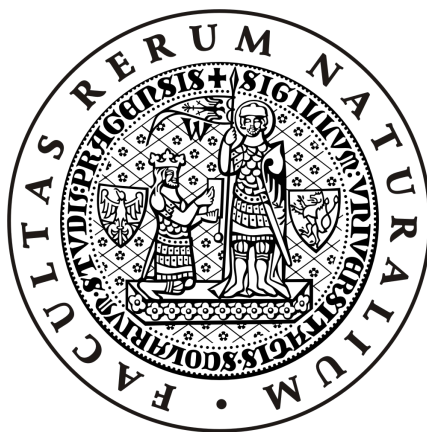


Charles University in Prague  
Faculty of Science  
Department of Physical and Macromolecular Chemistry



Bachelor Diploma Thesis

# Structure and Stabilization Energy of Phenol Dimer

Michal Kolář

2007

Advisor: Prof. Ing. Pavel Hobza, DrSc.  
Academy of Science, Czech Republic  
Institute of Organic Chemistry and Biochemistry  
Center for Biomolecules and Complex Molecular Systems

Čestně prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně, pouze za použití citované literatury. Dále prohlašuji, že v této bakalářské práci prezentuji výsledky své práce a nebo práce, na níž jsem se podílel.

Tato bakalářská práce ani žádná její část nebyly dříve použity k získání jakéhokoliv akademického titulu.

V Praze, 8. června, 2007

Michal Kolář

I would like to thank Prof. Pavel Hobza for introducing me into the field of computational chemistry as well as for his patient guidance through all delights and problems I have been discovering.

Many thanks go to Dr. Tomáš Kubař for his endurance in answering thousands of my questions and to the people in Canon building for the comfortable working environment.

I also thank my family, which has been supporting me during my studies.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Computational chemistry . . . . .	5
1.2	Motivation . . . . .	5
<b>2</b>	<b>Methods</b>	<b>7</b>
2.1	Essentials of quantum mechanics . . . . .	7
2.2	Hartree-Fock method, SCF procedure . . . . .	8
2.3	Møller-Plesset perturbation theory . . . . .	10
2.4	Coupled Clusters . . . . .	11
2.5	Density functional theory . . . . .	12
2.5.1	Exchange-Correlation functional . . . . .	14
2.5.2	DFT-D . . . . .	15
2.6	Basis sets . . . . .	15
2.7	Stabilization energy, interaction energy and BSSE . . . . .	16
<b>3</b>	<b>Computation</b>	<b>17</b>
3.1	HF, MP2 and DFT calculations . . . . .	17
3.2	DFT-D calculations . . . . .	17
3.3	Geometrical properties . . . . .	18
3.4	CCSD(T) geometry scan . . . . .	19
<b>4</b>	<b>Results and discussion</b>	<b>20</b>
4.1	Structure of phenol dimer . . . . .	20
4.2	Phenol dimer optimization . . . . .	20
4.2.1	HF, MP2, B3LYP . . . . .	20
4.2.2	DFT-D optimization . . . . .	21
4.3	CCSD(T) scan results . . . . .	21
4.4	Stabilization energy . . . . .	23
<b>5</b>	<b>Conclusions</b>	<b>24</b>
<b>6</b>	<b>Data section</b>	<b>25</b>

# 1 Introduction

## 1.1 Computational chemistry

Chemistry, as an exact science, has a wide theoretical equipment to describe the chemically interesting processes. For a long time, it had not been nearly possible to theoretically describe chemical substances on the atomistic nor sub-atomistic level. Combination of quantum mechanics and electronic computers led to introduction of a new branch of chemistry – computational chemistry.

## 1.2 Motivation

“The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.”

*P. Dirac, 1929*

These for theoretical chemistry hopeless words forced chemists to use approximations, which are more or less suitable to describe the chemical systems. But if we are not able to solve the exact equations how we can be sure that approximate equations describe the system accurately? The only way is to compare theoretical data with experiment.

Comparing the theoretical and experimental data can tell us which approximations are worth using. Therefore the structural data obtained by experimental techniques are very important for theoretical chemists.

One of the target in our group is determination of structure and properties of biomolecules. Phenol dimer is an example of a system which is qualitatively very similar to e.g. base pairs in DNA. Structure is, in both cases, determined by non-covalent interactions. Both systems contain H-bond as well as  $\pi - \pi$  interactions. The former interaction is of electrostatic origin while the latter one is due to London dispersion energy. Calculation of electrostatic energy is straightforward while determination of dispersion energy is demanding.

Phenol dimer had been already studied [1] but the other paper supporting us with more accurate data was published [2] recently. This give us an opportunity to re-investigate the phenol dimer problem. By using different *ab initio* methods we determined optimized structures in a gas phase and interaction energies and compare these with the experimental results [2].

## 2 Methods

To determine phenol dimer structure we used various quantum chemical methods. These methods are based on few axioms postulated in the beginning of 20th century. The brief description of the methods follows.

### 2.1 Essentials of quantum mechanics

The physical state of any particle or system of particles can be described by WAVEFUNCTION  $\psi$ , which is a function of coordinates and time. This function needs to be continuous, to have continuous first and second derivatives, to be square integrable and to have norm equal 1. The wavefunction appears in the essential law of quantum mechanics – the SCHRÖDINGER EQUATION, which describes, how the wavefunction evolves in time.

