

CHARLES UNIVERSITY IN PRAGUE
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICAL AND MACROMOLECULAR CHEMISTRY

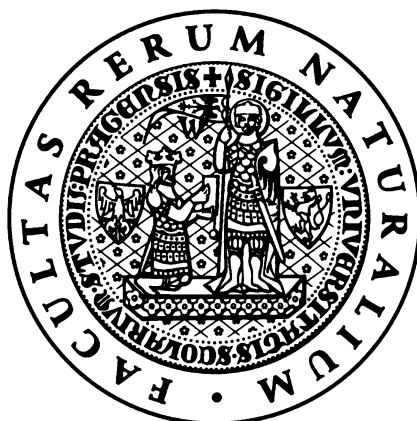
BACHELOR DIPLOMA THESIS

Michal KOLÁŘ

2007



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

1	Introduction	5
1.1	Computational chemistry	5
1.2	Motivation	5
2	Methods	7
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
3	Computation	17
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
4	Results and discussion	20
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
5	Conclusions	24
6	Data section	25

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 ψ , 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, π is pi – Ludolph's number, ϵ is permittivity, Δ_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 than 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 Φ_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 ϕ_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 λ 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”* form

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

where Φ_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

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

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

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

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

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

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

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

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

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

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

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

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 ρ 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 ψ is a wavefunction which depends on spin ζ_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

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

where \hat{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 ψ_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

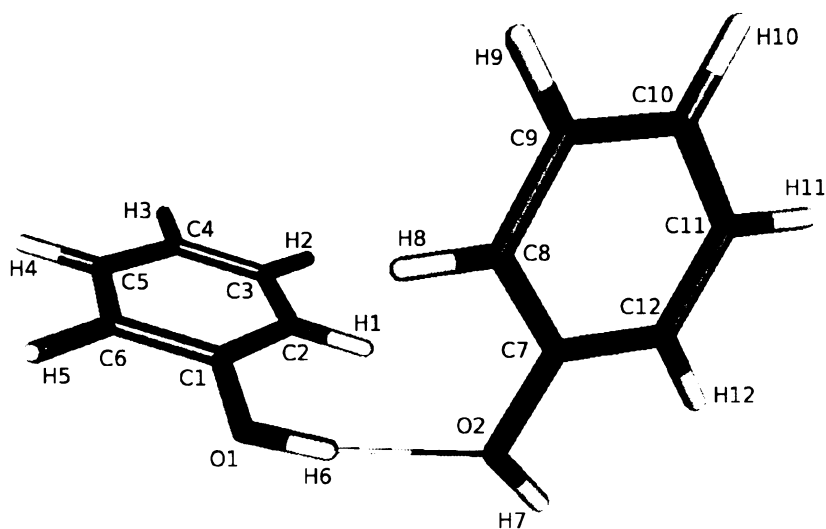
$$\delta X = \frac{X_{calc} - X_{exp}}{X_{exp}} \cdot 100, \quad (40)$$

where X_{calc} and X_{exp} is a rotational constant obtained by calculation and experiment, respectively. The total error Δ 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 α ($C_1 - O_1 - O_2 - C_7$)), 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 α in a range from 47° to 77° with the step 3° and two dimensional scan over the dihedral angle α ($47^\circ - 68^\circ$; step 3°) and distance between two oxygen atoms ($O_1 - O_2$) in a range from 2.891 \AA to 3.381 \AA with a step 0.07 \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 α (angle between aromatic rings) and distance between centers of mass amount to 3.329, 0.792, and 0.686 GHz, 63° 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 α 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 α 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 α is shown on Figure 2.

The energy minimum of -4.789 kcal/mol corresponds to the angle 59° . The MP2 energy minimum with the same basis set is -5.404 kcal/mol and corresponds to the angle 53° . 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 α) 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 α .

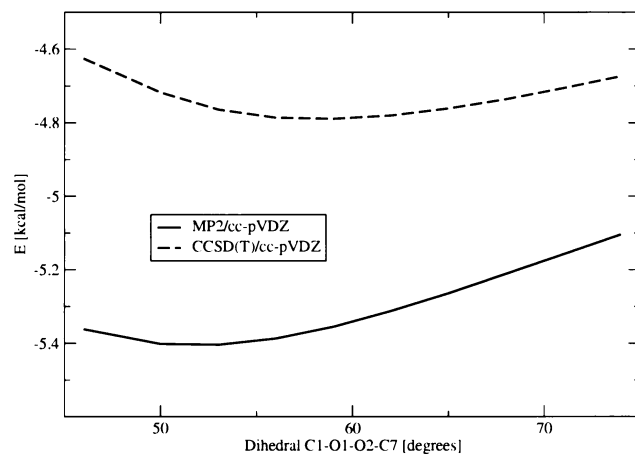
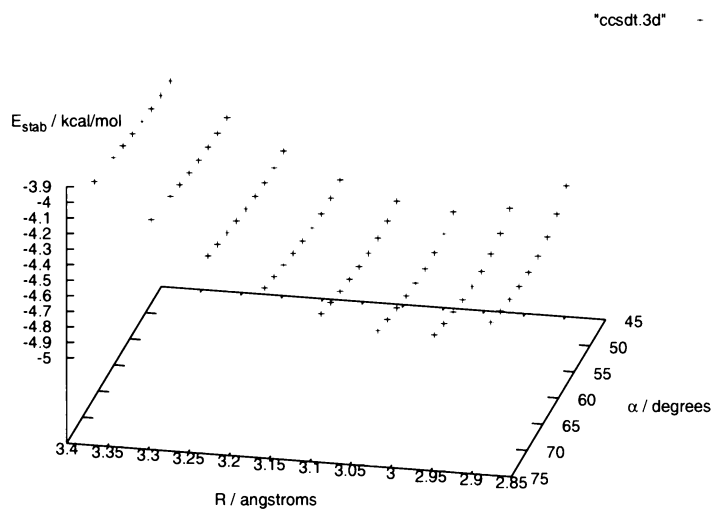


