Charles University in Prague
Faculty of Science
Department of Physical and Macromolecular Chemistry

Doctoral thesis

Non-covalent interactions in fundamental biological processes

Mgr. Vojtěch Klusák

Supervisor:
Mgr. Lubomír Rulíšek, CSc.

Institute of Organic Chemistry and Biochemistry AS CR
Center for Biomolecules and Complex Molecular Systems

Universita Karlova v Praze

Přírodovědecká fakulta

Katedra fyzikální a makromolekulární chemie

Disertační práce

Nekovalentní interakce v základních biologických procesech

Mgr. Vojtěch Klusák

Školitel:

Mgr. Lubomír Rulíšek, CSc.

Ústav organické chemie a biochemie AV ČR

Centrum biomolekul a komplexních molekulárních systémů

I hereby declare that I have written the presented thesis by myself, all the literature is properly cited and that I have not been yet awarded any other academic degree or diploma for this thesis or its substantial part.

Prague, 6th April 2010 Vojtěch Klusák

Acknowledgement

I would like to thank my thesis supervisor Dr. Lubomír Rulíšek for his sensitive guidance and support. Also, I would like to thank Dr. Jiří Vondrášek and Dr. Zdeněk Havlas for their support in the initial stages of my research. I am also grateful for the help and inspiration I obtained from Prof. Pavel Hobza and Doc. Jan Konvalinka. I thank my closest laboratory friends and colleagues, Dr. Martin Lepšík, Dr. Martin Švec, Mgr. Jakub Chalupský, RNDr. Martin Srnec, Mgr. Václav Navrátil, Dr. Mojmír Kývala, Dr. Daniel Svozil, Dr. Lada Bendová and RNDr. Ivana Beššeová for encouragement and many inspiring discussions.

Last but not least, I appreciate friendship of all the members of the laboratories of biochemistry and theoretical chemistry. My great thanks go to my family for continuous support.

Contents

1. Introduction

Recent advances in computational technology and methodology have allowed us to apply methods of quantum chemistry to answer questions of biological relevance. Nowadays, high-level quantum chemical methods describe molecular systems with a satisfactory degree of accuracy so that even weak, non-covalent interactions in biomolecules and their complexes can be conveniently studied.

The presented thesis is focused on accurate description of molecular properties, especially non-covalent interactions and interpretation of their role in various biological processes.

The first part of the thesis is focused on the stabilization of biomolecular structures, exemplified in a receptor-ligand complex and a network of interactions in the core of the protein. This includes the discussion of general phenomena resulting from the studied and analyzed characteristics, as well as of some pitfalls related to their theoretical treatment.

In the second part of the thesis the key structures (reaction intermediates and transition states) of a prominent enzymatic reaction (peptide hydrolysis) in glutamate carboxypeptidase II (GCPII) and the interactions important for their stabilization have been described. It allowed us to discuss and suggest mutation experiments carried out in our collaborator's laboratory and relate the calculated data to the reactions of other hydrolytic enzymes. Finally, important theoretical aspects of the quantum chemical treatment are discussed as well.

1.1. Is function of a protein coded in its structure?

Understanding the inter- and intra-molecular interactions is the key for our insight into the physicochemical properties of biomolecular systems which, in turn, maintain and govern virtually all the processes in biology. Attempts to draw the structure-function relationship at the atomistic or electronic level bring us quite often beyond experimental resolution and capabilities. Modern tools of computational chemistry enable us to focus on the details of the studied process, gather additional (often complementary) information and ascribe particular structural features to measurable quantities. [Nelson D. L. 2004; Müller-Dethlefs, K. 2000; Hobza P. 2009]

The atoms of molecules are maintained together by covalent interactions. The energies of covalent bonds are in the order of ~400 kJ/mol. Under normal conditions, in the processes related to biomolecular reorganization or assembly (folding or unfolding of protein or nucleic acid complexes or membrane formation), the covalent bonds do not break. Such processes are controlled by non-covalent interactions. Despite individual non-covalent interactions are relatively weak, they are

fairly numerous. Hence the total stabilization energy resulting from these interactions can be large. Thus, a large number of small non-covalent forces govern the equilibria between folding or unfolding of biomolecules, ligand binding or release, quaternary protein structure assembly or disassembly, etc. Most of these equilibria are subtle and bio-systems are held in delicate balance between large number of countervailing forces and it is the small difference between these large numbers (repulsive and attractive interactions) that determines direction of the specific process. [Daniel R. M. 2000; Cooper A. 1999] It can be noted that this delicate balance between numerous, complex, intertwined forces, and hence the structure, can be specifically changed by small changes in pH, temperature or any other condition, including interaction with other proteins, small molecules or ions. The prerequisite for gaining the control of protein secondary and tertiary structure is a deep understanding of how the non-covalent interactions provide both stabilization and specificity. [Tatko C.D. 2004]

Besides governing the shape and interactions of biomolecules and hence determining the structures of cells and organisms, the weak non-covalent interactions are crucial in the chemical transformations of molecules - enzymatic reactions. Out of a tremendous number of enzyme catalyzed processes we will further focus on hydrolytic reactions. In organisms, these need to be performed efficiently and selectively to sustain the biological functions. In many cases the evolution of the enzyme functionality led to metalloenzymes, often putting several metal ions close to each other in the active site of the particular enzyme. These co-catalytic ions work together as a single unit. The catalytic site and the surrounding enzyme macromolecule can be viewed as a device fine tuned for its reaction rate and its temporal and spatial position in the body.

There is an enormous amount of experimental data available concerning the structure and function of enzymes. However, a mechanistic description of their reaction mechanism is limited by the fact that the enzymatic reactions are fast enough to prevent most of the experimental techniques from following the individual steps of the reactions. Even the structures of complexes of the enzymes with inhibitors, designed to mimic the reaction transition state or intermediate, must be interpreted with care, because the inhibitors are often tightly bound to the residues in the active site, which might not be the case for the native substrate. Hence its binding mode might be different. [Weston J. 2005]

Therefore, one must often resort to computational modeling to gain some insight into the reaction mechanism, to verify experimental hypothesis and to merge the experimental information into a complex and comprehensive picture.

The most important source of data used for the structure-function relationship is the collection of three dimensional structures of bio-macromolecules, which are deposited in the Protein Data Bank (PDB). PDB is nowadays the largest repository of information about the 3D structures of large biological molecules worldwide. The database was founded in 1971 at Brookhaven National Laboratories as database for structures obtained by X-ray crystallography. Later, the structures solved by means of nuclear magnetic resonance (NMR) have been deposited as well. Nowadays, it is also the neutron and electron diffraction, and the cryo-electron spectroscopy that contribute to the fast growth of the database. [Berman H.M. 2000; Berman H.M. 2003]

1.2. Intermolecular and intra-molecular interactions in biology

As mentioned above, the structure of a protein (bio-polymer) is a result of an equilibrium; equilibrium of interactions inside the molecule, interactions between the molecule and the solvent and the interactions within the solvent. The fold of the protein is primarily determined by the primary amino acid sequence and is maintained in its native environment, usually polar aqueous one. Due to its enormous complexity, the process of protein folding is not fully understood and has only very limited predictability. It has been studied theoretically and experimentally by several scientific disciplines without succeeding to find a robust universal procedure how to read the primary sequence and translate it into the secondary and tertiary structure. [Dill A. K. 2008]

The most important interaction is assumed to be the hydrogen bond. The understanding of its role has been evolving through the years. [Desiraju G. R. 2001; Nishio M. 1998; Steiner T. 2001; Müller-Dethlefs, K. 2000; Hobza P. 1980; Hobza P. 2009] Besides hydrogen bonds between "conventional" O-H, N-H and S-H groups as donors and N, O, S atoms as acceptors,[Desiraju G. R. 2001] it has been shown that there are other possibilities represented by "non-conventional" functionalities: X-H … π interactions (where X stands for: O, N, S, C)[Steiner T. 2001; Brandl M. 2001], C-H … Y interactions (Y: O, N) [Derewenda Z. S. 1995] notably the C_{α}-H … O=C interaction [Vargas R. 2000], and the stacking interactions (the stabilization of the parallel aromatic functional groups) [Burley S.K. 1985; Baldwin R. L. 2002; Meyer, E.A. 2003]. Therefore, several amino acids have been recognized as having an indispensable role in the protein structure stabilization and formation. Their assumed role changed from the simple space filling within the hydrophobic protein core to formation of the enthalpically-driven, oriented non-covalent attractive interactions between these residues (side chains).

The effect of the aqueous environment is denoted the hydrophobic effect. [Martin T. 1998] The manifestation of hydrophobic effect is exemplified by the immiscibility of oil and water which

can be considered as its definition.[Southall N.T. 2002] It was the mixtures of oil and water and the separation coefficients of several substances between these two liquid phases on which it was first described [Meyer H. 1899; Overton E. 1901]. For biological substances this behavior was first described for case of lipid molecules. Based on careful experiments with surface concentration Goerter and Grendel suggested that lipid molecules can be organized into bilayers. [Goerter E. 1925] There is certain similarity observable already in the first structures of proteins obtained from X-ray crystallographic measurements in 1958. [Kendrew J.C. 1958] The surface of proteins is formed predominantly by the polar residues whereas the non-polar residues prevail in the interior.

In 1954 Walter Kauzmann used term hydrophobic interaction to describe tendency of oil liquids to associate in aqueous solutions. [Kauzmann W. 1954] The *hydrophobic effect* refers to unfavorable interaction of non-polar substance (particle, moiety) and water. The *hydrophobic interaction* then refers to the interaction of two (or more) non-polar substances (particles, moieties) in water. In the first approximation the two particles are drawn together by the tendency of water to minimize the contact surface and their interaction energy is considered as negligible.

In the case of small hydrophobic solute (e.g. methane molecule) the solvent (water) hydrogen bonds are not broken (i.e., the mean number of the bonds is not reduced by the inclusion of the solute). The dominant term in the solvation energy is the reorganization of the solvent hydrogen bonds. In the case of large solutes it is impossible for the adjacent water molecules to maintain a complete hydrogen-bonding network. As a result water tends to move away from the large solute and forms an interface around it, similar to the one between liquid and vapor. The lipophilic solutes represent regions of space where water hydrogen bonding cannot occur, resembling thus cavities in the water. Although the interactions between solute particles are too weak to affect the existence of interfaces in water, they do affect the shape of the interface. [Stillinger F.H. 1973; Chandler D. 2005]

1.3. Enzyme kinetics and modeling approach

Enzymes are biocatalysts mediating chemical transformations in living organisms under mild conditions, at specific spatial and temporal location in the body, and on a specific substrate. These two features, reaction rate enhancement and high specificity, make enzymes essential elements in all living organisms. It is of great importance to understand how enzymes work in detail. As has been mentioned, the rate of enzyme-catalyzed reactions, the complexity of the enzymatic structure and of the reaction mechanisms often leaves us with the modern computational chemistry as the only tools. Nowadays, quantum chemical (*ab initio* or DFT) methods can

conveniently deal with few hundreds of atoms. However, they are still far from modeling the whole protein, typically consisting of several thousands of atoms. Fortunately, enzymatic reactions are mostly confined to the active site which is usually formed by several well-defined functional groups (amino acid side chains, cofactors, metal ions, etc.) and the effect of the surrounding is assumed to be less important. In this thesis the approach combining quantum chemical (QM) description of the active site with more approximate description of the surrounding - molecular mechanical (MM) description, or continuum solvation model - has been adopted. Such methodology has been found appropriate and successfully applied to a large number of enzymes. [Warshel A. 2006; Senn H. M. 2007; Siegbahn P. E. M. 2009]

The true reason behind the catalytic power of enzymes is still a vigorously debated issue. Historically, the first hypothesis is the familiar "lock and key" model, analogy first postulated in 1894 by Hermann Emil Fischer [Fischer H. E. 1894; Lichtenthaler, F. W. 1994], which proposed that the enzyme and the substrate posses complementary geometric shapes that fit together. While this model explains receptor-ligand specificity it fails to explain the stabilization of the substrate transition state by the enzyme. Nowadays accepted "induced fit" model [Jorgensen W.L. 1991] allows for the flexibility of substrate and enzyme and suggests that the substrate binding is process involving continuous reshaping of both structures. The key catalytic contribution of the enzyme is that the active site provides preorganized polar environment that stabilizes the transition state much more than the corresponding environment in the bulk solvent [Warshel A. 2006]. A wide array of alternative proposals, and key factors have also been suggested by other authors: ground state destabilization, dynamical effects, near attack conformations, entropy guidance, desolvation effects, low-barrier hydrogen bond, or covalent catalysis. [For the citations see Warshel A. 2006] We believe that there is no simple general explanation of the catalytic power of enzymes. It is necessary to study each enzyme individually, using a clear energy-based theoretical and/or experimental analysis.

An ultimate goal of mechanistic investigations of the enzymatic catalysis is to find out how enzymes interact with their substrates in the transition state complex. In order to examine the origin of the enzymatic power and to discriminate between different mechanistic proposals and reaction pathways, it has been shown that quantum chemical methods are useful tool. [Siegbahn P. E.M.2009; Warshel A.2006] By calculating the potential energy surface (activation barriers and energies of intermediates), one can characterize the preferred catalytic pathway and analyze the role of various residues in the active site and their contribution to the catalytic power.

1.3.1. An overview approaches in computational modeling of enzymes

Construction of the active site. For understanding and modeling of the enzymatic reaction mechanisms the structural information about the active site is crucial. It enables to identify the important residues, water molecules, cofactors, and some other essential chemical compounds forming the catalytic center. Localization of the enzyme catalysis in the active site makes it possible to construct its quantum chemical model reflecting all chemically important aspects in the real process. In order to reduce size of the QM model, the residues are usually truncated so that only their side chains or important backbone peptide chains are included. Finally, a structure of the substrate is added to the active site model, either using a chemical intuition or utilizing the crystal structure of the enzyme complexed with the substrate analog (e.g., inhibitor). In case of polar character of the substrate functional groups and high similarity of the native substrate and its analog, the principal orientation or the specific binding mode of the substrate can be estimated with a high degree of reliability. If the substrate position is ambiguous, several alternatives need to be investigated in parallel in order to identify the most probable one. The most favorable situation occurs if the crystal structure of the inactive enzyme (enzyme with crucial active site residue mutated) with native substrate can be used.

Effect of the surroundings. The surroundings of the active site that is not included in the quantum chemical model interacts with it via electrostatic polarization effects and anchors the active site residues by covalent bonds linking the site to the rest of the protein. At least two approaches describing the chemistry in the active site and including effects of protein surroundings can be distinguished: (i) the cluster model which neglects the embedding completely, or uses polarized continuum model to mimic the protein dielectric environment and (ii) QM/MM model combining quantum mechanics (QM) description of the active site with classical molecular mechanics (MM) description of the embedding.

Cluster model. In case of cluster model the surrounding can be either neglected completely and substituted by vacuum, or approximated by a homogeneous polarizable medium. The solvation effects can be either calculated via the single point calculations using the *in vacuo* optimized geometries or incorporated in the calculation in the course of the geometry optimization. The polarizable medium is usually modeled by dielectric cavity techniques. [Cossi M. 2003; Klamt A. 1993] In addition, the neglected surrounding also imposes steric effect on the active site residues. In order to mimic these restraints, certain atoms in the model, typically the atoms, at which the truncation is made, are kept fixed during the calculations. This approach can ensure the structural

integrity of the model and at the same time brings a certain degree of flexibility. In general the model becomes more accurate with the increase of the model size. This modeling scheme has been proven to be accurate enough for distinguishing and ranking different mechanistic proposals and has been successfully applied to a wide variety of enzymes. [Siegbahn P.E.M. 2009]

QM/MM model

More realistic model chemistry is represented by the QM/MM scheme. In one possible approach, the protein and solvent are split into three subsystems: the QM region (system 1) contains the most relevant atoms for the chemical process under consideration (typically the active site and its vicinity) and is relaxed by QM/MM forces. System 2 consists of selected residues (e.g. within some distance from system 1) and is relaxed by MM minimization. Finally, system 3 contains the remaining part of the protein and surrounding solvent molecules and is kept fixed at the original (crystallographic) coordinates. For the more detailed description of this scheme see the chapter 2.7.

2. Theoretical Background and Methods

In the following, basic principles of the most important experimental method for the determination of biomolecular structures - the X-ray crystallography - and the fundamentals of enzyme kinetics are briefly described.

2.1. X-ray crystallographic structure determination and interpretation.

Besides the three dimensional structure of the molecule itself, there are several other parameters significant for the structural analysis based on X-ray diffraction. The structure is characterized by the resolution [Å], dimensionless (reliability) factors R and R_{free}. Moreover, each atom is characterized by its atomic displacement parameter (ADP, formerly denoted the B-factor) [$Å^2$].

The resolution is the shortest distance between the crystal lattice planes which yielded detectable diffraction during the data collection. It mostly correlates with the quality of the crystal, but can be also influenced by the quality of the used equipment (e.g., the light intensity).

The R factor can be expressed as follows:

$$R = \frac{\sum_{hkl} \| F_{obs} | - k | F_{calc} \|}{\sum_{hkl} | F_{obs} |}$$

R_{free} differs from R in that it refers to the fraction of the experimental date randomly chosen and excluded from the structure determination process.

The ADP (B-factor, temperature factor) is quantity proportional to mean quadratic deviation of an atomic vibration $\langle u \rangle^2$,

$B = 8\pi \times \langle u \rangle^2$.

The ADP reflects dynamic fluctuations caused by vibrations whose magnitude is dependent on temperature. Besides the dynamic perturbation there is also a static perturbation present in the crystal. The molecules or their parts in different crystal cells do not occupy the identical position. These two effects cannot be distinguished unless a comparison of data sets collected at different temperatures is available.

All these factors including the raw data from the diffraction measurement are nowadays requested during the deposition of a new structure.

2.2. Thermodynamic description of system behavior. Gibbs energy and partition function

The measure of the spontaneity of chemical and biological processes is the Gibbs free energy (assuming the constant pressure and temperature boundary conditions). It is a thermodynamic state function.

However, it is only the potential energy which is the primary result of the quantum mechanical calculation. The effect of translational energy, connected with non-zero thermodynamic temperature is missing. Owing to the statistical thermodynamics, it is possible to connect "the worlds" of thermodynamics, quantum mechanics and statistics. This enables us to derive the macroscopic thermodynamic quantities (functions) from the intrinsic molecular properties. The key function (quantity) of statistical thermodynamics is the partition function Q. The partition function describes the statistical properties of a system and can be decomposed into product of partial partition functions, $Q = q_{trans} q_{rot} q_{vib} q_{vib}$. If the partition function is known, it can be used to calculate the thermodynamic state functions. Detailed introduction to statistical thermodynamics can be found in [Chandler D. 1987; McQuarrie D. 2000].

The analytical expression of the partition functions can be derived for simple models only. For molecular partition function several approximations must be introduced. Assuming the rigid rotor approximation (no distortion due to rotation), the harmonic approximation (vibrations can be described by harmonic oscillator models) and an ideal gas behavior of the particle, it is possible to express the Gibbs energy as

$$G = E_{el} + E_{ZPVE} - RTln(q_{trans}q_{rot}q_{vib})$$

where E_{el} is the electronic energy, E_{ZPVE} is the zero point vibration energy, q_{trans}, q_{rot}, and q_{vib} stand for translation, rotation and vibration partition functions, respectively.

