Lecture Notes

Three Lectures in Computational Chemistry

Spanget-Larsen, Jens

DOI:
10.13140/RG.2.1.1157.1282

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Spanget-Larsen, J. (2016, Jan 6). Lecture Notes: Three Lectures in Computational Chemistry.
https://doi.org/10.13140/RG.2.1.1157.1282

Lecture Notes: Computational Chemistry

Data · January 2016

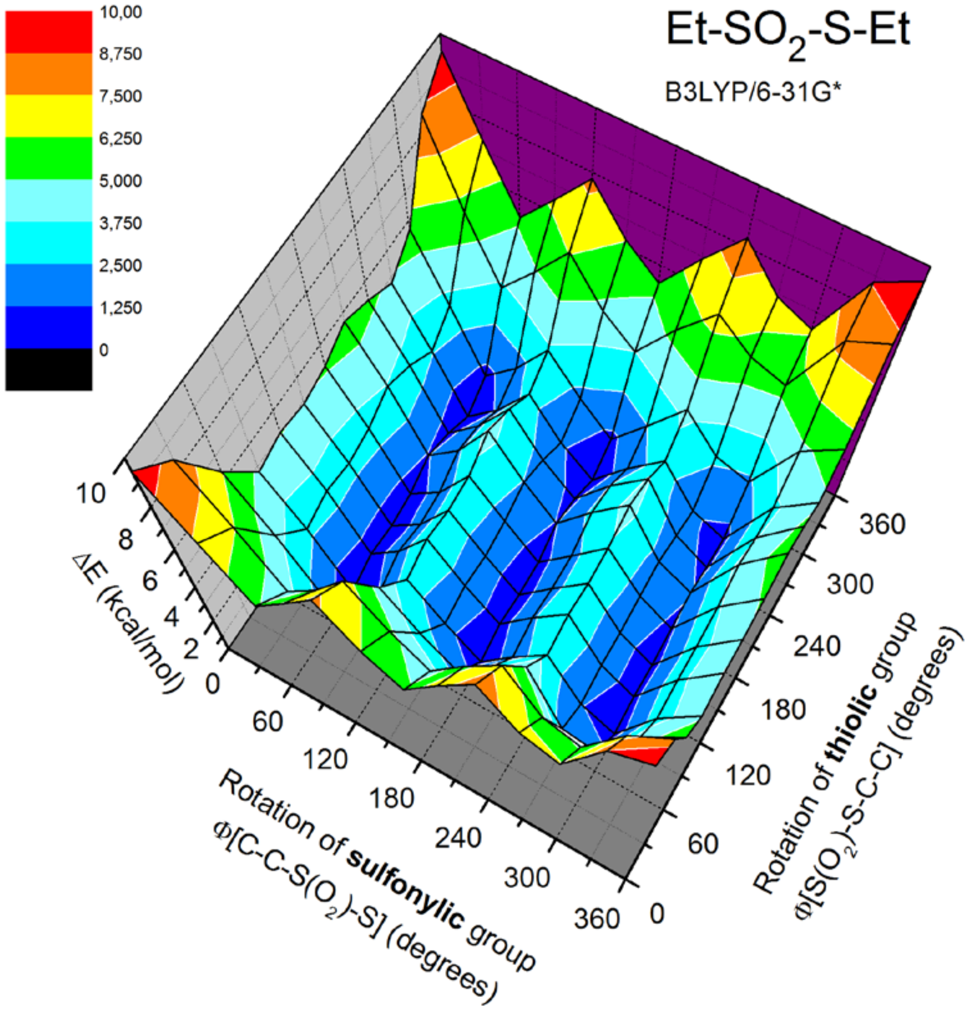
DOI: 10.13140/RG.2.1.1157.1282

Jens Spanget-Larsen
Roskilde University

Lecture Notes:
Three Lectures in Computational Chemistry

Jens Spanget-Larsen
RUC 2012-16

Largely based on a chapter by Warren J. Hehre in the textbook by Thomas Engel: "Quantum Chemistry and Spectroscopy", Pearson–Benjamin-Cummings, 2006.



Computational Chemistry
Lecture Notes 1

(RUC, NSM, September 2014)

Isolated molecule
Born-Oppenheimer approximation
    Potential energy surface
Nuclear eigenvalue problem
Electronic eigenvalue problem
    Molecular Mechanics
    MO theory
    The LCAO-MO procedure

Isolated molecule

At this point, the system under consideration is a collection of mutually interacting particles, *i.e.*, nuclei and electrons, a structureless "plasma".

We seek solutions to the time-independent, non-relativistic Schrödinger equation, $\hat{H}_{en}\Psi_{en}(\mathbf{r},\mathbf{R}) = E_{en}\Psi_{en}(\mathbf{r},\mathbf{R})$, involving the molecular Hamilton operator (atomic units, au):

$$\hat{H}_{en} = -\sum_{i} \nabla_i^2/2 - \sum_{u,i} Z_u/r_{ui} + \sum_{i<j} 1/r_{ij} + \sum_{u<v} Z_u Z_v/r_{uv} - \sum_{u} \nabla_u^2/2m_u$$

kinetic energy of elec-trons	attraction between electrons and nuclei	repulsion between electrons	repulsion between nuclei	kinetic energy of nuclei

The molecular wavefunction $\Psi_{en}(\mathbf{r},\mathbf{R})$ is a dynamical function of the coordinates of all electrons (**r**) and all nuclei (**R**). The eigenvalue problem is a many-body problem and cannot be solved exactly.

Born-Oppenheimer approximation

The nuclei are much heavier than the electrons. Hence, the electrons move very much faster than the nuclei. In the Born-Oppenheimer approximation, the motion of the electrons is decoupled from that of the nuclei, and the molecular eigenvalue problem is divided into two separate problems: One involving the motion of the electrons, and another involving the motions of the nuclei. The molecular Hamilton operator is divided into two parts, $\hat{H}_{en} = \hat{H}_e + \hat{H}_n$:

$$\hat{H}_{en} = \underbrace{-\sum_{i} \nabla_i^2/2 - \sum_{u,i} Z_u/r_{ui} + \sum_{i<j} 1/r_{ij} + \sum_{u<v} Z_u Z_v/r_{uv}}_{\hat{H}_e} \underbrace{- \sum_{u} \nabla_u^2/2m_u}_{\hat{H}_n}$$

In the electronic eigenvalue problem, the nuclei are considered as classical point charges at fixed positions in space. The nuclear coordinates **R** are input parameters to the formulation of the *electronic eigenvalue problem* (see later), involving the electronic Hamiltonian, $\hat{H}_e$.

$$\hat{H}_e \, \Psi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R}) \, \Psi_e(\mathbf{r};\mathbf{R})$$

The solutions $\Psi_e(\mathbf{r};\mathbf{R})$ and $E_e(\mathbf{R})$ are called the *electronic wavefunction* and *electronic energy*, respectively. They depend parametrically on the nuclear input coordinates **R**. Solution of the electronic eigenvalue problem for a particular set of input coordinates **R** is called a *single point calculation*, providing a "single point" on the potential energy surface (see below). In general, there are numerous solutions, corresponding to the electronic ground state and excited electronic states. In most cases, we are only interested in the ground state and a few of the lowest excited states.

Potential energy surface

A mapping of the electronic energy $E_e(\mathbf{R})$ as a function of **R** describes the *potential energy surface* for the molecule in the electronic state in question. In the case of a diatomic molecule, the surface is a *potential energy curve*. As an example is shown some results for different electronic states of NO:

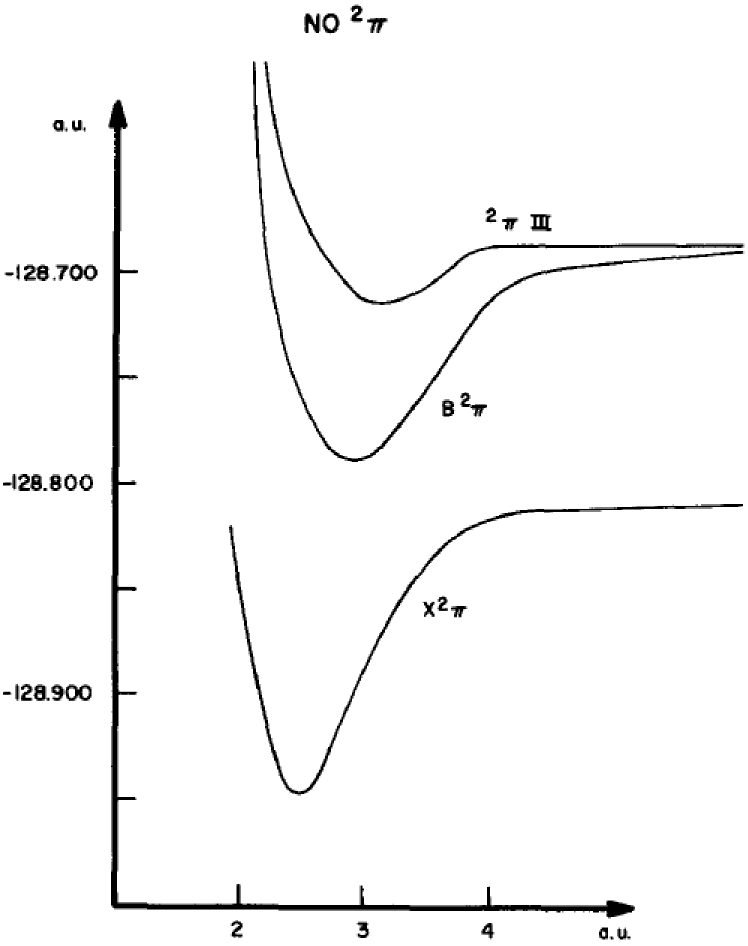
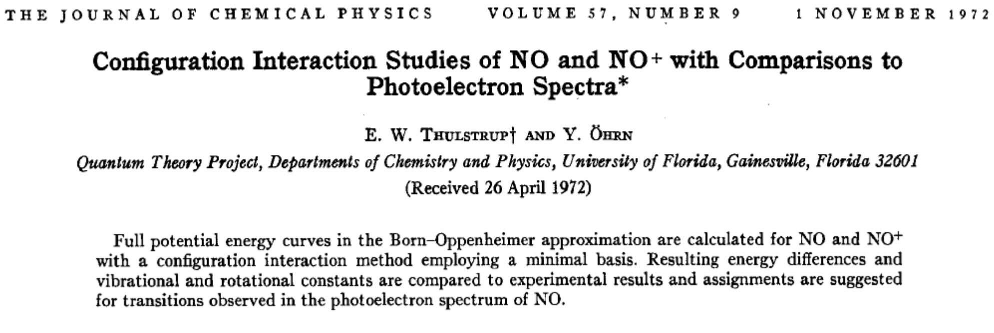


FIG. 1. $^2\Pi$ states of NO.



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 57, NUMBER 9 1 NOVEMBER 1972

Configuration Interaction Studies of NO and NO+ with Comparisons to Photoelectron Spectra*

E. W. THULSTRUP† AND Y. ÖHRN

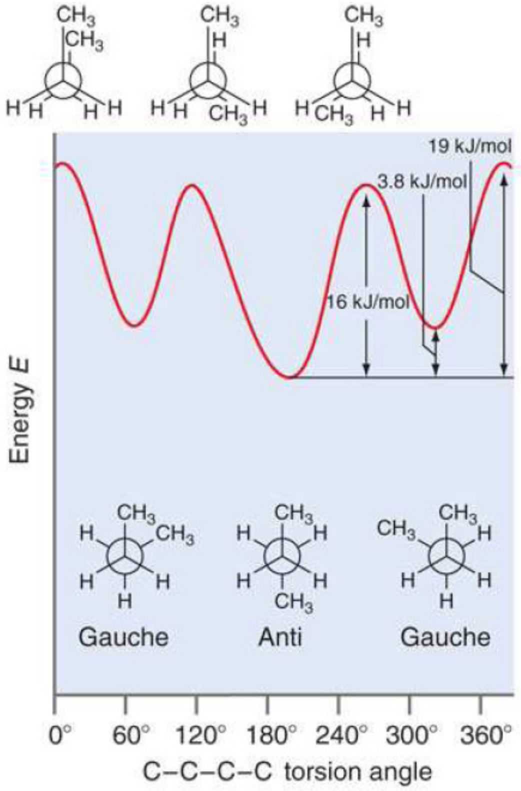
Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32601

(Received 26 April 1972)

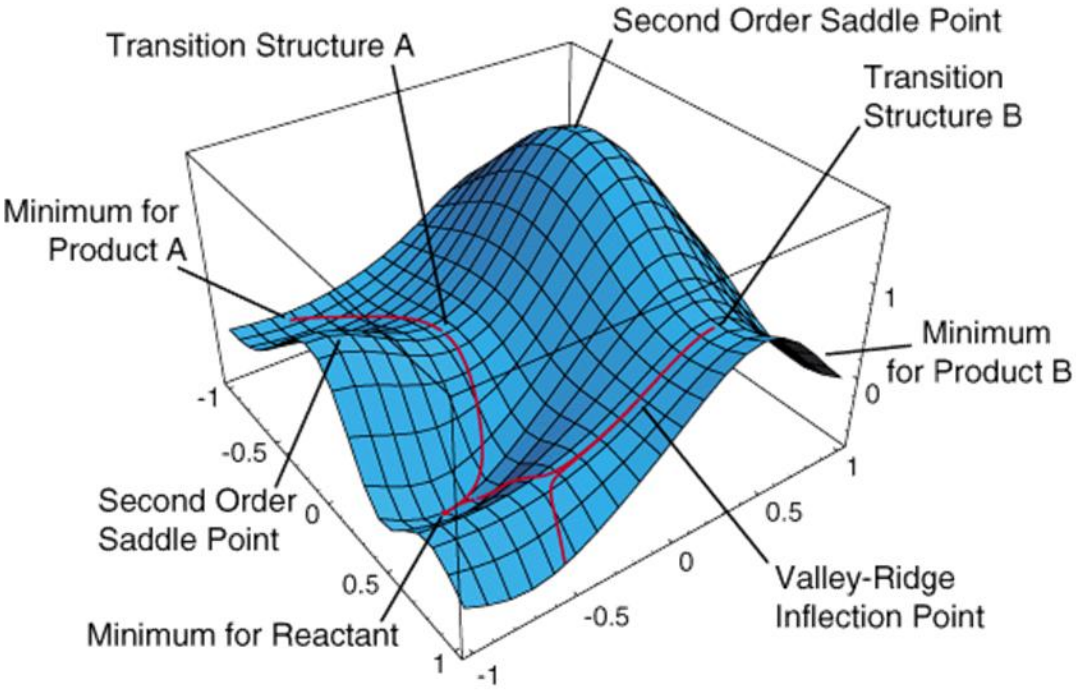
Full potential energy curves in the Born–Oppenheimer approximation are calculated for NO and NO^+ with a configuration interaction method employing a minimal basis. Resulting energy differences and vibrational and rotational constants are compared to experimental results and assignments are suggested for transitions observed in the photoelectron spectrum of NO.

In the general case, the potential energy is a function of many nuclear coordinates, and mapping of a multi-dimensional potential energy surface is less straight forward.

In most cases, the description is restricted to a few nuclear degrees of freedom that are of particular interest. In the following example (from the chapter by Hehre) a single degree of freedom is selected, corresponding to a single torsional angle, resulting in a torsional energy curve:



Two-dimensional surfaces can be represented by a contour map, or a “fishnet” diagram:



