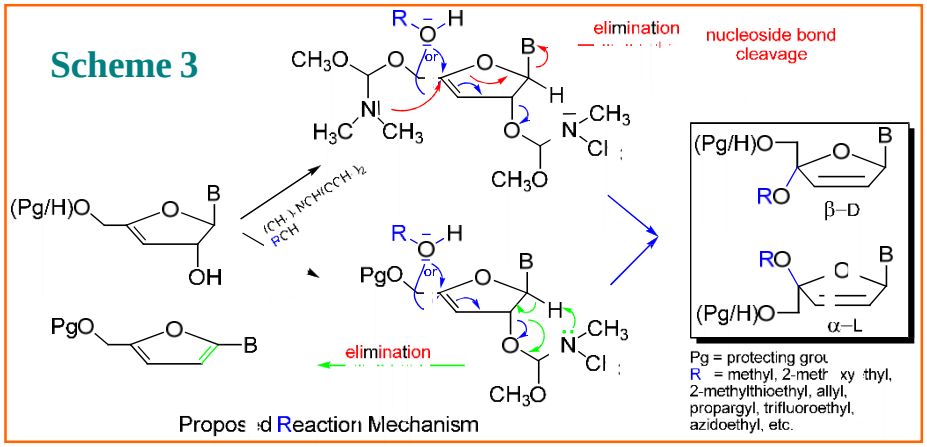
the sugar ring in the C4′ position were synthesized by multistep syntheses starting from L-pentoses and D-glucose (**Scheme 1**). The key synthons, *erythro*- and *threo*-configured tetrofuranosyl phosphonates in the D- and L- series were synthesised from the corresponding pentofuranosyl 5-hydroxyphosphonates via oxidative cleavage of the C1-C2 bond. These synthons provided the desired compounds after nucleosidation reaction with silylated pyrimidine nucleobases. Potential of these compounds is yet to be fully biologically evaluated. Nonetheless, one compound of this series inhibited human thymidine phosphorylase at low micromolar range ($K_i \approx 4$ μM). The results of NMR conformational studies showed that all calculated conformers have a maximum pucker in the range typical for nucleosides. In all four compounds, the S-type conformer (C2'-*endo*) is highly preferred (>80%) (**Fig. 1**). **Contribution of Group: design and synthesis of all compounds; thymidine phoshorylase assay.**

A-2. *3′,4′-Didehydronucleosides and corresponding 5′-hydroxyphosphonates.*[ii,iii] Within our study on suitable protection of the *cis* diol of nucleoside 5′-hydroxyphosphonates, we have discovered that the 2′,3′-*O*-orthoester derivatives of ribonucleoside-5′-aldehydes undergo,

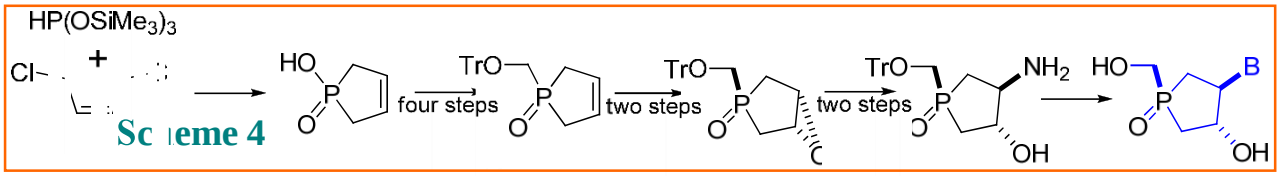


under mild conditions, a base-catalyzed elimination of the orthoester moiety resulting in the formation of 3′-deoxy-3′,4′-didehydronucleoside-5′-aldehydes. These compounds are otherwise very difficulty obtainable (**Scheme 2**). The 5′-aldehydo group can be smoothly and rapidly reduced using $NaBH_4$ in DMF with a limited amount of methanol to afford 3′-deoxy-3′,4′-didehydronucleosides. Moreover, the 3′-deoxy-3′,4′-didehydronucleoside-5′-aldehydes themselves represent useful synthons offering a wide range of chemical transformations. Thus, the nucleophilic addition of various dialkyl phosphites to 3′-deoxy-3′,4′-didehydronucleoside-5′-aldehydes provided epimeric hydroxyphosphonates in a quantitative yield (**Scheme 2**). Whereas direct NMR configuration assignment for the C5′ atom bearing the phosphoryl and hydroxy groups using the *J*(P,H4′) and *J*(H5′,H4′) coupling constants is impossible due to the absence of the H4′ atom, successful separation, crystallization and X-ray crystallographic analysis of a pair of epimeric 5′-*C*-phosphonates, followed by correlation with a series of NMR parameters, led to configuration assignment for individual epimers in the mixtures based on the difference of $\delta_{C3'}$ chemical shifts (**Fig. 2**). **Contribution of Group: design and synthesis of all compounds; crystallization of epimers.**

A-3. *4'-Alkoxy-2',3'-dideoxydidehydronucleosides.*[iv] We found that the exposure of above mentioned 3´-deoxy-3´,4´-didehydronucleosides (***A-2***) to DMF-dimethylacetal and an alcohol at room temperature resulted in a Ferrier-type allylic rearrangement upon formation of the C4´-epimeric 4´-alkoxy 2´,3´-didehydro-2´,3´-dideoxynucleo-sides (**Scheme 3**). Although the formation of these compounds was accompanied by elimination, which is a nucleoside bond cleavage and/or formation of a furan nucleoside, the β-D and α-L epimers were still obtained in pure form and yield by chromatography. In contrast to 2',3'-didehydrodideoxy-nucleosides (d4A, d4G), the 4'-alkoxy derivatives are very stable under acidic conditions. We expect that this simple approach towards a wide range of 4´-alkoxy-2´,3´-didehydro-2´,3´-dideoxynucleosides will become a useful tool in the search for biologically active compounds. Studies on antiviral and anticancer properties of the prepared nucleoside analogues are in progress. **Contribution of Group: design and synthesis of all compounds.**



A-4. *Novel phosphanucleosides.*[v] We have developed a multistep synthetic route leading to a new class of nucleoside mimics, the phosphanucleosides containing phospholane ring (**Scheme 4**). In all cases we employed both purine and pyrimidine nucleobase construction on the aminophospholane derivative. The prepared phospholane nucleosides were neither substrates nor inhibitors of purine and thymidine phosphorylases *in vitro*. Thorough studies on the antiviral, antibacterial, and antitumor activities of phosphonucleosides are in progress. This new trend in the chemistry of sugar-modified nucleosides (C-nucleosides) is expected to provide the number of valuable results. **Contribution of Group: design and synthesis of all compounds.**



A-5. *Synthesis and conformational behaviour of azanucleoside phosphonates*[vi,vii,viii]

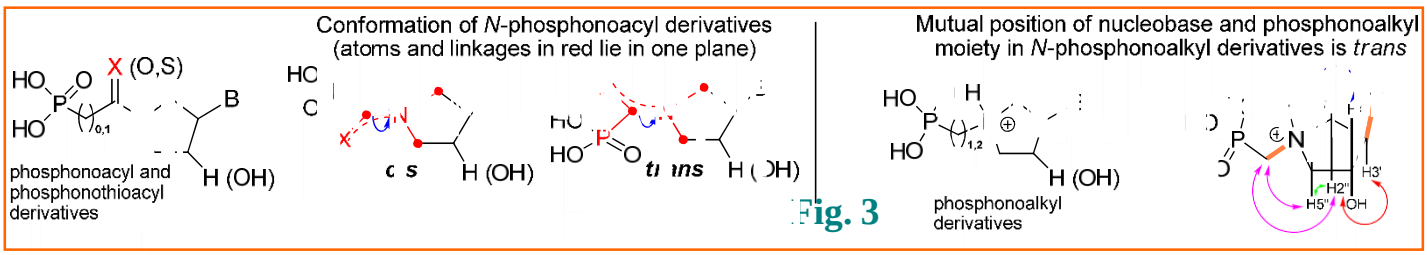


Fig. 3

Conformational preferences of the pyrrolidine ring in nucleotide analogs were investigated by means of NMR and molecular modeling. The effect of the relative configuration of hydroxy and nucleobase substituents as well as the effect of the alkylation or acylation of the pyrrolidine nitrogen atom on the conformation of the pyrrolidine ring were studied. The results of a conformational analysis showed that the alkylation/acylation can be effectively used for tuning the pyrrolidine conformation over the whole pseudorotation cycle. Synthesis and conformational analysis of pyrrolidine and piperidine phosphonate nucleotides has been studied. Reactivity of recently developed reagent for introduction of pyrimidine nucleobases on pyrrolidine and piperidine scaffolds was studied by means of NMR and *in silic*o approaches.[ix] **Contribution of Group: design and synthesis of all compounds.**

A-6. *Inhibition of nucleoside phosphorylases.* Thymidine (TP) and purine nucleoside (PNP) phosphorylases are potential targets in cancer therapy. TP is involved in angiogenesis and thus, the inhibition of TP may result in reduction of tumor growth and metastasis and/or in considerable decrease of the TP-catalyzed phosphorolysis of nucleosides used in cancer therapy. PNP is one of the most important purine salvage pathway enzymes. In T-lymphocytes, the inhibition of PNP leads to the accumulation of dGTP followed by allosteric inhibition of NDP-reductase and finally to apoptosis. We have reasoned that among NPAs, bisubstrate-like inhibitors of nucleoside phosphorylases could be found. Since the NPAs possess the nucleoside and phosphorus moieties mimicking both natural substrates – nucleoside and inorganic phosphate – these compounds could occupy simultaneously both the phosphate and nucleoside binding sites of the enzymes.

