APPLICATIONS OF OPTIMIZATION TO QUANTUM CHEMISTRY

CORY C. PYE


CENTRE FOR NFLD. STUDIES
MEMORIAL UNIVERSITY
OF NEWFOUNDLAND

Applications of Optimization to Quantum Chemistry

By

©Cory C. Pye, B. Sc. (Hon.)

A Thesis

Submitted to the School of Graduate Studies in Partial Fulfillment

of the Requirements for the Degree of Doctor of Philosophy

Department of Chemistry

Memorial University of Newfoundland

January 1997

St. John's, Newfoundland

Abstract

The behavior of the geometry optimization of a large set of molecules has been examined. Particular failures and difficulties are noted for some systems and are shown to correspond to certain conditions. In cyclic systems, difficulties which can arise from using Z-matrix coordinates, which necessarily leave one or more bonds undefined, can be eliminated by using natural internal coordinates. The construction and definition of natural internal coordinates is extended to weakly bound systems, and possible definitions for fused polycyclic ring assemblies are discussed.

5-Substituted cyclopentadienes were chosen for their relevance to studies of facial selectivity in the Diels-Alder reaction. The central atom of the substituents chosen were of p-block elements, with any remaining valences filled with hydrogen. Basis set effects on the structures were examined, as well as the role of conformation on the various geometric parameters. The relative stabilities of the various conformers and the changes in geometry upon change in conformation are predictable from hyperconjugation arguments.

The symmetric transition state of the degenerate substituent migration in 5-substituted cyclopentadienes has been studied. The activation barrier to migration was shown to correlate well with a dimensionless 'stretching' parameter.

Several metal-aquo complexes were studied, including Li^+, Be^{2+}, Mg^{2+}, Al^{3+}, Sc^{3+}, Zn^{2+}, Ga^{3+}, Cd^{2+}, and In^{3+}. For $Li(H_2O)_5^+$, the character of stationary points can change with different basis sets. The vibrational frequency of the symmetric M-O stretching mode is always underestimated with extended basis sets, but can be improved by explicit inclusion of a second solvation sphere.

The bimolecular complexes HF...HF, HF...H_2O, HF...NH_3, HF...CO, HF...NN, HF...NCH, and H_2O...H_2O have been studied. A rough correlation exists between the inverse of the hydrogen-bond distance and the bond strength. For HF complexes, the HF bond lengthens in accordance with the strength of the intermolecular interaction.

The valence tautomerism between benzene oxide and oxepin has been studied. The enthalpies and barriers to tautomerization were very sensitive to basis set and method of correlation. The inversion barriers of the oxepin forms are reported. The effect of simultaneous methyl substitution at the 2 and 7 positions, of protonation, and of the replacement of oxygen with sulfur is examined. Some corrections to photoelectron band assignments are pointed out. The anti selectivity in the Diels-Alder reaction of benzene oxide is determined by a steric effect.

Several 5,6-disubstituted 1,3-cyclohexadienes were studied. The potential surfaces of monocyclic structures were very basis set dependent, with minima disappearing and reappearing as one progressed to higher levels. A full comparison between experimental and theoretical structural and vibrational properties is made for 1,3-cyclohexadiene. The effect of conformation on the structure of cis-3,5-cyclohexadiene-1,2-diol is studied. The predicted photoelectron spectra are compared with experiment and assignments are made.

The Z-matrix optimizations involved in the above studies were examined carefully. For 1,3-cyclopentadienes, some optimization problems could be traced to linearization of atoms involved in a bending coordinate, which rendered a corresponding torsion undefined. The transition state of the substituent shift in 1,3-cyclopentadienes, can be found by minimization in the totally symmetric subspace. The optimization of these species took more function evaluations on average than the corresponding reactant. Indications of the inability of Davidon's Optimally Conditioned method to deal with nearly converged structures may be a problem with either the method itself or the quality of the Hessian. For metal-aquo complexes, the symmetry of the species can be used to simplify the search for minima and/or transition states. In many cases, there are no totally symmetric modes corresponding to water librations, in which case the optimization proceeds smoothly. Bimolecular complexes gave optimization problems when symmetry could not be used to remove coordinates corresponding to the relative orientation of the molecules. For the valence tautomerisations, the monocyclic species

optimized more quickly than the bicyclic species. The use of linear angles, and the resulting nearly singular Hessian gave problems. Poor Hessian updates for transition state optimization resulted in several failures which were corrected by resetting the Hessian. Many problems are related to both the coordinate system and the poor Hessian guess used for high amplitude modes. Proper internal coordinates in general speed up the optimization.

To my wife, Michelle

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Raymond A. Poirier, for his constant encouragement and guidance during my studies. I would also like to thank the other members of my supervisory committee, Drs. D. Jean Burnell and Robert W. Davis.

It is a pleasure to thank my colleagues, Youliang Wang and James Xidos, for their assistance and technical support, and to thank Drs. P. Golding and W. Rudolph for many interesting discussions.

I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Memorial University for financial support. I would also like to thank the Computing and Communications, Chemistry and Physics Departments for the use of their computational facilities.

Finally, I would like to thank my wife, Michelle, for her patience, understanding, and constant encouragement during the writing of this thesis.

Contents

List of Tables

Chapter 4

Appendix 3

List of Figures

Chapter 4

List of Abbreviations and Symbols

AM1	Austin Model 1
Aq	water or H_2O
B	error: failure of back-transformation to Cartesian coordinates
BFGS	Broyden-Goldfarb-Fletcher-Shanno Hessian update
BSSE	basis set superposition error
Ch	1,3-cyclohexadiene
CHA	Mayer's chemical Hamiltonian approach
CISD	configuration interaction with single and double excitations
CNDO	complete neglect of differential overlap
Cp	1,3-cyclopentadiene
DFP	Davidon-Fletcher-Powell Hessian update
DIIS	Pulay's direct inversion in the iterative subspace method
DOC	Davidon's derivative-free optimally conditioned method
E	error: failure of OC method by energy criteria
FC	frozen core approximation

GVB	Generalized Valence Bond
HF	Hartree-Fock
M	error: exceeded maximum iterations
MESQUAC	mixed electrostatic and quantum calculation
MINDO/3	modified incomplete neglect of differential overlap
MM	molecular mechanics
MNDO	minimal neglect of differential overlap
MPn	Møller-Plesset perturbation calculation, n-th order
MP4SDTQ	MP4 with singles, doubles, triples and quadruples
NMR	nuclear magnetic resonance
OC	Davidon's optimally conditioned optimization method
OSIPE	Open Structured Interfaceable Programming Environment
PAH	polycyclic aromatic hydrocarbon
PE	photoelectron
PES	potential energy surface
PSB	Powell-symmetric-Broyden Hessian update
PIC	proper internal coordinates
QCISD(T)	quadratic CISD with perturbative triples correction
RFO	rational function optimization
SCF	self-consistent field
SCRF	self-consistent reaction field

TS transition state

UV ultraviolet

VA Powell's hybrid minimization of least squares subroutine VA05AD

ZPE zero point energy

Chapter 1

Theoretical Background

1.1 Ab Initio Quantum Theory

In quantum chemistry, the energy $\mathcal{E}$ of a system may be regarded as an eigenvalue of a Hamiltonian operator $\mathcal{H}$, for which the corresponding eigenvalue relationship yields the time-independent Schrödinger equation,

$$\mathcal{H}|\Phi\rangle = \mathcal{E}|\Phi\rangle \tag{1}$$

where $|\Phi\rangle$ is some wavefunction (eigenfunction) describing the system of nuclei and electrons. The non-relativistic Hamiltonian describing the physical interactions can be written (in atomic units) as

$$\begin{aligned}\mathcal{H} = &-\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \\ &+\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}},\end{aligned} \tag{2}$$

where N is the number of electrons and M is the number of nuclei[1]. If we let $\mathbf{r}_i$ and $\mathbf{R}_A$ denote the positions of an electron or nucleus of charge Z_A with respect to

some origin, then $r_{iA} = |\mathbf{r_i} - \mathbf{R_A}|$, $r_{ij} = |\mathbf{r_i} - \mathbf{r_j}|$, and $R_{AB} = |\mathbf{R_A} - \mathbf{R_B}|$. In our notation, ∇_i^2 and ∇_A^2 are the Laplacian operators resulting from differentiation with respect to the coordinates of the electron and nucleus, respectively. These five terms are clearly the operators representing the kinetic energy of the electrons, the kinetic energy of the nuclei, the electron-nuclear attraction, the electron-electron repulsion, and the nuclear-nuclear repulsion, respectively.

1.1.1 The Born-Oppenheimer Approximation

Nuclei, being much more massive than electrons, move much more slowly than do electrons. The Born-Oppenheimer approximation treats the electrons as moving in a field of fixed nuclei. This causes the second term of Equation 2 to disappear, and the final term is then a constant, which shifts the eigenvalue spectrum but leaves the eigenfunctions unaffected. We can then solve

$$\mathcal{H}_{elec}|\Phi_{elec}\rangle = \mathcal{E}_{elec}|\Phi_{elec}\rangle, \tag{3}$$

using the electronic Hamiltonian

$$\mathcal{H}_{elec} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}, \tag{4}$$

where $|\Phi_{elec}\rangle$ depends parametrically on the nuclear coordinates. Adding in the constant nuclear-nuclear repulsion term gives the total energy as

$$\mathcal{E}_{tot} = \mathcal{E}_{elec} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}. \tag{5}$$

The Born-Oppenheimer approximation thus allows us to solve the electronic and nuclear parts of the Schrödinger equation separately by introducing the parametric dependence of the total energy on the nuclear coordinates. In other words, for a given set of nuclear configurations, the electronic Schrödinger equation is solved to find the total energy, allowing the energy to be studied as a function of the nuclear coordinates. This becomes important when discussing different regions of the potential energy surface (PES) corresponding to different classical chemical structures. In some cases, the interest lies in determining those points x_* where all points near to x_* possess a higher total energy, as these represent stable equilibrium chemical structures (minima, in mathematical parlance).

1.1.2 Electron Spin and the Antisymmetry Principle

The electronic Hamiltonian in Equation 4 contains no terms that relate to the fact that an electron has an intrinsic spin ω associated with it, unlike a proper relativistic Hamiltonian[2] for which spin is an integral part of the theory. Electron spin is therefore introduced in an *ad hoc* fashion by defining a set of two orthonormal spin functions $\alpha(\omega)$ and $\beta(\omega)$ and defining an electron's coordinate as $\mathbf{x} = \{\mathbf{r}, \omega\}$. The wavefunction for N-electrons can then be written as $\Phi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$. Spin becomes useful only when requiring that the interchange of any two electrons results in a sign change of the wavefunction, which is introduced as a postulate.

1.1.3 Slater Determinants

Except for the electron-electron repulsion term, the Hamiltonian can be written as a sum of one-electron operators,

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}}, \tag{6}$$

each of which will have a set of eigenfunctions $\{\chi_j\}$ and eigenvalues $\{\varepsilon_j\}$ as

$$h(i)\chi_j(\mathbf{x}_i) = \varepsilon_j \chi_j(\mathbf{x}_i). \tag{7}$$

In this case, the operator $\sum h(i)$ will have an eigenfunction $\prod \chi_j$, the product of the one-electron eigenfunctions with an eigenvalue $E = \sum \varepsilon_j$. These eigenfunctions may be regarded as spin orbitals, taken by multiplying a spatial orbital $\psi_j(\mathbf{r}_i)$ with a spin function $\alpha(\omega)$ or $\beta(\omega)$. Unfortunately, this so-called Hartree product is not antisymmetric with respect to the exchange of two electrons. We may antisymmetrize our Hartree product by converting it to a Slater determinant,

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \vdots & & \vdots \\ \chi_i(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix}, \tag{8}$$

which does not change the eigenvalues. Here, we emphasize that the electrons and the spin orbitals are not in correspondence by using numerical indices for the electrons and lowercase alphabetical indices for the spin orbitals.

1.1.4 The Hartree-Fock Approximation

The above analysis only applies to the case where the Hamiltonian can be written as a sum of one-electron operators. The presence of the interelectronic repulsion term invalidates the procedure. However the complicating two-electron operator can be approximated by an 'effective' one-electron Fock operator,

$$f(i) = h(i) + v^{HF}(i), \tag{9}$$

satisfying

$$f(i)\chi_j(\mathbf{x}_i) = \varepsilon_j \chi_j(\mathbf{x}_i). \tag{10}$$

Here, $v^{HF}(i)$ is the average potential of electron i in the field of all the other electrons and can be written as

$$v_a^{HF}(1) = \sum_{b \neq a} \mathcal{J}_b(1) - \sum_{b \neq a} \mathcal{K}_b(1), \tag{11}$$

where the coulomb operator $\mathcal{J}$ is defined as

$$\mathcal{J}_b(1)\chi_a(1) = \left[\int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_b(2)\right] \chi_a(1) \tag{12}$$

and the exchange operator $\mathcal{K}$ is defined as

$$\mathcal{K}_b(1)\chi_a(1) = \left[\int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_a(2)\right] \chi_b(1). \tag{13}$$

We note that the coulomb operator corresponds to the classical coulombic interaction, whereas the exchange operator has no classical counterpart, since it arises from the antisymmetric nature of the wavefunction. The exchange operator is nonlocal since the spin orbitals a and b are exchanged.

1.1.5 Basis Sets

In order to carry out calculations on molecular systems, it is necessary to further specify the form of the spatial orbitals $\psi_i(\mathbf{r})$. This may be done by expanding into a set of K known basis functions as

$$\psi_i(\mathbf{r}) = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu(\mathbf{r}). \tag{14}$$

The weighting coefficients are determined variationally, that is, in such a way as to minimize the energy. The self-consistent-field (SCF) procedure that determines these $C_{\mu i}$ is thus an optimization problem as well, being carried out for each new nuclear configuration.

It is fairly standard procedure to expand the orbitals into linear combinations of atom-centered Gaussian basis functions

$$Nx^l y^m z^n \exp(-\alpha_\mu r^2), \tag{15}$$

because of the ease with which the various integrals can be evaluated. The determination of α_μ and the weights of each ϕ_μ for atoms is also an optimization procedure, the results of which are used as-is for molecular calculations.

1.2 Nonlinear Optimization

1.2.1 Newton's Method

The general problem of unconstrained optimization may be represented mathematically as

$$\text{Given } f : \Re^n \rightarrow \Re$$

$$\text{find } x_* \in \Re^n \text{ for which } f(x_*) \leq f(x), \forall x \in \Re^n, \tag{16}$$

where $\Re^n$ denotes Euclidean space of n dimensions, and f a scalar function of elements x of this space[3]. This can be abbreviated as

$$\min_{x \in \Re^n} f : \Re^n \rightarrow \Re. \tag{17}$$

We normally require f to be sufficiently smooth (differentiable). For any open subset of $\Re^n$, we have the condition that, for a minimizer x_*,

$$\nabla f(x_*) = 0. \tag{18}$$

In the case where f is twice continuously differentiable, we may write a Taylor series expansion in the variable p, about some current point x_c as

$$f(x_c + p) = f(x_c) + \nabla f(x_c)^T p + \frac{1}{2} p^T \nabla^2 f(x_c) p + \mathcal{O}(p^3). \tag{19}$$

We can model the function f at x_c as a quadratic m_c by omitting the cubic and higher order terms from Equation 19, and then solve for the point $x_+ = x_c + s_N$,

where $\nabla m_c(x_+) = 0$. This is called Newton's method, and suggests the iteration scheme

$$\text{Solve } \nabla^2 f(x_k) s_k^N = -\nabla f(x_k)$$

$$x_{k+1} = x_k + s_k^N. \tag{20}$$

The advantages of Newton's method are that, for a sufficiently good initial guess x_0, to a minimizer x_*, with a nonsingular $\nabla^2 f(x_*)$, the sequence $\{x_k\}$ generated by Newton's method converges q-quadratically, i.e.

$$|x_{k+1} - x_*| \leq c|x_k - x_*|^2,$$

and if f is a strictly convex quadratic, then the solution is given in one step. Unfortunately, Newton's method is not globally convergent, requires the solution of a linear system of equations, and requires the first and second derivatives at all iterates.

1.2.2 Modifications of Newton's Method

In some cases, the Hessian $\nabla^2 f(x_c)$ is not positive definite, and as a result, Newton's method is not guaranteed to converge to a minimum. One could proceed along the directions of negative curvature to decrease f [4], or one could change the model Hessian by adding $\mu_c I$ to the actual Hessian, choosing μ_c such that the resulting model Hessian is positive definite[5]. This, in effect, modifies the step toward the direction of steepest descent and makes the model bounded below.

1.2.3 Quasi-Newton Methods

If Newton's method fails to lower the value of f, then another approach must be taken. One can take a step suggested by a global method, choosing the step length judiciously so as to ensure lowering the value of the objective f. This combination is called a quasi-Newton method. A descent direction must be chosen, such as the Newtonian or Cauchy (steepest descent) directions. One can then perform a line search in that direction, or perhaps define a trust region. The trust region approach is equivalent to adding some $\mu_c I$ to the actual Hessian. If δ_c is the trust radius, then $\mu \geq 0$ is found by solving $||s(\mu)|| = \delta_c$, where

$$s(\mu) = -(H_c + \mu I)^{-1} \nabla f(x_c). \tag{21}$$

If the Newton step is within the trust region, then $\mu = 0$. The two most common applications of this approach are the hook step, in which an approximate μ is used in Equation 21, and the double dogleg step, in which a piecewise linear approximation to $s(\mu)$ is constructed.

1.2.4 Secant Methods

A secant method is a quasi-Newton algorithm that uses an approximate Hessian at each point rather than the exact Hessian. This becomes advantageous in cases where the exact Hessian is expensive to compute. Secant methods require no more function and gradient evaluations per iteration than Newton's method and generally update the Hessian approximation at each step. These updated Hessians must satisfy

the secant equation

$$H_+ s_c = y_c \tag{22}$$

where $s_c = x_+ - x_c$ and $y_c = \nabla f(x_+) - \nabla f(x_c)$. Ideally, the updated H_+ should be symmetric.

Several secant updates have been proposed. Powell's symmetric secant update[6] (also known as Powell-symmetric-Broyden or PSB) is

$$H_+ = H_c + \frac{(y_c - H_c s_c)s_c^T + s_c(y_c - H_c s_c)^T}{s_c^T s_c} - \frac{(y_c - H_c s_c)^T s_c s_c s_c^T}{(s_c^T s_c)^2}. \tag{23}$$

A disadvantage of this formula for minimization is that this update does not preserve positive-definiteness. One which does preserve positive-definiteness is the Davidon-Fletcher-Powell (DFP) update[7, 8],

$$H_+ = H_c + \frac{(y_c - H_c s_c)y_c^T + y_c(y_c - H_c s_c)^T}{y_c^T s_c} - \frac{(y_c - H_c s_c)^T s_c y_c y_c^T}{(y_c^T s_c)^2}, \tag{24}$$

also known as the inverse positive definite secant update, which was the first secant update to be discovered. An update which performs better is the complementary DFP update, or the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update[9, 10, 11, 12, 13], which may be written as

$$H_+ = H_c + \frac{y_c y_c^T}{y_c^T s_c} - \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c} \tag{25}$$

or more usefully, in its inverse form,

$$H_+^{-1} = H_c^{-1} + \frac{(s_c - H_c^{-1} y_c)s_c^T + s_c(s_c - H_c^{-1} y_c)^T}{y_c^T s_c} - \frac{(s_c - H_c^{-1} y_c)^T y_c s_c s_c^T}{(y_c^T s_c)^2}. \tag{26}$$

Scaling of the variables in the quadratic model results in no change to the DFP and BFGS, and so one can use a different scaling at each iteration. These methods are therefore referred to as variable metric methods.

The choice of the initial Hessian H_0 is of some importance. In some cases, setting $H_0 = \nabla^2 f(x_0)$ would not be a good choice, since it is not necessarily positive definite. Starting with an identity Hessian guarantees positive definiteness, but ignores the scale. Using a diagonally-scaled Hessian is better. Pre-multiplication of the initial Hessian

$$\hat{H}_0 = \frac{y_0^T s_0}{s_0^T H_0 s_0} H_0 \tag{27}$$

before performing any updates was found to improve the performance of secant algorithms[14]. Some very recent studies have shown that a proper scaling of the Hessian at each iteration can improve the robustness of both the DFP and BFGS trust region methods[15] if a centered Oren-Luenberger scale,

$$\gamma(\theta_k) = \frac{(1-\theta_k)(y_-^T s_-)/(s_-^T s_-) + (\theta_k)(y^T s)/(s^T s)}{(1-\theta_k)(s_-^T B_k s_-)/(s_-^T s_-) + (\theta_k)(s^T B_k s)/(s^T s)} \tag{28}$$

is used after the first iteration. The subscript $-$ denotes that quantity of the previous iterate, and $\theta_k = \min(\tau_1, \tau_2 ||s_k||)$. The BFGS strategy for scaling sets $\tau_1 = 1/2$ and τ_2 large, but the DFP strategy sets $\tau_1 = 1$. Another update that was found to work well is Davidon's optimally conditioned (OC) method[16, 17], which was found to be competitive with BFGS. Comparisons of many variable metric updating schemes have been made recently[18, 19].

An interesting extension to Newton's method was made[20] in which the function

was modelled by a fourth-order function. To keep storage costs low, the third and fourth derivatives were approximated by symmetric, low-rank tensors. Reasonable iteration reduction was obtained.

1.3 The Determination of Saddle Points

1.3.1 Newton's Method with Nonlinear Least Squares

A stationary point is a point x_* at which Equation 18 is satisfied. Clearly, all minima are stationary points. However, not all stationary points are minima. Stationary points which are neither maxima nor minima are called saddle points. The Hessian matrix evaluated at such a point is necessarily indefinite.

Since, in such a case, one cannot apply a minimization algorithm to f, one must find other ways of determining these saddle points. An obvious approach is to solve Equation 18, which is equivalent to solving the minimization problem

$$\min_{x \in \Re^n} \mathcal{F}(x) = \frac{1}{2} R^T(x) R(x) = \frac{1}{2} \sum_{i=1}^{m} r_i(x)^2, \tag{29}$$

where the residual function $R(x) : \Re^n \rightarrow \Re^m$ is equal to the gradient of f, and thus $m = n$. This is a particular case of the zero-residual nonlinear least-squares problem[3].

Let the first derivative of $R(x) = (r_i(x))$ be the matrix $J(x) \in \Re^{n \times n}$, where the Jacobian, $J(x)_{ij} = \partial r_i(x)/\partial x_j$. We can model $R(x)$ about x_c as

$$M_c(x) = R(x_c) + J(x_c)(x - x_c). \tag{30}$$

Now the first derivative of $\mathcal{F}(x)$ is

$$\nabla \mathcal{F}(x) = J(x)^T R(x) \tag{31}$$

and the second derivative of $\mathcal{F}(x)$ is

$$\nabla^2 \mathcal{F}(x) = J(x)^T J(x) + S(x) \tag{32}$$

where

$$S(x) = \sum_{i=1}^{m} r_i(x) \cdot \nabla^2 r_i(x). \tag{33}$$

Newton's method applied to Equation 29 is

$$x_+ = x_c - (J(x_c)^T J(x_c) + S(x_c))^{-1} J(x_c)^T R(x_c). \tag{34}$$

Some problems are that $S(x_c)$ is generally difficult to calculate, and Newton's method can optimize to minima where the gradient is not zero.

