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

Doctoral Thesis

Noncovalent Interactions in the Gas Phase and Aqueous Solution: Theoretical Study

Lucie Zendlová

Advisor: Prof. Ing. Pavel Hobza, DrSc.
Institute of Organic Chemistry and Biochemistry AS CR
Center for Biomolecules and Complex Molecular Systems

Univerzita Karlova v Praze
Přírodovědecká Fakulta
Katedra fyzikální a makromolekulární chemie

Disertační práce

Nekovalentní interakce v plynné fázi a vodném roztoku: Teoretické studium

Lucie Zendlová

Vedoucí disertační práce: Prof. Ing. Pavel Hobza, DrSc.
Ústav organické chemie a biochemie AV ČR, v.v.i.
Centrum biomolekul a komplexních molekulárních systémů

I confirm that the presented PhD thesis was worked out solely by myself and all the literature used is properly cited. Neither the thesis nor its parts were used previously for obtaining any academic degree.

In Prague, 4th April, 2008 Lucie Zendlová

Acknowledgement

Firstly, I would like to thank prof. Hobza for his patiently supervising of my doctoral work. I thank to Dr. Michal Hocek for introducing me to the topic of unnatural base analogues. Finally, I would like to thank to all members of the Center for Biomolecules and Complex Molecular Systems who worked with me on the projects or helped me with a good advice, namely Dr. David Řeha, Dr. Martin Kabeláč, Dr. Lubomír Rulíšek, Dr. Jiří Černý, Dr. Petr Jurečka and Dr. Tomáš Kubař.

Contents

1 Introduction 6

1.1 Subject of Thesis 6

1.2 The Noncovalent Interactions 7

1.3 Electrochemical Properties of the Derivatives of Nucleobases Linked with Transition Metals Complexes 9

2 Computational Methods 10

2.1 Molecular Mechanics (MM) 10

2.1.1 Force-field Parameterization 10

2.1.2 Molecular Dynamics (MD) 11

2.1.3 Molecular Dynamics/Quenching (MD/Q) 12

2.2 Implicit Solvation Model 12

2.3 Quantum Mechanical (QM) Calculations 14

2.3.1 Hartree-Fock (HF) 15

2.3.2 Moller-Plesset (MP) Method 15

2.3.3 Coupled Clusters (CC) 15

2.3.4 Density Functional Theory with Empirical Dispersion Term (DFT-D) 16

2.3.5 Self-consistent Charges Density-functional Tight-binding Method Augmented with Empirical Dispersion Term (SCC-DFTB-D) 17

2.3.6 Local Method and Resolutions of Identity (RI) 18

2.4 Total Interaction Energy 18

2.4.1 Basis Set Superposition Error (BSSE) 19

2.4.2 Deformation Energy 19

2.4.3 Complete Basis Sets Limit (CBS) 20

2.4.4 Higher Order Correlation Contributions 20

3 Model Systems 22

3.1 Microsolvated Nucleic Acid Base Pairs 22

3.2 Unnatural Base Analogues 25

3.2.1 Hydrophobic Base Analogues 25

3.2.2 The Derivatives of Nucleobases Linked with Transition Metals Complexes 29

4 Results 31

4.1 Microsolvated Nucleic Acid Base Pairs 31

4.2 Hydrophobic Base Analogues 34

4.3 The Derivatives of Nucleobases Linked with Transition Metals Complexes 39

5 Conclusions 42

References 44

Molecular Modeling Packages Used 49

Appendices 50

1 Introduction

1.1 Subject of Thesis

The noncovalent interactions in biomolecules are essential for their functioning. To understand them properly is crucial for understanding the biological processes in living organisms.

The aim of this thesis is to study the nature of noncovalent (namely H-bonding and stacking) interactions in isolated nucleic acids base pairs in the gas phase and in chemically modified (unnatural) base pairs accommodated in the DNA duplex in aqueous solution. Further, apart from these, we have also investigated the electrochemical properties of the derivatives of nucleobases linked with transition metals complexes.

For these purposes the *ab initio*, semiempirical quantum chemical and empirical methods can be used. Although the *ab initio* quantum chemical methods are very powerful tool for study of noncovalent interactions, they suffer from the large requirement on the CPU time, especially at the highest levels of theory.

To overcome this disadvantage, the introduction of some approximations is needed. Depending on the approximation implemented, the systems of different size can be treated. Passing to the semiempirical and empirical methods, the calculations become significantly less CPU time demanding and the larger systems can be calculated, however, with lower accuracy.

This thesis is divided to two main parts:

i) a) Investigation of the noncovalent interactions in nucleic acid base pairs (adenine…thymine, guanine…cytosine and their methylated analogues) in a presence of few water or organic solvent molecules in the gas phase without sugar-phosphate backbone. The results of this study were already published in *J. Phys. Chem. B* and *Chem.Phys.Chem.* journals and they are attached to this thesis as Appendices A, B and C.

b) Investigation of the noncovalent interactions in the DNA duplex modified by a series of unnatural base analogues. The results of this study are in preparation and attached to this thesis as Appendix D.

ii) Investigation of the electrochemical properties of the derivatives of nucleobases linked with transition metals (Ru^{2+}, Os^{2+}) complexes. The results of the first part of this study were already published in *Eur. J. Inorg. Chem.* journal and is attached to this thesis as Appendix E. The second part was submitted to *J.Am.Chem.Soc.* journal and is attached to this thesis as Appendix F.

1.2 The Noncovalent Interactions

The noncovalent interactions play important roles in many biological systems. They control, among others, the base-base interactions leading to the double helical structure of DNA, secondary and tertiary protein structure, enzyme recognition of the substrate, etc. There exist many types of noncovalent interactions. In our work we are going to discusse two most important types of noncovalent interactions: H-bonding and stacking ones (cf. Figure 1).

Hydrogen Bond: The hydrogen bond is a special type of dipole-dipole interaction that exists between an electronegative atom and a hydrogen atom bonded to another electronegative atom. This interaction type always involves a hydrogen atom and the energy of this bond can be in some cases close to that of weak covalent bonds (1.9 kcal/mol for N-H…O, 6.9 kcal/mol for O-H…N, 5 kcal/mol for O-H…O, but 40 kcal/mol for F-H…F^-). These bonds exist between molecules (intermolecular H-bond), i.e. in the nucleic acid base pairs or within different parts of a single molecule (intramolecular H-bond), i.e. in peptides or proteins. The H-bonding interactions are very common in nature and play important role not only in the biology but also in the inorganic world. The theoretical study of the H-bonds is easy and H-bonding is properly described at the empirical, semiempirical and at any ab inito level.

Stacking interactions: The stacking interaction (π-π interaction) is a noncovalent interaction between compounds containing aromatic moieties. The π-π interactions have a origin in the intermolecular overlap of p-orbitals in π-conjugated systems, so they become stronger as the number of π-electrons increases. The stacking interactions mainly originate from London dispersion forces. For example, in DNA,

π-stacking occurs between adjacent nucleotides and contributes significantly to the stability of the molecular structure. The study of the stacked systems is more difficult than that of H-bonded ones since it needs a proper description of the London dispersion forces. Qualitatively and quantitatively corrected results can be achieved by using the high-level correlated *ab initio* quantum methods (like MP2 or CCSD(T)). On the contrary, the DFT method does not cover the London dispersion energy and their use in biological systems (where the dispersion plays significant role) is not recommendable.[1] The proper choice of the atomic orbitals basis set is also very important for proper description of noncovalent interactions and for stacking interaction in particular, the inclusion of diffuse polarization function is inevitable. The empirical force fields should be used with care. They can work quite well for some rigid systems (e.g. nucleobases), but they do not work for some floppy systems (e.g. peptides and proteins).

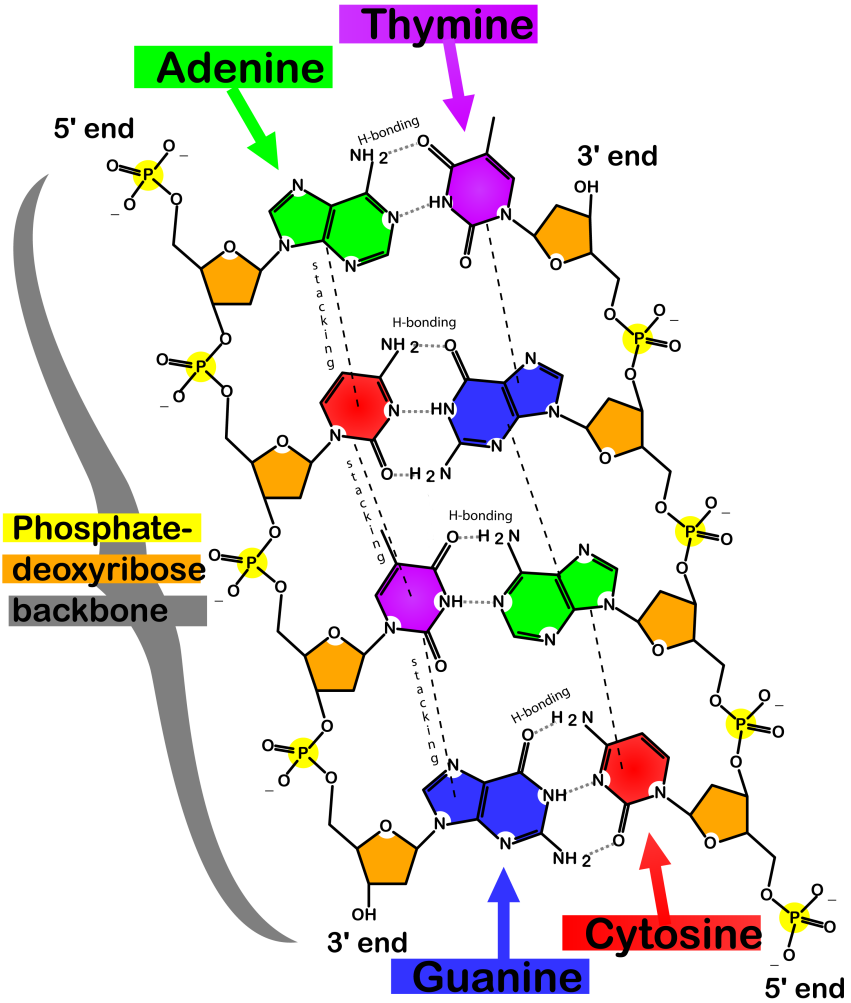


Figure 1. The H-bonding and stacking interactions in the DNA duplex. The figure was adopted from wikipedia.org web page.

1.3 Electrochemical Properties of the Derivatives of Nucleobases Linked with Transition Metals Complexes

An investigation of some biological processes (taking part in e.g. Parkinson's disease, schizophrenia, diabetes) and modeling of their mechanisms can be facilitated by the use of electrochemically-modified compounds and following their electrochemical properties. This approach relies on mutual electronic interactions of two parts of a molecular system where an incorporated redox probe is able to reflect electronic changes that occur in the other part of the molecule (induced by, e.g., a redox process, host-guest interaction, acid-base equilibria, ion pairing) by its changed electrochemical response. The prerequisite is that the redox probe interacts electronically with the parent molecule, forming usually an electronically delocalized system via a conjugated or coordination linker. Then, the redox properties of the probe are perturbed by changes taking place at the parent molecule and can be quantified by electrochemical methods. The complexes of transition metals (Ru, Os etc.) represent very useful electrochemically active label for important classes of biologically relevant compounds, and conjugates of nucleobases, nucleosides and nucleic acids.

2 Computational Methods

The main methods and approaches, which were used in our calculations are presented in this chapter. It starts with brief overview of molecular mechanics methods, continues with description of implicit solvation model and quantum chemical methods. The end of this chapter is dedicated to description of calculations of the interaction energy.

2.1 Molecular Mechanics (MM)

The MM methods do not consider electrons and calculate the energy of a system as a function of the nuclear positions only. Compare to the quantum chemical methods, the molecular mechanics method provides the results orders of magnitude faster, in some cases with accuracy at high-level quantum chemical calculations. Thus, the molecular mechanics methods can be used to perform calculations on the systems containing significant numbers of atoms.