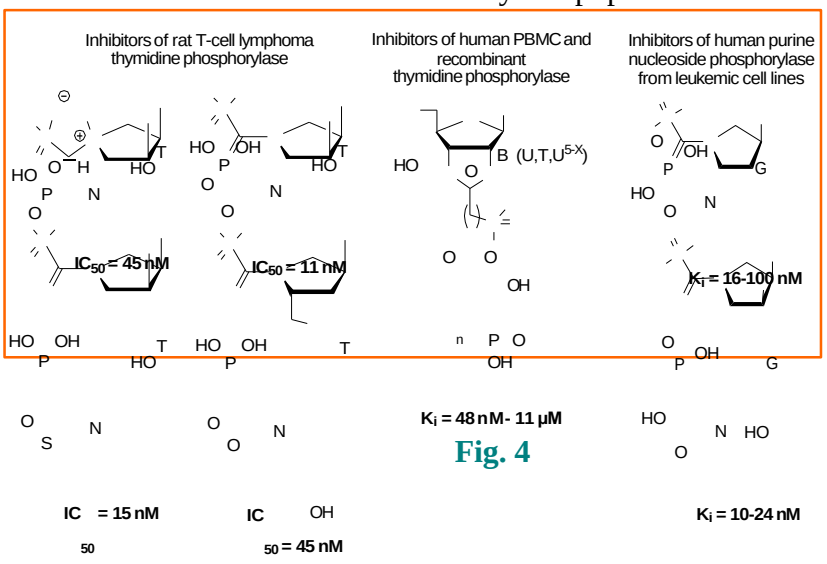


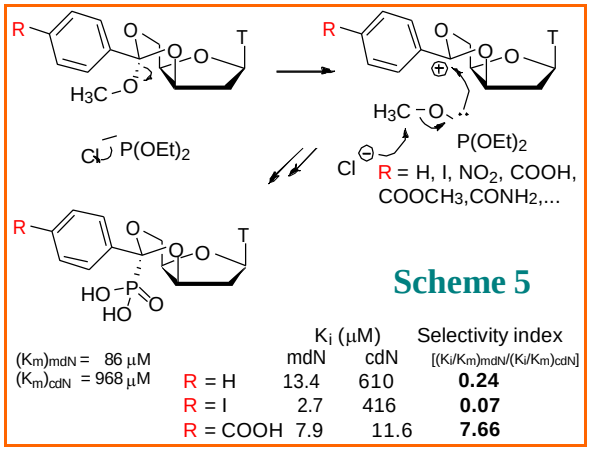
Fig. 4

Among the large pool of NPAs prepared and tested in our laboratories, we found four potent inhibitors of rat T-cell lymphoma thymidine phosphorylase (IC_{50} = 11 – 45 nM) (**Fig. 4**).[x] Interestingly, neither TP from healthy rat liver nor human TP were inhibited by these phosphonates. Recently, we published a thorough study on inhibition of the human TP with nucleoside phosphonic acids with the aim to find potent bisubstrate-like inhibitors.[xi] Out of a large set of tested compounds, several nucleoside phosphonic acids were found as inhibitors with the K_i values (related to both substrates) lying in a range of 48 nM–11 μM (**Fig. 4**). The inhibition potency of the studied compounds strongly depended on the degree of conformational flexibility of the phosphonate moiety, the stereochemical arrangement of the sugar-phosphonate part, and on the substituent in the position 5 of the pyrimidine nucleobase. At present, we are working on co-crystalization of human recombinant TP with our best inhibitor to facilitate the search for more potent compounds.

A large set of guanine-based NPAs was tested as potential inhibitors of human PNP isolated from peripheral blood mononuclear cells (PBMC) and cell lines of myeloid and lymphoid origin. We have selected two potent inhibitors (**Fig. 4**).[xii] Prodrug forms of these inhibitors were evaluated in the *in vitro* experiments, however, without any remarkable effects (even a common application with 2'-deoxyguanosine did not increase cellular toxicity). Also in this case we attempt to prepare a mixed crystal of PNP with inhibitor to facilitate the search for more potent compounds. **Contribution of Group: design and synthesis of all compounds; all experiments with human thymidine phosphorylase and purine nucleoside phosphorylase.**

A-7. *Inhibitors of 6-oxopurinephosphoribosyltransferases (HGXPPRT) as potential antimalarial and anti Mycobacterium tuberculosis drugs.* [xiii] Based on pyrrolidine scaffold a series of monophosphonate and later bis-phosphonate nucleotide inhibitors of HGXPRT from *P. falciparum* and *M. tuberculosis* has been discovered. Prodrugs of selected inhibitors have shown promising activity in cell based assays. Currently preparation of patent application is underway. **Contribution of Group: synthesis of all pyrrolidine nucleoside phosphonic acids.**

A-8. *Inhibition of pyrimidine specific 5'(3')-nucleotidases.*[xiv,xv] The biological function of nucleotidases (seven different types + m7GMP-specific) consists in regulation of cellular nucleotide pool *via* dephosphorylation of nucleoside 5'(3')-monophosphates. Also the 5'-phosphoesters of nucleoside analogs used as anti-viral and anti-cancer drugs may be dephosphorylated, and thus, their pharmacological efficacy may be considerably affected. The nucleoside analog resistance seems to be linked to a high expression of cytosolic 5′-nucleotidases. Therefore, the development of potent 5'-nucleotidase inhibitors may reverse drug resistance and increase the efficacy of clinically used nucleoside analogs. The search for potent and selective inhibitors of mitochondrial (mdN) and/or cytosolic (cdN) pyrimidine specific 5'-nucleotidases revealed a new group of compounds, the 3',5'- phosphonobenzylidene-*xylo*-thymidines substituted in *para* position of the phenyl ring by various substituents (**Scheme 5**).[xiv] The screening showed significant improvements of inhibitory activity of *para* substituted compounds (R = I, COOH) compared to the leading structure (R = H).



Scheme 5

	K_i (μM) mdN	K_i (μM) cdN	Selectivity index $[(K_i/K_m)_{mdN}/(K_i/K_m)_{cdN}]$
$(K_m)_{mdN}$ = 86 μM, $(K_m)_{cdN}$ = 968 μM			
R = H	13.4	610	**0.24**
R = I	2.7	416	**0.07**
R = COOH	7.9	11.6	**7.66**

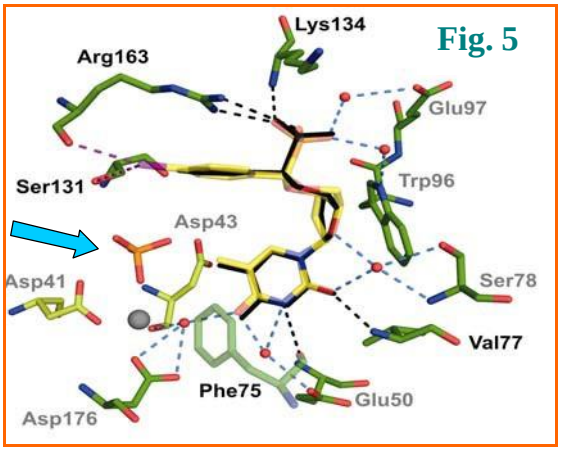


Fig. 5

This project was performed in close collaboration with the Department of Structural Biology of our Institute. The crystal structure of mdN in complex with the most potent compound so far known (**Scheme 5**, R = I; **Fig. 5,** structure in yellow) proved that the introduced modification in the *para* position did not change the inhibitor binding mode in comparison with the lead structure (**Scheme 5**, R = H; **Fig. 5,** structure in black). Moreover the iodo substituent forms iodo bond with carbonyl of Arg163. The obtained structural data can be used for further design of compounds with increased selectivity (**Scheme 5,** compare R = I and R = COOH) and/or improved inhibitory properties in relation to the specific enzyme isoform. Furthermore, the proximity of inorganic phosphate binding site (**Fig. 5**, blue arrow) and 5-methyl group of thymine provided an opportunity to design new mdN/cdN inhibitors based on the 5-phosphonoalkyl derivatives which should interact much strongly in the mdN/cdN catalytic site than the thymine ones. This assumption was proven by preliminary results which showed the decrease of K_i value by about one order of magnitude. The research in this area, while undoubtedly fruitful, is far from finished. The combination of substituents

in the 5-position of uracil moiety and on the phenyl ring will be further studied, as well as the search for inhibitors of purine specific nucleotidases. **Contribution of Group: design and synthesis of all nucleoside phosphonic acid.**

A-9. *Lipophosphonoxins.*[xvi,xvii] Novel compounds termed *lipophosphonoxins* (LLPO) were prepared using a simple and efficient synthetic approach. The general structure of these compounds consists of four constituent parts: (i) nucleoside module, (ii) iminosugar module, (iii) hydrophobic module (lipophilic alkyl chain), and (iv) phosphonate module (**Fig. 6**). LLPO displayed significant antibacterial properties against a panel of Gram-positive species, including multiresistant strains. The minimum inhibitory concentration (MIC) values of the best inhibitors were in the 1−12 μg/mL range, while their cytotoxic concentrations against human cell lines were significantly above this range. Mechanism of action has been elucidated (manuscript in preparation). A second generation of LPPO with extended antibacterial activities has been developed. **Contribution of Group: design and synthesis of all LLPO.**

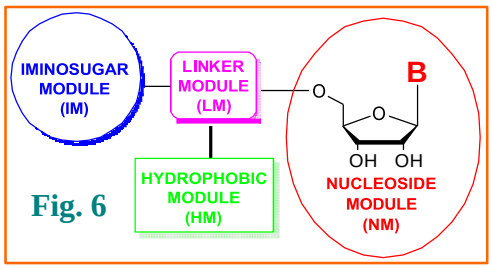


Fig. 6

Our investigation in the area of the chemistry of NPAs has provided, over last five years, an impressive number of novel structurally diverse compounds. Among them, several potent inhibitors of selected enzymes of nucleoside and nucleotide metabolism were found. These results prompted us to devote, in the near future, more attention to the biochemical and biological evaluation of our compounds within broadly-based collaborations. In addition, the knowledge obtained with the synthesis of NPAs is unique as such and will be fully utilized in new projects directed to a new generation of nucleotide analogs.

(B) Phosphonate Oligonucleotide Analogs. In the area of modified oligonucleotides we have elaborated the synthesis of several types of nucleoside phosphonate monomers in ribo and deoxyribo series suitable for the incorporation into oligonucleotide chains using the advanced phophotriester method. The incorporation of phosphonate units gave rise to several types of isopolar isosteric and nonisosteric phosphonate-based internucleotide linkages, such as C3'-O-P-CH_2-O-C4", C3'-O-CH_2-P-O-C5", C3'-O-P-CH_2-O-C5", C3'-O-P-(HO)C5", and C2',3'-di-O-CH(R)-P-O-C5". The evaluation of the several types of modified oligodeoxynucleotides as compounds aimed to influence gene expression is underway. The results have not yet been published due to the pending patent.

B-1. *Phosphonate 2',5'-linked oligoadenylates – potential agonists/antagonists of human ribonuclease L (RNase L).*[xviii] RNase L plays a significant role in interferon-induced antiviral defence mechanism of cells. The enzyme is activated by unique oligonucleotides – a short 5'-phosphorylated 2',5'-linked oligoadenylates. To determine the influence of internucleotide linkage conformation and sugar ring configuration, and the role of 5′-terminal phosphate, on the activation of human RNase L, a series of the 5′-*O*-methylphosphonate-modified tetramers were synthesized from appropriate monomeric units and evaluated for their ability to activate human RNase L (**Fig. 7**). Tetramers pApApAp_cX modified by *ribo, arabino* or *xylo* 5′-phosphonate unit p_cX, activated RNase L with efficiency comparable to that of natural activator. Moreover, incorporation of phosphonate linkages ensured the stability against cleavage by nucleases. The substitution of 5′-terminal phosphate for 5′- terminal phosphonate in tetramer p_cXpApApA afforded tetramers with excellent activation

efficiency and with complete stability against the cleavage by phosphomonoesterases. The study on tetramers composed of the adenosine-2'-phosphonate units showed inefficiency of these regioisomeric tetramers to activate RNase L. The selected tetramers containing the 5'-phosphonate units will be tested, in the form of prodrugs, for their cytostatic and antiviral properties. **Contribution of Group: design and synthesis of all compounds and RNase L FRET assays.**

B-3. *Phosphodiester 4'-alkoxyoligodeoxy-nucleotides.*[xix] We found a new very promising structural alternative to the existing analogs of the natural phosphodiester oligonucleotides. Modifications involving the presence of either a 4´-methoxy or 4´-(2-methoxyethoxy) group in oligothymidylates, represented by the ($_{4'\text{-MeO}}$T)$_{14}$T or ($_{4'\text{-MeOEtO}}$T)$_{14}$T oligonucleotides, respectively, were found to display markedly better hybridization properties with both the dA_{15} and rA_{15} counterparts as compared to unmodified dT_{15} (**Fig. 9**).