1.3.2 The Gauss-Newton Method

The Gauss-Newton method simply ignores $S(x_c)$, giving

$$x_+ = x_c - (J(x_c)^T J(x_c))^{-1} J(x_c)^T R(x_c). \tag{35}$$

For the case where, for Equation 29, $n = m$ this simplifies considerably. If we have a zero-residual problem, then the Gauss-Newton method would be q-quadratically convergent. A disadvantage is that the method becomes ill-defined when $J(x_c)$ does not have full column rank.

Some modifications to the Gauss-Newton method are to make use of a line search in the Gauss-Newton direction[21] or to choose x_+ by a trust-region approach, which, when applied to this problem, is known as the Levenberg-Marquardt[22, 23] method.

1.3.3 Secant Methods

A successful approach is to approximate $S(x_c)$ by a secant solution A_c. One that has been used successfully[24] is

$$A_+ = A_c + \frac{(y_c^\# - A_c s_c)y_c^T + y_c(y_c^\# - A_c s_c)^T}{y_c^T s_c} - \frac{(y_c^\# - A_c s_c)^T s_c y_c y_c^T}{(y_c^T s_c)^2}, \qquad (36)$$

where $s_c = x_+ - x_c$, $y_c = J(x_+)^T R(x_+) - J(x_c)^T R(x_c)$ and $y_c^\# = J(x_+)^T R(x_+) - J(x_c)^T R(x_+)$.

1.3.4 Powell's Method

All of the above methods assume that the Jacobian matrix J is readily available. This is not always the case, however, since J may have to be approximated by finite differencing of (an expensive) $\mathcal{F}$. This may occur in the case of the determination of saddle points as mentioned previously. Powell has proposed and tested a method which is efficient under these circumstances[25]. The method is similar to the Levenberg-Marquardt method with an approximate J updated with Broyden's formula[26]. A closely related version of this method[27] has been successfully implemented in modern quantum chemistry codes[28, 29, 30].

1.4 Applications of Optimization to Chemistry

As mentioned before, the Born-Oppenheimer approximation allows one to study the total energy as a function of the nuclear coordinates. One can apply mathematical optimization techniques as also discussed above to this potential energy function. These are now discussed in the following sections.

1.4.1 Energy Minimization of Geometry

The application of a minimization algorithm to the potential energy as a function of some set of coordinates will ideally result in convergence to a local minimum with respect to these coordinates. If these coordinates span Cartesian space, then these minima would correspond to (locally) stable chemical structures which may be a reactant, product or intermediate of some series of chemical reactions. The application of gradient methods to geometry determination was swift. Poppinger[31] used methods due to Fletcher[32, 33] and to Murtagh and Sargent[34] with an approximate derivative and concluded that gradient methods were superior to direct search methods. Payne essentially uses a Newton-Raphson method with finite differentiation to find the second derivatives and a clever step-choice strategy[35], avoiding the calculation of the gradient. The use of the exact gradient was found to be computationally feasible[36] with the use of the Murtagh-Sargent update. The BFGS update was found to be superior to the MS update, especially for systems with large-amplitude motions[37, 38].

A rather novel method for minimization, the geometric Direct Inversion in the Iterative Subspace (DIIS) was given by Császár and Pulay[40]. One tries to minimize a linear combination of error estimates of previous iterates, and, in so doing, one obtains an interpolated parameter and gradient vector upon which a Newton step is performed. The DIIS method has been combined with the BFGS Hessian update, improving the convergence[41, 42]. Our research group has used a variant of the DIIS method to refine structures for which other methods had failed to satisfy completely the gradient norm tolerance criterion[43, 44, 45].

Another minimizer that has been applied to the minimization of the potential energy surface is the method of conjugate gradients[46]. This was incorporated into Gaussian 80[47]. Further improvements can be made by using an empirical Hessian guess[48, 49].

One does not have to use a quadratic approximation to the potential energy surface. A rational function approximation[50] has been used which is essentially a Padé approximant, as

$$f(x_c + p) = f(x_c) + \frac{(\nabla f)^T p + \frac{1}{2} p^T H p}{1 + p^T S p} \tag{37}$$

for some matrix S. The effect of S is in the third order, but Equation 37 has the added advantage that it results in a finite model for large p. S is chosen to be the identity matrix in the absence of scaling information. The stationary condition requires that

$$\begin{pmatrix} H & \nabla f \\ (\nabla f)^T & 0 \end{pmatrix} \begin{pmatrix} p \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ 1 \end{pmatrix}, \tag{38}$$

which is a generalized eigenvalue equation. The particular application with $S = I$ is similar to a trust region approach with $\mu_c = -\lambda$. For minimization one chooses the smallest λ which ensures a descent direction.

Others have suggested using a cubic approximation to the surface[39, 51, 52]. Pulay corrects for the major anharmonic terms by modifying the Newton step[39]. Stanton gives a multiplicative correction to the Newton-Raphson step based on a Morse potential model[51]. Vogel gives a procedure for extending the DIIS method to third order, and includes update formulas for the Hessian and third-derivative tensor[52].

Pulay[39] suggests that one use a coordinate system that facilitates transfer and comparison between related molecules, and which allows a simple representation of the dominant anharmonic terms. An example of such a coordinate system is then given. He also gives formulae for estimating the Hessian and diagonal and semidiagonal third derivatives (see Appendix) and for converting between Cartesian and internal coordinates.

For the purposes of geometry determination, the use of internal coordinates is recommended[53, 54] over Cartesian coordinates, since the dominant anharmonic terms are more easily represented[54], but this recommendation is not unanimous[55]. A particular scheme for generating 'natural' internal coordinates was introduced which is based on the principles of locality, local pseudosymmetry, and elimination of redundancy[54]. For complex polycyclic molecules, redundancy is difficult

to eliminate, but this is not a problem since optimization in redundant coordinates is feasible[56]. Cartesian coordinates are simpler (but not necessarily cheaper) to use in unconstrained optimization[57] but are somewhat more cumbersome when constraints are introduced[58, 59]. Cartesian and internal coordinates have been successfully combined[60, 61].

1.4.2 Determination of Transition States

On potential energy surfaces, minima are linked by paths which pass through a transition state structure having exactly one direction of negative curvature (along that path) and a zero gradient. The terms transition state and transition structure are hereafter used interchangeably. Since the transition state is not a minimum of the PES in a spanning coordinate system, they are generally more difficult to locate, and direct energy evaluations of the interesting part of the PES are only feasible for a small number of atoms[62, 63]. Their importance is manifested by the amount of attention they receive[64, 65, 66]. In some cases transition states must obey some symmetry criteria[67, 68, 69, 70, 71, 72], which tend to reduce the number of independent variables. One can then minimize the energy in this subspace, for which the direction of negative curvature is orthogonal. Examples include the transition state (D_{3h}) to ammonia (C_{3v}) inversion, and our work on 1,2-heterotropic shifts.

One general way of finding transition states is by a least squares minimization of the gradient norm, which must have a zero minimum. Some early calculations

of transition states by McIver and Komornicki[73] used a Gauss-Newton step with line search as suggested by Powell[21], applied to a semiempirical PES (MINDO/2) for the cyclobutene–butadiene conversion[73], cyclohexane inversion[74], and 1,3-cyclohexadiene–1,3,5-hexatriene conversion[75]. A better version of Powell's least-squares minimizer[25] has been used with an ab initio wavefunction[76]. A straightforward Newton step with an estimated $S(x)$ has also been used[77].

Another method of finding a transition state is to partition the coordinates such that one search direction corresponds to a direction of negative curvature[78, 79]. The initial search direction is chosen to correspond to a direction of negative curvature, which is maximized, and then a set of conjugate directions are generated in which a quasi-Newton minimization (or conjugate gradient[80, 81]) is performed. The Hessian is updated using a BFGS formula. This method was expanded to define a quadratic search path and a changing direction of negative curvature[82]. Scharfenberg proposed a coupled iteration scheme in which successive minimizations and maximizations are performed[83]. This was extended to the case where several coordinates define the maximization space[84]. Culot combined the decomposition with a trust-region approach[85]. This was improved by Bofill, who generalizes the Hessian update and checks the Hessian spectra[86]. One unusual method of finding a TS is to model the surface with a hyperbolic paraboloid and to define the 'constant energy' lines which must intersect at the transition state (the X-method)[87].

A related technique involves 'walking' up a valley from the minimum to a saddle

point[88, 89, 90]. Cerjan and Miller's method is reminiscent of a trust-region approach with a specified step direction and length[91], but requires a Hessian. This idea can be extended to use an approximate Hessian with a BFGS or Powell updating formula[92]. For larger molecules, Cartesian coordinates can be used, as long as the six zero eigenvalues corresponding to translation and rotation are removed (by level shifting), a better step length is used[93], and an intelligent search direction is used[94]. Head, Weiner and Zerner use a similar approach[95], but switch to a least-squares technique once the Hessian develops a negative eigenvalue.

Presumably, the structures of the reactant and the product are known. In this case a synchronous transit method can be used to locate approximately the transition state[96]. The linear synchronous transit (LST) method finds a maximum along a line connecting reactants and products and then performs an optimization orthogonal to that path. With the new point, a quadratic synchronous path (QST) is constructed passing through the three points. A somewhat similar approach without gradients using a simplex method has also been proposed[97, 98]. A MINIMAX/MINIMI procedure has been developed which replaces the orthogonal optimization and allows one to find stable intermediates[99]. A related scheme is Dewar's hypersphere method, in which both paths to the TS are followed[100]. This method has been further developed[101]. Schlegel has combined the LST and QST methods in the early stages of an optimization with a quasi-Newton or EF method once an approximate TS is located[102]. The approach works as well when redundant internal coordinates are

used[103]. Another approach in this case is the 'ridge' method, in which a minimum is found on the ridge separating the valleys from the TS to reactant and product[104]. This method has been recently combined with the DIIS method[105].

The RFO method can also be used to find transition states[50] by partitioning the space into a space over which the energy is to be minimized and an orthogonal space over which the function is to be maximized (for transition states, of dimension 1). Powell's Hessian update formula can be successfully used. This 'eigenvector following' method has been successfully implemented[106] in Gaussian 82[107].

Chapter 2

Computational Developments and Programming Considerations

2.1 Discussion of MUNGAUSS 0.0

The original version of MUNGAUSS[29] was rather versatile in that several methods of optimization were utilized. These included 'Z-matrix' optimization (univariate search), Davidon's Optimally Conditioned (OC) method[16] with and without derivatives (DOC), a version of the BFGS method, and Pulay's DIIS method[40]. For transition states the VA method[27], similar to Powell's[25], was used. Both basis set and geometry optimization could be carried out, and the geometry optimization could be done either in Cartesian, Z-matrix, or proper internal coordinates (PIC)[39] with the PIC input specified similar to TEXAS format[108]. The initial Hessian was usually set to be diagonal, with the elements being estimated by Badger's rule[109] (for stretches) or a constant default value. These diagonal values could also be read

in as input. The VA method had the most flexible choice of initial Hessian, since in addition to these, Hessian elements of selected coordinates could be evaluated by a forward difference formula, the Hessian could be supplied in its entirety, and the entire Hessian could be estimated by a central difference formula.

2.2 OSIPE Considerations

Since the early 1990's our group's efforts have been directed to the maintenance and extension of our ab initio program, MUNGAUSS 1.0[28], which is coded according to the Open Structured Interfaceable Programming Environment (OSIPE)[110]. OSIPE treats a program as a collection of routines which create objects that can correspond to desired quantities, such as an energy or dipole moment, and to quantities required to build these desired objects. A successful extension to and simplification of the Generalized Valence Bond (GVB) code[111, 112] in MUNGAUSS and its subsequent application[113] illustrates the power and flexibility of OSIPE.

2.3 The Input Menu: the Example of Hessian Setup

In order to set up the type of calculation to perform, one must specify such parameters as the molecular system to be studied, the basis set, the wavefunction representation, and the optimization method (if needed). Our input menu takes on the form of a series of commands, delineated by the *end*, within which are either subcommands or assignments to some variable. Typically, only the first few characters

are significant, and these are capitalized in the following text. An example of where a menu is useful would be the specification of a general Hessian approximation for use within an optimizer.

As mentioned before, in MUNGAUSS 0.0 the choice of initial Hessian was restricted to a diagonal Hessian, with the stretching contributions determined from Badger's rule, or some other supplied diagonal guess. This was made to be the default in MUNGAUSS 1.0. In the input menu, the command HESSian is used to control the determination of the initial Hessian to be used. This command has several sub-commands: AN1F calculates rows of the Hessian by forward differentiation of the analytic first derivative; AN1B and AN1C are similar but use a backward and central difference formula, respectively; IDENtity sets the Hessian to the identity matrix: BADGer emphasizes the default. For finite differentiation, the step size is controlled by setting the variable STEP, and the choice to use a symmetric Hessian is controlled by setting SYM within the HESSian command. Within these subcommands, individual parameters may be set to that type by using SET = (list of parameters), or whole groups of parameters may be set by typing one of the descriptors BONds, ANGles, TORsions, ALL, GEOMetry, or BASIs. The list of parameters can be either the name of the parameter (preferred) or the number to which it corresponds. In contrast to the non-OSIPE MUNGAUSS, all current optimization methods can now make use of this Hessian (OC, DOC, BFGS, DIIS, VA). An example is shown in Figure 1.

Figure 1: Hessian construction menu

```
HESSian
  AN1F                          ! Forward difference, analytical 1st deriv.
    SET = (CH1 OH1 P03)         ! As defined in Z-matrix or PIC input
  end
  AN1B SET = ANGles end         ! Backward difference
  AN1C SET = TORsions end       ! Central difference
  STEP=0.01
end
```

2.4 Improvements to Optimization Methods

The first major step taken with the optimizers was to ensure that all worthwhile optimizers were made OSIPE-compliant. The OC, DOC, and BFGS optimizers were straightforward. The VA transition state routine was separated into the Hessian evaluation portion (which was made more flexible as mentioned above) and the optimizer.

A comparison of the various methods for the STO-3G optimization of methanol, given in Table 1, illustrates the improvements that a Hessian can make to the various optimization methods, especially to VA.

The most important change to the DIIS routine was the addition of an optional BFGS updating scheme to the Hessian. This ensured that the formerly static Hessian will be updated as the optimization proceeds. Another addition was the optional cubic correction to the step length if third derivative information is available. The final change was the addition of more throwaway strategies. If the maximum error matrix dimension is exceeded, then one must choose which parameter set to discard. The four options available are 'OLDEST' (the default), 'ENERGY' (throw away

Table 1: Z-matrix optimization of methanol, local C_{3v} symmetry, 5 parameters

Optimizer	Hessian	iterations	function evaluations
VA	diagonal	12	12
VA	forward difference	6	11
VA	central difference	6	16
VA	angles - forward difference	6	8
BFGS	diagonal	4	7
BFGS	angles - forward difference	3	9
OC	diagonal	4	5
DIIS	diagonal	5	5
DIIS	diagonal + update	5	5
DIIS	angles - forward difference	4	6
DIIS	identity	26	26
DIIS	identity + update	11	11

the parameter set corresponding to the highest energy), 'GRAD' (throw away the parameter set corresponding to the largest gradient), and 'ERROR' (throw away the parameter set corresponding to the largest error). A safeguard was added in case the current point was chosen as the throwaway point. The BFGS Hessian update improved significantly the DIIS optimization, especially if a poor starting Hessian is used, as seen in Table 1.

2.5 Optimization Parameters

The geometry can be optimized in either Cartesian, Z-matrix or proper internal coordinates. The proper internal coordinates are defined by the menu command PIC. Each new coordinate is specified by the NEWCoordinate command, in which the NAME, the SCale factor and TYpe of the coordinate (STREtch, BEND, TORSion, or INVErsion) are specified. The components of the coordinate are specified by the

Figure 2: Proper internal coordinate menu

```
PIC
  NEWCoordinate          ! Signals each new coordinate
    NAMe = SCISSOR1      ! Name (Used by Hessian setup)
    SCale = 1.0d0        ! Scaling of each coordinate
    TYpe = BEND          ! STREtch, BEND, TORSion, etc.
    ADDcomp COeff= 4.0d0 AToms= ( H1 C2 H2 ) end
    ADDcomp COeff=-1.0d0 AToms= ( H1 C2 C1 ) end
    ADDcomp COeff=-1.0d0 AToms= ( H1 C2 C3 ) end
    ADDcomp COeff=-1.0d0 AToms= ( H2 C2 C1 ) end
    ADDcomp COeff=-1.0d0 AToms= ( H2 C2 C3 ) end
  end
end
```

ADDcomponent command, for which the coefficient and the atom list defining the coordinate is specified. An example is shown in Figure 2.

The basis set optimization can be specified by the command BOPT. The centers to be optimized together are defined in the command GROUP by either listing the names of the centers (CENters = list) or by specifying that all atoms of a particular element are to be optimized together (ELEment = name). The command SHEll then specifies which attributes of the basis set are to be optimized for the specified TYpe (1S, 3SP, D-POL), such as the SCAle factor, the Gaussian EXPonent, the SDF-Coefficients, or the P-Coefficient in a P or SP shell. An example is shown in Figure 3.

Figure 3: Basis set optimization menu

```
BOPT
  GROup ! What atoms should we treat identically for opt?
! If all carbon atoms have same basis set, use ELEment=C
    CENters = ( C1 C2 C3 ) ! Group centers C1, C2, C3.
    SHELL                  ! Optimize Shell on Group
      TYpe = 3SP           ! As denoted by 3SP
      SCAle = false        ! Don't optimize scale factor
      EXPonent = true      ! Optimize exponent of shell
      SDF-Coef = true      ! Optimize s-contraction coeff.
      P-Coeff = true       ! Optimize p-contraction coeff.
    end
  end
end
```

2.6 Connectivity

One seeming disadvantage of internal coordinates has been the need to define them explicitly[114]. Model builders such as SPARTAN[115] typically generate Cartesian coordinates which are both less intuitive and less efficient[54] than internal coordinates. In order to be able to generate 'natural' internal coordinates automatically, one must define the molecular topology. The most intuitive way of carrying this out is to define which atoms are connected. A molecule in this sense may be regarded as a **graph** $\mathcal{G}$, which consists of a nonempty set of **vertices** $\mathcal{V} \neq \emptyset$ (atoms) and a distinct set of **edges** $\mathcal{E}$ (bonds), where each element of $\mathcal{E}$ consists of an unordered pair of distinct elements of $\mathcal{V}$. Two vertices (or edges) sharing a common edge (or vertex) are said to be **adjacent**. An edge vw is **incident** with the vertices v and w. The number of edges incident with v is called the **degree** of v. Multiple edges vw

or edges incident with only one vertex vv (**loops**) are not permitted, but these are allowed in the more general object known as a **pseudograph**. A **subgraph** $\mathcal{G}'$of $\mathcal{G}$ is a graph whose vertex and edge sets are subsets of those of $\mathcal{G}$.

Considerable flexibility in constructing the graph associated with a molecule exists since we are mapping a multidimensional collection of real ordered triples (Cartesians) to an integer set of connections. We choose to use the formula

$$C_{ij} = \left| \frac{\text{sign}(d_{ij} - \lambda_{BS}(R_i^{BS} + R_j^{BS})) - 1}{2} \right|,$$

where C_{ij} is the **adjacency matrix** of the graph, d_{ij} is the distance between atoms i and j, R^{BS} is the Bragg-Slater radius, and λ_{BS} is a scale factor (default value = 1.2). With this definition one can find most normal bonds of a chemical structure, from which we can define angles and torsions. The default specification should not find extremely long 'bonds' corresponding to transition vectors, hydrogen-bonds or van der Waals' contacts. In general, the graph it generates will consist of several disjoint **components** corresponding to 'fragments' of the inputted structure. This may consist, for example, of just one molecule where the graph is **connected**, a bimolecular transition state consisting of two (e.g. Diels-Alder) or three fragments (e.g. S_N2), or a collection of molecules (e.g. lithium ion surrounded by eight water molecules). The utility of such a definition is apparent in the study of intermolecular interactions as well. A very efficient graph theoretical algorithm is used for determining the molecular connectivity[116]. The adjacency matrix is also known as the **connection matrix**.

For describing weaker connections such as hydrogen-bonds and bonds corresponding to transition vectors, we need a more relaxed distance criterion. In this case, we use another definition to find those bonds not yet found by the above equation.

$$C_{ij}^{*} = \left| \frac{\mathrm{sign}(d_{ij} - \lambda_{vdW}(R_i^{vdW} + R_j^{vdW})) - 1}{2} \right| - C_{ij} - (C^2)_{ij,i\neq j}.$$

In this case we use the van der Waals' radius, which is typically larger than the Bragg-Slater radius, and a different scale factor λ_{vdW}. Subtracting C_{ij} ensures that we do not count normal bonds twice, and subtracting $(C^2)_{ij,i\neq j}$ ensures that atoms which are bonded to a mutual third atom (which are usually within each others van der Waals' radii) are not counted. The default value of λ_{vdW} is 0.8. Larger values sometimes resulted in finding 'connections' between vicinal cis or gauche atoms. Forseeable problems with this definition are that transition states involving the creation/destruction of three-membered rings are excluded *a priori*, and that for atoms whose van der Waals radii are only slightly greater than their Bragg-Slater radii, the element of the **close contact matrix** C_{ij}^{*} may actually be negative because of the choice of the various λ.

While the above two algorithms are usually sufficient to result in a connected molecular graph, for the case of extremely weakly bound species, or for studies of intermolecular interactions at large intermolecular distances it is desirable to have a failsafe mechanism for connecting these widely separated fragments. If the extended graph consisting of the edge sets determined in the above two formulas is still not connected, we find the closest two atoms belonging to different components (where

the van der Waals radii has been subtracted) and force a connection. We add this edge to the total edge set and iterate until the graph is connected. This gives the **intercomponent connection matrix.**

An example of the application of these ideas to the pentahydrate of lithium ion is given in Figure 4. The solid lines represent the connections, the short-dashed lines the close contacts, and the long-dashed lines the intercomponent connections.

For controlling the generation of these connection matrices we have defined an input menu initiated by the command CONNectivity which allows for great flexibility. Default values will not necessarily work in all cases, so we allow for the customization of the scale factors by setting BSCale and VDWScale. In addition one can change the individual radii from within the SETBragg or SETVdw commands using MODify as shown in Figure 5.