Figure 3: CCSD(T) scan over the dihedral angle α (in degrees) and intersystem distance (in Å). The interaction energy dependence on the distance is apparently stronger.



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

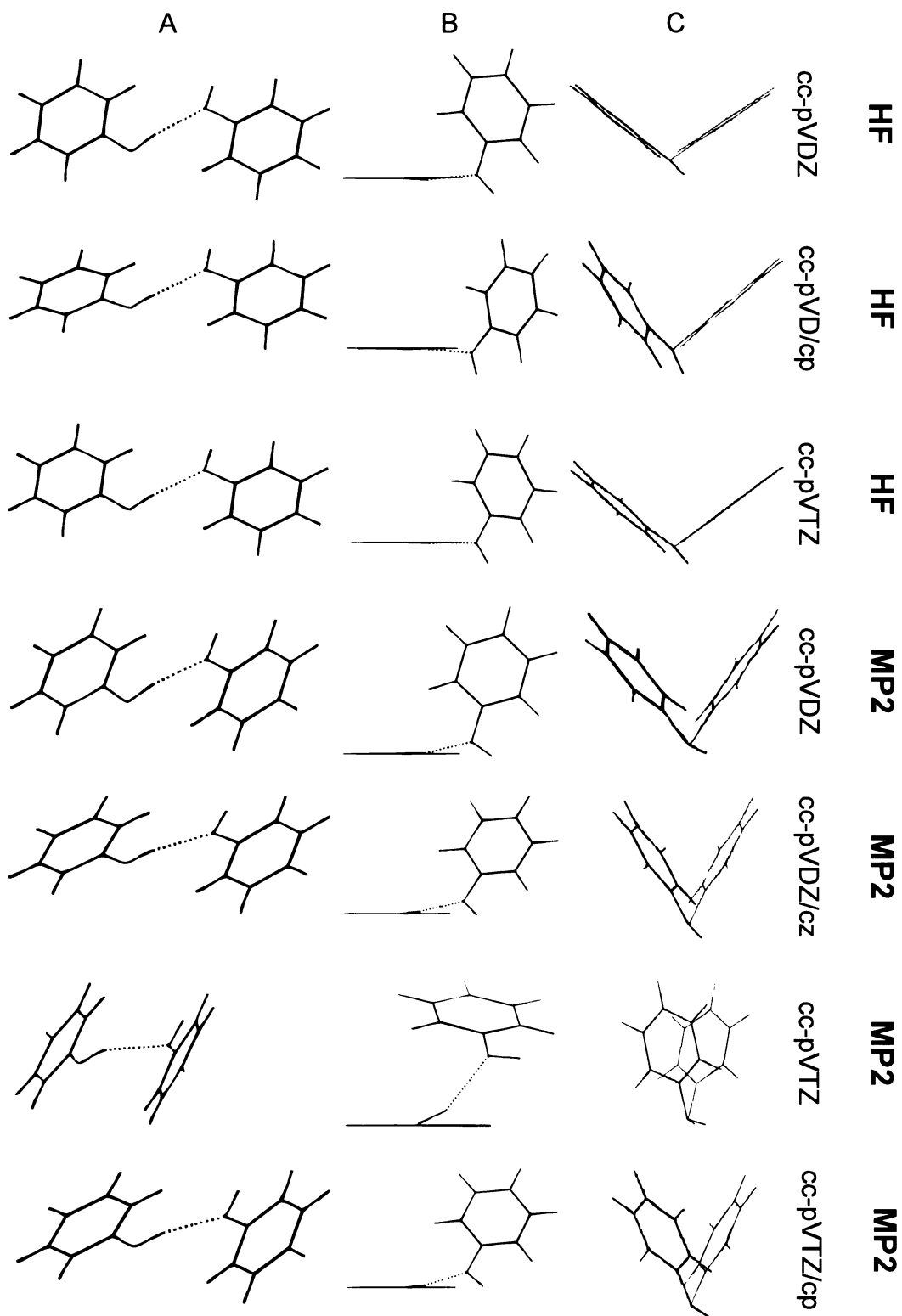
Table 1: Structural characteristics and stabilization energies. Dihedral angle α is in degrees, distance R of centers of mass in Å, stabilization energy E_{stab} in kcal/mol