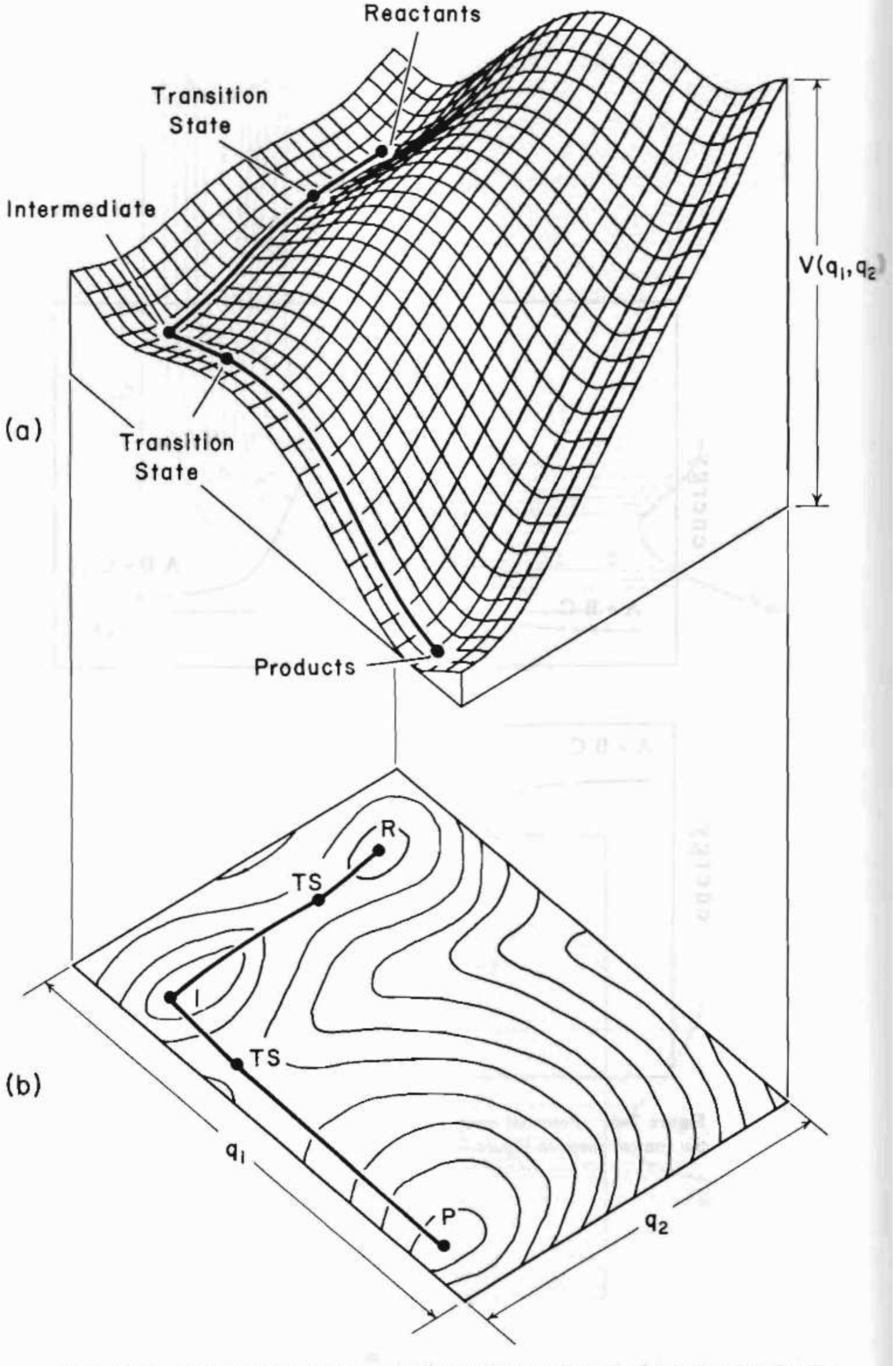
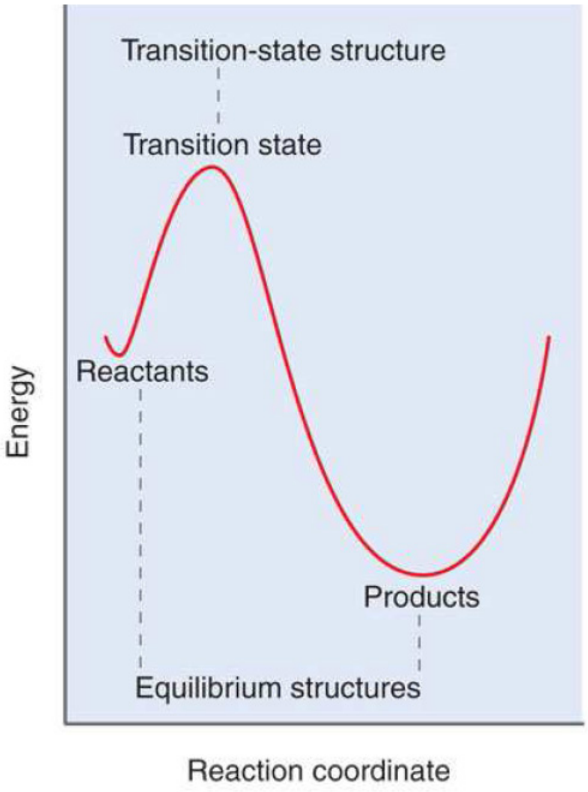
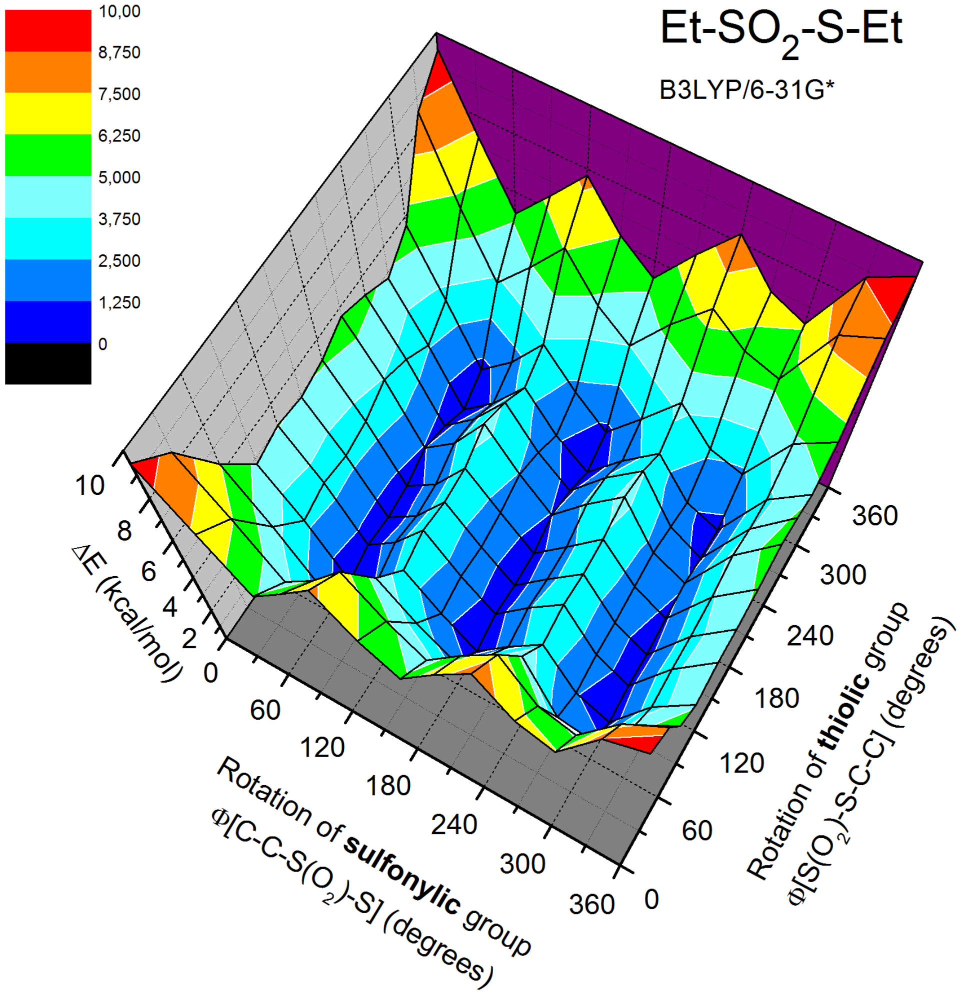


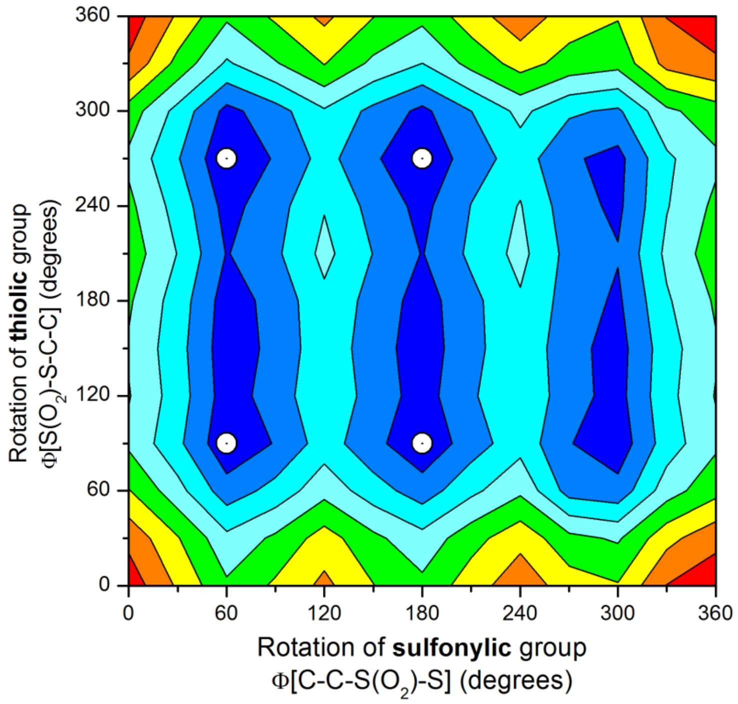
Figure 7–7. A three-dimensional perspective (a) and contour map (b) for a model chemical reaction. The solid line is the reaction path. [Adapted from G. M. Maggiora and R. E. Christofferson, in *Transition States of Biochemical Processes*, ed. R. D. Gandour and R. L. Schowen (New York: Plenum, 1978).]

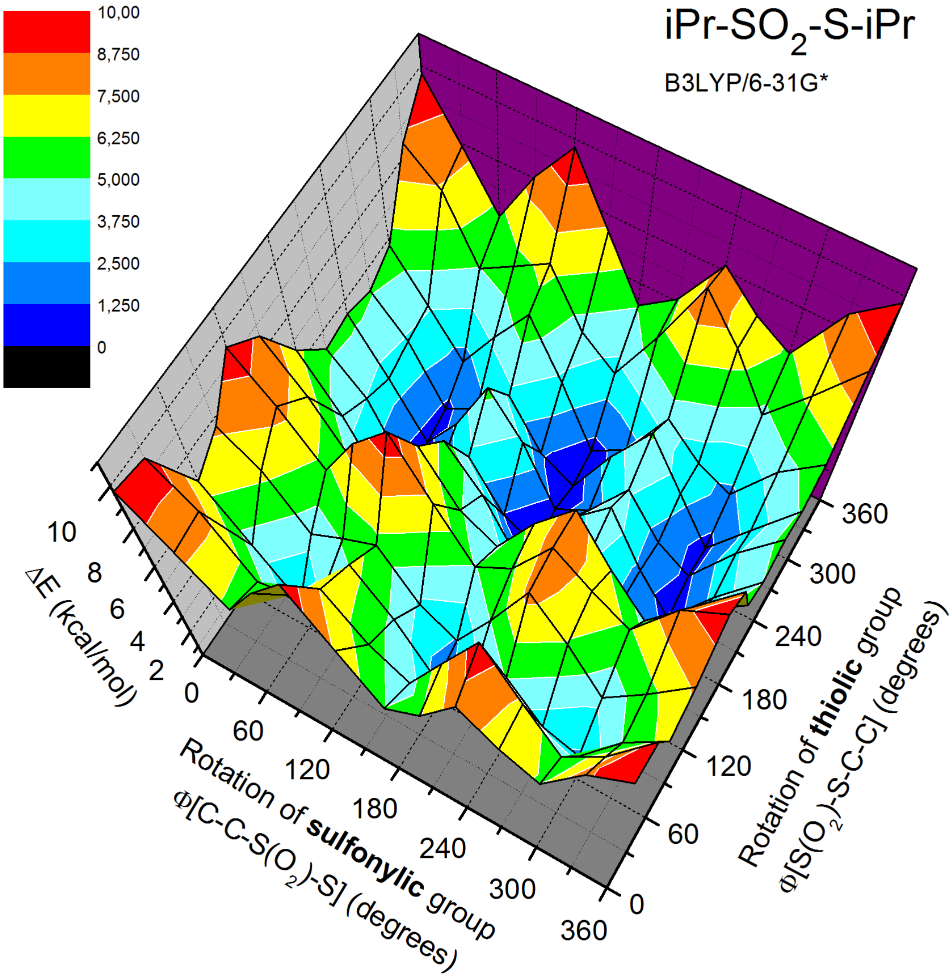
Of particular importance is the location of stationary points on the surface, *i.e.*, points where the gradient of $E_e(\mathbf{R})$ is zero (also called extrema), corresponding to *minima*, *maxima*, and *saddle points* (first and higher order saddle points). Efficient computer algorithms are developed to locate these points. The global minimum of the surface, *i.e.*, the point with the lowest energy, defines the nuclear *equilibrium configuration* for the molecule in the specific electronic state. Other minima correspond to local equilibria, indicating, e.g., rotamers or isomers. First order saddle points indicate *transition structures*, interrelating different minima. The *reaction coordinate*, describing the transition from one minimum, via a transition structure, to another minimum, amounts to the description of a molecular rearrangement or a chemical reaction:

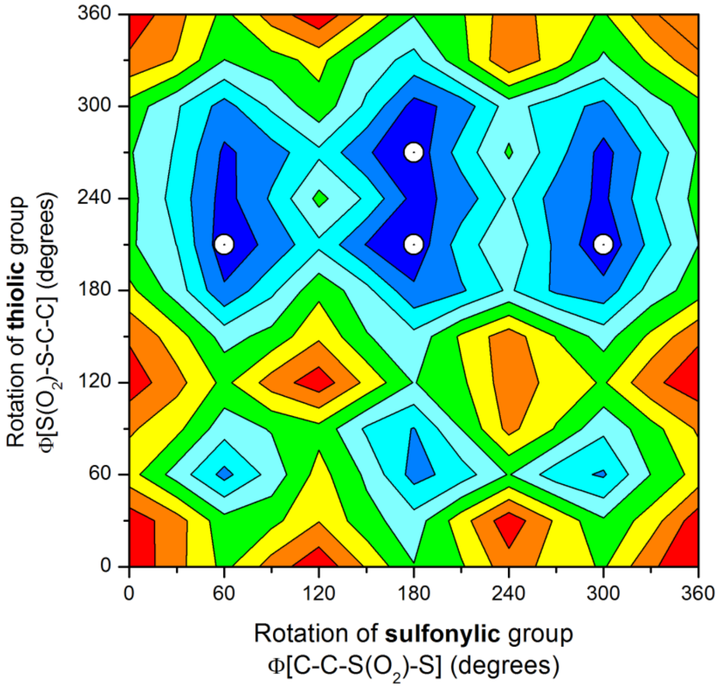


As a final example, the next two pages show color-coded representations of the computed two-dimensional energy surfaces spanned by two torsional angles of the compounds *S*-ethyl ethanethiosulfonate and *S*-isopropyl propane-2-thiosulfonate. The analyses revealed that these thiosulfonates are present as equilibria between several rotamers, giving rise to distinctly different IR spectra (T.X.T. Luu, F. Duus, J. Spanget-Larsen, *J. Mol. Struct.* **1049**, 165-171 (2013)).


Et-SO$_2$-S-Et
B3LYP/6-31G*
10,00
8,750
7,500
6,250
5,000
3,750
2,500
1,250
0
ΔE (kcal/mol)
0
2
4
6
8
10
60
120
180
240
300
360
Rotation of **sulfonylic** group
Φ[C-C-S(O$_2$)-S] (degrees)
0
60
120
180
240
300
360
Rotation of **thiolic** group
Φ[S(O$_2$)-S-C-C] (degrees)


360
300
240
180
120
60
0
Rotation of **thiolic** group
Φ[S(O$_2$)-S-C-C] (degrees)
0
60
120
180
240
300
360
Rotation of **sulfonylic** group
Φ[C-C-S(O$_2$)-S] (degrees)


10,00
8,750
7,500
6,250
5,000
3,750
2,500
1,250
0
iPr-SO$_2$-S-iPr
B3LYP/6-31G*
ΔE (kcal/mol)
10
8
6
4
2
0
60
120
180
240
300
360
Rotation of **sulfonylic** group
Φ[C-C-S(O$_2$)-S] (degrees)
0
60
120
180
240
300
360
Rotation of **thiolic** group
Φ[S(O$_2$)-S-C-C] (degrees)


360
300
240
180
120
60
0
Rotation of **thiolic** group
Φ[S(O$_2$)-S-C-C] (degrees)
0
60
120
180
240
300
360
Rotation of **sulfonylic** group
Φ[C-C-S(O$_2$)-S] (degrees)

Nuclear eigenvalue problem

The nuclear eigenvalue problem for the motion of the nuclei of a molecule with state wavefunction $\Psi_e(\mathbf{r};\mathbf{R})$ and electronic energy $E_e(\mathbf{R})$ is given by

$$[\hat{H}_n + E_e(\mathbf{R})]\ \Phi_n(\mathbf{R}) = E_{en}\ \Phi_n(\mathbf{R})$$

Note that the electronic energy function $E_e(\mathbf{R})$, determined by solution of the electronic eigenvalue problem, serves as potential energy operator in the nuclear problem. The total Born-Oppenheimer wavefunction, which is an approximate eigenfunction to the molecular hamiltonian operator $\hat{H}_{en}$, is written as the product

$$\Psi_{en}(\mathbf{r},\mathbf{R}) = \Psi_e(\mathbf{r};\mathbf{R})\cdot\Phi_n(\mathbf{R})$$

with total molecular energy E_{en}.

During the treatment of the nuclear problem, it is convenient to work with a “zeroed” potential energy function $V_n(\mathbf{R})$ defined as $E_e(\mathbf{R}) - E_e(\mathbf{R}_{eq})$, where $\mathbf{R}_{eq}$ indicates the nuclear equilibrium configuration:

$$[\hat{H}_n + \hat{V}_n(\mathbf{R})]\ \Phi_n(\mathbf{R}) = E_n\ \Phi_n(\mathbf{R})$$

$$E_n = E_{en} - E_e(\mathbf{R}_{eq})$$

The total energy of the molecule can then be written as the sum of an electronic and a nuclear energy, $E_{en} = E_e + E_n$, where E_e is taken as $E_e(\mathbf{R}_{eq})$.

As a first approximation, the motions of the nuclei can be separated into molecular *vibration*, *rotation*, and *translation* and the nuclear wavefunction $\Phi_n(\mathbf{R})$ can be factorized as

$$\Phi_n = \Phi_{vib}\cdot\Phi_{rot}\cdot\Phi_{trans}$$

with associated energies $E_e = E_{vib} + E_{rot} + E_{trans}$. The complete molecular wavefunction is then approximated by $\Psi_{en} = \Psi_e\cdot\Phi_n = \Psi_e\cdot\Phi_{vib}\cdot\Phi_{rot}\cdot\Phi_{trans}$, and the total molecular energy is obtained as

$$E_{en} = E_e + E_{vib} + E_{rot} + E_{trans}$$

In general, we have $E_n >> E_{vib} >> E_{rot} >> E_{trans}$. In a description of an isolated molecule we can neglect translation. As a starting point, molecular rotation is approximately described within the *rigid rotor approximation*, and molecular vibration within the *harmonic approximation*. In the latter, the function $V_n(\mathbf{R})$ is replaced by a second order polynomial in the displacement coordinates $\mathbf{R} - \mathbf{R}_{eq}$. This allows reformulation of the problem in terms of independent one-dimensional harmonic oscillators, one oscillator for each *normal mode* of vibration. The harmonic approximation leads to important *selection rules* for vibrational transitions ($\Delta v = \pm 1$, *etc*).