2.7 Topological Examination of Molecules

Once a molecular graph $\mathcal{G}$ has been constructed, several related graphs may be derived. A **pruned** subgraph of $\mathcal{G}$ is the graph remaining when all vertices of degree 0 or 1 are recursively removed until no more remain. A **homeomorphically reduced graph** associated with $\mathcal{G}$ is a graph produced by recursively replacing vertices of degree two (and thus their incident edges) with an edge. It is clear that the two adjacent vertices must be distinct, otherwise a pseudograph with multiple edges may be produced. If this procedure is continued, the resulting object we define as the

Figure 4: Application of connection analysis to $Li(H_2O)_5^+$

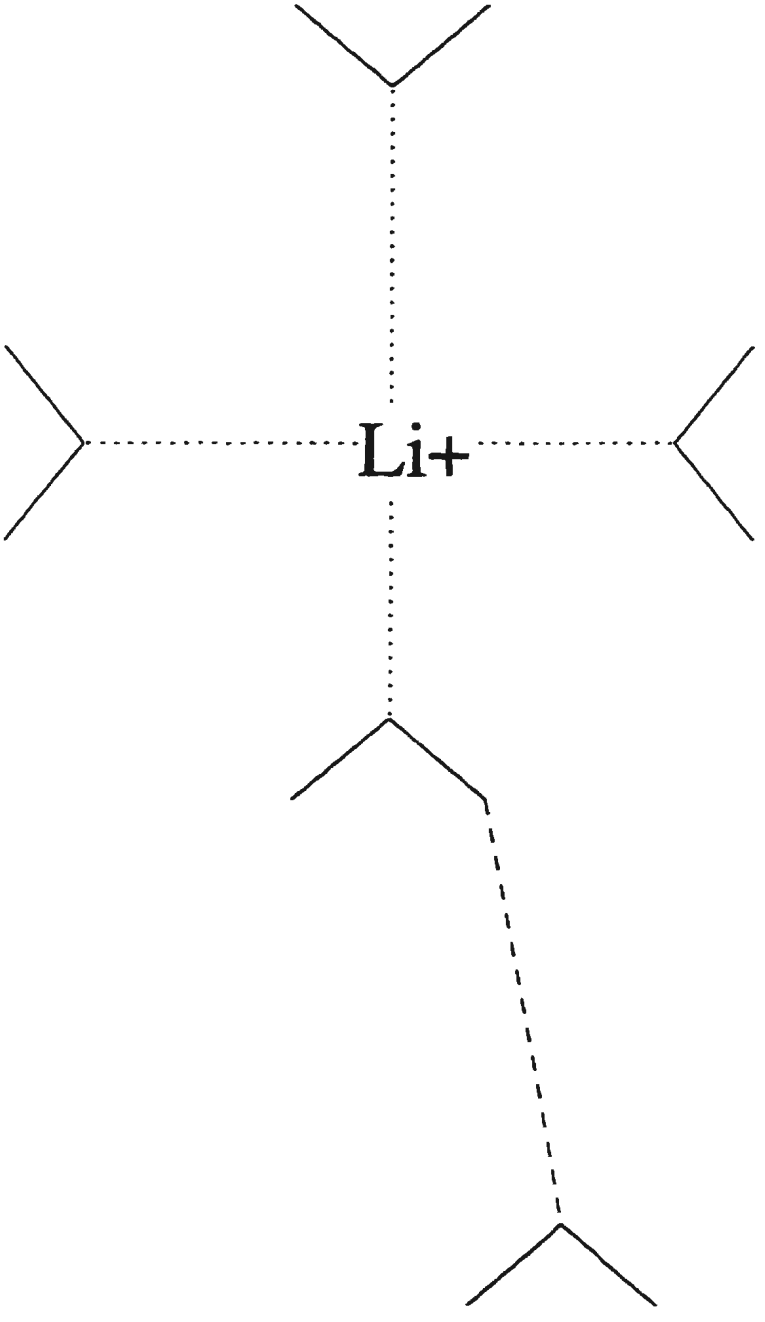


Figure 5: Connectivity menu

```
CONNectivity
  BSCale = 1.25   ! Scale factor for Bragg-Slater radii
  VDWscale = 0.86 ! Scale factor for van der Waals radii
  SETBragg
    MODify
      ATOM=C RADius = 0.75 ! Carbon B-S <= 0.75 Angstroms
    end
  end
  SETVdw
    MODify
      ATOM=O RADius = 1.45 ! Oxygen vdW <= 1.45 Angstroms
    end  ! Modify
  end ! SetVdw
end ! CONN
```

homeomorphically reduced pseudograph which contains multiple edges but no loops. Continuing once more, some vertices of degree two would be converted to a loop, which we will call a **minimal homeomorphically reduced** pseudograph. If a loop is considered as adding two to the degree of a vertex, then for isolated loops, the next iteration would annihilate the loop. The resulting object (which may be the empty set) is called the **fully homeomorphically reduced** pseudograph. It is clear that the minimal homeomorphically reduced pseudograph preserves the ring structure of the full molecule.

There are other concepts of graph theory that are useful for us. A **walk** is an alternating sequence of vertices and edges, starting and ending with vertices, in which each object is incident to those immediately preceding and following it in the sequence. If the edges are distinct, the walk is a **trail**, if the vertices are distinct, the walk is a

path. If the first and last vertices in a walk are identical, we say the walk is **closed**. otherwise it is **open**. A closed trail is called a **circuit**. A circuit whose vertices are distinct (except the first and last) is called a **cycle**, which may be prefixed by the number of vertices (which would equal the number of edges). A cycle is also known as a **ring** in the chemical literature. Finding a representation of all cyclic subgraphs is an important area of research. A **ring assembly** is a set of rings which cannot be subdivided into two or more sets whose edge sets do not intersect (fused ring system). A **separable** graph is a connected graph for which the removal of one vertex disconnects the graph. Molecules with simple spiro centers fall into this category. The **vertex connectivity** of a connected graph is the minimum number of removal of vertices required to disconnect the resulting subgraph. If the vertices are then replaced to each resulting component, the resulting graph is said to be **1-isomorphic** to the original. If the vertex connectivity is 2, we can define an analogous **2-isomorphism**. If the two cut-vertices of a graph of vertex connectivity 2 are adjacent, then we have a fused ring system such as in propellane or decalin. If these two vertices are not adjacent, then the atoms are called bridgeheads.

In order to find the ring systems, we use an algorithm similar to that of Matyska[117]. First, the molecular graph is pruned to remove all paths not terminated at both ends by a cycle. Secondly, from the pruned graph, the minimal homeomorphically reduced pseudograph is constructed. Once this graph is determined, a fundamental set of rings is constructed and the ring assemblies determined. We do not use the binary

representation of Matyska since difficulties arose with the spanning tree construction if loops were present. Once the fundamental rings and ring assemblies are found, a mapping array from each edge to membership of a ring assembly is constructed, with a map to zero indicating that the bond is not a member of any ring assembly. A pictorial representation of the application of the pruning, homeomorphic reduction and ring finding procedure to an organic molecule ($Cp(PhNO_2)(PhNMe_2)$) is given in Figure 6.

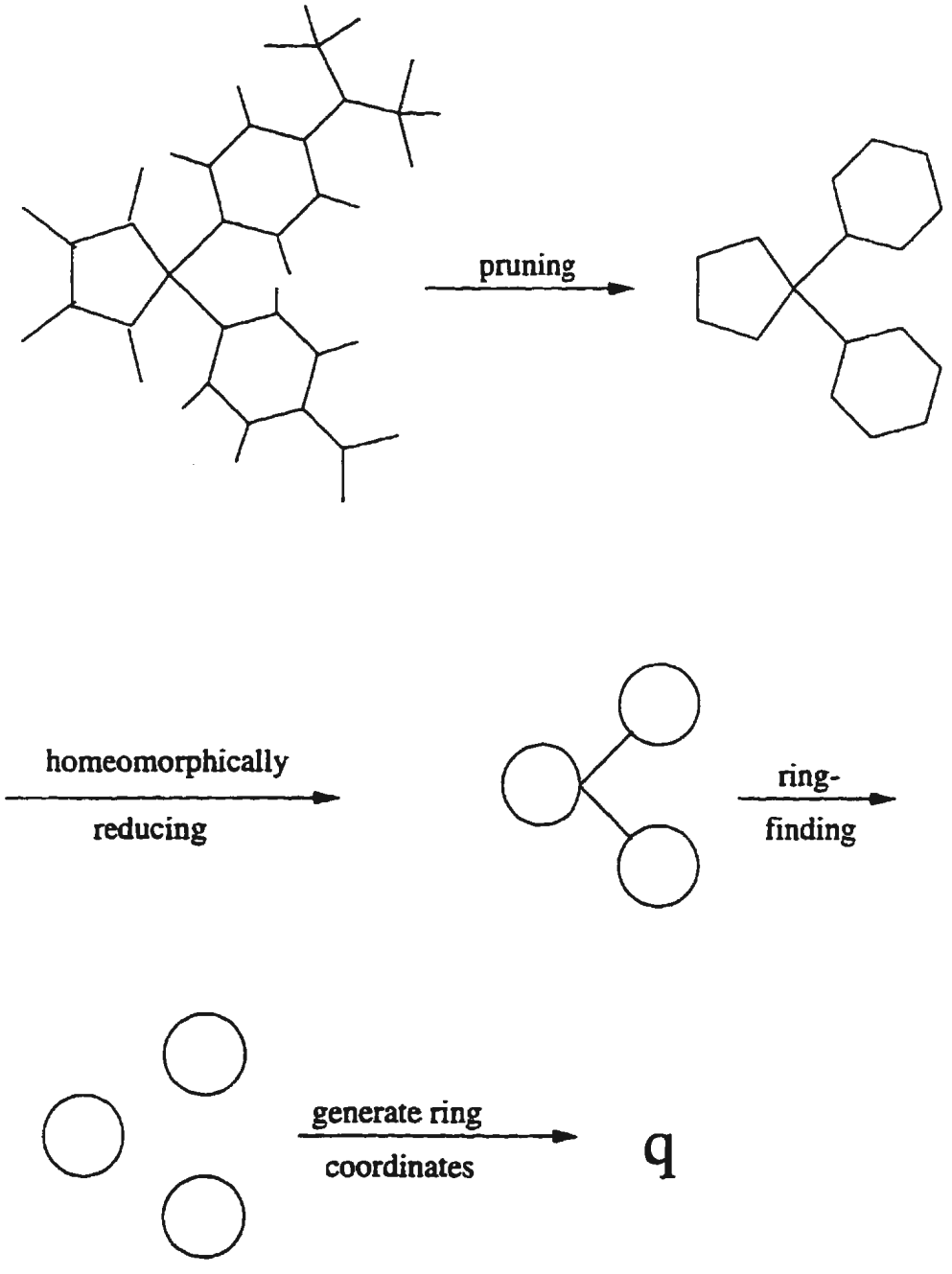
2.8 Automatic Generation of Coordinates

One of the disadvantages of using proper internal coordinates[39] has been the need to define them explicitly. Recently, their definition was extended to handle more complicated cases, and, more importantly, their construction was automated[54]. We have written a natural internal coordinate generation program for MUNGAUSS 1.0 which is similar to that of Pulay.

2.8.1 Internal Coordinates of Chains

All true stretches (from the connectivity matrix) are defined as individual coordinates. Non-ring torsional modes (with a central bond determined from the connectivity matrix) are defined as the sum of all torsions about this bond. Non-ring bending modes are defined as symmetrized linear combinations about a given atom. Any valence bend which is totally enclosed in one ring assembly is not considered.

Figure 6: Application of topological analysis to $Cp(PhNO_2)(PhNMe_2)$



However, the relative motion of spiro rings are included in this category (rock, wag, twist) since the two bonds comprising each bending component belong to two different ring assemblies. Out-of-plane bending modes are also handled.

2.8.2 Ring Systems

The deformation modes of simple rings are formed as linear combinations of the valence bends of the ring according to

$$S_m^a = \sum_{k=1}^{n} \cos\left\{\frac{(k-1)m2\pi}{n}\right\} q_k$$

and

$$S_m^b = \sum_{k=1}^{n} \sin\left\{\frac{(k-1)m2\pi}{n}\right\} q_k,$$

where q_k is a ring bend or torsion, and $m = 2, \ldots, [n/2]$. The $n-3$ bends and torsions are chosen from the above set.

More complicated ring assemblies pose special problems. Ring assemblies not composed of a single ring must have bridgehead atoms. If a pair of bridgehead atoms are adjacent, then it may be possible to decompose the ring assemblies further into subassemblies sharing a common edge (2-isomorphic to the original). When this can be done, the relative motion of the subassemblies can be described by valence parameters involving this edge or the incident atoms. Pulay treats the bicyclic and propellane cases separately, but these are simply subcases of two and three subassemblies, respectively. This insight also allows us to treat the subassemblies separately, combining them with the appropriate coordinate. Pulay only explicitly mentions the

cases where the subassemblies themselves are simple rings, but this can clearly be extended to more general cases. The use of the torsion components for this mode (as recommended by Pulay) is possible, but one could also use symmetrized bending components as well, although more coupling to the subassembly deformation modes would be expected. The coordinate generation for these systems has not yet been automated.

For more complicated cases than these, Pulay recommends the use of redundant coordinates. It would be convenient to construct topologically-derived coordinates for any system, regardless of the molecular topology. Numerical approaches to generating coordinates (the 'delocalized' internal coordinates[118]) look promising but lack chemical interpretation. For bicyclic systems of the form [n.m.k] the ring coordinates of the two rings with (n+k+2) and (m+k+2) atoms are used, which introduces redundancy. It is implied that one should use the two smallest rings but there is no reason *a priori* to choose these two of the (redundant) set of three possible rings (the third containing (m+n+2) atoms). In addition, if $m = k$ then the choice becomes arbitrary and would destroy symmetry. Inspired by the propellane coordinates above, one could 'pretend' that a bond exists between the two bridgeheads and write the appropriate propellane distortions, or perhaps as suggested above, write the propellane distortions in terms of bendings rather than torsions, in which case no 'pretending' is necessary.

The ideas were tested on bicyclo[1.1.1]propane (C_5H_8). The 13 atoms require

33 internal coordinates, of which 14 are simple stretches. Each of the three CH_2 bridges requires a scissor, rock, wag and twist (4 coordinates) to describe the hydrogen deformation motions (total of 12), and each of the two bridgehead hydrogens requires a degenerate rocking mode (2) for a total of 4 and a grand total of 30 coordinates. The propellane-like distortions (which we call ring wag and scissor) are defined as appropriate bending combinations about the bridgeheads, and the final coordinate (ring 'squish') as either the sum of the bridge bends or as an appropriate combination of all nine skeletal angles, as shown in Table 2. The angles in Table 2 follow the order: angles about bridgehead 1, corresponding angles about bridgehead 2, opposite bridge angles. We note the remarkable coincidence between the coordinates used and the symmetrized linear combinations of the ring deformations. This suggests that some double symmetry principle may be used to generate suitable coordinates, whereby the symmetry about the bridgeheads is used on coordinates which are themselves symmetric combinations of ring deformations. This type of coordinate generation has not been automated. Use of these coordinates were tested with the OC optimization method starting from a point 0.868 Hartrees (2280 kJ/mol) above the final geometry. The results were impressive. The Z-Matrix optimization converged in 29 function evaluations whereas optimization with these coordinates converged in 11 function evaluations.

Table 2: Deformation coordinates of bicyclo[1.1.1]pentane

	Angle Coefficients								
Description	Bridgehead 1			Bridgehead 2			Bridges		
Ring A def	1	0	0	1	0	0	0	-1	-1
Ring B def	0	1	0	0	1	0	-1	0	-1
Ring C def	0	0	1	0	0	1	-1	-1	0
Ass. sciss	2	-1	-1	2	-1	-1	0	0	0
Ass. rock	0	1	-1	0	1	-1	0	0	0
Ass. sq (1)	0	0	0	0	0	0	1	1	1
Ass. sq (2)	1	1	1	1	1	1	-2	-2	-2
A+B+C	1	1	1	1	1	1	-2	-2	-2
2A-B-C	2	-1	-1	2	-1	-1	0	0	2
B-C	0	1	-1	0	1	-1	1	-1	0

2.8.3 Weak Connections

As discussed before, weak connections such as H-bonds and breaking/forming bonds in transition states may need special treatment. The components of the inputted molecular structure are connected by close contacts (weak bonds). First, the close contacts are examined to ensure that there are no rings of weak bonds involving three or more components. If there are, then one of the weak bonds is excluded from further consideration. Weak bonds between atoms of the same component are likewise not considered. Of the remaining bonds, all stretches are generated. If for each pair of components there is exactly one weak bond under consideration then the five coordinates relating to their relative position/orientation are generated. If there are greater than three weak bonds between a pair of components, then the first three encountered are chosen and the rest dropped from further consideration. For two weak bonds with noncoincident atoms, the coordinates recommended by Pulay

are suggested, namely, the '4-ring' deformation, and the two 'butterfly' coordinates. For coincident atoms on component A, we suggest the use of the butterfly motion of component B, and three deformation coordinates of A. For the case of three connections, we may have a threefold coincidence on A or B, a 2 fold concidence on both A and B, a 2 fold coincidence on either A or B, or no coincidences. In these cases we would suggest

- **3-fold A**: B-twisting, B-deformation (2)
- **2 × 2-fold**: 4-ring pucker, A,B-butterfly
- **2-fold A** : Shearing along A, rocking A, twisting
- **None**: Twisting, Shearing (2).

Only the single-connection deformations are currently automatically generated. By construction, the intercomponent connections (very weak bonds) are treated in the same manner as the single close contacts.

2.9 Application Section

Much of the applications-related material from this work has been computed using the pre-OSIPE version of MUNGAUSS. These many types of systems serve as a performance benchmark to compare with the newest version. The discussion of the optimization behavior itself will be discussed in a later chapter. We now discuss several systems of chemical interest, namely, 5-substituted cyclopentadienes and their

conformations, the 1,2-heterotropic shift of this substituent, metal-water complexes. weak bimolecular complexes, the benzene oxide–oxepin valence tautomerism, and 5,6-disubstituted 1,3-cyclohexadienes.

Chapter 3

Applications

As stated before, local minima are easily characterised on a PES by the conditions of Equation 18 and by insisting that $\nabla^2 f(x_*)$ is positive definite, i.e. for all vectors p,

$$p^T \nabla^2 f(x_*) p > 0,$$

and in particular, for those p which are eigenvectors of $\nabla^2 f(x_*)$, i.e. for which

$$\nabla^2 f(x_*) p = \lambda p, \text{ where } \lambda > 0.$$

We let $\{v_i\}, i = 1, N$ be a (not necessarily unique) set of eigenvectors, and $\{\lambda_i\}, i = 1, N$ their corresponding eigenvalues. Although characterising a single local minimum is relatively straightforward, finding the global minimum, or finding *all* of the minima of a large system is a much more daunting task.

Transition states are characterized by the condition that only one λ_i is negative, in addition to Equation 18. If the corresponding v_i has no component corresponding to a bond stretch, then the transition state must be a barrier to conformation change

or inversion, such as rotation about a $C(sp^3)$-$N(sp^3)$ bond or the umbrella motion of ammonia. Otherwise, bonds are being broken or formed, such as in the familiar S_N2. Diels-Alder, and 1,3-hydrogen shift reactions.

A variety of reactions are presented in the following sections. Each of the structures discussed in the six subsections have merits for testing optimization. 5-Substituted 1,3-cyclopentadienes possess a 5-membered ring, and in some cases, a loose torsional mode corresponding to internal rotation. The transition states to 1,2-heterotropic shift have two bonds undergoing cleavage and some have loose deformational modes. The metal-water complexes have both relatively strong metal-water interactions, and in some cases, a hydrogen bond network. The bimolecular complexes possess hydrogen bonds and weak librational modes. The benzene oxide - oxepin valence tautomers exhibit 3, 6, and 7-membered rings, and the transition states for their interconversion show the conversion from a bicyclic to a monocyclic structure. The cyclohexadienes give, in some cases, considerable basis set dependence in their conformational behavior, and also give a relatively loose bicyclic ring system.

3.1 5-Substituted Cyclopentadienes and Conformation

The study of 5-substituted cyclopentadienes is central to the subsequent study of 1,2-substituent shifts within these molecules and to their Diels-Alder reactions. We therefore discuss the structure of 5-substituted cyclopentadienes and the energetics

of conformational change that some of these can undergo. [1]

3.1.1 Geometries

In most cases, geometries were optimized within the C_s point group. For the gauche minima and eclipsed gauche maxima, no symmetry (C_1) was imposed. Cyclopentadiene itself optimized to a C_{2v} structure. In all tables, bond lengths are given in Angstroms and angles in degrees. A general trend noticed is that d-polarization functions added to heteroatoms of the first and second row tend to shorten bonds to adjacent atoms, whereas, when added to heteroatoms to the third and fourth row, tend to leave unchanged or to increase bond lengths to adjacent atoms. Of course, for first and second row atoms, the d-functions are genuine polarization functions in the sense that they augment the basis set with functions of higher symmetry. For third and fourth row atoms, however, these functions may partly serve to describe the more diffuse region of the occupied d-shells in addition to polarizing the atom in the molecular environment.

1,3-Cyclopentadiene

The structure varied little with basis set, and the largest structural changes were about 0.015 Å in the C-C bond lengths, 0.012 Å in the C-H bond lengths, and about 1° in the angles (see Table 1). The geometry is reasonably close to previous ab initio

[1] I thank Mr. James D. Xidos for permission to use the results of many of the 6-31G* optimizations on the staggered, eclipsed and gauche structures, which are pertinent to an independent investigation[119].

Table 1: Selected geometric parameters of 1,3-cyclopentadiene

	Basis Set		
Parameter	STO-3G	3-21G	6-31G*
C_2-C_3[a]	1.4901	1.4847	1.4764
C_2-C_1	1.3188	1.3293	1.3285
C_1-C_5	1.5223	1.5193	1.5064
C_2-H_2	1.0812	1.0691	1.0742
C_1-H_1	1.0806	1.0689	1.0735
C_5-H_5	1.0914	1.0866	1.0890
C_3-C_2-C_1	109.27	109.29	109.17
C_5-C_1-C_2	109.90	109.66	109.59
C_4-C_5-C_1	101.66	102.10	102.47
H_2-C_2-C_3	123.62	124.14	124.39
H_1-C_1-C_2	127.45	126.93	126.70
H_5-C_5-H_5'	107.37	108.28	106.72
H_5-C_5-C_1	111.96	111.61	111.94

[a]The carbons of the cyclopentadiene are numbered cyclically with the sp^3 carbon being given the label C_5.

calculations[120, 121] and to the experimental temperature-independent microwave structures[122, 123, 124], but somewhat further from the temperature-dependent X-ray[125] and electron diffraction[126] structures. The 6-31G** structure differed from the 6-31G* structure only in the last reported significant figure.

Changes of similar magnitude occur when one of the hydrogens at C_5 is replaced by a series of substituents within a given basis set. The largest changes should be in the bond lengths and angles adjacent to the substituent. The substituents considered are F, Cl, Br, I, OH, SH, SeH, TeH, NH_2, PH_2, AsH_2, SbH_2, CH_3, SiH_3, GeH_3, and SnH_3.[2]

[2]Other substituents have also been investigated by our group, such as $CH{=}CH_2$, NH_3^+, OH_2^+, SH_2^+, O^-, S^-, CH_2OH, and CH_2OCH_3[119], and CCH, CN, Li, Na, OCH_3, SCH_3, BH_2, OCHO and SO_2H.

5-Halocyclopentadienes

Accurate geometries for 5-halocyclopentadienes have not yet been determined experimentally, though all have been prepared[127, 128, 129]. The 3-21G C-C and C-H bond lengths agree better with the 6-31G* results than STO-3G, but the reverse is true for C-F and C-Cl bond lengths (see Table 2). Polarization functions on chlorine are essential in describing the C-Cl bond length at the split-valence level. The adjacent C-C and C-H bonds shorten as one goes down the series, consistent with the diminishing electron withdrawal. The smaller basis sets tend to overestimate the magnitude of the shortening relative to 6-31G*.

Of the three basis sets employed, for CpF and CpCl the STO-3G and 6-31G* bond angles are closest, whereas for CpBr and CpI the 3-21G and 6-31G* results are closest. The X-C-H angle decreases down the group at the 6-31G* level, whereas no discernible trend is noted at the other basis sets. The X-C-C and H-C-C angles decrease and increase, respectively, upon going down the group at this level.

5-Chalcocyclopentadienes and Conformation

No 5-chalcocyclopentadienes have been isolated, presumably because of a rapid [1,5]-sigmatropic rearrangement (i.e. 1,2-hydrogen shift) to give an enol analogue (see Figure 1) followed by rapid conversion to the corresponding ketone analogue (see Figure 2).