$$\hat{H}\psi(\vec{r}, t) = i\hbar \frac{\partial\psi(\vec{r}, t)}{\partial t}, \quad (1)$$

where  $\hbar$  is the Dirac's constant,  $i$  is the imaginary unit and  $\hat{H}$  is a linear hermit operator called HAMILTONIAN, given by

$$\hat{H} = - \frac{\hbar^2}{2m} \sum_i \Delta_i - \quad (2)$$

$$- \frac{\hbar^2}{2} \sum_I \frac{\Delta_I}{M_I} - \quad (3)$$

$$- \frac{e^2}{4\pi\epsilon} \sum_{i,I} \frac{Z_I}{r_{i,I}} + \quad (4)$$

$$+ \frac{e^2}{4\pi\epsilon} \sum_{i<j} \frac{1}{r_{i,j}} + \quad (5)$$

$$+ \frac{e^2}{4\pi\epsilon} \sum_{I<J} \frac{Z_I Z_J}{r_{I,J}}, \quad (6)$$

where  $m$  is the mass of electron,  $M_I$  is the mass of  $I$ -th nucleus,  $e$  is the elementary charge,  $\pi$  is pi – Ludolph's number,  $\epsilon$  is permittivity,  $\Delta_i$  is the Laplace operator of  $i$ -th electron,  $Z_I$  is the proton number of  $I$ -th nucleus and  $r$  is the distance between two particles of the system.  $I, J$  are counting indices of nuclei and  $i, j$  are counting indices of electrons. The terms (2)

and (3) describe the kinetic energy of electrons and nuclei, respectively, and the terms (4), (5) and (6) describe coulombic interactions.

If we consider the time independent hamiltonian, the Schrödinger equation defined as Eqn. 1 simplifies into the TIME-INDEPENDENT SCHRÖDINGER EQUATION

$$\hat{H}\Phi = E\Phi, \quad (7)$$

which is a characteristic problem of the hamiltonian. Because the momentum of nuclei is negligible compared to the momentum of electrons, due to their very different masses, we can consider that the electrons move around *fixed* nuclei. This is the principle of the BORN-OPPENHEIMER APPROXIMATION (BOA). Having applied, we can define the ELECTRONIC HAMILTONIAN

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{en} + \hat{V}_{nn}, \quad (8)$$

where operators  $\hat{T}_e$ ,  $\hat{V}_{ee}$ ,  $\hat{V}_{en}$  and  $\hat{V}_{nn}$  correspond to the terms (2) – (6) (the only term of kinetic energy of nuclei is missing), and solve Schrödinger equation for given geometry of the nuclei

$$\hat{H}_e\phi_k(r, R) = E_k^e(R)\phi_k(r, R), \quad (9)$$

where  $\phi_k(r, R)$  is the wavefunction of  $k$ -th electronic state, depending on  $r$  and  $R$  – the coordinates of electrons and fixed nuclei, respectively, and where  $E_k^e(R)$  is electronic energy of  $k$ -th state according to the fixed nuclear geometry. By solving Eqn. 9 for different nuclear geometries we obtain parametric energy dependence on the molecular geometry. This dependence is called POTENTIAL ENERGY SURFACE.\* Born-Oppenheimer approximation is considered in all the following methods.

The commonly used approximate ab initio methods can be divided into two groups – variational and perturbation methods. Both method were used to study phenol dimer problem.

## 2.2 Hartree-Fock method, SCF procedure

The Hartree-Fock (HF) method is the example of variational method. Its basis had been set up in the late 1920's by Douglas Hartree [3] and Vladimir

---

\*also called *hypersurface* which emphasis the fact, that the energy depends on more variables



Alexandrovic Fock [4]. It was used to approximately solve the Schrödinger equation for many-body systems. These days there are many other methods, which are based on HF, trying to overcome its most important drawbacks.

Apart from Born-Oppenheimer approximation, HF method assumes the MEAN-FIELD APPROXIMATION. We suppose all the electrons are independent, so each electron moves in the effective mean potential field of the other particles (nuclei – which is constant due to BOA, and electrons).

Since it is a variational method, the energy we obtain from *any* guess wavefunction is larger then the exact energy corresponding to the exact wavefunction.

$$E_0 \leq \langle \phi_{guess} | \hat{H} | \phi_{guess} \rangle \quad (10)$$

In the HF method, we solve HARTREE-FOCK EQUATIONS

$$\hat{F}\Phi_k = \varepsilon_k \Phi_k, \quad (11)$$

where  $\Phi_k$  is the wavefunction of the  $k$ -th electron in the form of a SLATER DETERMINANT, which meets the Pauli exclusion principle

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \dots & \dots & \phi_N(N) \end{vmatrix} \quad (12)$$

where  $\phi_i$  is one-electron wavefunction.  $\hat{F}$  is FOCK OPERATOR. Fock operator is an effective hamiltonian for an electron in many-body system. It is defined as

$$\hat{F} = \hat{h}(1) + \sum_j [2J_j(1) - K_j(1)], \quad (13)$$

where  $\hat{h}(1)$  is one-electron CORE HAMILTONIAN,  $J_j(1)$  is the COULOMB OPERATOR and  $K_j(1)$  the EXCHANGE OPERATOR. The latter two operators include the mean-field principle and are constructed from the wavefunction of all electrons. This important feature leads to the iterative method for solution of HF equations. It can be summarized in a few points:

- We need a guess initial wavefunction to construct the Coulomb and exchange operators, as a part of the Fock operator.
- We solve the HF equations getting a new wavefunction and electronic energy.
- From the new wavefunction we construct a new Fock operator and solve HF equations
- This proceeds until the change of electronic energy is smaller than some pre-defined threshold. In that case, we say that the electronic field is *self-consistent*.

This is called the SELF-CONSISTENT FIELD (SCF) procedure. Unfortunately, the mean-field approximation in the SCF tends not to include energy coming from electron correlation, which seems to be very important. The correlation energy is defined as a difference between SCF energy (calculated by using the infinite basis set) and the exact electronic energy. Therefore other methods have been developed which take the correlation energy into account. These methods are collectively called POST HARTREE-FOCK METHODS.