2.1.1 Force-field Parameterization

Empirical force field defines the dependence of potential energy on the atomic coordinates of molecules in the system and it is usually represented (implemented in program code) by a set of atom types (to define the atoms in a molecule), parameters (for bond lengths, bond angles, dihedral angles, etc.) and equations (to calculate the energy of a molecule). The majority of the empirical forcefields use the following equation in order to calculate the energy of a molecule:

$$E_{total}(r^N) = \sum K_l(l - l_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals n=1,2,3,4,5,6} \sum \frac{V_n}{2}[1 + \cos(n\phi - \gamma_n)] + \sum_{i<j} [\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} - \frac{q_i}{\varepsilon r_{ij}}] \quad (1)$$

where the symbols have following meaning: l - bond length, l_{eq} - equilibrium bond length, K_l - bond stretching force constant, θ -bending angle, θ_{eq} - equilibrium angle, K_θ - bending force constant, V_n - amplitude ("force constant") of the dihedral, ϕ - dihedral angle, γ_n - phase offset, r_{ij} - interatomic distance, A_{ij} and B_{ij} - vdW parameters, q_i - atomic charge, ε- permitivity.

For our simulations we used parm94 force field implemented in AMBER package,[2] which performs well for modeling various organic biomolecules. For the study of modified biomolecules, which are not covered up in parm94 library of fragments it is necessary to estimate force-field parameters and atomic charges. The missing force-field parameters can be obtained, for example, from general amber force field (gaff),[3] which is designed for majority of organic molecules. The atomic charges can be derivated using The Restrained ElectroStatic Potentials (RESP)[4] procedure, which fits the quantum chemically calculated electrostatic potential at molecular surface using an atom centered point charge model. Since we performed our calculations in a solvent, we evaluated (following recommended procedure[4]) the atomic charges at HF/6-31G** level. This method overestimates dipole moments of most molecules by about 10-20%, this representation compensates the lack of polarization in an effective two-body force-field. The explicit water simulations were performed with TIP3P model of water molecules, which despite its simplicity gives good description of solvent and is often used in biomolecular simulations.

2.1.2 Molecular Dynamics (MD)

The MD method calculates the time dependent behavior of a molecular system using the Newton's laws of motions.

$$F_i = m_i \frac{d^2 q_i}{dt^2} \qquad (2)$$

where m_i is the mass of a particle i, q_i its position. The force F_i is the derivative of the potential.

The result of integration of Newton's equations is a trajectory that specifies how the position and velocities of atoms vary in time. There exist many methods solving the equations of motion, for example Verlet's method or Gear's method of predictor and corrector. Both algorithms use the positions and accelerations of atoms at time t, and the positions from the previous step r(t-δt) to calculate the new positions at (t+δt), r(t+δt). In contrast to Verlet's method the Gear's method contains more steps: First, the new positions, velocities and accelerations are predicted according to the Taylor expansion of r(t+δt). Second, the forces are evaluated at the new positions

to give accelerations a(t+δt). These accelerations are compared with the predicted ones from the Taylor series expansion, $a^c(t+\delta t)$, and the difference between them is used to 'correct' new positions and velocities.

The initial velocities at time t=0 and certain temperature are established using Maxwell-Boltzman distribution of velocities.

2.1.3 Molecular Dynamics/Quenching (MD/Q)

The MD/Q method is, in a fact, combination of MD and energy minimalization, which allows explore potential energy surface (PES), collects the various minima and estimates their populations. After certain time the MD run is interrupted, the structures are optimized to their minima. The minimized structures are saved, and the MD run continues from the point where it was interrupted. From the saved structures are selected just those, which are geometrical distinct. Thus, the populations obtained during MD correspond to the free energy surface.

2.2 Implicit Solvation Model

Implicit solvation (also known as continuum solvation) is a method representing solvent as a continuous medium, where the numbers of the degrees of freedom of the solvent molecules are described in a continuous way, usually by means of a radial distribution function.

The solvation free energy (ΔG_{solv}) is the free energy change accompanying transfer of molecule from vacuo to solvent. This energy consist of three components:

$$\Delta G_{solv} = \Delta G_{elec} + \Delta G_{vdw} + \Delta G_{cav} \qquad (3)$$

where ΔG_{elec} is electrostatic component comprising the polarization of a solvent which is modeled as continuum of constant dielectric ε. ΔG_{vdw} represents the van der Waals interaction between solute and solvent and ΔG_{cav} is the free energy required to form the cavity within the solvent.

There exist number of methods to describe solvation effects. One of the simplest is considered Born-Onsager model,[5] which places the solute in a spherical

cavity. The solute-dipole in the cavity induces dipole in the solvent, which in turn induces an electric field in the cavity - the reaction field.

Since the molecules are rarely rigorously spherical the more realistic cavity shape is created using van der Waals radii of atoms of solute (cf. Figure 2). The cavity surface is divided into a large number of small surface elements, and there is a point charge associated with each surface element. The total electrostatic potential at each surface element equals the sum of the potential due to the solute $\phi_\rho(r)$ and the potential due to the other surface charges $\phi_\sigma(r)$:

$$\phi(r) = \phi_\rho(r) + \phi_\sigma(r) \quad (4)$$

The reaction field can be incorporated into solute Hamiltonian (H) (the self-consistent reaction field (SCRF) method). In this approach the reaction field is considered as a perturbation[6] of the Hamiltonian of an isolated molecule.

$$H_{tot} = H_{id} + H_{per} \quad (5)$$

For our calculations we used a method derivable from polarized continuum model (PCM)[7] called conductor-like screening (COSMO)[8] model. This method is based on assumption of ideal solvent conductor: the electric potential of the cavity surface is zero. If the distribution of the electric charges in the molecule is known, then is possible to calculate the charge q* on the surface. For real solvents can we assume that the charge q is lower by a factor f:

$$q^* = q\frac{\varepsilon_r - 1}{\varepsilon_r + 0.5} \quad (6)$$

where the ε_r means relative permittivity.

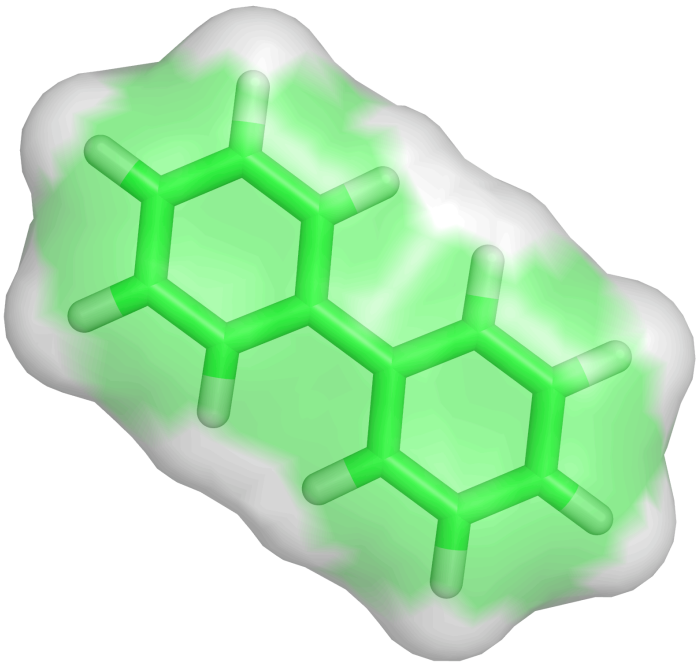


Figure 2. The cavity shape of biphenyl molecule.

2.3 Quantum Mechanical (QM) Calculations

Quantum mechanics is a mathematical theory that describes the behaviour of microscopic particles. In principle, quantum mechanics is able to provide any property of a molecule with a precision only limited by the uncertainty principle. The QM calculations use for description of the system the Schrödinger equation, which describes the space- and time-dependence of quantum mechanical systems. In order to calculate time-independent Schrödinger equation, the Born-Oppenheimer approximation is necessary to be established. In basic terms, it allows the wavefunction of a molecule to be separated into its electronic and nuclear (vibrational, rotational) components**.**

$$\Psi_{tot} = \Psi_{nucleic} \text{ x } \Psi_{electronic} \qquad (7)$$

However, there are two complications. Firstly, the analytical solution of Schrödinger equation is known only for a few simplest systems e.g. H, He^+, H_2^+. Thus a numerical approximation has to be used. Secondly, these numerical calculations are very

computer time demanding. To overcome this, numerous approximate solutions of the fundamental Schrödinger equation were suggested to apply.

2.3.1 Hartree-Fock (HF)

The HF method[9],[10] is the basic *ab inito* method. Its main assumption and weakness is the consideration of independent motion of electrons. The HF molecular wavefunction is built as a product of monoelectronic functions (orbitals) and, consequently, the correlation energy (dynamical correlation energy) is not included. The natural way how to improve the results of HF approximation is to involve the electron correlation.

2.3.2 Møller-Plesset (MP) Method

The MP perturbation theory[11] improves the HF method by adding electron correlation as a perturbation to the Fock operator. The unperturbated HF Hamiltonian operator H_o (created as a sum over the Fock operators) is extended by adding small perturbation V:

$$H = H_o + \lambda V \qquad (8)$$

where λ is arbitrary parameter.

The MP method is computationally very effective, size consistent, and estimates reasonably well the correlation energy. The main disadvantage is the fact that the MP method is nonvariational. It is usually applied to systems where the perturbation is relatively small (typically 0.5% in terms of energy) and the perturbation treatment is stable and well behaved. The MP approach is mostly used at the second order and is referred as the second-order many-body perturbation theory (MP2).

2.3.3 Coupled Clusters (CC)

The CC method[12],[13] is a numerical technique used for describing many-body systems. It starts from the HF molecular orbital method and adds a correction term to take into account electron correlation using cluster expansion of the wavefunction. The (intermediate normalized) coupled cluster wave function is written

as:

$$\Psi_{cc} = e^{T}\phi_0 \quad (9)$$

$$e^{T} = 1 + T + \frac{1}{2}T^2 + \frac{1}{6}T^3 + \ldots = \sum_{k=0}^{\infty}\frac{1}{k!}T^k \quad (10)$$

where the cluster operator T is given by:

$$T = T_1 + T_2 + T_3 \ldots + T_N \quad (11)$$

The idea in CC methods is to include all corrections of given type (S, D, T, Q etc.) to infinite order. A very promising method from the group of CC methods is the CCSD(T), which fully covers single and double excitations and partially (perturbatively) triple excitations. This method gives a very accurate estimation of the correlation energy, in the case of small systems (close to full configuration interaction (full CI) method[14-16]) at a relatively low cost (in comparison to full CI). Using this method is possible to calculate small, up to medium size systems.

2.3.4 Density Functional Theory with Empirical Dispersion Term (DFT-D)

The DFT[17] method is the based on the assumption that the electronic energy is determined completely by the electron density ρ contrary to wavefunction of the whole system. The main advantage of this assumption is that the electron density depends just on three coordinates (x,y,z), while many-electron wavefunction of the whole system depends on 3N coordinates (where N is number of electrons). However, there is a complication: although it was proven that each different density yields in the only one wavefunction, the functional connecting these two quantities is not known. Thus, the goal of DFT methods is to design functionals connecting the electron density with the energy.

The energy functionals can be divided in three main parts:

$$E_{DFT} = T[\rho] + E_{ne}[\rho] + E_{ee} \quad (12a)$$

$$E_{ee}[\rho] = J[\rho] + E_x[\rho] + E_c[\rho] \quad (12b)$$

where the symbols have following meaning: kinetic energy T[ρ], attraction between the nuclei and electrons $E_{ne}[\rho]$ and the repulsion between electrons $E_{ee}[\rho]$ (the nuclear repulsion is the constant in Born-Oppenheimer approximation). Furthermore, the $E_{ee}[\rho]$ can be divided into Coulomb J[ρ], exchange $E_x[\rho]$ and correlation $E_c[\rho]$ part, implicitly including correlations energy.

As it was mentioned above, the main problem is to design the functionals, mainly the exchange and correlation part. The exact form of these parts of functional is known only for the free electron gas. However, the approximation of non-interacting uniform electron gas does not work very well for atomic and molecular systems. Thus, many of exchange-correlations functionals were developed.