Electronic eigenvalue problem

This is the eigenvalue problem of the electronic Hamilton operator, as defined above:

$$\hat{H}_e \ \Psi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R}) \ \Psi_e(\mathbf{r};\mathbf{R})$$

Solution of this problem is central to computational chemistry, but it is a very difficult task, and it can only be solved approximately.

Molecular Mechanics
The most radical procedure is to avoid solution of the electronic eigenvalue problem and proceed directly to the nuclear potential energy $V_n(\mathbf{R})$, which is *parameterized empirically*. This amounts to considering the molecule as a system of atomic centers linked by bonds with prescribed mechanical properties, corresponding to a so-called *molecular mechanics* or *force field* model (containing carefully adjusted force fields for bond lengths, bond angles, and torsional angles, and possibly for steric, electrostatic, and other effects). The essential input data to the calculation is the molecular constitution, *i.e.*, a molecular "graph" with indication of the types of bonds between the atomic centers. There is no explicit representation of the electrons in the model; hence, it cannot easily describe the formation and breaking of covalent bonds during a chemical reaction. However, several highly refined models have been developed, and they are very useful in the study of the molecular structure of large systems, such as polymers, proteins, *etc*.

MO Theory
A frequently applied and very useful model of the electronic system is the *molecular orbital model*. In the orbital model, the electronic ground state is described by a many-electron wavefunction Ψ_g, which is defined as a product of one-electron wavefunctions ψ_i. Hence, for a system with *n* electrons:

$$\Psi_g = \psi_1 \cdot \psi_2 \cdot \psi_3 \cdot \ \dots \ \cdot \psi_n$$

The product wavefunction is an approximate eigenfunction to the electronic hamiltonian operator $\hat{H}_\mathrm{e}$ for the molecule, with eigenvalue E_g. The one-electron wavefunctions ψ_i are called *molecular orbitals* (MOs). MOs play the same role in the description of molecules, as *atomic orbitals* (AOs) do in the description of atoms. In principal, the only difference between AOs and MOs is that an AO is a one-electron wavefunction for an electron system in the field of one nucleus, while an MO is a one-electron wavefunction for an electron system in the field of two or several nuclei.

A wavefunction like Ψ_g which is constructed as a product of orbital wavefunctions is called a *configuration wavefunction*. More precisely, the wavefunction must be written as an anti-symmetrised product (a Slater-determinant), which means that the wavefunction Ψ_g changes sign "if two electrons are interchanged" (the Pauli principle). Here we shall not consider this aspect in detail, but it has the important consequence that two electrons can only be described by the same spatial MO if they have different spin. There can "be" at most two electrons in each MO, one with α and one with β spin.

The MOs ψ_i and their energies ε_i are determined as the eigenfunctions and eigenvalues of an effective one-electron energy operator $\hat{h}$, which can be written as

$$\hat{h} = \hat{t} + \hat{v}_{\mathrm{n}} + \hat{\bar{v}}_{\mathrm{e}}$$

Here $\hat{t}$ represents the kinetic energy of the electron, $\hat{v}_{\mathrm{n}}$ represents the electrostatic attraction from the nuclei in the molecule, and $\hat{\bar{v}}_{\mathrm{e}}$ represents the repulsion from the other electrons in the molecule. The last term $\hat{\bar{v}}_{\mathrm{e}}$ is problematic. In order to formulate $\hat{h}$ as a one-electron operator, the electron-electron interaction must be described by an effective *mean field approximation*: Each electron does not "see" the other electrons as individual particles; the field from the other electrons is represented by an averaged, static charge distribution. This charge distribution depends on the many-electron wavefunction Ψ_g, and thus on the MOs ψ_i. This means that the operator $\hat{h}$ depends on its own eigenfunctions! These are not known by the start of the calculation, and in general, the solution of the eigenvalue problem requires an iterative technique, the *Self Consistent Field* (SCF) procedure. The variationally best solution, i.e., the solution of the MO model that leads to the lowest possible total energy E_g, is called the **Hartree-Fock** (HF) solution. For a mathematical definition, see the chapter by Hehre.

The orbital model of a many-electron system involves a decoupling of the motions of the electrons (like the Born-Oppenheimer approximation separates the motions of nuclei and electrons). The model does not describe the instantaneous *correlation* of the individual motions of the electrons. It is important to understand that the model is only an approximation. Orbitals are one-electron wavefunctions that serve as useful elements in an approximate description of a many-electron system, but they do not, in principle, correspond to physically observable quantities! But it can be shown that in a model where electron correlation and reorganization effects are neglected, the negative of the orbital energy ε_i is equal to the ionization energy I_i required to remove an electron from the orbital ψ_i. This result is known as *Koopmans' theorem*. The approximate relation $I_i \approx -\varepsilon_i$ is of great importance in the assignment of photoelectron spectra. However, the usefulness of the relation depends on the circumstance that the errors due to neglect of electron correlation and neglect reorganization effects have opposite sign and tend to cancel each other out.

In the MO model, the total energy of the electrons can be written

$$E = \sum_i \varepsilon_i - V_{\mathrm{ee}}$$

where the sum is over the MO energies for all electrons i and V_{ee} is the total electronic repulsion energy. The typical *error* due to the fact that the model neglects electron correlation amounts to about 1%. Note that the total energy is not equal to the sum of the MO energies. This is because that in this sum, the contribution from electron interaction is counted twice (the repulsion between the *i*'th and the *j*'th electron contributes to ε_i as well as to ε_j). Therefore, V_{ee} must be subtracted. Nevertheless, the sum of the MO energies is frequently useful as a qualitative measure of the energy of the molecule. This can be rationalized by a consideration of the total energy of the molecule, which can be written

$$E_{\mathrm{e}} = \sum_i \varepsilon_i - V_{\mathrm{ee}} + V_{\mathrm{nn}}$$

where V_{nn} is the mutual repulsion of the nuclei. In many qualitative considerations, the difference $V_{\mathrm{nn}} - V_{\mathrm{ee}}$ can be set equal to zero or be taken as an approximately constant quantity (f. inst. during variation of a bond angle). In the *"Free electrons"*, *Hückel*, and *Extended Hückel* MO models (see later) where electron interaction is not considered explicitly, the "total energy" is traditionally taken as the sum of the MO energies, $E = \Sigma\,\varepsilon_i$.

The LCAO-MO procedure

Most MO calculations are based on the LCAO-MO procedure (Linear Combinations of Atomic Orbitals). In this model, an MO ψ is written as a weighted sum of atomic orbitals (AOs) φ_ν from the atoms in the molecule:

$$\psi = \sum_\nu c_\nu \varphi_\nu$$

There are great advantages associated with a description of MOs on the basis of AOs. In the first place, it must be expected that the AOs are particularly suitable; one can imagine that an electron close to a particular nucleus in a molecule is primarily affected by the field from this nucleus, and the MO of the electron should thus be related to the AOs of the isolated atom. Secondly, the circumstance that the AOs are orbitals for the constituting parts of the molecular system enables simplifications, not only what concerns calculational techniques, but also in a more conceptual vein (f.ex. in connection with the population analysis, see later).

Computational Chemistry
Lecture Notes 2

(RUC, NSM, September 2012)

The LCAO-MO Procedure
Secular Equations
The Hartree-Fock Solution
Basis Sets
Minimal Basis Sets
Split-Valence Basis Sets
Polarization Basis Sets
Diffuse Functions
Approximate Hartree-Fock models

The LCAO-MO Procedure

Most MO calculations are based on the LCAO-MO procedure (Linear Combinations of Atomic Orbitals). In this model, an MO ψ_i is written as a weighted sum of atomic orbitals (AOs) φ_ν from the atoms in the molecule:

$$\psi_i = \sum_\nu c_{\nu i}\varphi_\nu$$

There are great advantages associated with a description of MOs on the basis of AOs. In the first place, it must be expected that the AOs are particularly suitable; one can imagine that an electron close to a particular nucleus in a molecule is primarily affected by the field from this nucleus, and the MO of the electron should thus be related to the AOs of the isolated atom. Secondly, the circumstance that the AOs are orbitals for the constituting parts of the molecular system enables simplifications, not only what concerns calculational techniques, but also in a more conceptual vein (f.ex. in connection with the population analysis, see later).

Secular Equations

The adopted set of AOs φ_ν is referred to as the *basis set* $\{\varphi\}$ in which the MOs are expanded, and the number of AOs included is the size (N) of the basis set. The weight factors c_ν are called *LCAO coefficients*. The coefficient $c_{\nu i}$ indicates how much the AO φ_ν contributes to the MO ψ_i (the contribution may be positive, negative, or zero). The task now consists in determination of the LCAO coefficients and the corresponding MO energy. The MO ψ is an eigenfunction of the effective one-electron operator $\hat{h}$:

$$\hat{h}\psi = \varepsilon\psi \qquad \Leftrightarrow \qquad (\hat{h}-\varepsilon)\psi = 0$$

As previously mentioned, the operator $\hat{h}$ depends on its own eigenfunctions, and they are unknown at the start of the calculation; we must thus adopt a suitable starting approximation for $\hat{h}$. Introduction of the expansion $\psi = \sum_\nu c_\nu \varphi_\nu$ leads to

$$(\hat{h}-\varepsilon)\sum_\nu^N c_\nu\varphi_\nu = \sum_\nu^N (\hat{h}\varphi_\nu - \varepsilon\varphi_\nu)c_\nu = 0$$

(for convenience, the MO index *i* is omitted). This equation can be transformed into system of linear equations. Multiplication from the left with an arbitrary AO, say φ_μ,

$$\sum_\nu^N (\varphi_\mu\hat{h}\varphi_\nu - \varepsilon\varphi_\mu\varphi_\nu)c_\nu = 0$$

and integration over all space leads to

$$\sum_{\nu}^{N}\left(\int \varphi_{\mu}\hat{h}\varphi_{\nu}dq-\varepsilon\int \varphi_{\mu}\varphi_{\nu}dq\right)c_{\nu}=0$$

where $dq = dxdydz$. For simplicity, we introduce the notation $h_{\mu\nu}$ for the integral $\int\varphi_{\mu}\hat{h}\,\varphi_{\nu}dq$. This integral is frequently referred to as a *matrix element* of the operator $\hat{h}$ in the basis $\{\varphi\}$, and we have $h_{\mu\nu} = h_{\nu\mu}$ (because of the properties of physically acceptable operators). We further introduce the symbol $S_{\mu\nu}$ for the integral $\int\varphi_{\mu}\varphi_{\nu}dq$. The integral $S_{\mu\nu} = S_{\nu\mu}$ is called the *overlap integral* of the AOs φ_{μ} and φ_{ν}. With these symbols the equation is written

$$\sum_{\nu}^{N}(h_{\mu\nu}-\varepsilon S_{\mu\nu})c_{\nu}=0$$

Above we multiplied from the left by an arbitrary AO φ_{μ}. Hence, we can generate as many different equations as the number of AOs in the basis set $\{\varphi\}$. The resulting system of N linear equations is termed the *secular equations*:

$$\begin{array}{l}
(h_{11}-\varepsilon S_{11})c_1+(h_{12}-\varepsilon S_{12})c_2+\quad\dots\quad+(h_{1N}-\varepsilon S_{1N})c_N=0\\
(h_{21}-\varepsilon S_{21})c_1+(h_{22}-\varepsilon S_{22})c_2+\quad\dots\quad+(h_{2N}-\varepsilon S_{2N})c_N=0\\
\qquad\vdots\qquad\qquad\qquad\vdots\qquad\qquad\qquad\qquad\qquad\vdots\\
(h_{N1}-\varepsilon S_{N1})c_1+(h_{N2}-\varepsilon S_{N2})c_2+\ \dots\ +(h_{NN}-\varepsilon S_{NN})c_N=0
\end{array}$$

The AOs can be assumed to be normalized wavefunctions, which means that all ‘diagonal’ overlap integrals are equal to unity, $S_{\mu\mu} = 1$. The secular equations are often written in matrix form,
$(\boldsymbol{h} - \varepsilon\boldsymbol{S})\boldsymbol{c} = \mathbf{0}$:

$$\begin{pmatrix}
h_{11}-\varepsilon & h_{12}-\varepsilon S_{12} & \cdots & h_{1\mathrm{N}}-\varepsilon S_{1\mathrm{N}}\\
h_{21}-\varepsilon S_{21} & h_{22}-\varepsilon & \dots & h_{2\mathrm{N}}-\varepsilon S_{2\mathrm{N}}\\
\vdots & \vdots & \ddots & \vdots\\
h_{\mathrm{N}1}-\varepsilon S_{\mathrm{N}1} & h_{\mathrm{N}2}-\varepsilon S_{\mathrm{N}2} & \cdots & h_{\mathrm{NN}}-\varepsilon
\end{pmatrix}
\begin{pmatrix} c_1\\ c_2\\ \vdots\\ c_N\end{pmatrix}=
\begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix}$$

The equation system has a trivial solution, namely the one where all coefficients c_{ν} are equal to zero. The equations only have a non-trivial solution if the *secular determinant* is equal to zero, $|\boldsymbol{h} - \varepsilon\boldsymbol{S}| = 0$:

$$\begin{vmatrix}
h_{11}-\varepsilon & h_{12}-\varepsilon S_{12} & \cdots & h_{1\mathrm{N}}-\varepsilon S_{1\mathrm{N}}\\
h_{21}-\varepsilon S_{21} & h_{22}-\varepsilon & \dots & h_{2\mathrm{N}}-\varepsilon S_{2\mathrm{N}}\\
\vdots & \vdots & \ddots & \vdots\\
h_{\mathrm{N}1}-\varepsilon S_{\mathrm{N}1} & h_{\mathrm{N}2}-\varepsilon S_{\mathrm{N}2} & \cdots & h_{\mathrm{NN}}-\varepsilon
\end{vmatrix}=0$$

Expansion of the determinant yields an N'th order polynomium in ε, and the N roots of this polynomium are the MO energies ε_i. The LCAO coefficients for the i'th MO ψ_i are obtained by solving the secular equations for $\varepsilon = \varepsilon_i$, and then normalize ψ_i.

Solution of the secular problem yields just as many MOs as the number of AOs in the basis set (*i.e.*, N), and they are all mutually orthogonal: $\int \psi_i \psi_j \, dq = 0$. The many-electron wavefunction Ψ_g is obtained by populating the lowest MOs according to the aufbau-principle (max. two electrons per orbital). On the basis of this wavefunction, the electron distribution in the molecular can be calculated, and a new and better approximation of the effective one-electron operator $\hat{h}$ can be set up, and the procedure is repeated until self-consistency is achieved (SCF procedure).

The Hartree-Fock Solution

If no further approximations are introduced into the model, and a large basis set is applied, the MO calculation is called an *ab initio* **Hartree-fock (HF)** calculation. Within this framework, the effective one-electron energy operator $\hat{h}$ is called the Fock-operator $\hat{f}$. For a closed-shell molecule, where all MOs are either doubly occupied or empty, the matrix elements of the Fock-operator take the following form:

$$f_{\mu\nu} = t_{\mu\nu} + v_{\mu\nu} + \sum_{\rho\sigma} p_{\rho\sigma}\left[(\mu\nu \mid \rho\sigma) - ½(\mu\rho \mid \nu\sigma)\right]$$

Here $t_{\mu\nu}$ is a matrix element of the kinetic energy operator, $v_{\mu\nu}$ is an element of the potential energy operator in the field of all nuclei, $(\mu\nu \mid \rho\sigma)$ are two-electron interaction integrals in Mulliken's notation, and $p_{\rho\sigma}$ are elements of the molecular density matrix (a.u.):

$$t_{\mu\nu} = \int \varphi_\mu(\mathbf{r}_1)\left[\frac{-\nabla^2}{2}\right]\varphi_\nu(\mathbf{r}_1)\,d\mathbf{r}_1$$

$$v_{\mu\nu} = \int \varphi_\mu(\mathbf{r}_1)\left[\sum_u \frac{-Z_u}{r_{1u}}\right]\varphi_\nu(\mathbf{r}_1)\,d\mathbf{r}_1$$

$$(\mu\nu \mid \rho\sigma) = \iint \varphi_\mu(\mathbf{r}_1)\varphi_\nu(\mathbf{r}_1)\left[\frac{1}{r_{12}}\right]\varphi_\rho(\mathbf{r}_2)\varphi_\sigma(\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2$$

$$p_{\rho\sigma} = \sum_i^{\text{MO}} n_i \, c_{\rho i} c_{\sigma i} \,,$$

n_i = occupation number: 2 for doubly occupied MOs, 0 for empty

The number of two-electron integrals $(\mu\nu \mid \rho\sigma)$ increases as the fourth power of the number of basis functions, N. The cost of a calculation thus rises rapidly with the size of the basis set.

Basis Sets

The LCAO approximation requires a basis set of well-defined functions centered on each atomic nucleus in the molecule. In the early years of computational chemistry, many procedures applied so-called Stater-type orbitals (STOs), which are approximate atomic orbitals (a.u.):

$$\varphi(r,\theta,\phi) = N \cdot r^{n-1} e^{-\zeta r} Y_l^m(\theta,\phi)$$

Here, N is a normalization constant, n, l, and m are the usual atomic quantum numbers, and ζ ("zeta") is an effective nuclear charge: $\zeta = (Z - S)/n$, where Z is the nuclear charge and S a screening constant. $Y_l^m(\theta,\phi)$ is a spherical harmonic function, defining the angular part of the basis function. STO-type functions are no longer used in *ab initio* calculations because they have been found to be computationally disadvantageous, particularly in the evaluation of the numerous two-electron repulsion integrals $(\mu\nu|\rho\sigma)$ which represent the 'bottle-neck' in accurate computations (but STOs are still applied in many approximate, semi-empirical procedures). Today, most computational chemistry procedures apply functions expanded in terms of *Gaussians*:

$$g_{ijk}(\mathbf{r}) = N \cdot x^i y^j z^k e^{-\alpha r^2}$$

In this function, x, y, z are the Cartesian coordinates of the electron with respect to the nucleus of the atom, i, j, and k are positive integers, and α is an orbital exponent.