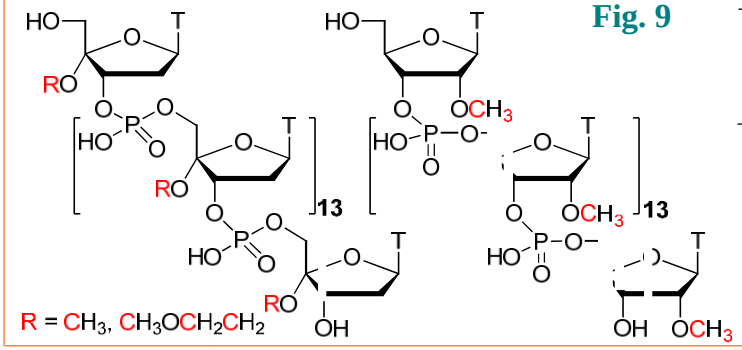


Fig. 9

Modified dT_{15}	Melting point T_m [°C] Complex type[b] dA_{15} Na^+	dA_{15} Mg^{2+}	rA_{15} Na^+	rA_{15} Mg^{2+}
dT_{15}	38 duplex	48 duplex	36 duplex	36 triplex
$(T_{2'\text{-OMe}})_{14}T$	35 --	54 duplex	48 duplex	55 duplex
$(_{4'\text{-MeOEtO}}T)_{14}T$	20 --	49 duplex	48 duplex	57 triplex
$(_{4'\text{-MeO}}T)_{14}T$	39 --	55 triplex	51 duplex	55 duplex

The novel oligomers containing $_{4'\text{-MeOEtO}}$T and $_{4'\text{-MeO}}$T units matched the T_m enhancement of the corresponding 2´-*O*-methyl oligomer $(T_{2'\text{-OMe}})_{14}T$ used as a reference. It suggests that the 4'-alkoxynucleotide units adopt predominantly the C3'-*endo* sugar conformation (*N*-type), similarly as the 2´-*O*-methylnucleoside units in 2´-*O*-methyloligonucleotides. We are aware of the importance of creating the mixed purine-pyrimidine 4´-alkoxyoligonucleotides (i.e., containing all of the U, C, A, and G nucleobases) for their direct use in biological experiments. The purine 4'-alkoxy-2'-deoxynucleosides have not been available in a preparative scale so far. Our effort in this direction has led a straightforward synthesis of 4´-alkoxy-2´deoxynucleosides in the purine series to make these compounds practically available on a preparative scale. These compounds are more than one order of magnitude stable under acidic conditions than corresponding 2'-deoxynucleosides. We believe that the 4'-alkoxy-2'-deoxynucleosides will be interesting building units for the assembly of the phosphodiester 4'-alkoxyoligodeoxynucleotides as new RNA chimeras. **Contribution of Group: design and synthesis of all compounds and Tm values measurement.**

B-4. *Non-isosteric phosphonate oligoribonucleotides.*[xx] Two types of phosphonate oligoribonucleotide analogs differing in the position of the methylene bridge in the isopolar nonisosteric internucleotide linkages (C3'-O-CH_2-P-O-C5" and C3'-O-P-CH_2-O-C5") were described. These regioisomeric oligonucleotides were synthesized from the 3' and 5'-

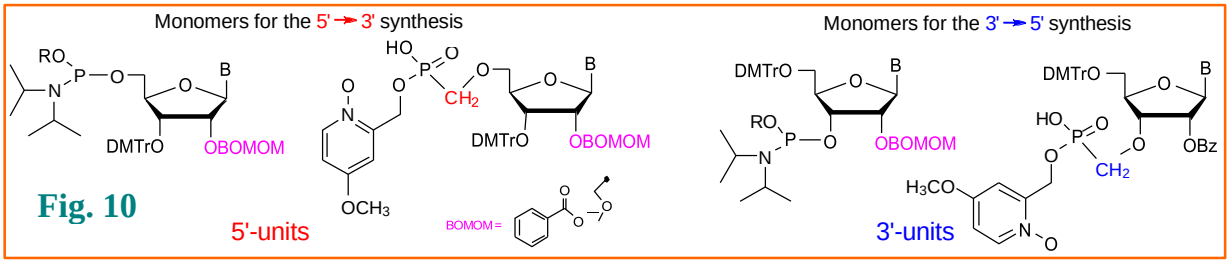


Fig. 10

monomeric units using advanced phosphotriester method. For the protection of the 2'-hydroxy group in the 5'-phosphonate and 5'- and 3'-phosphoramidite monomers, we developed a new 2´-*O*-benzoyloxymethoxymethyl group (BOMOM), the first non-silyl protecting group used in the solid phase synthesis of oligoribonucleotides in the reverse direction (**Fig. 10**). The use of gaseous ammonia allowed us for a simple on-column deprotection protocol. Both types of phosphonate linkages exhibited resistance against ribonuclease A, phosphodiesterase I and phosphodiesterase II. Concerning the hybridization properties of oligoribonucleotides modified by regioisomeric 3'- and 5'-phosphonate linkages, we found that the introduction of the 5'-phosphonate units resulted in moderate destabilization of the RNA/RNA duplexes (ΔT_m -1.8 °C/mod.), whereas the introduction of the 3'-phosphonate units resulted in considerably higher destabilization of the duplexes (ΔT_m -5.7 °C/mod.).

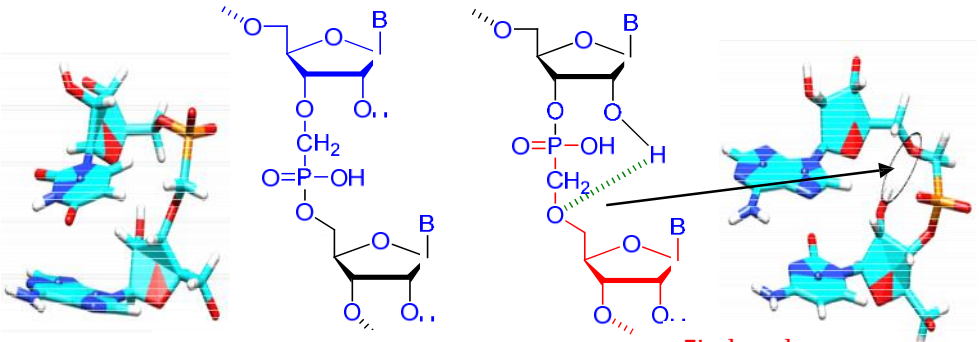


Fig. 11

The higher thermal stability of the duplexes modified with the 5'-phosphonate C3'-O-P-CH_2-O-C5" linkage was explained by molecular dynamic simulations which revealed a possible formation of intrastrand hydrogen bond stabilizing the conformation of the sugar-phosphonate backbone (**Fig. 11**). The 5'-phosphonate linkage (C3'-O-P-CH_2-O-C5") proved to have still sufficient hybridization potential to be used in biochemical applications, e.g., in siRNA assembly. This study is underway. **Contribution of Group: design and synthesis of all compounds and Tm values measurement.**

B-5. Antisense oligonucleotides working by RNase H mechanism.[xxi] Modified oligothymidylates containing various ratios of phosphodiester and unique isopolar 5'-O-methylphosphonate internucleotide linkages (in contrast to 3'-O-methylphosphonate and 5'-hydroxyphosphonate linkages[xxii]) were examined with respect to their hybridization properties with oligoriboadenylates and their ability to induce RNA cleavage by ribonuclease H (RNAse H), the key enzyme allowing the cleavage pathogenic mRNA strand in hybrid duplex antisense-DNA*RNA. The results demonstrated that the increasing number of 5'-O-methylphosphonate units in antisense DNA strand increases slightly the thermal stability of the hybrid duplexes and also up to 3-fold faster their cleavage by *E. coli* RNase H (depending on the ratio of phosphodiester and phosphonate linkages) in comparison with the natural duplex (data not shown).

Table 1. RNase H relative cleavage rate of miRNA191

AO	Gap size[a]	Sequence (5' - 3') of antisense oligonucleotide (AO)	T_m[b] [°C]	RNase H relative cleavage rate	
				SPR[c]	HPLC[d]
Amir-1	1	d(CAG CTG CTT TTG GGA TTC CGT TG)	72.0 (58.9)	0.7	1.1
Amir-2	2	d(CAG CTG CTT TTG GGA TTC CGT TG)	70.5 (61.1)	1.3	1.8
Amir-3	3	d(CAG CTG CTT TTG GGA TTC CGT TG)	70.2 (63.0)	2.9	2.0
Amir-4	4	d(CAG CTG CTT TTG GGA TTC CGT TG)	70.6 (63.6)	2.1	0.4
Amir-0	-	d(CAG CTG CTT TTG GGA TTC CGT TG)	66.2 (66.8)	1	1

A, **G**, **C**, and **T** denote 2'-deoxynucleoside-5'-*O*-methylphosphonate units; [a]number of consecutive natural units; [b]Melting temperature of the Amir*miRNA191 heteroduplex (the numbers in parentheses correspond to the melting temperatures of the duplexes formed by Amir AOs with DNA that was isosequential to miRNA191); [c] $K_{cat}^{AO}/K_{cat}^{natural\ DNA}$; [d] $t_{50}^{natural\ DNA}/t_{50}^{AO}$, t_{50} is time in which 50% of miRNA191 in the heteroduplexes were cleaved.

The most important fact is the finding a similar increase in RNase H cleavage activity for heteroduplexes composed of miRNA191 and complementary antisense oligonucleotides (AOs) containing 5'-O-methylphosphonate units (**Table 1**). This mofification allows the synthesis of efficient AOs working by RNase H mechanism to cleave a pathogenic RNA. We propose for this type of AOs, working via the RNase H mechanism, the abbreviation MEPNA (MEthylPhosphonate Nucleic Acids). RNase cleavage rate was determined by two independent methods – surface plasmon resonance (SPR) and HPLC.

