2.5.1 Research area "Ab initio methods" (W. Thiel)

Involved: J. Breidung, H. Lin, M. Schreiber, A. Yachmenev, S. N. Yurchenko, J. Zheng

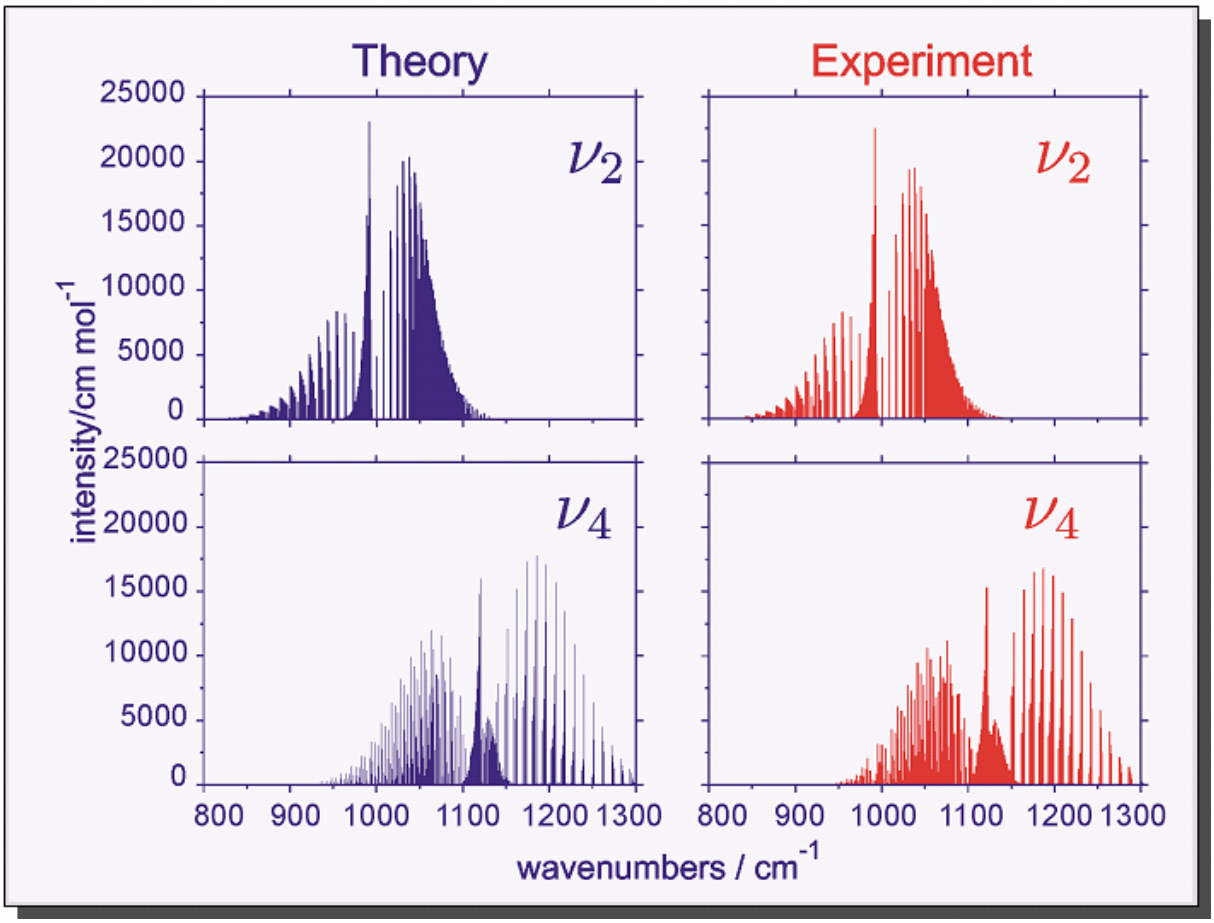
Objective: Vibration-rotation and electronic spectra of small molecules are computed with high accuracy using high-level ab initio calculations with large basis sets. This includes the development of a general variational treatment of nuclear motion that allows the prediction of rovibrational energies and intensities not only for semirigid molecules, but also for molecules with large amplitude motions. Highly correlated ab initio methods are used to provide theoretical benchmark data for the electronically excited states of representative organic chromophores.

Results: The theoretical prediction of vibration-rotation spectra requires the generation of accurate potential energy and dipole moment surfaces, followed by the variational calculation of rovibrational energies and intensities. For the former task, we employ ab initio electronic structure methods, typically coupled cluster theory with large basis sets (e.g., CCSD(T)/aug-cc-pVQZ in standard notation), possibly with complete basis set (CBS) extrapolation and corrections for core-valence correlation and relativististic effects. For the latter, we have developed and coded a variational treatment of nuclear motion that is based on the Hougen-Bunker-Johns approach with an Eckart-frame kinetic energy operator and thus also handles large amplitude motion. The corresponding theory for the energies [204] and intensities [205] was first formulated for pyramidal XY_3 molecules and later generalized [551]. A distinctive feature of our approach is that the kinetic energy operator is not derived explicitly, but generated numerically via recursion relations to any desired accuracy. As a result of this development, we now have a general and robust variational code for nuclear motion which has been shown to yield converged results [551].