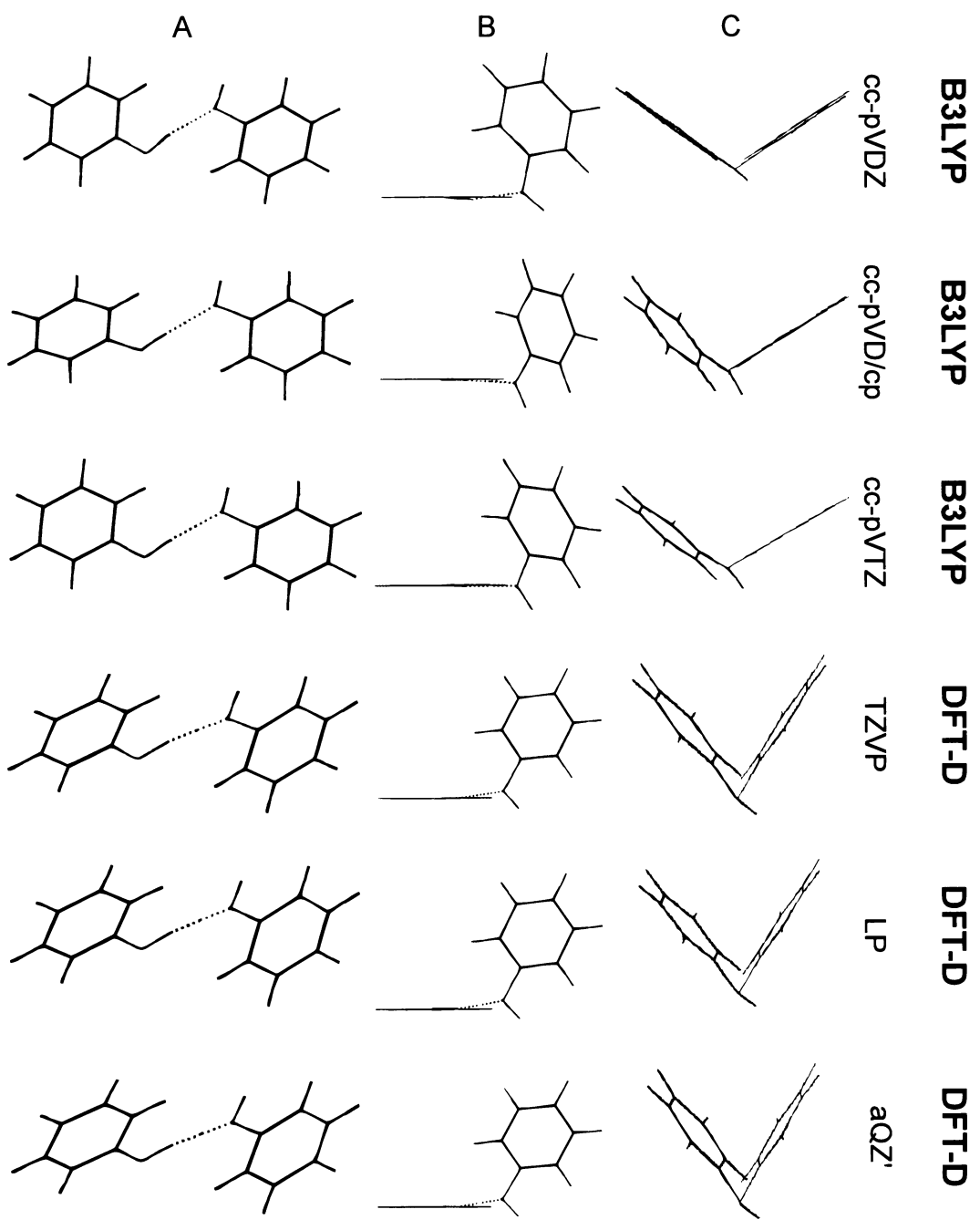
method	basis set	CP correction	α	R	E_{stab}
HF	cc-pVDZ	no	109.2	5.990	–
		yes	98.7	6.239	3.734
	cc-pVTZ	no	108.1	6.174	–
		yes	–	–	–
MP2	cc-pVDZ	no	70.6	5.035	–
		yes	54.0	5.328	4.959
	cc-pVTZ	no	39.9	3.601	–
		yes	46.7	4.894	6.518
DFT/B3LYP	cc-pVDZ	no	112.1	5.856	–
		yes	114.7	6.242	4.133
	cc-pVTZ	no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

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

method	basis set	CP correction	A (δA)	B (δB)	C (δC)	Δ
HF	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	-	-	-	-
MP2	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%
B3LYP	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	-	-	-	-
DFT-D	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%
exp06 [2]			1.41699	0.31351	0.28811	-

Figure 4:





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. Ratzner, 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. Sponer, 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. Sponer, 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. Sponer, 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 Uchimaru, 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