In the light of these findings we have started the synthesis of nucleoside phosphonates (closely related to "active" 2'-deoxynucleoside-5'-*O*-methylphosphonates) bearing modifications in phosphonomethoxy P-CH_2-O- moiety; earlier prepared 5'-deoxynucleoside-5'-*S*-methylphosphonates[xxiii] containing P-CH_2-S- moiety will be evaluated in AOs for their ability to work by RNase H mechanism. Both *E. coli* and human recombinant RNase H will be used for further study. **Contribution of Group: design and synthesis of all monomers and modified oligonucleotides and RNase H cleavage HPLC assay.**

The pool of our nucleoside phosphonic acids provided several types of these compounds which were evaluated as monomers for the synthesis of modified oligonucleotides. We elaborated the synthesis of protected monomers in deoxyribo and ribo series suitable for the 3'→5' and 5'→3' directions of the solid phase synthesis using advanced phosphotriester method, and developed the first non-silyl protecting group used in the solid phase synthesis of modified oligoribonucleotides in the reverse direction. Superior nuclease stability of the phosphonate internucleotide linkages, their abilitiy to enhance hybridization together with ability to elicit RNase H activity more efficiently then the natural oligodeoxynucleotides, may clasify these compounds for their use in biochemistry and biology. Series of several types of modified oligonucleotides with isopolar isosteric and non-isosteric phosphonate-based internucleotide linkages have been prepared and evaluated as antisense compounds and siRNAs.

Research Report of the team in the period 2010–2014

Institute	Institute of Organic Chemistry and Biochemistry of the CAS, v. v. i.

Scientific team	Ullrich Jahn - Chemistry of Natural Products

The major research areas of the Jahn group are the total synthesis of natural products and their analogs as well as their biological evaluation in collaboration with partner groups. This requires a strong methodology development component to enable unique and efficient total synthesis approaches. In this field the group is especially active in electron-transfer chemistry, the development of new tandem processes, metal catalysis and more recently also photocatalysis. The group also pursues research to address fundamental questions of organic chemistry. At the title lines of the individual projects, external financial support if any is given (GSRC = Gilead Sciences & IOCB Research Centre; GA ČR = Grant Agency of the Czech Republic). Publications detailing the results during the evaluation period are cited at the end on an additional page.

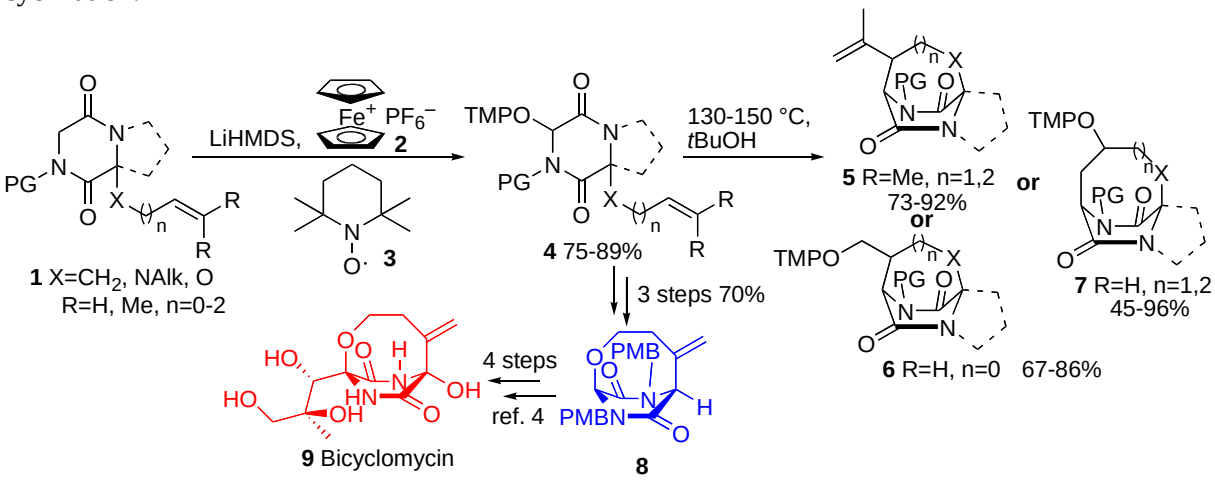
Contents:

2.1. Total synthesis of natural products and their biological evaluation

2.1.1. Total synthesis of bridged diketopiperazine alkaloids (GSRC)

Diketopiperazine (DKP) alkaloids are a diverse and interesting class of secondary metabolites with a number of important biological activities.[i] Our aim is the development of a unified synthetic approach to bridged DKP alkaloids (Scheme 1). The synthesis started with oxygenation of diketopiperazine **1** mediated by ferrocenium hexafluorophosphate **2** and TEMPO **3**, which provides oxygenated DKPs **4** in high yields (*cf.* 2.2.2.). Their thermal radical cyclizations using the persistent radical effect[ii] proceeded in very good yields efficiently furnishing bridged DKPs **5-7** with varying bridge sizes.[iii] Remarkably, medium-sized rings are assembled with extraordinary efficiency. The methodology was applied in a formal total synthesis of the effective antibiotic bicyclomycin **9** by accessing DKP **8** in 6 steps and an overall yield of 38%, from which **9** is accessible in four steps.[iv]

Diketopiperazines **1** are also suitable for approaches to bridged prenylated indole alkaloids, where the final ring is envisaged to be closed by adaption of a recently developed iron-mediated radical cyclization.[v]



Scheme 1. Approaches to the total synthesis of DKP alkaloids.

On the way to DKPs **1**, an efficient and simple preparation of the pyrazino[2,1-b]quinazoline-3,6-dione scaffold was found, which was applied to the successful synthesis of a number of alkaloids **14-16** with this core (Figure 1). Currently this methodology is applied to the multidrug resistance reversal-inducing alkaloid (–)-ardeemin **17**.

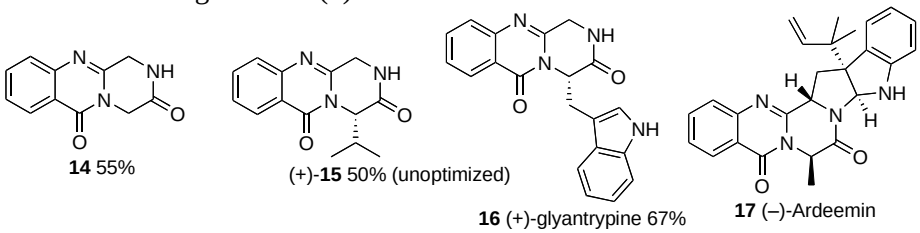
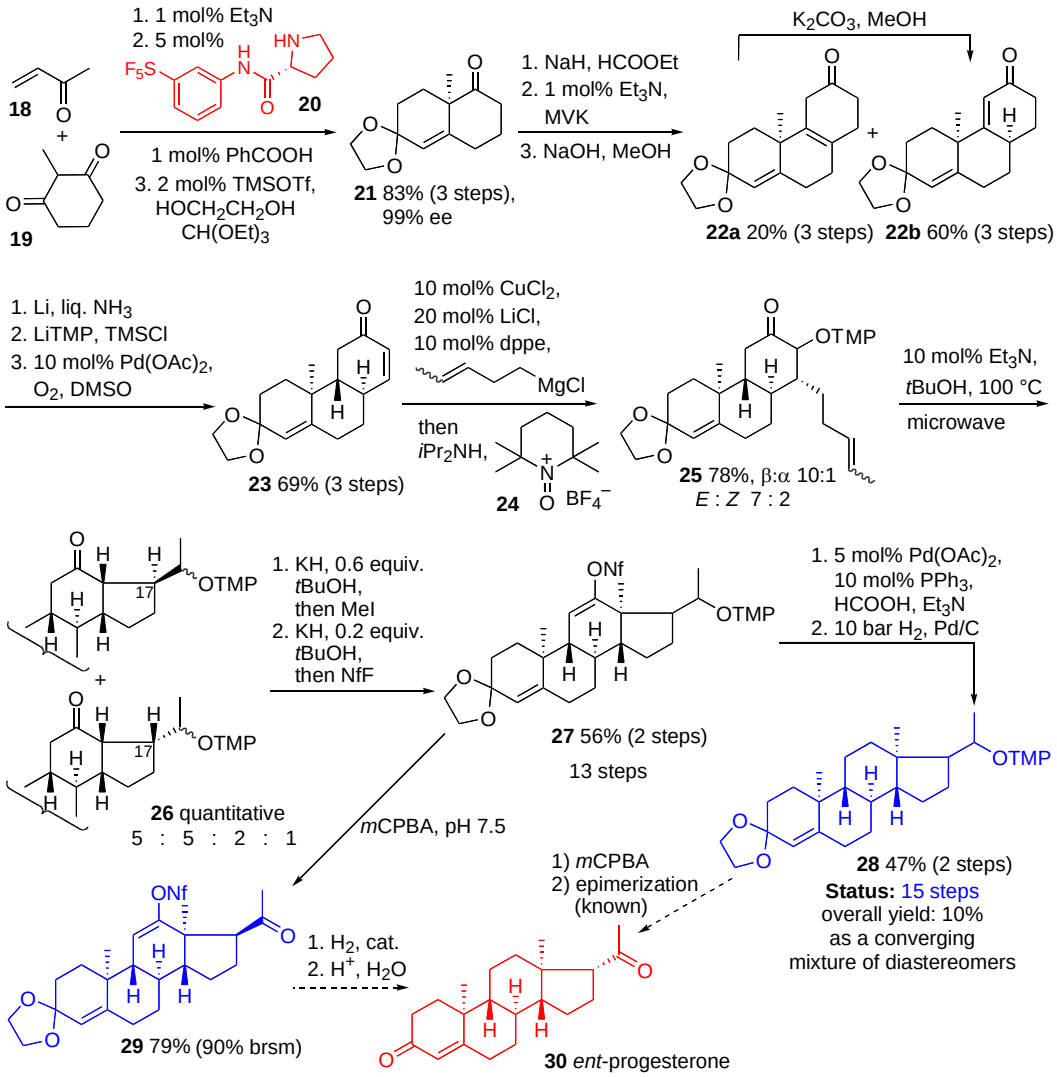


Figure 1. DKP alkaloids with pyrazino[2,1-b]quinazoline-3,6-dione cores.

2.1.2. Total synthesis of *ent*-steroids and analogs (GSRC)

The NMDA receptor mediates most of the excitatory synaptic transmissions in the brain. Its malfunction by blockage is connected to serious pathologic disorders, such as epilepsy or schizophrenia, and impairment of memory functions. A number of steroids exhibit neuroprotective actions, even that they partially block the receptor. Little is known about binding of neurosteroids to the receptor, a binding site was so far not identified.[vi] *ent*-Steroids are envisaged to be important tools for the elucidation of the mode of action, since potential binding sites may be identified. Since steroids are produced in Nature strictly in one enantiomeric form, only total synthesis provides access to *ent*-steroids. *ent*-Progesterone **30** was selected as the target, since it can be easily diversified to other steroid derivatives. Moreover an access to truncated derivatives, which can be evaluated for their ability to exert neuroprotection, is highly desirable. The minimum binding requirements will be

mapped by such an approach. This requires a synthetic strategy, in which the rings are subsequently appended.