### 2.3 Møller-Plesset perturbation theory

One of the ways, how to cover the correlation energy is to utilize the principles of perturbation theory. We consider that we are able to construct hamiltonian  $\widehat{H}$  as a sum of an unperturbed hamiltonian  $\widehat{H}_0$  and small perturbation  $\widehat{W}$

$$\widehat{H} = \widehat{H}_0 + \lambda\widehat{W}. \quad (14)$$

We know the solution for Schrödinger equation *Eqn. 7* for unperturbed hamiltonian, and we also know the small perturbation. Then we are able to write down energy and wavefunction as a sum of unperturbed term and perturbed terms.

$$E = E_0 + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} \dots \quad (15)$$

$$\phi = \phi_0 + \lambda\phi^{(1)} + \lambda^2\phi^{(2)} + \lambda^3\phi^{(3)} \dots, \quad (16)$$

where  $\lambda$  is a perturbation parameter. From these equations, the terms for corrections of energy and wavefunction can be derived.

The hamiltonian, in the Møller-Plesset perturbation theory [5], is represented as the sum of one-electron Fock operators (in a form of *Eqn. 13*)

$$\hat{H}_0 = \sum_i \hat{f}_i = \sum_i \hat{h}_i + \sum_{ij} (2J_{ij} - K_{ij}) \quad (17)$$

and the perturbation is derived as a difference between the exact and unperturbed hamiltonian.

With this approach the sum of zero- and first-order corrections yields the HF energy and to cover the correlation we have to add second-order correction. MP2 method invokes second-order corrections, likewise MP3 third-, MP4 fourth- etc. However, the second order energy correction already covers, depending on the system, ca 80 – 90 % of the correlation energy. Since it is not the variational method, we can not say anything about the energy we obtain. It can be higher or even lower than the exact energy corresponding to the exact wavefunction. Contributions of the higher order corrections somehow oscillate, so the third correction is much smaller, fourth higher etc.

## 2.4 Coupled Clusters

Currently, the most reliable method for involving the correlation energy, which yields highly accurate data while consuming acceptable computational time, is coupled clusters method. The theory, originally developed by *Coster* and *Kümmel* in 1950s, was introduced to the chemistry by *Čížek* and *Paldus* in late 1960s [6].

The theory adopts the wavefunction in the so-called “exponential ansatz”<sup>\*</sup> form

$$\Psi = e^{\hat{T}} \Phi_0, \quad (18)$$

where  $\Phi_0$  is a Slater determinant constructed from HF molecular orbitals and  $\hat{T}$  is a cluster operator which produces excited Slater determinants. The cluster operator is given as a sum of mono-, bi-, tri- etc. excitation operators

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots \quad (19)$$

---

<sup>\*</sup>in common use expression for the guess wavefunction comes from German “der Ansatz” – (basic) approach, statement, attempt

The operators  $\widehat{T}_i$  utilize a model of creation ( $\hat{a}$ ) and annihilation ( $\hat{a}^\dagger$ ) operators, and the first two are given as

$$\widehat{T}_1 = \sum_i \hat{t}_i = \sum_{ia} t_i^a \hat{a}_i^\dagger \hat{a}_a \quad (20)$$

$$\widehat{T}_2 = \frac{1}{2} \sum_{ij} \hat{t}_i \hat{t}_j = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_a \hat{a}_b. \quad (21)$$

The exponential factor in *Eqn. 18* is expanded into the series

$$e^{\widehat{T}} = 1 + \widehat{T} + \frac{\widehat{T}^2}{2!} + \frac{\widehat{T}^3}{3!} + \dots \quad (22)$$

In the practical calculations the single excitations yields, due to Brillouin's theorem, which says that the matrix element of hamiltonian between ground state and single excited state is equal zero, the HF energy. The double excitations (method CCD), single and double excitations (method CCSD) or triple excitations are usually applied. It means the cluster operator is usually in the form

$$\widehat{T} = \widehat{T}_2 \quad \text{CCD} \quad (23)$$

$$\widehat{T} = \widehat{T}_1 + \widehat{T}_2 \quad \text{CCSD} \quad (24)$$

$$\widehat{T} = \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 \quad \text{CCSDT} \quad (25)$$

As the best balance between accuracy and computational demands seems to be the CCSD(T) method, where the parenthesis signify, that the triple excitations are calculate by perturbation theory, while the single and double excitations are considered in the common way.

## 2.5 Density functional theory

Since the above mentioned methods use the wavefunction as a fundamental element of the interest, the density functional theory [7] describes a system with the electron density. The electron density  $\rho$  in the place  $\vec{r}$  is defined as

$$\rho(\vec{r}) = N \int \psi^*(\tau_1, \dots, \tau_N) \psi(\tau_1 \dots \tau_N) d\zeta_1 d\tau_2 \dots d\tau_N, \quad (26)$$

where  $\psi$  is a wavefunction which depends on spin  $\zeta_i$  and position  $\vec{r}_i$  of all particles of the system ( $\tau_i = \zeta_i \vec{r}_i$ ) and  $N$  is a normalization constant. Apparently, we convert the problem of  $3N$  variables to the problem of 3 variables, which dramatically decrease computational demands.

The theory is based on few theorems. In the first theorem Hohenberg and Kohn said that the electron density  $\rho(\vec{r})$  unambiguously determine the electrostatic potential  $v(\vec{r})$ . It means that the electron density defines  $\hat{H}$ , wavefunction, energy and all the other physical properties. Sometimes, this theorem is formulated in a different way: The energy is a functional of the electron density.

$$E = E[\rho(\vec{r})] \quad (27)$$

which means that there is an unique assignment between function (electron density) and number (energy). The energy functional is often written down as

$$E[\rho(\vec{r})] = \int \rho(\vec{r})v(\vec{r})d\vec{r} + F[\rho(\vec{r})], \quad (28)$$

where  $F[\rho(\vec{r})]$  is a functional which covers kinetic energy of electrons and their interactions. However, this functional  $F[\rho(\vec{r})]$  is not exactly known.