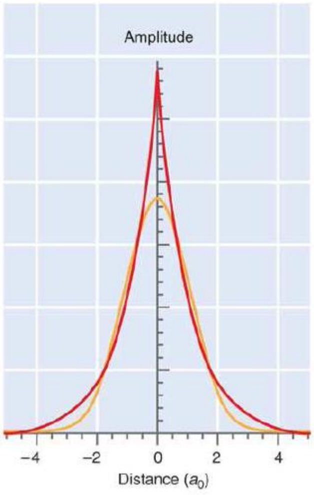
s-type: $g_{000}(\mathrm{r}) = N \cdot e^{-\alpha r^2}$

p-type: $g_{100}(\mathrm{r}) = N \cdot x\, e^{-\alpha r^2}$

$g_{010}(\mathrm{r}) = N \cdot y\, e^{-\alpha r^2}$

$g_{001}(\mathrm{r}) = N \cdot z\, e^{-\alpha r^2}$

d-type functions are generated by combinations that give $i + j + k = 2$, and so forth. The Gaussian functions lead to integrals that are easily computed, but the radial dependence $e^{-\alpha r^2}$ means that a single Gaussian is a rather poor representation of an atomic orbital; in this respect the STO is much better. The solution is to describe each atomic orbital by a normalized linear combination of several Gaussians:



$$\varphi_\mu = \sum_p a_{\mu p}\, g_p(\mathbf{r})$$

The components g_p in this expansion are called *primitive Gaussians*, and in general, these components are not considered as individual members of the basis set. The expansion coefficients $a_{\mu p}$ are determined according to some optimization criterion, and the resulting linear combination is used as a basis function. The resulting function with fixed coefficients is called a *contracted Gaussian*.

Minimal Basis Sets

A minimal basis set contains the minimum number of basis functions needed to hold the electrons of each atom, as in these examples:

H – He: 1s
Li – Ne: 1s, 2s, 2p_x, 2p_y, 2p_z
Na – Ar: 1s, 2s, 2p_x, 2p_y, 2p_z, 3s, 3p_x, 3p_y, 3p_z

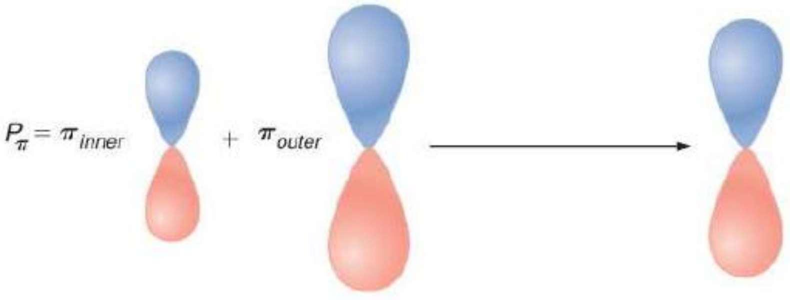
One of the most well-known minimal basis sets is the so-called STO-3G basis, where each basis function is expanded in terms of three primitive Gaussians (the expansion coefficients are optimized to produce the best fit to Slater-type exponential functions). However, this and other minimal basis sets suffer from insufficient flexibility to describe distortion of the electron density around the atomic centers in a molecular environment.

Split-Valence Basis Sets

The first way to enlarge the basis set is to increase the number of basis functions per atom. Simple split-valence basis sets, such as 3-21G and 6-31G, have two sizes of basis function for each valence orbital:

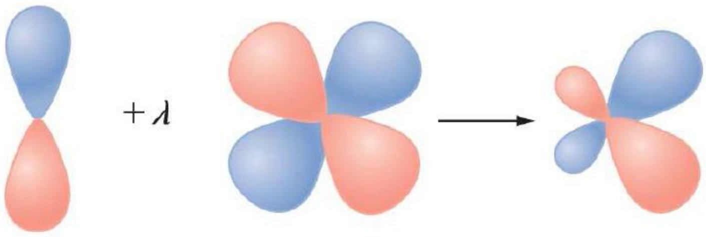
H – He: 1s, 1s'
Li – Ne: 1s, 2s, 2s', 2p_x, 2p_y, 2p_z, 2p_x', 2p_y', 2p_z'
Na – Ar: 1s, 2s, 2p_x, 2p_y, 2p_z, 3s, 3s', 3p_x, 3p_y, 3p_z, 3p_x', 3p_y', 3p_z'

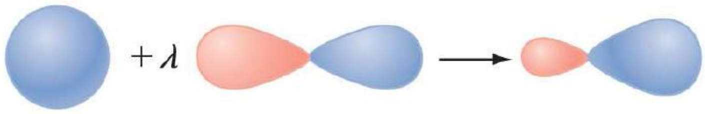
where the primed and unprimed orbitals differ in size. These basis sets are called *split valence double zeta* basis sets (remember the use of the symbol zeta (ζ) for the orbital exponent). Similarly, *split-valence triple zeta* basis sets, like 6-311G, uses three sizes of contracted functions for each orbital type. A split-valence basis set enables the electron density about an atom to expand differently in different directions, corresponding to a non-spherical electron distribution:



Polarization Basis Sets

Split-valence basis sets allow orbitals to change size, but not to change shape. This limitation is removed by adding *polarization functions*, i.e., orbital functions with angular momentum beyond what is required for the ground state, to the description of each atom. For example, polarization basis sets add *d* functions to carbon and *f* functions to transitions metals, and some of them add *p* functions to hydrogen atoms. Polarization functions provide flexibility for the orbitals to modify their shape, and, for example, shift the center of the electron distribution to the bonding region between atoms:





A simple and robust polarization basis set is 6-31G(d), also known as 6-31G*. Its name indicates that it is the 6-31G basis set with *d* functions added to heavy (non-hydrogen) atoms. This basis set is very popular in calculations on medium-sized molecular systems. Another popular polarization basis set is 6-31G(d,p), also known as 6-31G**, which adds *p* functions to hydrogen in addition to the *d* functions on heavy atoms.

Diffuse Functions

Diffuse functions are large-size versions of *s*- and *p*-type functions (as opposed to the standard valence-size functions). They allow orbitals to occupy a larger region of space, which is important in the description of systems where the electrons are relatively far from the nuclei. Typical examples are anions and molecules in excited electronic states. The inclusion of diffuse functions in the basis set is designated with a plus sign: The 6-31+G(d) basis set is the 6-31G(d) basis set with diffuse functions added to heavy atoms. The double plus version, 6-31++G(d), adds diffuse functions to the hydrogen atoms as well (diffuse functions on hydrogen atoms seldom make a significant difference in calculational accuracy).

Approximate Hartree-Fock models

For large systems, or if fast results are required, it may be desirable to adopt one of the many available approximate HF models. These models were particularly popular in the decades before the development of the present-day powerful computers. They involve a series of simplifications of the HF problem:

(1) The *all-valence-electrons* procedures include only a minimal basis of valence orbitals in the description, considering the inner shells as part of a fixed *core*. The basis functions are typically taken as a set of Slater-type orbitals (STOs).

(2) The mathematical expression for the elements of the Fock operator is simplified in order to avoid calculation of the many integrals. In the *neglect of differential overlap* (NDO) procedure, the formalism is simplified by the approximation $\varphi_\mu(\mathbf{r})\varphi_\nu(\mathbf{r}) = \delta_{\mu\nu}\,\varphi_\mu(\mathbf{r})\varphi_\mu(\mathbf{r})$, leading to the important simplifications $S_{\mu\nu} = \delta_{\mu\nu}$ and $(\mu\nu|\rho\sigma) = \delta_{\mu\nu}\,\delta_{\rho\sigma}\,(\mu\mu|\rho\rho)$. The latter result means that all three- and four-center electron repulsion integrals vanish, which makes calculations on very large systems feasible. Several NDO models (like AM1 and PM3) make exceptions from the strict NDO approximation for one-center terms, and maintain, e.g., one-center electron repulsion integrals of the type $(\mu\nu|\mu\nu)$.

(3) Most of the remaining integrals are estimated by *semi-empirical procedures*. This means that they are not properly calculated, but they are assigned values by reference to empirical data.

Popular semi-empirical all-valence-electrons NDO procedures are *AM1* (“Austin Model 1”) and *PM3* (“Parametric Model 3”) which are included in commercial software packages like HYPERCHEM® and GAUSSIAN®. The *PM*-model is still being expanded and refined, the current version is *PM7*.

The *Extended Hückel Method* (EHM) occupies a particular position among the semi-empirical all-valence-electrons methods. It should probably not be considered as an approximate HF model, since it is not easily derived from the formal expression for the Fock operator. The elements $h_{\mu\nu}$ are approximated in the following manner:

$$\begin{aligned} h_{\mu\mu} &= -I_\mu \\ h_{\mu\nu} &= -1.75 \cdot S_{\mu\nu}\,(I_\mu + I_\nu)/2 \qquad (\mu \neq \nu) \end{aligned}$$

Here I_μ is the ionization energy for an electron in the valence AO φ_μ, $S_{\mu\nu}$ is the overlap integral $\int\varphi_\mu(\mathbf{r})\varphi_\nu(\mathbf{r})d\mathbf{r}$, and the factor 1.75 is a semi-empirical parameter. In contrast to the NDO procedures, EHM includes overlap integrals in the secular equations. As the EHM does not involve explicit electron interaction terms, it does not require an iterative procedure. It must be considered as a qualitative MO model, and it works best for non-polar molecules (e.g., hydrocarbons). On the other hand, it is simple and economic: it can be applied to very large molecules.

Computational Chemistry
Lecture Notes 3

(RUC, NSM, October 2012)

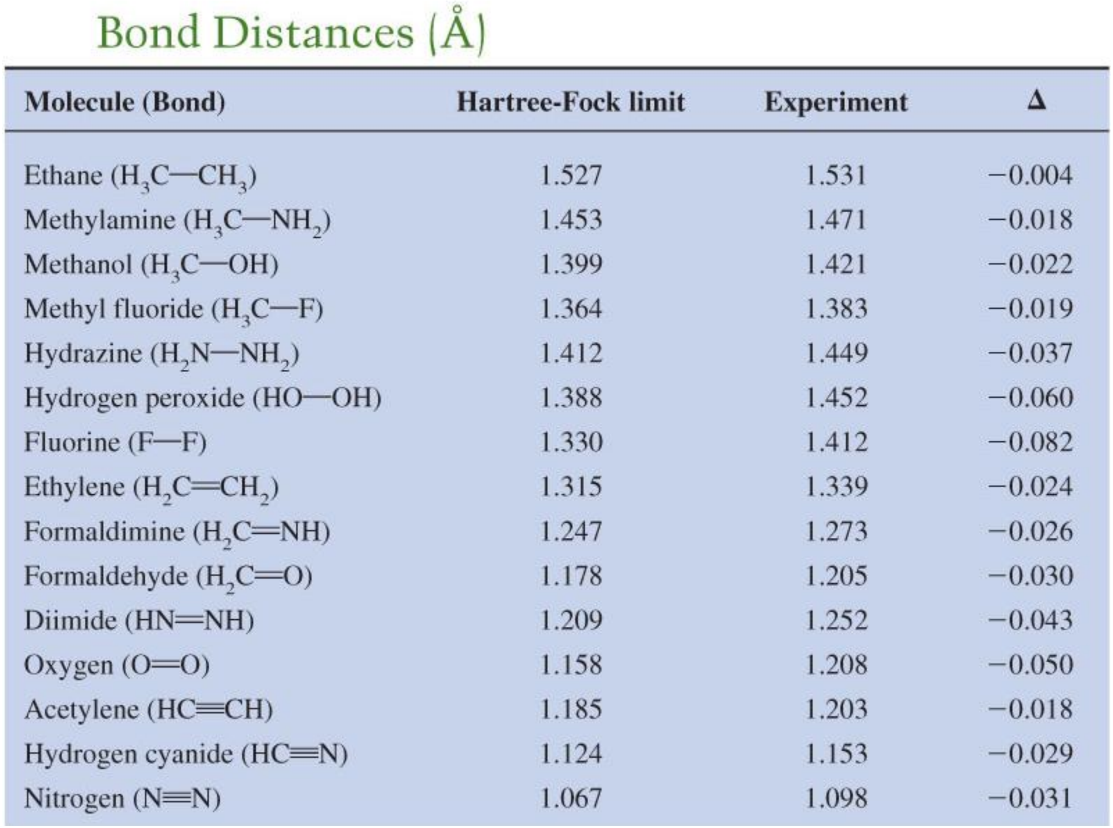
Limitations of the HF Procedure
Going Beyond the HF Model
- Configuration Interaction (CISD, CID)
- Møller-Plesset perturbation Theory (MP2)

Density Functional Theory (DFT)
Excited Electronic States
- Electronic transitions
- Configuration Interaction “Singles” (CIS)
- Time-Dependent (TD) DFT

Solvation Models
- Polarized Continuum Model (PCM)
- Mixed Discreet + Continuum Solvation Model

Limitations of the HF Procedure

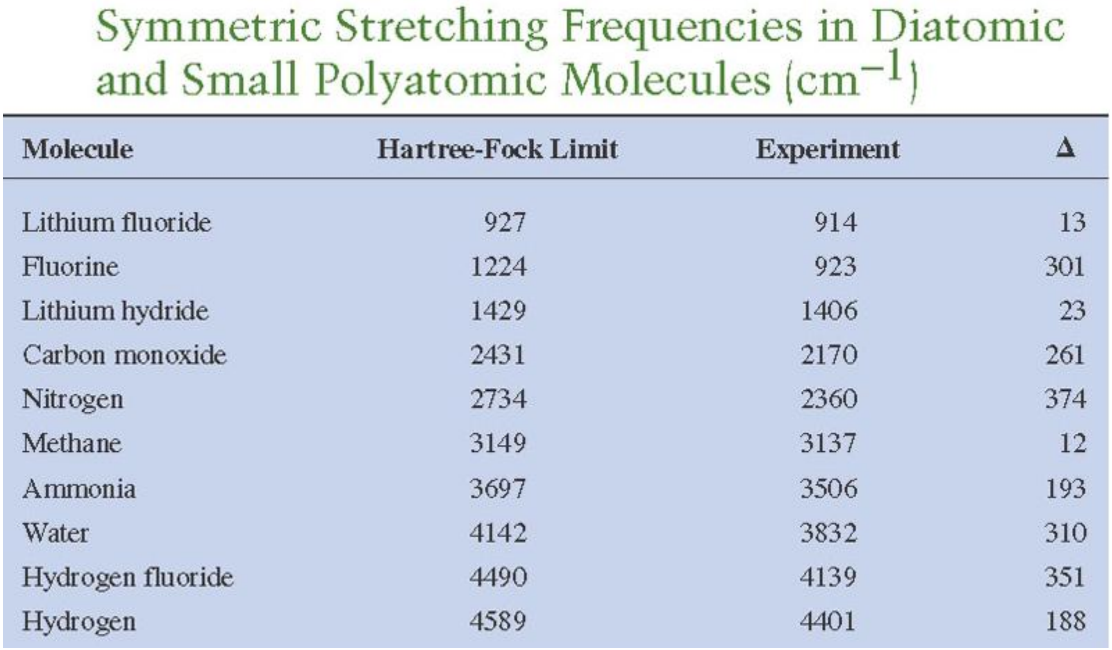
With few exceptions, equilibrium *bond lengths* are somewhat underestimated by the HF procedure. This can be explained by the fact that electron correlation effects are not accounted for in this model (see later).



Bond Distances (Å)

Molecule (Bond)	Hartree-Fock limit	Experiment	Δ
Ethane ($H_3C—CH_3$)	1.527	1.531	−0.004
Methylamine ($H_3C—NH_2$)	1.453	1.471	−0.018
Methanol ($H_3C—OH$)	1.399	1.421	−0.022
Methyl fluoride ($H_3C—F$)	1.364	1.383	−0.019
Hydrazine ($H_2N—NH_2$)	1.412	1.449	−0.037
Hydrogen peroxide (HO—OH)	1.388	1.452	−0.060
Fluorine (F—F)	1.330	1.412	−0.082
Ethylene ($H_2C{=}CH_2$)	1.315	1.339	−0.024
Formaldimine ($H_2C{=}NH$)	1.247	1.273	−0.026
Formaldehyde ($H_2C{=}O$)	1.178	1.205	−0.030
Diimide ($HN{=}NH$)	1.209	1.252	−0.043
Oxygen ($O{=}O$)	1.158	1.208	−0.050
Acetylene ($HC{\equiv}CH$)	1.185	1.203	−0.018
Hydrogen cyanide ($HC{\equiv}N$)	1.124	1.153	−0.029
Nitrogen ($N{\equiv}N$)	1.067	1.098	−0.031

Vibrational frequencies computed by the HF model within the harmonic approximation tend to be overestimated by roughly 10%. This is due in part to errors inherent in the HF model, and in part to limitations of the harmonic approximation. Neglect of electronic correlation effects influences the shape of the computed potential energy surface, leading in general to slightly too large force constants. Neglect of anharmonic effects further contributes to the sum of errors.