The main source of errors is the description of those vibrations in which the anharmonicity can be significant. The complication mostly comes from the low energy vibration modes which significantly contribute to the Gibbs energy (through the entropic component). In case of large systems the situation is also complicated by conformational variability which can be seen as low frequency vibrational motion.

2.3. Enzyme kinetics

In this chapter we introduce two important experimental quantities, k_{cat} and K_m. Let's assume a simple type of enzymatic reaction

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} EP \xrightarrow{k_3} E + P$$

where E stands for free enzyme, S for free substrate, ES is the enzyme-substrate complex, EP enzyme-product complex, P is the free product and k_1, k_{-1}, k_2, k_3 are the rate constants corresponding to the individual steps.

In the first approximation the reverse reaction ($EP \rightarrow ES$) is assumed to be very slow and steady state approximation ([ES] = *const*) is adopted:

$$k_1[E][S] = (k_{-1} + k_2)[ES]$$

Finally, we neglect the difference between the actual and total concentration of substrate. Based on these approximations we arrive to the Michaelis-Menten equation which describes the rate of an enzyme catalyzed reaction on a single substrate:

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

The K_m is the Michaelis constant, defined as $(k_{-1} + k_2)/k_1$. The reaction rate reaches its maximum (V_{max}) when the enzyme is fully saturated with substrate, $V_{max} = k_2 [E_t]$ (E_t stands for the total concentration of the enzyme, which is the sum of the free molecules and those complexed with substrate). When the substrate concentration is equal to K_m, the reaction rate is half of the maximum *(V_{max})*. In the case, when the second step (k_2) is the rate limiting, ($k_{-1} >> k_2$), then the K_m can be assumed to have direct connection to the enzyme substrate affinity (expression $k_{-1}+k_2/k_1$ can be simplified as k_{-1}/k_1, which corresponds to the equilibrium constant for the ES complex formation.

In practice the k_{cat} value is often used (instead of k_2) which can be interpreted as the rate constant for the rate determining step (RDS). If the slowest step (RDS) is the product dissociation, then k_{cat} corresponds to k_3 of the model reaction above. Experimentally, k_{cat} is measured using the following equation [Copeland R. A. 2000]:

$$k_{cat} = \frac{V_{max}}{[E_t]}.$$

2.4. Theory of the activated complex and reaction rates

The effect of catalytic acceleration of an enzymatic reaction is a consequence of lowering the activation energy (barrier) that is the $\Delta G^{\ddagger}$ value. This is accomplished by stabilization of the highest transition state on the reaction pathway.

The interpretation of reaction rates and reactivity of chemical processes is based on the theory of the activated complex (or transition state theory). In this theory, developed by Eyring, Polanyi and Evans [Eyring, H.1931; Eyring, H. 1935; Evans M. 1935] it is postulated that: (i) there exists a potential energy surface (PES) dividing the space into a reactant region and product region, (ii) there is a transition state between the two regions which is characterized by a single imaginary vibration frequency of sufficiently low magnitude (iii) the distribution of the energies in the reactant and transition state follows the Boltzmann distribution law, (iv) the nuclear motion along the reaction coordinate is electronically adiabatic and can be described by classical mechanics. The reaction rate constant (k) of the reaction

$$A + B \rightleftharpoons TS \longrightarrow P$$

can be expressed by the equation

$$k = \frac{k_B T}{h} \exp\left[-\frac{\Delta G^{\ddagger}}{RT}\right]$$

where k_B is the Boltzmann constant, h is the Planck's constant, the pre-exponential factor $k_B T/h$ at

room temperature equals to 6.248 10^{12} s^{-1}, $\Delta G^{\ddagger}$ is the Gibbs energy difference between transition state and reactants, R is the universal gas constant (8.314 J K^{-1} mol^{-1}) and T is the temperature. From this equation, we can see that at room temperature (298.15 K) the rate of 1 s^{-1} corresponds to the barrier of 17.4 kcal/mol, and an increase of $\Delta G^{\ddagger}$ by 1.4 kcal/mol results in the decrease by one order of magnitude in the reaction rate.

The Gibbs free energy barrier is used for the calculation of the reaction rate according to the above equation. In general, free energy can be decomposed into the enthalpic and entropic components,

$$\Delta G = \Delta H - T\Delta S.$$

Sometimes, free energy barrier is approximated by the enthalpic term only (ΔH). This is a plausible approximation in case when the entropic effects (which are very difficult to calculate in case of complex enzymatic reaction) are small.

The enthalpic term in the above equation can be further expressed as

$$\Delta H = \Delta U + p\,\Delta V,$$

where p is the pressure, V is the volume and U stands for internal energy. In the ideal gas approximation the second term is equal to ΔnRT, where Δn is the difference in the moles (e.g., between the reactants and products). In the case of condensed phase, the volume changes are considered as negligible, and hence it is plausible to assume $\Delta H = \Delta U$.

In the case of an enzymatic catalysis the substrate binds in the active site in the conformation suited for the reaction. Therefore, it can be assumed that the change of the conformational entropy is rather small [Nelson D. L. 2004, Liao R. Z. 2009] whereas the change in vibrational entropy can be estimated from calculations of the corresponding small model system in vacuum.

On the other hand, in case of ligand-receptor binding the entropic contribution can be significant or even dominating the overall affinity. [Perozzo R. 2000; Kožíšek M. 2004] Because the energy levels of different conformations are close together, the events such as ligand-binding may change population of different conformation states and hence change the entropy. It can be, however, neglected when comparison of similar ligands is made.

2.5. Quantum chemical calculations, wavefunction methods and DFT methods.

Quantum chemistry is a discipline in which the fundamental principles of physics - quantum mechanics and quantum field theory - are applied for solving chemical problems. The behavior of all chemical species can, in principle, be obtained mathematically by solving the time-dependent Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t)=\hat{H}\Psi(x,t)$$

which describes how the quantum state of a physical system evolves in time. If the Hamiltonian operator does not explicitly contain the time variable (stationary state), the Schrödinger equation in the time-independent form,

$$\hat{H}\,\Psi = E\,\Psi\,,$$

can be used. Furthermore, we can decouple this equation into electronic and nuclear components, because of the high ratio between the masses of nuclei and electrons which is known as the Born-Oppenheimer approximation. The Schrödinger equation can be then solved separately for the electrons with the positions of nuclei used as the parameters. Still, the two-body electron-electron potential energy term makes the exact, analytical solution to Schrödinger equation so far unreachable and it must be solved approximately by different methods.

Hartree-Fock method (HF): The approximation behind the HF method is the representation of the *N*-electron wave function by the antisymmetrized product of one-electron wave functions (atomic or molecular orbitals), Slater determinant. The orbitals are constructed from a linear combination of atomic basis functions. HF method is physically equivalent to the model of independent electrons, where the many body problem of solving the electron-electron interaction in an *N*-electron system is reduced to a set of *N* one-electron problems (i.e., the specific electron interacts with the mean field Hartree-Fock potential). The formulation of the problem leads to the set of *N* Hartree-Fock equations:

$$\hat{F}\varphi_i(i)=\varepsilon_i\varphi_i(i)$$

These equations form a set of generalized eigenvalue equations. The HF equations can be solved iteratively by the procedure called the self-consistent field (SCF) method.

The advantage of the HF method is that it is a variational method which means that the energy corresponding to the approximate solution is greater than the energy corresponding to the exact solution. The disadvantage of the HF method, which is, in principle, one-electron approximation, is the lack of the electron-electron correlation. This error increases with the

increasing number of electrons and is taken into account in the so-called post-SCF methods described below.

The advantage of the *Møller-Plesset perturbation method* is that it accounts for rather large portion of the electron-electron correlation energy. The wave function obtained from the HF calculation is taken as the zero-order approximation to the exact wave function and the HF energy is the first-order approximation to the exact energy. From this solution the higher order perturbations, both in energy and the wave function (MP2, MP3 and higher) are derived using Rayleigh-Schrödinger perturbation theory. The disadvantage is that the method is not variational, thus the resulting energies might oscillate around the exact value with the increasing order of perturbation.

The *coupled cluster (CC)* method also describes the electron correlation. In this case a cluster expansion of the wavefunction is used. It is a sum of individual contributions of singly, doubly, and higher excited determinants. Thus, for example, the CCSDT method determines iteratively (up to the infinitive perturbation order), all single, double, triple, and also part of higher electron excitations. The CCSDT energies are very close to the energies obtained from the full configuration interaction calculations and are thus approaching the exact non-relativistic energies. The CCSDT (and higher) method is, however, prohibitively expensive which leads to the formulation of popular CCSD(T) method in which the triple excitations are determined up to the fourth order of the Møller-Plesset perturbation theory. It has been shown that the enormous savings of the computational expenses are accompanied by only a small loss in the accuracy and therefore, the CCSD(T) method is often used for benchmark calculations. It is, of courses, still very expensive method and only systems containing up to ~30 atoms can be studied using larger basis sets (which is an imperative for obtaining the accurate values).

Both methods, MP2 and CCSD(T) do cover to some extent the electron correlation which is needed for proper description of the non-covalent interactions. MP2 method is computationally less demanding, but it considerably overestimates the interaction energy for some weakly bound complexes. For complexes with delocalized electrons the overestimation is greater than for H-bonded species. [Hobza P. 1996; Jaffe R.L. 1996; Tsuzuki S. 1999; Tsuzuki S. 2000a; Tsuzuki S. 2000b]. As mentioned above, the performance of the CCSD(T) method in this regard is satisfactory.

While the electronic wavefunction depends on $4N$ variables ($3N$ spatial and N spin coordinates, where N is the number of electrons), in *density function theory (DFT)*, the electron density $\rho[\mathbf{r}]$ is used as the basic variable. It depends only on 3 spatial coordinates regardless of the number of electrons. [Parr R. G. 1994, Koch W. 2001]

The theoretical background of DFT comprises two Hohenberg-Kohn theorems. The first one states that the nondegenerate ground state electron density (ρ_0) of a many electron system uniquely determines the external potential, V_{ext} (which is, in the absence of external fields, represented by the set of the positions and the atomic numbers of the nuclei). The second theorem defines the functional of energy of the system and shows, that in the ground state it obeys the variational theorem [Hohenberg P. 1964].

By applying the model of non-interacting electrons in an effective potential it is possible to derive the Kohn-Sham equations [Kohn 1965], which are analogous to the HF equations. Using the atomic orbital basis sets leads to the formally same solution as in the case of HF method. The Kohn-Sham equations are also solved iteratively.

The popularity of DFT for chemistry started with the introduction of the Generalized Gradient Approximation (GGA) which addresses the inhomogeneity of the electron density by considering both the density at a given point and its first derivative. An example of GGA based method is the Perdew, Burke and Ernzerhof (PBE) functional. Some of the modern functionals are the meta-GGA functionals, which consider electron density, its first and second derivative and the density of the kinetic energy. An example of such functional is the Tao-Perdew-Staroverov-Scuseria (TPSS) functional. [Tao J. M. 2003].

The last category – hybrid functionals - mix a fraction of the exact exchange from the HF method into the DFT exchange-correlation functionals. The most successful hybrid functional in practice is B3LYP.

The advantage of DFT methods is their efficiency whereas one of few disadvantages is poor description of the interactions stabilized by dispersion forces (London dispersion energy). This disadvantage can be tackled by adding *a posteriori* an empirical term, parameterized to compensate the error [Grimme, S. 2004; Grimme, S. 2006] or is accounted for in some recent functionals [Zhao Y. 2008].

DFT calculations can be further expedited by the introduction of the resolution of identity (RI, sometimes denoted the density fitting, DF) approximation which accounts for the Coulomb term in the DFT functional. It is done by using an auxiliary basis set for expression of pair products of atomic orbital basis functions [Feyereisen, M.W. 1993; Eichkorn K. 1995; Eichkorn K. 1997]. This approximation reduces four-centre two-electron integrals into three-centre two-electron integrals. It results in the reduction of computational expenses by more than one order of magnitude

and only a small loss in the accuracy in computed data. Applicability of the method for the non-covalent interactions description has been also extensively studied. The results indicate that the RI approximation is fully comparable to the original (non-RI) solution. [Jurečka P. 2001; Hobza P. 2002; Jurečka P. 2003; Šponer J. 2004]

In carrying out the quantum chemical calculations, it is very important to select the appropriate basis set. The optimal selection (which is always a compromise between the expenses and accuracy) has been discussed in thousands of computational articles. [e.g., Martin J.M.L. 1998; Tsuzuki S. 2000a; Tsuzuki S. 2000b; Pitoňák, M. 2008 to mentioned only few of them] Interestingly, it has been found that in some systems the energy difference between MP2 and CCSD(T) method does not vary significantly with the increasing size of the basis set whereas the overall interaction energy does. In light of this finding it is possible to devise various extrapolation schemes, such as the popular ΔCCSD(T) correction scheme:

$$E_{\mathrm{CCSD(T)/CBS}} = E_{\mathrm{MP2/CBS}} + \Delta E_{\mathrm{CCSD(T)/small}},$$

Where $E_{\mathrm{CCSD(T)/CBS}}$ is the estimate of the CCSD(T) energy in the complete basis set limit, $E_{\mathrm{MP2/CBS}}$ is the energy obtained by using an extrapolation scheme [Fast P. L. 1999; Varandas A. J. C. 2000], and $\Delta E_{\mathrm{CCSD(T)/small}}$ is the difference between the MP2 and CCSD(T) energies in the smaller basis set.

2.6. Practical aspects of the physico-chemical description of biomolecular systems

The systems of interest, biomolecules, are too large for a high-level quantum chemical description. In some cases, we may study a specific feature of the whole process that is localized in a part of the macromolecule and thus, it is possible to design a model system of a reasonable size. Such feature can be an enzymatic reaction localized in the active site, or interactions between specific amino acids stabilizing the protein structure. The impact of surrounding can be approximated by computationally cheaper methods as was mentioned in the introduction.

In case of interactions stabilizing the protein structure, the overall stabilization can be decomposed into a sum of pair interactions of individual residues. Contribution of each pair interaction can be computed individually on an isolated model. This approach would be exact, if the interacting pair would not be influenced by the other residues of the protein. Such assumption is

conceivable for the distant residues. To what extent is a specific interaction influenced (e.g., polarized) by a close residue depends on the type of interaction and the type of residue. The mutual influence can be quantified in form of the "cooperative effect". In case of three-body complex (ABC) the cooperative effect would be expressed as follows:

$$E^3 = E^{ABC} - E^{AB} - E^{BC} - E^{AC} + E^A + E^B + E^C$$

The total interaction energy can then be expressed as the sum of pair contributions and the cooperative effect,

$$E^{ABC} = E^{AB} + E^{BC} + E^{AC} + E^3 .$$

The structure, geometry, and interaction energy of molecular clusters are mostly obtained by variational (supermolecular) calculations. The stabilization energy is determined as the difference between the energy of the complex and the sum of the energies of the isolated subsystems (molecules).[Szabo A. 1982] This simple and straightforward procedure is complicated by a problem of an unequal description of the supersystem and the subsystems in terms of the size of the basis set which gives rise to the *basis set superposition error (BSSE)*.[Jansen H. B. 1969; Liu B. 1973] This error becomes significant especially when small basis sets or basis sets with diffuse functions are used. [Hobza P. 1988] The BSSE decreases with the increasing size of the basis set and vanish in an infinite basis set (or in the extrapolation to the complete basis set limit). [Dunning T. H. 2000] In the case of finite basis sets, the BSSE is now routinely eliminated by using the counterpoise correction method introduced by Boys and Bernardi. [Boys S. F. 1970] The correction is calculated by subtracting the fragment energy calculated in the regular basis set of the fragment from the fragment energy calculated in the full basis set for the whole complex. For cluster of two molecules, A and B, the counterpoise correction is defined as

$$E_{CP} = E_A(bas_A) - E_A(bas_{AB}) + E_B(bas_B) - E_B(bas_{AB})$$

where $E_A(\mathrm{bas}_A)$ is the energy of molecule A in the regular (atom-centered) basis set, $E_A(\mathrm{bas}_{AB})$ is the energy of molecule A in the basis set of the whole complex, etc.

The counterpoise-correction can be also implemented into the gradient optimization.[Simon S. 1996] Provided that a highly correlated method with extended basis set is used, the counterpoise-corrected (cc) gradient optimization yields accurate geometries, stabilization energies and various characteristics (e.g., vibrational frequencies) of a molecular cluster. While in the standard gradient optimization the gradient is calculated only once in each optimization cycle, the cc gradient

optimization involves the calculation of five gradients [i.e., ***g***(A...B), ***g***(A), ***g***(B), ***g***(A...Bghost), ***g***(B...Aghost)] in each optimization cycle, which makes the cc gradient optimization much more time-consuming.

Recently Jensen introduced a theoretical solution, which estimates the BSSE as a sum of atomic contributions,

$$E_{CP} = \sum_{A}^{atoms} E_A(bas_A) - E_A(bas_{AS})$$

Here, the $E_A(\mathrm{bas}_A)$ stands for the energy of atom A in the regular basis set bas_A, while $E_A(\mathrm{bas}_{AS})$ stands for the energy of atom A in a subset of the whole system basis set. In the intermolecular case the subset includes basis functions on all the atoms in the other molecule, while the basis functions on atoms within the same fragment are excluded. In the intra-molecular case, the subset includes basis functions on atoms separated from atom A in terms of bonding and distance. This method can be applied with equal ease to estimate both inter- and intra-molecular BSSE. The method requires typically double the amount of the computer time in comparison to the BSSE uncorrected energy calculation. [Jensen F. 2010]

2.7. Combined Quantum Mechanics and Molecular Mechanics (QM/MM) methods

The general discussion of QM/MM methods and various approaches can be found in [Senn H.M. 2007]. In the following, the standard (Oniom-like) approach used in the ComQum program [Ryde 1996] will be described in detail.

The total QM/MM energy is calculated as:

$$E_{\mathrm{tot}} = E_{\mathrm{QM}} + E_{\mathrm{MM123}} - E_{\mathrm{MM1}}$$

where E_{QM} is the QM energy of the quantum system truncated by hydrogen atoms, E_{MM1} is the MM energy of the quantum system, still truncated by hydrogen atoms but without any electrostatic interactions (the partial charges set to zero). Finally, E_{MM123} is the classical energy of all atoms in the system with original atoms at the junctions and with the charges of the quantum system set to zero (to avoid double-counting of the electrostatic interactions). By using this approach which is similar to the one used in the Oniom method [Maseras M. 1995; Svensson M. 1996; Vreven T. 2006] errors caused by the truncation of the quantum system should cancel out.

In the quantum chemical calculations the QM system is represented by a wave function, whereas all the atoms from MM region are represented by an array of partial point charges, one for each atom, taken for example from the standard Amber libraries. Thereby, the polarization of the QM system by the surroundings is included in a self-consistent manner (so called "electrostatic embedding"). In the MM calculations of the QM/MM forces and energies, all atoms are represented by the force field.