Accurate Theoretical Determination of the Structure of Aromatic Complexes Is Complicated: The Phenol Dimer and Phenol···Methanol Cases

Michal Kolář and Pavel Hobza*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Molecular Systems, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

Received: February 22, 2007; In Final Form: April 16, 2007

The structure of the phenol dimer and phenol···methanol complexes was determined by gradient optimization using the Hartree–Fock (HF), MP2, DFT, and RI-DFT-D methods with various basis sets. Theoretical rotational constants were compared with experimental values and the following conclusions were made: (1) HF and DFT methods fail to predict cluster geometries; (2) MP2 with a medium basis set yields reliable cluster geometries but only because of a compensation for errors; (3) when the AO basis set is enlarged, the geometry becomes incorrect, and the theoretical geometry becomes reliable only when the higher correlation energy contributions (CCSD(T)) are included; and (4) the RI-DFT-D procedure covering the dispersion energy provides excellent geometries.

Introduction

Recent years have brought significant progress to the calculation of interaction energies of extended complexes containing aromatic subsystems. It is now well-known that MP2 stabilization energy when determined with extended basis sets (or even at the complete basis set (CBS) limit) is overestimated, which is especially true for stacking structures.^{1,2} To reduce the stabilization energy of these structures, it is necessary to include the CCSD(T) term. The CCSD(T) correction term is definitely not negligible, and in our database,³ consisting of more than 160 complexes, it reaches up to 4 kcal/mol (~25% of interaction energy). It is necessary to include this correction term not only in the case of stacked structures of aromatic hydrocarbons and DNA bases but also in that of complexes of amino acids.⁴ An important question thus arises, namely, how to determine the structure of these systems. When using the MP2/extended basis set treatment (which has nowadays become feasible even for extended complexes), overestimated stabilization energies are expected to yield incorrect geometries. The only straightforward way to prove it would be to compare the structures determined at various theoretical levels with experimental data. The problem lies in the fact that experimental gas-phase structures of extended (aromatic) complexes are extremely rare. Phenol···methanol⁵ and phenol···phenol^{6,7} complexes would provide unique experimental data.

The latter complex is highly interesting as the structure of the dimer is characterized by two different types of noncovalent interactions: the O–H···O hydrogen bonding and the stacking of two aromatic rings. The hydrogen bonding is mainly of an electrostatic origin while the aromatic stacking is mainly due to London dispersion interactions, both of which participate in determining the equilibrium structure.

In our previous study,⁸ we calculated the structure of the phenol dimer using the MP2 method combined with 6-31G-(d,f) and TZVPP basis sets, and the latter calculations agreed relatively well with the experimental rotational constants (with the average error being about 1.5%). A recently published paper⁷ on the phenol dimer which contained in addition to the ground-state geometries also the geometries in the first excited-state

prompted us to reinvestigate the structure of the dimer. We used a wider set of computational techniques (HF, MP2, DFT, CC2), and unlike in our previous study, we systematically used the Gaussian 03⁹ code, which uses stricter convergence criteria (maximum force = 4.5×10^{-4} ; root-mean-square force = 3.0×10^{-4} ; maximum displacement = 1.8×10^{-3} ; root-mean-square displacement = 1.2×10^{-3} which approximately corresponds to an energy change of about 10^{-8}) than the previously used Turbomole¹⁰ code (energy change 10^{-6} , gradient change 10^{-4}). In the case of a flat potential energy surface, tighter convergence criteria can yield a different structure.

The presented study shows that geometries determined with the MP2 procedure employing an extended AO basis set can be less reliable than those calculated with the same method but using only medium AO basis sets. An improvement in the former procedure can only be obtained by considering the CCSD(T) corrections. Two aromatic complexes will be investigated, the phenol dimer and the phenol···methanol complex.

Computations

Methods. Geometry and energy characteristics of all the systems investigated were determined by the Hartree–Fock, MP2, DFT/B3LYP, and CC2 procedures using the cc-pVDZ and cc-pVTZ AO basis sets and TZVP auxiliary basis set for all RI methods.¹¹ Geometry optimizations were performed using the standard procedure as well as the counterpoise-corrected gradient optimization.

Besides the widely used standard methods mentioned, we also used the recently introduced RI-DFT-D method,¹² which covers the London dispersion energy. Because of the parametrization of the method to known, accurate CCSD(T) energies (via the damping function associated with dispersion energy), the method yields accurate interaction energies comparable to the CCSD(T) values. In this study, we used the TPSS functional¹⁴ and three different AO basis sets: (i) the TZVP basis set, (ii) the Pople 6-311++G(3df,3pd), for which we used the abbreviation LP, and (iii) the modified aug-cc-pVQZ basis set, where both g-functions and the most diffuse f-functions were removed from