The second Hohenberg-Kohn theorem is, in fact, a variational principle for electron density. It says that the energy which belongs to *any* guess electron density  $\tilde{\rho}$  is higher than the exact energy corresponding to the exact electron density.

$$E_0 \leq E[\tilde{\rho}]. \quad (29)$$

This helped to Kohn and Sham to suggest the way, how to practically solve Hohenberg-Kohn problem. They define the functional  $F[\rho(\vec{r})]$  as

$$F[\rho(\vec{r})] = E_{kin}[\rho(\vec{r})] + E_H[\rho(\vec{r})] + E_{exc}[\rho(\vec{r})], \quad (30)$$

where  $E_{kin}[\rho(\vec{r})]$  describes kinetic energy of electrons,  $E_H[\rho(\vec{r})]$  describes Coulomb interaction between two electrons and the  $E_{exc}[\rho(\vec{r})]$  describes exchange and correlation effects. In the Kohn-Sham approach the electron density can be written by using orthonormal orbitals  $\phi_i(\vec{r})$  as

$$\rho(\vec{r}) = \sum_i^N |\phi_i(\vec{r})|^2, \quad (31)$$

and we solve Kohn-Sham equations

$$\widehat{H}_{KS}\phi_i = \varepsilon_i\phi_i, \quad (32)$$

where  $\widehat{H}_{KS}$  is an operator resulting from equations above. This approach tends to, similarly to SCF, to iterative process, which yields energy of the system.

### 2.5.1 Exchange-Correlation functional

The last term in *Eqn. 30* is called EXCHANGE-CORRELATION FUNCTIONAL. Because its exact form is not known, variations of approximations are applied.

The most simple one is the LOCAL DENSITY APPROXIMATION (LDA) [8], which, surprisingly, sometimes support meaningful results. The exchange-correlation functional in LDA assumes model of homogeneous electron gas, where the electron density is constant throughout all space. The electron density in the studied system is not constant, but for a small volume element around  $\vec{r}$  it has value which is obtain from the homogeneous gas model. The functional is only dependent on the electron density.

In the GENERALIZED GRADIENT APPROXIMATION (GGA) [9] the functional depends also on the first derivative of the electron density. Since the functional depends on quantities in the particular space, the GGA is also local approximation.

The aim of the recent development in the field is to find the functional, which can describe the exchange-correlation effects in an optimal way. Unfortunately, in many cases it leads to addition of another terms to LDA, unwillingly preserving its locality. Due to the local character of these approximations, they are not able to successfully describe dispersion interaction, which is typically of non-local character. The problem mentioned represents without doubts one of the most challenging problem in further application and development of DFT procedure.

There are different approaches how to take dispersion into account [10], [11]. Some of them use empirical parameters and in the following part we describe an approach introduced by *Jurečka et al.*

### 2.5.2 DFT-D

In DFT-D method [12] is the standard DFT augmented with an empirical term, which covers dispersion energy. “The dispersion energy is calculated as a sum of all possible pairwise atomic contributions”, write *Jurečka et al.* in the paper. The total DFT energy is therefore simply determined as a sum of DFT and dispersion energy. The dispersion energy is calculated as follows

$$E_{dis} = - \sum_{ij} f_{damp}(r_{ij}, R_{ij}^0) C_{6ij} r_{ij}^{-6}, \quad (33)$$

where  $f_{damp}(r_{ij}, R_{ij}^0)$  is a damping function, which depends on the inter-atomic distance  $r_{ij}$  and the equilibrium van der Waals separation  $R_{ij}^0$ , which is derived from the atomic van der Waals radii.  $C_{6ij}$  in the Eqn. 33 is the set of atomic dispersion coefficients. There are two parameters,  $R_{ij}^0$  and  $C_{6ij}$ , and the damping function, which were adopted from *Grimme* [13]. The damping function was optimized on the set of 22 training structures by comparing total DFT energies with highly accurate CCSD(T) data [14]. In the second step the validity of the parametrization was verified on the set of 58 testing structures [14], [15], [16] and [17].

As mentioned in the paper [12], the use of larger basis sets (triple zeta quality) and diffuse functions\* is recommended.

## 2.6 Basis sets

Because all the wavefunctions (as the eigenfunctions of the hermitian operator [18]) create the complete vector space of functions, we are able to express each of the wavefunctions as a infinite expansion of basis functions

$$\Psi = \sum_i^{\infty} c_i \psi_i, \quad (34)$$

where  $c_i$  are expansion coefficients and  $\psi_i$  are basis functions of the vector space, called orbitals. In the practical calculations we are not able to use infinite expansion of basis function, so we “cut” the expansion. Slater’s orbitals,

$$y(r) = KY(\theta, \varphi)e^{-kr}, \quad (35)$$

---

\*exponential functions with a small exponent (see later)

where  $K$  and  $k$  are constants,  $Y(\theta, \varphi)$  is a spherical function and  $r$  is the distance of the electron from the nucleus which is the orbital centered on, had been overcome by Gaussian orbitals,

$$y(r) = KY(\theta, \varphi)e^{-kr^2}, \quad (36)$$

which provide much higher computational efficiency [19]. The Slater orbitals can be approximately expressed as a linear combination of the Gaussian orbitals.

## 2.7 Stabilization energy, interaction energy and BSSE

When the supersystem is formed from subsystems, the released energy is called stabilization energy. It has opposite sign than interaction energy which is defined as a difference between the total energy of the supersystem and total energies of its subsystems.

$$\Delta E_{int} = E_{sys} - \sum_i^N e_i, \quad (37)$$

in the case of phenol dimer interaction energy is given as

$$\Delta E_{int} = E_{dimer} - 2 \cdot E_{monomer}. \quad (38)$$

The problem appears when the finite basis set is used. Supersystem and subsystems are described by non-equal basis sets. The resulting purely mathematical error we make, due to using the qualitatively different basis sets, is called BASIS SET SUPERPOSITION ERROR (BSSE). There are several approaches to obtain the BSSE-free stabilization energy.