A "broken" covalent bond between systems 1 and 2 (a junction) is capped by hydrogen atom. The positions of the capping hydrogens are linearly related to the corresponding atoms in the full system (the hydrogen link approach). [Reuter N.I. 2000] In order to avoid over-polarization of the quantum system, point charges on atoms in the MM region bound to junction atoms are set to zero, and the remaining charges on the truncated amino acid are adjusted to keep the fragment neutral.

QM/MM forces used for the optimization are taken as the negative gradients of individual energies:

$$-\nabla E_{tot}\left(r_1, r_{H_j}, r_{2-3}\right) = -\nabla E_{QM}\left(r_1, r_{H_j}\right) - \nabla E_{MM123}\left(r_1, r_{C_j}, r_{2-3}\right) + \nabla E_{MM1}\left(r_1, r_{H_j}\right)$$

The positional vector of the junction atoms is expressed by the following relation:

$$r_{C_j} = r_{X_j} + k\left(r_{H_j} - r_{X_j}\right),$$

where

$$k = \left(r_{C_j} - r_{X_j}\right) / \left(r_{H_j} - r_{X_j}\right),$$

varies with method, functional and basis set used. Symbols r_1 denote positional vector of the atoms in system 1, $r_{2\text{-}3}$ are for atoms of systems 2 and 3, symbols r_{C_j}, r_{X_j} and r_{H_j} denote positional vectors of the atoms C_j, X_j, H_j. The described scheme with position of the atoms is depicted in the Figure 2.1.

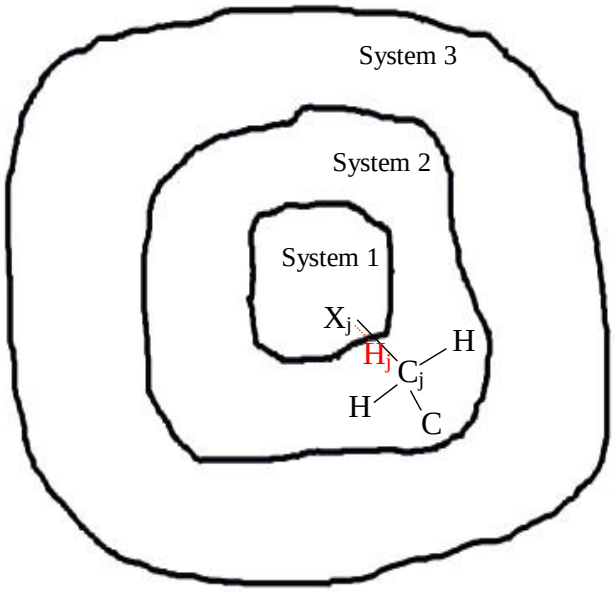


Figure 2.1: Scheme of the multi-layer treatment of the system by the QM/MM methods.

The computational algorithm can be outlined as follows:

```
Evaluate QM wavefunction
Repeat
   Evaluate the QM forces (from S1-S3 onto S1)
   Evaluate the MM forces (from S2-S3 onto S1)
   Add the forces
   Relax the geometry of S1 using these forces
   Change the coordinates of S1 in MM representation
   If S2 is to be relaxed
      Calculate the QM charges of S1
      Insert them into the MM representation
      Relax S2 by MM minimization with S1,S3 fixed
      Change the coordinates of S2 in QM representation
   Evaluate the QM wavefunction (and energy) of S1
   Evaluate the MM potential energy (S2-S1)
   Add the energies
until convergence
```

[Ryde U., http://www.teokem.lu.se/~ulf/Methods/comqum.html]

3. Results and Discussion

The ultimate goal of the presented studies is a reliable description of the strength of various interactions in biomolecules including an understanding to the enzymatic action. It is assumed that an accurate description is a prerequisite for any qualitative or quantitative discussions of the relation between the protein structure and protein function. In the first part, the interactions in biomolecules and their complexes in their minimum energy are studied whereas in the second part the structures on the reaction pathways, not always in their energy minima, are investigated.

3.1 Noncovalent interactions stabilizing biomolecular structures

3.1.1. Van der Waals interactions in hydrophobic ligand recognition [1]

Interactions of insects with their surroundings are typically mediated by the chemical signals. Mature females of moths ready to have offspring emit a sexual pheromone from their abdomen to attract males for mating. The "single-pheromone molecule" tuned detection system [Kaissling K.E. 1970] of males is located in branches of males' antennae. On these antennae there are olfactory hairs, which are filled with sensillar lymph surrounding specialized dendritic cells innervated to insect's brain. For sexual communication, the silkworm moth uses sex pheromone composed of three molecular species: (10E,12Z)-hexadeca-10,12-dien-1-ol, bombykol; (10E,12E)-hexadeca-10,12-dien-1-ol; and (10E,12Z)-hexadeca-10,12-dienal [Butenandt A.1956; Kasang G. 1978a; Kasang G. 1978b]. The first compound is by far the most abundant.

The sensillar lymph olfactory hairs contain a high concentration (10mM) of water-soluble pheromone binding protein (PBP), whose role is not fully understood. A series of experimental studies [Maida, R.1993; Sandler, B.H. 2000; Oldham, N.J. 2000; Oldham, N.J. 2001; Wojtasek, H. 1999; Horst 2001] have shown that bombykol is tightly bound in a flask-like pocket of BmPBP. Based on the studies several possible scenarios have been suggested [Kaissling K.E. 2001]: (i) PBP acts as a carrier that shuffles the lipophilic pheromone through the sensillar lymph to the pheromone receptor where it is released from PBP; (ii) whole PBP-pheromone complex is recognized by the pheromone receptor; and (iii) PBP is involved in the pheromone "cleaning" process after its recognition on the receptor.

The aim of our analysis was to elucidate the character of the bmPBP-pheromone interaction. Despite the experimental structure of bmPBP-pheromone complex was determined, the exact nature of binding was not clear, except for a hydrogen bond between the polar OH group of the bombykol and the side chain of Ser56 that was quite apparent from the analysis of the crystal structure. (see

Figure 3.1.1.1) The character of other binding interactions was rather speculative, and other contacts were described as unspecific hydrophobic interactions. Only the residues Phe12 and Phe118 were described to form a sandwich-like structure with the double bonds of the pheromone, but their role was also assumed to be nonspecific [Sandler, B.H. 2000].

Using simple model fragments as the representatives of the residues, the interaction energies of their complexes with bombykol were calculated. A size-consistent correlated *ab initio* method (RI-MP2/aug-SVP) was applied, which is capable of satisfactory description of all the types of weak interactions present in the biomolecules, including those with a major dispersion energy component. The anticipated systematic errors of MP2 method (overestimation of the studied types of interactions) have been corrected by the benchmark CCSD(T)/cc-pVDZ calculations. Owing to huge computational demands of this method, a set of simplified complexes representing the studied structures and interactions has been selected. It was shown that the MP2 method systematically overestimates interaction energies by 5-17%.

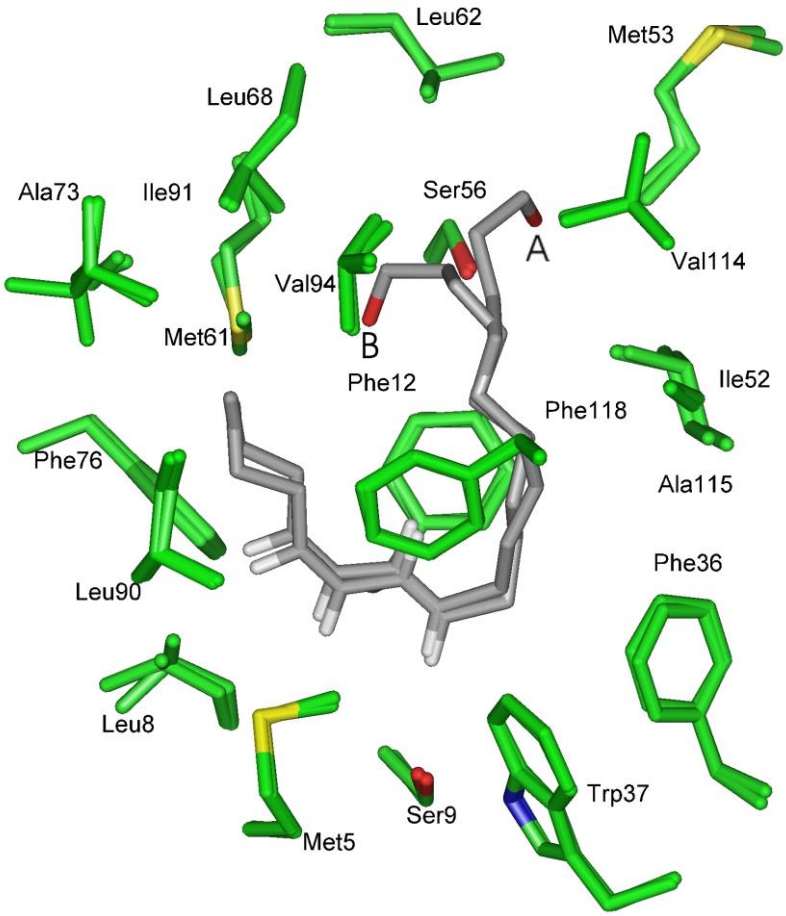


Figure 3.1.1.1: Schematic representation of the binding cavity with superimposed bombykol and amino acid residues from both structures(A, B), present in the single asymmetric crystal unit. Only hydrogen atoms which are connected to the bombykol double bonds are shown.

It has been shown that the pheromone molecule is not just expelled into the binding cavity from the outer environment (polar sensilar lymph) due to its hydrophobicity. On the contrary, the pheromone is attracted by several aromatic residues in the cavity (via X-H…π and π… π interactions) that interact with practically the whole hydrocarbon unsaturated chain of the pheromone (see Figure 3.1.1.2 and Figure 3.1.1.3).

The thorough understanding of the interactions involved in the bombykol binding provided us with solid basis for the discussion and proposal of engineering of both molecules (pheromone and receptor) and of the role of the [receptor…ligand] complex in the process of pheromone recognition.

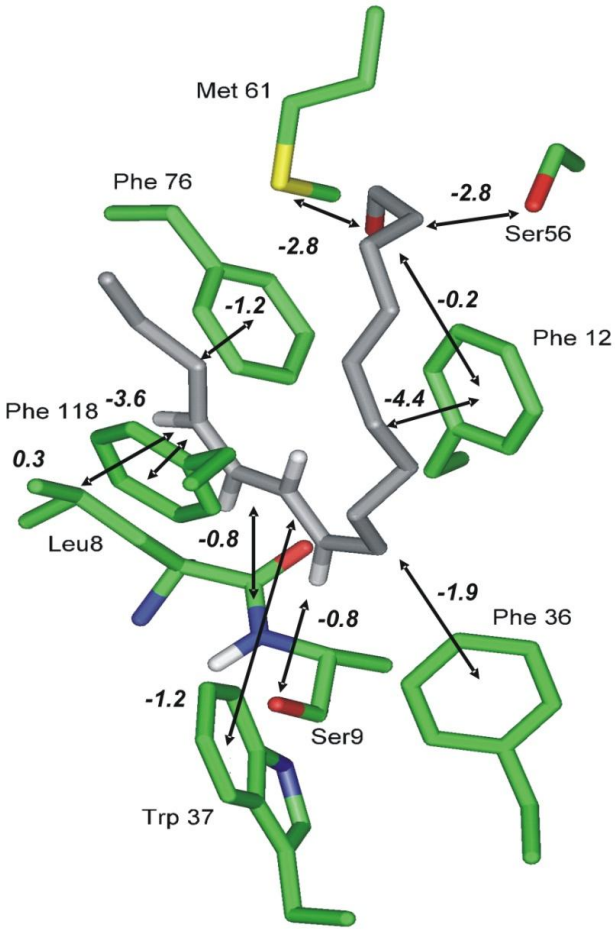


Figure 3.1.1.2: Main interactions (in kcal/mol) of the pheromone bound in the binding pocket.

However, determining experimentally binding constants of air-borne lipophilic pheromone ligands in aqueous solutions of proteins is not a trivial task. It is difficult to keep the immiscible, water-insoluble ligand in solution at the higher concentrations as needed for NMR or calorimetric assays. For PBP-pheromone complex, several attempts including the utilization of novel methodological

approaches have been carried out. [Campanacci J. 2001; Leal W. S. 2005] Unfortunately, the results were not unambiguous. It clearly demonstrates that the experimental studies of such systems are at the cutting edge of experimental feasibility (e.g., calorimetric measurements for the separation of enthalpy and entropy contribution are still thought to be unconceivable). The theoretical study enabled us to assess the dominant part of interaction enthalpy, the interaction energy, and has shown, that it can be the driving force of the process. Generally, contrary to widely accepted classical assumption, the interaction between hydrophobic residues can be considered as enthalpically driven.

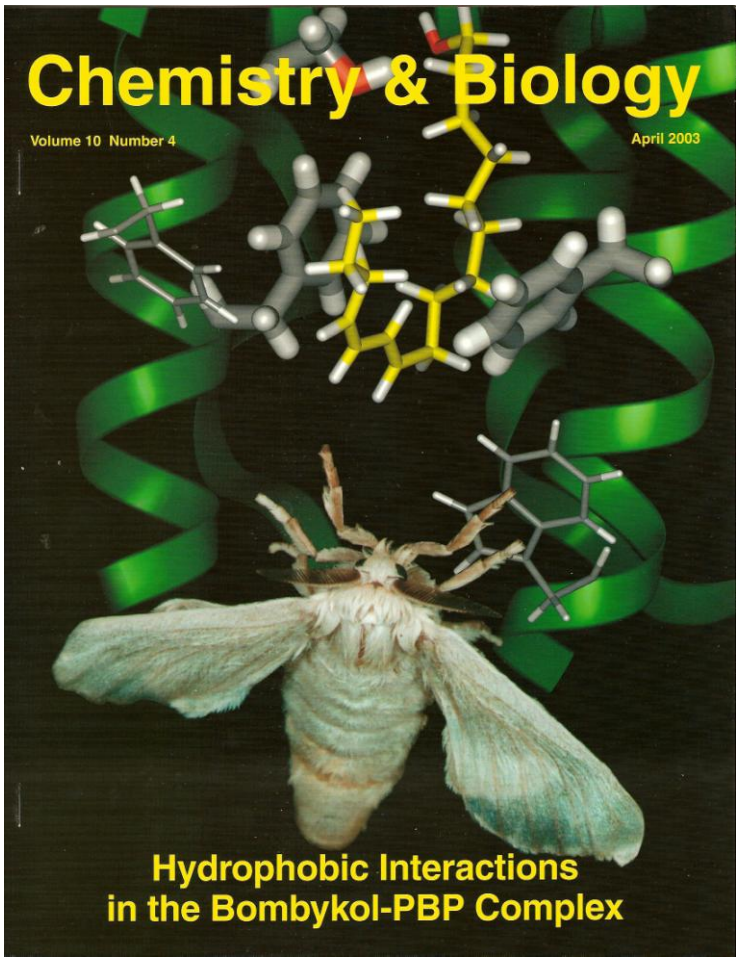


Figure 3.1.1.3 Male of silkworm moth, *bombyx mori*, and schema of the PBP-bombykol complex with the key binding residues highlighted. Hydrogen atoms are white, carbon atoms gray (yellow in case of the pheromone) and oxygen atoms are red. The protein back bone is represented by a green ribbon.

[1] Klusák V., Havlas Z., Rulíšek L., Vondrášek J., Svatoš A. (2003) Sexual attraction in the silkworm moth: Nature of binding of bombykol in the pheromone binding protein – an ab initio study. *Chem. Biol. 10,* 331-340.; see Appendix A for the full paper

3.1.2. Rational design of pheromone analogs [2]

As a follow up of the above described work the mono-, difluorinated, and thioanalogues of *Bombyx mori* female sex pheromone (bombykol, **1**) were designed according to small-model *ab initio* calculations. A new simplified binding assay based on nanoLC-linear ion trap ESI-MS for quantifying complexation of the *B. mori* pheromone-binding protein (BmPBP) with native and newly synthesized analogues (see Figure 3.1.2.1 for their chemical structures) was developed. This experimental approach can speed up evaluation of biological assays and minimize sample handling. For the improvement of the assay accuracy, there is a need for an internal standard with similar chemical structure as the native bombykol but with a different molecular mass. It would allow the binding constants of diverse ligands to be determined in a standardized way.

To this aim, using small model molecules in an analogous way as in the previous study [Klusák V. 2003], we characterized changes in the energetics of binding caused by the atom isosteric replacement. Specifically we have estimated the effect of pheromone structure isosteric change on the interaction energy for the Phe12 and Phe118 residues that were found previously as the most important interacting partners.[Klusák V. 2003] Using the X-ray structure of the BmPBP–bombykol complex, the individual CH_2 groups were replaced (one by one) by CF_2 group or S atom and the structures of pheromone together with the two phenylalanine benzene rings were optimized for the positions of hydrogen and fluorine atoms only. We concluded that in the geometry confined to the X-ray structure none of the studied modification resulted in a significant modification of the overall interaction energy. However, several positions seemed to be slightly preferential. (see Table 3.1.2.1)

Table 3.1.2.1: The effect of the substitution of CF_2 or S group for the CH_2 group in bombykol (**1**). The interaction energies (kcal/mol) of pheromone with both Phe12 and Phe118 in native crystallographic conformation. Energies are relative to the native pheromone interaction.

Substituted Position in **1**	**C4**	**C5**	**C6**	**C7**	**C16**
CF2	-0.4	0.9	-0.2	0.2	0.0
S	-0.6	-0.3	-0.3	-0.6	0.0

In the next step, we studied in detail interactions of three model molecules (2-fluoropropane, 2,2- difluoropropane, and dimethylsulfide) with benzene and compared them with propane (bombykol) with the aim to elucidate the orientation preference of the tested functional groups towards the aromatic residues. Dissociation curves were calculated at the RI-MP2/aug-cc-pVDZ level of theory for three different orientations. In addition, unrestrained optimizations of the complexes starting from structures corresponding to the calculated optimum benzene ligand distance were performed. (see Fig. 3.1.2.2) The effect of solvation has been assessed using a polarized continuum solvation model.

The calculated solvation energies confirmed our assumption that the CH_2 group replacement by more polar groups systematically enhances solubility of the ligand. The changes in ΔG_{solv} ranged between -2.2 and -3 kcal/mol. In order to have similar or stronger affinity in comparison to the native ligand, the modified ligands need to compensate for the unfavorable effect of solvation. The calculations on model molecules showed that the studied functional groups can slightly enhance the interaction energy. Hence, the compensation to the energy of ligand–protein interaction may arise from the enthalpic contribution. (Table 3.1.2.2). Though the conformations after optimization without any restraints are different, the interaction energies are surprisingly similar (Table 3.1.2.2).