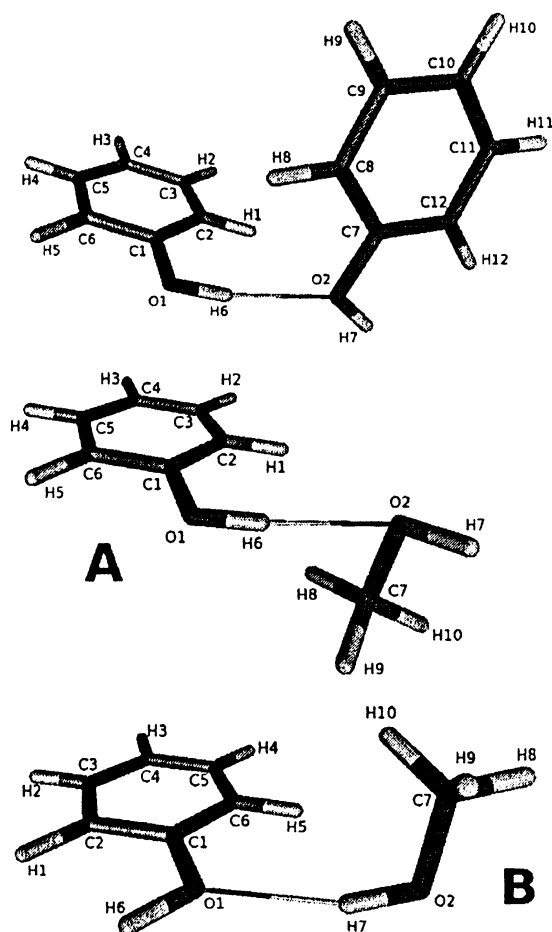


Figure 1. Optimized structures and numbering of phenol dimer and phenol...methanol complexes; the latter complex possesses two structures: A (phenolic hydrogen participates in H-bonding) and B (methanolic hydrogen in H-bonding).

TABLE 1: Experimental and Theoretical (HF, MP2, B3LYP) Characteristics of the Phenol Dimer^a

	exptl ^b	basis set	HF	MP2	B3LYP
α	63	cc-pVDZ	99/109	54/71	115/112
		cc-pVTZ	<i>c</i> /108	47/40	<i>c</i> /119
<i>R</i>	5.251	cc-pVDZ	6.239/5.990	5.328/5.035	6.242/5.856
		cc-pVTZ	<i>c</i> /6.174	4.894/3.601	<i>c</i> /6.146
<i>E</i>		cc-pVDZ	-3.734	-4.959	-4.133
		cc-pVTZ	<i>c</i>	-6.518	<i>c</i>
<i>A</i>	1.41699	cc-pVDZ	1.9166/1.8334	1.3893/1.3321	1.9838/1.7914
		cc-pVTZ	<i>c</i> /1.9246	1.2723/1.0007	<i>c</i> /1.9933
<i>B</i>	0.31351	cc-pVDZ	0.2268/0.2422	0.3039/0.3307	0.2230/0.2517
		cc-pVTZ	<i>c</i> /0.2291	0.3603/0.6230	<i>c</i> /0.2313
<i>C</i>	0.28811	cc-pVDZ	0.2212/0.2374	0.2752/0.3118	0.2217/0.2442
		cc-pVTZ	<i>c</i> /0.2272	0.3203/0.4905	<i>c</i> /0.2258
FE		cc-pVDZ	29.8/27.8	3.1/6.6	30.6/20.4
		cc-pVTZ	<i>c</i> /27.9	12.0/66.2	<i>c</i> /29.5

^aTheoretical calculations were performed using the cc-pVDZ and cc-pVTZ basis sets. The values in the nominator and denominator correspond to counterpoise-corrected and counterpoise-uncorrected characteristics. The dihedral angle $\alpha(\text{C1}-\text{O1}-\text{O2}-\text{C7})$ is given in deg, the distance *R* of the centers of mass is given in Å, interaction energies are given in kcal/mol, rotational constants are given in GHz, and error (FE) is given in %. ^bReference 7. ^cNot calculated

various theoretical procedures in conjunction with the cc-pVDZ and cc-pVTZ basis sets, of which the rotational constants will be investigated first. The final error, FE, is prohibitively large for the HF and DFT calculations, and no improvement is obtained when passing from cc-pVDZ to cc-pVTZ basis set. The same is true for passing from standard to counterpoise-corrected gradient optimization. The explanation is straightforward: neither method covers the London dispersion energy, which is responsible for attraction between the aromatic rings. Consequently, the distance *R* between the aromatic rings is overestimated by almost 1 Å (cf. Table 1), and the intersystem dihedral angle α is too large (by more than 40°).