By the COUNTERPOISE CORRECTION presented by Boys and Bernardi [20] both energies – of the subsystems and the supersystem – are determined by using the basis sets of the supersystem.

$$\Delta E_{AB}^{CP} = E_{AB}(AB) - E_A(AB) - E_B(AB) \quad (39)$$



## 3 Computation

### 3.1 HF, MP2 and DFT calculations

The structure of phenol dimer was optimized by Gaussian 03 program package [21] on the Hartree-Fock level employing cc-pVDZ basis set [22]. The optimized structure was used as a starting structure for the other methods. Geometry optimizations were performed without and with the Boys-Bernardi counterpoise correction [20] for two standard basis sets – cc-pVDZ and cc-pVTZ. Optimized cc-pVDZ structures represented the starting structures for the cc-pVTZ [22] optimizations. Within all calculations the Gaussian default convergence criteria had been invoked and all optimizations were running till the all four convergence criteria \* were *fulfilled*.

Stabilization energies were calculated considering cp-correction. The DFT calculations were performed with B3LYP functional [23] on the cc-pVDZ and cc-pVTZ level using Gaussian 03 program package, as well.

### 3.2 DFT-D calculations

To show that the dispersion is the leading and important force determining the phenol dimer structure, we more extensively applied the DFT-D method [12]. implemented by *Jurečka et. al.* into TurboMole program package [24].

Thanks to the method’s low computational demands larger basis sets were applied. Besides the standard TZVP basis set also the LP basis set was used. The LP basis set is equal to Pople’s 6-311++G(3df,3pd) basis set [25]. Further also the aQZ’ basis set, which is modified aug-cc-pVQZ [22], where the g-functions and the most diffuse functions had been removed from heavy atoms and hydrogens, was used. For all DFT-D calculations the TPSS functional [26] was applied. All the DFT-D optimizations utilized resolution of identity (RI) procedure [27], which notably decreases computational demands. The RI was in all cases performed with TZVP auxiliary basis set [28]. The DFT-D method using basis sets described is practically BSSE-free.

---

\*maximum force =  $4.5 \cdot 10^{-4}$ ; root-mean-square force =  $3.0 \cdot 10^{-4}$ ; maximum displacement =  $1.8 \cdot 10^{-3}$ ; root-mean-squar displacement =  $1.8 \cdot 10^{-3}$

### 3.3 Geometrical properties

The experiment [2] provides a wide range of geometrical properties of phenol dimer. The rotational constants well characterize the molecular structure (given mainly by the positions of nuclei), so we compare the experimental and calculated results and define the deviations as

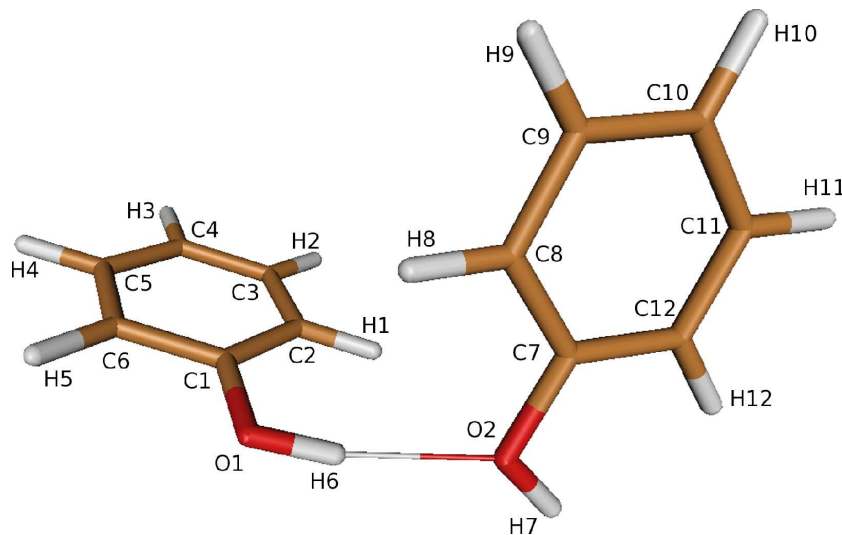
$$\delta X = \frac{X_{calc} - X_{exp}}{X_{exp}} \cdot 100, \quad (40)$$

where  $X_{calc}$  and  $X_{exp}$  is a rotational constant obtained by calculation and experiment, respectively. The total error  $\Delta$  was defined as

$$\Delta = \frac{1}{3} \sum_{X=A}^C |\delta X|. \quad (41)$$

Rotational constants were calculated by Gaussian 03 program package.

Figure 1: Phenol dimer structure, numbering of the atoms



Angle between the aromatic rings' planes (described by the corresponding dihedral angle  $\alpha$  ( $C_1 - O_1 - O_2 - C_7$ )), and the distance  $R$  between the centers of mass of each phenol molecule were evaluated. The errors were not calculated. The structure of phenol dimer is shown on Figure 1.

### 3.4 CCSD(T) geometry scan

To get some idea about the potential energy surface around the energy minimum obtained by optimization, we performed scan over dihedral angle  $\alpha$  in a range from  $47^\circ$  to  $77^\circ$  with the step  $3^\circ$  and two dimensional scan over the dihedral angle  $\alpha$  ( $47^\circ - 68^\circ$ ; step  $3^\circ$ ) and distance between two oxygen atoms ( $O_1 - O_2$ ) in a range from  $2.891 \text{ \AA}$  to  $3.381 \text{ \AA}$  with a step  $0.07 \text{ \AA}$ .

The scan have been performed as a set of single point CCSD(T)/cc-pVDZ energy calculations. For the comparison the MP2 energies were also determined. For the scan Molpro 02 program package [29] was used.

## 4 Results and discussion

### 4.1 Structure of phenol dimer

The monomers in the dimer are non-covalently bonded via H-bond and also by London dispersion energy between two aromatic rings. Experimental rotational constants, angle  $\alpha$  (angle between aromatic rings) and distance between centers of mass amount to 3.329, 0.792, and 0.686 GHz,  $63^\circ$  and 5.251 Å, respectively.