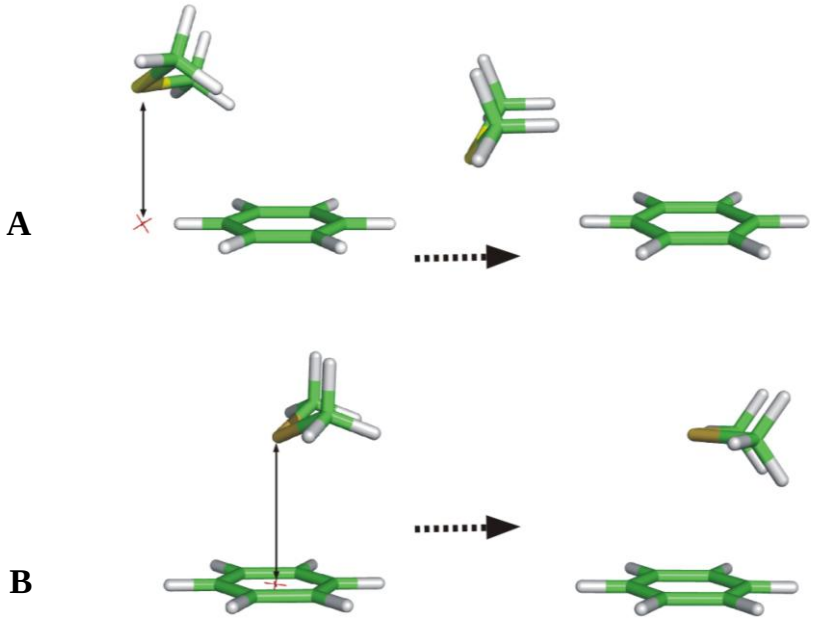


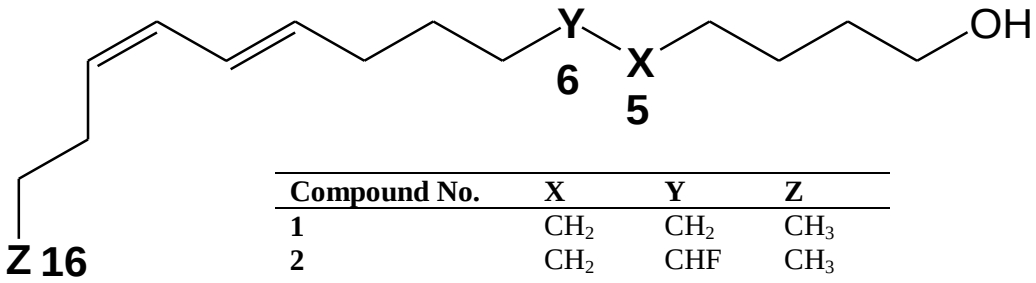
Fig. 3.1.2.2 A coarse scan of the interaction energy minima for the model molecules placed above the plain of benzene molecule. In the figure is example of dimethylsufide. Two starting geometries (A, above benzen centre; B, above the centre, 2A shifted aside), corresponding to minimun energy of a one dimensional scan, were subjected to unrestrained optimization.

Table 3.1.2.2: A coarse scan of the interaction energy minima for the model molecules placed above the plain of benzene molecule. Two starting orientations were considered corresponding to the calculated optimum benzene ligand distances. (cf. Fig. 3.1.2.2) The calculated interaction energies (in kcal/mol) of the complexes result from unrestrained optimization.

Starting orientation	*n*-propane	2,2-difluoropropane	2-fluoropropane	dimethylsulfide
Above centre	-3.1	-3.4[a]	-3.5	-2.6
Above, 2 Å away	-2.9	-3.3[a]	-3.4	-3.8

[a] The optimization from both orientations ended in a similar conformation.

In addition, the difference in behavior of 2-fluoropropane and 2, 2-difluoro-propane is also remarkable, indicating that the interaction enthalpy increases in response to contact of the aromatic ring and the C–H bond, polarized by the adjacent C–F group. This led us to consider testing the racemic mixture of monofluorinated bombykol. We have chosen to test the following substitutions in the bombykol molecule (Fig. 2): C6 for –CF2– and for the racemic mixture of –CFH–, C16 for –CH_2F–, and C5 for –S–.



Compound No.	X	Y	Z
1	CH_2	CH_2	CH_3
2	CH_2	CHF	CH_3
3	CH_2	CF_2	CH_3
4	S	CH_2	CH_3
5	CH_2	CH_2	CH_2F

Figure 3.1.2.2 Chemical structures of rationally designed bombykol (**1**) analogues.

The binding properties of native and thioanalogue (cf. Fig. 3.1.2.2) with PBP were studied in detail. The dissociation constant (K_D) of **1** and **4** was determined to be 2.1×10^{-6} M and 2.4×10^{-6} M. Almost identical experimental values obtained for both ligands correlate with qualitative conclusion from the *ab initio* calculations. The fluorinated pheromone analogs were not resolved in the experiment, which can be due to their low affinity.

The theoretical study can be extended by using more advanced and time consuming methods (MD, QM/MM) providing further support for the experiments. In future, this may help to clarify the mechanisms of the peri-receptor events in insect olfaction and might provide an answer to the question of how important is the PBP–pheromone interaction for the pheromone recognition.

[2] Mansurova M., Klusák V., Nešněrová P., Muck A., Doubský J., Svatoš A. (2009) Design and synthesis of bombykol analogues for probing pheromone-binding protein–ligand interactions. *Tetrahedron 65*, 1069-1076.; see Appendix B for the full paper

3.1.3. Van der Waals interactions in hydrophobic core packing and stabilization [3]

Formation of the protein core is believed to be the consequence of hydrophobic forces of entropic nature.[Tanford, C. 1978; Rose, G. D. 1985] The classical hydrophobic effect [Meyer, E. A. 2003] is characterized by small contribution (repulsive or attractive) of complexation enthalpy to the overall value of free energy. This, together with low occurrence of hydrogen bonds in the protein core, leads to the assumption that the energy (enthalpy) contribution of the core formation to protein folding is small or negligible.

Recent theoretical and experimental investigations of various types of noncovalent interactions have shown [Müller-Dethlefs, K. 2000] that a rather large attraction is provided not only by hydrogen bonding but also by other types of noncovalent interactions. Thus it is highly interesting to find out the strength of the stabilizing contributions of amino acids in a hydrophobic core of proteins. This question is of key importance for understanding the mechanism of protein folding as well as understanding protein secondary and tertiary structure.

The aim of the presented work was to evaluate the stabilization energy provided by interactions of two phenylalanines in a model protein hydrophobic core based on a high-resolution X-ray (0.92 Å) structure of rubredoxin, a small soluble FeS protein (PDB code 1RB9, c.f. Figure 3.1.3.1). Stabilization energy was determined using high-level correlated ab initio calculations, as a sum of the estimate of the complete basis set limit of the MP2 stabilization energy and CCSD(T) correction term.

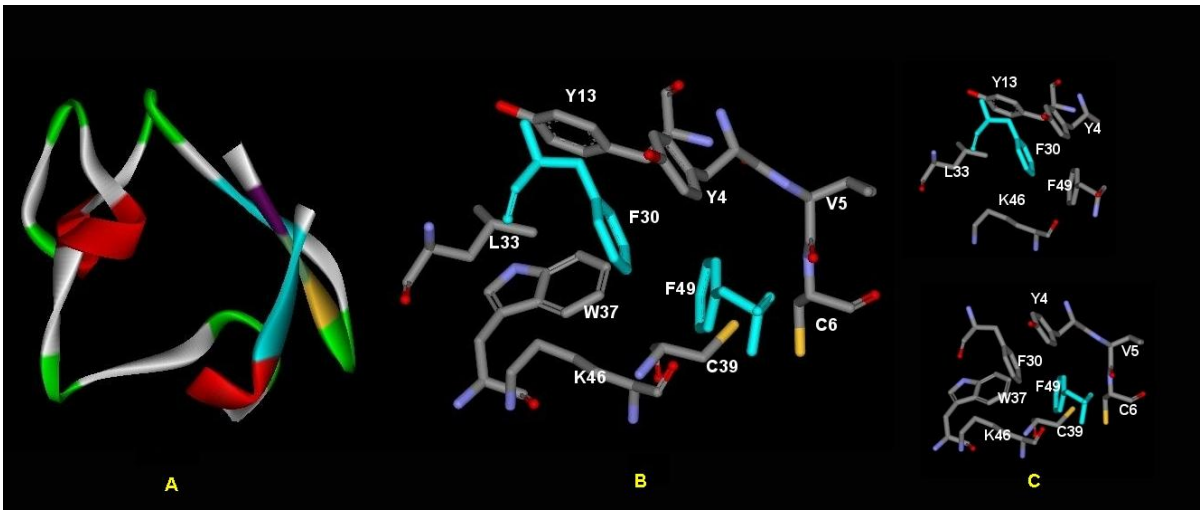


Figure 3.1.3.1. The rubredoxin. A) Schematic view of the whole protein (52 amino acids) represented by a ribbon; B) super cluster of F30 and F49; C) both sub clusters individually.

We have shown that stabilization inside the hydrophobic core of a small protein, rubredoxin, determined by means of high-level correlated ab initio calculations (complete basis set limit estimate of MP2 stabilization energy + CCSD(T) correction term), is surprisingly strong (see Figure 3.1.3.2 and Table 3.1.3.1). The sum of pair interaction energies amounted to almost 24 and 28 kcal/mol in case of Phe30 and Phe49, respectively. This attraction originates in London dispersion energy between aromatic rings or between an aromatic ring and an aliphatic chain, and is comparable to classical H-bonding. Moreover, residues of aromatic nature can participate in several strong interactions at once, which may be crucial for the role of key residues in establishing networks inside a protein.[Vendruscolo, M. 2001]

Table 3.1.3.1. The pair interaction energies of the selected residues clustered around F30. The CCSD(T)/CBS estimates were calculated as the sum of energy calculated using the RI-MP2/CBS extrapolation scheme and the ΔCCSD(T) quantity, which is the difference of the CCSD(T)/6-31G*(0.25) and RI-MP2/6-31G*(0.25) energies.

		RI-MP2		**ΔCCSD(T)**	**CCSD(T)**
	aug-cc-pVDZ	aug-cc-pVTZ	CBS estimate	6-31G*(0.25)	CBS estimate
residue	(kcal/mol)	(kcal/mol)	(kcal/mol)	(kcal/mol)	(kcal/mol)
F49	-3.1	-3.3	-3.3	-	-
K46	-3.1	-3.3	-3.4	0.3	-3.1
L33	-4.9	-5.3	-5.5	0.5	-5.0
Y13	-4.2	-4.4	-4.5	0.6	-3.9
Y4	-6.5	-6.8	-7.0	-	-
sum	-21.8	-23.2	-23.7	-	-

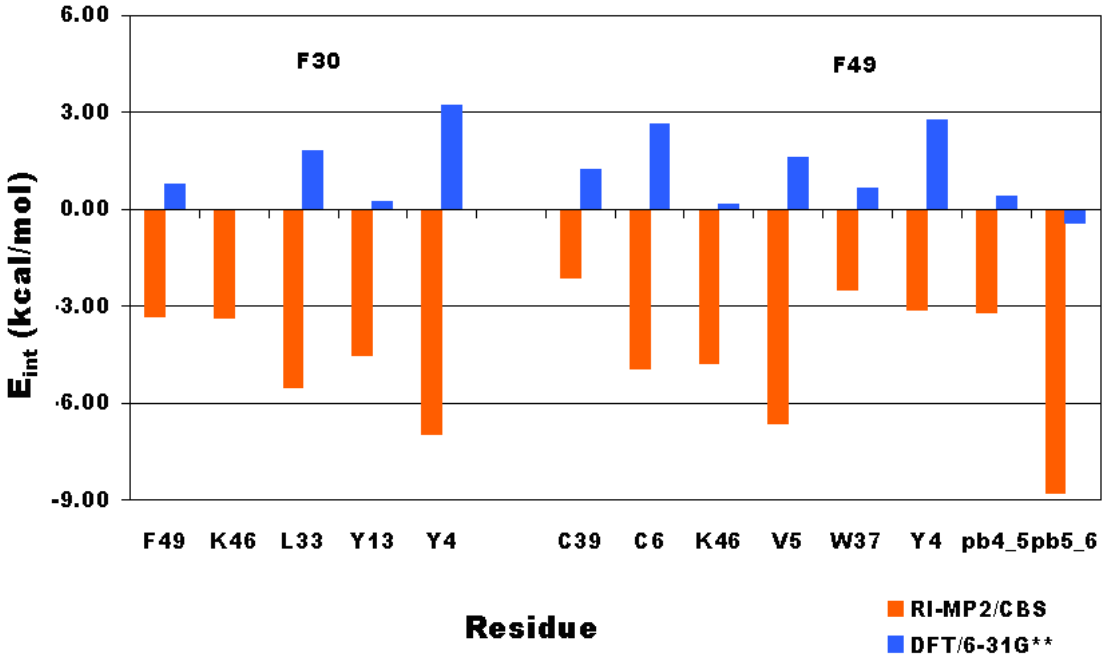


Figure 3.1.3.2. DFT and MP2/CBS(*est.*) interaction energies of F30 and F49 phenylalanines with selected amino acids from the rubredoxin core; DFT interaction energy of F30…K46 pair amounts to 0. pb4_5pb5_6 stands for back bone fragment containing two peptide bonds between residues 4, 5 and 5, 6.

These results clearly demonstrate that there is a strong attraction inside a hydrophobic core. This finding may lead to substantial changes in the current view of protein folding. Hydrophobic nature of a protein core implies that hydrophobic interactions can initiate the folding process. Present results also indicate a decisive role of stabilization energy (enthalpy). Hence it is possible to assume that the enthalpy (not only the entropy) can play a significant role during the early stages of protein folding. We have also pointed out the inability of the DFT/B3LYP method to describe a strong attraction between studied amino acids (cf., Figure 3.1.3.2).

[3] Vondrášek J., Bendová L., Klusák V., Hobza P. (2005) Unexpectedly strong energy stabilization inside the hydrophobic core of small protein rubredoxin mediated by aromatic residues: correlated ab initio quantum chemical calculations. *J. Am. Chem. Soc. 127*, 2615-2619.; see Appendix C for the full paper

3.1.4. Accounting for intramolecular Basis Set Superposition Error in complex molecules [4]

Correlated ab initio calculations, such as the popular MP2 (or RI-MP2) method, applied on large systems, suffer from the intramolecular basis set superposition error (BSSE). For a long time, it was believed that the basis set extension effect was characteristic for molecular clusters and did not affect isolated systems. Only recently, it has been pointed out that the same effect, i.e. improving the basis set of one part of a system by orbitals localized on the atoms of the other part, affects the calculation of the relative energies of isolated systems, especially in very flexible molecules such as peptides.[van Mourik, T. 2006; Jensen, F. 1996; Senet, M. L. 2001; Han, Y. 2002] Apparently, non-covalent interactions act between atoms or groups of atoms regardless whether they are placed in one or more chemical entities. In analogy with the BSSE in molecular clusters, this effect is referred to as intramolecular BSSE.

What intermolecular BSSE has in common with intramolecular BSSE is that it is also repulsive, and unless it is eliminated, the stabilization energy is artificially too high. The use of the counterpoise procedure, which fully eliminates the intermolecular BSSE, is impractical for isolated systems. The only possibility of eliminating intramolecular BSSE is to use procedures, which are BSSE-free or at least, which have only small values of the BSSE.

This practically means either to use a correlated method with an extremely large basis set, which may eliminate the BSSE, or to use a suitable DFT procedure (covering the London dispersion energy), which is characterized by a small BSSE. The former procedure is impractical (or rather

impossible) for extended systems, and the practical solution to the problem is represented by the second approach.

Since most of the popular DFT functionals do not cover London dispersion energy, their use for the study of weakly bound molecular clusters is limited. [Valdes, H. 2006; Dabkowska, I. 2005; Piacenza, M. 2005; Hobza, P. 1995; Rappe, A. K. 2000; Allen, M. J. 2002] A practical solution of this problem has been recently accomplished by empirically augmenting the DFT energy by London dispersion energy terms [Grimme, S. 2004; Grimme, S. 2006; Jurečka P. 2007; Elstner M. 2001; Zimmerli U. 2004; von Lilienfeld A. O. 2004; Becke A. D. 2005; Gerber I. C. 2007; Grimme S. 2007] or by devising a new set of DFT functionals [Zhao Y. 2008]. It can be noted that empirical correction has virtually no effect on the computer time.

The aim of this presented study is to critically assess the intramolecular BSSE. Its consequences for the calculations of larger molecular systems are demonstrated on two model systems: (i) [*n*]helicenes consisting of *all-ortho* annulated benzene rings [Urbano A. 2003] and (ii) phenylalanyl-glycyl-phenylalanine (FGF) tripepdide. The helicene molecules are characterized by a very large number of intramolecular π–π stacking interactions (perhaps one of the largest when normalized by the number of atoms in the given molecular system) and are important prototypical molecules in the field of functional materials. The FGF model peptide is an excellent model system for the aromatic interactions in proteins. The importance of such interaction has been illustrated in the previous chapters of this thesis.

In this study, we demonstrated that the intramolecular BSSE can influence the energies of folded molecular systems dramatically and cause predictions of erroneous geometries as well (see Figure 3.1.4.1.). Namely, in the case of [*n*]helicenes, it has been shown that the intramolecular BSSE can be so large that the calculations predict clearly absurd results, such as the higher stability of [*n*]helicene as compared with [*n*]phenacene for $n > 6$ (using using MP2 and medium to large basis sets, such as TZVP or aug-cc-pVDZ). (see Table 3.1.4.1)

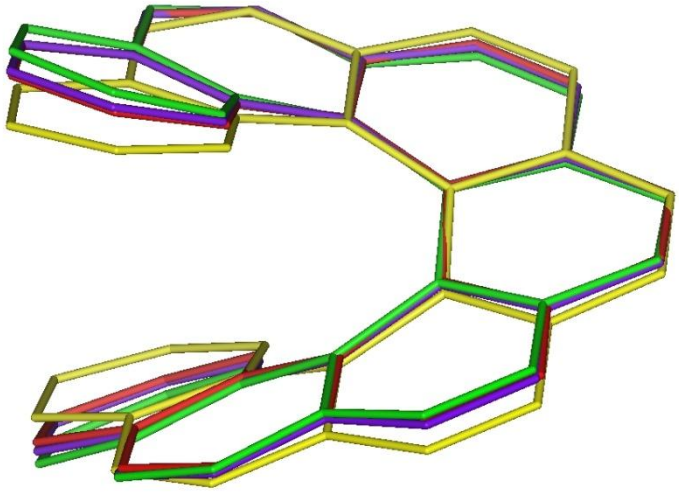


Figure 3.1.4.1. A comparison of the equilibrium structures of [7]helicene obtained using RI-DFT (green), RI-DFT-D (purple), and RI-MP2 (yellow) methods with the experimental X-ray structure (red).

Table 3.1.4.1: The calculated reaction energies (ΔE_n) for the homodesmic reaction (the addition of naphthalene to [*n*-1]helicene yielding [*n*]helicene and benzene molecules) used for estimating the stabilities of growing helical systems. The 6-31G(d) basis set was used for geometry optimization (in both DFT and MP2 methods). The TZVP was used to calculate the single point energies. All values are in kJ.mol^{-1}.