Symmetric Stretching Frequencies in Diatomic and Small Polyatomic Molecules (cm^{-1})

Molecule	Hartree-Fock Limit	Experiment	Δ
Lithium fluoride	927	914	13
Fluorine	1224	923	301
Lithium hydride	1429	1406	23
Carbon monoxide	2431	2170	261
Nitrogen	2734	2360	374
Methane	3149	3137	12
Ammonia	3697	3506	193
Water	4142	3832	310
Hydrogen fluoride	4490	4139	351
Hydrogen	4589	4401	188

It is customary to transform the computed vibrational wavenumbers into empirical scale by multiplication by a scale factor α. Scott and Radom have evaluated the effective scaling constants and resulting rms errors for a number of procedures, based on the results for a test set comprising 1066 experimental wavenumbers. For HF calculations with a variety of basis sets, α is around 0.9, with rms errors close to 50 cm^{-1}:

Anthony P. Scott and Leo Radom*, *J. Phys. Chem.* **100** (1996)16502-16513:
“Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Møller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors”

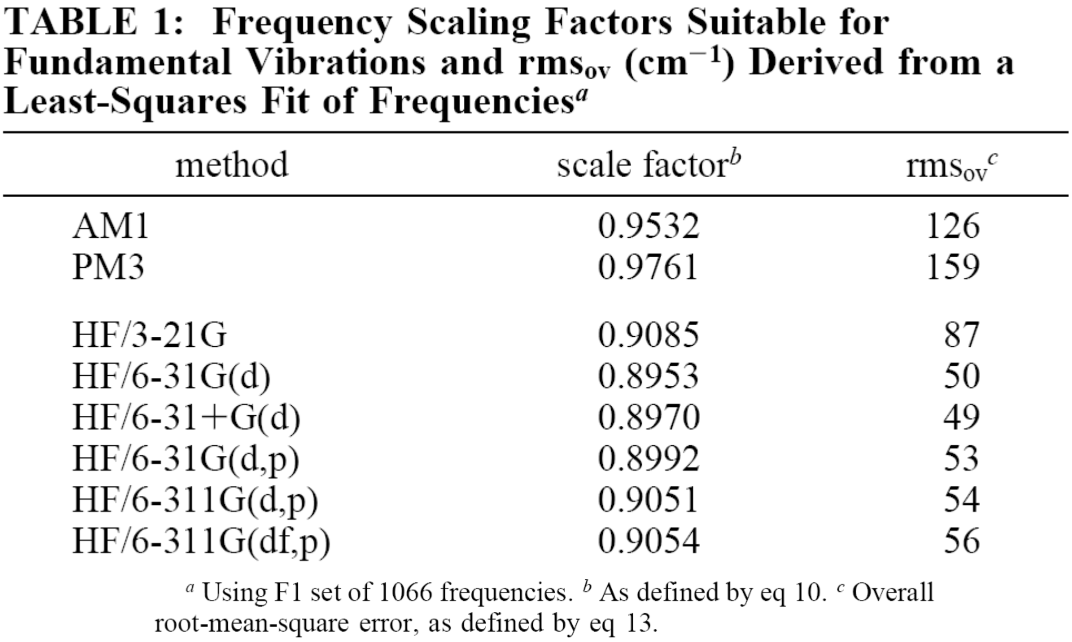
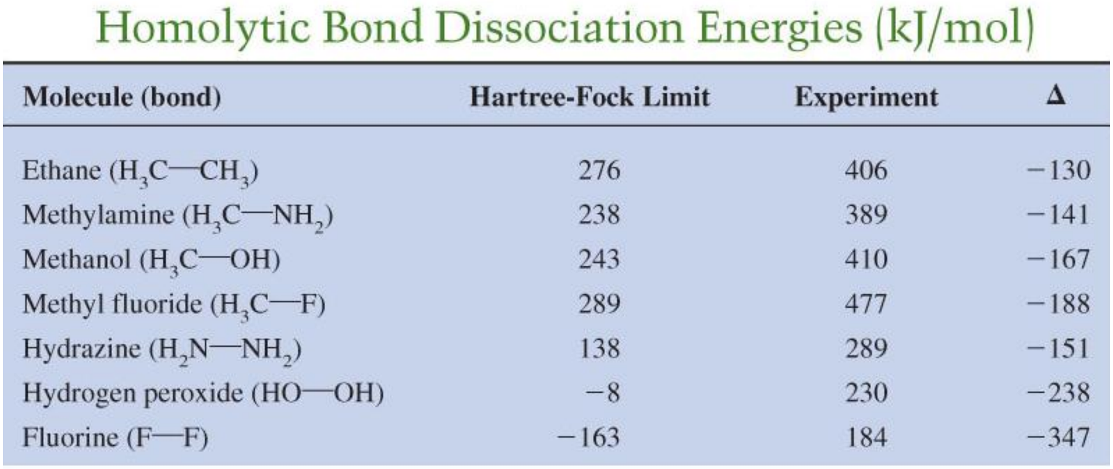


TABLE 1: Frequency Scaling Factors Suitable for Fundamental Vibrations and $\mathrm{rms}_{\mathrm{ov}}$ (cm^{-1}) Derived from a Least-Squares Fit of Frequencies[a]

method	scale factor[b]	$\mathrm{rms}_{\mathrm{ov}}$[c]
AM1	0.9532	126
PM3	0.9761	159
HF/3-21G	0.9085	87
HF/6-31G(d)	0.8953	50
HF/6-31+G(d)	0.8970	49
HF/6-31G(d,p)	0.8992	53
HF/6-311G(d,p)	0.9051	54
HF/6-311G(df,p)	0.9054	56

[a] Using F1 set of 1066 frequencies. [b] As defined by eq 10. [c] Overall root-mean-square error, as defined by eq 13.

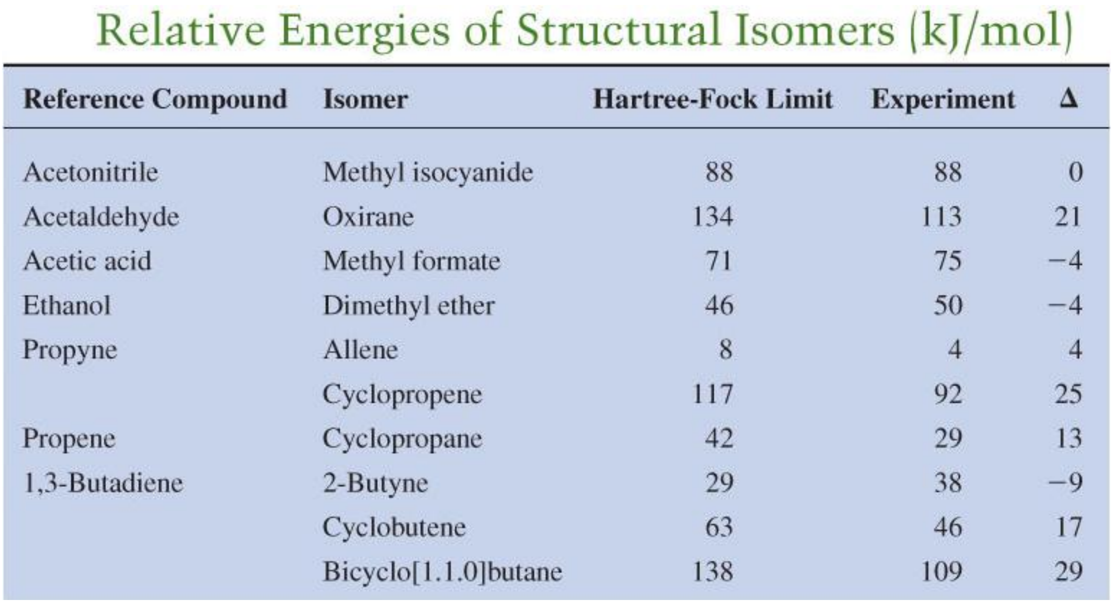
In general, *homolytic bond dissociation energies* are significantly underestimated by the HF procedure. This can be explained by noticing that the products have one electron pair less than the reactants, leading to larger HF error for the latter.



Homolytic Bond Dissociation Energies (kJ/mol)

Molecule (bond)	Hartree-Fock Limit	Experiment	Δ
Ethane ($H_3C—CH_3$)	276	406	−130
Methylamine ($H_3C—NH_2$)	238	389	−141
Methanol ($H_3C—OH$)	243	410	−167
Methyl fluoride ($H_3C—F$)	289	477	−188
Hydrazine ($H_2N—NH_2$)	138	289	−151
Hydrogen peroxide (HO—OH)	−8	230	−238
Fluorine (F—F)	−163	184	−347

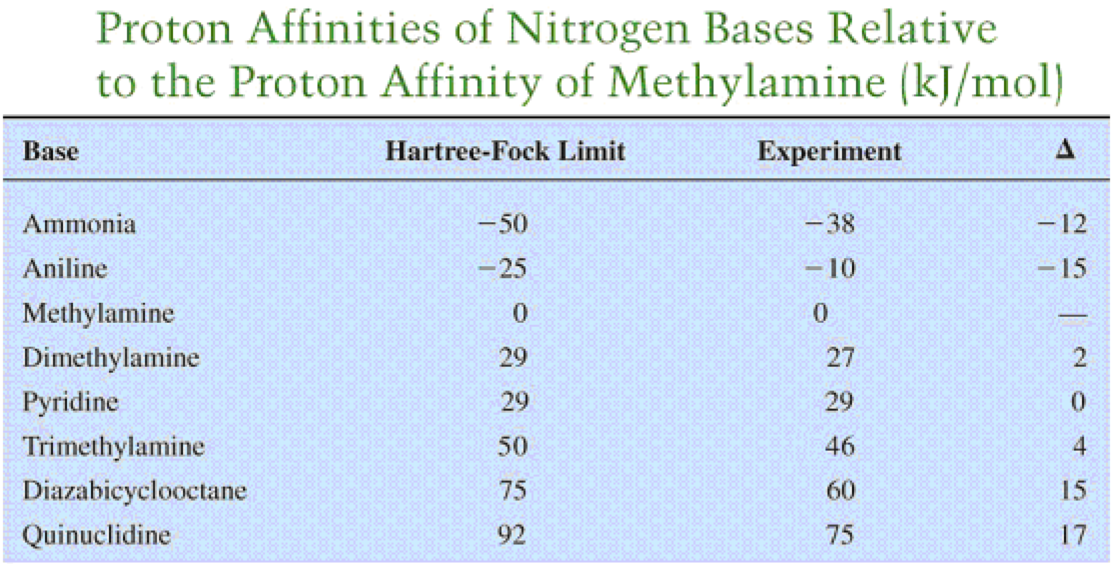
Relative energies of structural isomers are predicted with errors that are an order of magnitude lower than those obtained for homolytic bond dissociation energies. This is due in part to cancelation of errors.



Relative Energies of Structural Isomers (kJ/mol)

Reference Compound	Isomer	Hartree-Fock Limit	Experiment	Δ
Acetonitrile	Methyl isocyanide	88	88	0
Acetaldehyde	Oxirane	134	113	21
Acetic acid	Methyl formate	71	75	−4
Ethanol	Dimethyl ether	46	50	−4
Propyne	Allene	8	4	4
	Cyclopropene	117	92	25
Propene	Cyclopropane	42	29	13
1,3-Butadiene	2-Butyne	29	38	−9
	Cyclobutene	63	46	17
	Bicyclo[1.1.0]butane	138	109	29

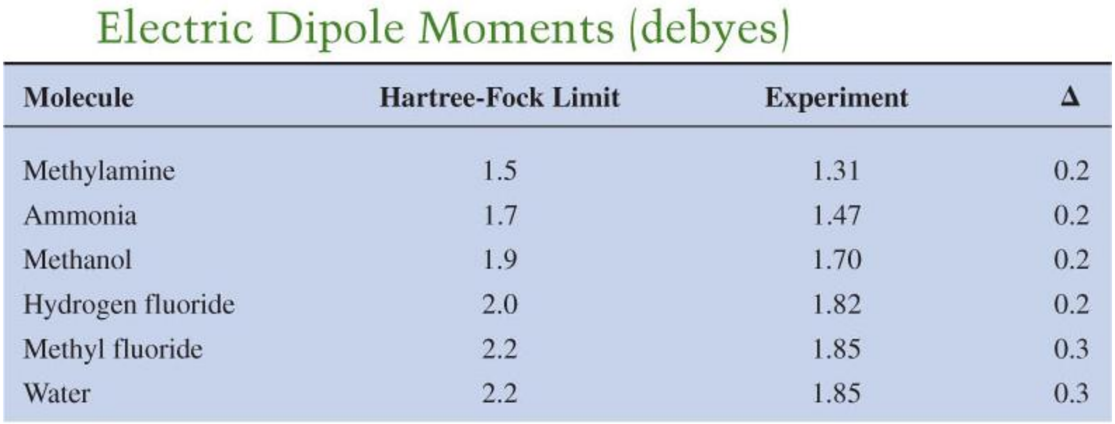
Also the *relative proton affinities of nitrogen bases* are quite well predicted:



Proton Affinities of Nitrogen Bases Relative to the Proton Affinity of Methylamine (kJ/mol)

Base	Hartree-Fock Limit	Experiment	Δ
Ammonia	−50	−38	−12
Aniline	−25	−10	−15
Methylamine	0	0	—
Dimethylamine	29	27	2
Pyridine	29	29	0
Trimethylamine	50	46	4
Diazabicyclooctane	75	60	15
Quinuclidine	92	75	17

In general, *electric dipole moments* are reasonably well predicted by HF calculations (being slightly too large), indicating that the computed charge distributions are realistic:



Electric Dipole Moments (debyes)

Molecule	Hartree-Fock Limit	Experiment	Δ
Methylamine	1.5	1.31	0.2
Ammonia	1.7	1.47	0.2
Methanol	1.9	1.70	0.2
Hydrogen fluoride	2.0	1.82	0.2
Methyl fluoride	2.2	1.85	0.3
Water	2.2	1.85	0.3

Post HF procedures

The most important so-called post HF procedures are the configuration interaction (CI) model and Møller-Plesset perturbation theory.

Configuration Interaction (CI)

In this procedure, the electronic wavefunction is described by using contributions from several configurations (not by a single configurational wavefunction like in HF):

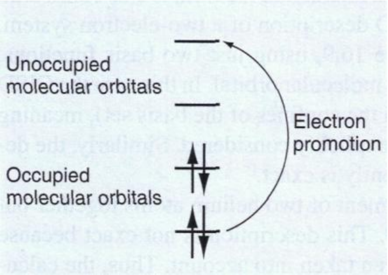
$$\Psi = a_0\Psi_0 + \sum_{s>0} a_s\Psi_s$$

In principle, this leads to a CI secular problem within matrix elements of the many-electron Hamilton operator:

$$H_{st} = \int \ldots \int \Psi_s \hat{H} \Psi_t d\tau_1 d\tau_2 \ldots d\tau_n$$

The lowest energy obtained from solution of the CI secular problem corresponds to the energy of the electronic ground state. Limiting the configurations to be considered in the CI calculation to single-electron promotions only (the CIS procedure) leads to no improvement of the HF ground state. In order to introduce ground-state correlation effects, double-electron or multiple-electron promotions must be considered. In the CID procedure, only double-electron promotions are included:

$$\Psi_{CID} = a_0\Psi_0 + \overset{\text{molecular orbitals}}{\overset{occ}{\sum_{i<j}\sum} \quad \overset{unocc}{\sum_{a<b}\sum}} a_{ij}^{ab}\Psi_{ij}^{ab}$$



Typically, this amounts to admixture of contributions from unoccupied MOs with anti-bonding character into the ground state wavefunction, leading to prediction of weaker and longer bonds, thereby obtaining improved agreement with experimental data. More extensive CI schemes include additional promotions, such as CISD which considers single- as well as double-electron promotions, and CISDT which includes also triple-electron promotions. In principle, extending the CI expansion to include all promotions, correponding to a *full CI*, leads to an exact solution of the electronic eigenvalue problem, with full account of electron correlation (provided the basis set is complete!). However, the CI expansion converges very slowly, and the full CI solution can been approached only for small systems.

Møller-Plesset

theory is based on a mathematical approximation formalism called perturbation theory. The exact Hamiltonian is written as the sum of the HF Hamiltonian and a small perturbation representing the correlation energy. In the so-called *second order* Møller-Plesset model (MP2), the approximate correlation energy correction to the HF ground state energy is obtained as

$$E^{(2)} = \overset{\text{molecular orbitals}}{\overset{occ}{\sum_{i<j}\sum} \quad \overset{unocc}{\sum_{a<b}\sum}} \frac{\left[(ij\|ab)\right]^2}{(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j)}$$

where

$$(ij\|ab) = (ij|ab) - (ib|ja),$$

$$(ij|ab) = \int \psi_i(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\left[\frac{1}{r_{12}}\right]\psi_j(\mathbf{r}_2)\psi_b(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

The integrals ($(ij|ab)$) involve orbitals over occupied and un-occupied MOs, and ε_i, ε_a, etc. are the corresponding MO energies. – It is apparent that the presence of low-lying un-occupied MOs is likely to produce large contributions to $E^{(2)}$.