Table 2: Selected geometric parameters of halocyclopentadienes

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(*)	6-31G*
F	F-C	1.3884	1.3990		1.3694
	C-C	1.5400	1.5207		1.5093
	H-C	1.1035	1.0845		1.0865
	F-C-C	113.24	114.60		113.42
	H-C-C	110.26	108.61		110.06
	F-C-H	108.49	107.57		106.85
Cl	Cl-C	1.8272	1.8944	1.8212	1.8027
	C-C	1.5266	1.5101	1.5167	1.5062
	H-C	1.0947	1.0760	1.0780	1.0822
	Cl-C-C	112.15	109.78	110.38	112.78
	H-C-C	112.45	114.71	113.49	111.57
	Cl-C-H	105.61	104.43	106.37	105.18
Br	Br-C	1.9294	1.9989		1.9755
	C-C	1.5260	1.5097		1.5022
	H-C	1.0937	1.0772		1.0797
	Br-C-C	112.91	110.65		111.81
	H-C-C	111.27	113.93		113.07
	Br-C-H	106.85	104.92		103.84
I	I-C	2.1424	2.2068		2.2004
	C-C	1.5221	1.5101		1.4992
	H-C	1.0917	1.0785		1.0794
	I-C-C	112.84	111.09		111.36
	H-C-C	111.61	113.64		113.96
	I-C-H	106.27	104.51		102.83

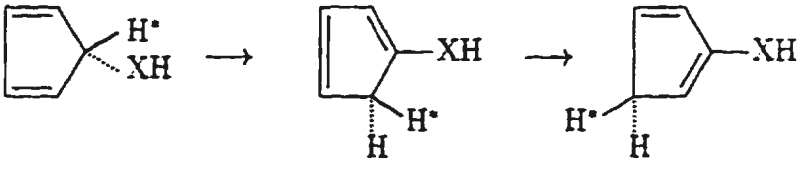


Figure 1: 1,2-hydrogen shift of 5-chalcosubstituted cyclopentadienes

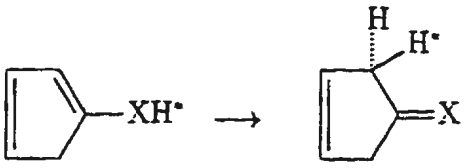


Figure 2: 1,3-hydrogen shift of isomeric cyclopentadiene

5-Chalcocyclopentadienes have been implicated in the decomposition of substituted dicyclopentadienes (Cp_2OX, X=H, $COCH_3$)[130, 131]. The MNDO method estimates that 2,4-cyclopentadien-1-ol is about 30 kJ/mol less stable than its isomeric enol forms[132], as suggested by the acetate derivative ratios[133]. The derivatives of 5-chalcocyclopentadienes are useful in synthesis[134] and for other more fundamental experimental[135, 140, 141, 142, 137, 138, 139] and theoretical[136, 137, 138, 139, 140] investigations.

Tables 3, 4, 5, and 6 show that the C-C bond is lengthened relative to cyclopentadiene, but this difference decreases quickly down the series. The O-C-C, S-C-C, Se-C-C and Te-C-C angles are all larger than the corresponding H-C-C angle of cyclopentadiene, as was noted previously for CpF, but this difference also decreases down the series. The H-C-C angle is smaller in CpOH and CpSH than in CpH, but the other two members of the series show little difference.

For CpOH, the O-C and O-H bond lengths decrease upon progressing along our basis set series, and the presence of oxygen polarization functions seem important. With the exception of H-O-C, the bond angles seem insensitive to basis set. For CpSH, the S-C and S-H bond lengths are very sensitive to basis set, and the necessity

Table 3: Selected geometric parameters of 2,4-cyclopentadien-1-ol

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
OH antiperiplanar	O-C	1.4365	1.4351	1.4014
	H-O	0.9915	0.9671	0.9484
	C-C	1.5382	1.5266	1.5158
	H-C	1.0989	1.0830	1.0864
	O-C-C	115.23	115.53	115.39
	H-O-C	103.91	109.16	108.64
	H-C-C	109.75	109.14	109.21
	O-C-H	105.77	105.47	105.34
OH synperiplanar	O-C	1.4419	1.4383	1.4052
	H-O	0.9892	0.9643	0.9446
	C-C	1.5382	1.5245	1.5139
	H-C	1.0994	1.0891	1.0911
	O-C-C	113.32	113.86	113.93
	H-O-C	104.78	111.07	110.10
	H-C-C	108.90	107.71	108.13
	O-C-H	110.89	111.03	110.00
OH synclinal	O-C	1.4384	1.4375	1.4027
	H-O	0.9917	0.9665	0.9471
s[a]	C-C	1.5381	1.5271	1.5156
a	C-C	1.5339	1.5202	1.5099
	H-C	1.1027	1.0894	1.0927
s	O-C-C	115.59	116.63	116.03
a	O-C-C	110.98	110.54	111.22
	H-O-C	103.74	110.27	109.38
s	H-C-C	109.23	108.11	108.58
a	H-C-C	109.36	108.27	108.42
	O-C-H	110.14	110.71	109.89
	H-O-C-H	60.01	61.33	56.53
OH anticlinal	O-C	1.4438	1.4409	1.4065
	H-O	0.9895	0.9650	0.9454
s	C-C	1.5343	1.5277	1.5143
a	C-C	1.5376	1.5207	1.5115
	H-C	1.1025	1.0877	1.0911
s	O-C-C	116.41	116.82	116.32
a	O-C-C	112.60	111.18	112.20
	H-O-C	104.57	110.70	110.00
s	H-C-C	109.19	108.15	108.54
a	H-C-C	108.95	108.67	108.59
	O-C-H	108.35	109.70	108.70
	H-O-C-H	113.64	95.42	103.98

[a]The 's' denotes that the particular parameter is syn to XH (synperiplanar or synclinal), and 'a' denotes anti to XH (antiperiplanar or anticlinal).

Table 4: Selected geometric parameters of 2,4-cyclopentadiene-1-thiol

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(*)	6-31G*
SH antiperiplanar	S-C	1.8123	1.8951	1.8298	1.8300
	H-S	1.3308	1.3515	1.3258	1.3270
	C-C	1.5268	1.5123	1.5185	1.5078
	H-C	1.0928	1.0810	1.0845	1.0867
	S-C-C	114.93	112.28	112.35	114.65
	H-S-C	95.53	96.62	96.31	97.23
	H-C-C	110.02	112.83	111.26	110.71
	S-C-H	105.74	104.09	105.47	103.65
SH synperiplanar	S-C	1.8205	1.9053	1.8408	1.8424
	H-S	1.3302	1.3522	1.3250	1.3251
	C-C	1.5276	1.5137	1.5204	1.5090
	H-C	1.0919	1.0809	1.0831	1.0848
	S-C-C	114.03	111.00	112.01	113.03
	H-S-C	95.63	97.78	97.66	97.99
	H-C-C	109.67	112.48	111.07	110.80
	S-C-H	108.17	107.26	108.54	106.77
SH synclinal	S-C	1.8165	1.8992	1.8333	1.8331
	H-S	1.3315	1.3544	1.3275	1.3277
s	C-C	1.5260	1.5110	1.5168	1.5065
a	C-C	1.5253	1.5143	1.5220	1.5085
	H-C	1.0938	1.0800	1.0825	1.0853
s	S-C-C	114.37	111.83	112.70	114.55
a	S-C-C	111.23	107.97	108.71	109.98
	H-S-C	94.96	97.33	97.16	97.45
s	H-C-C	110.36	113.48	112.24	111.29
a	H-C-C	110.25	112.82	111.34	110.64
	S-C-H	109.23	107.97	109.56	107.83
	H-S-C-H	58.98	55.41	54.33	54.92
SH anticlinal	S-C	1.8235	1.9112	1.8468	1.8450
	H-S	1.3298	1.3500	1.3235	1.3243
s	C-C	1.5242	1.5133	1.5190	1.5072
a	C-C	1.5273	1.5133	1.5202	1.5083
	H-C	1.0940	1.0790	1.0821	1.0848
s	S-C-C	114.43	112.25	113.02	114.20
a	S-C-C	112.69	109.42	110.80	112.18
	H-S-C	96.09	98.26	97.91	98.23
s	H-C-C	110.09	112.90	111.30	110.60
a	H-C-C	109.84	113.08	111.43	110.68
	S-C-H	108.50	106.71	108.28	106.83
	H-S-C-H	113.03	108.40	111.33	111.31

Table 5: Selected geometric parameters of 2,4-cyclopentadiene-1-selenol

		Basis Set		
Subst.	Param.	STO-3G	3-21G	Huz
SeH antiperiplanar	Se-C	1.9519	1.9802	1.9784
	H-Se	1.4409	1.4716	1.4640
	C-C	1.5238	1.5107	1.5030
	H-C	1.0915	1.0835	1.0851
	Se-C-C	114.33	113.66	113.68
	H-Se-C	94.98	94.90	95.47
	H-C-C	110.77	111.58	112.06
	Se-C-H	105.24	103.98	102.71
SeH synperiplanar	Se-C	1.9594	1.9894	1.9906
	H-Se	1.4406	1.4742	1.4633
	C-C	1.5244	1.5123	1.5041
	H-C	1.0913	1.0824	1.0827
	Se-C-C	113.92	112.73	112.42
	H-Se-C	94.72	95.45	95.97
	H-C-C	110.52	111.29	112.15
	Se-C-H	106.58	106.43	105.08
SeH synclinal	Se-C	1.9562	1.9863	1.9826
	H-Se	1.4415	1.4760	1.4662
s	C-C	1.5228	1.5085	1.5010
a	C-C	1.5228	1.5136	1.5042
	H-C	1.0921	1.0808	1.0823
s	Se-C-C	113.62	112.78	113.21
a	Se-C-C	111.22	108.75	109.06
	H-Se-C	94.29	95.23	95.65
s	H-C-C	111.29	112.83	113.01
a	H-C-C	111.12	112.16	112.35
	Se-C-H	108.05	107.77	106.43
	H-Se-C-H	59.48	56.11	54.71
SeH anticlinal	Se-C	1.9627	1.9963	1.9940
	H-Se	1.4400	1.4704	1.4612
s	C-C	1.5222	1.5115	1.5025
a	C-C	1.5240	1.5121	1.5039
	H-C	1.0919	1.0804	1.0822
s	Se-C-C	113.31	112.89	112.99
a	Se-C-C	112.38	111.04	111.57
	H-Se-C	95.47	96.23	96.55
s	H-C-C	111.09	111.86	112.07
a	H-C-C	110.86	112.08	112.18
	Se-C-H	107.75	106.72	105.48
	H-Se-C-H	111.97	110.84	113.10

Table 6: Selected geometric parameters of 2,4-cyclopentadiene-1-tellurol

		Basis Set		
Subst.	Param.	STO-3G	3-21G	Huz
TeH antiperiplanar	Te-C	2.1565	2.1933	2.1957
	H-Te	1.6232	1.6775	1.6694
	C-C	1.5204	1.5105	1.4993
	H-C	1.0905	1.0842	1.0850
	Te-C-C	114.40	113.19	112.64
	H-Te-C	94.24	93.72	94.21
	H-C-C	110.73	112.06	113.21
	Te-C-H	105.14	103.96	102.34
TeH synperiplanar	Te-C	2.1617	2.1992	2.2029
	H-Te	1.6242	1.6834	1.6708
	C-C	1.5205	1.5116	1.5000
	H-C	1.0906	1.0835	1.0833
	Te-C-C	114.54	112.92	112.53
	H-Te-C	93.08	93.33	93.96
	H-C-C	110.55	111.83	112.92
	Te-C-H	105.26	105.00	103.22
TeH synclinal	Te-C	2.1616	2.1990	2.1987
	H-Te	1.6242	1.6836	1.6733
s	C-C	1.5182	1.5056	1.4959
a	C-C	1.5197	1.5131	1.5001
	H-C	1.0902	1.0805	1.0817
s	Te-C-C	113.12	111.79	112.22
a	Te-C-C	111.35	108.08	108.46
	H-Te-C	93.14	93.45	94.17
s	H-C-C	111.75	113.82	114.18
a	H-C-C	111.59	113.34	113.69
	Te-C-H	107.50	107.12	105.27
	H-Te-C-H	60.97	56.55	55.49
TeH anticlinal	Te-C	2.1670	2.2089	2.2097
	H-Te	1.6231	1.6777	1.6679
s	C-C	1.5191	1.5106	1.4991
a	C-C	1.5199	1.5107	1.4997
	H-C	1.0899	1.0804	1.0817
s	Te-C-C	112.45	111.26	111.57
a	Te-C-C	112.63	111.02	111.59
	H-Te-C	94.49	94.90	95.42
s	H-C-C	111.46	112.85	113.15
a	H-C-C	111.38	113.09	113.30
	Te-C-H	107.45	106.22	104.48
	H-Te-C-H	114.50	114.81	116.14

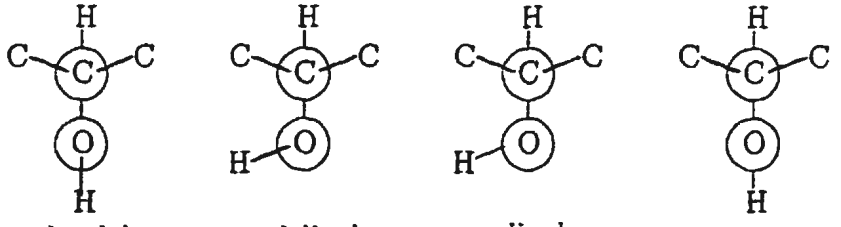


Figure 3: Conformation of 5-chalcosubstituted cyclopentadienes

of including polarization functions on sulfur is evident[143]. The bond angles are somewhat more sensitive to basis set than was the case for oxygen. For CpSeH, there is some change upon proceeding from the minimal to the split valence basis set, but the addition of polarization functions is not nearly as important, as was the case for sulfur[144]. Basis set trends for CpTeH are similar to CpSeH. The X-H and X-C bond lengths, and H-X-C angles are similar to those calculated before at the STO-3G[145] and 3-21G levels[146, 147] for the H_2X and CH_3XH species. It should be noted that there are relatively few ab initio calculations for organic selenium and tellurium compounds[146, 148, 149].

The conformation of chalcogenocyclopentadienes is shown in Figure 3. The conformational changes resemble in many respects that presented in Reference [147]. In general, when there is an eclipsing interaction between the X-H bond and either a C-C or C-H bond, the X-H bond shortens (by up to 0.006 Å) and the connecting X-C bond lengthens (by up to 0.017 Å). In general the H-X-C bond angle increases by at most 1.6° upon eclipsing, but as one goes down a group this becomes smaller, especially when comparing the two asymmetric structures. The X-C-H bond angle is

anomalously small in the staggered conformers (by up to 5°), but this becomes less pronounced as one progresses down the group. The H-C-C angle changes by at most 2° but in no obvious pattern. The X-C-C angle undergoes large changes of up to 5°, with the smallest angle consistently being for the gauche conformer, in which the C-C bond is trans to the X-H bond. For CpOH, the largest X-C-C angle is found for the eclipsed gauche conformer, in which the C-C bond is cis to the X-H bond, whereas for the others the staggered conformer gives the largest angle. For the larger atoms, the relation between torsion and X-C-C is fairly flat near the maximum X-C-C angle.

5-Pnictocyclopentadienes

None of the parent pnictocyclopentadienes are known experimentally. Several substituted derivatives are known, such as the pentamethylcyclopentadienyl amine[150, 141, 142], phosphine[151], and dichlorostibine[152], and cyclopentadienyl difluorophosphine[153, 154, 155]. In addition, a MNDO study of the [1,5]-sigmatropic rearrangement involving the motion of the phosphorus group has been carried out[156, 157].

In general, the adjacent C-C bond shortens, the X-C-C and X-C-H angles become smaller, and the H-C-C angle becomes larger as one goes down this group. The trend is the same as that for the chalcosubstituted cyclopentadienes, but unlike these, the angular parameters are not disposed predominantly to one side of those of CpH. The parameters of $CpPH_2$ most closely resemble CpH.

Table 7: Selected geometric parameters of 2,4-cyclopentadien-1-amine

Subst.	Param.	Basis Set			
		STO-3G	3-21G	3-21G(N)	6-31G*
NH_2 antiperiplanar	N-C	1.4951	1.4738	1.4729	1.4585
	H-N	1.0343	1.0046	1.0113	1.0023
	C-C	1.5309	1.5238	1.5256	1.5119
	H-C	1.1003	1.0927	1.0928	1.0962
	N-C-C	112.28	111.59	111.81	112.85
	H-N-C	106.88	113.20	108.44	110.19
	H-N-H	104.31	111.11	105.69	106.64
	H-C-C	109.65	108.91	108.65	108.12
	N-C-H	111.38	113.61	113.74	112.42
	H-N-C-H	55.61	63.80	57.15	58.70
NH_2 synperiplanar	N-C	1.5043	1.4852	1.4888	1.4689
	H-N	1.0325	1.0029	1.0090	1.0001
	C-C	1.5318	1.5264	1.5278	1.5138
	N-C-C	114.75	113.68	113.73	115.13
	H-N-C	107.60	112.99	108.28	110.80
	H-N-H	104.16	110.80	105.16	106.61
	N-C-H	107.77	109.73	109.20	108.32
	H-N-C-H	124.15	116.60	123.24	120.94

Table 8: Selected geometric parameters of 2,4-cyclopentadien-1-amine

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(N)	6-31G*
NH_2 synclinal	N-C	1.4922	1.4692	1.4685	1.4558
s	H-N	1.0339	1.0049	1.0109	1.0026
a	H-N	1.0342	1.0046	1.0110	1.0021
s	C-C	1.5307	1.5231	1.5247	1.5118
a	C-C	1.5356	1.5324	1.5319	1.5187
s	N-C-C	112.25	111.68	111.60	112.54
a	N-C-C	116.48	117.23	117.50	117.94
s	H-N-C	107.19	112.10	107.73	109.95
a	H-N-C	107.04	113.46	108.94	110.57
	H-N-H	104.50	111.33	106.32	107.12
	N-C-H	107.02	108.12	108.19	107.08
s	H-N-C-H	185.40	178.80	183.75	182.77
a	H-N-C-H	-62.94	-54.04	-61.41	-59.15
NH_2 anticlinal	N-C	1.4994	1.4688	1.4755	1.4618
s	H-N	1.0315	1.0013	1.0085	0.9993
a	H-N	1.0313	0.9997	1.0073	0.9990
s	C-C	1.5352	1.5261	1.5275	1.5148
a	C-C	1.5316	1.5266	1.5281	1.5140
s	N-C-C	113.01	113.26	112.93	113.93
a	N-C-C	114.73	113.13	114.54	115.11
s	H-N-C	108.10	115.19	109.77	111.68
a	H-N-C	108.11	115.05	109.47	111.57
	H-N-H	104.82	113.17	106.76	107.80
	N-C-H	109.73	112.16	111.90	110.39
s	H-N-C-H	-3.03	-31.65	-14.42	-12.34
a	H-N-C-H	109.93	102.75	102.45	108.38

Table 9: Selected geometric parameters of 2,4-cyclopentadiene-1-phosphine

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(*)	6-31G*
PH_2 antiperiplanar	P-C	1.8638	1.9242	1.8679	1.8797
	H-P	1.3819	1.4286	1.4054	1.4060
	C-C	1.5207	1.5117	1.5179	1.5044
	H-C	1.0908	1.0820	1.0850	1.0860
	P-C-C	112.10	108.50	110.34	110.23
	H-P-C	95.46	97.24	97.12	97.90
	H-P-H	93.52	94.94	93.72	94.56
	H-C-C	110.74	113.40	111.44	112.13
	P-C-H	109.79	110.34	111.08	109.67
	H-P-C-H	47.04	47.97	47.34	47.88
PH_2 synperiplanar	P-C	1.8736	1.9396	1.8895	1.8968
	H-P	1.3807	1.4216	1.3989	1.4014
	C-C	1.5213	1.5134	1.5188	1.5067
	P-C-C	114.21	112.33	113.43	114.29
	H-P-C	96.88	98.59	98.73	99.22
	H-P-H	93.31	95.13	94.06	94.80
	P-C-H	107.31	106.29	107.29	105.65
	H-P-C-H	132.90	131.72	132.23	131.78

Table 10: Selected geometric parameters of 2,4-cyclopentadiene-1-phosphine

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(*)	6-31G*
PH_2 synclinal	P-C	1.8614	1.9222	1.8716	1.8776
s	H-P	1.3811	1.4216	1.3993	1.4024
a	H-P	1.3817	1.4263	1.4035	1.4049
s	C-C	1.5212	1.5142	1.5204	1.5072
a	C-C	1.5204	1.5089	1.5132	1.5031
s	P-C-C	112.31	109.01	109.34	110.93
a	P-C-C	115.19	112.79	113.42	115.49
s	H-P-C	96.28	97.02	97.20	98.19
a	H-P-C	95.71	97.75	97.94	98.29
	H-P-H	93.77	95.91	94.90	95.49
	P-C-H	107.11	106.90	107.96	105.64
s	H-P-C-H	166.79	169.26	168.72	168.04
a	H-P-C-H	72.35	72.28	72.72	71.22
PH_2 anticlinal	P-C	1.8728	1.9357	1.8870	1.8944
s	H-P	1.3811	1.4267	1.4029	1.4042
a	H-P	1.3805	1.4229	1.4000	1.4017
s	C-C	1.5234	1.5152	1.5215	1.5084
a	C-C	1.5214	1.5143	1.5192	1.5070
s	P-C-C	112.80	110.69	111.26	112.38
a	P-C-C	114.59	112.59	112.98	114.17
s	H-P-C	96.17	97.92	98.09	98.57
a	H-P-C	96.73	98.63	98.81	99.13
	H-P-H	93.19	94.98	93.76	94.42
	P-C-H	108.46	108.12	109.11	107.24
s	H-P-C-H	10.88	8.51	10.36	9.86
a	H-P-C-H	104.85	104.82	105.46	105.78

Table 11: Selected geometric parameters of 2,4-cyclopentadien-1-arsine

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
AsH_2 antiperiplanar	As-C	1.9612	1.9836	2.0052
	H-As	1.4588	1.5308	1.5219
	C-C	1.5194	1.5094	1.4982
	H-C	1.0908	1.0840	1.0831
	As-C-C	113.24	111.51	109.15
	H-As-C	94.85	95.47	96.01
	H-As-H	93.53	94.28	92.84
	H-C-C	110.27	111.37	113.84
	As-C-H	108.45	108.76	108.01
	H-As-C-H	46.98	47.42	46.75
AsH_2 synperiplanar	As-C	1.9706	1.9990	2.0203
	H-As	1.4566	1.5212	1.5155
	C-C	1.5205	1.5112	1.5012
	As-C-C	114.59	114.03	113.39
	H-As-C	96.68	97.89	97.53
	H-As-H	93.68	94.69	93.21
	As-C-H	107.08	105.45	104.23
	H-As-C-H	132.74	132.06	132.87