This program has been applied to study NH_3 [203, 207], PH_3 [206, 399], SbH_3 [400], and BiH_3 [202, 400]. In the protypical case of ammonia, the challenge is to compute the complete vibration-rotation-inversion spectrum as accurately as possible. This involves the calculation of the six-dimensional (6D) potential energy surface at the CCSD(T)/CBS level with corrections for core-valence correlation and relativististic effects (using up to 51816 grid points), the computation of the 6D dipole moment surface at the CCSD(T)/aug-cc-pVTZ level, the fitting of these surfaces to a suitable analytical form, and the 6D variational treatment of nuclear motion, with definition of appropriate basis functions and construction and diagonalization of the corresponding

Hamiltonian matrices. This purely ab initio approach leads to errors of 1.9 cm^{-1} for 24 inversion splittings and 4.3 cm^{-1} for 34 vibrational term values which are mostly due the neglect of higher excitations in the coupled cluster calculations. Since these errors are rather systematic, they can be reduced significantly (to 0.8 and 0.4 cm^{-1}, respectively) by a slight empirical adjustment of the ab initio potential. The computed rovibrational intensities agree very well with the available experimental data, and they are accurate enough to assist in line-by-line assignments of high-resolution spectra.

To illustrate the results obtained in a similar manner for phosphine, the computed and observed spectra for the ν_2 and ν_4 fundamental bands are shown in the Figure. It is obvious that the ab initio calculations reproduce the experimental spectra remarkably well.



Intensities of the ν_2 and ν_4 bands in phosphine: Ab initio theory vs. experiment

In the case of PH_3 [206] and its higher homologues SbH_3 and BiH_3 [400], our variational calculations predict rotational energy clusters with sixfold degeneracy at high rotational excitation. The corresponding energy-level scheme is qualitatively

different from the standard pattern for symmetric tops, due to centrifugal-force-induced dynamic symmetry breaking beyond a certain J threshold value, which in a semiclassical picture is associated with predominant rotation around any of the three X-H bonds. This phenomenon has been predicted theoretically and later confirmed experimentally in the case of XH_2 molecules. Our present calculations indicate that it should also be observable for XH_3 molecules that are close to the local mode limit.

Concerning further ab initio studies of vibration-rotation spectra, we have investigated the two lowest electronic states of the PH_2 radical to guide the assignment of the disperse fluorescence spectra measured by our experimental partners [290], and identified rotational energy clustering in its electronic ground state through variational calculations [401]. For the two lowest electronic states of the CH_2^+ cation, we have computed improved potential energy surfaces and a spin-orbit coupling surface, as input to variational calculations that yield excellent agreement with recently observed high-resolution spectra [425]. Using second-order rovibrational perturbation theory in combination with an ab initio (CCSD(T)/aug-cc-pVQZ) anharmonic potential, we have determined the relevant spectroscopic parameters of SbD_3 which are fully consistent with the analysis of the experimental high-resolution spectra [32, 245].

Our focus in the ab initio area is on vibration-rotation spectroscopy, but we also use ab initio methods in other projects (see below) for validation purposes. One noteworthy example is a benchmark study on electronically excited states that covers typical organic chromophores (28 molecules, 213 singlet and triplet excited states). Extensive ab initio calculations using multi-configuration perturbation theory (CASPT2) and coupled cluster theory (CC2, CCSD, CC3) have been performed to generate a reference data base (similar to the established ground-state thermochemistry data bases) that can be used for the assessment and development of lower-level theoretical methods.

Publications resulting from this research area: 32, 202, 203, 204, 205, 206, 207, 235, 245, 290, 341, 399, 400, 401, 402, 425, 551

External funding: European Research Training Network QUASAAR (MRTN-CT-2004-512202), German Research Council (DFG, SFB 663, project C4)

Cooperations: V. Boudon (Dijon, FR), P. R. Bunker (Ottawa, CA), H. Bürger (Wuppertal, DE), L. Fusina (Bologna, IT), P. Jensen (Wuppertal, DE), S. P. A. Sauer (Copenhagen, DK); other QUASAAR partners include A. Campargue (Grenoble, FR), J.-M. Flaud (Paris, FR), L. Halonen (Helsinki, FI), M. Herman (Brussels, BE), T. Rizzo (Lausanne, CH) and J. Tennyson (London, UK).