Density Functional Theory (DFT)

The starting point of this approach is the electron density, $\rho(\mathbf{r})$, rather than the wavefunction Ψ. It has been shown by Hohenberg and Kohn that the total electronic energy E is uniquely determined by the electron density $\rho(\mathbf{r})$:

$$E = E\,[\rho(\mathbf{r})]$$

The electron density is a function of the coordinates $\mathbf{r}$, defined such that $\rho(\mathbf{r})d\mathbf{r}$ is the number of electrons inside a small volume of space $d\mathbf{r}$ (this is what is measured in X-ray diffraction experiments). $\rho(\mathbf{r})$ is a function of only three coordinates, namely the three coordinates of space; in this sense, $\rho(\mathbf{r})$ is much simpler than the wavefunction Ψ which is a function of the coordinates of all the electrons in the system. In current DFT procedures, $\rho(\mathbf{r})$ is expressed in terms of the so-called Kohn-Sham MOs ψ_i, which enables a simple expression for the density:

$$\rho(\mathbf{r}) = 2\sum_{i}^{orbitals} \left|\psi_i(\mathbf{r})\right|^2$$

E is thus a function of $\rho(\mathbf{r})$, which is a function of $\mathbf{r}$. A 'function of a function' is called a *functional*. The problem is that the exact form of the energy functional is not known. The approximate functional is generally written as a sum of several contributions:

$$E = E_T + E_V + E_J + E_{XC}$$

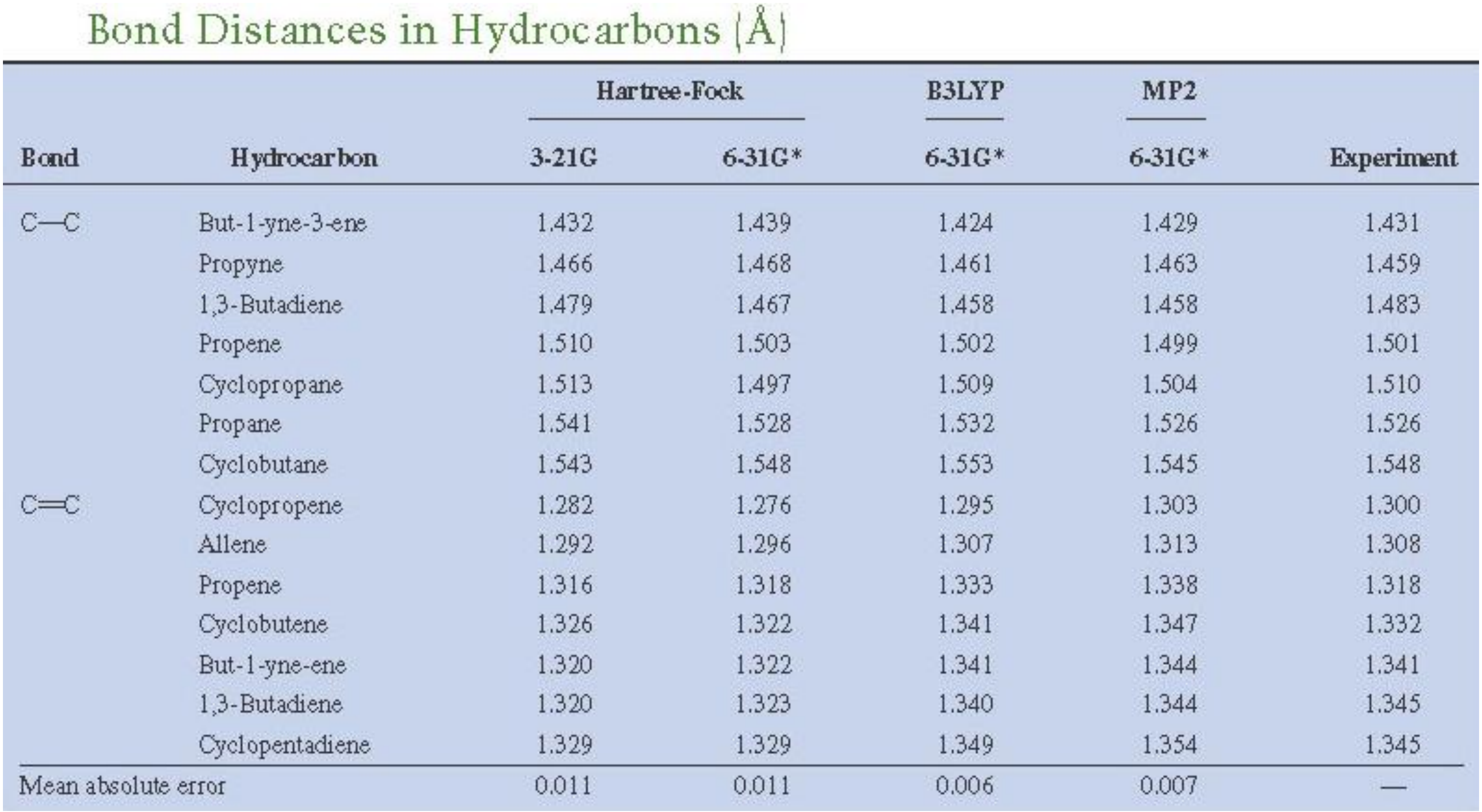
where E_T is the kinetic energy term, E_V the nuclear attraction term, E_J the electron repulsion term (also called the Coulomb term), and E_{XC} is the so-called *exchange/correlation term*. The terms E_T, E_V, and E_J are carried over from HF theory, but the exchange/correlation term E_{XC} accounts for the remaining terms in the energy:

1. The exchange energy arising from the quantum mechanical anti-symmetry requirement.

2. Dynamic correlation in the motions of the individual electrons.

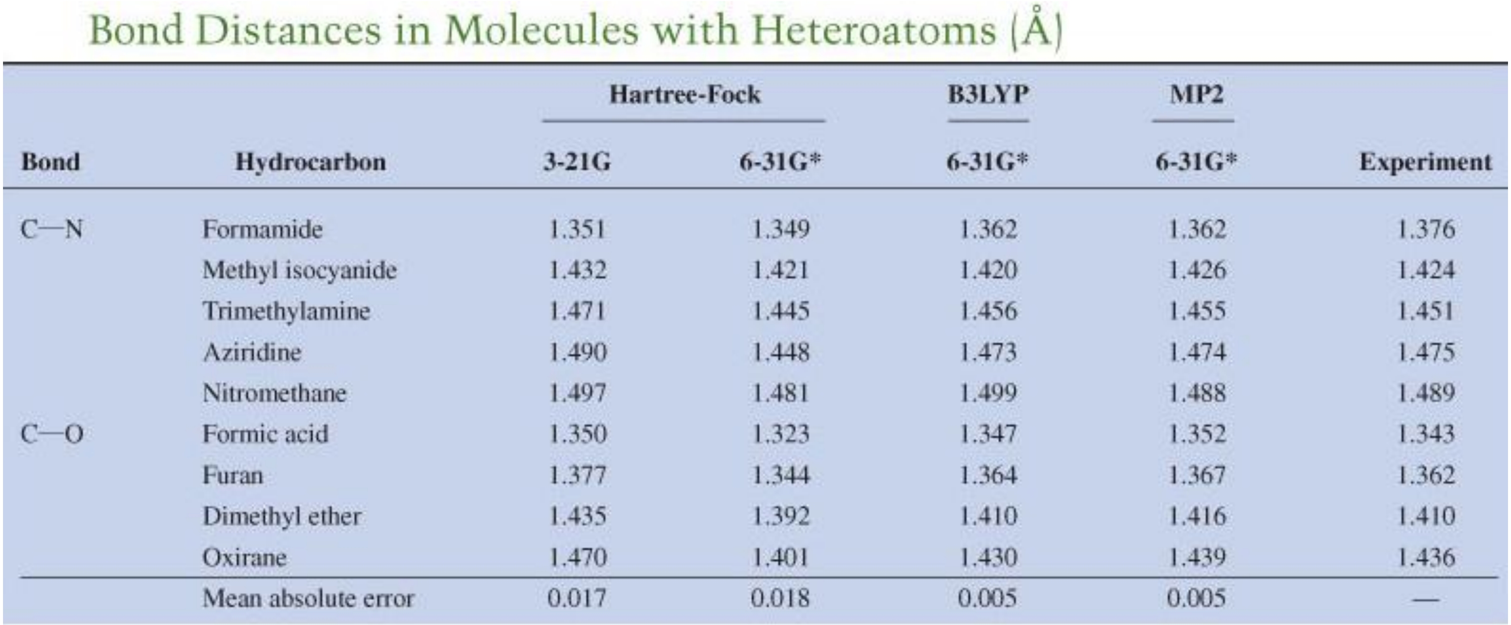
Many varieties of approximate DFT procedures have been developed, involving different models of the exchange/correlation functional $E_{XC}[\rho(\mathbf{r})]$. This functional is usually divided into two parts, referred to as the *exchange* and *correlation* functionals, $E_{XC}[\rho(\mathbf{r})] = E_X[\rho(\mathbf{r})] + E_C[\rho(\mathbf{r})]$. So-called *local* model functionals depend only on the electron density $\rho(\mathbf{r})$, while *gradient corrected* functionals depend also on the gradient of $\rho(\mathbf{r})$. Some DFT procedures involve *hybrid exchange functionals*, which include a mixture of HF-type exhange into the DFT exchange functional E_{XC}. A popular hybrid DFT procedure is B3LYP, which is based on Becke's three-parameter hybrid exchange functional and the correlation functional of Lee, Yang, and Parr. – In actual practice, self-consistent Kohn-Sham DFT calculations are carried out in an iterative manner very similar to the HF SCF procedure, and the computing time is only slightly larger than that of a HF calculation. DFT calculations thus tend to much faster than CI and Møller-Plesset calculations, and they are frequently at least as accurate. DFT procedures have thus gained considerable popularity in recent years.

Comparisons of the results of HF, B3LYP, and MP2 calculations:



Bond Distances in Hydrocarbons (Å)

		Hartree-Fock		B3LYP	MP2	
Bond	**Hydrocarbon**	**3-21G**	**6-31G***	**6-31G***	**6-31G***	**Experiment**
C—C	But-1-yne-3-ene	1.432	1.439	1.424	1.429	1.431
	Propyne	1.466	1.468	1.461	1.463	1.459
	1,3-Butadiene	1.479	1.467	1.458	1.458	1.483
	Propene	1.510	1.503	1.502	1.499	1.501
	Cyclopropane	1.513	1.497	1.509	1.504	1.510
	Propane	1.541	1.528	1.532	1.526	1.526
	Cyclobutane	1.543	1.548	1.553	1.545	1.548
C═C	Cyclopropene	1.282	1.276	1.295	1.303	1.300
	Allene	1.292	1.296	1.307	1.313	1.308
	Propene	1.316	1.318	1.333	1.338	1.318
	Cyclobutene	1.326	1.322	1.341	1.347	1.332
	But-1-yne-ene	1.320	1.322	1.341	1.344	1.341
	1,3-Butadiene	1.320	1.323	1.340	1.344	1.345
	Cyclopentadiene	1.329	1.329	1.349	1.354	1.345
Mean absolute error		0.011	0.011	0.006	0.007	—



Bond Distances in Molecules with Heteroatoms (Å)

		Hartree-Fock		B3LYP	MP2	
Bond	**Hydrocarbon**	**3-21G**	**6-31G***	**6-31G***	**6-31G***	**Experiment**
C—N	Formamide	1.351	1.349	1.362	1.362	1.376
	Methyl isocyanide	1.432	1.421	1.420	1.426	1.424
	Trimethylamine	1.471	1.445	1.456	1.455	1.451
	Aziridine	1.490	1.448	1.473	1.474	1.475
	Nitromethane	1.497	1.481	1.499	1.488	1.489
C—O	Formic acid	1.350	1.323	1.347	1.352	1.343
	Furan	1.377	1.344	1.364	1.367	1.362
	Dimethyl ether	1.435	1.392	1.410	1.416	1.410
	Oxirane	1.470	1.401	1.430	1.439	1.436
	Mean absolute error	0.017	0.018	0.005	0.005	—

Scott and Radom (loc. cit.):

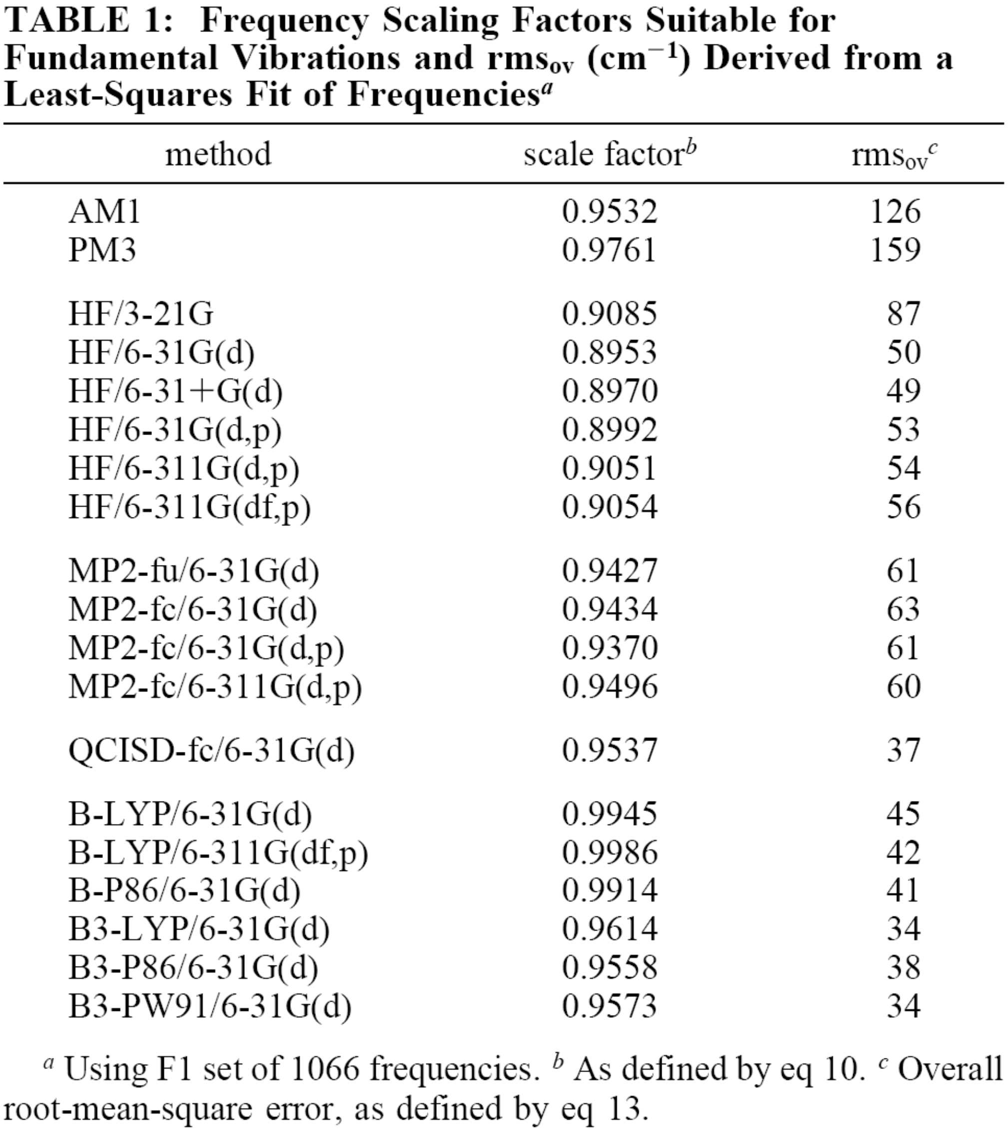
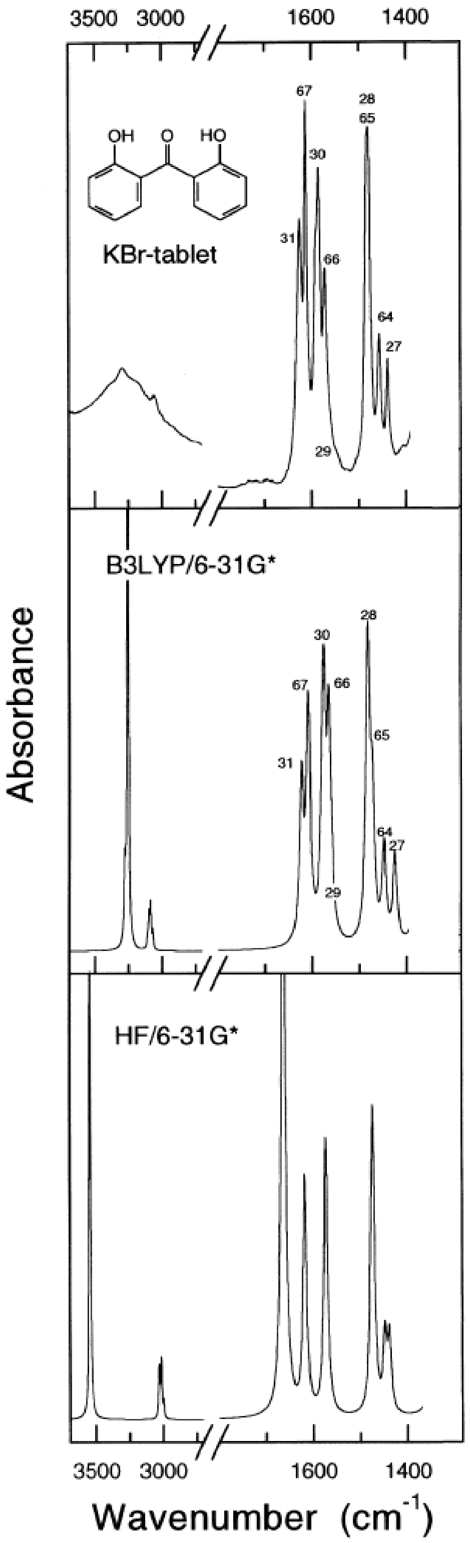


TABLE 1: Frequency Scaling Factors Suitable for Fundamental Vibrations and $\mathrm{rms}_{\mathrm{ov}}$ (cm^{-1}) Derived from a Least-Squares Fit of Frequencies[a]

method	scale factor[b]	$\mathrm{rms}_{\mathrm{ov}}$[c]
AM1	0.9532	126
PM3	0.9761	159
HF/3-21G	0.9085	87
HF/6-31G(d)	0.8953	50
HF/6-31+G(d)	0.8970	49
HF/6-31G(d,p)	0.8992	53
HF/6-311G(d,p)	0.9051	54
HF/6-311G(df,p)	0.9054	56
MP2-fu/6-31G(d)	0.9427	61
MP2-fc/6-31G(d)	0.9434	63
MP2-fc/6-31G(d,p)	0.9370	61
MP2-fc/6-311G(d,p)	0.9496	60
QCISD-fc/6-31G(d)	0.9537	37
B-LYP/6-31G(d)	0.9945	45
B-LYP/6-311G(df,p)	0.9986	42
B-P86/6-31G(d)	0.9914	41
B3-LYP/6-31G(d)	0.9614	34
B3-P86/6-31G(d)	0.9558	38
B3-PW91/6-31G(d)	0.9573	34

[a] Using F1 set of 1066 frequencies. [b] As defined by eq 10. [c] Overall root-mean-square error, as defined by eq 13.