Method	RI-DFT/ PBE	DFT/ B3LYP	RI-MP2	RI-DFT-D/ PBE	DFT-D/ B3LYP
n \ BS	TZVP	TZVP	TZVP	TZVP	TZVP
3	-4.9	-5.3	-11.9	-5.6	-6.1
4	23.7	27.0	-1.4	21.0	23.3
5	20.6	24.0	-8.5	15.3	17.3
6	16.1	19.6	-19.5	7.8	9.1
7	20.9	25.6	-27.4	7.1	9.8
8	21.5	27.0	-35.7	4.6	6.9
9	20.2	25.8	-38.7	3.6	5.4
10	22.8	28.3	-34.3	5.9	8.0
11	23.3	29.1	-37.2	5.8	8.0
12	23.0	28.8	-38.0	5.3	7.1
13	24.0	29.6	-36.3	5.7	7.3
14	24.1	30.2	-38.0	5.7	7.5

Regarding the FGF tripeptide, it has been shown that intramolecular BSSE can lead to the wrong order of the stability of the conformers. Therefore, extreme caution must be exercised when interpreting the results of quantum chemical calculations using correlated methods, such as MP2 and CCSD(T), in systems for which a large intramolecular BSSE is anticipated. On the other hand the (RI-)DFT-D method has been shown that it is much less susceptible to BSSE than MP2 and CCSD(T) and for the geometry optimized structures properly accounts for the dispersion energy.

Finally, it shall be noted that very recently, Jensen introduced a theoretical solution, which estimates the BSSE as a sum of atomic contributions. The method requires typically double the amount of the computer time in comparison to the calculation of the uncorrected energy. [Jensen F. 2010] It remains, however, to be shown whether this method will be practical and stable in calculations of stabilization energies of extended molecular systems.

[4] Valdés H., Klusák V., Pitoňák M., Exner O., Starý I., Hobza P., Rulíšek L. (2007) Evaluation of the intramolecular basis set superposition error in the calculations of larger molecules: [n]helicenes and Phe-Gly-Phe tripeptide. *J. Comput. Chem. 29*, 861-870.; see Appendix D for the full paper

3.2. Reaction mechanism of dinuclear zinc enzymes and significance of the active site interactions

3.2.1. Reaction mechanism of GCPII in comparison to other mono and dizinc metalloenzymes[5]

Human glutamate carboxypeptidase II (GCPII, EC 3.4.17.21) is a membrane-bound zinc-dependent exopeptidase [Rawlings, N. D. 1997]. Other members of this family structurally related to GCPII include prokaryotic amino peptidases from *Aeromonas proteolytica* [Chevrier, B. 1994] (AAP) and *Streptomyces griseus* [Greenblatt H. M. 1997] (SGAP). Human GCPII is expressed in a wide range of tissues, including brain, prostate, small intestine, and kidney [Kinoshita Y. 2006; Troyer J. K. 1995; Šácha P. 2007]. One of its functions is hydrolysis of endogenous *N*-acetyl-L-aspartyl-L-glutamate (*N*-Ac-Asp-Glu, also known as NAAG) with release of *N*-acetyl-L-aspartate and free glutamate [Robinson M. B. 1987], the latter being a potent excitatory neurotransmitter. Several GCPII inhibitors have been shown to be neuroprotective in animal models of neurological disorders associated with high levels of glutamate, such as stroke and neuropathic pain [Slusher B. 1999; Zhou J. 2005]. The enzyme is an important therapeutic target for neurodegeneration and prostate cancer as well.

The presented study was initiated by the availability of several GCPII crystal structures, including the unliganded form of the enzyme [Davis M.I. 2005; Bařinka C. 2007a] and complexes with transition state analogues [Mesters J. R. 2006; Bařinka, C. 2008], glutamate as a product of *N*-Ac-Asp-Glu hydrolysis [Mesters J. R. 2006], and several glutamate mimetics/derivatives [Bařinka C. 2007b]. The structural results have been further complemented by the site directed mutagenesis studies focused on the effects of mutations in both the S1 and S1' pockets of GCPII on the K_m and k_{cat} values [Speno, H. S. 1999; Mlčochová, P. 2007].

Despite the above-mentioned efforts, the reaction mechanism of GCPII was not understood in detail. Although for the similar enzyme - AAP - experimental (X-ray crystallography, kinetics, spectroscopy, and site-directed mutagenesis) and computational data were available [Chevrier B. 1994; Harding, M. M. 1999; Stamper, C.2001; Lowther W. T. 2002; Holz R. C. 2002; Schürer G. 2004; Desmarais W. 2006] and reaction cycle has been proposed, its atomistic details were still a matter of debate. Probably the most comprehensive view in case of AAP was obtained in the computational study of Chen *et al.* [Chen S. 2008], based on cluster model approach.

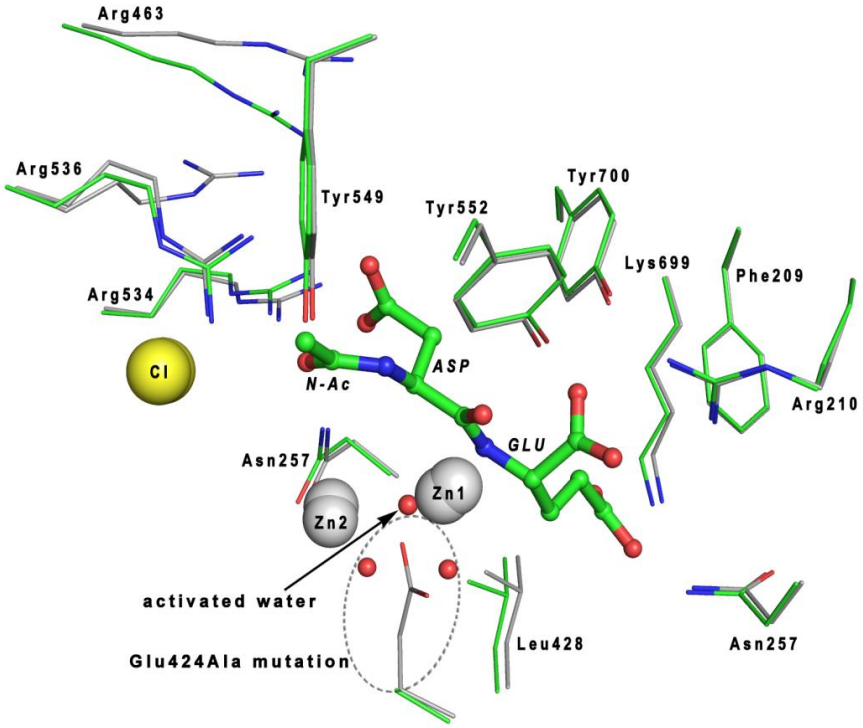


Figure 3.2.1.1: The structure of the [GCPII(E424A)…*N*-Ac-Asp-Glu] complex (carbon atoms colored green) was superimposed on the structure of the wild-type [GCPII…inhibitor(SPE)] complex (carbon atoms colored gray, the inhibitor was omitted for clarity; PDB code 3BHX) NAAG is in ball-and-stick representation, protein residues are in sticks Zn^{2+} and Cl^{-} ions are shown as gray and yellow spheres, respectively. Note the two water molecules at the approximate positions of the 'missing' Glu424 γ-carboxylate group.

The joint crystallographic, biochemical, and computational evidence, concentrated in our study, allowed us finally to propose a detailed reaction mechanism of substrate hydrolysis by human glutamate carboxypeptidase II. It has been, to our best knowledge, the first QM/MM study of the dizinc carboxypeptidase that showed a good agreement with the experimental (both structural and kinetic data). Moreover, the crystal structure of the substrate bound in the active site of the GCPII (E424A) mutant (at 1.7 Å resolution) was the first experimental description of a GCPII Michaelis complex. An agreement between this structure and our QM/MM model was obtained which gave us confidence in the calculated results. It can be noted that the model starting structure for the calculations (Michaelis complex) was constructed from an early structure of [GCPII…inhibitor] complex (PDB code 2C6C). The X-ray structure of the substrate bound in the mutant active site with an intact peptide bond also suggested that Glu424 was directly involved in the hydrolysis from its early stages. (c.f, Fig. 3.2.1.1)

Using the QM/MM approach we have identified one acceptable intermediate and two

corresponding transition states, one leading from the intermediate to the structure of Michaelis complex, and the other leading to the peptide bond cleavage. (c.f., Fig. 3.2.1.2)

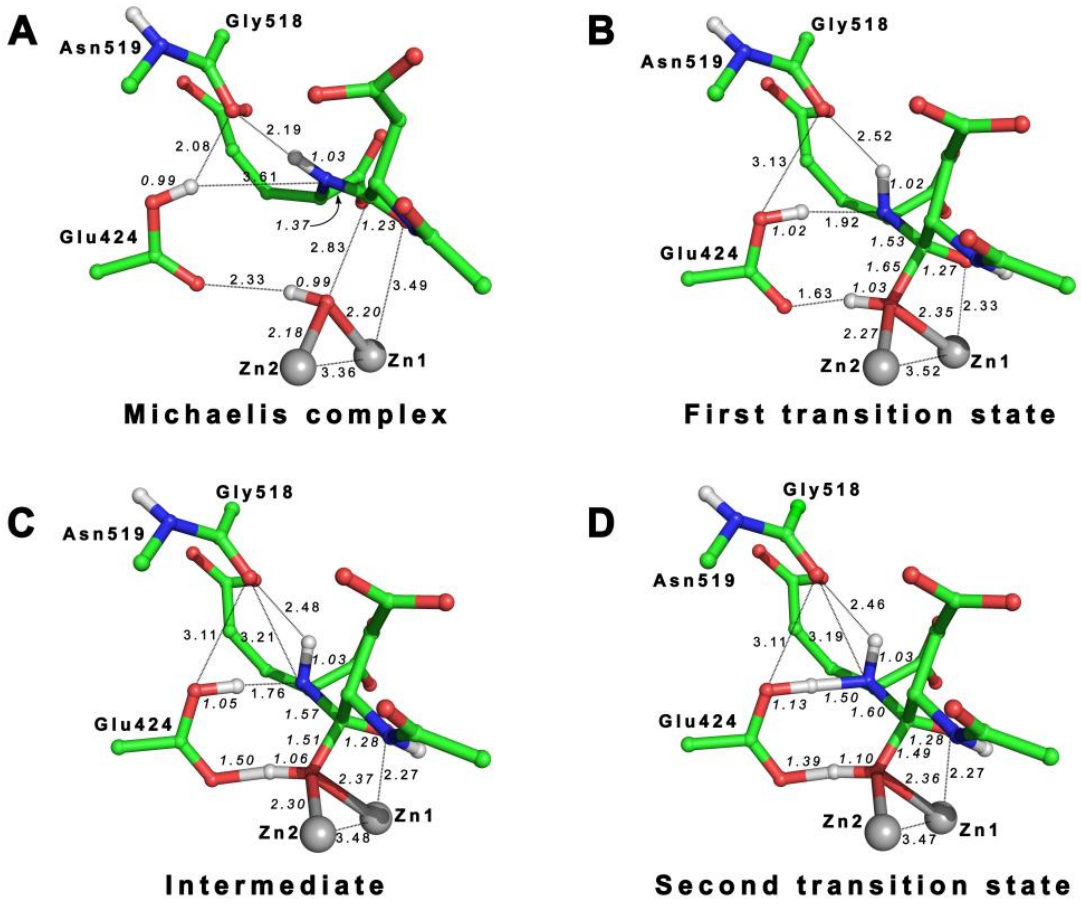


Figure 3.2.1.2: The QM/MM equilibrium structures of (A) the Michaelis complex (substrate bound in the active site), (B) the first transition state (TS1), (C) the reaction intermediate, and (D) the second, rate-determining, transition state (TS2) leading to the products. Important bond lengths and other inter-atomic distances are shown in italic and regular font, respectively.

The OH^- anion that is coordinated between two Zn^{2+} ions in the GCPII active site is considerably less nucleophilic than a free hydroxide ion, because two of its lone pairs are engaged in coordination bonds to the zinc ions. Indeed, the QM/MM calculations indicate that a direct nucleophilic attack of the hydroxide on the peptide bond carbon is an unlikely alternative. The attack is more likely to proceed concertedly with the proton bound to the Glu424 approaching the peptide bond nitrogen atom, forming a hydrogen bond and leading to pyramidalization of the nitrogen and concomitant formation of a metastable tetrahedral intermediate. As was proposed already by Monzingo and Matthews [Monzigo A. F. 1984] for thermolysin, and for GCPII by Mesters [Davis M. I. 2005; Mesters J. R. 2006], this study supports that Glu424 acts as a general

base/acid, shuttling the first proton from the activated water molecule to the peptide bond nitrogen and then accepting the second proton from the bridging hydroxide anion (which turns into the carboxylate oxygen at the end of the reaction).

A crossing of the first barrier of the reaction can be considered as the stabilization of the less populated resonance structure of the peptide bond (-NH-C^{+}-O^{-} *vs.* more stable -NH-C=O-) by its proximity to the Zn1 ion, its interaction with the Tyr552 hydroxyl group, and the bridging hydroxide moiety. According to the QM/MM calculations, the key feature of the hydrolytic reaction is the peptide bond nitrogen stabilization in the sp^3 hybridization state via interaction with the proton on the carboxylic group of Glu424. As a result, the electrophilicity of the peptide carbon is increased, thus facilitating the formation of a bond between the carbon and the oxygen from the OH^- ion. A crossing of the second barrier is triggered by a proton transfer from the protonated Glu424 to the nitrogen of the peptide bond followed by the transfer of the second proton from the OH^- moiety (already bound to the peptide bond carbon) back to the Glu424 carboxylic group. At the same time, the peptide bond is cleaved. (see Figure 3.2.1.2 and Figure 3.2.1.3)

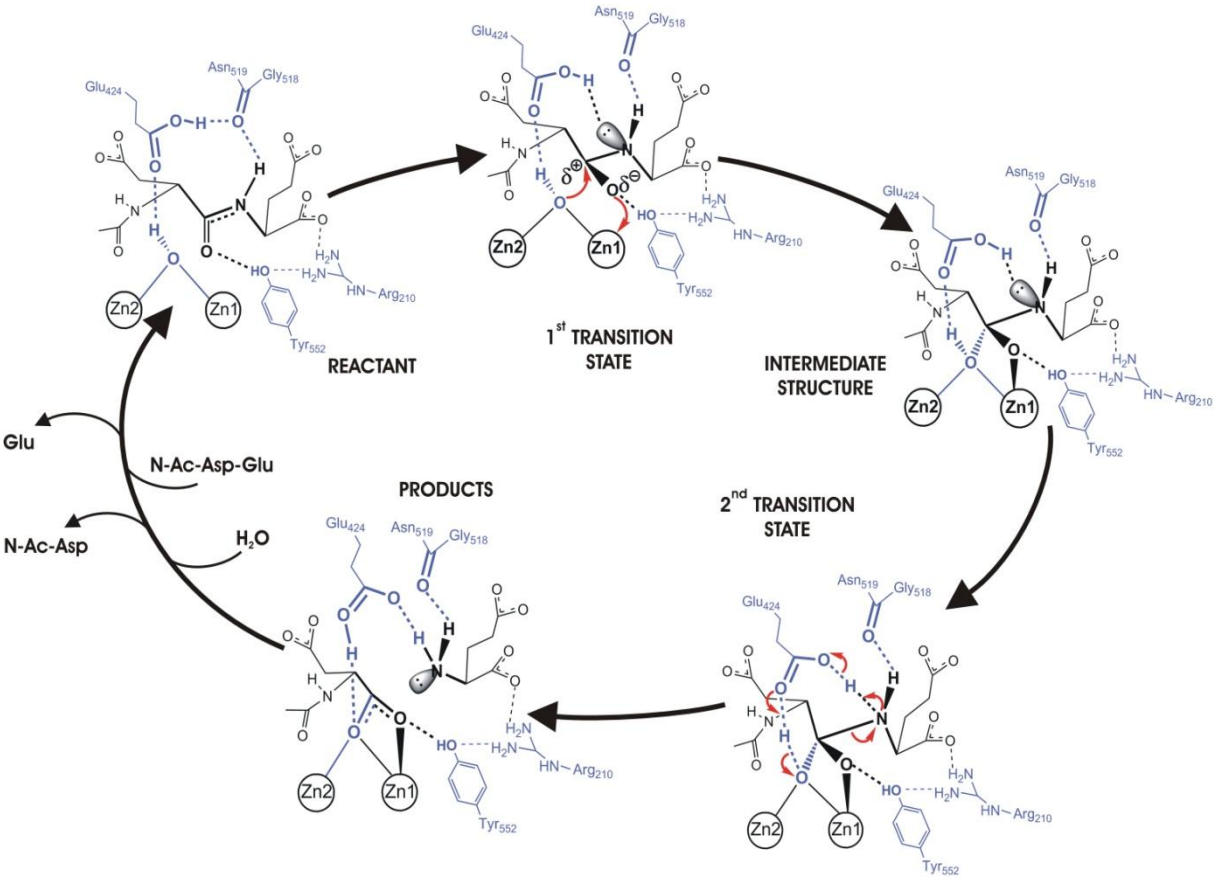


Figure 3.2.1.3: *The GCPII reaction cycle.* This cycle depicts important steps/structures throughout the reaction pathway. The suggested reaction mechanism was obtained by QM/MM calculations. It is fully consistent with the crystal structure of the substrate bound in the active site of the E424A

mutant of GCPII reported here as well as with the structures of the transition state analogue, and enzyme-product complexes reported previously [Mesters J. R. 2006; Bařinka, C. 2008]

In simple terms, we proposed to describe the *N*-Ac-Asp-Glu hydrolysis by GCPII as the reaction between a deactivated hydroxide with an activated peptide bond assisted by glutamate as a general acid/base. The fascinating and unique role of di-zinc center with regard to the reaction mechanism described is in making the nucleophile, the OH^- anion, available, but at the same time enabling control of its nucleophilicity. The course and rate of the reaction depends on proper orientation of the substrate and of the proton shuttle (Glu424) provided by the interactions with the enzyme active site.

Our best estimate of the reaction barrier was calculated to be $\Delta G^{\#} \approx 22$ kcal.mol^{-1}, which is in a good agreement with the experimentally observed reaction rate constant at 37ºC $k_{cat} = 1.15$ s^{-1} (which translates into $\Delta G^{\#} = 18.1$ kcal/mol).

We have also discussed the importance of the Gly518 carbonyl group stabilizing both the electrophilic hydroxyl group of Glu424 and the scissile peptide bond nitrogen in close proximity. We pointed out that this structural feature is conserved in other di-zinc metalopeptidases as well. Thus, we can hypothesize that the reaction mechanisms of these enzymes have a common pattern characterized by cooperation between a protonated glutamate residue, and a backbone carbonyl group of an enzyme. While the glutamate accommodates and shuttles the proton from the activated water molecule, the backbone carbonyl (Gly518 in the case of GCPII) is not only hydrogen bond acceptor for the protonated glutamate but at the same time also forms a hydrogen bond with the amide group of the scissile peptide bond of the substrate, assisting in a “productive orientation” of the proton shuttle and the substrate.