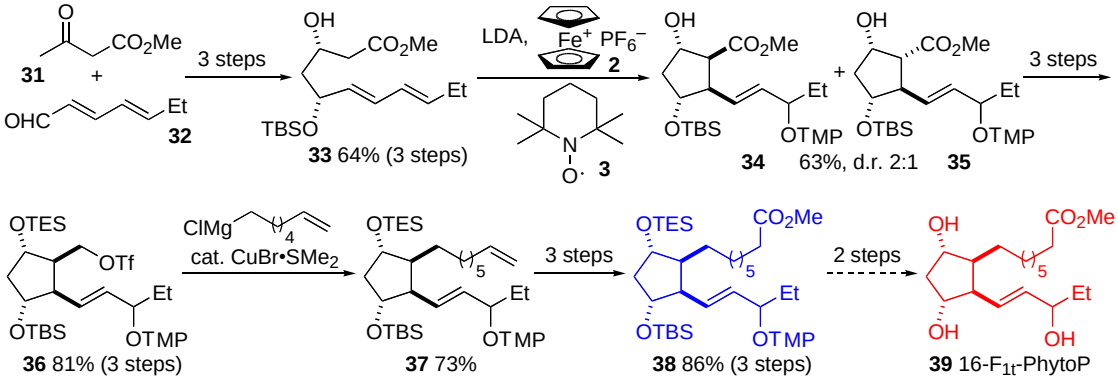


Scheme 2. Total synthesis of *ent*-steroids

The synthesis commenced with an asymmetric organocatalytic Robinson annulation of diketone **19** with vinyl ketone **18**. Since this reaction is traditionally slow, the new organocatalyst **20** was designed, which indeed accelerates the reaction compared with previous catalysts. The resulting bicycle **21** was subjected to a diastereoselective substrate-controlled Robinson annulation giving a mixture of tricycles **22a** and **22b**, which converged to conjugated **22b** on treatment with base. A three-step transposition of the enone unit afforded enone **23**. This compound was subjected to a new tandem conjugate addition/ oxygenation reaction developed in the group, which gave compound **25** in good yield and diastereoselectivity accompanied by 6% of cyclized **26**. The D-ring was appended by a thermal radical cyclization according to the persistent radical effect in very good yield, with exclusive *cis*-selectivity at the ring junction and a 3:1 β/α ratio at the 17-position of **26**. This is however inconsequential, since the center will converge later as desired. The diastereomeric mixture was methylated via the enolate at the 13-position to introduce the remaining carbon atom and subsequently converted to the enol nonaflate **27**. Compound **27** was further transformed via two strategies. The first consisted of transfer hydrogenation of the nonaflate and subsequent hydrogenation providing **28** in unoptimized 47% yield in 15 steps and good 10% overall yield. Alternatively, the alkoxyamine unit was first oxidatively removed to give **29** in good yield. From both precursors the target molecule *ent*-progesterone **30** can be accessed by standard steps, which will be performed in due course. Besides that truncated analogs (not shown) were synthesized from **21-23** to investigate their neuroprotective properties.[vii]

2.1.3. Total synthesis of autoxidatively formed lipid metabolites

Iso-, phyto- and neuroprostanes are major autoxidatively formed metabolites in mammals and plants. They are the gold standard for quantifying lipid oxidative stress and exert significant biological effects, ranging from cardiovascular to apoptosis-inducing effects.[viii] Since isolation from biological material is not practical, total synthesis is the only way to study biological effects in detail. In the evaluation period several potential metabolites of 15-E_2-IsoP were synthesized in 11 steps and 1.8% overall yield (not shown).[ix] More recently an approach to all isomeric and ring-substituted phytoprostanes was devised (Scheme 3). It starts with a vinylogous aldol addition giving the C8-18 unit. Functional group adjustment provided ester **33**. The second keystep was an oxidative radical cyclization with oxygenation in 16-position giving predominately the cyclopentanecarboxylate **34**. By switching to a magnesium base, the derivative **35** with prostaglandin configuration can be selectively synthesized (not shown). Ester **34** was manipulated efficiently to triflate **36**. The subsequent copper-catalyzed alkyl-alkyl cross-coupling to compound **37** with both side chains in place was a challenge since halide exchanges complicated the process. Transformation of **37** to ester **38** via hydroboration and two-step oxidation proceeded uneventfully. The remaining two steps to phytoprostane **39** concern global deprotection of the silyl and alkoxyamine functions.

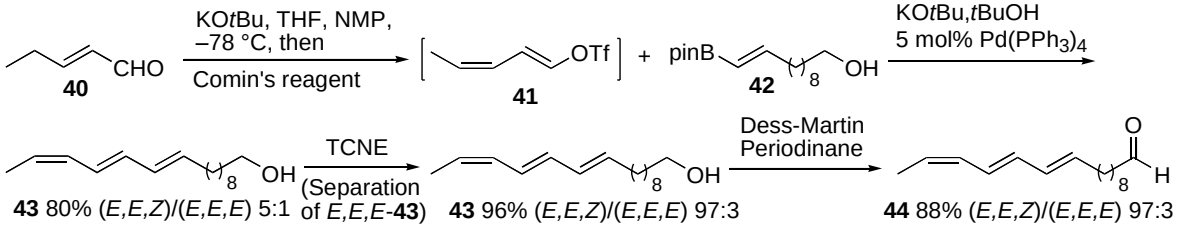


Scheme 3. Total synthesis of phytoprostanes.

The access to ester **35** enables an approach to „phytoglandins“, which have so far never been isolated from natural sources or synthesized, but which may also display biological activity. The orthogonal protecting group pattern will enable approaches to the different ring-substituted A-E- and J-type PhytoP. This synthesis also paves the way for the total synthesis of the other ω-3 fatty acid-derived isoprostanes and neuroprostanes.

2.1.4. Structure elucidation and total synthesis of insect pheromones (GA ČR 2010-2012)

Manduca sexta pheromone (10E,12E,14Z)-hexadecatrienal): In a collaborative project aimed at the elucidation of the desaturase enzymes involved in the biosynthesis of *Manduca sexta* sex pheromones the trienal sex pheromone component **44** was synthesized using the new Suzuki-Miyaura-type coupling of pentenal **40** and vinylboronate **42** (*cf.* 2.2.5.2.), the shortest total synthesis in four steps and 53% overall yield (Scheme 14). The Jahn group performed the syntheses of the pheromones, the Pichova and Svatos groups the biochemical work.[3]



Scheme 4. Total synthesis of the *Manduca sexta* sex pheromone hexadecatrienal **44**.

Structure elucidation and total synthesis of the Stylops muelleri sex pheromone (3R,5R,9R)-trimethyldodecanal: These bee-parasites are a mysterious genus about which only little was known including their reproduction. In a collaborative project with the Straka, Cvacka and Valterova groups

the sex pheromone was identified based on diastereodivergent total syntheses of several proposed structures (Fig. 2). The pheromone was proven to be (*3R,5R,9R*)-trimethyldodecanal **45** by the first total synthesis. The synthesis features two different asymmetric copper-catalyzed conjugate additions, which allowed the access to all enantio- and diastereomers of **45** and an unequivocal comparison with the natural product. The synthesis was completed in 11 steps (longest linear sequence) and 16% overall yield. The biosynthesis of the pheromone was elucidated and the efficacy was proven in field tests.[x] The Jahn group performed all synthetic work, the Cvacka group the analytical chemistry and the Straka and Valterova groups the biological testing.

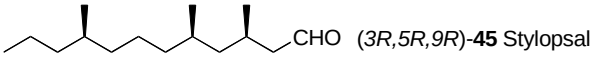


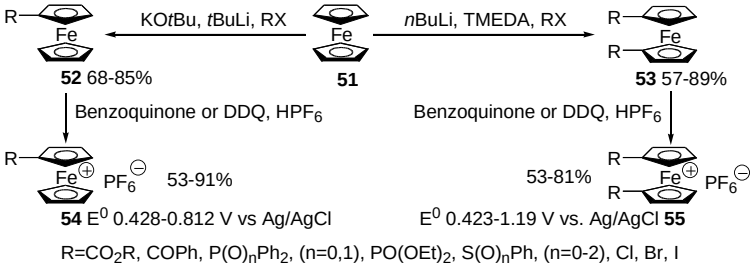
Figure 2. Structure of stylopsal, the first identified sex pheromone of *Strepsiptera*.

Identification of the trail pheromone of the unusual termite Hodotermes mossambicus: These harvester termites, being the only above-ground foraging and therefore having eyes, secrete in contrast to all other termites, who mainly use 3,6,8-dodecatrienol (for which we also developed a short total synthesis, not shown) as trail pheromones, one of unknown complex structure. In collaboration with the Hanus group a structural proposal **50** was developed. Since neither the constitution nor the configuration is proven, a regio-, enantio- and diastereodivergent total synthesis is devised, which will on one hand hopefully confirm the constitution and on the other hand provide enough information on the absolute configuration to minimize the synthetic effort to confirm the correct structure.

2.2. Development of synthetic methodology as an enabling technology

2.2.1. Novel ferrocenium salt oxidants

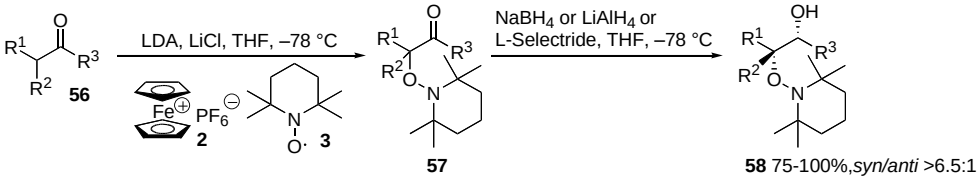
The Jahn group introduced ferrocenium hexafluorophosphate **2** as a convenient SET oxidant in organic chemistry and developed it to general applicability (*cf.* 2.1.1., 2.1.3., 2.2.3.). The ferrocene core offers in contrast to most other SET oxidants the unique opportunity to tune the oxidation power over a wide potential range. This has an unexplored potential in organic chemistry, but simple approaches to stronger ferrocenium oxidants represent a challenge. A number of mono- and diacceptor-substituted ferrocenes **52** and **53** were prepared in good yields using a unified approach (Scheme 6), in which monosubstituted **52** are selectively produced with Schlosser's base, while disubstituted **53** are generated by dilithiation using BuLi/TMEDA. The oxidation to ferrocenium salts **54** or **55** is most conveniently achieved with benzoquinone or DDQ. Their redox potentials were determined in collaboration with the Michl group, whereas the Rulisek group provided the theoretical insight for the prediction of the redox properties.[xi]



Scheme 6. New ferrocenium salts as oxidants.