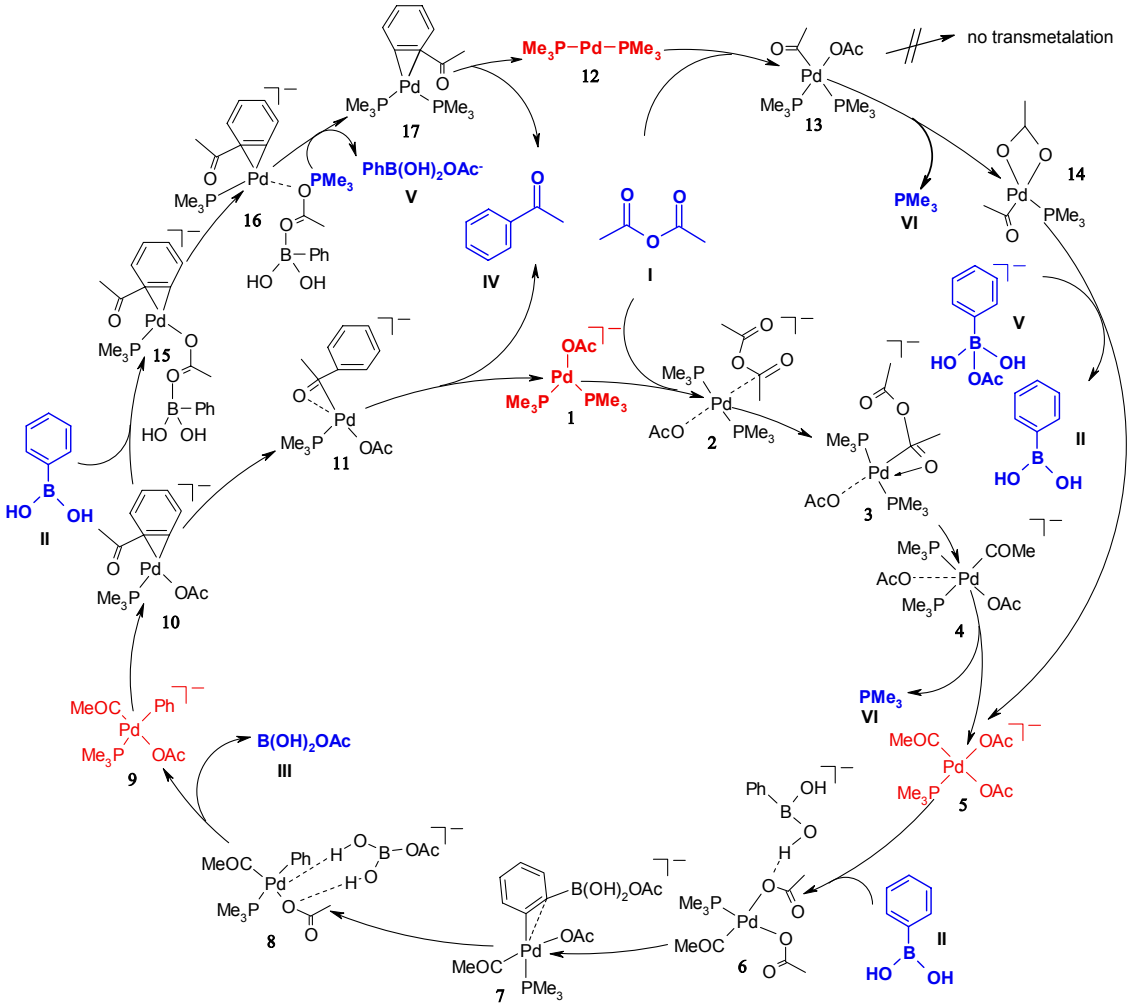
2.5.2 Research Area "Density Functional Methods" (W. Thiel)

Involved: A. Anoop, Z. Chen, A. Fu, M. Graf, H. Hermann, M. N. Jagadeesh, D. Koley, T. Tuttle, S. F. Vyboishchikov , D. Wang

Objective: Density functional methods are applied in studies of transition metal and other compounds in order to understand and predict their properties. Much of the work on homogeneous transition metal catalysis and organocatalysis involves a close collaboration with experimental groups at our Institute and aims at a detailed mechanistic understanding of the reactions studied experimentally.

Results: Many of our applications of density functional theory (DFT) focus on transition metal compounds. Based on previous extensive validation our standard DFT approach normally employs the Becke-Perdew functional with an effective core potential at the metal and with medium-sized polarized basis sets. In the case of organic compounds, we normally use the B3LYP hybrid functional.

Palladium-catalyzed cross coupling reactions have been investigated in cooperation with the Gooßen group [62, 63, 278]. The DFT calculations provide a plausible mechanism for the initial step, i.e., the oxidative addition of aryl halides to anionic three-coordinate Pd(0) complexes: the halide coordinates linearly to the palladium to form a hypervalent four-coordinate complex, that subsequently rearranges to the desired addition product without ever forming a five-coordinate Pd(II) intermediate as previously assumed [62]. More extensive DFT computations have been performed for the cross coupling reaction between phenylboronic acid and acetic anhydride, considering altogether five interconnected catalytic cycles that start from neutral PdL_2, anionic PdL_2X^-, and anionic $PdLX^-$ complexes (L=phosphine, X=acetate); two of these are sketched in the Figure. According to the calculations, the anionic pathways are overall more favorable. The initial oxidative addition of acetic anhydride to the anionic complexes leads to anionic Pd(II)monophosphine species with two acetate ligands (cis or trans, slight preference for cis variant) which then undergo transmetalation and reductive elimination [63, 278]. The computed catalytic cycles offer detailed mechanistic insight.



Main intermediates in two of the calculated catalytic cycles for
Pd-catalyzed cross coupling of phenylboronic acid to acetic anhydride

The joint studies with the Fink group on propene polymerization have been completed. Previous work had established the connection between the microstructure of zirconocene-based catalysts and the tacticity of the formed polymer, with excellent agreement between computed and observed pentad distributions at low temperatures. The DFT calculations indicate that the larger deviations at higher temperatures are due to the onset of (originally neglected) epimerization (back-skip) processes that become more favored at high temperatures for entropic reasons [279]. Car-Parrinello molecular dynamics (CPMD) simulations show that the activation of other internal motions does not play a significant role in this regard, but support the fundamental assumption of statistical propagation models that each insertion is independent of the preceding ones [81]. Concerning the zirconocene-catalyzed oligomerization of norbornene, the proposed σ-bond metathesis mechanism has been corroborated, and the helical microstructure of polynorbornene has been simulated [464]. Finally, the single-center

two-state model for the observed broken rate order in zirconocene-catalyzed ethylene polymerization has been investigated through DFT calculations of the relevant agostic conformers and the corresponding transition states [82].