[5] Klusák, V., Bařinka, C., Plechanovová, A., Mlčochová, P., Konvalinka, J., Rulíšek, L., Lubkowski, J. (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry, 48*, 4126-4138; see Appendix E for the full paper

3.2.2. Significance of the GCPII active site residues, theoretical mutation study

As a follow-up of the above study of GCPII reaction mechanism an attempt has been made to contribute to elucidation of the reaction mechanism of several mutants of GCPII (E424A, N519D, and Y700F) [Navrátil V. 2009] and correlate the calculated activation barriers with the experimentally determined enzymatic rate constants (k_{cat}). [Mlčochová P. 2007]

Once the satisfactory correlation between the theoretical and experimental data is obtained it not only demonstrates the usefulness of the applied methods of computational chemistry for addressing the problems of biological relevance, but also provides us with further supporting evidence in favor of the recently proposed reaction mechanism. [Klusák V. 2009] An ultimate goal of these efforts is to develop theoretical chemistry methods to the level of accuracy so that they can be used as the predictive tools in 'rational' site-directed mutational studies. [Ishikita H. 2008]

Two main reasons can be mentioned why the results of the standard QM/MM approach [Senn H. M. 2009] may not be always reliable. First, it is the deficiency of the quantum chemical method used for electronic structure calculations. This is discussed in detail in chapter 3.2.3 of this thesis. Second, it is the need for a sampling of conformational landscape of an enzyme upon the enzymatic catalysis invoked by the fact that the enzymatic reaction is a dynamical process occurring in an ensemble of pathways leading from the reactant to the product valley. [Warshel A. 1991, Benkovic S. J. 2008]

The latter phenomenon is addressed poorly by standard QM/MM approaches or the so-called cluster model approach. [Siegbahn P.E.M. 2009] There are certainly more issues related to the accuracy of the computational approaches used for the description of the catalysis and for further discussion the reader is pointed to recent reviews. [Ryde U. 2003, Hu H. 2008, Ryde U. 2009, Harvey J.N. 2009] Altogether it shows that it is still very difficult to achieve a 'biochemical' accuracy of 1-2 kcal.mol^{-1} to be able to describe the difference of one order of magnitude in k_{cat} values (or various equilibrium constants).

In this study, using QM/MM modeling scheme, three mutants (Asn519Asp, Tyr700Phe, Glu424Ala) with single mutations in the vicinity of the active site of GCPII protein were studied and compared with the previously published experimental (kinetic) data (c.f. Table 3.2.2.1). [Mlčochová P. 2007] The mutants (Asn519Asp, Tyr700Phe, Glu424Ala) were selected so that their k_{cat} values for *N*-Ac-Asp-Glu hydrolysis were at least an order of magnitude lower than the k_{cat} value of the GCPII wild type (corresponding to the change in the Gibbs energy reaction barrier of 1.4 kcal.mol^{-1} at t=37 °C). Furthermore, the original and mutated residues were required to be in the vicinity of the active site of GCPII protein and were sterically similar and at the same time

representing two chemically distinct side chains.

We adopted an approximation that the activation barrier can be estimated as the difference between the energy of the tetrahedral intermediate and the Michaelis complex. It was assumed that this approximation is plausible because the previously studied transition states in the wild type GCPII reaction displayed close similarity, both energetic and geometric, to both transition states (TS1 and TS2, leading to the reactants and products, respectively).[Klusák V. 2009] Similar findings were also reported for AAPs [Chen S. 2008]. Following an attack of the bridging hydroxide moiety on the scissile peptide bond carbon (C_{pept}) a stable tetrahedral intermediate, accompanied by a change in the hybridization of C_{pept} and N_{pept} atoms from sp^2 to sp^3 is formed in all three mutants. The structures of all the intermediates were similar to the tetrahedral intermediate in the wild type GCPII.

The energy of the QM system also included an estimate of ZPVE and thermal correction to the enthalpy, thus yielding a QM/MM estimate of activation enthalpy. The calculated energy difference between the tetrahedral intermediate and the Michaelis complex, approximating the reaction activation barrier is for the GCPII wild type 25.3 kcal.mol^{-1}, compared to 27.7 kcal.mol^{-1} for Tyr700Phe mutant, 27.8 kcal.mol^{-1} for the Asn519Asp mutant, and 28.8 kcal.mol^{-1} for Glu424Ala mutant. [Navrátil V. 2009] (c.f. Table 3.2.2.1)

Therefore, a quantitative agreement between the differences in the calculated barriers and the differences in experimentally determined values of enzymatic rate constants, k_{cat} was obtained. Although the calculated QM/MM barriers are in all cases systematically shifted by ~10 kcal.mol^{-1} (being closer to 30 kcal.mol^{-1} rather than expected 20 kcal.mol^{-1} as can be inferred from the experimental k_{cat} values) we consider the agreement between the relative values as satisfactory. In our opinion, it *a posteriori* justifies the plausibility of the adopted approximations and supports the earlier proposed reaction mechanism of GCPII.

Table 3.2.2.1: Comparison of the reaction free energy barriers, $\Delta G^{\ddagger}$, corresponding to the experimental k_{cat} values, and the calculated estimate of the enthalpy barriers, $\Delta H^{\ddagger}$. [Mlčochová P. 2007, Navrátil V. 2009]

Mutation	k_{cat}(s^{-1})	$\Delta G^{\ddagger}$	$\Delta G^{\ddagger}$ - $\Delta G^{\ddagger}$(WT)	$\Delta H^{\ddagger}$	$\Delta H^{\ddagger}$ - $\Delta H^{\ddagger}$(WT)
WT	1.1 ± 0.2	18.1	0	25.3	0
Asn519Asp	0.078 ± 0.005	20.6	2.5	27.8	2.5
Tyr700Phe	0.075 ± 0.003	19.8	1.7	27.7	2.4
Glu424Ala	<0.001	>22.4	>4.3	28.8	3.5

3.2.3. Theoretical aspects of peptide hydrolysis in metallopeptidases [6]

This study was motivated by discrepancies between experimental and calculated quantities and by differences between various theoretical methods that can be used for the description of enzymatic reactions. In order to estimate the accuracy and deficiencies of the available theoretical approaches we attempted to provide a fairly broad comparison, including the performance of the methods, the size of the basis sets used and the size of the systems studied.

Most of the calculated data for metallopeptidases (or in general for metalloenzymes) are obtained using DFT methods. Typically, popular functionals such as B3LYP are used [Harvey J.N. 2009; Ryde U. 2009]. Since DFT methods offer the most favorable price/performance ratio, the question of their accuracy for a particular set of chemical processes is clearly a highly relevant question to ask in order to validate their performance for quantitative predictions.

It is generally accepted that for the systems with closed shell electronic structure there exists a hierarchy of standard *ab initio* quantum chemical methods that can be conveniently used for systems containing at least few tens of atoms. It starts from Hartree-Fock methods through popular second-order Møller-Plesset perturbation theory up to the coupled cluster methods. In most applications, CCSD(T) calculations using larger basis set are considered as a solid reference (benchmark) for the cheaper methods, including approaches based on DFT. [Hobza P. 2009]

The aim of the study is to critically address the accuracy of several (selected) quantum chemical methods on the model systems representing important intermediates and transition states in amide hydrolysis reactions. Concomitantly, we compare the un-catalyzed peptide bond (amide) hydrolysis of two model systems with the same reaction in the small cluster models representing the metalloenzyme active sites. Finally, we study the convergence of the calculated values with the increasing size of the system representing the enzyme active site.

Four model systems were used for the assessment of accuracy of quantum chemical methods. The set consisted of the models of metal-free (i) formamide and (ii) Ala-Ala hydrolysis, (iii) model of single metal (Zn^{2+}) assisted peptide hydrolysis based on thermolysin active site structure [Blumberger J. 2007] and (iv) model of the dizinc metallopeptidase active site (based on the GCPII transition state found earlier [Klusák 2009]). The reactants (Michaelis complexes) are depicted in Figure 3.2.3.1. In case of minimal models for thermolysine and GCPII, we attempted to include all essential features of the active site: metal ions, first-shell metal-binding residues (with imidazoles modeled by ammonia and carboxylates modeled by formate moieties), second-sphere glutamate residue acting as the “proton shuttle” (modeled by formate) and minimal model for the substrate. According to the described reaction mechanisms of the studied enzymes,[Blumberger J.

2007; Chen S. 2008; Klusák V. 2009] the truncated systems should posses all the principal physicochemical characteristics of the enzyme active sites.

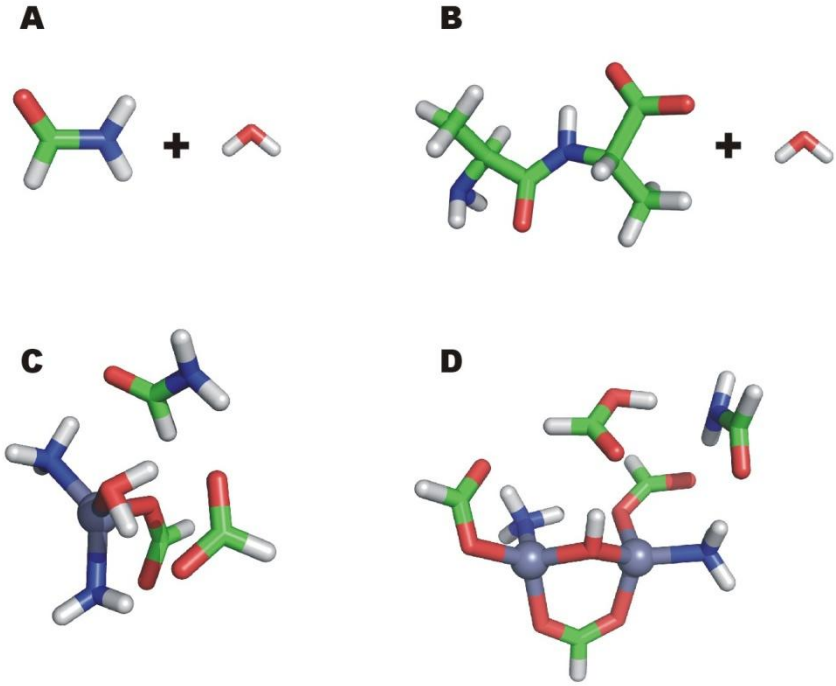


Figure 3.2.3.1: The structures of reactants (Michaelis complexes) of the studied systems: (A) the hydrolysis of formamide (in water-assisted mechanism two water molecules are engaged), (B) the hydrolysis of Ala-Ala dipeptide, (C) hydrolysis of formamid in the thermolysine–like active site, and (D) hydrolysis of formamide in the GCPII–like active site.

For all these fully optimized models we have obtained the transition states by following the imaginary vibration mode corresponding to the "reactive coordinate(s)".

The height of barrier in case of thermolysin model is remarkable. It is rather surprising that the 'catalyzed' reaction barrier is by one third higher than in the case of uncatalyzed process. However, it should be noted that the conformational space has not been extensively inspected and that the structures on this reaction coordinate can be influenced by the unrestrained optimization applied for the purpose of the transition state search (not closely corresponding to the reaction in protein). Apparently, it is the protein scaffold that drives the reaction through the lower barrier pathway. More thorough study would be further necessary to confirm this assumption. However, for the purpose of this work which is a method comparison we consider the presented structures sufficient. The key feature of our comparison is correctly described *in vacuo* reaction pathway and the true nature of the transition states.

The resulting comparison of method performance, described in detail in the attached

manuscript and illustrated in the Table 1, shows that from *ab initio* methods the (RI-)MP2 and SCS-MP2 perform reasonably well and can be recommended for the modeling of the medium sized closed shell systems including zinc ions. Regarding DFT, the studied non-hybrid functionals underestimate the barriers whereas the B3LYP exhibit an overestimation of the barriers. For the def2-TZVP basis set in the case of formamide and thermolysin it is only a slight overestimation, by 0.4 and 1.2 kcal/mol, respectively. However, in case of Ala-Ala hydrolysis and GCPII model system the barrier is overestimated by 3.6 kcal/mol and 3.5 kcal/mol, respectively (it can be noted that in case of GCPII only CCSD(T)/TZVP calculations were feasible). (see Table 3.2.3.1)

It was also shown that one probably cannot expect better than 3-4 kcal.mol^{-1} accuracy in the calculated TS barriers in studying the systems of realistic size (50-200 atoms which is the typical size of the cluster or quantum core in e.g. QM/MM modeling of enzymatic reactions).

It is interesting to mention the results of DFT-D calculations. The empirical term does not yield such a reliable and systematic correction to the DFT values so that they can approach the CCSD(T) benchmark values. In all but one case (thermolysin) the correction was lowering the barrier by 2.5 to 5.5 kcal/mol. This had further deteriorated the performance of non-hybrid functionals (PBE, TPSS). Only in case of the GCPII activation energy and the B3LYP method the empirical correction almost canceled out with the error in B3LYP method. On the contrary, for thermolysin the empirical correction lead to even greater overestimation of the barrier in B3LYP method. (see Table 3.2.3.1)

This unreliable behavior is not surprising once we realize that the dispersion correction is parameterized for medium size systems and their non-covalent interactions in the electronic ground state and equilibrium geometry [Grimme S. 2004, Jurečka P. 2007, Zimmerli U. 2004]. It is not guaranteed to perform equally well for the saddle points (transition state geometries).

[6] Klusák, V., Navrátil, V., Rulíšek, L. (2010) Theoretical Aspects of Peptide Bond Hydrolysis by Zinc Metallopeptidases, *submitted for publication*; see Appendix F for the full paper

Table 3.2.3.1: Comparison of activation energies ($\Delta E^{\neq}$) for the RDS of hydrolytic reactions in the four model systems studied. The basis set used was def2-TZVP with the exception of GCPII model where TZVP basis set was used.

System		CCSD(T)	CC2	MP2	SCS-MP2	PBE	PBE-D	TPSS	TPSS-D	B3LYP	B3LYP-D
formamide	$\Delta E^{\neq}$	**28.7**	22.7	25.8	31.5	17.2	13.8	21.5	17.0	29.1	24.4
Ala-Ala	$\Delta E^{\neq}$	**28.2**	23.1	26.6	30.2	23.7	19.8	25.8	20.6	31.8	26.3
thermolysin	$\Delta E^{\neq}$**(TS1)**	**39.4**	37.7	38.0	39.7	35.4	36.5	35.7	37.1	40.6	42.2
GCPII	$\Delta E^{\neq}$**(TS1)**	**20.8**	16.8	20.3	23.3	17.6	15.2	17.1	14.0	24.5	21.2

4. Concluding Remarks

The aim of the presented thesis was to extend our understanding of molecular properties, especially non-covalent interactions and to improve interpretation of their role in various biological processes. Using several examples of biologically relevant systems, the essential role of quantum mechanics and theoretical chemistry in our understanding of the biomolecule structure stabilization and structural specificity has been illustrated.

In the first part of the thesis we demonstrated the importance of dispersion energy for an accurate description of selected non-covalent interactions of aromatic residues, both in inter-molecular, protein – ligand, interaction, and intra-molecular interactions inside a protein core. It has been shown that such interactions are competitive and sometimes can prevail over the hydrogen bonding. Concomitantly, it has been shown that an extreme care must be taken when modeling larger molecular systems. One of the phenomena that is not present in the small molecular systems but can dramatically influence the description of larger molecules is the intra-molecular basis set superposition error (BSSE). Therefore, a good method for the theoretical description of biomolecules and nanomaterials should account for the dispersion energy and must avoid the BSSE at the same time, which is a fairly strict requirement. One such candidate and practical solution to many of these problems is the DFT-D method proposed recently and shown to provide very good results for the geometry optimized structures.

A different situation and methodological requirements are encountered in the computational modeling of enzymatic reactions. While quantum chemistry is essentially the only approach that can provide us with the mechanistic details of enzymatic reactions, an extreme care must be taken when calculating and interpreting the theoretical transition state (TS) barriers. Various quantum chemical methods (including the highly correlated methods) can fail to reproduce the activation barriers within the desired 2 kcal/mol accuracy. For example, DFT-D method cannot be considered as the method of choice not even for the case of the closed shell systems. On the other hand the standard DFT methods, lacking proper dispersion description, performed reasonably well. It is a matter of debate whether the missing component in TS structure energy description is the dispersion energy.

In conclusion, it was shown that an accurate modeling of biological processes is a complex problem and it is not always possible to *a priori* neglect the role of the more distant or weekly bound components (residues), such as the protein back bone (Gly518) in the case of GCPII hydrolysis. The same holds true in the stabilization of biomolecular systems, such as the studied (and not anticipated) remarkable attraction between the protein back bone and aromatic residue side chain in the model of rubredoxin hydrophobic core.

List of Publications

Klusák V., Havlas Z., Rulíšek L., Vondrášek J., Svatoš A. (2003) Sexual attraction in the silkworm moth: Nature of binding of bombykol in the pheromone binding protein – an ab initio study. *Chem. Biol. 10*, 331-340.

Vondrášek J., Bendová L., Klusák V., Hobza P. (2005) Unexpectedly strong energy stabilization inside the hydrophobic core of small protein rubredoxin mediated by aromatic residues: correlated ab initio quantum chemical calculations. *J. Am. Chem. Soc. 127*, 2615-2619.

Hlouchová, K., Bařinka, C., Klusák, V., Šácha, P., Mlčochová, P., Majer, P., Rulíšek, L., Konvalinka, J. (2007) Biochemical characterization of human glutamate carboxypeptidase III. *J. Neurochem. 101*, 682-696.

Valdés H., Klusák V., Pitoňák M., Exner O., Starý I., Hobza P., Rulíšek L. (2007) Evaluation of the intramolecular basis set superposition error in the calculations of larger molecules: [n]helicenes and Phe-Gly-Phe tripeptide. *J. Comput. Chem. 29*, 861-870.

Mansurova M., Klusák V., Nešněrová P., Muck A., Doubský J., Svatoš A. (2009) Design and synthesis of bombykol analogues for probing pheromone-binding protein–ligand interactions. *Tetrahedron 65*, 1069-1076.

Klusák, V., Bařinka, C., Plechanovová, A., Mlčochová, P., Konvalinka, J., Rulíšek, L., Lubkowski, J. (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry, 48*, 4126-4138.

References

Ahlrichs, R., Bar, M., Haser, M., Horn, H. and Kolmel, C. (1989) Electronic-Structure Calculations on Workstation Computers - the Program System Turbomole. *Chemical Physics Letters 162*, 165-169.

Allen, M. J., Tozer, D. J. (2002) Helium dimer dispersion forces and correlation potentials in density functional theory. *J. Chem. Phys. 117*, 11113-11120.

Artymiuk P.J., Grindley A.M., Park J.E., Rice D.W., Willett P. (1992) Three-dimensional structural resemblance between Leucine Aminopeptidase and Carboxypeptidase-A revealed by graph-theoretical techniques. *FEBS Lett*. *303*, 48-52.