The correlated MP2 method covers the London dispersion energy, hence, closer agreement with the experiment is expected. When investigating the cc-pVDZ results, we found that the FE error was considerably smaller. This error was further reduced when the counterpoise-corrected constants were taken into consideration, and these values basically agreed with our previous results (ref 8). Good agreement with experimentally proved values was also demonstrated when the angle α was considered where the standard and counterpoise-corrected cc-pVDZ values bracketed the experimental value. Similarly, the uncorrected and counterpoise-corrected distances *R* also bracketed the experimental value, but the error was still not negligible (about 0.1 Å). We had expected both errors to be reduced significantly when passing to the considerably larger cc-pVTZ basis set. To our surprise, the opposite was true, and the FE error dramatically increased: in the case of the standard optimization to 66% and in the case of the counterpoise-corrected one to a still considerably large value of 12%. (The situation is, however, more complicated, and it is a priori not clear that the counterpoise-corrected gradient optimization yields geometry which is closer to accurate geometry.) The intersystem angle α told us that the dihedral angle was too small, which was confirmed by the significantly underestimated distance *R*. The incorrect results from the MP2/cc-pVTZ calculations can only be explained by the overestimated stabilization energy, which was confirmed by the fact that the FE error was smaller when using the counterpoise-corrected optimization. In this case, the geometry was optimized on the basis of the BSSE-corrected total energy while in the other case it was optimized on the basis of uncorrected total energy. It was mentioned above that the overestimation of the MP2 stabilization energy was removed

heavy atoms and analogical modifications were made for hydrogens (the abbreviation aQZ'). The RI-DFT-D energies are practically basis set superposition error (BSSE)-free thanks to both the large basis set and the method employed. This means that with this technique, it is not necessary to perform the time-consuming counterpoise-corrected gradient optimization.

Interaction Energies. Interaction energy was systematically determined, including the basis set superposition error (BSSE). In the case of counterpoise-corrected optimization, the final interaction energies covered the deformation energies.

Rotational Constants, Intersystem Angle, and Intersystem Distance. Rotational constants were calculated from the optimized geometries and compared with the experimental values. The final error (FE; in %) was defined as an arithmetic mean of errors of three rotational constants (the difference between the theoretical and experimental values divided by the experimental one) multiplied by 100. The meaning of the intersystem dihedral angle $\alpha(\text{C1}-\text{O1}-\text{O2}-\text{C7})$ is apparent from Figure 1. The intersystem distance *R* is the distance between the centers of the mass of the two subsystems.

Throughout the study we used Gaussian 03⁹, Turbomole,¹⁰ and Molpro¹³ suites.

Results and Discussion

Phenol Dimer. Table 1 shows the experimental and theoretical characteristics of the dimer in question determined utilizing

Determination of Aromatic Complex Structures

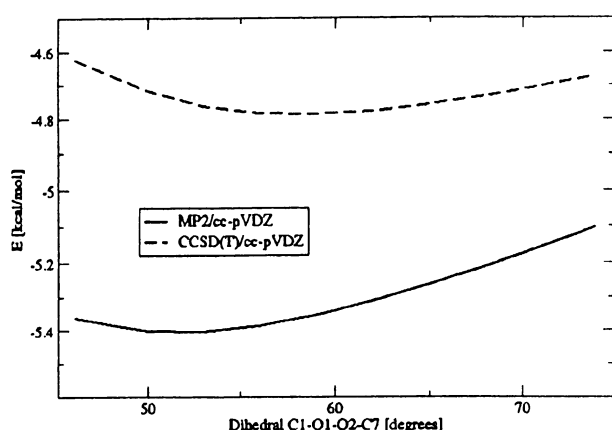


Figure 2. Dependence of CCSD(T) and MP2 interaction energies of phenol dimer on the dihedral angle.

159 by considering the higher correlation energy terms. Passing from
160 the MP2 to the CC2 procedure did not provide any improvement
161 (not shown), which leads us to the conclusion that an explicit
162 consideration of the CCSD(T) term is inevitable.

163 The CCSD(T) calculations are even with medium basis sets
164 time-consuming, and for a cluster of the present size gradient
165 optimization is clearly impractical. The only possible way of
166 estimating the role of higher correlation contributions is thus
167 to perform a series of single-point calculations. Figure 2 shows
168 a one-dimensional scan of the dihedral angle α (C1–O1–O2–
169 C7) in dependence on the MP2 and CCSD(T) uncorrected
170 interaction energies (cc-pVDZ basis set). Evidently, both energy
171 curves are quite flat, and the rather large change in the dihedral
172 angle is connected with only a small energy change. Figure 2
173 shows that the CCSD(T) energy curve is shifted to lower
174 stabilization energies and the higher dihedral angle α , which
175 better agrees with the experimental value (in comparison with
176 the MP2 values). The minima found at the CCSD(T) and MP2
177 energy curves correspond to 59° and –4.789 kcal/mol, and 53°
178 and –5.404 kcal/mol, respectively. These results clearly indicate
179 the need to determine the geometry of the phenol dimer (and
180 also of any complexes with aromatic electrons) at higher levels
181 than the MP2 level. It is true that when using a medium basis
182 set, the MP2 data agree reasonably well with the experimental
183 values, but this is evidently caused by the compensation of error
184 (the increase in MP2 stabilization energy when the extended
185 basis set is used and its decrease when a method beyond MP2
186 is applied). The problem remains that this error compensation
187 cannot be relied upon to provide valid results.