2.2.2. Electron transfer-induced oxygenation of carbonyl compounds

The scope of α-oxygenations of carbonyl compounds **56** using ferrocenium hexafluorophosphate **2** and TEMPO **3** was determined (Scheme 7). Mechanistic studies proved that the reactions occur by a radical mechanism. Surprisingly, for some substrate classes, dimerization competed to a certain extent (not shown). However, the application of LiCl suppressed this undesired process, which was traced to enolate aggregate modification.[xii] Protected compounds **57** were reduced with good to excellent diastereoselectivity to monoprotected diols **58** using a number of reducing agents. Especially high selectivity (>20:1) was achieved with L-Selectride. This two-step protocol allows thus the stereoselective access to monoprotected *syn*-diols from ketones.[3]

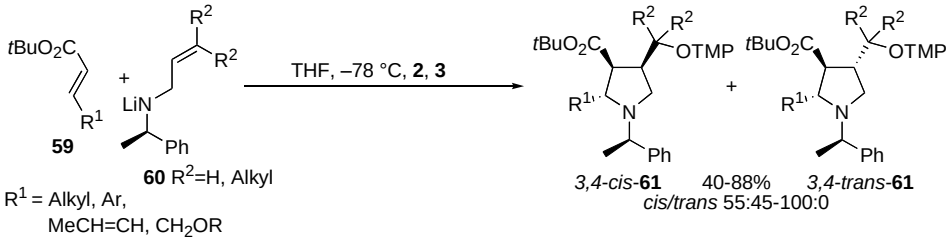


Scheme 7. Enolate oxygenation by ferrocenium hexafluorophosphate/TEMPO and reduction.

2.2.3. Tandem organometallic addition/radical cyclization reactions (GSRC)

2.2.3.1. Tandem lithium amide conjugate addition/radical cyclization/oxygenation

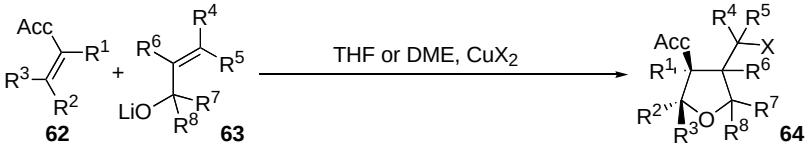
Tandem addition/cyclization reactions are an attractive strategy to obtain pyrrolidines. However, neither purely enolate nor radical sequences are possible. Therefore, asymmetric tandem conjugate additions of lithium *N*-(1-phenylethyl)-*N*-allylamides **59** to *tert*-butyl enoates **60**, followed by SET oxidation by **2**, radical cyclization and radical coupling with **3** were devised (Scheme 8). In this tandem process a C-N, a C-C and a C-O bond are formed and pyrrolidines **61** were isolated in good yields. The initial anionic aza-Michael addition step occurs essentially with complete diastereoselectivity. The 2,3-diastereoselectivity in the radical cyclization is completely *trans*, while the simple cyclization diastereoselectivity is dependent on the substituent at the allyl terminus. With R^2 = H diastereomeric ratios of >3:1 are observed irrespective of the nature of R^1, while low diastereoselectivity was found for R^2 = Alkyl. Products **61** can be transformed to useful derivatives in a number of follow-up reactions, which makes them useful scaffolds for medicinal chemistry. The strategy will serve as a basis for approaches to natural products (*see 3.1.*).



Scheme 8. Tandem aza-Michael addition/radical 5-exo cyclization/oxygenation reactions.

2.2.3.2. Tandem lithium alkoxide conjugate addition/radical cyclization reactions

Similar oxidative tandem reactions of alkoxides **63** to acceptor-substituted olefins **62** lead to tetrahydrofurans **64**. In most cases only two diastereomers are formed in the sequences (Scheme 9). Here copper(II) halides often serve as convenient oxidants and ligand transfer reagents at the same time.[xiii] Asymmetric diastereoselective Michael additions are possible if R^2 is chiral. Biological testing revealed that one of the tetrahydrofuran derivatives **64** is a low nanomolar inhibitor for a panel of hepatitis C viruses having only a low cytotoxicity. A number of other compounds are similarly active against individual HCV genotypes. Some derivatives of **64** will be applied as non-natural constrained amino acids.

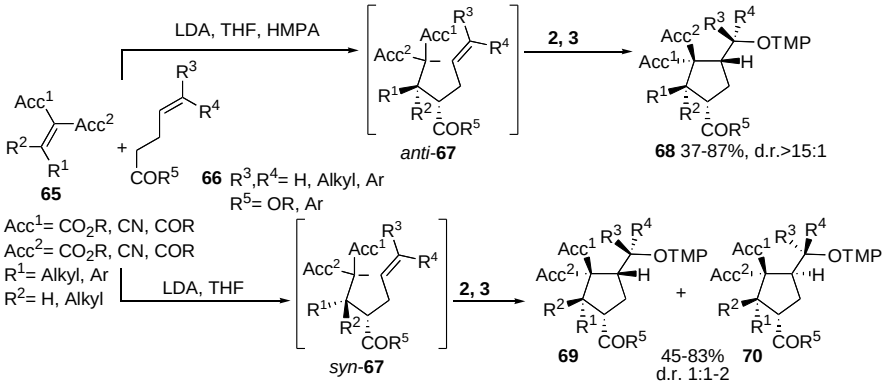


Scheme 9. Tandem alkoxide conjugate addition/radical 5-exo cyclization/coupling sequences.

2.2.3.3. Tandem Michael addition/radical cyclization sequences

Michael additions are convenient steps in tandem processes, since at the end of the reaction another reactive intermediate is generated that can be further transformed. Enolate Michael additions proved to be indeed very useful in tandem organometallic/radical processes (Scheme 10).[xiv] When β-dicarbonyl compounds **65** were treated with the (*E*)- or (*Z*)-enolates of **66** either *syn*- or *anti*-Michael adducts result as was shown in an individual study of these reactions. Enolates **67** are configurationally stable (irreversible addition) and are subject to oxidative cyclizations with **2** and **3**. From (*Z*)-enolates of **66**, cyclopentanes **68** result in mostly good yields and very good diastereoselectivity. (*E*)-Enolates of **66**

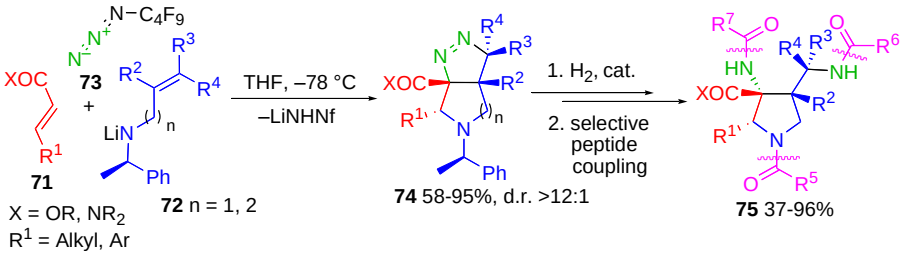
provide in contrast products **69** and **70** in almost equal proportions. The stereochemical divergence can be traced to alternative Beckwith-Houk transition states of the radical cyclization. While only one energetically feasible transition state can be reached from *anti*-**67**, two are accessible for the radicals derived from *syn*-**67**. In case the cyclized radical is prochiral ($R^3 \neq R^4$), even good diastereoselectivity in the radical coupling with **3** can be achieved.



Scheme 10. Tandem enolate Michael addition/radical cyclization/oxygenation reactions.

2.2.4. Tandem lithium amide conjugate addition/[3+2] cycloaddition reactions (GSRC)

Not only organometallic addition/radical cyclization sequences are feasible, also organometallic additions in sequence with 1,3-dipolar cycloadditions. A first example consists of tandem lithium amide conjugate addition/diazo transfer/[3+2] cycloadditions (Scheme 11).[xv] Unsaturated esters or amides **71**, chiral lithium amides **72**, and nonaflyl azide **73** as stable and perfectly suited diazo transfer reagent gave pyrrolidinopyrazolines **74** in mostly high yield and excellent diastereoselectivity. It is noteworthy that piperidinopyrazolines are similarly accessible, although the reaction times are considerably longer. Compounds **74** represent isoproline-based protected α,β,γ-triamino acid derivatives. Indeed, after selective hydrogenolysis all amino functions can be selectively addressed in peptide couplings giving derivatives **75**.

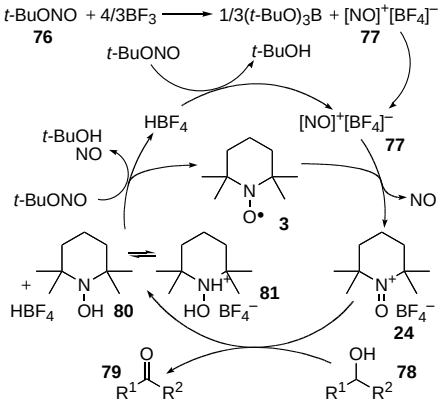


Scheme 11. Tandem amide conjugate addition/diazo transfer/1,3-dipolar cycloadditions.

2.2.5. Catalysis

2.2.5.1. Oxidative organocatalysis (GA ČR 2013-2015)

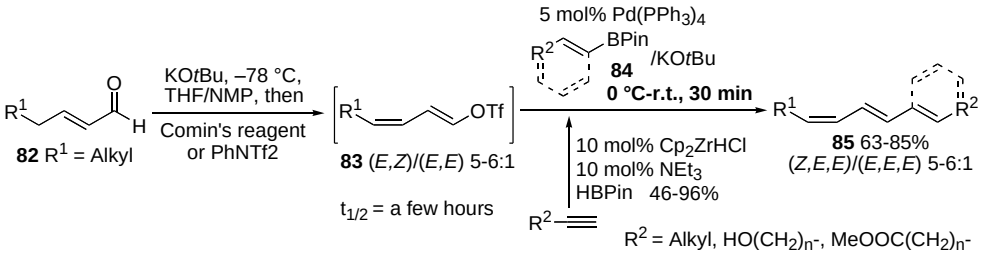
The extensive experience with TEMPO **3** and the aim to design traceless oxidative single electron catalysis (see 2.2.5.3.), required the development of methods to access *N*-oxopiperidinium ions under catalytic conditions. It was found that *tert*-butyl nitrite **76** in the presence of BF_3•OEt_2 is an efficient oxidation system for TEMPO **3** (Scheme 12). It generates the nitrosonium ion **77** in situ, which is the active oxidant for **3**. The resulting *N*-oxopiperidinium ion **24** is the active oxidant in the nitroxide-catalyzed oxidation of alcohols **78** to carbonyl compounds **79**. The nitrite serves subsequently also to reoxidize the equilibrium mixture of *N*-hydroxypiperidine **80** and its salt **81** to **3**. The scope of this oxidation method for alcohols is wide.[xvi] Significant advantages are that only very volatile byproducts, such as *t*BuOH and NO result and product isolation is very easy, which is especially important for sensitive compounds.



Scheme 12. TEMPO-catalyzed alcohol oxidation with the *t*BuONO/BF_3 system.