Experimental work in the Reetz group has shown that rhodium catalysts with chiral monodentate phosphorous ligands can achieve asymmetric hydrogenation with high efficiency and enantioselectivity, and may thus serve as an economic alternative to the classical catalysts with bidentate ligands. Unlike the latter, the new catalysts follow the "lock-and-key" principle, i.e., the major enantiomer of the product is formed from the more stable of the two diastereomeric prochiral catalyst-substrate complexes [145]. We have performed a detailed DFT study of the enantioselective hydrogenation of itaconic acid using a chiral Rh(phosponite)$_2$ catalyst, combined with kinetic Monte Carlo simulations, to explain the differences between the two types of catalysts (see also section 2.5.5).

There are several other collaborative DFT projects within the Institute. The ring-closing metathesis reaction of large sterically hindered α,ω-olefins has been studied to understand the stereochemistry (Z/E preferences) observed in the Fürstner group during the synthesis of salicylihalamid and related model compounds [186]. The mechanism of the proline- and 2-methylproline-catalyzed alpha-alkylation of aldehydes has been investigated in cooperation with the List group, focusing on the crucial intramolecular nucleophilic substitution in the enamine intermediate: the added base accelerates the reaction through the electrostatic activation of the leaving group and affects the stereoselectivity by stabilizing the anti and syn transition states to a different extent, which offers an explanation for the differences found experimentally between proline and 2-methylproline [261]. An ongoing project with the List group concerns asymmetric organocatalytic reactions mediated by ion pairs. Motivated by experimental work in the Schüth group, we have studied the properties of the proton sponge 4,9-dichloroquino[7,8-h]quinoline and the role of its catalytically active palladium complexes in the Heck reaction.

Several DFT studies have been carried out without involvement of experimental groups from the Institute, for example on the noncovalent catalysis of the the Diels-Alder reaction by the neutral hydrogen bond donors thiourea and urea [262] and on ruthenium-based colorimetric sensors for fluoride [462]. In an external cooperation with an industrial partner, we have computed the mechanism of olefin hydrosilylation catalyzed by $RuCl_2(CO)_2(PPh_3)_2$ and $[RuCl(NCCH_3)_5]^+$ complexes [388, 540]: in both

cases, a σ-bond metathesis mechanism (with formation of a hydride intermediate) is favored over the conventional Chalk-Harrod mechanism that operates in the case of Pt-based catalysts. The computational results for these and other Ru-based catalysts are consistent with the available experimental evidence, and the mechanistic insight gained may be helpful for further optimization of these catalysts in an industrial setting.

Publications resulting from this research area: 40, 62, 63, 78, 81, 82, 145, 186, 261, 262, 278, 279, 388, 454, 462, 464, 540

External funding: Consortium für elektrochemische Industrie GmbH (München); European COST program (working group D17/010/02).

Cooperations: K. Angermund (Mülheim/Ruhr, DE), G. Fink (Mülheim/Ruhr, DE), A. Fürstner (Mülheim/Ruhr, DE), B. Ganguly (Gujarat, IN), L. Gooßen (Mülheim/Ruhr, DE), V. R. Jensen (Bergen, NO), B. List (Mülheim/Ruhr, DE), M. T. Reetz (Mülheim/Ruhr, DE), P. v. R. Schleyer (Athens, USA), F. Schüth (Mülheim/Ruhr, DE)

2.5.3 Research Area "Semiempirical Methods" (W. Thiel)

Involved: E. Fabiano, T. Keal, A. Koslowski, Y. Lu, M. Ramos da Silva, M. Scholten, R. Steiger

Objective: This long-term project aims at the development of improved semiempirical quantum-chemical methods that can be employed to study ever larger molecules with useful accuracy. This includes the development of more efficient algorithms and computer programs. Our current focus in this area is on electronically excited states.

Results: Over the past years, we have developed semiempirical methods that go beyond the standard MNDO model by including orthogonalization corrections at the NDDO level. This has led to three new approaches labelled OM1, OM2 and OM3 (orthogonalization models 1-3) which offer significant improvements over established MNDO-type methods in several areas, including conformational properties, hydrogen bonds, reaction barriers, and electronically excited states.

Motivated by our participation in SFB 663 (Molecular Response after Electronic Excitation) we have concentrated on semiempirical methods for electronically excited states during the reporting period. Our code provides an implementation of the GUGACI method in a semiempirical context such that general CI expansions can be handled efficiently in a relatively small active space. It allows for CI calculations with about 100000 configurations, for all excitation classes up to full CI. We have now included a semi-analytic GUGACI gradient code for the OMx methods which executes all the expensive steps analytically and evaluates only the integral derivatives numerically, resulting in the same scaling behaviour as in the case of a fully analytic implementation [132].