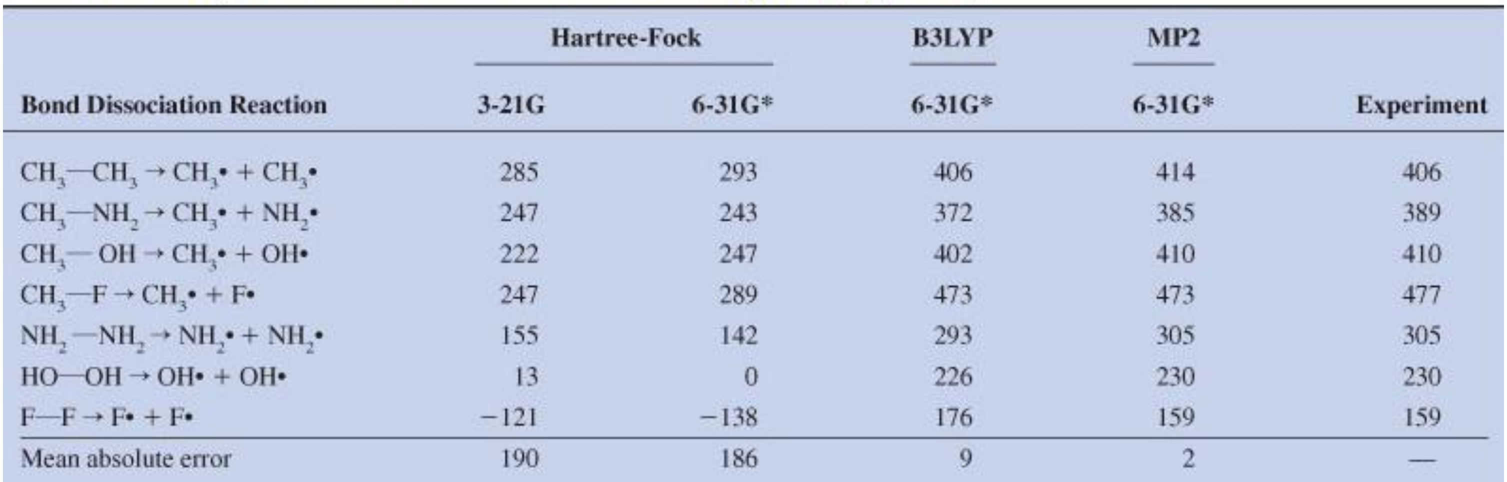


K.B. Andersen, M. Langgård[1], J. Spanget-Larsen*

Journal of Molecular Structure 509 (1999) 153–163

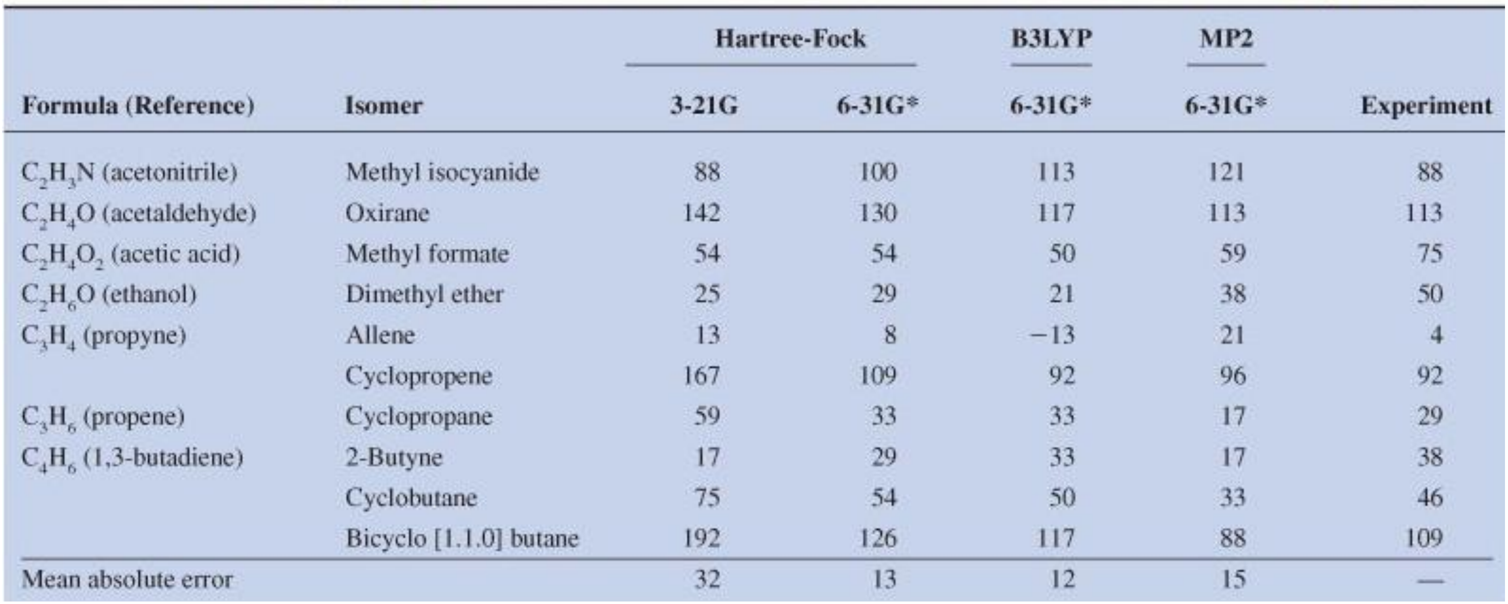
Molecular and vibrational structure of 2,2′-dihydroxybenzophenone: infrared linear dichroism and quantum chemical calculations☆

Homolytic Bond Dissociation Energies (kJ/mol)



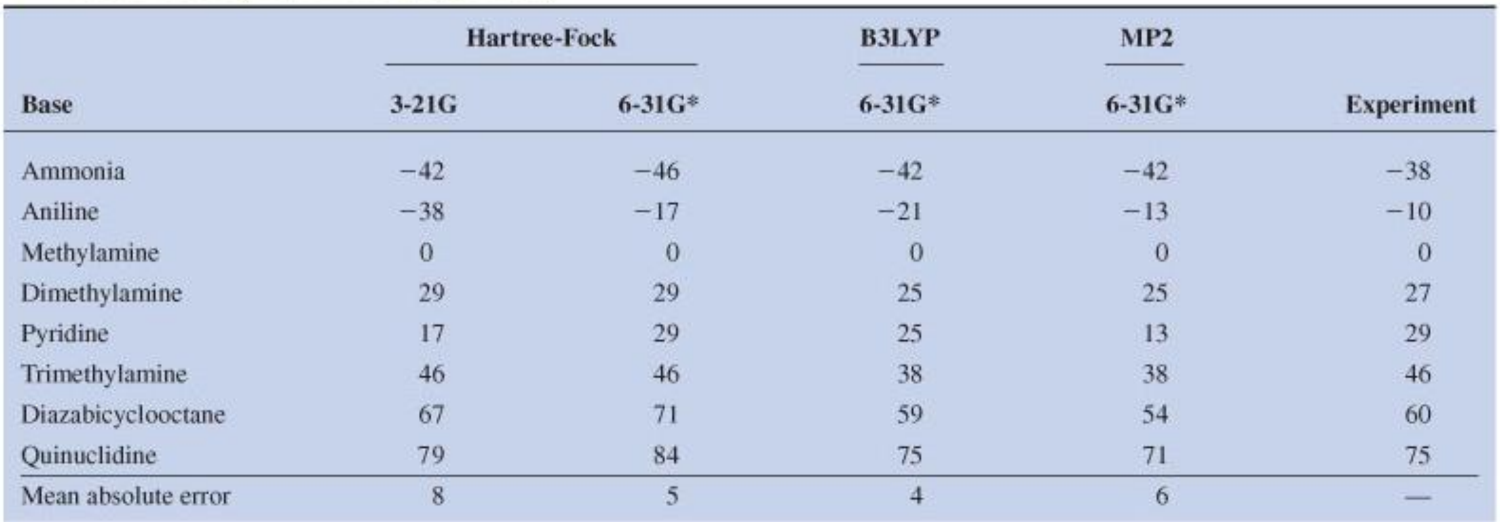
Bond Dissociation Reaction	Hartree-Fock 3-21G	Hartree-Fock 6-31G*	B3LYP 6-31G*	MP2 6-31G*	Experiment
$CH_3—CH_3 \rightarrow CH_3\bullet + CH_3\bullet$	285	293	406	414	406
$CH_3—NH_2 \rightarrow CH_3\bullet + NH_2\bullet$	247	243	372	385	389
$CH_3—OH \rightarrow CH_3\bullet + OH\bullet$	222	247	402	410	410
$CH_3—F \rightarrow CH_3\bullet + F\bullet$	247	289	473	473	477
$NH_2—NH_2 \rightarrow NH_2\bullet + NH_2\bullet$	155	142	293	305	305
$HO—OH \rightarrow OH\bullet + OH\bullet$	13	0	226	230	230
$F—F \rightarrow F\bullet + F\bullet$	−121	−138	176	159	159
Mean absolute error	190	186	9	2	—

Relative Energies (Column 1 − Column 2) of Structural Isomers (kJ/mol)



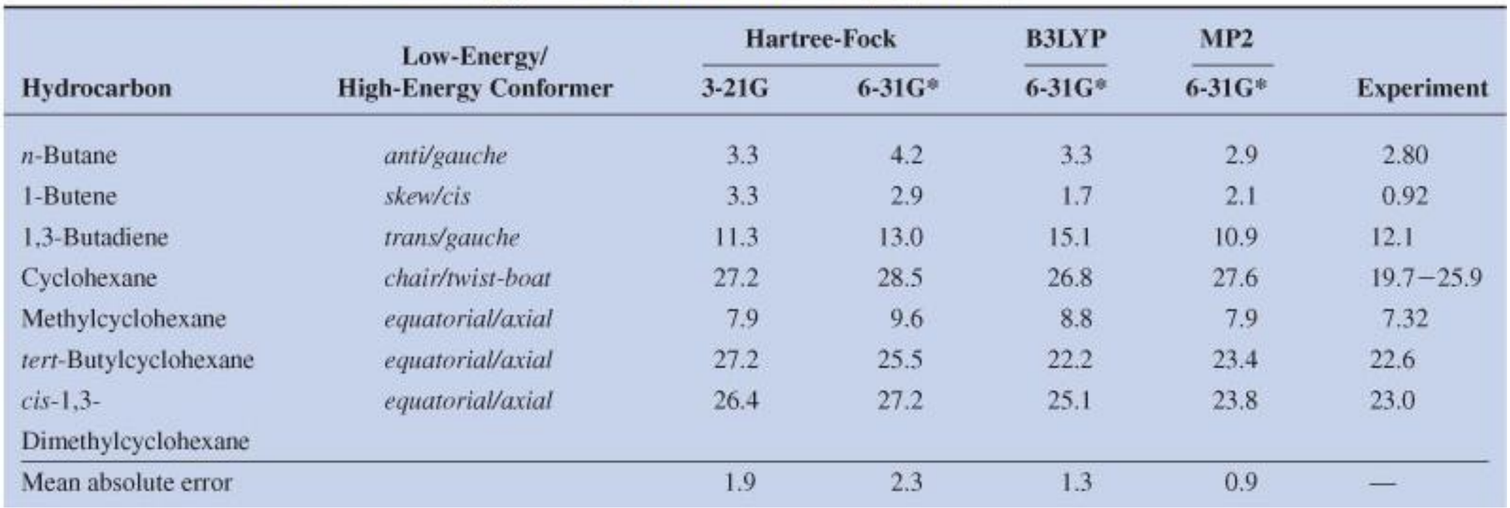
Formula (Reference)	Isomer	Hartree-Fock 3-21G	Hartree-Fock 6-31G*	B3LYP 6-31G*	MP2 6-31G*	Experiment
C_2H_3N (acetonitrile)	Methyl isocyanide	88	100	113	121	88
C_2H_4O (acetaldehyde)	Oxirane	142	130	117	113	113
$C_2H_4O_2$ (acetic acid)	Methyl formate	54	54	50	59	75
C_2H_6O (ethanol)	Dimethyl ether	25	29	21	38	50
C_3H_4 (propyne)	Allene	13	8	−13	21	4
	Cyclopropene	167	109	92	96	92
C_3H_6 (propene)	Cyclopropane	59	33	33	17	29
C_4H_6 (1,3-butadiene)	2-Butyne	17	29	33	17	38
	Cyclobutane	75	54	50	33	46
	Bicyclo [1.1.0] butane	192	126	117	88	109
Mean absolute error		32	13	12	15	—

Proton Affinities of Nitrogen Bases Relative to the Proton Affinity of Methylamine (kJ/mol)

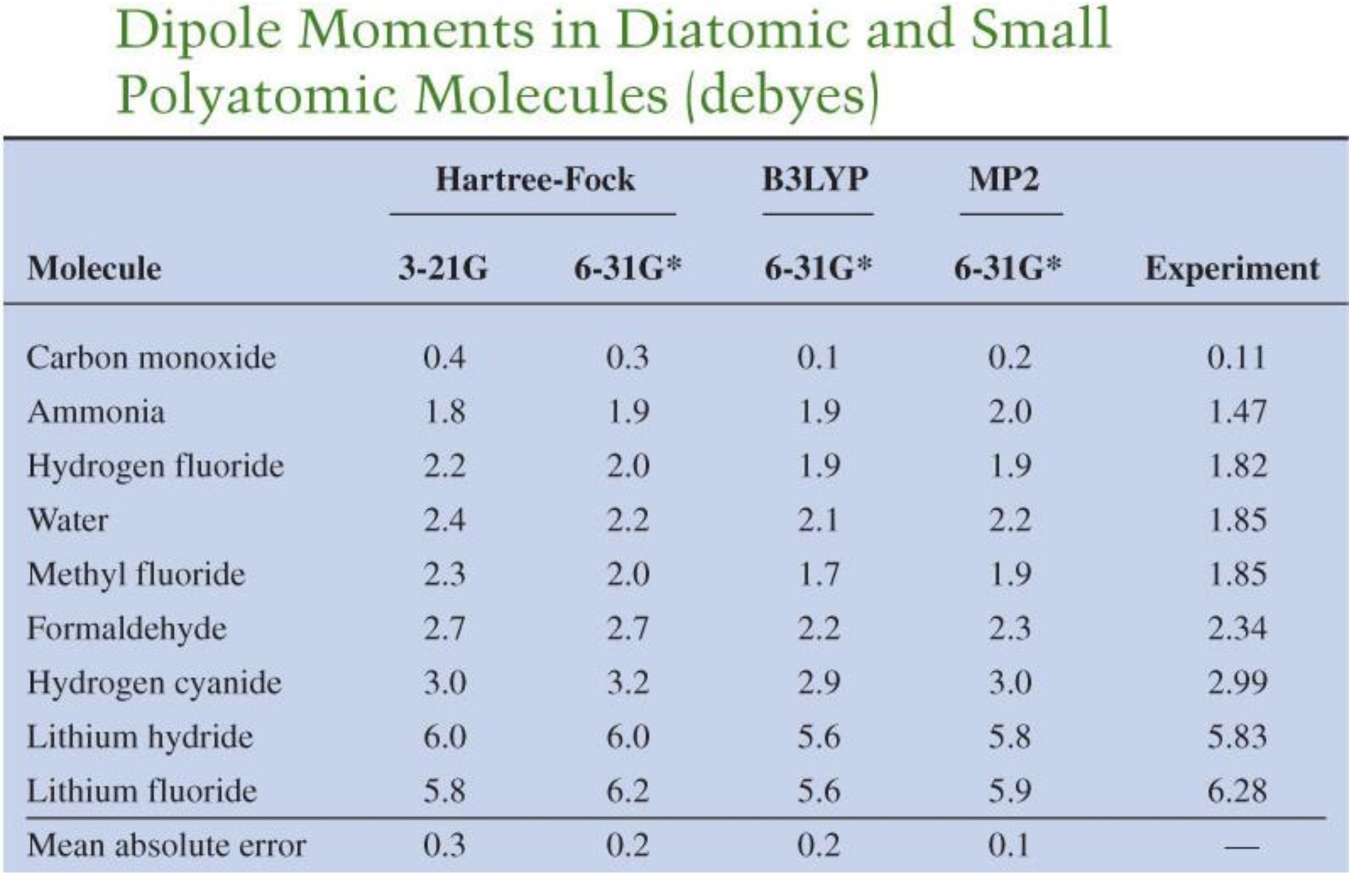


Base	Hartree-Fock 3-21G	Hartree-Fock 6-31G*	B3LYP 6-31G*	MP2 6-31G*	Experiment
Ammonia	−42	−46	−42	−42	−38
Aniline	−38	−17	−21	−13	−10
Methylamine	0	0	0	0	0
Dimethylamine	29	29	25	25	27
Pyridine	17	29	25	13	29
Trimethylamine	46	46	38	38	46
Diazabicyclooctane	67	71	59	54	60
Quinuclidine	79	84	75	71	75
Mean absolute error	8	5	4	6	—

Conformational Energy in Hydrocarbons (kJ/mol)



Hydrocarbon	Low-Energy/ High-Energy Conformer	Hartree-Fock 3-21G	Hartree-Fock 6-31G*	B3LYP 6-31G*	MP2 6-31G*	Experiment
n-Butane	*anti/gauche*	3.3	4.2	3.3	2.9	2.80
1-Butene	*skew/cis*	3.3	2.9	1.7	2.1	0.92
1,3-Butadiene	*trans/gauche*	11.3	13.0	15.1	10.9	12.1
Cyclohexane	*chair/twist-boat*	27.2	28.5	26.8	27.6	19.7−25.9
Methylcyclohexane	*equatorial/axial*	7.9	9.6	8.8	7.9	7.32
tert-Butylcyclohexane	*equatorial/axial*	27.2	25.5	22.2	23.4	22.6
cis-1,3-Dimethylcyclohexane	*equatorial/axial*	26.4	27.2	25.1	23.8	23.0
Mean absolute error		1.9	2.3	1.3	0.9	—