The DFT method itself fails in the description of noncovalent complexes where dispersion energy plays a dominant role. Thus, the improvement by adding the empirical dispersion term and establishing of DFT-D method was introduced by Grimme[18] and Jurečka *et al*.[19] For our calculations we used the DFT-D method with following functionals: PBE,[20] B3LYP[21] and TPSS.[22]

2.3.5 Self-consistent Charges Density-functional Tight-binding Method Augmented with Empirical Dispersion Term (SCC-DFTB-D)

The SCC-DFTB-D method combines approximated self-consistent charge density functional based tight-binding method with empirical dispersion energy treatment.

Originally, the non-self consistent tight-binding (TB) approach itself was being applied mostly to solid state and cluster physics and it was not working well for the molecular systems. New improvements like self-consistent charges (SCC) and addition of empirical dispersion term in order to reach the accuracy necessary for description of molecular systems, were introduced by Elstner.[23]

The SCC-DFTB method belongs to the group of traditional semi-empirical methods in quantum chemistry derived from HF method (AM1, PM3, MNDO). Formally, it has some similarity with extended Hückel theory or CNDO theory. However, the SCC-DFTB method is not semi-empirical method in a strict sense, since the parameterization procedure is completely based on density function theory (DFT) calculations (using self-consistent field-local density approximation (SCF-LDA) functional). SCC-DFTB method is roughly 2-3 orders of magnitude faster than

standard DFT calculations, the main cost is the solution of the generalized eigenvalue problem in a minimal basis, no integrals have to be evaluated during the run-time of the program. The inclusion of an empirical dispersion term (by Hobza[24]) establishing SCC-DFTB-D (parameterization using PBE functional) method, removes the major deficiency of DFT methods, namely the lack of dispersion energy.

2.3.6 Local Method and Resolutions of Identity (RI)

The speed of calculations is one of the properties of quantum chemical code, which influences its applicability. There are two different directions how to speed up the calculations. First, build more efficient codes (direct computation of electronic integrals, the parallelization of the calculations, the use of symmetry). Second, improve theoretical procedures (local treatment of the correlation energy or the use of the RI method). The local treatment[25],[26] of the correlation energy is based on the fact that the correlation is mainly a local quantity. In this type of approach, the molecular orbitals are transformed into localized ones and only those excitations from the geometrically closest orbitals are considered. RI[27] approximation is based on approximate evaluation of four-index two-electron integrals which is a critical component of many *ab initio* calculations. The four-index integrals are transformed into three-index integrals, which are computed considerably faster. The RI method is only another way to evaluate necessary integrals and no additional approximation to the theoretical method are introduced. Therefore RI modifications of standard methods (like self consistent field (SCF), MP2 and DFT) yield results qualitatively equivalent to original ones, but by an order of magnitude faster.

2.4 Total Interaction Energy

The interaction energy of the complex is generally defined as:

$$E^{int}_{AB} = E_{AB} - E_A - E_B \qquad (13)$$

where E^{int}_{AB} is a interaction energy of complex AB, E_{AB} is a calculated energy of complex AB and E_A, E_B are calculated energies of monomers A, B. However, in dependence on the approach used for calculation of the total interaction energy, the

general equation can be modified by other terms or procedures.

2.4.1 Basis Set Superposition Error (BSSE)

In quantum chemical calculations, only the finite basis sets can be used. Subsequently, the calculations of interaction energies are susceptible BSSE. As the atoms of interacting molecules (or of different parts of the same molecule) or two molecules approach one another, their basis functions overlap. Each monomer "borrows" functions from other nearby components, effectively increasing its basis set and bring artificial improvement in the calculation of derived properties such as energy. The BSSE decreases with the increasing size of basis sets used. One evident solution is the use of extremely large bases sets or even the infinite basis set. In the latter case the BSSE should disappeared. In the real situation the interaction energies should be corrected for the BSSE using the counterpoise method (CP) developed by Boys and Bernardi.[28] In this approach, the complex and the isolated monomers are described by the same basis set.

$$E_{int}AB = E_{(AB)}AB - E_{(AB)}A - E_{(AB)}B \qquad (14)$$

where $E_{(AB)}AB$ is energy of AB complex in AB basis set, $E_{(AB)}A$ is energy of A monomer in AB basis set and $E_{(AB)}B$ is energy of B monomer in AB basis set.

2.4.2 Deformation Energy

The deformation energy is always repulsive and is connected with the fact that the BSSE is not considered in the gradient geometry optimization. It is determined as the energy difference between the monomers adopting the final deformed geometry (as adjusted in the complex) and relaxed isolated monomers, all evaluated with the monomer basis set.

$$E_{def}AB = (E_{def}A - E_{mon}A) + (E_{def}B - E_{mon}B) \qquad (15)$$

where $E_{def}AB$ is energy of the optimized complex, $E_{mon}A$ and $E_{mon}B$ are energies of optimized monomers, $E_{def}A$ and $E_{def}B$ are energies of monomers in a final deformed geometries.

2.4.3 Complete Basis Sets Limit (CBS)

To obtain the most accurate interaction energies, the interaction energies can be extrapolated to the CBS using energies determined by systematically improved AO basis sets (e.g. the Dunning basis sets: cc-pVXZ, X= D, T, Q, … [29]) The HF energy converges with respect to the one-electron basis set already for relatively small basis sets. Contrary, the correlation energy converges unacceptably slowly. Thus, both energies must be extrapolated separately. In our work, we used the Helgaker *et al.* extrapolation scheme.[30],[31] In this extrapolation scheme, the HF (E_X^{HF}) and correlation (E_X^{corr}) energies are calculated as follows:

$$E_X^{HF} = E_{CBS}^{HF} + A^{(-\alpha X)} \tag{16a}$$

$$E_X^{corr} = E_{CBS}^{corr} + BX^{-3} \tag{16b}$$

where E_X and E_{CBS} are energies for the basis set with the largest angular momentum X and for the complete basis set respectively, and α is the parameter fitted in the original work.[30],[31]

2.4.4 Higher Order Correlation Contributions

To overcome problems native from high-level-method/large-basis-set combination can be solved by this way:[32] high order correlation effects can be evaluated using small basis set and basis set dependence can be estimated separately on a lower, less demanding level. One of the possibilities (also used in our calculations) is using the MP2 together with coupled clusters (CC) method, since the MP2 method itself overestimates the correction energy (namely its dispersion part). The CC and MP2 methods converge to the basis set limit in the same manner, what makes the difference between their results almost basis set independent. This approach has to two advantages: i) CCSD(T) method (coupled clusters with single and double excitations treated iteratively and perturbative triple excitations taken from MP4) is a method giving results close to full CI limit, and ii) MP2 method is of moderate computational demands, and can be further speed-up by resolution of

identity (RI) approximation. Finally, the CCSD(T) CBS energy can be approximated as:

$$E_{CBSCCSD(T)} \approx E_{CBSMP2} + \Delta E_{(CCSD(T))} \quad (17)$$

where the symbols have following meaning: $E_{CBSCCSD(T)}$ is the final CBS energy, E_{CBSMP2} is MP2 energy extrapolated to CBS limit and $\Delta E_{(CCSD(T))}$ is the CCSD(T) correction term.

3 Model Systems

The model systems, which we have studied can be separated in two main groups:

i) Isolated Microsolvated Nucleic Acid Base Pairs in Vacuo

ii) Unnatural Base Analogues

3.1 Microsolvated Nucleic Acid Base Pairs

A study of the isolated nucleic acid bases base pairs in the gas phase allows exploration of their intrinsic properties independently on the influence of the sugar-phosphate backbone. Consequently, the direct interaction between the bases and exact amount of ions or solvents molecules (water, organic solvent etc.) can be studied.

The aim of this study was to examine the dynamic structure, the potential and free energy surfaces of adenine…thymine, guanine…cytosine and their methyl derivative 9-methyladenine…1-methylthymine and 9-methylguanine…1-methylcytosine exposed to a small number (1,2 in the study of all systems using MD/Q or 0,1,2,4,8,16 in the study of organic solvent using MD) of water (W) molecules and organic solvents (methanol (CH_3OH), chloroform ($CHCl_3$) and dimethylsulfoxide (DMSO)), cf. Figures 3, 4. The methylation of the bases should simulate the presence of sugar molecules, by blocking the N1 and N9 atoms through them the bases are connected to the DNA backbone.

We have also focused our calculations on the description of specific interaction between the bases and solvent molecules.

The various theoretical approaches starting from the simple empirical methods (MD, MD/Q) to highly accurate *ab initio* quantum chemical calculations were used. Accurate results from the *ab initio* calculations could also serve as benchmarks to verify calculations at a lower level of theory. For more details see Appendices A, B and C.

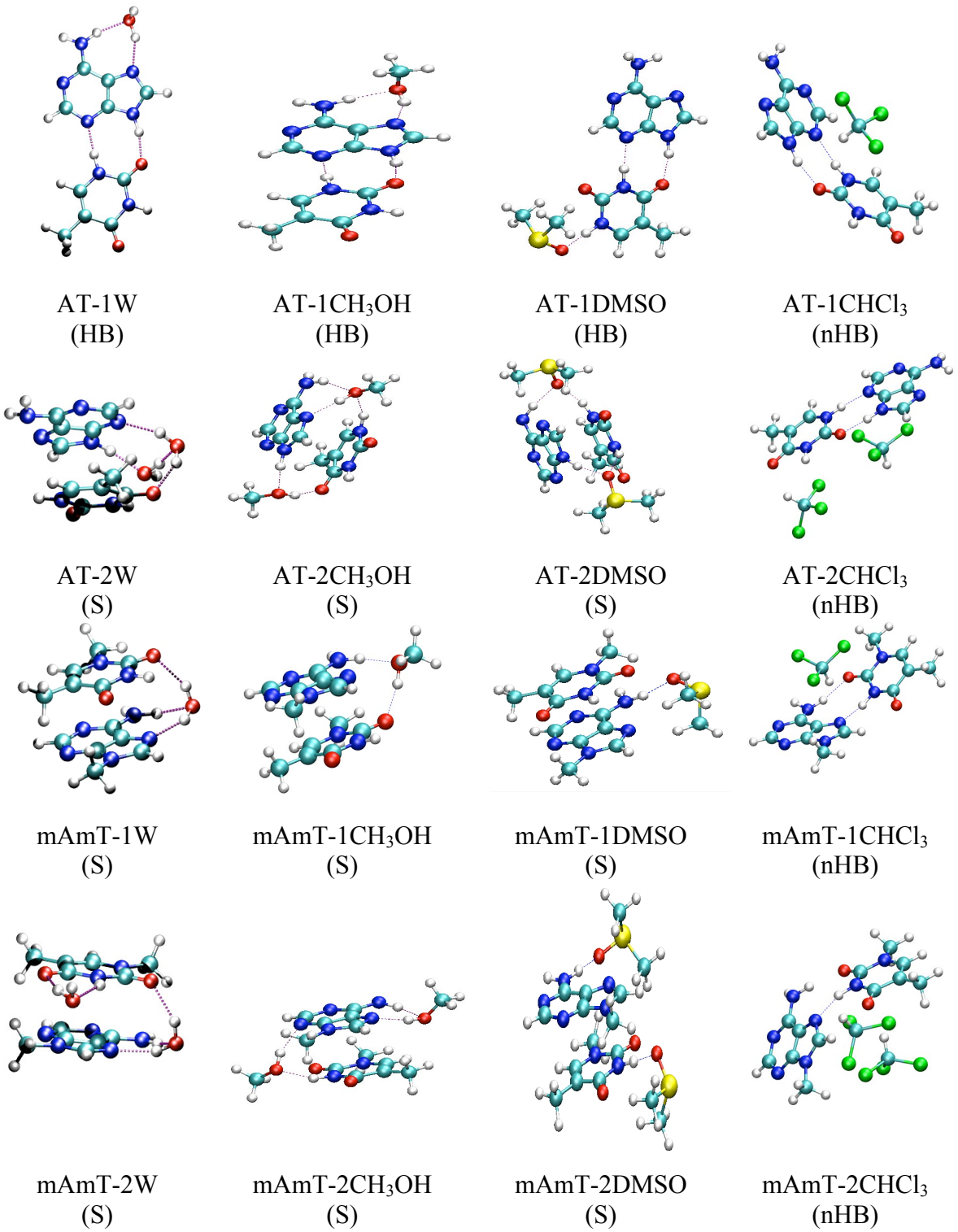


Figure 3. The most stable (at *ab initio* PES) nonmethylated adenine...thymine (AT) complexes with a one or two water, CH_3OH, DMSO and $CHCl_3$ molecule(s) and methylated adenine...thymine (mAmT) complexes with one or two water, CH_3OH, DMSO and $CHCl_3$ molecule(s). The letters in the parentheses mean: HB and nHB planar and nonplanar H-bonded, S stacked, T T-shape arrangement of the bases.