Baker, E.N., Hubbard, R.E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol. 44*, 97-179.

Baldwin, R.L. (2002) Making a network of hydrophobic clusters. *Science 295*, 1957-1958.

Bařinka, C., Hlouchová, K., Rovenská, M., Majer, P., Dauter, M., Hin, N., Ko, Y., Tsukamoto, T., Slusher, B. S., Konvalinka, J., and Lubkowski, J. (2008) Structural basis of interactions between human glutamate carboxypeptidase II and its substrate analogs. *J. Mol. Biol. 376*, 1438-1450.

Bařinka, C., Rovenská, M., Mlčochová, P., Hlouchová, K., Plechanovová, A., Majer, P., Tsukamoto, T., Slusher, B. S., Konvalinka, J., and Lubkowski, J. (2007b) Structural insight into the pharmacophore pocket of human glutamate carboxypeptidase II. *J. Med. Chem. 50*, 3267-3273.

Bařinka, C., Starková, J., Konvalinka, J., and Lubkowski, J. (2007a) A high-resolution structure of ligand-free human glutamate carboxypeptidase II. Acta Crystallogr., Sect. F: *Struct. Biol. Cryst. Commun. 63*, 150-153.

Bayliss, M. E., Prescott, J. M. (1986) Modified activity of Aeromonas aminopeptidase: metal ion substitutions and role of substrates. *Biochemistry 25*, 8113-8117.

Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A 38*, 3098-3100.

Becke, A. D. (1993) Density-functional thermochemistry. iii. the role of exact exchange, *J. Chem. Phys. 98*, 5648-5652.

Becke, A. D., Johnson, E. R. (2005) A density-functional model of the dispersion interaction. *J. Chem. Phys. 123*, 154101.

Benkovic, S. J., Hammes, G. G. Hammes-Schiffer, S. (2008) Free-Energy Landscape of Enzyme Catalysis. *Biochemistry 47*, 3317-3321.

Ben-Meir, D., Spungin, A., Ashkenazi, R., Blumberg, S. (1993) Specificity of *Streptomyces griseus* aminopeptidase and modulation of activity by divalent metal ion binding and substitution. *Eur. J. Biochem. 212*, 107-112.

Berman, H.M., Henrick, K., Nakamura, H. (2003) Announcing the worldwide Protein Data Bank. *Nature Structural Biology 10*, 980-980.

Berman, H.M., Westbrook, J., Feng Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Research, 28*, 235-242.

Bertini, I., Luchinat, C., Rosi, M., Sgamellotti, A., Tarantelli, F. (1990) pKa of zinc-bound water and nucleophilicity of hydroxo-containing species. Ab initio calculations on models for zinc-enzymes. *Inorg.Chem. 29*, 1460-1463.

Blumberger, J., Lamoureux, G., Klein, M. L. (2007) Peptide hydrolysis in thermolysin: Ab initio QM/MM investigation of the Glu143-assisted water addition mechanism. *Journal of Chemical Theory and Computation 3*, 1837-1850.

Boys, S. F., Bernardi, F. (1970) The calculations of small molecular interaction by the difference of separate total energies. Some procedures with reduced error. *Mol Phys, 19*, 553-566.

Brandl, M.,Weiss, M. S., Jabs, A., Sühnel, J., Hilgenfeld, R. (2001) C-H···π-Interactions in Proteins. *J. Mol. Biol. 307*, 357-377.

Burley, S.K., Pesko, G.A. (1985) Aromatic-Aromatic interaction: A mechanism of protein structure stabilisation. *Science 229*, 23-28.

Butenandt, A., Beckmann, R., Stamm, D., and Hevker, E. (1959) On the sex pheromone of the silkworm moth Bombyx mori. Isolation and structure. *Z. Naturforsch. B 14*, 283-284.

Campanacci, V., Krieger, J., Bette, S., Sturgis, J. N., Lartigue, A., Cambillau, C., Breer, H., Tegoni, M. (2001) Revisiting the specificity of Mamestra brassicae and Antheraea polyphemus pheromone-binding proteins with a fluorescence binding assay. *J. Biol. Chem. 276*, 20078-20084.

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B. and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668-1688.

Cauwe, B., Van den Steen, P. E., Opdenakker, G. (2007) The biochemical, biological, and pathological kaleidoscope of cell surface substrates processed by matrix metalloproteinases. *Crit. Rev. Biochem. Mol. Biol. 42*, 113-185.

Chandler, D. (1987) Introduction to Modern Statistical Mechanics, *Oxford University Press*.

Chandler, D. (2005) Interfaces and the driving force of hydrophobic assembly, *Nature 437*, 640-647.

Chen, S., Marino, T., Fang, W., Russo, N., Himo, F. (2008) Peptide Hydrolysis by the Binuclear Zinc Enzyme Aminopeptidase from *Aeromonas proteolytica*: A Density Functional Theory Study. *J. Phys. Chem. B 112*, 2494-2500.

Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M., Moras, D., Tarnus, C. (1994) Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure 2*, 283-291.

Cooper, A. (1999) Thermodynamic analysis of biomolecular interactions. *Curr. Opin. Cell Biol. 3*, 557-563.

Copeland, R. A. (2000) Enzymes: A Practical Introduction to Structure, Mechanism and Data analysis., *Wiley-VCH*.

Cossi, M., Rega, N., Scalmani, G., Barone, V. (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem. 24*, 669-681.

Dabkowska, I., Valdés-Gonzalez, H., Jurečka, P., Hobza, P. (2005) Stabilization energies of the hydrogen-bonded and stacked structures of nucleic acid base pairs in the crystal geometries of CG, AT, and AC DNA steps and in the NMR geometry of the 5'-d(GCGAAGC)-3' hairpin: complete basis set calculations at the MP2 and CCSD(T) levels. *J. Phys. Chem. A, 109*, 1131-1136.

Daniel, R. M., Cowan, R. M. (2000) Biomolecular stability and life at high temperatures. *Cell. Mol. Life. Sci. 57*, 250-264.

Davis, A.M., Teague, S.J. (1999) Hydrogen bonding, Hydrophobic intraction, and failure of the rigid receptor hypothesis. *Angewandte Chem. Int. ed. 38*, 736-749.

Davis, M. I., Bennett, M. J., Thomas, L. M., and Bjorkman, P. J. (2005) Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase. *Proc. Natl. Acad. Sci. U.S.A. 102*, 5981-5986.

Derewenda, Z.S., Lee, L., Derewenda, U. (1995) The ocurence of C-H□□□O hydrogen bonds in proteins. *J.Mol.Biol. 252*, 248-262.

Desiraju, G.R., Steiner, T. (2001) The weak hydrogen bond in structural chemistry and biology. *Oxford University Press*, New York.

Desmarais, W., Bienvenue, D. L., Bzymek, K. P., Petsko, G. A., Ringe, D., Holz, R. C. (2006) The high-resolution structures of the neutral and the low pH crystals of aminopeptidase from Aeromonas proteolytica. *J. Biol. Inorg. Chem. 11*, 398-408.

Dill, A.K., Ozkan, S.B., Shell, S.M., Weikl, T.R. (2008) The Protein Folding Problem. *Annual review of biophysics 37*, 289-316.

Dunning, T. H. Jr. (2000) A Road Map for the Calculation of Molecular Binding Energies *J. Phys.*

Chem. A, 104, 9062-9080.

Eichkorn, K., Treutler, O., Ohm, H., Haser, M., Ahlrichs, R. (1995) Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett. 240,* 283-289.

Eichkorn, K., Weigen, F., Treutler, O., Ahlrichs, R. (1997) Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chim. Acta 97,* 119-124.

Elstner, M., Hobza, P., Frauenheim, T., Suhai, S., Kaxiras, E. (2001) Hydrogen bonding and stacking interactions of nucleic acid base pairs: a density-functional-theory based treatment. *J. Chem. Phys. 114,* 5149-5155.

Ercan, A., Park H. I., Ming L. (2000) Remarkable enhancement of the hydrolyses of phosphoesters by dinuclear centers: Streptomyces aminopeptidase as a natural model system. *Chem. Commun. 24,* 2501-2502.

Erhardt, S., Jaime, E., Weston, J., (2005) A water sluice is generated in the active site of bovine lens leucine aminopeptidase. *J. Am. Chem. Soc. 127,* 3654-3655.

Evans, M.G., Polanyi, M. (1935). "Some applications of the transition state method to the calculation of reaction velocities, especially in solution". *Trans. Faraday Soc. 31,* 875-894.

Eyring, H., Polanyi, M. (1931) Über einfache Gasreaktionen. *Z. Phys. Chem. Abt. B* 12: 279-311.

Eyring, H. (1935). The Activated Complex in Chemical Reactions. *J. Chem. Phys. 3,* 107-115.

Fast, P. L., Sanchez, M. L., Truhlar, D. G. (1999) Infinite basis limits in electronic structure theory. *J. Chem.Phys. 111,* 2921-2926.

Fersht, A.R., ShiJ, P., Knill-Jones, J., Lowe, D.M., Wilkinson, A.J., Blow, D.M., Brick, P., Carter, P.,Wayne, M.M.Y., Winter, G. (1985) Hydrogen-bonding and biological specificity analyzed by protein engineering. *Nature 314,* 235-238.

Feyereisen, M.W., Feller, D., Dixon, D.A. (1996) Hydrogen bond energy of water dimer. *J. Phys. Chem. 100,* 2993-2997.

Feyereisen, M.W., Fitzgerald, G., Komornicki, A. (1993) Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Letters 208,* 359.

Fischer, E. (1894) Einfluβ der Konfiguration auf die Wirkung der Enzyme. *Ber. Ges. Dtsch. Chem.* ***27*** , 2984-2993.

Gerber, I. C., Angyan, J. G. (2007) London dispersion forces by range separated hybrid density functional with second order perturbational corrections (RSH+MP2): the case of rare gas complexes. *J. Chem. Phys. 126,* 044103.

Gorter, E., Grendel, F. (1925) On biomolecular layers of lipoids on the chromocytes of the blood. *J. Exp. Med. 41,* 439-443.

Gorter, E., Grendel, F.(1925) On bimolecular layers of lipoids on the chromocytes of the blood. *J. Exp. Med. 41,* 439-443.

Greenblatt, H. M., Almog, O., Maras, B., Spungin-Bialik, A., Barra, D., Blumberg, S., and Shoham, G. (1997) Streptomyces griseus aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution. *J. Mol. Biol. 265,* 620-636.

Grimme, S. (2004) Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem. 25*, 1463-1476.

Grimme, S. (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem. 27*, 1787-1799.

Grimme, S., Antony, J., Schwabe, T., Mück-Lichtenfeld, C. (2007) Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio)organic molecules. *Org. Biomol. Chem. 5*, 741-758.

Han, Y., Kim, K. H., Son, S., Lee, Y. S. (2002) Effects of intramolecular basis set superposition error on conformational energy difference of 1,2-difluoroethane and 1,2-dimethoxyethane. *Bull. Korean Chem. Soc. 23,* 1267-1271.

Hasselgren, C.,Park, H. I., Ming, L. J. (2001) Metal ion binding and activation of *Streptomyces griseus* dinuclear aminopeptidase: cadmium(II) binding as a model. *J. Biol. Inorg. Chem. 6,* 120-127.

Harding, M. M. (1999) The geometry of metal–ligand interactions relevant to proteins. *Acta Crystallogr. Sect. D-Biol. Crystallogr. 55*, 1432-1443.

Hershcovitz, Y.F., Gilboa, R., Reiland, V., Shoham, G., Shoham, Y. (2007) Catalytic mechanism of SGAP, double-zinc aminopeptidase from *Streptomyces griseus. The FEBS Journal 274,* 3864-3876.

Hertwig, R. H., Koch, W. (1997) On the parameterization of the local correlation functional: What is Becke-3-LYP? *Chem. Phys. Lett. 268,* 345-351.

Harvey, J. N., Mujika, J. I., Mulholland A. J. (2009) "Computational Methods: Modeling of Reactivity in Zn-Containing Enzymes", in Computational Inorganic and Bioinorganic Chemistry, E. I. Solomon, R. B. King & R. A. Scott, eds. J. Wiley & Sons, Ltd. Chichester, UK, pp. 343–352.

Hobza, P., Müller-Dethlefs, K. (2009) Non-Covalent Interactions: Theory and Experiment. *Royal Society of Chemistry*, Cambridge.

Hobza, P., Selzle, H.L., Schlag, *E.W. (1996) Potentia*l Energy Surface of the Benzen Dimer: Ab Initio Theoretical Study. *J. Phys. Chem. 100*, 18790-18794.

Hobza, P., Šponer, J. (2002) Toward True DNA Base-Stacking Energies: MP2, CCSD(T), and

Complete Basis Set Calculations. *J. Am. Chem. Soc. 124*, 11802-11808.

Hobza, P., Šponer, J., Reschel, T. (1995) Density functional theory and molecular clusters. *J. Comput. Chem. 16*, 1315-1325.

Hobza, P., Zahradnik, R. (1980) *Weak Intermolecular Interactions in Chemistry* and Biology. *Elsevier*, Amsterdam.

Hobza, P., Zahradník, R. (1988) Intermolecular interactions between medium-sized systems. Nonempirical and empirical calculations of interaction energies. Successes and failures. *Chem Rev, 88*, 871-897.

Hohenberg, P., Kohn, W. (1964) Inhomogeneous Electron Gas. *Physical Review B 136*, B864-B871.

Holz, R. C. (2002) The aminopeptidase from *Aeromonas proteolytica*: structure and mechanism of co-catalytic metal centers involved in peptide hydrolysis. *Coord. Chem. Rev. 232*, 5-26.

Holz, R.C., Bzymek, K.P., Swierczek, S.I. (2003) Co-catalytic metallopeptidases as pharmaceutical targets. *Curr. Opin. Chem. Biol. 7*, 197-206.

Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L., Leal, W.S., Wuthrich, K. (2001). NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl. Acad. Sci. USA 98*, 14374-14379.

Hu, H., Yang, W. T. (2008) Free Energies of Chemical Reactions in Solution and in Enzymes with *Ab Initio* Quantum Mechanics/Molecular Mechanics Methods. *Annu. Rev. Phys. Chem. 59*, 573-601.

Ippolito, J.A., Alexander, R.S., Christianson, D.W. (1990) Hydrogen bond stereochemistry in protein structure and function. *J. Mol. Biol. 215*, 457-471.

Ishikita, H., Warshel, A. (2008) Predicting Drug-Resistant Mutations of HIV Protease. *Angew. Chem. Int. Ed. 47*, 697-700.

Jaime, E., Kluge, S., Weston, J. (2007) On the origin of the broad-bend selectivity of bovine-lens-leucine-aminopeptidase. *ARKIVOC 3*, 77-95.

Jansen, H. B., Ros, P. (1969) Non-empirical molecular orbital calculations on the protonation of carbon monoxide. *Chem. Phys. Lett. 3*, 140-143.

Jeffrey, G.A., Sanger, W. (1991) Hydrogen bonding in biological systems, *Springer-Verlag*, Berlin,.

Jensen, F. (1996) The magnitude of intramolecular basis set superposition error. *Chem. Phys. Lett. 261*, 633-636.

Jensen, F. (1999) Introduction to Computational Chemistry. *John Wiley & Sons*, New York.

Jensen F. (2010) An Atomic Counterpoise Method for Estimating Inter- and Intramolecular Basis Set Superposition Errors. *J.Chem.Theory Comput.* 6, 100-106.

Jorgensen, W.L. (1991) Rusting of the lock and key model for protein-ligand binding. *Science 254*, 954-955.

Jurečka, P., Nachtigall, P., Hobza, P. (2001) RI-MP2 calculations with extended basis sets—a promising tool for study of H-bonded and stacked DNA base pairs. *Phys. Chem. Chem. Phys. 3*, 4578-4582.

Jurečka, P., Hobza, P. (2003) True Stabilization Energies for the Optimal Planar Hydrogen-Bonded and Stacked Structures of Guanine···Cytosine, Adenine···Thymine, and Their 9- and 1-Methyl Derivatives: Complete Basis Set Calculations at the MP2 and CCSD(T) Levels and Comparison with Experiment. *J. Am. Chem. Soc. 125*, 15608-15613.

Jurečka, P., Černý, J., Hobza, P., Salahub, D. R. (2007) Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *J. Comput. Chem. 28*, 555-569.

Kaissling, K.E. (2001). Olfactory perireceptor and receptor events in moths: A kinetic model. *Chem. Senses 26*, 125-150.

Kaissling, K.E., and Priesner, E. (1970) Smell threshold of the silkworm. *Naturwissenschaften 57*, 23-28.

Kasang, G., Kaissling, K.E., Vostrowsky, O., and Bestmann, J. (1978a) Bombykal the second component of the Bombyx mori L. silkworm pheromone. *Angew. Chem. 90*, 74-75.

Kasang, G., Schneider, D., and Schafer, W. (1978b) The silkworm moth Bombyx mori. Presence of the (E,E)-Stereoisomer of Bombykol in the female pheromone gland. *Naturwissenschaften 65*, 337-338.

Kauzmann, W., (1954) In: The mechanism of enzyme action. Eds.: McElroy, W., Glass E. *The John Hopkins Press*, Baltimore, pp 70-110.

Kendall, R. A., Dunning, T. H., Jr., Harrison, R. J. (1992) Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys. 96*, 6796-6806.

Kendrew, J.C., Bodo, G., Dintzis, H. M., Parrish, R.G.,Wyckof, H., Phillips, D.C. (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature 181*, 662-666.

Kinoshita, Y., Kuratsukuri, K., Landas, S., Imaida, K., Rovito, P. M.Jr., Wang, C. Y., and Haas, G. P. (2006) Expression of prostate-specific membrane antigen in normal and malignant human tissues. *World J. Surg. 30*, 628-636.

Klamt, A., Schüürmann, G. (1993) 'COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc.-Perkin*

Trans. 2, 799-805.

Klamt, A., Schuurmann, G. (1993) Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc.Perkin. Trans. 2,* 799-805.

Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., Schwalbe, H. *(2002)* Long-range interactions within a nonnative protein. *Science 295*, 1719-1722.

Klusák V., Havlas Z., Rulíšek L., Vondrášek J., Svatoš A. (2003) Sexual attraction in the silkwork moth: Nature of binding of bombykol in the pehromone binding protein – an ab initio study. *Chem. Biol. 10*, 331-340.

Klusák, V., Bařinka, C., Plechanovová, A., Mlčochová, P., Konvalinka, J., Rulíšek, L., Lubkowski, J. (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry 48,* 4126-4138.

Koch, W. and Holthausen, M. C. (2001) A Chemist's Guide to Density Functional Theory, *Wiley-VCH*.

Kollman, P., McKelvey, J., Johansson, A., Rothenberg, S. (1975) Theoretical studi*es of hyd*rogen-bonded dimers. Complexes ivolving HF, H_2O, NH_3, HCl, H_2S, Ph_3, HCN, HNC, HCP, CH_2NH, H_2CS, H_2CO, CH_4, CF_3H, C_2H_2, C_2H_4, C_6H_6, F^- and H_3O^+. *J. Am. Chem. Soc. 97*, 955-965.