### 4.2 Phenol dimer optimization

#### 4.2.1 HF, MP2, B3LYP

Table 1 shows the geometrical characteristics of the phenol dimer, obtained by methods mentioned. Table 2 presents the rotational constants with deviations (cf. Section 3.3) and total errors.

Structure of the dimer optimized on HF/cc-pVDZ level is not correct since the monomers are far-away. The distance  $R$  is by more than 1 Å larger than the experimental value. The dihedral angle and the total error of the rotational constants are, at this level, unacceptable high. Passing from cc-pVDZ to cc-pVTZ basis set doesn't bring any improvement. The explanation lies evidently in the correlation energy, which is *not* covered in the HF theory.

Similar structures were obtained at the DFT/B3LYP level. As the dispersion is also not covered, the distance  $R$ , dihedral angle  $\alpha$  and the total error are still too high. Enlarging the basis set as well as considering counterpoise correction do not lead to any improvement.

Surprising results provides the MP2 optimization. As it is a widely used method for correlated ab initio calculations, we expected good agreement with the experiment. With the cc-pVDZ basis set the total error around 6 % is acceptable. When the counterpoise correction was applied, good phenol dimer geometry resulted (total error about 3 %). However, the dihedral angle  $\alpha$  shows that the aromatic rings are closer than in the experiment structure.

Passing from cc-pVDZ to cc-pVTZ basis set we expected better agreement with experimental geometry. Contrary to expectation, the difference

between both geometries were larger. The total error is the highest one and the dihedral angle became too small. We could say that the two phenols get during the optimization into some “stacked-like” structure which is well known in DNA [30].

The MP2 method is known to overestimate the stabilization energy [31], [32]. This is clearly reason for the wrong results obtained on triple-zeta level. Fortunately the BSSE correction improves the uncorrected results significantly, but they are still worse than the MP2/cc-pVDZ results. It must be mentioned, that the good results of MP2/cc-pVDZ optimization are the results of compensation of errors. The error of overestimated stabilization energy is, in this case favorably, compensated by smaller basis set.

#### 4.2.2 DFT-D optimization

The DFT/B3LYP method fails in the description of the phenol dimer, the DFT-D method provides, on the other hand, excellent results. The structural characteristics and the rotational constants are introduced in the Table 1 and Table 2. With all basis sets the total error is lower than at any previously mentioned level. Also the dihedral angle and intermolecular distance  $R$  is in a good agreement with the experiment. The best phenol dimer geometry is obtained employing the LP basis set (c. f. Section 3.2) with the total error about 1 %.

### 4.3 CCSD(T) scan results

The CCSD(T)/cc-pVDZ energy scan over the dihedral angle  $\alpha$  is shown on Figure 2.

The energy minimum of  $-4.789$  kcal/mol corresponds to the angle  $59^\circ$ . The MP2 energy minimum with the same basis set is  $-5.404$  kcal/mol and corresponds to the angle  $53^\circ$ . All energies are counterpoise corrected. The MP2 stabilization energy is much higher and shows that for proper description we need to use higher level of theory than MP2.

From the Figure 2 it is clear, that the potential (interaction) energy surface, with one variable (angle  $\alpha$ ) is very flat. Hence, the small change in energy can cause significant change in the geometry. The two dimensional scan

Figure 2: CCSD(T) scan. Interaction energy dependence on the dihedral angle  $\alpha$ .

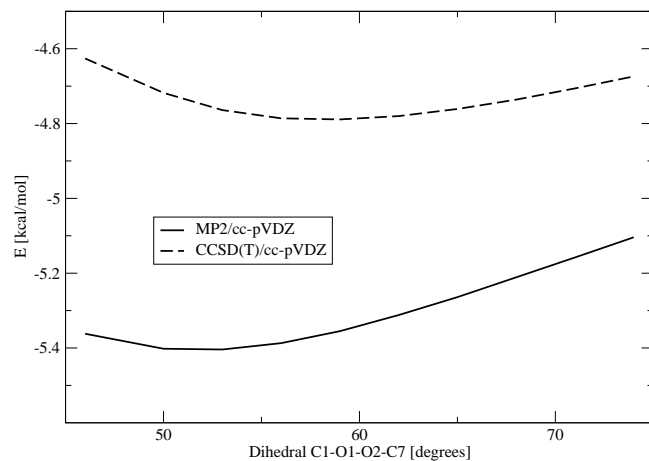
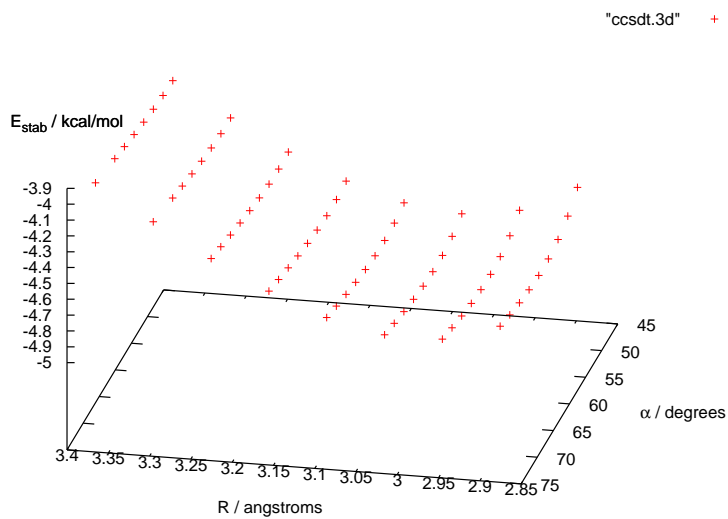


Figure 3: CCSD(T) scan over the dihedral angle  $\alpha$  (in degrees) and intersystem distance (in Å). The interaction energy dependence on the distance is apparently stronger.



over the dihedral angle and the intersystem distance (see Section 3.4), which is shown on Figure 3, confirms the flatness of the interaction energy surface. The dependence on the intermolecular distance is apparently stronger.