Table 12: Selected geometric parameters of 2,4-cyclopentadien-1-arsine

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
AsH_2 synclinal	As-C	1.9618	1.9871	2.0021
s	H-As	1.4575	1.5212	1.5159
a	H-As	1.4582	1.5272	1.5204
s	C-C	1.5196	1.5118	1.5016
a	C-C	1.5189	1.5054	1.4975
s	As-C-C	112.32	109.90	110.16
a	As-C-C	114.47	113.85	114.13
s	H-As-C	95.99	96.38	96.34
a	H-As-C	95.40	96.60	96.49
	H-As-H	93.95	95.36	93.92
	As-C-H	106.93	106.29	104.64
s	H-As-C-H	166.63	168.40	166.97
a	H-As-C-H	72.08	72.22	72.27
AsH_2 anticlinal	As-C	1.9726	1.9991	2.0187
s	H-As	1.4576	1.5283	1.5197
a	H-As	1.4569	1.5234	1.5163
s	C-C	1.5211	1.5129	1.5027
a	C-C	1.5200	1.5109	1.5007
s	As-C-C	113.07	111.38	111.14
a	As-C-C	114.03	113.46	113.11
s	H-As-C	95.69	96.56	96.52
a	H-As-C	96.44	97.76	97.50
	H-As-H	93.39	94.17	92.72
	As-C-H	107.68	107.46	105.79
s	H-As-C-H	10.71	10.23	12.07
a	H-As-C-H	104.79	105.37	105.70

Table 13: Selected geometric parameters of 2,4-cyclopentadiene-1-stibine

		Basis Set		
Subst.	Saram.	STO-3G	3-21G	6-31G*
SbH_2 antiperiplanar	Sb-C	2.1632	2.2035	2.2158
	H-Sb	1.6444	1.7344	1.7277
	C-C	1.5143	1.5021	1.4909
	H-C	1.0889	1.0805	1.0825
	Sb-C-C	111.91	106.85	107.61
	H-Sb-C	94.29	94.39	94.85
	H-Sb-H	93.88	92.84	92.40
	H-C-C	111.84	115.49	115.66
	Sb-C-H	107.95	108.80	106.75
	H-Sb-C-H	47.11	46.60	46.41
SbH_2 synperiplanar	Sb-C	2.1722	2.2145	2.2288
	H-Sb	1.6432	1.7248	1.7196
	C-C	1.5157	1.5075	1.4961
	Sb-C-C	112.80	111.47	111.98
	H-Sb-C	96.17	96.80	96.92
	H-Sb-H	93.95	93.37	93.27
	Sb-C-H	107.50	105.70	103.89
	H-Sb-C-H	132.67	132.88	132.92

Table 14: Selected geometric parameters of 2,4-cyclopentadiene-1-stibine

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
SbH_2 synclinal	Sb-C	2.1628	2.2024	2.2146
s	H-Sb	1.6434	1.7232	1.7190
a	H-Sb	1.6443	1.7321	1.7259
s	C-C	1.5155	1.5079	1.4962
a	C-C	1.5144	1.5014	1.4917
s	Sb-C-C	112.17	108.08	109.03
a	Sb-C-C	113.39	111.26	112.17
s	H-Sb-C	95.48	95.02	95.40
a	H-Sb-C	94.70	95.10	95.43
	H-Sb-H	94.15	93.77	93.64
	Sb-C-H	106.89	106.64	104.39
s	H-Sb-C-H	166.94	167.44	166.81
a	H-Sb-C-H	72.31	73.19	72.62
SbH_2 anticlinal	Sb-C	2.1709	2.2128	2.2272
s	H-Sb	1.6441	1.7339	1.7254
a	H-Sb	1.6432	1.7276	1.7209
s	C-C	1.5164	1.5086	1.4963
a	C-C	1.5159	1.5068	1.4950
s	Sb-C-C	113.26	109.94	110.50
a	Sb-C-C	113.25	111.07	111.58
s	H-Sb-C	94.65	94.73	94.88
a	H-Sb-C	95.90	96.62	96.87
	H-Sb-H	93.67	92.48	92.34
	Sb-C-H	106.78	106.59	104.41
s	H-Sb-C-H	11.62	15.32	17.28
a	H-Sb-C-H	105.81	108.38	110.23

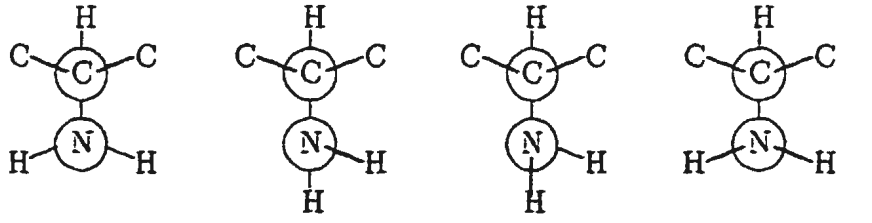


Figure 4: Conformation of 5-pnictoosubstituted cyclopentadienes

As one progresses along our basis set series, some clear trends emerge. The C-N bond length decreases as one increases the sp portion of the basis set, whereas the C-P bond length behavior is similar to that of C-S. The C-As and C-Sb bond lengths all increase with increasing basis set size. There is a large increase in X-H bond length between the STO-3G and 3-21G basis set (decrease for N-H), and a smaller decrease between 3-21G and 6-31G*. Polarization functions on nitrogen do not affect the N-C or N-H bond lengths much (3-21G to 3-21G(N)). The angles about the heteroatom are not sensitive to basis set, with the exception of nitrogen, for which the 3-21G basis set predicts an amino group which is too planar.

The conformation of pnictocyclopentadienes is shown in Figure 4, in which the conformational designation refers to the pnicogen lone pair. If there exists an eclipsing interaction with an X-H bond, then the X-C bond lengthens by up to 0.02 Å, and the X-H bond shortens by as much as 0.008 Å in some cases. An eclipsing interaction can also increase the H-X-C angle by up to 2°. The X-C-H angle is anomalously *larger* in the staggered conformation by up to 5°. The largest X-C-C angle always occurs in the gauche structure, in which the nitrogen lone pair is trans to the C-C bond in

question. The smallest X-C-C angle usually occurs in the staggered structure.

Group IV Cyclopentadienes

The group IV cyclopentadienes have been the most extensively studied of our series. 5-Methylcyclopentadiene was first separated from its more stable isomers by Csicsery[158] using vapor-phase chromatography, and it constituted about 3 % of the total equilibrium mixture. It can also be prepared by the methylation of CpMgBr at 263 K, but this isomerizes to the 1-isomer at 298 K in about 3 hours, and further isomerization to an equilibrium mixture of the 1- and 2-isomers occurs within a couple of days[159]. The isomerization was shown to proceed by a 1,2-hydrogen shift and an activation barrier of 83.3 $\pm$ 1.3 kJ/mol was determined from NMR[160, 161]. Evidence for methyl migration was also seen in trimethylcyclopentadienes, with a barrier of about 185 kJ/mol[162], which agreed well with the MINDO/3 result of 194.3 kJ/mol for CpMe[163].

Silylcyclopentadiene was first prepared by the reaction of potassium cyclopentadienide with silyl bromide at 77 K[164]. The structure was determined by electron diffraction soon after[165]. $CpSiH_3$ undergoes a rapid 1,2-silyl migration (248 K) and a slower 1,2-hydrogen shift (339 K)[166]. A joint photoelectron and ab initio study suggested that the fluxional behavior of $CpSiH_3$ may be related to the delocalization of the Si-C bonding orbital with the π-system[167]. The activation barrier for silyl

Table 15: Selected geometric parameters of 5-methyl-1,3-cyclopentadiene

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
CH_3 stag	C-C	1.5476	1.5458	1.5346
i[a]	H-C	1.0859	1.0826	1.0847
o	H-C	1.0863	1.0843	1.0856
	C-C	1.5272	1.5202	1.5101
	H-C	1.0938	1.0870	1.0911
	C-C-C	113.59	111.86	113.81
i	H-C-C	110.74	109.59	110.62
o	H-C-C	110.38	110.67	110.89
io	H-C-H	108.46	108.82	108.25
oo	H-C-H	108.34	108.23	107.81
	H-C-C	110.01	110.64	109.36
	C-C-H	108.38	109.99	108.50
	H-C-C-H	59.87	59.99	59.87
CH_3 ecli	C-C	1.5597	1.5620	1.5511
i	H-C	1.0854	1.0832	1.0843
o	H-C	1.0854	1.0819	1.0837
	C-C	1.5285	1.5226	1.5120
	H-C	1.0935	1.0875	1.0905
	C-C-C	114.47	113.83	115.14
i	H-C-C	110.85	111.20	111.49
o	H-C-C	111.04	110.72	111.38
io	H-C-H	107.96	108.17	107.56
oo	H-C-H	107.85	107.73	107.27
	H-C-C	109.05	108.90	108.01
	C-C-H	108.60	109.67	108.58
	H-C-C-H	120.01	120.29	120.15

[a] 'i' refers to the hydrogen contained in the symmetry plane, 'o' refers to the hydrogens out of the plane

Table 16: Selected geometric parameters of 2,4-cyclopentadiene-1-silane

		Basis Set			
Subst.	Param.	STO-3G	3-21G	3-21G(*)	6-31G*
SiH_3 stag	Si-C	1.8815	1.9355	1.9012	1.9084
i	H-Si	1.4219	1.4834	1.4696	1.4735
o	H-Si	1.4227	1.4900	1.4758	1.4786
	C-C	1.5152	1.5088	1.5119	1.4995
	H-C	1.0897	1.0848	1.0859	1.0891
	Si-C-C	112.46	108.18	107.79	109.94
i	H-Si-C	110.14	109.57	109.71	110.33
o	H-Si-C	109.42	110.03	110.11	109.89
io	H-Si-H	109.39	109.43	109.37	109.30
oo	H-Si-H	109.06	108.34	108.15	108.11
	H-C-C	111.25	114.09	114.00	113.35
	Si-C-H	108.15	109.55	110.58	107.72
	H-Si-C-H	59.72	59.65	59.58	59.42
SiH_3 ecli	Si-C	1.8906	1.9448	1.9134	1.9231
i	H-Si	1.4229	1.4912	1.4762	1.4788
o	H-Si	1.4223	1.4863	1.4719	1.4752
	C-C	1.5190	1.5141	1.5177	1.5041
	H-C	1.0898	1.0864	1.0879	1.0896
	Si-C-C	113.47	111.42	111.37	112.60
i	H-Si-C	109.39	109.85	109.73	109.53
o	H-Si-C	110.46	111.01	111.11	111.20
io	H-Si-H	108.75	108.08	108.05	108.03
oo	H-Si-H	108.99	108.71	108.68	108.75
	H-C-C	110.42	111.99	111.73	111.58
	Si-C-H	107.93	108.09	108.85	106.53
	H-Si-C-H	119.67	119.48	119.43	119.33

Table 17: Selected geometric parameters of 2,4-cyclopentadiene-1-germane

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
GeH_3 stag	Ge-C	1.9368	1.9693	1.9922
i	H-Ge	1.4323	1.5361	1.5379
o	H-Ge	1.4333	1.5438	1.5441
	C-C	1.5170	1.5045	1.4956
	H-C	1.0895	1.0842	1.0867
	Ge-C-C	112.61	109.04	109.64
i	H-Ge-C	109.51	109.29	109.88
o	H-Ge-C	108.76	109.45	109.51
io	H-Ge-H	110.03	109.85	109.74
oo	H-Ge-H	109.73	108.95	108.43
	H-C-C	111.10	113.80	114.14
	Ge-C-H	108.18	108.52	106.49
	H-Ge-C-H	59.73	59.67	59.39
GeH_3 ecli	Ge-C	1.9459	1.9769	2.0045
i	H-Ge	1.4336	1.5460	1.5444
o	H-Ge	1.4323	1.5394	1.5398
	C-C	1.5184	1.5093	1.4994
	H-C	1.0894	1.0858	1.0869
	Ge-C-C	113.41	111.79	111.73
i	H-Ge-C	108.65	108.94	108.87
o	H-Ge-C	109.85	110.83	110.94
io	H-Ge-H	109.41	108.55	108.45
oo	H-Ge-H	109.63	109.08	109.12
	H-C-C	110.50	111.98	112.81
	Ge-C-H	107.90	107.30	105.53
	H-Ge-C-H	119.66	119.37	119.27

Table 18: Selected geometric parameters of 2,4-cyclopentadiene-1-stannane

		Basis Set		
Subst.	Param.	STO-3G	3-21G	6-31G*
SnH_3 stag	Sn-C	2.1442	2.1898	2.2134
i	H-Sn	1.6300	1.7341	1.7374
o	H-Sn	1.6314	1.7437	1.7456
	C-C	1.5102	1.4977	1.4872
	H-C	1.0884	1.0818	1.0848
	Sn-C-C	111.50	105.67	106.93
i	H-Sn-C	109.18	109.30	109.72
o	H-Sn-C	108.13	109.39	109.25
io	H-Sn-H	110.58	110.08	110.07
oo	H-Sn-H	110.15	108.58	108.47
	H-C-C	112.26	116.23	116.22
	Sn-C-H	107.96	109.18	106.65
	H-Sn-C-H	59.63	59.41	59.25
SnH_3 ecli	Sn-C	2.1500	2.1944	2.2206
i	H-Sn	1.6320	1.7453	1.7460
o	H-Sn	1.6305	1.7382	1.7401
	C-C	1.5117	1.5035	1.4913
	H-C	1.0885	1.0834	1.0854
	Sn-C-C	112.33	108.69	109.16
i	H-Sn-C	107.73	108.60	108.19
o	H-Sn-C	109.43	110.86	110.86
io	H-Sn-H	109.97	108.52	108.71
oo	H-Sn-H	110.27	109.41	109.46
	H-C-C	111.71	114.55	115.04
	Sn-C-H	107.55	107.61	105.39
	H-Sn-C-H	119.54	119.14	119.11

migration was determined both from proton (59.4 ± 1.0 kJ/mol) and ^{13}C NMR lineshape analysis (57.7 ± 0.1 kJ/mol)[168]. An MNDO study overestimates this barrier (99.1 kJ/mol)[169].

Germylcyclopentadiene was prepared in a similar fashion to silylcyclopentadiene and the NMR behavior suggested a fluxional molecule[170]. The X-ray and electron diffraction structure of this compound has been published[171]. The trimethyl derivative undergoes a 1,2-shift with a barrier of 38.5 ± 4.2 kJ/mol, and the corresponding stannane, 32.6 ± 4.2 kJ/mol[172].

Tables 15, 16, 17, and 18 show that, upon substitution, the C-C bond length becomes slightly longer in the case of $CpCH_3$, but becomes shorter for the higher analogues. The X-C-C and X-C-H angles decrease upon going down a group, with only the $CpCH_3$ angle being larger than cyclopentadiene, whereas the H-C-C angle increases.

The basis set trend is similar to that of the pnictogens, with the exception that polarization functions on the substituent do not make a significant difference in the bond angles, presumably due to the isotropic environment about the pseudo-C_3 axis along the X-C bond.

The existence of only two conformational possibilities simplifies our analysis considerably. The X-C and C-C bonds in the eclipsed conformer are longer than in the staggered by as much as 0.015 Å and 0.005 Å, respectively, whereas no clear trend exists for the H-X bond length. The X-C-C and H-C-C increase and decrease by

Table 19: Conformational energies of chalcocyclopentadienes (kJ/mol)

		Basis Set			
System	Energy	STO-3G	3-21G	6-31G*//3-21G	6-31G*
CpOH	$E_{stag \rightarrow eg}$	12.42	17.58	8.67	13.36
	$E_{stag \rightarrow gau}$	6.18	16.48	10.45	9.96
	$E_{stag \rightarrow ecli}$	13.84	21.49	14.18	13.91
	$E_{gau \rightarrow eg}$	6.24	1.10	-1.78	3.40
	$E_{gau \rightarrow ecli}$	7.66	5.01	3.73	3.95
CpSH[a]	$E_{stag \rightarrow eg}$	10.77	13.82	13.25	13.14
	$E_{stag \rightarrow gau}$	7.18	7.65	5.33	5.05
	$E_{stag \rightarrow ecli}$	10.08	12.59	10.03	10.32
	$E_{gau \rightarrow eg}$	3.59	6.17	7.91	8.09
	$E_{gau \rightarrow ecli}$	2.90	4.94	4.69	5.27
CpSeH	$E_{stag \rightarrow eg}$	8.39	12.70	10.70	10.65
	$E_{stag \rightarrow gau}$	4.23	5.55	2.99	2.86
	$E_{stag \rightarrow ecli}$	8.30	9.92	7.59	7.46
	$E_{gau \rightarrow eg}$	4.16	7.15	7.71	7.79
	$E_{gau \rightarrow ecli}$	4.07	4.37	4.60	4.60
CpTeH	$E_{stag \rightarrow eg}$	5.45	7.93	7.29	7.27
	$E_{stag \rightarrow gau}$	2.78	1.32	-0.04	-0.09
	$E_{stag \rightarrow ecli}$	5.83	6.07	4.11	3.87
	$E_{gau \rightarrow eg}$	2.67	6.61	7.33	7.36
	$E_{gau \rightarrow ecli}$	3.05	4.75	4.15	3.96

[a] 3-21G(*): $E_{stag \rightarrow eg} = 15.32$, $E_{stag \rightarrow gau} = 5.95$, $E_{stag \rightarrow ecli} = 11.55$, $E_{gau \rightarrow eg} = 9.37$, $E_{gau \rightarrow ecli} = 5.60$; 6-31G*//3-21G(*): $E_{stag \rightarrow eg} = 12.86$, $E_{stag \rightarrow gau} = 4.94$, $E_{stag \rightarrow ecli} = 10.12$, $E_{gau \rightarrow eg} = 7.92$, $E_{gau \rightarrow ecli} = 5.18$

as much as 2.7° and 1.8°, respectively, upon eclipsing. The X-C-H angle always decreases (except for $CpCH_3$) by up to 1.4°. No persistent trend for the angles about the substituent emerges.

3.1.2 Conformation Energies

Chalcocyclopentadienes

Table 20: Barriers to internal rotation in chalcomethanes (kJ/mol)

	Basis Set			
System	STO-3G	3-21G	3-21G(*)	6-31G*
MeOH	8.42	6.19	6.3	5.9
MeSH	6.09	4.54	5.72	5.9
MeSeH	4.31	4.29	4.56	4.57*
MeTeH	2.55*	2.78	2.85	3.01*

The barriers to rotation are shown in Table 19. The barriers are somewhat basis set dependent, but the 6-31G* and 6-31G*//3-21G (and 6-31G*//3-21G(*) for CpSH) are usually very close to each other. The anomalous difference between the rotation barriers for CpOH, as determined by 6-31G* and 6-31G*//3-21G is clearly a result of the poor description of the eclipsed gauche conformer by the 3-21G basis set. The addition of polarization functions to the sulfur 3-21G basis set improves significantly the energies relative to 6-31G*. The barriers $E_{gau \rightarrow ecli}$ are compared to the corresponding results for CH_3XH at STO-3G[146], 3-21G[147] and 6-31G*[145] (See Table 20. The '*' represents our calculated value). In general, the chalcomethane barriers are larger for O and S, similar for Se, and smaller for Te.

Pnictocyclopentadienes

The barriers to internal rotation in 5-pnictocyclopentadienes are shown in Table 21. In general the barriers are insensitive to basis set, with the following clarifications. Like for CpOH, the 3-21G basis set is inadequate for describing $E_{stag \rightarrow eg}$, because of the especially poor description of the eclipsed gauche conformer. At the 6-31G* level, the conformation barriers are insensitive to what geometry is used,

Table 21: Conformational energies of pnictocyclopentadienes (kJ/mol)

		Basis Set			
System	Energy	STO-3G	3-21G	6-31G*//3-21G	6-31G*
$CpNH_2$[a]	$E_{stag \rightarrow eg}$	8.28	1.25	5.61	5.52
	$E_{stag \rightarrow gau}$	-7.42	-17.09	-11.60	-10.79
	$E_{stag \rightarrow ecli}$	7.37	-0.90	4.94	5.67
	$E_{gau \rightarrow eg}$	15.70	18.34	17.21	16.31
	$E_{gau \rightarrow ecli}$	14.79	16.19	16.54	16.46
$CpPH_2$[b]	$E_{stag \rightarrow eg}$	8.62	11.48	12.24	12.67
	$E_{stag \rightarrow gau}$	-3.17	-4.31	-1.00	-1.01
	$E_{stag \rightarrow ecli}$	7.57	10.14	13.60	13.96
	$E_{gau \rightarrow eg}$	11.79	15.79	13.24	13.68
	$E_{gau \rightarrow ecli}$	10.74	14.45	14.60	14.97
$CpAsH_2$	$E_{stag \rightarrow eg}$	6.44	10.06	10.74	11.80
	$E_{stag \rightarrow gau}$	-3.44	-3.15	-0.22	1.01
	$E_{stag \rightarrow ecli}$	5.47	10.36	12.85	13.90
	$E_{gau \rightarrow eg}$	9.88	13.21	10.96	10.79
	$E_{gau \rightarrow ecli}$	8.91	13.51	13.07	12.89
$CpSbH_2$	$E_{stag \rightarrow eg}$	4.93	11.05	10.74	10.80
	$E_{stag \rightarrow gau}$	-1.80	1.93	3.89	3.91
	$E_{stag \rightarrow ecli}$	3.91	13.17	14.39	14.52
	$E_{gau \rightarrow eg}$	6.73	9.12	6.85	6.89
	$E_{gau \rightarrow ecli}$	5.71	11.24	10.50	10.61

[a] 3-21G(N): $E_{stag \rightarrow eg}$ = 4.76, $E_{stag \rightarrow gau}$ = -15.43, $E_{stag \rightarrow ecli}$ = 2.21, $E_{gau \rightarrow eg}$ = 20.19, $E_{gau \rightarrow ecli}$ = 17.64; 6-31G*//3-21G(N): $E_{stag \rightarrow eg}$ = 5.22, $E_{stag \rightarrow gau}$ = -11.03, $E_{stag \rightarrow ecli}$ = 6.60, $E_{gau \rightarrow eg}$ = 16.25, $E_{gau \rightarrow ecli}$ = 17.63

[b] 3-21G(*): $E_{stag \rightarrow eg}$ = 13.89, $E_{stag \rightarrow gau}$ = -2.35, $E_{stag \rightarrow ecli}$ = 14.86, $E_{gau \rightarrow eg}$ = 16.24, $E_{gau \rightarrow ecli}$ = 17.21; 6-31G*//3-21G(*): $E_{stag \rightarrow eg}$ = 12.53, $E_{stag \rightarrow gau}$ = -0.84, $E_{stag \rightarrow ecli}$ = 13.63, $E_{gau \rightarrow eg}$ = 13.37, $E_{gau \rightarrow ecli}$ = 14.47

Table 22: Barriers to internal rotation in pnictomethanes (kJ/mol)

	Basis Set			
System	STO-3G	3-21G	3-21G(*)	6-31G*
$MeNH_2$	11.7	8.4	8.4	10.0
$MePH_2$	7.9	7.1	8.4	8.4
$MeAsH_2$	6.08*	5.98*		6.37*
$MeSbH_2$	3.58*	3.50*		3.86*

Table 23: Conformational energies of group IV cyclopentadienes (kJ/mol)

	Basis Set			
System	STO-3G	3-21G	6-31G*//3-21G	6-31G*
$CpCH_3$[a]	14.55	17.97	17.37	17.62
$CpSiH_3$[b]	13.43	13.76	13.39	13.71
$CpGeH_3$[c]	6.80	12.27	12.65	11.11
$CpSnH_3$	4.37	9.77	7.94	8.05

[a] 6-31G**//3-21G = 17.28
[b] 3-21G(*) = 16.02, 6-31G*//3-21G(*) = 13.44, 6-31G**//3-21G = 13.37
[c] BC//3-21G = 12.95

whether optimized, or from 3-21G or 3-21G(*). The 3-21G(*) barriers are slightly greater than 6-31G*. The relative energy of the staggered to gauche, however, is very sensitive to basis set in this system, since the numbers involved are rather small, and can even switch sign. The preference for a staggered conformer increases as one goes down a period, even though $CpNH_2$ itself assumes the gauche conformer at 6-31G*. In general the barriers in substituted cyclopentadienes are larger than the corresponding pnictomethanes, as seen from Table 22.