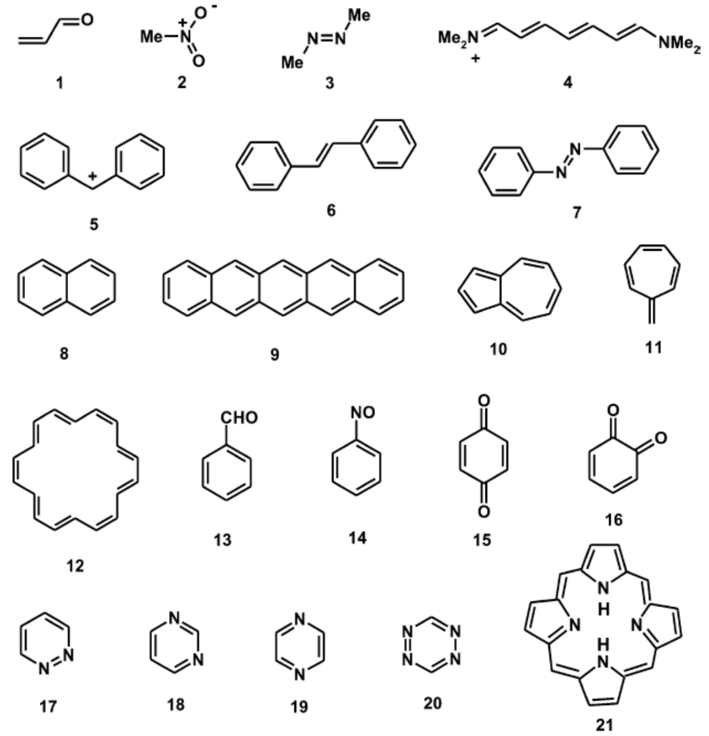
Dipole Moments in Diatomic and Small Polyatomic Molecules (debyes)

	Hartree-Fock		B3LYP	MP2	
Molecule	**3-21G**	**6-31G***	**6-31G***	**6-31G***	**Experiment**
Carbon monoxide	0.4	0.3	0.1	0.2	0.11
Ammonia	1.8	1.9	1.9	2.0	1.47
Hydrogen fluoride	2.2	2.0	1.9	1.9	1.82
Water	2.4	2.2	2.1	2.2	1.85
Methyl fluoride	2.3	2.0	1.7	1.9	1.85
Formaldehyde	2.7	2.7	2.2	2.3	2.34
Hydrogen cyanide	3.0	3.2	2.9	3.0	2.99
Lithium hydride	6.0	6.0	5.6	5.8	5.83
Lithium fluoride	5.8	6.2	5.6	5.9	6.28
Mean absolute error	0.3	0.2	0.2	0.1	—

Excited Electronic States
The prediction of the properties of excited electronic states is one of the most difficult tasks within quantum chemistry. HF CI calculations with inclusion only of single-electron promotions (CIS) tend to overestimate the transition energies seriously, because of different correlation effects for ground and excited states. A number of semi-empirical procedures are available, generally based on the CIS formalism, e.g., INDO/S and ZINDO/S. These methods are parameterized in order to predict transition energies of the right order of magnitude, but they should be used with caution. The so-called Time Dependent (TD) DFT procedure is very promising, and has performed remarkable well for a number of chromophores, but it has been reported to be problematic for transitions with a electron-transfer character. Fabian et al. have published a comparison of the performance of current calculational procedures to predict the UV-VIS absorbance spectra for a variety of organic chromophores:

J. Fabian et al.: *J. Mol. Struct. (Theochem)* **594** (2002) 41–53.
“Calculation of excitation energies of organic chromophores: A critical evaluation”

Sulfur-free chromophores:



Sulfur-containing chromophores:

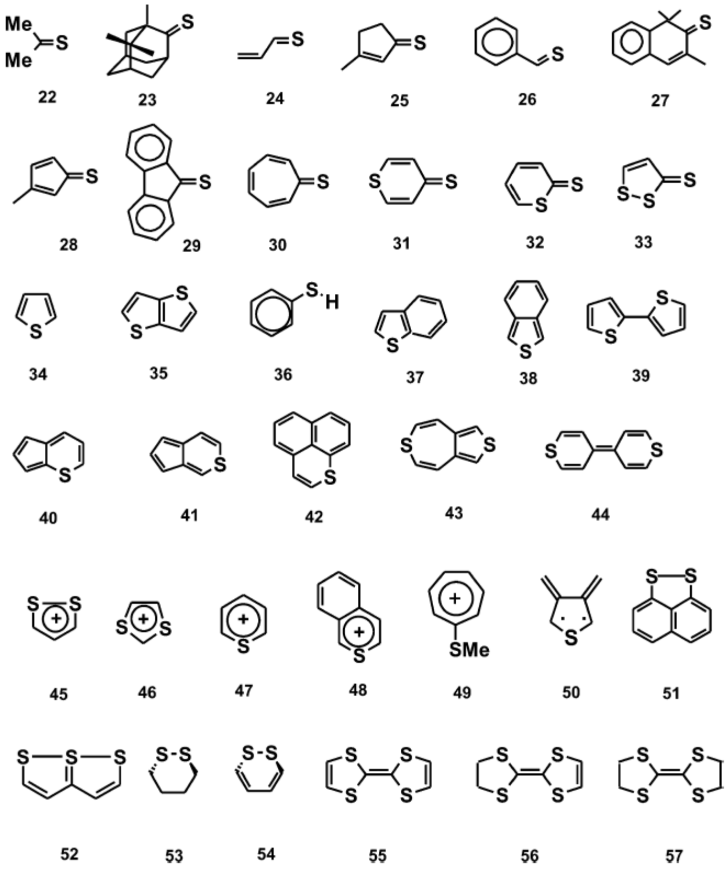


Table 5
Mean absolute deviations between theoretical and experimental excitation energies in eV

Response theory approaches			Semiempirical approximations	
Transition	TD-DFRT	TD-DFRT-TB	ZINDO/S	PM3/S
Sulfur-free compounds (**1**–**21**)[a]				
Total	0.27	0.36	0.31	0.40
$n \rightarrow \pi^*$	0.20	0.41	0.51	0.31
$\pi \rightarrow \pi^*$	0.29	0.34	0.25	0.42
Sulfur-containing compounds (**22**–**50**)[b]				
Total	0.22 (0.25)		0.33	0.25
$n \rightarrow \pi^*$	0.09 (0.09)		0.09	0.21
$\pi \rightarrow \pi^*$	0.24 (0.29)		0.38	0.26

Single-point calculations based on DFT B3-LYP/6-31 + G* optimum geometries.
(1 electron volt (eV) ≙ 8066 cm^{-1})

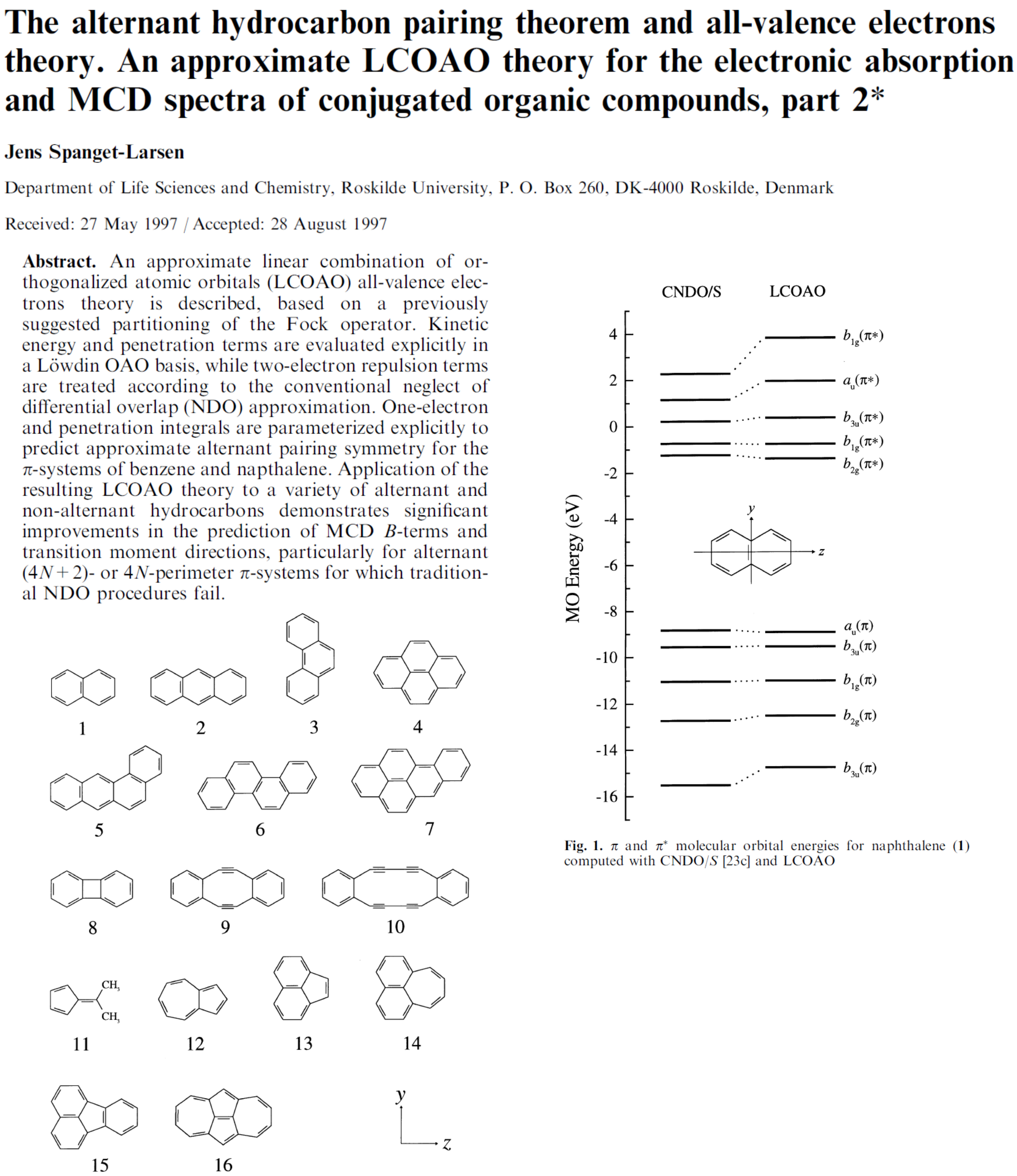
[a] 54 Electronic transitions altogether, 12 of the $n \rightarrow \pi^*$ and 42 of the $\pi \rightarrow \pi^*$-type.

[b] 70 Electronic transitions altogether, 12 of the $n \rightarrow \pi^*$ and 58 of the $\pi \rightarrow \pi^*$-type; Mean absolute deviations of TD-DFRT/B3-LYP/3-21 + G* results are given in parentheses. A change in the selection of the compounds explains the slight difference between the total deviations calculated in this study and the previous ones [27] (TD-DFRT B3-LYP/6-31 + G*: 0.21 and ZINDO/S: 0.35). The full list of data considered in the statistics is available on request.

4. Conclusions

According to the statistics TD-DFRT proved to be superior over semiempirical methods. The method allows a reasonable calculation of $n \rightarrow \pi^*$-, $n \rightarrow \sigma^*$-, $\pi \rightarrow \sigma^*$- and $\pi \rightarrow \pi^*$ -transitions which determine the spectral absorptions in the ultraviolet and visible region. This success, however, is paid by a higher demand in computer time. Unexpectedly, the mean average deviation is distinctly larger for sulfur-free than for sulfur-containing compounds. Poor results were obtained for heteroatom-rich small organic compounds, polyaza-substituted benzenes and polymethine dyes. The structures of the sulfur-free series are obviously more complex than in the sulfur-containing one.

The LCOAO procedure:



The alternant hydrocarbon pairing theorem and all-valence electrons theory. An approximate LCOAO theory for the electronic absorption and MCD spectra of conjugated organic compounds, part 2*

Jens Spanget-Larsen

Department of Life Sciences and Chemistry, Roskilde University, P. O. Box 260, DK-4000 Roskilde, Denmark

Received: 27 May 1997 / Accepted: 28 August 1997

Abstract. An approximate linear combination of orthogonalized atomic orbitals (LCOAO) all-valence electrons theory is described, based on a previously suggested partitioning of the Fock operator. Kinetic energy and penetration terms are evaluated explicitly in a Löwdin OAO basis, while two-electron repulsion terms are treated according to the conventional neglect of differential overlap (NDO) approximation. One-electron and penetration integrals are parameterized explicitly to predict approximate alternant pairing symmetry for the π-systems of benzene and napthalene. Application of the resulting LCOAO theory to a variety of alternant and non-alternant hydrocarbons demonstrates significant improvements in the prediction of MCD *B*-terms and transition moment directions, particularly for alternant $(4N+2)$- or $4N$-perimeter π-systems for which traditional NDO procedures fail.

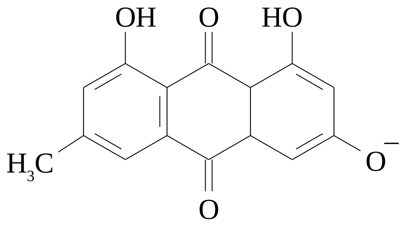
Fig. 1. π and π^* molecular orbital energies for naphthalene (**1**) computed with CNDO/*S* [23c] and LCOAO

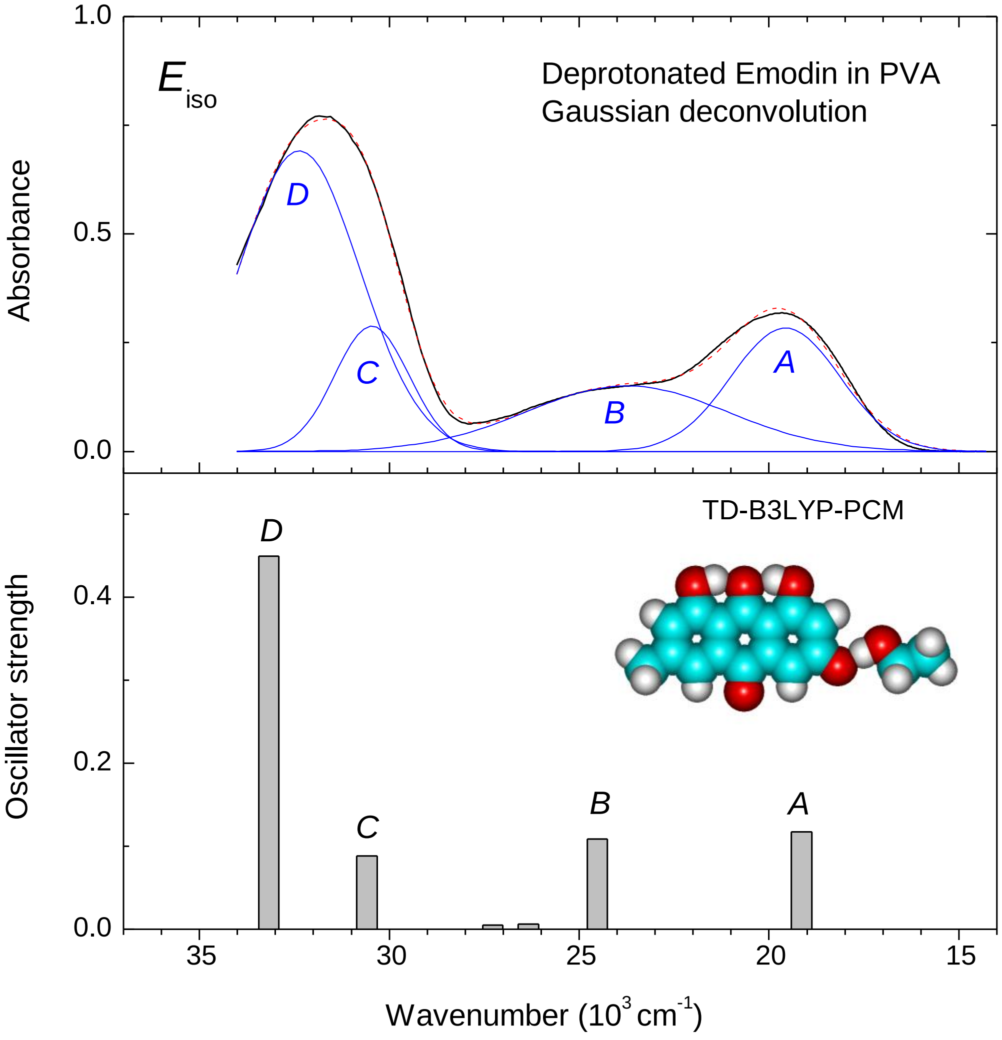
http://akira.ruc.dk/~spanget/LCOAO/PAH/!SUMMARY.htm

Solvent effects

The *Polarized Continuum Model* (PCM) represents the influence of a solvent by considering the molecule in a cavity submerged in a polarizable dielectric continuum. A variety of PCM procedures have been developed, differing particularly in the definition of the solute cavity.

Mixed discreet + continuum solvation model:
An example: In basic alcoholic solution, Emodin is deprotonated at the 3-position. The exposed, negatively charged oxygen center is likely to be specifically solvated by hydrogen bonding to a solvent molecule.





S.C. Nguyen, B.K.V. Hansen, S.V. Hoffmann, J. Spanget-Larsen: *Chem. Phys.* **352** (2008) 167–174
"Electronic states of emodin and its conjugate base. Synchrotron linear dichroism spectroscopy and quantum chemical calculations"