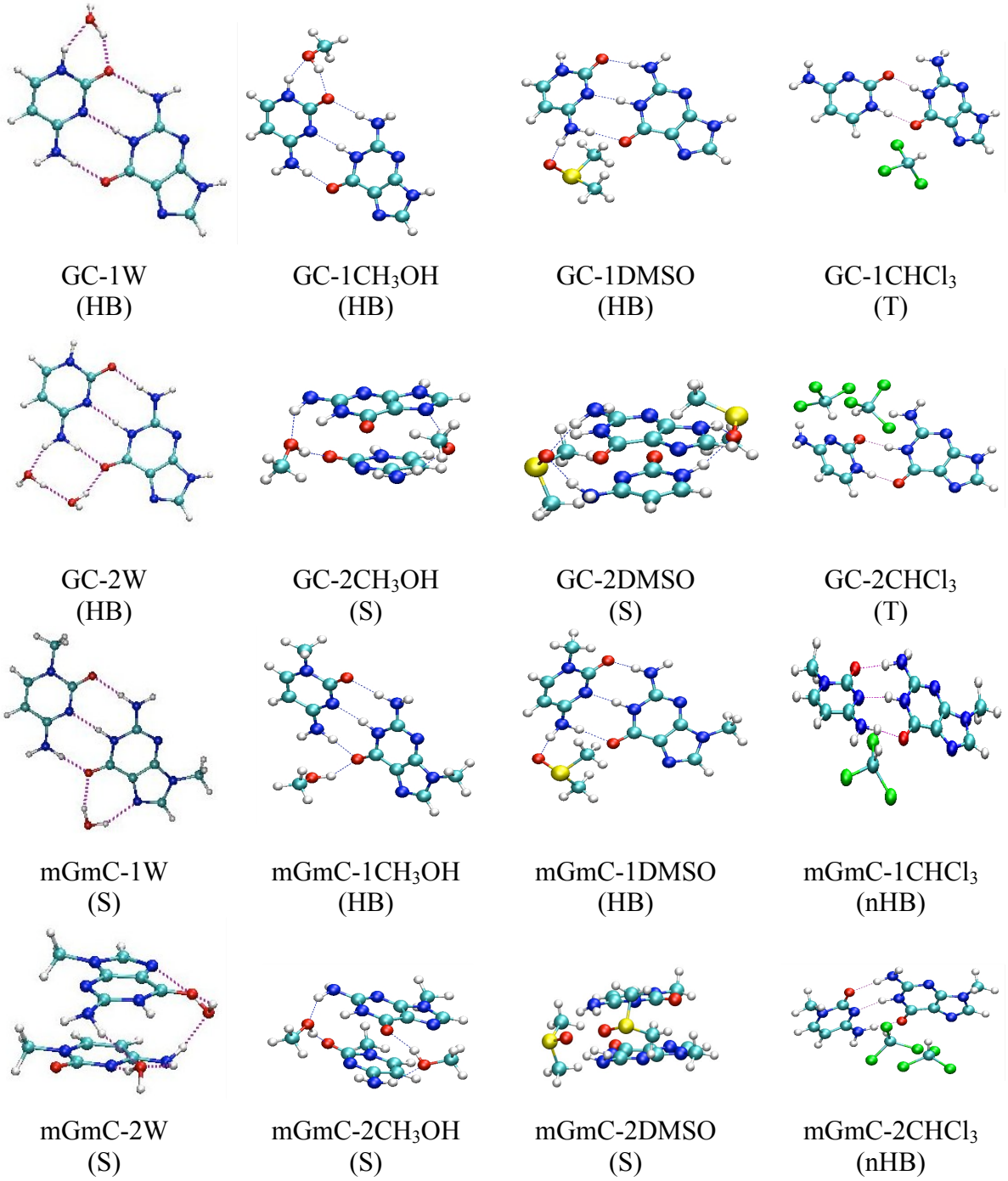


Figure 4. The most stable (at *ab initio* PES) nonmethylated guanine…cytosine (GC) complexes with a one or two water, CH_3OH, DMSO and $CHCl_3$ molecule(s) and methylated guanine…cytosine (mGmC) complexes with one or two water, CH_3OH, DMSO and $CHCl_3$ molecule(s). The letters in the parentheses mean: HB and nHB planar and nonplanar H-bonded, S stacked, T T-shape arrangement of the bases.

3.2 Unnatural Base Analogues

The natural DNA helix contains natural bases: adenine, thymine, cytosine and guanine. These bases are pairing together accordingly Watson-Crick rule:[33] adenine with thymine and cytosine with guanine. The DNA helix is held together via H-bonding and stacking interaction between these bases. The relative contribution of H-bonding and stacking interaction to the stability of the DNA duplex were discussed from the discovery of double helix.[33]

To investigate this phenomena, or even the possibility of the existence of stable DNA double helix, where natural bases are replaced by unnatural nucleobases, number of unnatural base analogues were designed.[34],[35],[36]

Several main groups of unnatural base analogues exist:

i) the derivatives of purine and pyrimidine bases, which posses the similar behavior to the natural ones, it means, that they are connected via H-bonds.[37],[38]

ii) hydrophobic base analogues, which are unable to create H-bonds and usually are too large to be accommodate in the DNA helix in one plane. Instead, they are stabilized by stacking interaction native from stacked arrangement of the bases in the base pair.[39],[40],[41]

iii) the bases derivatives linked with complexes of transition metals (e.g ferrocene[42] or bidentate N-ligands[43] with Ru, Rh, Ni, Cu, Co, Pt, Pd, Os etc.). These derivatives possess unique electrochemical and photophysical properties.

The base analogues from group ii) and iii) were investigated in our study.

3.2.1 Hydrophobic Base Analogues

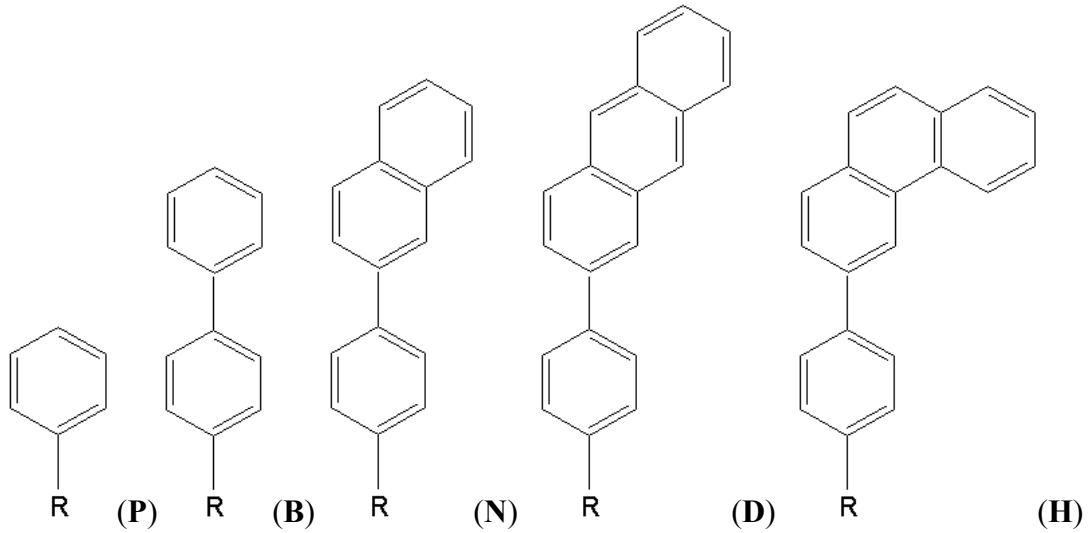
The hydrophobic unnatural base analogues are nonpolar and therefore in the polar solvent (water, in our calculations) show the higher affinity to other nonpolar bases than to polar ones. Also, the desolvation energies of unnatural base analogues are lower, than of the polar ones (in the case that the sizes of polar and nonpolar bases are comparable), which increase the stability of the duplex modified by these base analogues.

The recent experimental studies of nonpolar unnatural base analogues and their function in modified DNA observed the surprising promiscuity of DNA polymerase, and its ability to recognize and selectively incorporate the hydrophobic nucleobases into DNA duplex using the triphosphates of the particular base-modified nucleosides. It was also observed, that the hydrophobic unnatural base analogues tend to be selective to the same molecule and therefore form the self-pair. It follows, that the presence of H-bonds is not crutial for the selectivity of DNA building up.

The modified DNA can be used in genetic engineering, nanotechnology and molecular electronics. The extension of genetic alphabet via base analogues can lead to extension of aminoacids group and creation of artificial proteins or enzymes, which can catalyze new reaction.

The theoretical studies on the stability of potential unnatural nucleobases play an important role in their further design aimed at improving selectivity. The theoretical studies exploring the thermodynamic stability of the system are compared to the experimentally measured melting temperature of modified DNA.

The aim of our study of modified DNA duplexes is to investigate the nature of binding, high stability and selectivity of the base analogues. Also, the maximal size of base analogues which can be incorporated in DNA duplex were investigated. For our calculations we have chosen following series of base analogues (cf. Figure 5): phenyl, biphenyl, phenyl-naphtalene, phenyl-anthracene and phenyl-phenanthrene, which were subsequently incorporated instead of **AT** central base pair into (5'-GCGTAC**A**CATGCG-3') DNA duplex as a selfpair (A and T were both replaced) or a misspair (only A or T were replaced) (for example of DNA duplex modified by biphenyl cf. Figure 6). Since the **N**, **D** and **H** are not symmetrical along their axial axis, two of their orientations are possible: the nonsymetric part of modified nucleobase **X** oriented in direction to major, respectively minor groove of the DNA duplex (cf. Figure 7). Our studies has started with the MD simulations and continued with the calculations of H-bonded, stacked interstrand and intrastrand interaction, desolvation (solvation) energies and DNA duplex free stabilization. For more details see Appendix D.



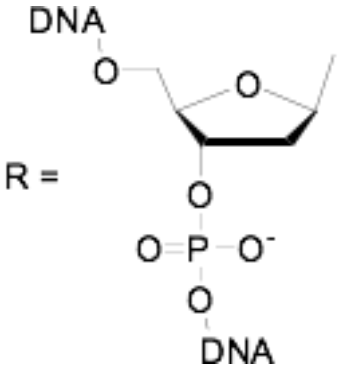
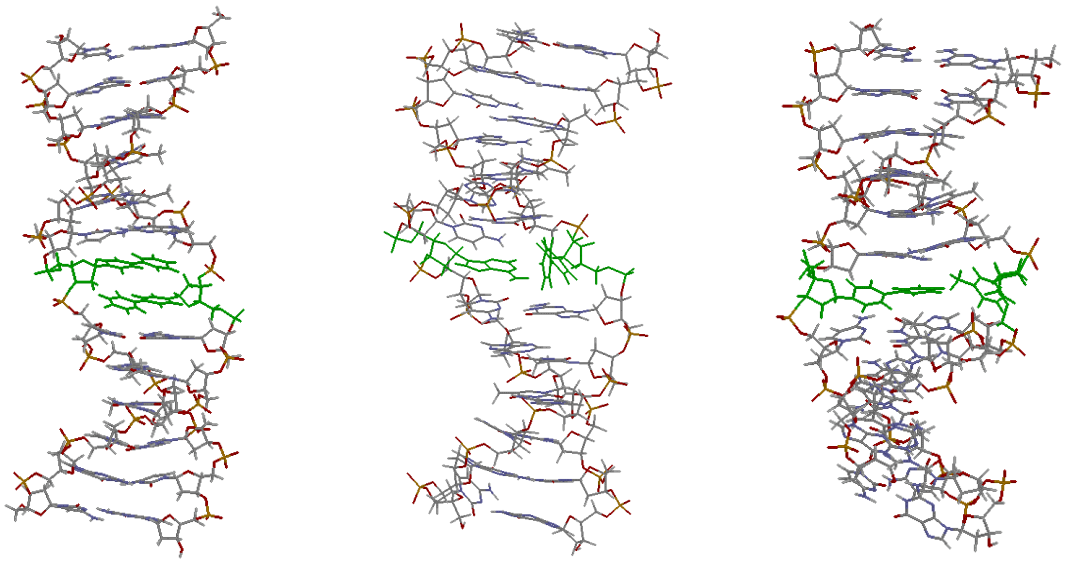


Figure 5. Chemical structures of modified nucleobases **X**: phenyl (**P**), biphenyl (**B**), phenyl-naphatelene (**N**), phenyl-anthracene (**D**), phenyl-phenanthrene (**H**)



(a) (b) (c)

Figure 6. 13mer of DNA duplex modified by a) biphenyl-biphenyl selfpair, b) adenine-biphenyl misspair, c) biphenyl-thymine misspair.