188 Among the various techniques enabling the achievement of
189 CCSD(T) performance at a much lower cost, the modified DFT
190 methods covering the London dispersion energy are in first
191 place. These computationally very favorable methods^{12,15,16}
192 make it possible to perform the gradient optimization even for
193 extended complexes. Table 2 summarizes the RI-DFT-D/TPSS
194 results for the phenol dimer when the TZVP, aQZ', and LP
195 basis sets were used. The calculated FE errors were small, and
196 when extended LP and aQZ' basis sets were used, the error was
197 close to 1%. The intersystem dihedral angle α determined using
198 all three basis sets agreed well with the experiment, which was
199 also true about the intersystem distance R . It should be
200 mentioned here that the performance of the rather small TZVP
201 basis set was very good, which is especially promising for
202 calculations of extended complexes.

203 **Phenol···Methanol Complex.** Table 3 summarizes both the
204 experimental and the theoretical characteristics of the complex

TABLE 2: Experimental and Theoretical (RI-DFT-D) Characteristics of the Phenol Dimer^a

	exptl ^b	RI-DFT-D/TPSS		
		basis set TZVP	basis set LP	basis set aQZ'
α	63	66	67	61
R	5.251	5.288	5.298	5.201
ΔE		–4.804	–4.062	–3.360
A	1.41699	1.4422	1.4262	1.3997
B	0.31351	0.2997	0.3055	0.3185
C	0.28811	0.2798	0.2865	0.2925
FE		3.0	1.3	1.4

^a The dihedral angle α (C1–O1–O2–C7) is given in degrees, the distance R of the centers of mass are in Å, interaction energies are in kcal/mol, rotational constants are in GHz, and error (FE) is in %; theoretical calculations were performed using the TZVP, LP, and aQZ' basis sets. ^b Reference 7.

TABLE 3: Experimental and Theoretical (MP2, RI-DFT-D) Characteristics of the Phenol···Methanol Complex^a

	exp ^b	MP2	
		basis set cc-pVDZ	basis set cc-pVTZ
ΔE		4.959	–6.581
A	3290.8	3233.7/3172.8	3226.4/2971.9
B	792.2	841.1/832.7	849.1/867.0
C	685.6	726.2/706.3	734.2/742.7
FE		4.6/3.9	5.4/9.2

	exptl ^b	RI-DFT-D/TPSS		
		basis set TZVP	basis set LP	basis set aQZ'
ΔE		–7.484/–3.995	–6.366/–3.156	–5.774/–2.560
A	3290.8	3601.8/3362.7	3332.6/3384.0	3312.6/3311.7
B	792.2	802.4/751.9	793.0/751.3	795.9/779.7
C	685.6	690.6/648.2	680.3/645.4	683.6/665.3
FE		3.8/4.2	0.7/4.6	0.5/1.7

^a The interaction energies are in kcal/mol, rotational constants are in MHz, and error (FE) is in %. Theoretical calculations were performed using the cc-pVDZ and cc-pVTZ and TZVP, LP, and aQZ' basis sets, respectively. Values in the nominator and denominator refer to structures A and B. ^b Reference 5.

205 in question. Unlike the previous complex, which possesses only
206 one structure, this one has two possible structures where either
207 phenolic or methanolic hydrogen can participate in H-bonding.
208 On the basis of different acidities of the two hydrogens, it is
209 expected that the former will be more probable. Stabilization
210 energies for this structure are systematically larger than stabi-
211 lization energies for the other structure, but the latter ones are
212 definitively not negligible; the CH₃OH···phenol structure will
213 thus coexist with the C₆H₅OH···methanol one (population of
214 both structures at room temperature will be about 1:200). When
215 investigating the FE values, we found again smaller values for
216 the C₆H₅OH···methanol structure, and the best agreement with
217 the experiment was achieved using the LP and aQZ' basis sets;
218 the TZVP values were slightly larger. It must be, however,
219 concluded that the assignment of theoretical constants to the
220 experimental ones is not unambiguous. The RI-DFT-D/aQZ'
221 values for structures A and B agree fairly well with the
222 experimental numbers, which prevents a clear decision from
223 being made.

224 Like in the case of the phenol dimer, passing to a larger basis
225 set is in the case of MP2 calculations not connected with
226 improving the accuracy of the FE error estimate but with its
227 worsening, which indicates the need to include the higher
228 correlation energy terms. The RI-DFT-D values, which ef-
229 fectively cover these terms, are systematically smaller than MP2/

230 cc-pVTZ ones. Further, the FE error for the C₆H₅OH...methanol
231 structure is systematically smaller than this value for the other
232 structure.