Group IV Cyclopentadienes

From Table 23, the barriers are seen to be relatively insensitive to basis set, with the values agreeing quite well at the split-valence level. The barriers are larger than

Table 24: Barriers to internal rotation in group IV methanes (kJ/mol)

	Basis Set			
System	STO-3G	3-21G	3-21G(*)	6-31G*
$MeCH_3$	12.1	11.3	11.3	12.6
$MeSiH_3$	5.4	4.6	5.9	5.9
$MeGeH_3$	4.15*	4.31*		4.75*
$MeSnH_3$	2.12*	1.79*		2.14*

in the corresponding methyl derivatives (see Table 24).

3.1.3 Discussion of Conformation

There are numerous ways of explaining the conformational preferences about a σ-bond, but the approach used in this work will be that taken by Brunck and Weinhold[173, 174], which states that the predominant conformational preferences can be explained by bond-antibond interactions (hyperconjugation). Other studies[175] have used Fourier analysis, but are only strictly applicable to molecules possessing a rough three-fold symmetry axis and thus cannot be applied meaningfully to molecules involving the higher pnictogen series, in which the H-X-H angles are closer to 90° than 109.5°. The justification for our chosen approach will lie in the ease of explanation of the major geometric and energetic conformational trends.

If we consider the antiperiplanar form of CpOH, then we can count four trans vicinal hyperconjugative interactions, giving an interaction energy

$$E_{antiperiplanar} = 2E_{n,\sigma^*_{CC}} + E_{\sigma_{OH},\sigma^*_{CH}} + E_{\sigma_{CH},\sigma^*_{OH}}.$$

Similarly, the synclinal form of CpOH gives

$$E_{synclinal} = E_{n,\sigma^*_{CC}} + E_{n,\sigma^*_{CH}} + E_{\sigma_{OH},\sigma^*_{CC}} + E_{\sigma_{CC},\sigma^*_{OH}}.$$

The strongest interactions are normally those involving lone pairs donating into vicinal antibonds polarized toward the rotation axis, i.e., corresponding to vicinal bonds polarized away from the axis, such as C-F, since in this case the overlap is greatest. The C-H bond is polarized toward the carbon atom, which has the higher electronegativity, and thus its 'antibond' would be polarized toward the hydrogen, which results in poorer overlap with the 'lone pair', than that for a C-C antibond. In this case the antiperiplanar conformer will be favored. Because the overlap depends on the distance between the axis atoms, the difference in energy between these conformers should decrease as oxygen is successively replaced by sulfur, selenium and tellurium. which it does. In fact the two conformers of CpTeH are essentially degenerate.

If we consider the antiperiplanar form of $CpNH_2$, then we can count five trans interactions, giving an interaction energy

$$E_{antiperiplanar} = E_{n,\sigma^*_{CH}} + 2E_{\sigma_{NH},\sigma^*_{CC}} + 2E_{\sigma_{CC},\sigma^*_{NH}}.$$

Similarly, the synclinal form of $CpNH_2$ gives

$$E_{synclinal} = E_{n,\sigma^*_{CC}} + E_{\sigma_{NH},\sigma^*_{CC}} + E_{\sigma_{NH},\sigma^*_{CH}} + E_{\sigma_{CC},\sigma^*_{NH}} + E_{\sigma_{CH},\sigma^*_{NH}}.$$

For similar reasons, the synclinal form of $CpNH_2$ will be favored in this case. Because of the quicker drop in electronegativity and longer bond lengths in this case, the preference does not persist down the group and other terms start playing an important

role, but the difference is still very slight, being less than 4.0 kJ/mol. Steric interactions between the antiperiplanar hydrogen and the two vicinal C-C bonds in the larger substituents may favor the antiperiplanar arrangement because of the closer match in the lengths of these bonds.

The geometry changes also agree with the idea of hyperconjugation. The shorter X-H bond in the eclipsed forms is a result of cis-hyperconjugation being weaker than trans-hyperconjugation, and thus the antibond contribution is lower. The longer X-C bond may simply be explained by stronger steric interactions in the Lewis structure. In the Group IV series, hardly any trend is noticed, because the X-H antibond, being polarized predominantly toward the hydrogen, is a poor acceptor.

For CpOH, the widening of H-O-C in an eclipsed arrangement is consistent with a simple steric effect, especially with the C-H bond. As one goes down the group, the steric interaction with the C-C bonds become greater, leading to the largest angles in the synclinal form. The 'anomalous' smallness of the X-C-H angle in the antiperiplanar conformer can be explained simply as the absence of any strong lone pair donation into the C-H antibond, whereas all of the other conformers have either a cis or trans donation to C-H. For the same reason the smallest X-C-C angle will be that containing no lone pair donation, i.e., trans to the X-H bond.

For $CpNH_2$, the 'anomalous' largeness of the X-C-H angle in the antiperiplanar conformer is consistent with a strong hyperconjugation of the nitrogen lone pair with the C-H antibond. Similarly, the largest X-C-C angle occurs for the gauche structure,

in which the nitrogen lone pair is trans to the C-C bond in question.

3.2 1,2-Heterotropic Shifts in 5-Substituted Cyclopentadienes

5-Substituted cyclopentadienes can undergo a degenerate 1,2-heterotropic shift as shown in Figure 5. This process has been described in many ways, for example, (hydrogen, metal, carousel) migration; (prototropic, metallotropic, walk, intramolecular, [1,5]-sigmatropic, circumambulatory) rearrangement; fluxional behavior; and stereochemical non-rigidity, to name a few. The process has been demonstrated for many substituents,* as shown in Table 25. The following quotation from Larrabee's 1974 review[178] leads naturally into our investigation,

> ...it is not yet possible to predict the type of bonding a new structure will exhibit. It may ultimately be possible to make such predictions from large-scale numerical computations. Such bonding descriptions are to be considered incomplete if they do not include a description of the remarkable property found in many σ-bonded organometallic compounds: their fluxional behavior.

After discussing the optimization behavior, we discuss the geometry of the transition states, followed by a comparison between the predicted and experimental activation energies.

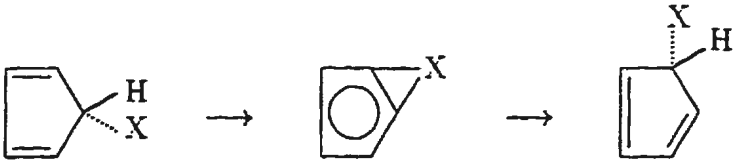


Figure 5: Degenerate 1,2-heterotropic shift of 5-substituted cyclopentadiene

Table 25: Experimental activation barriers for 1,2-heterotropic shift of 5-substituted cyclopentadienes

Substituent	Activation Barrier (kJ/mol)	Method	System
H	101.7 ± 2.1	^{1}H-VT NMR	-C_5D_5 [176]
CH_3	168.6-185.8	microflow/GLC	-$CpMe_2$ [162]
SiH_3	59.4 ± 1.0	^{1}H-VT NMR	-Cp[168]
	57.7 ± 0.1	^{13}C lineshape	-Cp[168]
$SiMe_3$	54.4 ± 4.2	^{1}H-VT NMR	-Cp[177]
$GeMe_3$	38.5 ± 4.2	^{1}H-VT NMR	-Cp[172]
$SnMe_3$	32.6 ± 4.2	^{1}H-VT NMR	-Cp[172]
PH_2	131.0 ± 7.5	^{1}H-VT NMR	-$CpMe_5$[151]
PMe_2	96.2 ± 2.1	^{1}H-VT NMR	-$CpMe_5$[151]

3.2.1 Geometries

In most cases, geometries were optimized within the C_s point group. For purposes of naming, the substituent is assumed to be migrating from C_5 to C_4.

Cyclopentadiene

Like the corresponding cyclopentadiene, the structure of the transition state for the 1,2-hydrogen shift changes little with basis set, with the angular range being about the same, but the bond range being larger (0.03 Å), predominantly about the reaction centers. The 6-31G** structure differed only in the last reported figure, except for the parameters about the reaction centers. The structures shown in Table 26 are in excellent agreement with those reported previously (3-21G [179] and 6-31G*[180]) and not much different than the MP2/6-31G* structure[180, 181]. The C-C bond lengths are between the single and double bond lengths of cyclopentadiene, as expected. One surprising feature is the shortening of the C-C bond bisected by the migrating carbon by over 0.03 Å, but this may be related to the lengthening of the migrating C-H bond by over 0.21 Å, making the reactant carbons more sp^2-like . The bond lengths of the hydrogens attached to the reacting carbons decrease by 0.017 Å to be very close to the usual sp^2 values. The C-C-C angles are within 2° of the idealized 108° of the cyclopentadienyl anion, and the non-migrating hydrogens are very nearly planar. The migrating H-C-C angle becomes smaller in the transition state.

Table 26: Geometry of the 1,2-hydrogen shift transition state

	Basis Set		
Parameter	STO-3G	3-21G	6-31G*
C_2-C_3	1.3874	1.3934	1.3903
C_3-C_4	1.3980	1.4002	1.4001
C_4-C_5	1.4971	1.4951	1.4703
C_2-H_2	1.0809	1.0690	1.0743
C_3-H_3	1.0775	1.0671	1.0719
C_4-H_4	1.0799	1.0672	1.0724
C_5-H'_4	1.3130	1.3319	1.3039
C_1-C_2-C_3	110.42	110.02	110.23
C_2-C_3-C_4	108.56	108.64	108.06
C_3-C_4-C_5	106.23	106.34	106.83
H_3-C_3-C_2	126.49	126.32	126.67
H_4-C_4-C_3	129.18	128.45	128.33
H'_4-C_4-C_5	55.24	55.86	55.68
H'_4-C_4-C_3	107.97	107.95	107.89

Halocyclopentadienes

The geometries of the 1,2-halogen shift transition states are given in Table 27. The 3-21G and 6-31G* structures are very similar whereas the STO-3G geometry underestimates the X-C forming/breaking bond lengths relative to the 6-31G* result by as much as 0.22 Å. There is not much difference between the geometry of the different basis sets in the carbon skeleton. Polarization functions on chlorine are shown to be important in describing the X-C bond length, as was shown for the reactant molecule.

The transition states are all remarkably similar in the carbon framework for all halogens, with the only major difference being in the X-C length. We can define a

Table 27: Geometry of the 1,2-halogen shift transition state

	Basis Set			
Parameter	STO-3G	3-21G	3-21G(*)	6-31G*
F-C	1.5943	1.7627		1.7770
C_1-C_2	1.3903	1.4015		1.4002
C_2-C_3	1.4094	1.3897		1.3852
C_3-C_4	1.5170	1.5200		1.4952
H-C_3	1.0885	1.0639		1.0685
F-C_3-C_4	61.59	64.46		65.12
F-C_3-C_2	112.10	113.40		114.26
H-C_3-C_2	127.29	127.49		127.38
Cl-C	2.1407	2.3410	2.2885	2.3061
C_1-C_2	1.3986	1.4067	1.4045	1.4029
C_2-C_3	1.3880	1.3757	1.3794	1.3754
C_3-C_4	1.5255	1.5350	1.5313	1.5089
H-C_3	1.0855	1.0651	1.0653	1.0689
Cl-C_3-C_4	69.13	70.86	70.45	70.90
Cl-C_3-C_2	113.78	113.18	113.08	115.04
H-C_3-C_2	127.71	128.31	128.16	127.69
Br-C	2.2117	2.3890		2.4321
C_1-C_2	1.3942	1.4023		1.4021
C_2-C_3	1.3945	1.3837		1.3767
C_3-C_4	1.5156	1.5229		1.5075
H-C_3	1.0841	1.0658		1.0692
Br-C_3-C_4	69.96	71.41		71.95
Br-C_3-C_2	114.52	113.89		115.09
H-C_3-C_2	126.85	127.65		127.55
I-C	2.4287	2.5885		2.6085
C_1-C_2	1.3946	1.4019		1.4003
C_2-C_3	1.3924	1.3846		1.3800
C_3-C_4	1.5161	1.5219		1.5047
H-C_3	1.0831	1.0667		1.0700
I-C_3-C_4	71.81	72.90		73.24
I-C_3-C_2	115.24	114.03		115.15
H-C_3-C_2	126.66	127.42		127.28

dimensionless measure of the stretching in the transition state as

$$\xi = \frac{r_{\text{transition state}} - r_{\text{reactant}}}{r_{\text{reactant}}},$$

where r is the X-C length (HF/6-31G*) at a particular point on the PES. The closer this value is to zero, the less 'stretched' the transition state is. For hydrogen, ξ = 0.1973. As one proceeds from fluorine to iodine, ξ becomes 0.2976, 0.2792, 0.2311, and 0.1854, showing that on a relative scale, the molecule does not need to stretch as much to reach the transition state. There is no trend seen in $r_{\text{transition state}} - r_{\text{reactant}}$. The C-C bond completing the triangle varies more than the corresponding bond in the reactant and shows no systematic trend, save that of being anomalously short for the fluorine case. The adjacent C_2-C_3 bond is longer in the fluorine species than the other halogen species.

Chalcocyclopentadienes

The geometries of the 1,2-chalcogen shift transition states are given in Table 28. The basis set trends are similar to those of the halogens. All the transition states are similar in the carbon framework, except for the anomalously short C_4-C_5 and long C_3-C_4 bonds for CpOH, which mirror the corresponding fluorine results. The stretching parameters ξ for this series are 0.2217, 0.2152, 0.1933, and 0.1600, showing that the parameter definition works well for the chalcogens as well.

Table 28: Geometry of the 1,2-chalcogen shift transition state

	Basis Set			
Parameter	STO-3G	3-21G	3-21G(*)	6-31G*
O-C	1.6114	1.7374		1.7121
C_2-C_3	1.3880	1.3972		1.3944
C_3-C_4	1.4199	1.4119		1.4113
C_4-C_5	1.5124	1.5044		1.4774
H-O	0.9936	0.9729		0.9527
O-C_4-C_5	62.01	64.34		64.44
O-C_4-C_3	113.16	113.86		115.20
H-O-C_4	105.07	105.46		105.37
S-C	2.0543	2.3026	2.2399	2.2238
C_2-C_3	1.3891	1.4002	1.3982	1.3952
C_3-C_4	1.4063	1.3926	1.3972	1.3961
C_4-C_5	1.4992	1.5097	1.5048	1.4836
H-S	1.3321	1.3492	1.3228	1.3244
S-C_4-C_5	68.60	70.86	70.37	70.51
S-C_4-C_3	115.07	113.28	113.20	115.40
H-S-C_4	95.40	92.87	92.23	94.20
Se-C	2.2253	2.3450		2.3609
C_2-C_3	1.3910	1.3981		1.3957
C_3-C_4	1.4004	1.3965		1.3937
C_4-C_5	1.5053	1.5015		1.4845
H-Se	1.4419	1.4666		1.4607
Se-C_4-C_5	70.23	71.33		71.68
Se-C_4-C_3	114.97	114.12		115.13
H-Se-C_4	94.35	92.10		92.41
Te-C	2.4338	2.5394		2.5471
C_2-C_3	1.3913	1.3980		1.3954
C_3-C_4	1.3980	1.3958		1.3935
C_4-C_5	1.5054	1.5015		1.4841
H-Te	1.6238	1.6726		1.6657
Te-C_4-C_5	115.26	113.46		114.39
Te-C_4-C_3	71.99	72.80		73.06
H-Te-C_4	93.37	91.28		91.57

Pnictocyclopentadienes

The geometries of the 1,2-pnictogen shift transition states are given in Table 29. Basis set trends are similar to those of the halogens. All the transition states are similar in the carbon framework, except for the anomalously *short* C_3-C_4 bond length in $CpNH_2$, opposite to the trend of F and OH. We emphasize that, to maintain C_s symmetry, the staggered conformation is enforced. The stretching parameters for this series are 0.3326, 0.1661, 0.1502, and 0.1091, but in this series there is a consistent downward trend in $\tau_{\text{transition state}} - \tau_{\text{reactant}}$ as well.

Group IV Cyclopentadienes

The geometries of the 1,2-Group IV shift transition states are given in Table 30. The basis set trends in this series are not as clear as before, except that polarization functions help the description of the Si-C and Si-H bond lengths. The C_4-C_5 bond is shorter in $CpCH_3$ than in the higher analogues. The stretching parameters for this series are 0.2706, 0.1320, 0.1355, 0.0973.

3.2.2 Activation Energies

The activation barriers for group migration are shown in Table 31. As a simplifying assumption, we assume that the activation barrier is independent of conformation. The difference between the staggered reactants and corresponding transition states is taken to be the activation barrier, regardless of whatever conformation the reactant

Table 29: Geometry of the 1,2-pnictogen shift transition state

	Basis Set			
Parameter	STO-3G	3-21G	3-21G(*)	6-31G*
N-C	1.8651	1.9454		1.9436
C_2-C_3	1.3909	1.3959		1.3938
C_3-C_4	1.3952	1.3931		1.3921
C_4-C_5	1.4813	1.4800		1.4615
H-N	1.0447	1.0140		1.0055
N-C_4-C_5	66.60	67.64		67.92
N-C_4-C_3	108.02	107.16		108.77
H-N-C_4	87.13	88.37		87.51
H-N-H	100.06	107.24		104.09
P-C	2.1509	2.2671	2.2014	2.1920
C_2-C_3	1.3892	1.3951	1.3940	1.3913
C_3-C_4	1.3978	1.3996	1.4041	1.4015
C_4-C_5	1.4797	1.4857	1.4781	1.4614
H-P	1.3872	1.4330	1.4082	1.4106
P-C_4-C_5	69.88	70.87	70.38	70.53
P-C_4-C_3	107.61	104.96	103.68	104.90
H-P-C_4	84.03	86.14	86.37	87.05
H-P-H	90.89	92.09	90.51	91.01
As-C	2.2378	2.2775		2.3064
C_2-C_3	1.3894	1.3948		1.3926
C_3-C_4	1.3980	1.4023		1.4007
C_4-C_5	1.4867	1.4793		1.4630
H-As	1.4628	1.5373		1.5302
As-C_4-C_5	70.60	71.05		71.51
As-C_4-C_3	108.94	104.81		104.65
H-As-C_4	84.12	86.66		86.47
H-As-H	91.20	90.60		89.81
Sb-C	2.3887	2.4463		2.4576
C_2-C_3	1.3891	1.3954		1.3928
C_3-C_4	1.3995	1.4047		1.4028
C_4-C_5	1.4895	1.4796		1.4635
H-Sb	1.6470	1.7412		1.7372
Sb-C_4-C_5	71.83	72.40		72.68
Sb-C_4-C_3	110.23	104.97		104.45
H-Sb-C_4	84.34	86.17		86.47
H-Sb-H	91.44	89.38		89.03

Table 30: Geometry of the 1,2-group IV shift transition state

	Basis Set			
Parameter	STO-3G	3-21G	3-21G(*)	6-31G*
C-C	1.8741	1.9865		1.9499
C_4-C_5	1.4755	1.4678		1.4479
H-C s	1.0838	1.0724		1.0744
H-C a	1.0894	1.0786		1.0799
C-C_4-C_5	66.82	68.32		68.21
C-C_4-C_3	109.76	107.96		109.87
H-C-C_4 s	105.79	100.27		102.83
H-C-C_4 a	92.01	91.11		91.34
H-C-H sa	109.54	112.69		111.93
H-C-H aa	105.88	109.66		108.25
Si-C	2.1108	2.2031	2.1582	2.1604
C_4-C_5	1.4797	1.4766	1.4714	1.4555
H-Si s	1.4197	1.4737	1.4597	1.4625
H-Si a	1.4267	1.4925	1.4787	1.4823
Si-C_4-C_5	69.48	70.42	70.07	70.31
Si-C_4-C_3	108.65	106.20	105.25	105.63
H-Si-C_4 s	105.61	104.06	105.13	105.52
H-Si-C_4 a	94.85	96.00	95.67	95.81
H-Si-H sa	109.97	110.93	111.02	111.03
H-Si-H aa	104.74	104.12	103.04	102.82
Ge-C	2.1891	2.2057		2.2618
C_4-C_5	1.4830	1.4737		1.4586
H-Ge s	1.4294	1.5248		1.5266
H-Ge a	1.4398	1.5509		1.5495
Ge-C_4-C_5	70.20	70.48		71.19
Ge-C_4-C_3	109.66	107.05		106.07
H-Ge-C_4 s	105.02	104.55		104.04
H-Ge-C_4 a	94.02	95.39		95.41
H-Ge-H sa	110.95	111.68		112.14
H-Ge-H aa	106.47	103.73		104.57
Sn-C		2.4031		2.4288
C_4-C_5		1.4763		1.4615
H-Sn s		1.7259		1.7274
H-Sn a		1.7482		1.7521
Sn-C_4-C_5		72.11		72.49
Sn-C_4-C_3		105.85		104.75
H-Sn-C_4 s		105.26		105.33
H-Sn-C_4 a		96.17		96.05
H-Sn-H sa		111.54		111.94
H-Sn-H aa		104.61		104.75

is most likely in. In the cases where the gauche is the preferred conformer in the reactant, we expect this assumed barrier to be slightly overestimated, because the energy of stabilization of one conformer relative to another would likely be lessened in the transition state, as the stabilization would inversely depend on some power of the distance between the substituent and the cyclopentadiene moiety. The error in the activation energy under these assumptions can only be as great as the *enthalpy* difference between the conformers. We do not view this as a problem since for most of these systems the activation barrier is either unknown or has a large margin of experimental error. The assumptions were tested on the eclipsed form of the chalcogen series at the STO-3G and 3-21G levels and were found to be reasonable.

A plot of the activation barrier versus the stretching parameter is shown in Figure 6. The activation barrier always decreases upon going down a group, but the trend across a row is not so clear. Fortunately, the correlation between the activation energy and the stretching parameter is well-pronounced, with the activation barrier generally increasing with the stretching parameter.