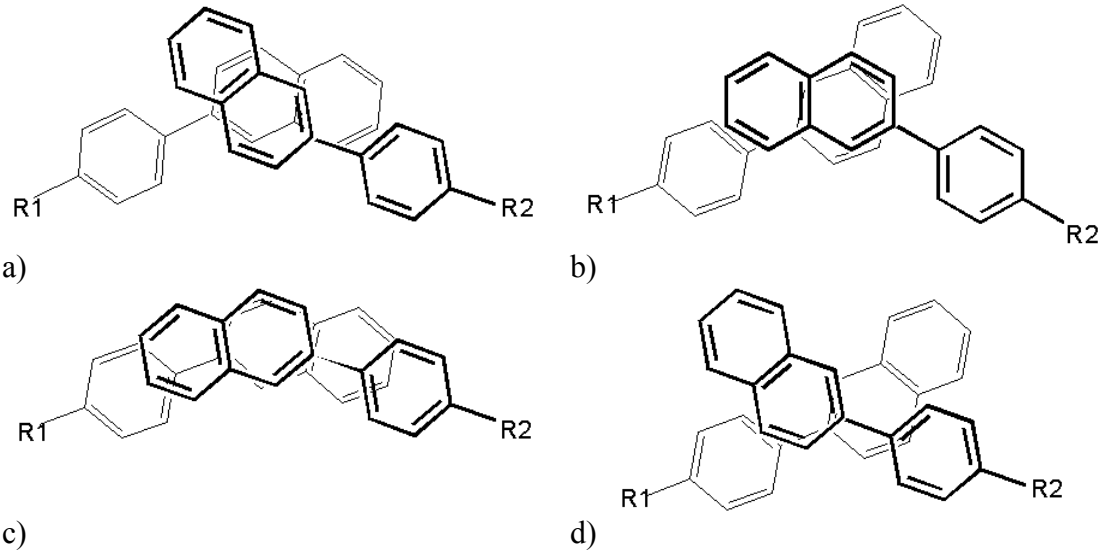


Figure 7. The schematic representation of four possible self-orientations of nonsymmetrical base analogue (**N** in this case) in the centre of DNA (top view).

a) Orientation of **N** (A, respectively) connected to strand 1 (R1) in direction of minor groove and **N** (T, respectively) connected to strand 2 (R2) in direction of major groove: $\mathbf{N}_{min}$-$\mathbf{N}_{maj}$ or A-$\mathbf{N}_{maj}$, $\mathbf{N}_{min}$-T, respectively.

b) $\mathbf{N}_{maj}$-$\mathbf{N}_{min}$ or A-$\mathbf{N}_{min}$, $\mathbf{N}_{maj}$-T, respectively.

c) $\mathbf{N}_{min}$-$\mathbf{N}_{min}$ or A-$\mathbf{N}_{min}$, $\mathbf{N}_{min}$-T, respectively.

d) $\mathbf{N}_{maj}$-$\mathbf{N}_{maj}$ or A-$\mathbf{N}_{maj}$, $\mathbf{N}_{maj}$-T, respectively.

3.2.2 The Derivatives of Nucleobases Linked with Transition Metals Complexes

The nucleic acids themselves are electroactive species producing oxidation and reduction signals at mercury or solid electrodes.[44],[45],[46] In addition to label-free DNA detection, different electroactive (or enzyme) tags connected to target DNAs or hybridization probes are used to improve sensitivity and/or specificity of analysis.[47],[48],[49],[50] As mentioned above, the transition metal complexes, such as ferrocene or complexes containing bidenate N-ligands (in particular phenanthrolines and bipyridines) with Ru, Rh, Ni, Cu, Co, Pt, Pd, Os[47],[48],[49], [50],[51],[52],[53] ions exhibit unique electrochemical and photophysical properties. Some of the phenanthroline complexes, which are also efficient DNA intercallators, have been extensively used as luminescent and electroactive DNA labels.[54],[55],[56],[57],[58] Attachment of probes based on metal complexes directly to a nucleobase via conjugate linkers should increase the efficiency of the charge transfer and thus enhance sensitivity.

The aim of our study was investigate the electrochemical properties of the labeled purines. The 9-benzyl-adenine linked with Ru and Os with bipyridine or phenantroline complexes (cf. Figure 8) were chosen as the models for our calculations. The DFT calculations have been performed in order to verify the potential utilization of these methods for theoretical prediction of redox potentials of the selected complexes. For more details see Appendices E, F.

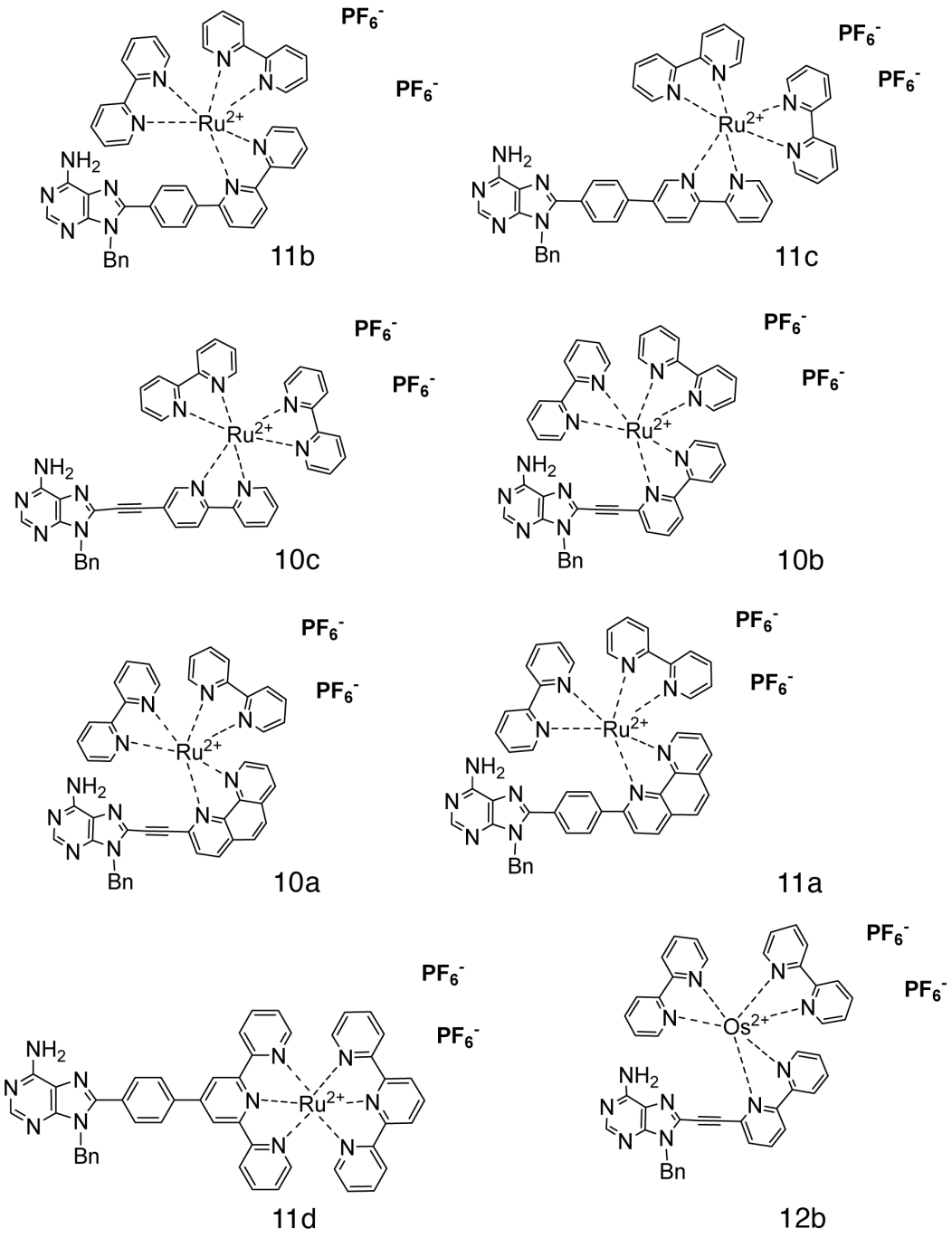


Figure 8. The studied 9-benzyl-adenine linked with Ru^{2+} and Os^{2+} with bipyridine or phenantroline complexes; each dication is compensated by two PF_6^- anions.

4 Results

4.1 Microsolvated Nucleic Acid Base Pairs

The potential and free energy surfaces (PES) and (FES) of all systems studied (adenine…thymine (AT), guanine…cytosine (GC) and their methyl derivative 9-methyladenine…1-methylthymine (mAmT) and 9-methylguanine…1-methylcytosine (mGmC) exposed to a small number (1,2) of molecules of water and organic solvents (CH_3OH, $CHCl_3$ and DMSO) with different dielectric constants (78, 33, 46.7, 4.8 at 25° respectively), were examined by the MD/Q method using the Cornell *et al*. force field[59] The most stable and populated structures found were fully reoptimized by the correlated *ab initio* RI-MP2 method.

We sorted the structures obtained by MD/Q to following three main classes: planar H-bonded (HB) and slightly nonplanar H-bonded (nHB), T-shaped (T), planar stacked (S) and slightly nonplanar stacked (nS), cf. Figures 3, 4.

The description of most stable structures of all systems studied can be found in following paragraphs. The interaction energies of the most stable structure of every system studied can be found in the Table 1.

The description of most stable structures at the AT and GC ab initio PESs

Complexes with water:The HB structures were observed as the most stable at both, AT-1W and GC-1W PESs. However, after addition of the second water molecule or/and methylation, the AT and GC systems behaved quite differently. In the case of nonmethylated AT complexes the presence of two water molecules led to the equivalent stability of S and HB structures (with difference in stabilization energy about 0.5 kcal/mol). Although, the increasing hydration number also lowers the difference between the stability of HB and S GC complexes (which means that S structures hydrates better than HB ones), the HB GC structure stays still the most stable (about 2.5 kcal/mol more stable than S one). Thus, two water molecules and also the methylation of the bases is needed to balance the stability of HB and S GC structures.

It is the methylation, what evokes the main difference in the stability of the most stable nonmethylated and methylated HB AT complexes (non Watson-Crick (WC) pattern), while on the HB GC complexes it has almost no influence. The originally most stable AT structures are eradicated due to blocking of two proton donors participated in the original H-bonds, while in the case of GC HB complexes (WC pattern), the methyl groups do not block the proton donors participated in the H-bonds between the bases.

The T-shaped structures were never located as the global minima at any PESs.

Complexes with CH_3OH: Although in the CH_3OH molecule absents the second hydrogen in comparison to water molecule, the behaviour of nonmethylated and methylated AT and GC complexes in the presence of CH_3OH molecule(s) is very similar to the behaviour of these complexes in water.

We found that for the monosolvated AT pair the global minimum corresponds to the HB structure. With addition of the second CH_3OH molecule the S structure is the most stable. Generally, the preference of the S structures is undoubted for all mAmT complexes studied. For GC and mGmC complexes with CH_3OH (except for GC-1CH_3OH, where the stabilization energy of HB is significantly higher than that of S), the stability of the most stable S and HB structures detected was comparable. In all the cases, T structures are less stable than the HB and S ones.