233 **Effect of Basis Set on the Geometry.** Accurate geometry
234 (and also stabilization energy) is obtained when using a highly
235 correlated method with extended basis set, and the CCSD(T)/
236 CBS level represents one possibility. For gradient geometry
237 optimization, this level is impractical, and some compromise
238 should be selected. The use of basis sets with diffuse functions
239 (like aug-cc-pVDZ) in combination with MP2 method is
240 generally not recommended because the respective stabilization
241 energy can be overestimated. Hence, an extended basis set
242 without diffuse functions is to be used. In this case we rely on
243 compensation of errors, and the question remains which basis
244 set should be used. In other words we try to find a level where
245 the compensation of errors is the most efficient. As shown in
246 this paper the MP2/cc-pVDZ level provides better geometries
247 than the MP2/cc-pVTZ level. We are certainly aware of the
248 fact that it is impossible to rely on the compensation of errors
249 mentioned and a computational procedure yielding accurate
250 geometries as well as stabilization energies should be selected.
251 If a method provides accurate stabilization energies, than it is
252 possible to expect that accurate geometries will also result. The
253 RI-DFT-D procedure represents such a level, and it was shown
254 repeatedly that RI-DFT-D stabilization energies agree very well
255 with the benchmark data. In this paper we have shown that the
256 method also yields excellent geometries. Contrary to wave
257 function theories (WFT), the DFT methods are less sensitive
258 to the presence of diffuse functions, and in the present paper
259 we have shown that the TZVP basis set yields interaction
260 energies and geometries similar to those of LP and aQZ' basis
261 sets containing diffuse functions. Let us close this paragraph
262 by stating that the RI-DFT-D method provides excellent
263 geometries also for isolated extended systems with folded
264 structures (e.g., helicenes) where standard WFT calculations
265 fail.¹⁷

266 Conclusions

267 (i) MP2 calculations with medium AO basis sets yield reliable
268 structures. Enlarging the AO basis set provides overestimated
269 MP2 stabilization energies and inaccurate MP2 geometries. The
270 geometrical characteristics are improved by the inclusion of the
271 higher correlation energy contributions. The reliable results from
272 the MP2/medium basis set calculations are thus a result of the
273 compensation of errors.

274 (ii) The CCSD(T) calculations are impractical for the
275 geometry optimization of the complexes of the present size. The
276 RI-DFT-D/TPSS calculations, which are also feasible in the case
277 of extended complexes, yield reliable geometrical constants even
278 with a rather small TZVP basis set. When larger LP and aQZ'
279 basis sets are adopted, excellent geometrical characteristics
280 results are achieved.

281 (iii) Conclusions made for the phenol dimer are also valid
282 for the phenol...methanol complex, and also here the best
283 agreement with the experimental rotational constants is obtained

284 employing the RI-DFT-D/TPSS/aQZ' procedure. The best
285 structure geometries are attached in Supporting Information.

286 **Acknowledgment.** This work is part of the Research Project
287 No. Z4055905 and was supported by Grants EC512 (from the
288 Ministry of Education of the Czech Republic) and A400550510
289 (from the Grant Agency of the Academy of Sciences of the
290 Czech Republic).

291 **Supporting Information Available:** Phenol...phenol DFT-
292 D/TPSS/LP, phenol...methanol A DFT-D/TPSS/aQZ', and
293 phenol...methanol B DFT-D/TPSS/aQZ' structure geometries.
294 This material is available free of charge via the Internet at [http://](http://pubs.acs.org)
295 pubs.acs.org.

296 References and Notes

- 297 (1) Sinnokrot, M. O.; Sherill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.
- 298 (2) Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608.
- 299 (3) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem.*
300 *Phys.* **2006**, *8*, 1985.
- 301 (4) Vondrášek, J.; Bendová, L.; Klusák, V.; Hobza, P. *J. Am. Chem.*
302 *Soc.* **2005**, *127*, 2615.
- 303 (5) Schmitt, M.; Kuepper, J.; Spangenberg, D.; Westpahl, A. *Chem.*
304 **2000**, *254*, 349.
- 305 (6) Weichert, A.; Riehn, C.; Weichert, A.; Brutschy, B. *J. Phys. Chem.*
306 *A* **2001**, *105*, 5679.
- 307 (7) Schmitt, M.; Boehm, M.; Ratzler, C.; Kruegler, D.; Kleinermanns,
308 K.; Kalman, I.; Berden, G.; Meerts, W. L. *ChemPhysChem* **2006**, *7*, 1241.
- 309 (8) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys.* **2002**,
310 *283*, 331.
- 311 (9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,
312 M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.
313 N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.;
314 Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.;
315 Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.;
316 Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li,
317 X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.;
318 Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.;
319 Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.;
320 Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich,
321 S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A.
322 D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A.
323 G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.;
324 Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham,
325 M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.;
326 Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian*
327 *03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
- 328 (10) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys.*
329 *Lett.* **1989**, *162*, 165.
- 330 (11) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.*
331 **1993**, *208*.
- 332 (12) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.*
333 **2007**, *28*, 555.
- 334 (13) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz,
335 M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardtsson, A.;
336 Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.;
337 Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura,
338 M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone,
339 A. J.; Tarroni, R.; Thorsteinsson, T. *MOLPRO, a package of ab initio*
340 *programs*, version 2006.1; see <http://www.molpro.net>.
- 341 (14) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys.*
342 *Rev. Lett.* **2003**, *91*, 146.
- 343 (15) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463; **2006**, *27*, 1787.
- 344 (16) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J.*
345 *Chem.* **2001**, *114*, 5149.
- 346 (17) Ruliššek, L.; Valdés, H.; Klusák, V.; Exner, O.; Starý, I.; Hobza, P.
347 *J. Chem. Phys.*, submitted.