#### 4.4 Stabilization energy

The stabilization energies of optimized structures are mentioned in the Table 1. To compare structure dependent stabilization energies, the single point calculations on the MP2/cc-pVDZ, CCSD(T)/cc-pVDZ and DFT-D/LP levels were performed with the structure, which is in the best agreement with the experimental one (DFT-D/LP, cf. Table 2). CCSD(T)/cc-pVDZ, MP2/cc-pVDZ and DFT-D/LP are 5.036, 5.569 and 4.062 kcal/mol, respectively. Evidently, all values are similar and in the case of MP2 this is due to mentioned fact, that medium basis set was used.

## 5 Conclusions

In the thesis the structure of the phenol dimer was optimized at different theoretical levels and the following conclusions were made:

- The structures of the complex obtained at various ab initio levels are shown on Figure 4.
- Hartree-Fock and DFT methods fail in the description of dispersion effects. These methods are thus not suitable for determination of the phenol dimer structure.
- Second order Møller-Plesset perturbation theory provides correct geometry only if the medium-size basis set is applied. This is, however, due to compensation of errors. The MP2 method overestimates stabilization energy. Consequently, the optimization with large basis set provides wrong geometry of phenol dimer. The similar failure occurs within MP2 optimization also for other complexes (cf. Appendix 1), where dispersion is significant.
- Density functional theory augmented by empirical term describing dispersion energy seems to be well suited method for geometry optimization of the phenol dimer. With resolution of identity approximation it is fast, accurate and accessible method for structure determination of non-covalent complexes.



## 6 Data section

Table 1: Structural characteristics and stabilization energies. Dihedral angle  $\alpha$  is in degrees, distance  $R$  of centers of mass in Å, stabilization energy  $E_{stab}$  in kcal/mol

method	basis set	CP correction	$\alpha$	$R$	$E_{stab}$
<b>HF</b>	cc-pVDZ	no	109.2	5.990	–
		yes	98.7	6.239	3.734
	cc-pVTZ	no	108.1	6.174	–
		yes	–	–	–
<b>MP2</b>	cc-pVDZ	no	70.6	5.035	–
		yes	54.0	5.328	4.959
	cc-pVTZ	no	39.9	3.601	–
		yes	46.7	4.894	6.518
<b>DFT/B3LYP</b>	cc-pVDZ	no	112.1	5.856	–
		yes	114.7	6.242	4.133
	cc-pVTZ	no	118.9	6.146	–
		yes	–	–	–
<b>DFT-D/tpss</b>	TZVP	no	66.1	5.288	4.804
	LP	no	67.1	5.298	4.062
	aQZ'	no	60.9	5.201	3.360
<b>exp06 [2]</b>			63.1	5.251	–

Table 2: Rotational constants (in GHz) with deviations and the total errors

method	basis set	CP correction	A ( $\delta A$ )	B ( $\delta B$ )	C ( $\delta C$ )	$\Delta$
<b>HF</b>	cc-pVDZ	no	1.8334(+29.4%)	0.2422(-22.7%)	0.2374(-17.6%)	27.80%
		yes	1.9166(+35.3%)	0.2268(-27.7%)	0.2212(-26.3%)	29.77%
	cc-pVTZ	no	1.9246(+35.8%)	0.2291(-26.9%)	0.2272(-21.1%)	27.93%
		yes	–	–	–	–
<b>MP2</b>	cc-pVDZ	no	1.3321(+6.0%)	0.3307(+5.5%)	0.3118(+8.2%)	6.57%
		yes	1.3893(-2.0%)	0.3039(-3.1%)	0.2752(-4.3%)	3.13%
	cc-pVTZ	no	1.0007(-29.4%)	0.6231(+98.7%)	0.4905(+70.4%)	66.17%
		yes	1.2723(-10.2%)	0.3603(+14.8%)	0.3203(+11.1%)	12.03%
<b>B3LYP</b>	cc-pVDZ	no	1.7914(+26.4%)	0.2517(-19.7%)	0.2442(-15.2%)	20.43%
		yes	1.9838(+40.0%)	0.2230(-28.9%)	0.2217(-23.0%)	30.63%
	cc-pVTZ	no	1.9933(+40.7%)	0.2313(-26.3%)	0.2258(-21.6%)	29.53%
		yes	–	–	–	–
<b>DFT-D</b>	TZVP	no	1.4422(+1.8%)	0.2997(-4.4%)	0.2798(-2.9%)	3.03%
	LP	no	1.4262(+0.6%)	0.3055(-2.6%)	0.2865(-0.6%)	1.27%
	aQZ'	no	1.3997(-1.2%)	0.3185(+1.6%)	0.2925(+1.5%)	1.43%
<b>exp06 [2]</b>			1.41699	0.31351	0.28811	–





## List of abbreviations

- aQZ'** modified aug-cc-pVQZ basis set [22], where the g-functions and the most diffuse functions had been removed from the heavy atoms
- B3LYP** Becke hybrid functional introduced by Gaussian inc. [23]
- BOA** Born-Oppenheimer approximation
- BSSE** basis set superposition error
- cc-pVDZ** correlation corrected double zeta basis set [22]
- cc-pVTZ** correlation corrected triple zeta basis set [22]
- CCD** coupled cluster ab initio method employing only double excitations
- cp** counterpoise (correction) [20]
- CCSD** coupled cluster ab initio method employing single and double excitations
- CCSD(T)** coupled cluster ab initio method employing triple excitations by perturbation theory and single and double excitation in common way
- DFT** density functional theory [7]
- DFT-D** density functional theory augmented by empirical term describing dispersion energy[12]
- DNA** deoxyribonucleic acid
- exp06** experimental determination of phenol dimer structure made by *Schmitt et al.* [2]
- GGA** generalized gradient approximation [9]
- HF** Hartree-Fock ab initio method
- LDA** local density approximation [8]
- LP** modified Pople's 6-311++G(3df,3pd) basis set [25]