Our program allows an exploration of excited-state potential energy surfaces at the OMx-GUGACI level, e.g., the location of minima and transition states. Since conical intersections play a central role in photochemistry, we have implemented three methods for finding them: for a test set of 12 well-characterized conical intersections, the Lagrange-Newton method (Yarkony) turns out to be most efficient, closely followed by the projected gradient method (Bearpark-Schlegel), while the penalty function method performs less well [465]. Given the short time scale of photophysical and photochemical events, it will often be of interest to study the dynamics after electronic excitation, and we have therefore also implemented the surface hopping method with

the fewest switches algorithm (Tully). The code offers different schemes for the computation of the required nonadiabatic couplings (analytic vs numerical vs approximate) and ensures proper orbital and state tracking. In three case studies (ethylene, methaniminium ion, and methanimine) the OM2-CI approach yields decay times and dynamics paths similar to high-level ab initio results.

In the context of our general validation project for electronically excited states (see section 2.5.1) we have continued the evaluation of the OMx-CI methods for vertical excitation energies, oscillator strengths, excited-state dipole moments, and excited-state geometries. The results are very satisfactory especially for the OM2-CI approach, in spite of the fact that the OM2 parameterization was performed with regard to ground-state properties only. We shall therefore check in the near future whether further improvements are possible by an OM2 parameterization that employs both ground-state and excited-state reference data.

In collaboration with the Elstner group, the OM2-CI method has been applied to the calculation of absorption shifts in retinal proteins. An initial validation study concludes that the response to external fields generated by the protein environment is not captured properly by a number of commonly used theoretical methods (including CASSCF and TDDFT), but is well represented by ab initio CI approaches (such as SORCI) and by OM2-CI [188]. These methods have therefore been used to study the mechanism of color tuning in the rhodopsin family of proteins, by comparing the optical properties of bacteriorhodopsin (bR) and sensory rhodopsin II (sRII) [287]. The results indicate that several sources contribute to the spectral shift between bR and sRII, the main factors being the counterion region at the extracellular side of retinal and the amino acid composition of the binding pocket. A number of other OM2-CI applications are currently in progress in the SFB project, especially QM/MM (quantum mechanical/molecular mechanical) calculations where OM2-CI serves as QM components. Examples include carotenoids, molecules with intramolecular charge transfer excitations, and proteins with LOV1 domain.

Turning to methodological advances for semiempirical ground-state treatments, a smooth solvation model SCOSMO has been developed in cooperation with the York group. SCOSMO has been implemented for semiempirical methods with an spd-basis [88]. It provides smooth energies and gradients by overcoming discretization errors and thus allows numerically more stable geometry optimizations and reaction path

calculations in solution. The MNDO/d-SCOSMO approach has been applied successfully to study transphosphorylation thio effects in solution [66].

The self-consistent-charge density functional tight binding method (SCC-DFTB) has been evaluated in comparison with MNDO-type and OMx methods using standard test sets [490]. SCC-DFTB is found to be a viable semiempirical method with specific strengths and weaknesses. The overall accuracy of SCC-DFTB and the other semiempirical methods is in the same range, with an overall tendency AM1 < SCC-DFTB < OM2, which may however vary depending on the properties and compounds considered.

Recent advances and applications of semiempirical methods have been reviewed [177].

Publications resulting from this research area: 66, 88, 132, 174, 177, 188, 287, 465, 490, 534

External funding: Fonds der Chemischen Industrie, German Research Council (DFG, SFB 663, project C4), DAAD (stipend to M. Ramos da Silva)

Cooperations: M. Elstner (Paderborn, DE), S. Patchkovskii (Ottawa, CA), K. Schulten (Urbana, USA), D. M. York (Minneapolis, USA)

2.5.4 Research area "Combined Quantum Mechanical/Molecular Mechanical Methods" (W. Thiel)

Involved: M. Altarsha, A. Altun, T. Benighaus, M. Bocola, M. Doerr, E. Fabiano, G. Gillies, Y. Hsiao, J. Kästner, T. Keal, D. Kumar, T. Leyssens, H. Lin, S. Metz, N. Otte, M. Parac, J. C. Schöneboom, H. M. Senn, S. Thiel, T. Tuttle, D. Wang, J. Zheng

Objective: This research focuses on hybrid approaches for large systems where the active center is treated by an appropriate quantum mechanical method, and the environment by a classical force field. It involves considerable method and code development. This approach allows a specific modeling of complex systems such that most of the computational effort is spent on the chemically important part. Current applications primarily address biocatalysis and aim at a better understanding of enzymatic reactions including the role of the protein environment.

Results: Combined quantum mechanical/molecular mechanical (QM/MM) methods have become a popular tool for studying reactions in complex systems such as enzymes. Typical applications make use of density functional theory (DFT) or semiempirical methods as QM component and a standard biomolecular force field (e.g., CHARMM or GROMOS) as MM component. Geometry optimization techniques are commonly employed to determine reaction paths and the relevant minima and transition states. For further improvements beyond this standard QM/MM level, one needs to consider the issues of accuracy and sampling.