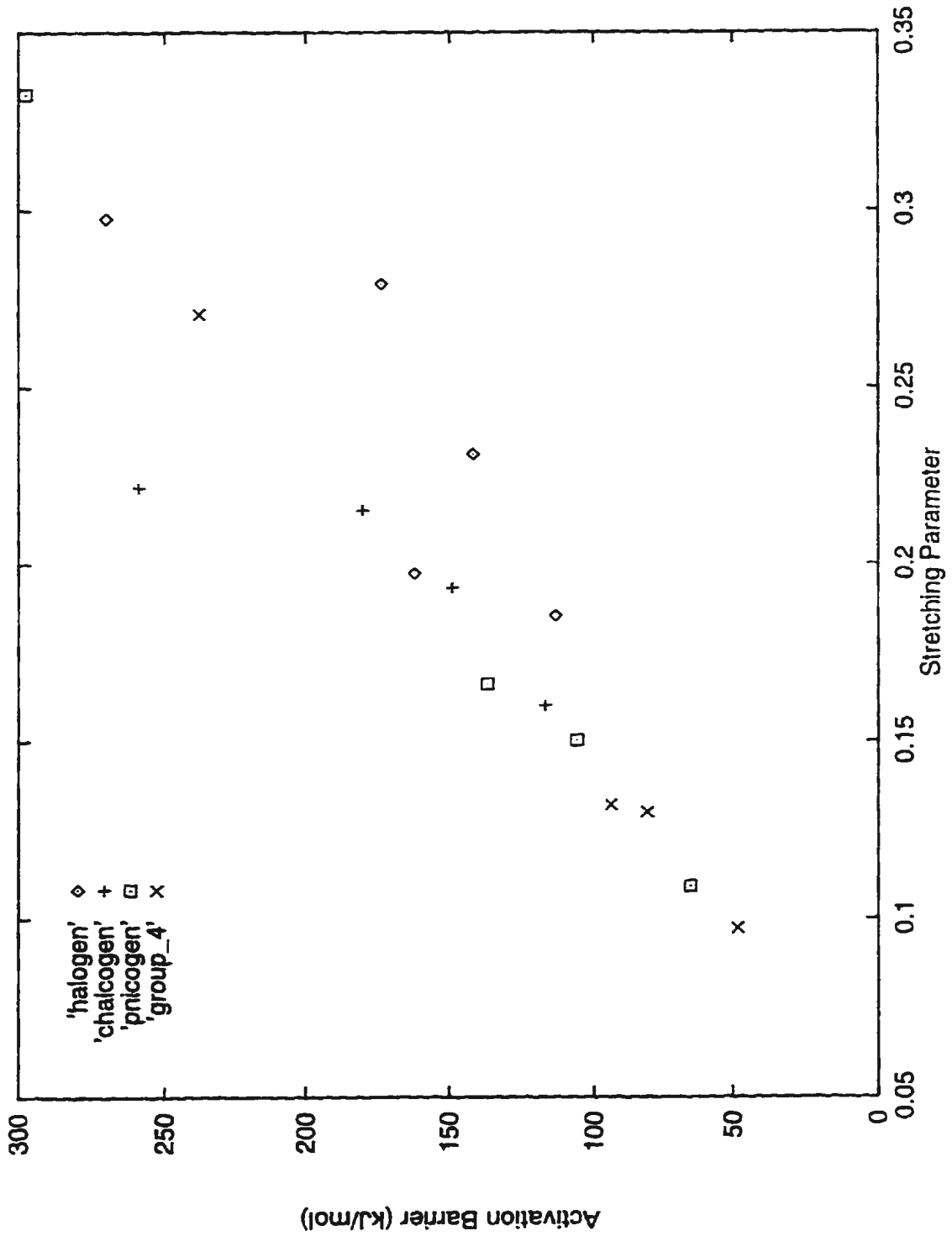
The only recent theoretical studies found were for cyclopentadiene itself. Some barriers found (in kJ/mol) were 162.3 (HF/6-31G*), 119.7 (MP2/6-31G*), 110.2 (MP2/6-31G* + HF/ZPE+therm)[180], 119.7 (MP2/6-31G* + MP2/ZPE+therm), 107.5 (MP2/6-311G**//MP2/6-31G*), 124.3 (QCISD(T)/6-311G**// MP2/6-31G*), and 122.2 (MP4SDTQ/6-311G**//MP2/6-31G*)[181]. Even at these sophisticated levels the barrier still has not converged on a reasonable value, although tunnelling

Table 31: Activation energies of 1,2-group shifts in cyclopentadienes (kJ/mol)

	Basis Set				
System	STO-3G	3-21G	6-31G*//3-21G	6-31G*	ξ
H[a]	243.04	176.27	163.72	162.26	0.1973
F[b]	304.35	255.53	268.94	270.33	0.2976
Cl[c]	213.58	142.55		174.03	0.2792
Br	196.62	140.09		141.89	0.2311
I	163.47	115.60		113.15	0.1854
OH	260.00	223.45	258.53	258.92	0.2217
SH[d]	206.61	154.61		180.56	0.2152
SeH	184.85	148.29	149.67	148.93	0.1933
TeH	154.59	117.45	116.99	116.73	0.1600
NH_2[e]	360.62	277.34	295.80	297.44	0.3326
PH_2[f]	210.57	132.97		136.80	0.1661
AsH_2	179.64	109.57	105.40	105.81	0.1502
SbH_2	128.81	69.64	64.82	64.82	0.1091
CH_3[g]	311.52	221.11	238.55	237.84	0.2706
SiH_3[h]	162.61	94.91		93.77	0.1320
GeH_3	149.53	84.38		86.85	0.1353
SnH_3		50.59	48.26	48.22	0.0973

[a] 6-31G**//3-21G = 157.73, 6-31G** = 156.31
[b] 6-31G**//3-21G = 268.68
[c] 3-21G(*) = 165.76, 6-31G*//3-21G(*) = 173.57, 6-31G**//3-21G = 167.86
[d] 3-21G(*) = 175.44, 6-31G*//3-21G(*) = 181.15
[e] 6-31G**//3-21G = 294.99
[f] 3-21G(*) = 133.27, 6-31G*//3-21G(*) = 136.88
[g] 6-31G**//3-21G = 237.23
[h] 3-21G(*) = 92.66, 6-31G*//3-21G(*) = 93.28, 6-31G**//3-21G = 94.07

Figure 6: Activation barrier vs. stretching parameter (HF/6-31G*)



effects can account for part of the discrepancy. Diffuse functions would likely help describe the transition state better, which would tend to lower the barrier.

From the experimental data presented in Table 25, it seems that methyl substitution has only a small effect on the activation barrier of the migrating group IV atoms. We can thus compare our calculated values with the methyl substituted experimental values. For this series the HF/6-31G* overestimates the enthalpy of activation by 20 - 50 kJ/mol, which is expected, given the importance of including the correlation energy in calculating the energy of transition states.

3.3 The Hydration of Metal Ions

The hydration of metal ions is of vital importance in understanding their solution behavior and of complex formation in general, since in dilute solution complexation necessarily results in the displacement of water from the inner coordination sphere. Raman spectroscopy has proven to be a powerful tool in the determination of the predominant species involved[182] and can be used to unravel the thermodynamics of these complex systems.[3] In addition, ab initio calculations have helped to understand the thermodynamics and assign spectra, both of which are essential to this work.

[3]I would like to thank Dr. Wolfram Rudolph for introducing me to this topic and for the fruitful collaboration that has ensued.

3.3.1 Lithium (I)

Lithium has been found in coordination numbers ranging from two to eight, but if any water is present, this range drops from four to six, and if only water is present as a ligand, then the coordination in the solid state is four[183]. Very recently it has been established by a concentration-dependent Raman spectroscopic study that lithium ion in aqueous solution is surrounded by only 4 waters, and the symmetric stretching mode of the Li-O stretch was found to be 255 cm^{-1}, in excellent agreement with our calculated HF/6-31G* frequency of 248 cm^{-1}[44]. This corrects the previous assignments of the Li-O mode by Moskovits (190 cm^{-1}) and of Nash (440 cm^{-1}) as discussed in Ref. [44], as well as those of Kameda (also 190 cm^{-1})[184] and Rull (382 cm^{-1})[185]. The calculated HF/6-31G* Li-O distance of 1.97 Å also agrees well with Kameda's neutron diffraction result[186] of 1.96 ± 0.02 Å.

Early theoretical work suggested that the hydration number of the lithium cation was either four (CNDO[187], CNDO/2[188], MM[189], MESQUAC[190], MNDO, MP2/6-31+G*//HF/6-31G* + ZPE[191], STO-3G[192] or six (extended huckel[193], HF/3-21G(d)[194]). It is only recently that the tetracoordinated structure has been shown to be favored in solution by purely computational means ((HF, MP2(FC))/(cc-pVDZ, aug-cc-pVDZ)[195], (HF,MP2)/6-31+G*[196]). Our later work[45] explored the role that the second hydration sphere has on the frequencies of the inner-sphere complex, and is the first full ab initio optimization of a second hydration sphere. The second sphere raised the frequency of the Li-O vibration by 18 cm^{-1} at both

the 6-31G* (266 cm^{-1}) and 6-31+G* (257 cm^{-1}) levels, the latter being in excellent agreement with experiment. Finally, the calculations of tetracoordinated $LiX(H_2O)_3$, X=F,Cl of Woon and Dunning[197] should be mentioned since these species exist in concentrated lithium halide solutions[44]. Unfortunately, no frequencies were reported for $LiX(H_2O)_3$, and so a direct comparison is not available.

Since most of the important work that we carried out is already presented in Ref. [45], readers are referred to this paper for such trends as Li-O bond length, symmetric stretch frequency and interaction energy. It is shown that the BSSE-corrected binding energies of the species studied, namely $Li(H_2O)_n^+, n = 1, \ldots, 6, 8, 12$, are independent of the CHA molecular partitioning. The trends observed are rationalized using both steric and electrostatic arguments. In addition, the second hydration sphere significantly modifies the predicted Raman spectrum. The focus hereon will be on aspects not covered by the published work.

We first begin our discussion with the directly bound penta-aquolithium species. At the STO-3G level, a C_2 structure seems to be favored which resembles a trigonal bipyramid. A frequency calculation confirms that the structure is a minimum at this level. The 'equatorial' Li-O bonds are about 0.05 Å longer than those of the tetraaquo species, whereas the 'axial' Li-O bonds are 0.18 Å longer. The 3-21G level gives a minimum C_2 structure which resembles a square pyramid. The Li-O bond lying on the C_2 axis is 0.05 Å shorter than that of the tetraaquo species, but the 4 Li-O distances to the base of the pyramid are between 0.17–0.21 Å longer. This

structure is probably preferred because of the overestimation of the strength of the water-water interaction at this level, as noted previously[195]. The 6-31G* level also gives a directly bound C_2 structure, which resembles the STO-3G structure, but at this level the structure is a transition state, with the imaginary mode corresponding to a libration of the water on the C_2 axis which breaks the symmetry but appears to form a hydrogen-bond with one of the 'axial' waters. This transition state serves as a model for the associative water exchange reaction. The Li-O distances were 0.03–0.07 Å longer for the equatorial ligands and 0.20 Å longer for the axial ligands. The hexaaquolithium species (T_h) possessed 9 and 3 imaginary frequencies at the 3-21G and 6-31G* levels, respectively.

3.3.2 Beryllium (II)

Beryllium is thought to have a hydration number of four, which is supported by most computations[190, 198], although the hexacoordinated structure has also been considered[187, 194]. Ab initio calculations at the HF/6-31G* and MP2/6-31G* level suggest that both the tetra- and the hexa-aquo species are minimal structures, but not the penta-aquo lithium structure, and that the [4+2] structure was the most stable at the HF/6-31G*, MP2/6-31G*, and MPn/6-31G*//HF/6-31G* (n=2,3,4SDQ) levels[198].

A recent extension of these results to include up to eight water molecules at MP2/6-311++G**//HF/6-31G* was recently published[199]. Three structures for

the $Be(H_2O)_8^{2+}$ [4+4] structure were proposed. Firstly, a D_{2d} structure contained four equivalent waters which formed two hydrogen-bonds each to the inner sphere. The other two C_2 structures gave two of the waters forming two hydrogen-bonds to the inner sphere, as in the first structure, and the other two waters forming just one hydrogen-bond ('dangling' waters, in the parlance of Bock and coworkers[199]). These latter structures are more stable, in spite of having fewer hydrogen-bonds.

Our calculations on the [4+0] and [4+4] structure are in agreement with those of Bock, giving, in [4+0], a Be-O distance of 1.6548 Å and an O-H distance of 0.9596 Å. Complexation with water via our S_4 structure shortens the Be-O distance to 1.6484 Å and lengthens the O-H distance to 0.9639 Å. A very major change occurs in the O-Be-O angles, with a significant flattening of the inner-sphere tetrahedron; for example, about the S_4 axis, the angle increases from 108.2° to 119.8°. This may suggest why the dangling structures are preferred, since the geometrical requirements for forming a second hydrogen-bond to the dangling water may impose deformational strain in the inner sphere.

If we compare our [4+4] structure with that of lithium, several differences emerge. Although the outer-sphere waters act as a hydrogen-bond acceptor in both cases, in lithium a weak hydrogen-bond is formed in which the second-sphere water acts as a hydrogen-bond donor, giving 8 strong and 4 weak hydrogen-bonds for a total of 12. No such structure exists for beryllium, presumably because the inner sphere waters, being next to a higher charge, cannot act as hydrogen-bond acceptors. In addition, the

hydrogens of the inner sphere, being more polarized, would act as stronger hydrogen-bond donors. Thus the cost of destroying the 4 weak H-bonds are compensated by the strengthening of the remaining H-bonds.

We must also be careful about stating with certainty what the structure of the second solvation sphere is in solution. In order to describe properly the structure of the n-th solvation sphere, the H-bonds to both the (n-1)-th and (n+1)-th solvation spheres must be taken into account. This was noted for lithium and beryllium, in which inclusion of a second solvation sphere increased the symmetry from S_4 to D_{2d}. However, for lithium, the second sphere waters lie in the σ_d planes, whereas for beryllium, the local C_2 axis of water coincides with the C_2 axes of the complex. This becomes even more critical when the energy separating several possible structures is small. The existence of a long-lasting second solvation sphere for lithium is doubtful, given the small difference in enthalpy between a second-sphere water bonded to tetraaquolithium and that of bulk water. The use of large water clusters to model bulk water must certainly have its limitations, since in the current capabilities of ab initio theory, only a small percentage of the waters in the cluster are sufficiently 'solvated' unless the cluster is large. However, it is certainly the only 'practical' approach, other than adopting either a standard 'total' energy of $H_2O_{(aq)}$ for a given basis set, or accepting an experimental measure of the solvation energy per water.

Finally we must compare the known experimental frequencies of aqueous beryllium (II) with our calculated frequencies. The experimental frequency of the Be-O

symmetric stretch is given as 535 cm^{-1}, in excellent agreement with the 3-21G frequency of 524 cm^{-1}, but in poorer agreement with the 6-31G* frequency of 477 cm^{-1}. The presence of the second sphere waters already corrects the 6-31G* frequency to 508 cm^{-1}, and thus is shown to have a dramatic effect on this mode, as previously noted for the case of lithium[45]. An experimental 'F_2' mode (in D_{2d}, $B_2 + E$) at 335 cm^{-1} is well predicted at HF/6-31G* by deformational modes at 331 cm^{-1} (B_2) and 345 cm^{-1} (E). Another experimental 'F_2' mode at 880 cm^{-1} corresponds to the unscaled HF/6-31G* E-mode at 1072 cm^{-1} (scale factor = 0.82).

3.3.3 Magnesium (II)

The thermodynamics and structure of aquated magnesium (II) has been studied at the same level as the beryllium counterpart[200, 199], and the [6+0] structure was found to be most stable, although the energy of the [5+1] and [4+2] structures were only 17 and 36 kJ/mol higher than the [6+0] structure, and thus entropy may contribute significantly to the corresponding ΔG of water rearrangement from [6+0] to [(6-k)+k].

The Mg-O bond lengthens upon progressing along the basis set series STO-3G (1.9301 Å), 3-21G (2.0429 Å), and 6-31G* (2.1059 Å). The latter value agrees well with the MP2/6-31G* value of 2.097 Å[200]. The experimental Mg-O symmetric stretch frequency of 359 cm^{-1} is in reasonable agreement with the 3-21G frequency of 375 cm^{-1} but in worse agreement with the 6-31G* frequency of 327 cm^{-1}. Again,

this is a consequence of the lack of hydrogen-bonding to the second solvation sphere. To determine whether the second solvation sphere affected M-O modes in general for an octahedral complex, as they certainly do for a tetrahedral complex as formed with Li^+ and Be^{2+}, we optimized a [6+12] structure for $Mg(H_2O)_{18}^{2+}$, retaining the full T_h point group. There are two possible structures for the second sphere which retain T_h symmetry, those in which the second sphere waters lie within a σ-plane of the molecule, or those in which the σ-plane bisects the second sphere waters. The 3-21G level favored the former, contrary to our expectations, but neither proved to be a minimum on the PES. We took both structures and performed a gradient optimization at the 6-31G* level. The Mg-O bond shortens significantly to 2.0822 Å (2nd-sphere HOH in plane) or 2.0829 Å (2nd-sphere HOH out-of-plane).

Because of the enormous size of this system, sufficient disk space was not initially available for analytic frequency evaluation at HF/6-31G*, so a finite difference scheme was used instead on the HOH in-plane structure. The Mg-O symmetric stretching mode increases in frequency to 398 cm^{-1}, a change of over 20 % . However, this mode is coupled to a second sphere H-O-H rocking motion, and thus the change probably is an overestimation. As well, 15 imaginary frequencies of magnitude less than 100 cm^{-1} clouded the characterization of this species as being a minimum on the PES and may have been a result of the finite differencing, but we must be alert to the possibility that the T_h structure simply is not a minimum.

When computational resources became available, we repeated the calculation using analytic frequency calculations. The in-plane structure possessed 15 imaginary frequencies and a symmetric stretching mode of 398 cm^{-1}, reproducing the finite-difference values. The out-of-plane structure possessed 6 imaginary frequencies and a symmetric stretching mode of 383 cm^{-1}, and was 0.27 kJ/mol lower in energy. The symmetric stretching mode does not couple with the wagging water libration. Since these are the only two possible T_h structures, and both possess imaginary frequencies. the true minimum must belong to some subgroup of T_h.

3.3.4 Other Metal Ions

Other hexacoordinated metal ions that we investigated are aluminum (III), gallium (III), indium (III), scandium (III), zinc (II), cadmium (II) and iron (II). Some of these have been investigated theoretically[194, 190, 187, 201, 202, 203, 204, 205], especially zinc (II)[206, 207]. For zinc, there is little energy difference between the [6+0], [5+1] and [4+2] structures at the MP2/6-31G*//HF/6-31G* level (for zinc a (53321/531*/41) basis set was employed)[206].

Our results are shown in Table 32. The geometries and frequencies given above for the 3-21G and 6-31G* bracket the earlier calculations of Sandstrom and coworkers[204]. A significant theoretical underestimation of the symmetric stretch frequency was noticed and correctly attributed to the second solvation sphere. A second solvation

Table 32: Metal-oxygen distances and symmetric stretch frequencies

Metal	M-O Distance (Å)				M-O Symmetric Stretch (cm^{-1})		
	STO-3G	3-21G	6-31G*	Expt.	3-21G	6-31G*	Expt.[182]
Al^{3+}	1.8298	1.9004	1.9347		520	462	526
Sc^{3+}	2.1042	2.1262	2.1817	2.18	447	397	440
Fe^{2+}	2.0493						370
Zn^{2+}		2.0505	2.1357		413	337	390
Ga^{3+}	1.9046	1.9504	2.0053		531	466	521
Cd^{2+}	2.1582	2.2914	2.3532	2.31	345	298	355
In^{3+}	2.0789	2.1394	2.2076		481	411	485

sphere was represented by point charges but only applied to the study of the Jahn-Teller effect[202].

3.4 Interaction Energies of Bimolecular Complexes

The ab initio calculation of intermolecular interaction energies is of general chemical and biological importance. It is well known, however, that interaction energies using small or medium-sized basis sets are plagued by basis set superposition error (BSSE), which results from the lack of balance between the description of the monomers and the complex. BSSE can be corrected by several methods, the most well-known being the Boys-Bernardi counterpoise approach. We have previously investigated the usefulness of an alternative scheme, Mayer's CHA formalism, in correcting for this artifact[43], and thus the conclusions of this paper will be summarized in Section 3.4.2. The systems that we studied were FH-FH, FH-OH_2, FH-NH_3, FH-CO, FH-NCH, FH-NN, H_2O-H_2O, and H_2O-CO_2.

3.4.1 Geometries

No geometries of bimolecular complexes were given in our paper[43], and so we present here the H-F and H···Donor bond lengths. In Tables 33, 34, and 35, the complexes are ordered in terms of increasing interaction energy. With the exception of the STO-3G (all) and N-21G (NN and CO) basis sets, the F-H bond length increases upon complexation, with the change paralleling the increasing interaction energy. This shows that, in general, a polarised split valence basis set is required to accurately predict geometric trends. The unpolarized split-valence basis sets overestimate the increase compared with our best results. Deviations from a linear hydrogen-bond can occur for hydrogen fluoride complexes with water and itself by as much as 25° (HF...HF 6-31G*), but the Hartree-Fock limit is suggested to be about 0° (water) and 10° (hydrogen fluoride). The monomers assume relative orientations that are indicative of classical lone-pair attractions. The HF dimer may also be influenced by the classical dipole-dipole interactions which may force nonlinearity of the hydrogen bonds. On the other hand, in water dimer, the only reasonable arrangement of water leads to a perpendicular dipole arrangement. The N-H bonds and H-N-H angles become slightly larger upon complexation with HF, for sufficiently large (N-31G) basis sets. The donating O-H bond length and O-H bond lengths of the hydrogen acceptor become larger and the geminal O-H bond becomes shorter upon complexation with HF, for sufficiently large (N-31G*) basis sets.

Unlike the overall trend for F-H bond length, we can only say that, in general,

Table 33: FH bond lengths as a function of basis set and complex

		HF ...X, X =						
Basis	Mono.	NN	CO	HF (1)	HF (2)	NCH	H_2O	NH_3
STO-3G	0.9555	0.9540	0.9539	0.9541	0.9529	0.9531	0.9558	0.9563
3-21G	0.9374	0.9364	0.9370	0.9419	0.9401	0.9425	0.9540	0.9644
4-21G	0.9390	0.9384	0.9390	0.9434	0.9415	0.9440	0.9551	0.9655
6-21G	0.9395	0.9390	0.9395	0.9437	0.9418	0.9445	0.9553	0.9656
4-31G	0.9222	0.9224	0.9239	0.9267	0.9244	0.9292	0.9388	0.9495
5-31G	0.9211	0.9216	0.9229	0.9256	0.9233	0.9280	0.9375	0.9482
6-31G	0.9209	0.9213	0.9227	0.9253	0.9230	0.9278	0.9372	0.9478
6-311G	0.9105	0.9117	0.9123	0.9146	0.9123	0.9176	0.9265	0.9363
3-21G*	0.9232	0.9223	0.9236	0.9272	0.9280	0.9271	0.9347	0.9404
4-21G*	0.9250	0.9246	0.9257	0.9290	0.9295	0.9291	0.9365	0.9420
6-21G*	0.9255	0.9252	0.9263	0.9294	0.9300	0.9296	0.9368	0.9423
4-31G*	0.9120	0.9125	0.9140	0.9163	0.9160	0.9178	0.9229	0.9310
5-31G*	0.9112	0.9116	0.9131	0.9153	0.9150	0.9168	0.9218	0.9300
6-31G*	0.9110	0.9114	0.9130	0.9151	0.9148	0.9166	0.9216	0.9299
6-311G*	0.8973	0.8979	0.8989	0.9008	0.8997	0.9023	0.9075	0.9146
3-21G**	0.9090	0.9088	0.9102	0.9138	0.9125	0.9137	0.9208	0.9265
4-21G**	0.9088	0.9092	0.9103	0.9138	0.9123	0.9137	0.9207	0.9259
6-21G**	0.9087	0.9092	0.9102	0.9137	0.9122	0.9136	0.9205	0.9255
4-31G**	0.8997	0.9005	0.9021	0.9040	0.9030	0.9054	0.9099	0.9173
5-31G**	0.9004	0.9011	0.9026	0.9045	0.9035	0.9060	0.9104	0.9180
6-31G**	0.9006	0.9013	0.9029	0.9047	0.9037	0.9062	0.9105	0.9183
6-311G**	0.8960	0.8972	0.8982	0.8999	0.8990	0.9025	0.9066	0.9143

hydrogen-bond distances correspond only roughly to the interaction energy, with shorter hydrogen-bonds suggesting a stronger interaction. If we restrict comparisons to systems having the same hydrogen-bond acceptor atoms, then the trend would be very clear (for example, HF ...NN, HF ...NCH, HF...NH_3). As is clear from Tables 34 and 35, trends here are relatively insensitive to what basis set is used.