Kožíšek, M., Bray, J., Řezácová, P., Šašková, K., Brynda, J., Pokorná, J., Mammano, F., Rulíšek, L. and Konvalinka, J. (2007) Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *Journal of Molecular Biology* **374**(4): 1005-1016.

Lee, C., Yang, W., Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B, 37*, 785-789.

Leal, W. S., Chen, A. M., Ishida, Y., Chiang, V. P., Erickson, M. L., Morgan, T. I., Tsuruda, J. M. (2005) Kinetics and molecular properties of pheromone binding and release. *Proc. Natl. Acad. Sci. 102*, 5386-5391.

Liao, R. Z. (2009). Mechanistic Insights into Dinuclear Zinc Enzymes from Density Functional Studies. *Licentiate Thesis, Department of Theoretical Chemistry.* Royal Institute of Technology, Stockholm.

Lichtenthaler, F. W. (1994) "100 years schlussel-schloss-prinzip": What made Emil Fischer use this analogy. *Angev. Chem. Int. Ed. Engl. 33*, 2364-2374.

Liu, B., McLean, A. D. (1973) Accurate calculation of the attractive interaction of two ground state

helium atoms. *J. Chem. Phys, 59*, 4557-4558.

Lowther, W. T., Matthews, B. W. (2002) Metalloaminopeptidases: Common Functional Themes in Disparate Structural Surroundings. *Chem. Rev. 102*, 4581-4608.

Maseras, M., Morokuma, K. (1995) IMOMM—a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comp. Chem. 16*, 1170–1179.

Maida, R., Steinbrecht, A., Ziegelberger, G., and Pelosi, P. (1993) The pheromone binding protein of Bombyx mori: purification, characterization and immunocytochemical localization. *Insect Biochem. Mol. Biol. 23*, 243-253.

Martin, T., Obst, U., Rebek, J., Jr. (1998) Molecular assembly and encapsulation directed by hydrogen-bonding preferences and the filling of space. *Science 281*, 1842-1845.

McQuarrie, D. (2000) Statistical Mechanics, *University Science books*.

Mesters, J. R., Bařinka, C., Li, W., Tsukamoto, T., Majer, P., Slusher, B. S., Konvalinka, J., and Hilgenfeld, R. (2006) Structure of glutamate carboxypeptidase II, a drug target in neuronal damage and prostate cancer. *EMBO J. 25*, 1375-1384.

Meyer, E. A., Castellano, R. K., Diederich, F. (2003) Interactions with Aromatic Rings in Chemical and Biological Recognition. *Angew. Chem., Int. Ed. 42*, 1210-1250.

Meyer, H.(1899) Zur Theorie der Alkoholnarkose, Erste Mittheilung. Welche Eigenschaft der Anästhetica bedingt ihre narkotische Wirkung? *Arch. Exptl. Pathol. Pharmakol. 42*, 109-118.

Mlčochová, P., Plechanovová, A., Bařinka, C., Mahadevan, D., Saldanha, J. W., Rulíšek, L., and Konvalinka, J. (2007) Mapping of the active site of glutamate carboxypeptidase II by site-directed mutagenesis. *FEBS J. 274*, 4731-4741.

Monzingo, A. F., and Matthews, B. W. (1984) Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by X-ray crystallography: a novel class of transition-state analogs for zinc peptidases. *Biochemistry 23*, 5724-5729.

Müller-Dethlefs, K., Hobza, P. (2000) Noncovalent interactions: a challenge for experiment and theory. *Chem. Rev. 100*, 143-168.

Navrátil, V. (2009) Studium variantních forem lidské glutamátkarboxypeptidasy II kombinací experimentálních a teoretických metod. *Diplomová práce* Přírodovědecká fakulta UK, Praha.

Nishio, M., Hirota, M., Umezava, Y. (1998) The CH/π interaction; evidence, nature and consequences, *Wiley-VCH*, New York.

Nelson, D. L. and Cox, M. M. (2004) Principles of Biochemistry, *W. H. Freeman*.

Oldham, N.J., Krieger, J., Breer, H., Fischedick, A., Hoskovec, M., and Svatoš, A. (2000) Analysis of the silkworm moth pheromone binding protein-pheromone complex by electrospray-

ionization mass spectrometry. *Angew. Chem. Int. Ed. Engl. 39*, 4341-4343.

Oldham, N.J., Krieger, J., Breer, H., Fischedick, A., Hoskovec, M., and Svatos, A. (2001). Detection and removal of an artefact fatty acid from the binding site of recombinant Bombyx mori pheromone-binding protein. *Chem. Senses 26*, 529-531.

Overton, E.(1901) Studien über die Narkose, *Fischer*, Jena.

Papir, G., Spungin-Bialik, A., Ben-Beir, D., Fudim, E., Gilboa, R., Greenblatt, H. M., Shoham, G., Lessel, U., Schomburg, D., Ashkenazi, R., Blumberg, S. (1998) Inhibition of *Streptomyces griseus* aminopeptidase and effects of calcium ions on catalysis and binding : Comparisons with the homologous enzyme Aeromonas proteolytica aminopeptidase. *Eur. J. Biochem. 258*, 313-319.

Park, H. I., Ming, L. J. (1999) A 10^{10} Rate Enhancement of Phosphodiester Hydrolysis by a Dinuclear Aminopeptidase - Transition-State Analogues as Substrates? *Angew. Chem., Int. Ed. Engl. 38*, 2914-2916.

Parr, R. G. and Yang, W. (1994) Density-Functional Theory of Atoms and Molecules, *Oxford University Press*.

Perdew, J. P., Burke, K., Ernzerhof, M. (1996) Generalized gradient approximation made simple. *Phys. Rev. Lett. 77*, 3865-3868.

Perozzo, R., Jelesarov, I., Bosshard, H. R., Folkers, G. and Scapozza, L. (2000) Compulsory order of substrate binding to herpes simplex virus type 1 thymidine kinase - A calorimetric study. *Journal of Biological Chemistry 275***,** 16139-16145.

Piacenza, M., Grimme, S. (2005) Van der Waals interactions in aromatic systems: structure and energetics of dimers and trimers of pyridine. *Chem. Phys. Chem. 6*, 1554-1558.

Pitoňák, M., Riley, K. E., Neogrády, P., Hobza, P. (2008) Highly Accurate CCSD(T) and DFT-SAPT Stabilization Energies of H-Bonded and Stacked Structures of the Uracil Dimer. *Chem. Phys. Chem. 9*, 1636-1644.

Rappé, A. K., Bernstein, E. R. (2000) Ab initio calculation of nonbonded interactions: Are we there yet? *J. Phys. Chem. A 104*, 6117-6128.

Rawlings, N. D., and Barrett, A. J. (1997) Structure of membrane glutamate carboxypeptidase. *Biochim. Biophys. Acta 1339*, 247-252.

Reuter, NI., Dejaegere, A., Maigret, B., Karplus, M. (2000) Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. 104*, 1720-1735.

Robinson, M. B., Blakely, R. D., Couto, R., and Coyle, J. T. (1987) Hydrolysis of the brain dipeptide N-acetyl-L-aspartyl-L-glutamate. Identification and characterization of a novel N-

acetylated alpha-linked acidic dipeptidase activity from rat brain. *J. Biol. Chem. 262*, 14498-14506.

Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., Zehfus, M. H. (1985) Hydrophobicity of amino acid residues in globular proteins. *Science 229*, 834-438.

Ryde, U. (1996) The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum chemical and molecular mechanical calculations. *J. Comput.-Aided Mol. Design 10*, 153-164.

Ryde, U. (2003) Combined quantum and molecular mechanics calculations on metalloproteins. *Curr. Opin. Chem. Biol. 7*, 136-142.

Ryde U., http://www.teokem.lu.se/~ulf/comqum.html

Ryde, U. (2009) "Quantum Mechanical/Molecular Mechanical (QM/MM) Methods and Applications in Bioinorganic Chemistry", in Computational Inorganic and Bioinorganic Chemistry, E. I. Solomon, R. B. King & R. A. Scott, eds. J. Wiley & Sons, Ltd. Chichester, UK, pp. 33-42.

Sandler, B.H., Nikonova, L., Leal, W.S., and Clardy, J. (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol. 7*, 143-151.

Senet, M. L., Wilson, S. (2001) Intramolecular basis set superposition errors. *Int. J. Quant. Chem. 82*, 282-292.

Senn, H. M., Thiel, W. (2007) QM/MM methods for biological systems. In: Atomistic approaches in modern biology: From quantum chemistry to molecular simulations. *Topics in current chemistry 268*, 173-290.

Senn, H. M., Thiel, W. (2009) QM/MM Methods for Biomolecular Systems. *Angew. Chem.-Int. Edit. 48*, 1198-1229.

Schäfer, A., Huber, C., Ahlrichs, R. (1994) Fully optimized contracted gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 100, 5829-5835.

Schäfer, A., Klamt, A., Sattel, D., Lohrenz, J. C. W., Eckert, F. (2000) COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems. *Phys. Chem. Chem. Phys. 2*, 2187-2193.

Schürer, G., Horn, A. H. C., Gedeck, P., Clark, T. (2002) The Reaction Mechanism of Bovine Lens Leucine Aminopeptidase. *J. Phys. Chem. B 106*, 8815-8830.

Schürer, G., Lanig, H., Clark, T. (2004) Aeromonas proteolytica aminopeptidase: an investigation of the mode of action using a QM/MM-approach. *Biochemistry 43*, 5414-5427.

Siegbahn, P.E.M., Himo, F. (2009) Recent developments of the quantum chemical cluster approach

for modeling enzyme reactions. *J. Biol. Inorg. Chem. 14*, 643-651.

Simon, S., Duran, M., Dannenberg, J. J. (1996) C-H/O H-Bonded complexes. How does basis set superposition error change their potential energy surfaces? *J. Chem. Phys. 105*, 11024-11031.

Slusher, B. S., Vornov, J. J., Thomas, A. G., Hurn, P. D., Harukuni, I., Bhardwaj, A., Traystman, R. J., Robinson, M. B., Britton, P., Lu, X. C., Tortella, F. C., Wozniak, K. M., Yudkoff, M., Potter, B. M., and Jackson, P. F. (1999) Selective inhibition of NAALADase, which converts NAAG to glutamate, reduces ischemic brain injury. *Nat. Med. 5*, 1396-1402.

Southall, N.T., Dill, K.A., Haymet, A. D.J. (2002) A view of the hydrophobic effect. *J. Phys. Chem. B 106*, 521-533.

Speno, H. S., Luthi-Carter, R., Macias, W. L., Valentine, S. L., Joshi, A. R., and Coyle, J. T. (1999) Site-directed mutagenesis of predicted active site residues in glutamate carboxypeptidase II. *Mol. Pharmacol. 55,* 179-185.

Spungin, A., Blumberg, S. (1989) *Streptomyces griseus* aminopeptidase is a calcium-activated zinc metalloprotein : Purification and properties of the enzyme. *Eur. J. Biochem. 183*, 471-477.

Stamper, C., Bennett, B., Edwards, T., Holz, R. C., Ringe, D., Petsko, G. (2001) Inhibition of the aminopeptidase from Aeromonas proteolytica by L-leucinephosphonic acid. Spectroscopic and crystallographic characterization of the transition state of peptide hydrolysis. *Biochemistry 40*, 7035-7046.

Steiner, T., Koellner, G. (2001) Hydrogen bonds with π-acceptors in proteins: Frequencies and role in stabilizing local 3D structures. *J. Mol. Biol. 305*, 535-557.

Stephens, P. J., Devlin, F. J., Frisch, M. J., Chabalowski, C. F. (1994) Ab-initio calculations of vibrational absorbtion and circular-dichroism spectra using density-functional force-fields. *J. Phys. Chem. 98*, 11623-11627.

Stillinger, F.H. (1973) Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled- particle theory, *J. Solution Chem. 2*, 141-158.

Sträter, N., Sun, L., Kantrowitz, E. R., Lipscomb, W.N. (1999) A bicarbonate ion as a general base in the mechanism of peptide hydrolysis by dizinc leucine aminopetidase. *Proc. Natl. Acad. Sci. 96*, 11151-11155.

Street I.P., Armstrong C.R., Withers C.G. (1986) hydrogen-bonding and specificity - fluorodeoxy sugars as probes of hydrogen - bonding in the glycogen - phosphorylase glucose complex. *Biochemistry 25*, 6021-6027.

Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., Morokuma, K. (1996) ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and $Pt(P(t\text{-}Bu)_3)_2 + H_2$

Oxidative Addition. *J. Phys. Chem. 100*, 19357-19363.

Šácha, P., Zámečník, J., Bařinka, C, Hlouchová, K., Vícha, A., Mlčochová, P., Hilgert, I., Eckschlager, T., and Konvalinka, J. (2007) Expression of glutamate carboxypeptidase II in human brain. *Neuroscience 144*, 1361-1372.

Štrajbl, M., Florián, J., Warshel,A (2000) Ab initio evaluation of the potential surface for general base-catalyzed methanolysis of formamide: A reference solution reaction for studies of serine proteases. J. Am. Chem. Soc. 122, 5354-5366.

Šponer, J., Jurečka, P., Hobza, P. (2003) Accurate Interaction Energies of Hydrogen-Bonded Nucleic Acid Base Pairs. *J. Am. Chem. Soc. 126*, 10142-10151.

Szabo A., Ostlund N. S. (1982) Modern quantum chemistry:Introduction to advanced electronic structure theory. *McGraw-Hill*, New York.

Tanford, C. (1978) The hydrophobic effect and the organization of living matter *Science 200*, 1012-1018.

Tao, J. M., Perdew, J. P., Staroverov, V. N., Scuseria, G. E. (2003) Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Physical Review Letters 91***,** 146401.

Tatko, C. D., Waters, M. L. (2004) Comparison of C−H···π and Hydrophobic Interactions in a β-Hairpin Peptide: Impact on Stability and Specificity. *J. Am. Chem. Soc. 126*, 2028-2034.

Troyer, J. K., Beckett, M. L., and Wright, G. L.Jr. (1995) Detection and characterization of the prostate-specific membrane antigen (PSMA) in tissue extracts and body fluids. *Int. J. Cancer 62*, 552-558.

Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., Tanabe, K. (2000a) Origin of the atraction and directionalizy of the NH/π intreaction: comparison with OH/π and CH/π interactions. *J. Am. Chem. Soc. 122*, 11450-11458.

Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., Tanabe, K. (2000b) The magnitude of the CH/π interactions between benzen and some model hydrocarbons. *J. Am. Chem. Soc. 122*, 3746-3753.

Tsuzuki, S., Uchimaru, T., Matsumara, K., Mikami, M., Tanabe, K. (1999) Effects of basis set and electron correlation on the calculated interaction energies of hydrogen bonding complexes: MP2/cc-pVD5Z calculations of H_2O-MeOH, H_2O- Me_2O, H_2O- H_2CO, MeOH-MeOH and HCOOH-HCOOH complexes. *J. Chem. Phys. 110*, 11906-11910.

Umezava, Y., Nishio, M. (1998) CH/π interactions as demonstrated in the crystal structure of guanine-nucleotide binding proteins, src hohology-2 domains and human grouth hormon in complex with their specific ligands. *Bioorganic & Medicinal Chemistry 6*, 493-504.

Umezava, Y., Nishio, M. (1998) CH/π interactions in crystal structures of class I MHC antigens and their complexes with peptides. *Bioorganic & Medicinal Chemistry 6*, 2507-2515.

Urbano, A. (2003) Recent developments in the synthesis of helicene-like molecules. *Angew. Chem. Int. Ed., 42*, 3986-3989.

V aldés, H., Řeha, D., Hobza, P. (2006) Structure of isolated tryptophyl-glycine dipeptide and tryptophyl-glycyl-glycine tripeptide: ab initio SCC-DFTB-D molecular dynamics simulations and high-level correlated ab initio quantum chemical calculations. *J. Phys. Chem. B 110*, 6385-6396.

van Mourik, T., Karamertzanis, P. G., Price, S. L. (2006) Molecular conformations and relative stabilities can be as demanding of the electronic structure method as intermolecular calculations. *J. Phys. Chem. A 110*, 8-12.

Varandas, A. J. C. (2000) Basis-set extrapolation of the correlation energy. *J. Chem. Phys. 113*, 8880-8887.

Vargas, R., Garza, J., Dixon, D.A., Hay, B.P. (2000) How strong is the C^{α}-H□□□O=C hydrogen bond? *J. Am. Chem. Soc. 122*, 4750-4755.

Vendruscolo, M., Paci, E., Dobson, C. M., Karplus, M. (2001) Three key residues form a critical contact network in a protein folding transition state. Nature, 409, 641-645.

von Lilienfeld, A. O., Tavernelli, I., Rothlisberger, U., Sebastiani, D. (2004) Optimization of effective atom centered potentials for London dispersion forces in density functional theory. *Phys. Rev. Lett. 93*, 153004.

Vreven, T., Byun, K. S., Komaromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K., Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput. 2***,** 815-826.

Warshel, A. (1991) *Computer Simulation of Chemical Reactions in Enzymes and Solutions*, John Wiley & Sons.

Warshel, A., Sharma, P.K., Kato, M., Xiang, Y., Liu, H., Olsson, M.H.M. (2006) Electrostatic basis for enzyme catalysis. *Chem. Rev. 106*, 3210-3235.

Weigend, F., Ahlrichs, R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys. 7*, 3297-3305.

Weston, J. (2005) Mode of action of bi- and trinuclear zinc hydrolases and their syntheticanalogues. *Chem. Rev. 105*, 2151-2174.

Williams, D.H., Searle, M.S., Mackay, J.P., Gerhard, U., Maplestone, R.A. (1993) Toward an estimation of binding constants in aqueous-solution - studies of associations of vancomycin

group antibiotics. *Proc. Natl. Acad. Sci. USA 90,* 1172-1178.

Williams, D.H., Westwell, M.S. (1998) Aspects of weak interactions. *Chem. Soc. Rev. 27,* 57-63.

Wojtasek, H., and Leal, S.W. (1999). Conformational change in the pheromone-binding protein from Bombyx mori induced by pH and by interaction with membranes. *J. Biol. Chem. 274,* 30950-30956.

Woon, D. E., Dunning, T. H., Jr. (1993) Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys. 98,* 1358-1371.

Wouters, M.A., Husain, A. (2001) Changes of zinc ligation promote remodeling of the active site in the zinc hydrolyase superfamily. *J. Mol. Biol. 314,* 1191-1207.

Zhao, Y., Truhlar, D. (2008) Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res. 41,* 157-167.

Zhou, J., Neale, J. H., Pomper, M. G., and Kozikowski, A. P. (2005) NAAG peptidase inhibitors and their potential for diagnosis and therapy. *Nat. Rev. Drug Discov. 4,* 1015-1026.

Zimmerli, U., Parrinello, M., Koumoutsakos, P. (2004) Dispersion corrections to density functionals for water aromatic interactions. *J. Chem. Phys. 120,* 2693-2699.

Appendices