Concerning accuracy, we have explored the use of correlated ab initio methods in QM/MM calculations and carried out a case study on the hydroxylation reaction catalyzed by p-hydroxybenzoate hydroxylase (PHBH), in collaboration with the Werner group [246]. This involved B3LYP/GROMOS optimizations of reaction paths and stationary points, single-point ab initio QM/MM energy evaluations using local correlation methods up to the LCCSD(T0) coupled cluster level, and AM1-based determinations of the small zero-point vibrational, thermal, and entropic corrections. Careful validation of the applied local correlation methods with regard to all computational parameters has established that the QM contribution to the barrier should be converged to within 1 kcal/mol, which is supported by the excellent agreement with the available experimental data [246].

The accuracy of the computed QM/MM barriers is mostly determined by the quality of the QM treatment (for sufficiently large QM regions), but the influence of the MM

contributions and the QM/MM interactions may be non-negligible. It is therefore desirable to move from the standard force fields with fixed MM charges to polarizable force fields. In a cooperation with the van Gunsteren group, we have implemented a particular variant of polarizable force fields, the charge-on-spring (COS) model, into our ChemShell QM/MM software [456]. The COS model has the advantage that it is entirely formulated in terms of point charges and is thus consistent with any QM code that can handle external point charges. In the initial QM/MM(COS) test, we studied the identity S_N2 reaction between NH_2Cl and the chloride anion in liquid dimethyl ether (MM solvent described by a polarizable force field). Including solvent polarization raised the barrier by 3 kcal/mol due to a better solvation of the separate reactants [456]. In subsequent work we have shown that semiempirical OM3/MM molecular dynamics (MD) simulations of a QM water solute in liquid MM water also benefit from the use of a polarized MM model for water.

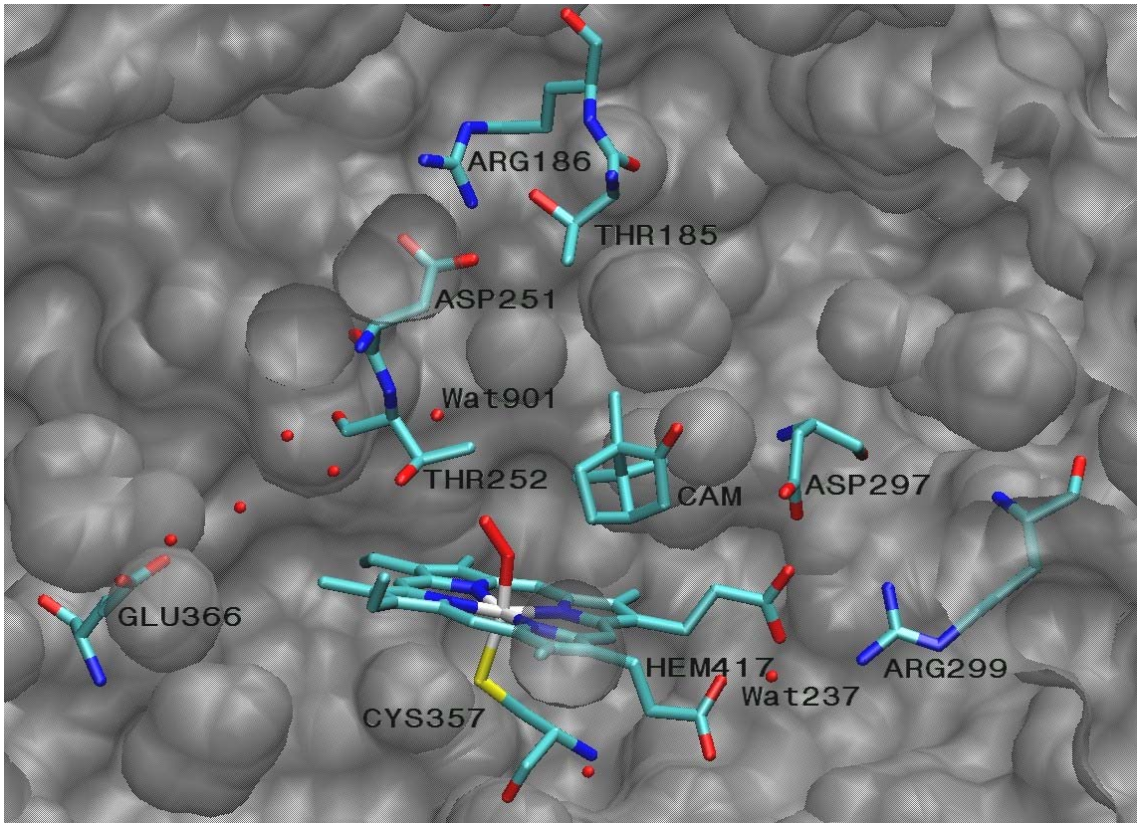
Concerning sampling, the conformational complexity of enzymes calls for extensive explorations of the underlying potential energy surfaces that go beyond performing geometry optimizations for some selected snapshots. When aiming for barriers, this may be achieved by MD simulations for points along a suitable reaction path, with the reaction coordinate being constrained to a fixed value (as in thermodynamic integration) or restrained by a harmonic bias potential (as in umbrella sampling). We have implemented both these techniques into the ChemShell QM/MM package (with proper attention to convergence and error control) and have shown for the test case of PHBH that both yield the same converged free energy barrier [84, 168]. In the course of this work, we noted some drawbacks in the weighted histogram analysis method (WHAM) that is commonly used in umbrella sampling, and we therefore introduced an alternative analysis method called umbrella integration which employs only quantities with easily controllable equilibration and greatly reduces the statistical error compared to WHAM [84]. As an additional advantage, it is possible to derive approximate expressions for the statistical errors which leads to rules for the choice of the bias potential and the sampling parameters [294]. The straightforward application of these MD-based free energy techniques to enzymatic reactions is computationally very demanding and currently only practical at the semiempirical QM/MM level. We have therefore implemented an approximate QM/MM free energy perturbation (FEP) method into ChemShell (originally proposed by Yang) and tested a number of options in the treatment of the link atoms and the QM/MM electrostatic interaction [293]. We find that it is adequate to approximate the QM density by electrostatic-potential-fitted point charges during the FEP-MD runs which makes QM/MM-FEP affordable for any QM