Complexes with DMSO: The complexes with DMSO exhibit the highest stabilization energies of all the systems studied. For comparison, each molecule of CH_3OH contributes to the stability of the complex by approximately 12 kcal/mol, DMSO by 15 kcal/mol, $CHCl_3$ by 9 kcal/mol and water by 5-10 kcal/mol. Surprisingly enough, apart from the fact that the solvents have different properties, the stability order of the most stable HB, S, and T structures remains similar as in CH_3OH. Even though the S structures are in most cases the most stable, they are less populated, and thus, their role in the global context is less important.

Complexes with $CHCl_3$: The weakest interaction between the base pair and solvent was found for $CHCl_3$. We found that at all PESs not the planar HB structures, but slightly nonplanar HB arrangement with the molecule(s) of the solvent situated above

or below of plane of bases corresponds to the global minimum. This arrangement of the solvent molecule(s) was not observed for any other system studied. The planar structure of the base pair with a hydrogen atom of $CHCl_3$ situated in the plane of the base pair is usually less stable by 1-4 kcal/mol than the nonplanar one (see the nonmethylated and methylated HB AT and GC structures in the presence of one or two molecules of $CHCl_3$ in Figure 4) The most stable T and S structures are less stable.

Table 1. The BSSE corrected interaction energies (deformation energies included) obtained at the RI-MP2/TZVPP//RI-MP2/cc-pVDZ level of theory (in kcal/mol) of the most stable structures of the nonmethylated and methylated adenine...thymine complexes with or two molecule(s) of solvent (water, CH_3OH, DMSO and $CHCl_3$).

mol	water	CH_3OH	DMSO	$CHCl_3$
AT-1mol	-28.05	-29.17	-31.87	-21.22
AT-2mol	-37.20	-39.33	-46.86	-31.77
mAmT-1mol	-26.73	-25.64	-27.02	-23.68
mAmT-2mol	-35.75	-37.49	-40.55	-32.31
GC-1mol	-34.87	-36.25	-38.40	-32.05
GC-2mol	-45.12	-45.68	-52.91	-41.91
mGmC-1mol	-33.68	-33.56	-37.99	-33.53
mGmC-2mol	-44.93	-46.18	-50.26	-46.85

Comparison between ab initio and empirical calculations

We have also compared the interaction energies calculated empirically and quantum chemically. We observed that the accurate quantum chemical calculations verified the empirical results qualitatively (the geometries obtained from MD/Q simulations did not change during the ab inito optimization significantly) as well as quantitatively (match between the stabilization energies).

There were observed only slight deplanarization (by 10-30°) of some planar HB structures and some minor changes in organization of solvent molecules during the *ab initio* optimization. In empirical potential geometries, all atoms of water or methyl group in CH_3OH molecules lie in planar HB structures in the bases planes, whereas after the optimization the water hydrogen atoms or methyl group of CH_3OH are often rotated outside the bases planes.

Also from the energetic point of view, there is a good agreement between the empirical and *ab initio* data. The empirical force field correctly detected the global minima at the PESs. The empirical stabilization energies lie between the cc-pVDZ and TZVPP *ab initio* values, which means that they are still underestimated with respect to the CBS limit.

The behaviour of AT and GC WC base pairs

Due to the major biological importance of AT and GC WC base pattern we focused on its behaviour after methylation and addition of second water or organic solvent molecule. The nonmethylated GC WC structure appeared to be the global minima at all GC PESs in presence of one water, CH_3OH or DMSO molecule. The methylation or/and addition of other solvent molecule leads to the balancing of the WC HB and S complex stability. At the GC-1$CHCl_3$ PESs the WC structure occurs as a global minimum only after methylation. Contrary to the behaviour of GC WC complexes, the AT WC structure was never observed as a global minimum at any nonmethylated or methylated AT PESs.

More detailed information concerning to this study can be found in the Appendices A, B, C.

4.2 Hydrophobic Base Analogues

The experimental crystal geometries of complexes studied do not exist, thus the geometries of 13mer of DNA double helix (5'-GCGTAC**A**CATGCG-3') were determined by MD simulations using Cornell et al. empirical force-field.[59] The bases in the central base pair (shown in bold) were replaced (one or both) by a series of hydrophobic base analogues (phenyl, biphenyl, phenyl-naphatalene, phenyl-anthracene and phenyl-phenanthrene) (as an example cf. Figures 5, 6).

However, several complications followed: the obtained geometries from MD simulations were too flexible to be reliably geometrically averaged. Thus, the calculations of interaction energies were performed on the sets of geometries obtained from MD by using SCC-DFTB-D, which underestimates slightly the H-bonded

interactions, but for stacked systems it provides reasonable results. The set of interaction energies were averaged subsequently.

To evaluate and analyze the nature of stability of modified DNA duplexes we calculated several types of interaction energies using two models:

Model A (three central steps of DNA duplex without sugar phosphate backbone)

i) total interaction energies (defined as a energy released upon separation of all six subsystems to infinity)

ii) particular pair interaction energies (defined as a energy between two bases, interstrand and intrastrand).

Model B (three central steps of DNA duplex including sugar phosphate backbone)

i) intrastrand interaction energies (defined as a interaction energy between two strands)

ii) intrastrand desolvation energies (defined as a necessary energy for desolvation of both solvated strands in order to form the solvated duplex)

Geometries and interaction, desolvation energies of systems studied

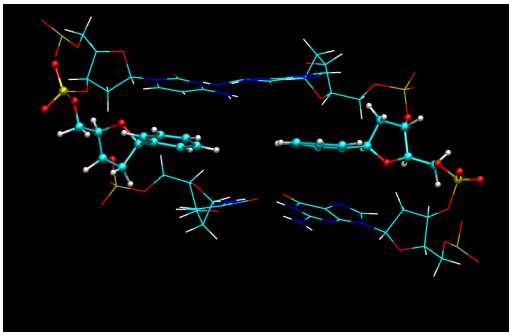
It is clear that the modified central base pair characterizes the systems. In the most of the systems studied the stack arrangement (as compensation of the lack of H-bonds) of the central base pair is created (for example cf. C-G/$\mathbf{N}_{min}$-$\mathbf{N}_{maj}$/C-G and C-G/$\mathbf{H}_{min}$-$\mathbf{H}_{maj}$/C-G selfpairs in Figure 9, symbol '/' dividing subsystems means stacking interaction, '-' symbol means H-bonding interaction). However, there are few exceptions. The systems with modified nucleobase **P** (cf. C-G/**P**-**P**/C-G in Figure 9), and half modified by modified nucleobase **B** (cf. Figures 6), are not large enough to form stack. Further, the DNA duplexes modified by **P** are only ones, where the bases in central base pairs lie in the plane in spite of lack of H-bonds. In the case of half modified DNA duplex by modified nucleobase **B**, the one of the bases in central pair of DNA is forced out from the duplex. In contrast, in the systems containing two **D** in specific mutual orientation, the modified nucleobase **D** are too long to be both accommodated inside the duplex and again one of the modified nucleobase **D** is forced out from the duplex (cf. C-G/$\mathbf{D}_{min}$-$\mathbf{D}_{maj}$/C-G in Figure 9). From these

observations we can conclude that the creation of stacked structures among others depends on the length of modified nucleobase **X**.

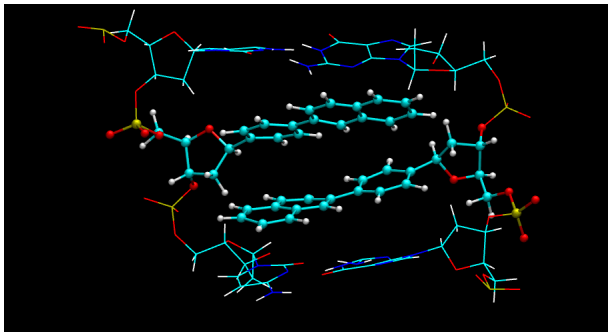
The important factor determining the strength of stacking interaction between the bases/modified nucleobases **X** is the extent of the overlap between stacked bases/modified nucleobases **X**. The larger (more aromatic rings) selfpairs show higher stability (with interaction energies from approx. 12 up to 16 kcal/mol), while the systems with the smaller modified nucleobase **B** or half substituted systems show lower stability (with stabilization energies about 9 kcal/mol), which is comparable to natural H-bonded A-T base pair. Apart from the number of aromatic rings, also their configuration plays important role for the stacking interaction of particular base analogue. For example, the DNA duplex with modified nucleobase **D** with linear configuration of aromatic rings shows lower stacking interaction then the DNA with modified nucleobase **H** with nonlinear configuration, even the both modified nucleobases **D** and **H** have the same number of aromatic rings. As the most stable system was observed the DNA with central part created from C-G/$\mathbf{H_{maj}}$-$\mathbf{H_{min}}$/C-G with the total stabilization energy -94.09 kcal/mol (model A).

Further, we estimated the overall stability of the DNA duplexes as a sum of interstrand interaction and desolvation energies (Model B). The system with modified nucleobases $\mathbf{H_{maj}}$-$\mathbf{H_{maj}}$ selfpair was observed as the most stable duplex with stabilization energy about 7 kcal/mol followed by systems with modified nucleobases $\mathbf{N_{maj}}$-$\mathbf{N_{maj}}$ and $\mathbf{D_{maj}}$-$\mathbf{D_{min}}$ selfpairs. For comparison, the stabilization energy of unmodified DNA duplex is about 1 kcal/mol.

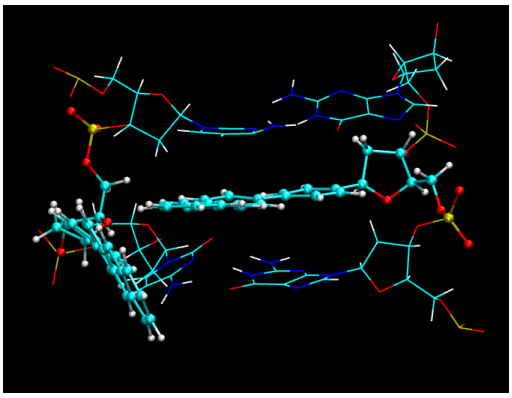
However, the most significant feature of the quality of duplex modified by base analogues is the thermodynamic selectivity to form the selfpair. We evaluated this as a difference between the stability of least stable selfpairs and most stable half-substituted misspairs. The largest difference and therefore the highest selectivity was observed for systems with modified nucleobase **D**. Thus, we suggested the modified nucleobase **D** as the best functional base analogue among all systems studied.



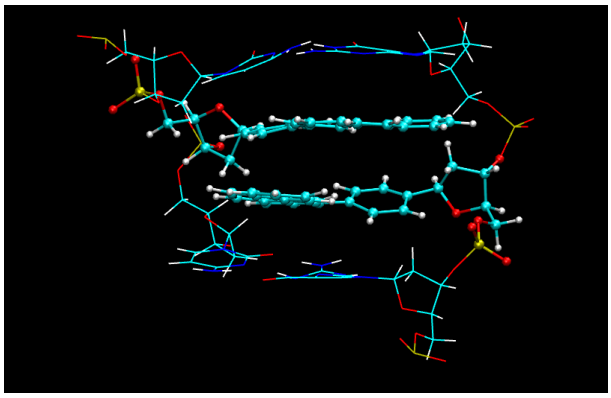
C-G/**P**-**P**/C-G



C-G/$\mathbf{N}_{min}$-$\mathbf{N}_{maj}$/C-G



C-G/$\mathbf{D}_{min}$-$\mathbf{D}_{maj}$/C-G



C-G/$\mathbf{H}_{min}$-$\mathbf{H}_{maj}$/C-G

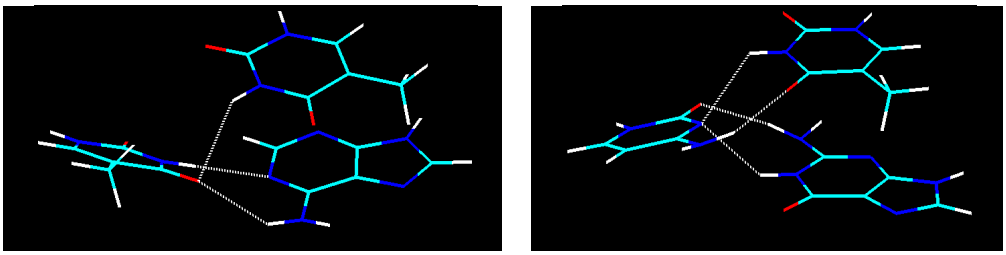
Figure 9. The geometries of central part of 13mer of chosen perturbated DNA duplexes after 8 ns of MD simulation.