**MP $n$**  Møller-Plesset perturbation theory on the  $n$ -th order corrections level

**RI** resolution of identity [27]

**SCF** self-consistent field

**TPSS** Tao, Perdew, Staroverov and Scuseria meta-GGA functional [26]

**TZVP** triple zeta valence polarized basis set [28]

## References

- [1] P. Hobza, Ch. Riehn, A. Weichert, and B. Brutschy. Structure and binding energy of the phenol dimer: correlated ab initio calculations compared with results from rotational coherence spectroscopy. *Chemical Physics*, 283:331–339, Oct 2002.
- [2] M. Schmitt, M. Böm, Ch. Ratzer, D. Krügler, K. Kleinermanns, I. Kalkman, G. Berden, and L. Meerts, W. Determining the intermolecular structure in the s0 and s1 states of the phenol dimer by rotationally resolved electronic spectroscopy. *ChemPhysChem*, 7:1241–1249, 2006.
- [3] R. D. Hartree. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. *Proc. Camb. Philos. Soc.*, 24(89), 1927.
- [4] V. Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörper problems. *Z. Phys.*, 61(126), 1930.
- [5] Chr. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, 46(7):618–622, Oct 1934.
- [6] J. Cizek and J. Paldus. Coupled cluster approach. *Physica Scripta*, 21(3-4):251–254, 1980.
- [7] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, Nov 1964.
- [8] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, Nov 1965.
- [9] David C. Langreth and M. J. Mehl. Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B*, 28(4):1809–1834, Aug 1983.
- [10] Y. Zhao, N.E. Schultz, and D.G. Truhlar. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent in-

- teractions. *Journal of Chemical Theory and Computation*, 2(2):364–382, 2006.
- [11] T. Kubar, P. Jurecka, J. Cerny, J. Rezac, M. Otyepka, H. Valdes, and P. Hobza. Density-functional, density-functional tight-binding, and wave function calculations on biomolecular systems. *Journal of Physical Chemistry A*, 2007.
- [12] P. Jurecka, J. Cerny, P. Hobza, and D. R. Salahub. 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*, 28:555–569, 2007.
- [13] S. Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25:1463–1473, SEP 2004.
- [14] P. Jurecka, J. Spöner, J. Cerny, and P. Hobza. Benchmark database of accurate (mp2 and ccsd(t) complete basis set limit) interaction energies of small model complexes, dna base pairs, and amino acid pairs. *Physical Chemistry Chemical Physics*, 8:1985–1993, 2006.
- [15] P. Jurecka and P. Hobza. True stabilization energies for the optimal planar hydrogen-bonded and stacked structures of guanine center dot center dot center dot cytosine, adenine center dot center dot center dot thymine, and their 9-and 1-methyl derivatives: Complete basis set calculations at the mp2 and ccsd(t) levels and comparison with experiment. *Journal of the American Chemical Society*, 125:15608–15613, DEC 17 2003.
- [16] J. Spöner, P. Jurecka, and P. Hobza. Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *Journal of the American Chemical Society*, 126:10142–10151, AUG 18 2004.
- [17] P. Jurecka, J. Spöner, and P. Hobza. Potential energy surface of the cytosine dimer: Mp2 complete basis set limit interaction energies, ccsd(t)



- correction term, and comparison with the amber force field. *Journal of Physical Chemistry B*, 108:5466–5471, APR 29 2004.
- [18] J. Fišer. *Úvod do kvantové chemie*. Academia, 1983.
- [19] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Dover, 1996.
- [20] F. Boys, S. and F. Bernardi. Calculation of small molecular interactions by differences of separate total energies - some procedure with reduces errors. *Molecular Physics*, 19:553, 1970.
- [21] M. J. Frisch, G. W. Trucks, 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 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
- [22] Jr. Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989.
- [23] Becke3lyp method references and general citation guidelines. *Gaussian NEWS*, 5:2, summer 1994.
- [24] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel. *Chem. Phys. Lett.*, 162:165, 1989.

- [25] M. J. Frisch, J. A. Pople, and J. S. Binkley. Self-consistent molecular orbital methods 25. supplementary functions for gaussian basis sets. *The Journal of Chemical Physics*, 80(7):3265–3269, 1984.
- [26] John P. Perdew, Jianmin Tao, Viktor N. Staroverov, and Gustavo E. Scuseria. Meta-generalized gradient approximation: Explanation of a realistic nonempirical density functional. *The Journal of Chemical Physics*, 120(15):6898–6911, 2004.
- [27] M. Feyereisen, G. Fitzgerald, and A. Komornicki. *Chem. Phys. Lett.*, 208:359, 1993.
- [28] R Ahlrichs and R. Taylor, P. The choice of gaussian basis sets for molecular electronic structure calculations. *J. Chim. Physique*, 78:315, 1981.
- [29] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, and T. Thorsteinsson. Molpro, version 2002.6, a package of ab initio programs, 2002. see <http://www.molpro.net>.
- [30] J. Watson and F. Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737, 1953.
- [31] M. Poblet, J. M. and Benard. Cuprophilicity, a still elusive concept: a theoretical analysis of the ligand-unsupported cu-i-cu-i interaction in two recently reported complexes. *Chemical Communications*, pages 1179–1180, JUN 7 1998.
- [32] Seiji Tsuzuki, Tadafumi Uchimarui, Ko ichi Sugawara, and Masuhiro Mikami. Energy profile of the interconversion path between t-shape and slipped-parallel benzene dimers. *The Journal of Chemical Physics*, 117(24):11216–11221, 2002.

## Appendix

M. Kolář, P. Hobza, “Accurate Theoretical Determination of the Structure of Aromatic Complexes Is Complicated: The Phenol Dimer and Phenol...Methanol Case”, *Journal of Physical Chemistry A*, 2007, Vol. 111, 0000