method. Concerning geometry optimization, we have further improved the microiterative approach in the case of electrostatic and polarized embedding [463].

In ongoing methodological work, we develop the tools that are needed for QM/MM investigations of electronically excited states. This includes the adaptation of conical intersection search routines and of surface hopping algorithms (see section 2.5.3) in a QM/MM framework and the definition of protocols for QM/MM calculations of condensed-phase electronic spectra.

Turning to QM/MM applications, cytochrome P450cam remains the enzyme that is studied most extensively in our group, usually at the B3LYP/CHARMM level (often in collaboration with the Shaik group). The earlier work has been summarized in a review [170]. In the reporting period, we have continued to characterize the intermediates in the catalytic cycle including the pentacoordinated ferric and ferrous complexes [5], the ferrous dioxygen and ferric peroxo complexes, the last experimentally accessible intermediate Cpd 0 [406], and the yet unobserved reactive species Cpd I [153]. In addition, we have investigated several key reactions, in particular the proton transfers that generate Cpd 0 and Cpd I [101, 406], and camphor hydroxylation by Cpd I [73, 211, 212, 408, 553]. In all these cases, the comparison between B3LYP/CHARMM calculations for the complete solvated enzyme (around 25000 atoms) and B3LYP calculations for the isolated QM region (typically 40-200 atoms) allows us to assess the role of the protein environment in P450cam. In the following, we outline a few selected results. For Cpd I, the earlier B3LYP/CHARMM prediction of an almost degenerate electronic ground state, with the doublet slightly below the quartet, has been confirmed by ab initio MRCI/CHARMM calculations, and the EPR and Mössbauer parameters have been predicted to facilitate the experimental search for this yet unknown species [153]. Excited states of penta-radical character are found to lie only 12-14 kcal/mol above the tri-radical ground state of Cpd I [153], and subsequent B3LYP calculations have shown that this gap diminishes along the reaction pathway of hydroxylation, due to the cumulative exchange stabilization of the more open-shell species, which might give rise to multi-state reactivity [73]. Contrary to recent suggestions by others, excited states of Cpd I with an Fe(V)-oxo moiety are too high in energy to be mechanistically relevant, and one-electron reduced species (Cpd II) show only sluggish reactivity compared with Cpd I [408]. The hydrogen abstraction from camphor by Cpd I has been investigated systematically to identify the factors that may affect this reaction [211,212]: the claim of remote transition state stabilization (via larger electrostatic interactions between the A-propionate side chain of the heme and the Arg299 residue) has been refuted, the crystallographic water903 has been found to act as a catalyst

(through differential hydrogen bonding), and the role of other environmental residues has been clarified (Asp297, His355). Different mechanisms have been considered for the proton transfer that transforms Cpd 0 into Cpd I [101, 406]. In the enzyme, where the proton can come from Glu366 and Asp251 (see Figure), the best computed mechanism starts with an initial O-O cleavage followed by a concomitant proton and electron transfer in the Asp251 channel to yield Cpd I and water. According to the QM/MM calculations, this is more favorable than the alternative textbook mechanism (protonation followed by O-O cleavage). Hydrogen bond networks play a crucial role in these proton transfer processes, and it is thus not surprising that the latter are affected by mutations of active-site residues which influence these networks. Ongoing QM/MM studies on proton transfer in five such mutants of P450cam reproduce the experimentally observed effects of the mutations and rationalize how they control activity and product distribution (uncoupling vs hydroxylation). As a final remark on P450cam, we mention a methodological study showing that it is feasible to treat the entire catalytic cycle with a common QM/MM setup [553].



Active-site environment in cytochrome P450cam. Red dots indicate crystal water molecules.

In continuation of the collaboration with the Reetz group on directed evolution in lipases, we have used classical MD and QM/MM calculations in order to understand the enantioselectivity in the lipase-catalyzed ester hydrolysis and its optimization through successive mutagenesis. In previous work, we proposed a relay mechanism involving only two out of six mutations in the "best" mutant, to account for the source of the enhanced enantioselectivity. The corresponding double mutant S53P/L162G has meanwhile been prepared and has indeed been found to be even more enantioselective [512]. In control experiments, we have also replaced His83 which is assumed to provide a crucial additional hydrogen bond in the proposed relay mechanism (only in the favored enantiomer). Enantioselectivity is suppressed (as expected) upon substitution by phenylalanine which is of similar size, but largely retained upon substitution by the smaller alanine. MD simulations indicate that a water molecule can assume the role of His83 in the latter case and provide the stabilizing hydrogen bond [512]. Related work on a lipase from *Bacillis subtilis* has demonstrated that the combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution: both QM/MM-based calculations and molecular biology experiments find His76 as a residue that strongly affects catalytic activity [59]. Further unpublished work on this lipase has identified a number of possible binding modes for the substrate 1-(2-naphthyl)-ethyl-acetate, which provide starting points for QM/MM calculations (DFT/CHARMM geometry optimizations and SCC-DFTB/CHARMM umbrella sampling MD runs) that give qualitative insight into the origin of the observed enantioselectivity.