2.2.5.2. Transition metal catalysis

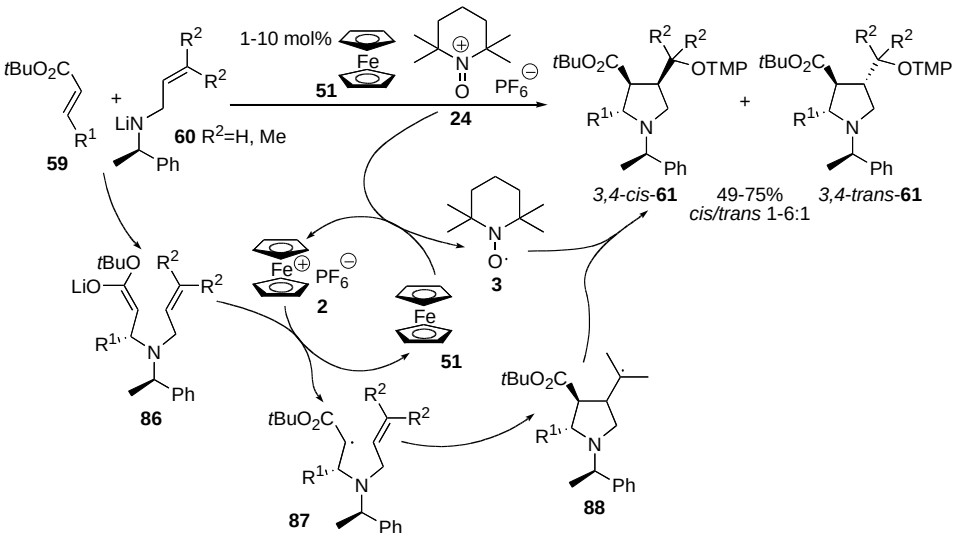
The Suzuki-Miyaura coupling is widely used in material-oriented chemistry, but often fails with sensitive substrates because of the relatively harsh conditions. Highly toxic thallium compounds must be used to avoid them. As a prerequisite for the total synthesis of the triene pheromone (*cf.* 2.1.4.), an extraordinarily mild and fast coupling was required. We discovered that α,β-unsaturated aldehydes **82** could be converted to the corresponding (*E,Z*)-dienyl triflates **83** (Scheme 13). These very unstable intermediates with a half-life of a few hours can be directly in situ coupled by treatment with a solution of the palladium(0) catalyst and a premixed mixture of vinyl or arylboronates **84** and KO*t*Bu. A clean cross-coupling to trienes, tetraenes, arylalkenes or aryldienes **85** occurred after very short reaction times at room temperature or below. This method belongs thus to the mildest Suzuki-Miyaura coupling methods. It tolerates very labile dienyl and also trienyl triflates, free alcohol and ester units.



Scheme 13. Mild Suzuki-Miyaura-coupling reactions of enals with vinyl- or arylboronates.

2.2.5.3. Electron transfer catalysis (GA ČR 2013-2015)

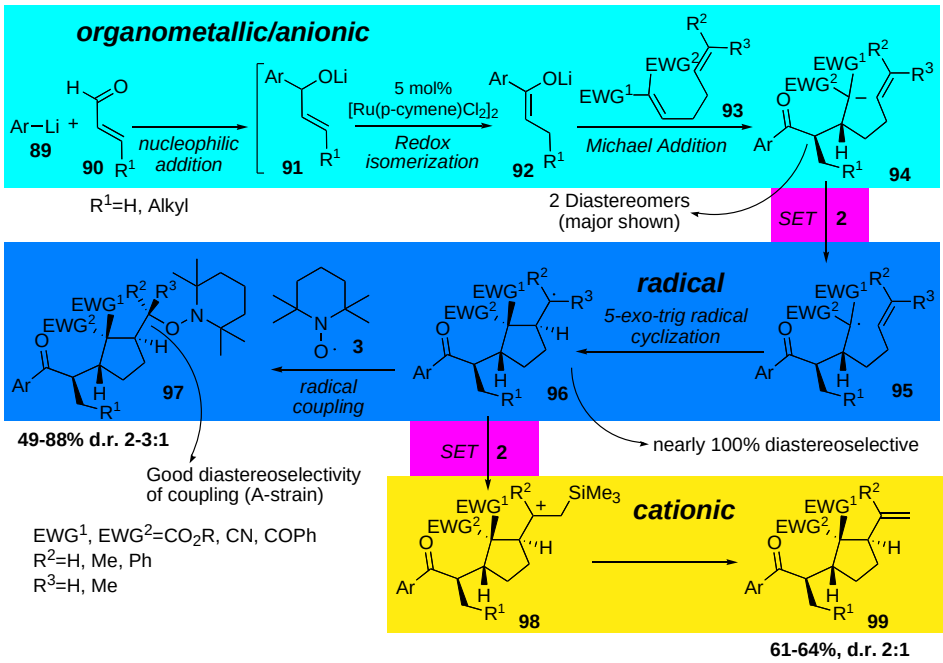
The oxidative (tandem) reactions or total synthesis efforts presented above (*cf.* 2.1.1., 2.1.3., 2.2.2., 2.2.3.) require as most other SET-mediated oxidative or reductive processes stoichiometric amounts of the SET oxidant **2**. It was found that **24** is a powerful oxidant for ferrocene **51** generating ferrocenium hexafluorophosphate **2** in situ. At the same time TEMPO **3** is formed, which can play its obvious role as a radical coupling reagent. The catalytic system **51**/**24** was probed initially in tandem lithium amide conjugate addition/radical cyclization/oxygenation sequences (Scheme 14) of **59** and **60** as well as in Michael addition/radical cyclizations (not shown).[xvii] The enolate **86** was formed as usual; the radical cyclization is triggered by addition of a catalytic amount of **51** followed by small portions of **24**. In this way the concentration of radical **87** and the persistent radical TEMPO **3** is kept at a low level, which prevents premature trapping. This was proved by including slower 6-endo radical cyclization steps. The progress of the reaction is easily monitored by the appearance of the blue color of **2** after **86** was completely consumed. The yields and diastereoselectivity of the process are more or less unchanged compared to the stoichiometric experiments (*cf.* 2.2.3.1.).



Scheme 14. Catalytic tandem amide conjugate addition/radical cyclization/oxygenation sequences.

2.2.5.4. Tandem transition metal catalysis/electron transfer reactions (GA ČR 2009-2011)

A completely unexplored field is the combination of two-electron transition metal catalysis with main group organometallic chemistry and radical reactions in domino processes.[xviii] The unprecedented coupling of a two-electron ruthenium-catalyzed redox isomerization (**91→92**) with SET processes enables the link between nucleophilic addition chemistry, Michael addition chemistry, radical chemistry and even carbocation generation, if the structural prerequisites exist (Scheme 15).[xix] In these sequences three C-C bonds and one C-O bond are formed. Only two out of eight or 16 diastereomers of compounds **97** or **99**, respectively, form. The Michael addition of enolate **92** to unsaturated carbonyl compound **93** is the stereoselectivity-limiting step, all others occur with high diastereoselectivity. Remarkably, a number of the products **97** displayed good antiviral activity against the Dengue virus. Similar sequences, in which enolates **92** were oxidatively dimerized with good to excellent *d,l*-diastereoselectivity, were also developed (not shown).

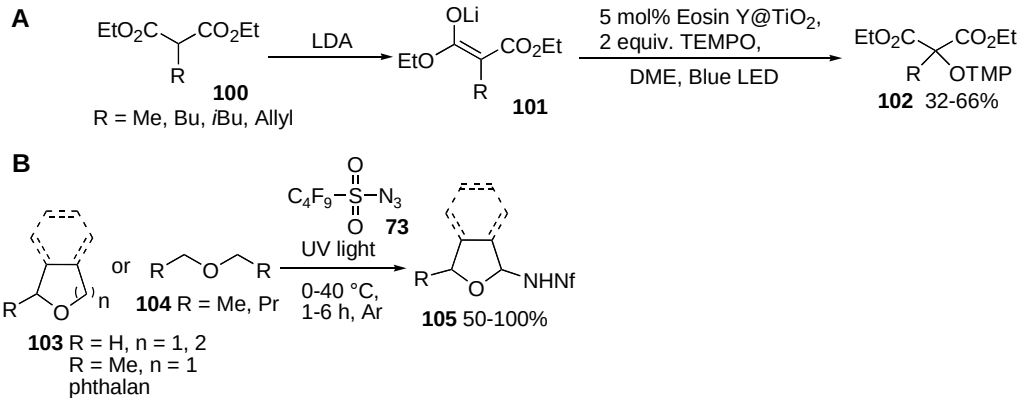


Scheme 15. Redox and intermediate diversity in tandem processes.

2.2.5.5. Photocatalysis

Since photo(redox)catalysis allows performing reactions under very mild conditions, it is an attractive strategy for the functionalization of organic compounds and to quickly generate complexity, by

triggering electron transfer processes.[xx] We initially aimed at photocatalytic oxidations of enolates using heterogeneous photocatalysts since they are less able to participate in undesired coupling reactions. Eosin Y-modified titanium dioxide indeed triggered the single-electron oxidation of malonate enolates **101** to the corresponding radicals under irradiation with blue LEDs, which were coupled with TEMPO to oxygenated malonates **102** in reasonable yields (Scheme 16A).[xxi] These conditions do not lead to degradation of the substrates as observed with TiO_2 alone or in the presence of UV light. Thus a new platform for selective orthogonal radical generation exists.

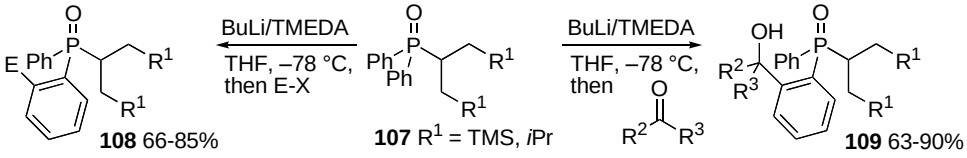


Scheme 16. Photocatalytic oxidative radical generation and photolytic amination of ethers.

In parallel, we serendipitously discovered that nonaflyl azide **73** reacts with THF in daylight with C-H amination, a very interesting reaction type (Scheme 16B). Based on this general conditions for the C-H amination of ethers **103** or **104** were developed. The scope and the mechanistic features are under investigation and recently also conditions for the direct amination of hydrocarbons, such as cyclohexane, were identified (not shown). These amination conditions may be very attractive in introducing amino functions into biomolecules.

2.2.5.6. New phosphine ligands for catalysis (GA ČR 2011-2013)

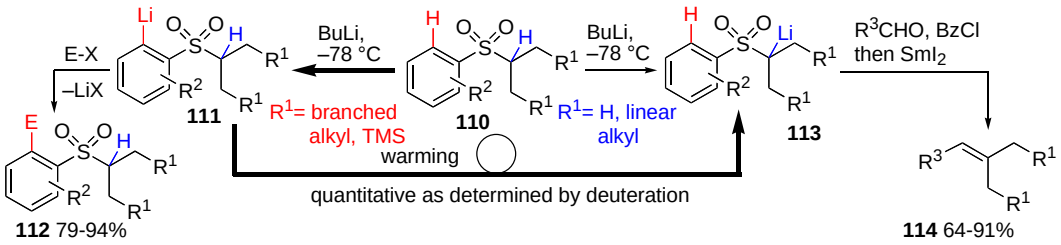
New and tailored ligands for catalysis are important for the discovery of catalytic processes. Branched phosphine oxides **107**, which were synthesized in high yield by double alkylation of $MeP(O)Ph_2$ with alkyl iodides undergo in contrast to most other phosphine oxides no α-deprotonation, but a directed *ortho*-metalation with BuLi/TMEDA (Scheme 17). In contrast to sulfones (see 2.2.6.) no rearrangement to the α-phosphinoyllithium compounds proceeds. The resulting aryllithium reacted with various electrophiles. This allows a very easy access to chiral, but so far racemic phosphine oxides **108** and **109**.[xxii] The reduction to phosphines succeeded in good yield (not shown) and thus new chiral phosphine ligand classes or Lewis bases can be envisaged.