Fork-like structures

We observed in central part of some DNA duplexes (namely where adenine in the central base pair was replaced by modified nucleobase **X**) an interesting geometrical feature: creation of so-called "fork-like" structures (cf. Figure 10). This effect comes up due the distortion caused by larger modified nucleobase **X** accommodated inside the DNA duplex. We have studied such systems in more details using C-G/**X**-T/C-G, C-G/**X**-T/T-A, C-G/**X**-T/G-C and C-G/**X**-T/A-T combination of bases (symbol '/' dividing subsystems means stacking interaction, '-' symbol means H-bonding interaction). The geometries of the studied system with modified nucleobase **N** were generated using MD simulation, where the startup geometry has already contained fork-like structure. The fork-like structures were stable during the whole MD simulations, when the pyrimidine base established two H-bonds with both bases on other strand (C-G/**N**-T/C-G and C-G/**N**-T/T-A).

We observed that the C-G/**N**-T/T-A structures benefit from "fork-like" arrangement since the newly created T-T interaction overbalanced the weakened T-A interaction. This is in the contrast to C-G/**N**-T/C-G, where the fork-like structures are

less stable than unperturbated ones. This can be explained by the fact that the H-bonded interaction of the unperturbated C-G pair is much stronger compare to T-A pair and the respective loose is not compensated by relatively weak newly created C-T pair interaction.



T-T/A C-T/G

Figure 10. The geometrical T-T/A and C-T/G "fork-like" structures obtained from MD simulations (superimposed by *ab initio* interbase geometries).

Comparison with experiment

The experimental measurements of melting point were provided only for DNA duplexes containing selfpair of modified nucleobases **B**-**B**[60],[61],[62] and A-**B** misspair.[63] The experimental modified DNA duplexes had different length and different composition than these of that investigated in our study. Experimental measures shows that the DNA duplex modified by **B**-**B** selfpair is less stable then the unmodified DNA (by about 5.4 °C), which is in disagreement with our calculations, where the modified **B**-**B** selfpair was calculated more stable then the unmodified DNA (by about 2 kcal/mol). However, the both experimental and theoretical results shows that DNA duplex modified by misspair A-**B** is significantly less stable then the unmodified DNA (by about 8.4 °C and 7.5 kcal/mol) and DNA modified by **B**-**B** selfpair.

More detailed information concerning to this study can be found in the Appendix D.

4.3 The Derivatives of Nucleobases Linked with Transition Metals Complexes

In this study we compared the experimental values of redox potentials of Ru^{2+} and Os^{2+} complexes with the corresponding quantum chemical calculations (using DFT as a quantum chemical method and COSMO as the solvation method).

The Gibbs free energy was calculated as the sum of these contributions:

$$G = E_{el} + G_{solv} + E_{ZPE} - RT \ln(q_{trans}\, q_{rot}\, q_{vib}) \quad (18)$$

where E_{el} is the energy of system in vacuo (at the B3LYP/TZVP level and the geometry optimized at the RI-PBE/def2-SVP level), G_{solv} is the solvation free energy (at the RI-PBE/def2-SVP level), E_{ZPE} is the zero-point energy, and $-RT \ln(q_{trans}\, q_{rot}\, q_{vib})$ accounts for the entropic terms and the thermal correction to the enthalpy obtained from a frequency calculation; at the RI-PBE/def2-SV(P) level, 298 K and 1 atm, using the ideal-gas approximation

The reduction potentials were calculated according to the equation:

$$E^0\,[\mathrm{V}] = 27.21(G_{ox}\,[\mathrm{a.u.}] - G_{red}\,[\mathrm{a.u.}]) - 4.43\ \mathrm{V} \quad (19)$$

where $G_{ox/red}$ are free energy values calculated according to Eq. (18), and 4.43 V is an absolute redox potential of the standard hydrogen electrode (SHE).

Apart from the calculation of reduction potentials we also investigated the structural changes associated with the oxidation and reduction of the studied complexes and spin densities that helped to assign the electronic origin of the redox process.

Oxidation: We have shown that the calculated and experimental results of reduction potentials for oxidation of Ru^{2+} and Os^{2+} complexes are in good agreement (with a standard deviation of 130 mV, considering that 100mV corresponds approx.

10kJ/mol), (cf. Table 2). The systematic (mean) error of –60 mV can be attributed to several factors:

i) neglect of spin-orbit coupling (SOC) in the nonrelativistic treatment which has been estimated to account for approximately +30 mV on the model $Ru^{2+/3+}$ complex. However, this effect is by one order of magnitude greater for Os^{2+} complexes (+350 mV) and cannot be neglected.

ii) the single reference treatment of the triple (quasi)degenerate ground state in $[Ru(bipy)_3]^{3+}$ complexes

iii) the uncertainty in the absolute potential of SHE

iv) the difference between the theoretical calculations and the experimental setup (presence of counterions and the adsorption of complexes on the electrode).

The analysis of the spin densities have shown that for two complexes (11c, 11d, cf. Figure 8), the unpaired electron is not localized on the Ru atom as in all other cases, but rather resides on the adenine moiety.

Reduction: The interpretation of calculated reduction potentials, i.e., the assignment of the formal redox states to subsystems in the studied complexes is more complicated. But, in general the calculated data were in the range of experimental values (cf. Table 2).

The spin densities are strongly delocalized on the ligands, which suggest that these redox potential values do not refer to the reduction of the metal center, in agreement with the experimental data.

Finally, it can be mentioned that we observed only very small geometry changes associated with the $Ru^{2+/3+}$ oxidation and the $[RuX_n]^{2+/1+}$ reduction.

Table 2. The calculated redox potentials of the synthesized compounds (vs. Ag/AgCl/3 M KCl reference electrode).

Compound	**Oxidation $Ru^{2+/3+}$ ($Os^{2+/3+}$)**			**Reduction**
	E^0_{calc} (V)[a]	ΔE^0_{calc} (mV)[b]	ΔE^0_{exp} (mV)	E^0_{calc} (V)
$[Ru(bipy)_3]^{2+}$	1.082a	0	*0*	–1.450
5c	1.074	–8	*80*	–1.557
10c	1.183	101	*115*	–1.144
10b	1.238	156	*115*	–1.357
6c	1.368	286	*70*	–1.429
11c	0.987	–95	*115*	–1.292
11b	1.130	48	*115*	–1.411
10a	1.165	83	*(125)*	–1.234
11a	1.189	107	*(120)*	–1.484
11d	0.960	–122	*(170)*	–1.389
12b	1.127	45	*(–305)*	–1.498

[a] E^0_{calc} were calculated using the B3LYP/TZVP//PBE/def2-SVP energies, PBE/def2-SV(P) frequencies, ideal gas approximation for thermodynamic functions, the COSMO solvation model, 4.34 V as the absolute potential of the SHE, and 0.207 V as the potential of the reference electrode used in the electrochemical measurements.[b] $\Delta E^0_{calc/exp}(X) = E^0_{calc/exp}(X) - E^0_{calc/exp}([Ru(bipy)_3]^{2+})$

More detailed information concerning to this study can be found in the Appendices E, F.

5 Conclusions

The aim of this study was twofold:

i) a) to investigate and compare stacked and H-bonding interaction between natural base pairs and few molecules of solvent in vacuo.

b) to investigate and compare stacked and H-bonding interaction between unnatural base analogues built in the DNA duplex in the presence of solvent.

ii) to investigate the electrochemical properties of the derivatives of nucleobases linked with transition metals (Ru^{2+}, Os^{2+}) complexes.

In the ia) part of present work we compared the stability and molecular interaction between adenine…thymine, guanine…cytosine and their methylated analogues with a small number of water and organic solvent molecules (CH_3OH, DMSO, $CHCl_3$).

We observed that for nonmethylated and methylated guanine…cytosine with water or organic solvent molecule(s), the HB WC structure appeared to be systematically a global minimum at almost all PESs (with exception of GC-1$CHCl_3$), whereas the WC structure has never been located as a global minimum on the PESs of AT complexes. We have also demonstrated easier hydration of stacked systems than H-bonded.

Further, we observed completely different interactions between the bases and the solvents studied. Whereas water and CH_3OH can stabilize the S structures of the base pairs by a higher number of H-bonds than is possible in H-bonded pairs, the $CHCl_3$ molecule lacks such a property, and the HB structures with molecule(s) of solvent situated above or below the base pair are preferred. The DMSO molecule is unique by its dimension in comparison with other solvents, and the T structures are the most abundant ones.

In the ib) part of present work based on the MD, SCC-DDFTB-D and COSMO calculations of modified nucleobases **X** incorporated to DNA duplex we made following conclusions:

i) Replacing nucleic acid base by modified nucleobase **X** leads mostly to structural changes of the central base pair (stack arrangement of central

modified base pairs). Only with the smallest modified nucleobase **P** the central base pairs (A-**P**, **P**-T) stay planar. In the case of **B**-T, A-**B** or **D**-**D** in specific orientation, one of the modified nuclebase was forced out from DNA duplex.

ii) Increasing aromaticity of modified nucleobases **X** increases the stacking stabilization.

iii) The highest selectivity among all modified nucleobases **X** studied was found for phenyl-anthracene.

In the second part of present work, we used for the calculation of reduction potentials of derivatives of nucleobases linked with transition metals (Ru^{2+}, Os^{2+}) complexes the DFT and COSMO methods and compared our results with experimental values. We observed, that in the case of Ru^{2+} complexes the calculated reduction potentials of oxidation agreed quite well with experimental values, whereas in the case of Os^{2+} complexes, the quite strong disagreement was observed. This can be explained by the spin orbit coupling, which was neglected in our calculations and plays in the Os^{2+} complexes more significant role, than in Ru^{2+} complexes. The interpretation of calculated reduction potentials was more complicated, but in general, the calculated data were in the range of experimental values.

References

1. P. Hobza, J. Šponer, and T. Reschel, *Density-Functional Theory and Molecular Clusters.* Journal of Computational Chemistry, 1995. 16(11): p. 1315-1325.
2. P.K. Weiner and P.A. Kollman, *Amber - Assisted Model-Building with Energy Refinement - a General Program for Modeling Molecules and Their Interactions.* Journal of Computational Chemistry, 1981. 2(3): p. 287-303.
3. J.M. Wang *et al*., *Development and testing of a general amber force field (vol 25, pg 1157, 2004).* Journal of Computational Chemistry, 2005. 26(1): p. 114-114.
4. C.I. Bayly *et al*., *A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model.* Journal of Physical Chemistry, 1993. 97(40): p. 10269-10280.
5. L.J. Onsager, *Born Onsager Cosmo Model.* Journal of American Chemical Society, 1936. 58(1486).
6. O. Tapia and O. Goscinski, *Self-Consistent Reaction Field-Theory of Solvent Effects.* Molecular Physics, 1975. 29(6): p. 1653-1661.
7. S. Miertus, E. Scrocco, and J. Tomasi, *Electrostatic Interaction of a Solute with a Continuum - a Direct Utilization of Abinitio Molecular Potentials for the Prevision of Solvent Effects.* Chemical Physics, 1981. 55(1): p. 117-129.
8. A. Klamt and G. Schuurmann, *Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient.* Journal of the Chemical Society-Perkin Transactions 2, 1993(5): p. 799-805.
9. a) D.R. Hartree, *The Wave Mechanics of An Atom with a Non-coulomb Central Field. Part I –Theory and Methods.* Proc. Cambridge Phil. Soc., 1928. 24(89): p. 111. b) D.R. Hartree, *The Wave Mechanics of An Atom with a Non-coulomb Central Field. Part II – Some Results and Discussion.* Proc. Cambridge Phil. Soc., 1928. 24(111): p. -.
10. V. Fock, - Z. Physic, 1930. 61(126).
11. C. Møller and M.S. Plesset, Physical Reviews, 1934. 46(7): p. 618-622.
12. J. Čížek, *On Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods.* Journal of Chemical Physics, 1966. 45(11): p. 4256-&.
13. J. Čížek and J. Paldus, *Coupled Cluster Approach.* Physica Scripta, 1980. 21(3-4): p. 251-254.
14. P.O. Löwdin, *Quantum Theory of Many-Particle Systems .1. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction.* Physical Review, 1955. 97(6): p. 1474-1489.
15. P.O. Löwdin, *Quantum Theory of Many-Particle Systems .2. Study of the Ordinary Hartree-Fock Approximation.* Physical Review, 1955. 97(6): p. 1490-1508.
16. P.O. Löwdin, *Quantum Theory of Many-Particle Systems .3. Extension of the Hartree-Fock Scheme to Include Degenerate Systems and Correlation Effects.* Physical Review, 1955. 97(6): p. 1509-1520.