A number of other biomolecular systems have been studied at the QM/MM level during the reporting period, often motivated by cooperations with other groups. We list ten such applications.

(a) Enzymatic C-F bond formation by a fluorination enzyme has been found to follow an S_N2 reaction mechanism, while the alternative elimination-addition mechanism can be excluded as well as the route via sulfur ylides [167].
(b) The enzymatic activity of 4-oxalocrotonate tautomerase and seven synthetic mutant analogues has been analyzed by classical MD simulations and QM/MM calculations. Replacing arginine (Arg) by citrulline (Cit) has only minor effects in the case of the Arg61Cit mutation (because of the flexibility of this residue), while the Arg11Cit and Arg39Cit mutations disrupt the binding site and strongly impair the catalytic activity [387].
(c) An in-depth QM/MM study of the wild-type 4-oxalocrotonate tautomerase has led us to propose a new model for the reactive substrate orientation which combines

favorable substrate binding geometries with reasonable barriers and is consistent with the experimental evidence from mutation studies concerning the catalytic ability of specific residues in the binding site [539].

(d) In cooperation with the Fürstner group, we have studied the binding of latrunculin and its synthetic analogues to actin and rationalized the differences between these compounds in terms of the possible hydrogen bond networks [446].

(e) In collaboration with the Engels group, we have examined the importance of the active site histidine residue for the activity of epoxide- and aziridine-based inhibitors of cystein proteases and the particular role of a water molecule that is required to establish an efficient relay system for proton transfer [482].

(f) To characterize the active site of vanadium chloroperoxidase, ten models with different protonation patterns have been optimized at the DFT/CHARMM level and then subjected to ^{51}V NMR computations (see section 2.5.6), which allowed us to identify the two most likely candidate structures on the basis of the computed anisotropic NMR chemical shifts [547].

(g) In cooperation with the Cremer group, we have studied the Bergman cyclization of dynemicin A in a DNA environment. According to the DFT/CHARMM calculations, acyclic enediynes can undergo this reaction in the minor groove with a much smaller barrier than in the gas phase, while the DNA environment has less effect in the case of cyclic enediynes [538].

(h) In collaboration with the Cao group, the deprotonation mechanism in the *Escherichia coli* ammonium transporter AmtB has been investigated using QM, QM/MM, and QM/MM MD techniques. We find a stepwise rather than a concerted mechanism, with the carboxylate group of Asp160 acting as the proton acceptor, which clarifies why mutations of the preserved residue Asp160 reduce or disable the activity of AmtB [426].

(i) In cooperation with the Hildebrandt group, the Raman spectra of the phycocyanobilin chromophore in α-C-phycocyanin have been computed at the B3LYP/CHARMM level. This provides a substantially improved description of the experimental resonance Raman spectra of the protein-bound cofactor (compared with gas-phase QM model calculations) and allows an assessment of the protein-cofactor interactions [483].

(j) Ongoing unpublished QM/MM work addresses the mechanism of oxidation reactions catalyzed by molybdenum-containing enzymes such as aldehyde oxidoreductase and xanthine oxidase.

The ChemShell software that has been used in all these applications is available under a license agreement (see www.chemshell.org). The QM/MM methodology and QM/MM applications to biological systems have been reviewed [526, 527].

Publications resulting from this research area: 5, 33, 59, 73, 84, 101, 153, 167, 168, 170, 211, 212, 246, 293, 294, 387, 406, 408, 426, 446, 456, 463, 482, 483, 512, 526, 527, 538, 539, 547, 553

External funding: Deutsch-Israelische Projektkooperation (project DIP-F.6.2), VolkswagenStiftung (project I/80454), German Research Council (DFG, SFB 663, project C4), Triple-M Research Initiative of the Max Planck Society, Fonds der chemischen Industrie (stipend to S. Metz), Alexander-von-Humboldt Foundation (fellowship to T. Leyssens)

Cooperations: M. Bühl (Mülheim/Ruhr, DE), Z. Cao (Xiamen, CN), D. Cremer (Göteborg, SE), B. Engels (Würzburg, DE), R. A. Friesner (New York, USA), J. N. Harvey (Bristol, UK), P. Hildebrandt (Berlin, DE), K.-E. Jaeger (Jülich, DE), E. Keinan (Haifa, IL), C. Marian (Düsseldorf, DE), A. J. Mulholland (Bristol, UK), F. Neese (Bonn, DE), D. O'Hagan (St. Andrews, UK), M. T. Reetz (Mülheim/Ruhr, DE), S. Shaik (Jerusalem, IL), P. Sherwood (Daresbury Laboratory, UK), W. F. van Gunsteren (Zürich, CH), H.-J. Werner (Stuttgart, DE), Y. D. Wu (Hong Kong, HK)