Scheme 17. Directed *ortho*-metalation of diphenylphosphine oxide **54**.

2.2.6. A new carbanion rearrangement and its application (GA ČR 2011-2013)

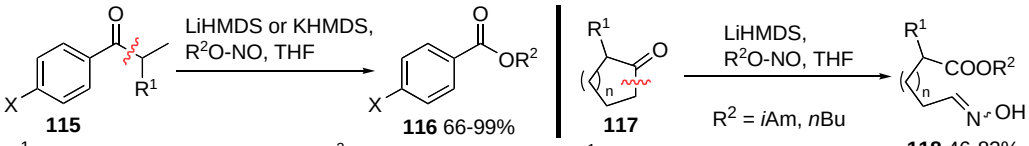
An attempted Julia olefination using sulfone **110** (R^1 = TMS) gave surprisingly an addition of the aryl unit. A closer inspection revealed that directed *ortho*-metalation to aryllithium **111** is kinetically favored over thermodynamic α-deprotonation to **113** (Scheme 18). Intermediate **111** has a sufficient lifetime to be intercepted by a number of electrophiles providing **112** in good to excellent yield. On warming a quantitative rearrangement of **111** to **113** proceeds, which can be subsequently subjected to Julia olefination to alkenes **114**. Nonbranched sulfones form in contrast **114** directly. The mechanism and the kinetics of the processes have been determined, so that it can be predictively applied in target-oriented synthesis.[xxiii]



Scheme 18. Divergent deprotonation of sulfones and their carbanion rearrangement.

2.2.7. Nitrosative cleavage of ketones (GA ČR 2013-2015)

In the course of identifying new traceless stoichiometric oxidants, which are incorporated into the products during catalytic SET oxidations, we discovered that alkyl nitrites react directly with enolates (Scheme 19). Aromatic ketones **115** as well as cyclic aliphatic ketones **117** undergo apparently first a nitrosation. The co-formed lithium or potassium alkoxide adds apparently in situ to the carbonyl group, which leads via a formal retro-Claisen reaction to products **116** and **118**, respectively. The value of this method lays in the cleavage of typically kinetic enolates, thus the method provides products with opposite regioselectivity than the classical Baeyer-Villiger oxidation or Beckmann rearrangement. Both ester and oxime functions in **116** or **118** provide ample further functionalization opportunities.



Scheme 19. Nitrosative cleavage of carbonyl compounds.

2.3. Fundamental organic chemistry research.

2.3.1. New organosuperbases rivaling the benchmark phsophazene bases

This project aimed at new imidazole superbases was performed after Dr. I. Lyapkalo's tragic death. New imidazole bases **119** proved to be stronger than Schwesinger's P_1-phosphazene bases (Fig. 3).[xxiv] Combining the imidazole element with guanidine units as in **120** leads to the currently strongest phosphorus-free organosuperbases being with pK_a values of up to 29.3 in THF stronger than P_2-bases.[xxv] Their base strength rests on the new design principles of steric strain release, aromatization on protonation as well as extensive charge delocalization on protonation.

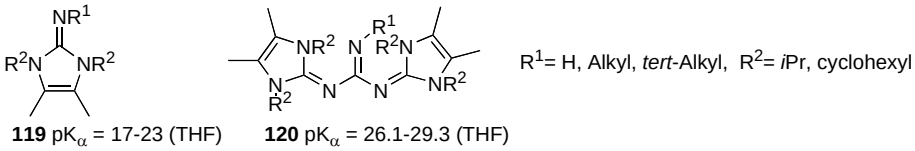


Figure 3. New imidazole organosuperbases.

2.3.2. New porphyrin-thiophene hybrid compounds

A collaborative project on the synthesis of new thiophene-porphyrin hybrid materials, such as **121**, was successfully completed in 2010 (Fig. 4). It concerned the synthesis of electropolymerizable constructs for the application in organic electronic materials.[xxvi]

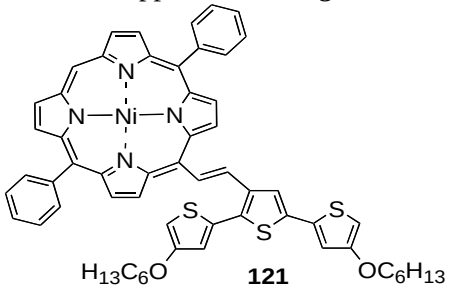


Figure 4. Electropolymerizable Porphyrin-thiophene constructs.

References

[i] Selected reviews: a) Li, S. M. *Nat. Prod. Rep.* **2010**, *27*, 57-78. b) Borthwick, A. D. *Chem. Rev.* **2012**, *112*, 3641-3716. c) Amatov, T.; Jahn, U. *Angew. Chem. Int. Ed.* **2014**, *53*, 3312-3314.
[ii] Review: Studer, A. *Chem. Soc. Rev.* **2004**, *33*, 267-273 and cited reviews.
[iii] Manuscript ready for submission.
[iv] Williams, R. M.; Armstrong, R. W.; Dung, J.-S. *J. Am. Chem. Soc.* **1984**, *106*, 5748-5750.
[v] Lo, J. C.; Yabe, Y.; Baran, P. S. *J. Am. Chem. Soc.* **2014**, *136*, 1304-1307.
[vi] Review: Covey, D. F. *Steroids* **2009**, *75*, 577-586.
[vii] Kudová, E.; Chodounská, H.; Kapras, V.; Vyklický, L.; Valeš, K.; Jahn, U. Czech patent application PV-2014-575 „Amfifilni sloučeniny s neuroprotektivnimi účinky“, 26.08.2014.
[viii] Selected Reviews: a) Jahn, U.; Galano, J. M.; Durand, T. *Angew. Chem. Int. Ed.* **2008**, *47*, 5894-5955. b) Jahn, U.; Galano, J. M.; Durand, T. *Prostaglandins Leukotrienes Essent. Fatty Acids* **2010**, *82*, 83-86. c) Jahn, E.; Durand, T.; Galano, J.-M.; Jahn, U. *Chem. Listy* **2014**, *108*, 301-319.
[ix] Jahn, U.; Dinca, E. *J. Org. Chem.* **2010**, *75*, 4480-4491.
[x] a) Cvačka, J.; Jiroš, P.; Kalinová, B.; Straka, J.; Černá, K.; Šebesta, P.; Tomčala, A.; Vasičková, S.; Jahn, U.; Šobotník, J. *J. Chem. Ecol.* **2012**, *38*, 1483-1491. b) Lagoutte, R.; Šebesta, P.; Jiroš, P.; Kalinová, B.; Jirošová, A.; Straka, J.; Černá, K.; Šobotník, J.; Cvačka, J.; Jahn, U. *Chem. Eur. J.* **2013**, *19*, 8515-8524.
[xi] Khobragade, D. A.; Mahamulkar, S. G.; Pospíšil, L.; Císařová, I.; Rulíšek, L.; Jahn, U. *Chem. Eur. J.* **2012**, *18*, 12267-12277.
[xii] Dinca, E.; Hartmann, P.; Smrček, J.; Dix, I.; Jones, P. G.; Jahn, U. *Eur. J. Org. Chem.* **2012**, 4461-4482.
[xiii] a) Jahn, U.; Rudakov, D.; Jones, P. G. *Tetrahedron* **2012**, *68*, 1521-1539. b) Jahn, U.; Rudakov, D.; Jones, P. G. *Tetrahedron* **2012**, *68*, 447-463.
[xiv] a) Holan, M.; Pohl, R.; Císařová, I.; Jahn, U. *Eur. J. Org. Chem.* **2012**, 3459-3475. b) Holan, M.; Pohl, R.; Císařová, I.; Klepetářová, B.; Jones, P. G.; Jahn, U. *Chem. Eur. J.* **2015**, *21*, *in press.*
[xv] Kapras, V.; Pohl, R.; Císařová, I.; Jahn, U. *Org. Lett.* **2014**, *16*, 1088-1091.
[xvi] Holan, M.; Jahn, U. *Org. Lett.* **2014**, *16*, 58-61.
[xvii] Kafka, F.; Holan, M.; Hidasová, D.; Pohl, R.; Císařová, I.; Klepetářová, B.; Jahn, U. *Angew. Chem. Int. Ed.* **2014**, *53*, 9944-9948.
[xviii] Review on radicals in transition metal catalysis: a) Jahn, U. *Top. Curr. Chem.* **2012**, *320*, 121-190. b) Jahn, U. *Top. Curr. Chem.* **2012**, *320*, 191-322. c) Jahn, U. *Top. Curr. Chem.* **2012**, *320*, 323-452.
[xix] Jagtap, P. R.; Ford, L.; Deister, E.; Pohl, R.; Císařová, I.; Hodek, J.; Weber, J.; Mackman, R.; Bahador, G.; Jahn, U. *Chem. Eur. J.* **2014**, *20*, 10298-10304.
[xx] Selected recent book: *Chemical Photocatalysis*; König, B., Ed.; deGruyter: Berlin, 2013.
[xxi] Salzl, S.; Master thesis, 2014.
[xxii] Mahamulkar, S. G.; Císařová, I.; Jahn, U. *Adv. Synth. Catal.* **2015**, *357*, 793-799.
[xxiii] a) Puget, B.; Jahn, U. *Synlett* **2010**, 2579-2582. b) Řehová, L.; Císařová, I.; Jahn, U. *Eur. J. Org. Chem.* **2014**, *2014*, 1461-1476. c) Řehová, L.; Jahn, U. *Eur. J. Org. Chem.* **2014**, 4610-4623.
[xxiv] Kunetskiy, R. A.; Polyakova, S. M.; Vávřik, J.; Císařová, I.; Saame, J.; Nerut, E. R.; Koppel, I.; Koppel, I. A.; Kütt, A.; Leito, I.; Lyapkalo, I. M. *Chem. Eur. J.* **2012**, *18*, 3621-3630.
[xxv] Vazdar, K.; Kunetskiy, R.; Saame, J.; Kaupmees, K.; Leito, I.; Jahn, U. *Angew. Chem. Int. Ed.* **2014**, *53*, 1435-1438.
[xxvi] a) Zöllner, M.; Frähmcke, J.; Elstner, M.; Jahn, U.; Jones, P.; Becker, E.; Kowalsky, W.; Johannes, H. *Macromol. Chem. Phys.* **2010**, *211*, 359-371. b) Zöllner, M. J.; Frähmcke, J. S.; Elstner, M.; Jahn, U.; Jones, P. G.; Becker, E.; Kowalsky, W.; Johannes, H. H. *Macromol. Chem. Phys.* **2010**, *211*, 359-371.