17. P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas.* Physical Review B, 1964. 136(3B): p. B864-&.
18. S. Grimme, *Accurate description of van der Waals complexes by density functional theory including empirical corrections.* Journal of Computational Chemistry, 2004. 25(12): p. 1463-1473.
19. P. Jurečka, *et al*., *Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations.* Journal of Computational Chemistry, 2007. 28(2): p. 555-569.
20. J.P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple.* Physical Review Letters, 1996. 77(18): p. 3865-3868.
21. A.D. Becke, *Density-Functional Thermochemistry .3. The Role of Exact Exchange.* Journal of Chemical Physics, 1993. 98(7): p. 5648-5652.
22. J.M. Tao, *et al*., *Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids.* Physical Review Letters, 2003. 91(14): p. -.
23. M. Elstner, *et al*., *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties.* Physical Review B, 1998. 58(11): p. 7260-7268.
24. M. Elstner, *et al*., *Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment.* Journal of Chemical Physics, 2001. 114(12): p. 5149-5155.
25. T.J. Martinez and E.A. Carter, *Pseudospectral Multireference Single and Double-Excitation Configuration-Interaction.* Journal of Chemical Physics, 1995. 102(19): p. 7564-7572.
26. C. Hampel and H.J. Werner, *Local treatment of electron correlation in coupled cluster theory.* Journal of Chemical Physics, 1996. 104(16): p. 6286-6297.
27. O. Vahtras, J. Almlof, and M.W. Feyereisen, *Integral Approximations for Lcao-Scf Calculations.* Chemical Physics Letters, 1993. 213(5-6): p. 514-518.
28. S.F. Boys and F. Bernardi, *Calculation of Small Molecular Interactions by Differences of Separate Total Energies - Some Procedures with Reduced Errors.* Molecular Physics, 1970. 19(4): p. 553-&.
29. T.H. Dunning, *Gaussian-Basis Sets for Use in Correlated Molecular Calculations .1. The Atoms Boron through Neon and Hydrogen.* Journal of Chemical Physics, 1989. 90(2): p. 1007-1023.
30. A. Halkier, *et al*., *Basis-set convergence in correlated calculations on Ne, N-2, and H2O.* Chemical Physics Letters, 1998. 286(3-4): p. 243-252.
31. A. Halkier, *et al*., *Basis-set convergence of the energy in molecular Hartree-Fock calculations.* Chemical Physics Letters, 1999. 302(5-6): p. 437-446.
32. T.H. Dunning and K.A. Peterson, *Approximating the basis set dependence of coupled cluster calculations: Evaluation of perturbation theory approximations for stable molecules.* Journal of Chemical Physics, 2000. 113(18): p. 7799-7808.
33. J.D. Watson and F.H.C. Crick, *Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid.* Nature, 1953. 171(4356): p. 737-738.
34. E.T. Kool, J.C. Morales, and K.M. Guckian, *Mimicking the structure and function of DNA: Insights into DNA stability and replication.* Angewandte Chemie-International Edition, 2000. 39(6): p. 990-1009.

35. E.T. Kool, *Replacing the nucleohases in DNA with designer molecules.* Accounts of Chemical Research, 2002. 35(11): p. 936-943.
36. A.A. Henry, C.Z. Yu, and F.E. Romesberg, *Determinants of unnatural nucleobase stability and polymerase recognition.* Journal of the American Chemical Society, 2003. 125(32): p. 9638-9646.
37. J.D. Bain, *et al*., *Ribosome-Mediated Incorporation of a Nonstandard Amino-Acid into a Peptide through Expansion of the Genetic-Code.* Nature, 1992. 356(6369): p. 537-539.
38. Y.Tor and P.B. Dervan, *Site-Specific Enzymatic Incorporation of an Unnatural Base, N(6)-(6-Aminohexyl)Isoguanosine, into Rna.* Journal of the American Chemical Society, 1993. 115(11): p. 4461-4467.
39. Y.Q. Wu, *et al*., *Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions.* Journal of the American Chemical Society, 2000. 122(32): p. 7621-7632.
40. A.K. Ogawa, *et al*., *Rational design of an unnatural base pair with increased kinetic selectivity.* Journal of the American Chemical Society, 2000. 122(36): p. 8803-8804.
41. K.M. Guckian, T.R. Krugh, and E.T. Kool, *Solution structure of a nonpolar, non-hydrogen-bonded base pair surrogate in DNA.* Journal of the American Chemical Society, 2000. 122(29): p. 6841-6847.
42. D.R. Staveren and N. Metzler-Nolte, Chemical Reviews, 2004. 104: p. 55.
43. V. Balzani, *et al*., *Luminescent and redox-active polynuclear transition metal complexes.* Chemical Reviews, 1996. 96(2): p. 759-833.
44. M. Fojta, *Electrochemical sensors for DNA interactions and damage.* Electroanalysis, 2002. 14(21): p. 1449-1463.
45. M. Fojta, *Mercury electrodes in nucleic acid electrochemistry: Sensitive analytical tools and probes of DNA structure. A review.* Collection of Czechoslovak Chemical Communications, 2004. 69(4): p. 715-747.
46. M. Fojta, *in Elektrochemistry of Nucleic Acids and Proteins. Towards Electrochemical Sensors for Genomics and Protetomics.* Elesevier, Amsterodam, 2005: p. 44.
47. E. Paleček and M. Fojta, Bioelectronics, 2005: p. 65.
48. E. Paleček and M. Fojta, *Detecting DNA hybridization and damage.* Analytical Chemistry, 2001. 73(3): p. 74A-83A.
49. N. Popovich and H. Thorp, Interface, 2002. 11: p. 5.
50. J. Wang, *in Electrochemistry of Nucleic Acids and Proteins. Towards electrochemical Sensors for Genomics and Proteomics.* Elsevier, Amsterodam, 2005: p. 16.
51. M. Fojta, *et al*., *Two-surface strategy in electrochemical DNA hybridization assays: Detection of osmium-labeled target DNA at carbon electrodes.* Electroanalysis, 2003. 15(5-6): p. 431-440.
52. M. Fojta, *et al*., *Multiply osmium-labeled reporter probes for electrochemical DNA hybridization assays: detection of trinucleotide repeats.* Biosensors & Bioelectronics, 2004. 20(5): p. 985-994.
53. E. Paleček, M. Fojta, and F. Jelen, *New approaches in the development of DNA sensors: hybridization and electrochemical detection of DNA and RNA at two different surfaces.* Bioelectrochemistry, 2002. 56(1-2): p. 85-90.
54. P.K. Bhattacharya, H.J. Lawson, and J.K. Barton, Inorganic Chemistry, 2003(42): p. 8.

55. C. Stinner, *et al*., *Synthesis and spectroelectrochemistry of Ir(bpy)(phen)(phi)(3+), a tris(heteroleptic) metallointercalator.* Inorganic Chemistry, 2001. 40(20): p. 5245-5250.
56. J.L. Kisko and J.K. Barton, *Recognition of DNA base pair mismatches by a cyclometalated Rh(III) intercalator.* Inorganic Chemistry, 2000. 39(21): p. 4942-4949.
57. R.E. Holmlin, J.A. Yao, and J.K. Barton, *Dipyridophenazine complexes of Os(II) as red-emitting DNA probes: Synthesis, characterization, and photophysical properties.* Inorganic Chemistry, 1999. 38(1): p. 174-189.
58. S.J. Franklin, C.R. Treadway, and J.K. Barton, *A reinvestigation by circular dichroism and NMR: Ruthenium(II) and rhodium(III) metallointercalators do not bind cooperatively to DNA.* Inorganic Chemistry, 1998. 37(20): p. 5198-5210.
59. W.D. Cornell, *et al*., *A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules.* Journal of the American Chemical Society, 1995. 117(19): p. 5179-5197.
60. C. Brotschi and C.J. Leumann, *RNA duplexes with biphenyl substituents as base replacements are less stable than DNA duplexes.* Chemical Communications, 2005(15): p. 2023-2025.
61. C. Brotschi and C.J. Leumann, *DNA with hydrophobic base substitutes: A stable, zipperlike recognition motif based on interstrand-stacking interactions.* Angewandte Chemie-International Edition, 2003. 42(14): p. 1655-1658.
62. A. Zahn, C. Brotschi, and C.J. Leumann, *Pentafluorophenyl-phenyl interactions in biphenyl-DNA.* Chemistry-a European Journal, 2005. 11(7): p. 2125-2129.
63. C. Beuck, *et al*., *Polycyclic aromatic DNA-Base surrogates: High-affinity binding to an adenine-specific base-flipping DNA methyltransferase.* Angewandte Chemie-International Edition, 2003. 42(33): p. 3958-3960.
64. R. Ahlrichs, *et al*., *Electronic-Structure Calculations on Workstation Computers - the Program System Turbomole.* Chemical Physics Letters, 1989. 162(3): p. 165-169.
65. M. J. Frisch, G.W.T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2003., *Gaussian 03, Revision A.1*.
66. H.-J. Werner, P.J.K., R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R.

Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and H.-J. Werner., *MOLPRO version 2002.1, a package of ab initio programs*.

Molecular Modeling Packages Used

The RI-MP2 and DFT calculations were performed in the TURBOMOLE (versions 6-9) program package.[64] The COSMO calculations were performed in the TURBOMOLE (versions 6-9) and Gaussian03 programs.[65] The CCSD(T) calculations were performed in the MOLPRO (version 2002.1) program package.[66] The SCC-DFTB-D calculations were performed in the DFTB program.[23] The empirical force-fields calculations were performed in the AMBER package (versions 5-8).[2]

Appendices

Appendix A: Kabeláč, M; Zendlová, L; Řeha, D; *et al*.:Potential energy surfaces of an adenine-thymine base pair and its methylated analogue in the presence of one and two water molecules: Molecular mechanics and correlated *ab initio* study
JOURNAL OF PHYSICAL CHEMISTRY B Volume: 109 Issue: 24
Pages: 12206-12213 Published: JUN 23 2005

Appendix B: Zendlová, L; Hobza, P; Kabeláč, M;:Potential energy surfaces of the microhydrated guanine(...)cytosine base pair and its methylated analogue
CHEMPHYSCHEM Volume: 7 Issue: 2 Pages: 439-447
Published: FEB 13 2006

Appendix C: Zendlová, L; Hobza, P; Kabeláč, M;: Stability of nucleic acid base pairs in organic solvents: Molecular dynamics, molecular dynamics/quenching, and correlated *ab initio* study
JOURNAL OF PHYSICAL CHEMISTRY B Volume: 111 Issue: 10
Pages: 2591-2609 Published: MAR 15 2007

Appendix D: Zendlová, L; Řeha, D; Hocek, M; Hobza, P;: Theoretical study of the stability of DNA duplexes modified by series of hydrophobic base analogues
CHEMISTRY- A EUROPEAN JOURNAL, in preparation

Appendix E: Vrábel, M; Hocek, M; Havran, L; Fojta, M; Votruba, I; Klepetářová, B; Pohl, R; Rulíšek, L; Zendlová, L; Hobza, P; Shih, I; Mabery, E; Mackman, R;: Purines bearing phenanthroline or bipyridine ligands and their Ru-II complexes in position 8 as model compounds for electrochemical DNA labeling - Synthesis, crystal structure, electrochemistry, quantum chemical calculations, cytostatic and antiviral activity
EUROPEAN JOURNAL OF INORGANIC CHEMISTRY Issue: 12
Pages: 1752-1769 Published: APR 2007

Appendix F: Srnec, M; Chalupský, J; Zendlová, L; Hocek, M; Havran, L; Fojta, M; Kývala, M; Rulíšek, L;: An effect of spin-orbit coupling on the values of reduction potentials of octahedral ruthenium (II/III) and osmium (II/III) complexes
submitted to JOURNAL OF AMERICAN CHEMICAL SOCIETY