Table 34: Hydrogen-bond lengths as a function of basis set and complex

	System						
Basis	HF-NN	HF-CO	HF-HF	$(H_2O)_2$	HF-NCH	HF-H_2O	HF-NH_3
STO-2G	2.2100	2.2262	1.5685	1.7073	2.0389	1.6911	1.8736
STO-3G	2.3233	2.3268	1.6189	1.7505	2.0749	1.6730	1.8137
STO-4G	2.3691	2.3540	1.6246	1.7411	2.0831	1.6544	1.7739
STO-5G	2.3727	2.3415	1.6143	1.7306	2.0727	1.6463	1.7676
STO-6G	2.3655	2.3291	1.6100	1.7290	2.0711	1.6451	1.7684
3-21G	2.0410	2.0902	1.6709	1.8236	1.8672	1.6149	1.7023
3-21G*	2.1230	2.1586	1.6771	1.8936	1.9391	1.6914	1.7795
3-21G**	2.1307	2.1699	1.6947	1.9088	1.9521	1.7063	1.7980
4-21G	2.0753	2.1527	1.6910	1.8443	1.8928	1.6336	1.7200
4-21G*	2.1714	2.2302	1.6923	1.9127	1.9718	1.7085	1.8003
4-21G**	2.1773	2.2443	1.7111	1.9273	1.9855	1.7255	1.8215
6-21G	2.0854	2.1745	1.6967	1.8496	1.9010	1.6394	1.7254
6-21G*	2.1855	2.2552	1.6978	1.9188	1.9815	1.7155	1.8070
6-21G**	2.1928	2.2691	1.7161	1.9326	1.9955	1.7320	1.8283
4-31G	2.1198	2.1366	1.7738	1.8727	1.9027	1.6661	1.7335
4-31G*	2.2697	2.2373	1.8344	2.0136	1.9932	1.7885	1.8216
4-31G**	2.2972	2.2447	1.8441	2.0274	2.0059	1.8047	1.8379
5-31G	2.1457	2.1758	1.7953	1.8845	1.9169	1.6765	1.7413
5-31G*	2.2884	2.2597	1.8515	2.0195	2.0041	1.8005	1.8251
5-31G**	2.3010	2.2610	1.8564	2.0371	2.0085	1.8127	1.8383
6-31G	2.1536	2.1854	1.7990	1.8852	1.9186	1.6792	1.7446
6-31G*	2.2889	2.2606	1.8547	2.0206	2.0048	1.8023	1.8254
6-31G**	2.3015	2.2616	1.8599	2.0366	2.0085	1.8142	1.8383
6-311G	2.1554	2.2753	1.8195	1.8621	1.9181	1.6667	1.7389
6-311G*	2.2836	2.3276	1.8601	1.9714	2.0087	1.7642	1.8253
6-311G**	2.2682	2.3295	1.8959	2.0326	2.0033	1.7950	1.8296
SV	2.1444	2.2441	1.7805	1.8551	1.8632	1.6222	1.6787
SV*	2.2355	2.3071	1.8430	1.9927	1.9662	1.7416	1.7772
SV**	2.2449	2.2837	1.8611	2.0244	1.9747	1.7859	1.8213
DZ	2.0446	2.1324	1.7501	1.8487	1.8370	1.6167	1.6698
DZ*	2.1581	2.2064	1.8068	1.9791	1.9447	1.7294	1.7673
DZ**	2.1551	2.1893	1.8295	2.0116	1.9585	1.7689	1.8121
TZ	2.1320	2.2452	1.8312	1.8772	1.8690	1.6457	1.6822
TZ*	2.2433	2.2929	1.9053	2.0088	1.9738	1.7658	1.7904
TZ**	2.2317	2.2828	1.9224	2.0372	1.9664	1.7940	1.8124
TQZ	2.1285	2.2316	1.8331	1.8779	1.8713	1.6327	1.6662
TQZ*	2.2287	2.2857	1.9063	2.0043	1.9613	1.7570	1.7765
TQZ**	2.2422	2.2793	1.9347	2.0568	1.9635	1.8007	1.8128

Table 35: Hydrogen-bond lengths as a function of basis set and complex

	System				
Basis	HF-NN	HF-HF	$(H_2O)_2$	HF-H_2O	HF-NH_3
QPZ	2.1295	1.8297	1.8761	1.6333	1.6801
QPZd	2.2513	1.9060	2.0006	1.7566	1.7851
QPZdp	2.2429	1.9397	2.0566	1.7991	1.8125
QPZdfp	2.2395	1.9313	2.0520	1.7976	1.8202
QPZ2dp	2.2845	1.9338	2.0877	1.8049	1.8128
QPZ2dfp	2.2436	1.9454	2.0792	1.8008	1.8207
QPZ2df2p	2.2473	1.9426	2.0915	1.8128	1.8204
QPZ2df2pd	2.2464	1.9404	2.0786	1.8100	1.8162

3.4.2 Interaction Energies

We concluded in our paper that the CHA/CE scheme is an excellent method for correcting for BSSE and converges to the Boys-Bernardi result faster than to the uncorrected SCF result. It was also shown that a polarized split-valence basis set is necessary when correcting for BSSE if energies close to the Hartree-Fock limit are needed.

3.5 Benzene Oxide and Oxepin Valence Tautomerism

The valence tautomerism between benzene oxide and oxepin[208] (see Figure 7 and similar systems[209] is of considerable experimental and theoretical[4] interest. Benzo-fused analogs of these are metabolites of polycyclic aromatic hydrocarbons (PAH's)

[4]Parts of this work were presented at the 17th and 19th Annual Atlantic Student Chemistry Conferences, held at Sir Wilfred Grenfell College, Corner Brook, NF 1992 and Mount Saint Vincent University, Halifax, NS May 19-21, 1994, respectively, and discussed in C. C. Pye, *Honours Diss.*, Memorial University of Newfoundland, St. John's, NF Canada.

Figure 7: Benzene oxide–oxepin valence tautomerism

and are thus of interest in mutagenicity and carcinogenicity studies[210]. The proportion of benzene oxide increases upon decreasing the temperature or increasing the polarity of the solvent. Substituting the 2 and 7 positions of oxepin (geminal to oxygen) with methyl groups favors the oxepin form (steric strain) whereas a propylene bridge at these positions favors the oxide form (ring strain). The ΔH^0 was determined from signal area ratios (low T) and from the coalescence position (high T) in proton NMR to be 7.1 ± 1.7 kJ/mol, and the forward and backward activation barriers, determined from an Arrhenius plot with rate constants derived from exchange broadening of the NMR signals, were 38.1 ± 3.3 and 30.1 ± 4.2 kJ/mol from which we estimate a ΔH^0 of 8.0 ± 5.3 kJ/mol. The entropy change was derived both from the ratio of the frequency factors (43.9 ± 34.7 J mol/K) and from the equilibrium constant (46.0 ± 20.9 J mol/K)[208].

Benzene oxide itself was first optimized at the STO-5G level[211] in an examination of the first step of metabolism of benzene, along with the peroxide. The structure was close to that obtained by Kollman[212], who used both STO-3G and the semi-empirical MINDO/3 method to investigate the valence tautomerism. The partial

STO-3G optimization in this work (C-H fixed at 1.09 Å) gave a planar oxepin moiety, similar to the near-planar MINDO/3 structure.[5] The ΔE obtained from this STO-3G calculation (and the HF/4-31G//STO-3G calculation which followed) can therefore be in significant error, as well as the estimates of the relative stability of the substituted cases. Using similar constraints, Cremer and coworkers found both the boat and planar forms of oxepin with the STO-3G basis set as well as unconstrained MNDO structures, and the barriers to inversion were determined to be 6.7 (STO-3G) and 6.3 kJ/mol (6-31G*//STO-3G)[213]. Fully optimized structures of benzene oxide and oxepin were first published by Schulman, who also carried out single point calculations with the STO-3G geometries at the HF/4-31G, MP2-FC/4-31G, HF/6-31G*, and MP2-FC/6-31G* levels[214]. Finally, Bock optimized the structures of benzene oxide and oxepin at the semiempirical AM1 and ab initio 6-31G and 6-31G* levels and carried out single point calculations at the MP2/6-31G* and MP3/6-31G* levels, demonstrating the extreme sensitivity to theoretical level[215]. Most recently, a HF/3-21G structure of benzene oxide was published in conjunction with a photoelectron and electron transmission study[216]. In addition, a suggested mechanism of decomposition of benzene oxide to phenol was published which involved protonation at the oxygen[217].

The sulfur analogues, benzene sulfide and thiepin, remain elusive synthetic targets. The only theoretical studies performed were at the rather crude MNDO[218] and

[5]Our first attempt at the optimization of oxepin at the STO-3G level also gave a planar structure which was later traced to an improper torsional symmetry constraint. The planar structure will be shown to be the transition state for the degenerate interconversion of the boat forms of oxepin.

STO-3G levels, the latter in conjunction with a photoelectron study[219].

3.5.1 Geometries

In most cases, geometries were optimized within the C_s point group. For the inversion barriers a C_{2v} symmetry was obtained, corresponding to a planar structure.

Benzene Oxide–Oxepin

The structures of species involved in the valence tautomerism between benzene oxide and oxepin, and of the inversion of oxepin, at several different theoretical levels, are presented in Tables 36, 37, 38, and 39. The lengths of double bonds are insensitive to basis set at the Hartree-Fock level, but increase by about 0.02-0.03 Å on going from HF/6-31G* to MP2/6-31G*. Single bonds vary considerably (0.07Å) over all theoretical levels, especially the polar C-O bond. The changes are most extreme in the three-membered ring. Angles are not sensitive to theoretical level except for those involving oxygen. The three torsions reported here vary more upon proceeding to the oxepin form, but α (defined in Table 36) is sensitive in all three species.

At the MP2 level, the curvature ($360° - \alpha - \beta$) is practically the same for both benzene oxide (85.13°) and oxepin (85.27°). The value of β of the transition state is closer to that of the oxide, whereas the value of α of the transition state is closer in value to that of oxepin. The result of this angular asynchronicity is that the transition state is more curved (92.87°) than either of the tautomers. The bond lengths involved

Table 36: Selected geometric parameters of benzene oxide

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-H_2[a]	1.0831	1.0721	1.0757	1.0874
C_1-H_1	1.0826	1.0719	1.0756	1.0872
C_5-H_5	1.0892	1.0701	1.0768	1.0897
C_2-C_3	1.4769	1.4606	1.4641	1.4443
C_1-C_2	1.3239	1.3284	1.3292	1.3602
C_1-C_6	1.4973	1.4699	1.4801	1.4598
C_5-C_6	1.5044	1.5022	1.4765	1.5188
C_5-O	1.4404	1.4791	1.4061	1.4423
C_1-C_2-C_3	121.77	121.50	121.53	121.42
C_2-C_1-C_6	120.79	120.88	120.51	120.21
C_1-C_6-C_5	117.15	117.26	117.74	117.40
C_1-C_6-O	117.18	115.85	116.96	116.88
C_5-C_6-O	58.52	59.48	58.33	58.23
C_5-O-C_6	62.96	61.04	63.34	63.54
$\alpha = C_2$-C_1-C_4-C_5	173.85	173.21	174.53	168.74
$\beta = C_1$-C_6-C_5-O	106.73	105.40	106.11	106.13
$\gamma = C_1$-C_6-C_5-H_5	150.49	153.54	152.97	152.76

[a]The carbons of benzene oxide (a 1,3-diene) are numbered cyclically and consistent with IUPAC nomenclature.

in the reaction coordinate (namely, C_2-C_3, C_1-C_2, C_1-C_6 and C_5-C_6) are usually closer to the slightly less stable oxepin form, in accord with the Hammond postulate. (The C_1-C_2 bond is an exception.) The planar form shows more bond alternation and much wider angles, the latter suggesting some ring strain.

Benzene Sulfide–Thiepin

The structural parameters of species involved in the valence tautomerism between benzene sulfide and thiepin, and of the inversion of thiepin, at several different theoretical levels, are presented in Tables 40, 41, 42, and 43.

Table 37: Selected geometric parameters of the benzene oxide–oxepin transition state

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-H_2	1.0833	1.0729	1.0759	1.0880
C_1-H_1	1.0823	1.0732	1.0761	1.0881
C_5-H_5	1.0899	1.0688	1.0749	1.0892
C_2-C_3	1.3768	1.4036	1.3953	1.3986
C_1-C_2	1.4056	1.3699	1.3824	1.4029
C_1-C_6	1.3747	1.3924	1.3873	1.3857
C_5-C_6	1.8910	1.8205	1.8191	1.9010
C_5-O	1.4163	1.4319	1.3701	1.3995
C_1-C_2-C_3	122.61	122.45	122.36	122.51
C_2-C_1-C_6	122.04	122.10	121.89	121.78
C_1-C_6-C_5	111.34	112.22	112.38	111.28
C_1-C_6-O	120.20	118.23	119.46	119.02
C_5-C_6-O	48.11	50.53	48.41	47.22
C_5-O-C_6	83.78	78.94	83.18	85.56
$\alpha = C_2$-C_1-C_4-C_5	157.86	160.04	159.59	156.71
$\beta = C_1$-C_6-C_5-O	112.02	109.00	110.23	110.42
$\gamma = C_1$-C_6-C_5-H_5	147.11	150.22	149.28	148.42

Table 38: Selected geometric parameters of oxepin

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-H_2	1.0839	1.0741	1.0766	1.0884
C_1-H_1	1.0828	1.0732	1.0760	1.0880
C_5-H_5	1.0901	1.0698	1.0743	1.0884
C_2-C_3	1.3237	1.3267	1.3300	1.3637
C_1-C_2	1.4833	1.4674	1.4670	1.4465
C_1-C_6	1.3203	1.3154	1.3201	1.3488
C_5-C_6	2.3153	2.3956	2.3175	2.2477
C_5-O	1.4143	1.4026	1.3696	1.3959
C_1-C_2-C_3	124.66	125.57	124.76	124.07
C_2-C_1-C_6	125.18	125.33	125.04	123.84
C_1-C_6-C_5	105.27	104.04	105.04	105.85
C_1-C_6-O	124.88	124.41	128.28	121.34
C_5-C_6-O	35.06	31.35	32.21	36.38
C_5-O-C_6	109.88	117.30	115.58	107.24
$\alpha = C_2$-C_1-C_4-C_5	156.17	155.76	155.53	152.98
$\beta = C_1$-C_6-C_5-O	130.01	135.15	131.88	121.75
$\gamma = C_1$-C_6-C_5-H_5	151.34	154.36	153.03	149.16

Table 39: Selected geometric parameters of planar oxepin

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-H_2	1.0832	1.0734	1.0757	1.0874
C_1-H_1	1.0815	1.0726	1.0751	1.0868
C_5-H_5	1.0895	1.0691	1.0723	1.0852
C_2-C_3	1.3162	1.3213	1.3220	1.3459
C_1-C_2	1.4894	1.4744	1.4757	1.4674
C_1-C_6	1.3179	1.3153	1.3196	1.3430
C_5-C_6	2.4303	2.4750	2.4154	2.4347
C_5-O	1.4017	1.3868	1.3603	1.3848
C_1-C_2-C_3	126.05	127.07	126.09	125.94
C_2-C_1-C_6	129.93	129.21	129.76	130.42
C_1-C_6-C_5	104.02	103.72	104.15	103.64
C_1-C_6-O	133.92	130.55	131.54	132.11
C_5-C_6-O	29.90	26.83	27.39	28.47
C_5-O-C_6	120.20	126.34	125.22	123.06
$\alpha = C_2$-C_1-C_4-C_5	180.00	180.00	180.00	180.00
$\beta = C_1$-C_6-C_5-O	180.00	180.00	180.00	180.00
$\gamma = C_1$-C_6-C_5-H_5	180.00	180.00	180.00	180.00

Similar trends with respect to theoretical level persist for the C-C and C=C bonds as for the oxepin system. The C-S bond is rather insensitive to correlation but very sensitive to basis set, especially in the three membered ring. The 3-21G basis set gives a ridiculously long C-S bond length, especially in benzene sulfide (2.0151 Å) compared with the MP2/6-31G* value of 1.8524 Å. The angles are relatively constant except those involving the heteroatom, and these are relatively constant once at or beyond the HF/3-21G(*), except for the decrease in the C-S-C angle in thiepin upon proceeding to the correlated level. The torsions α of benzene sulfide, and β of thiepin change significantly upon proceeding to the correlated level.

The curvature of benzene sulfide (72.68°) and thiepin (80.56°) are unequal, unlike those of the benzene oxide system, and the more stable tautomer is also the less curved. Like the oxygen analogue, the transition state is more curved than either (86.90°) for similar reasons. The bond length trends in the transition state are less clear. The planar form of thiepin also shows increased bond alternation, and the ring angles are larger than both the boat form and also those of planar oxepin. This suggests that the inversion barrier would be greater for thiepin than oxepin.

Xylene Oxide–Dimethyloxepin

The structural parameters of species involved in the valence tautomerism between *o*-xylene oxide and 2,7-dimethyloxepin, and of the inversion of 2,7-dimethyloxepin, at several different theoretical levels, are presented in Tables 44, 45, 46, and 47.

Table 40: Selected geometric parameters of benzene sulfide

	Basis Set					
Parameter	STO-3G	STO-3G*	3-21G	3-21G(*)	6-31G*	MP2/6-31G*
C_2-C_3	1.4776	1.4788	1.4498	1.4601	1.4639	1.4472
C_1-C_2	1.3227	1.3215	1.3335	1.3275	1.3283	1.3565
C_1-C_6	1.4976	1.5033	1.4556	1.4733	1.4790	1.4667
C_5-C_6	1.5174	1.5297	1.4554	1.4938	1.4770	1.4978
C_5-S	1.7928	1.7573	2.0151	1.8612	1.8449	1.8524
C_1-C_2-C_3	121.52	121.60	121.00	121.31	121.32	121.16
C_2-C_1-C_6	121.70	121.90	120.97	121.47	121.13	121.23
C_1-C_6-C_5	116.64	116.34	118.03	117.18	117.54	117.47
C_1-C_6-S	119.43	119.78	114.76	117.45	118.82	118.97
C_5-C_6-S	64.96	64.20	68.83	66.34	66.40	66.15
C_5-S-C_6	50.08	51.60	42.34	47.32	47.20	47.70
$\alpha = C_2$-C_1-C_4-C_5	175.92	175.52	178.93	177.97	179.15	175.78
$\beta = C_1$-C_6-C_5-S	111.86	112.10	107.62	109.93	111.44	111.54
$\gamma = C_1$-C_6-C_5-H_5	141.32	139.88	154.13	146.54	145.80	145.75

Table 41: Selected geometric parameters of the benzene sulfide–thiepin transition state

	Basis Set					
Parameter	STO-3G	STO-3G*	3-21G	3-21G(*)	6-31G*	MP2/6-31G*
C_2-C_3	1.3769	1.3828	1.3869	1.3972	1.3970	1.4047
C_1-C_2	1.4069	1.4006	1.3858	1.3764	1.3812	1.3966
C_1-C_6	1.3742	1.3823	1.3727	1.3882	1.3906	1.3961
C_5-C_6	2.0330	2.0553	2.0587	2.0301	2.0378	2.0576
C_5-S	1.7640	1.7317	1.8368	1.7614	1.7556	1.7647
C_1-C_2-C_3	122.71	123.12	123.07	123.25	123.25	123.01
C_2-C_1-C_6	123.85	123.83	124.68	124.40	124.59	124.15
C_1-C_6-C_5	108.33	108.08	107.83	108.41	108.31	108.13
C_1-C_6-S	122.71	122.53	121.14	121.44	122.57	121.80
C_5-C_6-S	54.81	53.60	55.92	54.81	54.52	54.34
C_5-S-C_6	70.38	72.80	68.16	70.38	70.96	71.32
$\alpha = C_2$-C_1-C_4-C_5	155.53	155.81	157.32	158.45	158.74	156.52
$\beta = C_1$-C_6-C_5-S	117.58	117.52	115.99	115.98	117.42	116.58
$\gamma = C_1$-C_6-C_5-H_5	137.90	136.78	140.39	138.69	138.59	138.74

Table 42: Selected geometric parameters of thiepin

	Basis Set					
Parameter	STO-3G	STO-3G*	3-21G	3-21G(*)	6-31G*	MP2/6-31G*
C_2-C_3	1.3212	1.3233	1.3293	1.3312	1.3322	1.3648
C_1-C_2	1.4892	1.4863	1.4642	1.4625	1.4662	1.4476
C_1-C_6	1.3169	1.3218	1.3169	1.3220	1.3240	1.3528
C_5-C_6	2.7096	2.6731	2.7764	2.6999	2.7301	2.6458
C_5-S	1.7749	1.7544	1.8342	1.7748	1.7762	1.7703
C_1-C_2-C_3	126.86	126.77	127.17	126.81	127.14	126.30
C_2-C_1-C_6	127.26	126.62	127.23	126.57	127.16	126.32
C_1-C_6-C_5	98.70	99.36	97.03	98.35	98.09	99.21
C_1-C_6-S	126.78	126.25	122.85	123.21	124.44	122.60
C_5-C_6-S	40.24	40.38	40.81	40.48	39.78	41.64
C_5-S-C_6	99.52	99.24	98.38	99.04	100.44	96.72
$\alpha = C_2$-C_1-C_4-C_5	151.67	151.38	149.14	149.54	150.81	149.72
$\beta = C_1$-C_6-C_5-S	139.18	136.99	133.91	132.90	136.23	129.72
$\gamma = C_1$-C_6-C_5-H_5	144.21	143.10	143.04	142.22	143.46	140.95

Table 43: Selected geometric parameters of planar thiepin

	Basis Set					
Parameter	STO-3G	STO-3G*	3-21G	3-21G(*)	6-31G*	MP2/6-31G*
C_2-C_3	1.3191	1.3188	1.3242	1.3231	1.3258	1.3511
C_1-C_2	1.4907	1.4914	1.4711	1.4731	1.4757	1.4655
C_1-C_6	1.3157	1.3191	1.3135	1.3177	1.3207	1.3458
C_5-C_6	2.8461	2.8066	2.9584	2.8648	2.8713	2.8595
C_5-S	1.7667	1.7449	1.8240	1.7631	1.7684	1.7671
C_1-C_2-C_3	130.09	129.90	131.41	130.57	130.49	130.25
C_2-C_1-C_6	131.32	130.81	131.77	131.26	131.44	131.51
C_1-C_6-C_5	98.59	99.29	96.82	98.17	98.07	98.23
C_1-C_6-S	134.93	135.75	132.63	133.84	133.80	134.23
C_5-C_6-S	36.34	36.46	35.81	35.66	35.72	35.99
C_5-S-C_6	107.32	107.08	108.38	108.68	108.56	108.02
$\alpha = C_2$-C_1-C_4-C_5	180.00	180.00	180.00	180.00	180.00	180.00
$\beta = C_1$-C_6-C_5-S	180.00	180.00	180.00	180.00	180.00	180.00
$\gamma = C_1$-C_6-C_5-H_5	180.00	180.00	180.00	180.00	180.00	180.00

Table 44: Selected geometric parameters of *o*-xylene oxide

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_5-CH_3	1.5360	1.5109	1.5117	1.5077
C_2-C_3	1.4745	1.4592	1.4617	1.4436
C_1-C_2	1.3223	1.3262	1.3270	1.3570
C_1-C_6	1.5045	1.4784	1.4893	1.4699
C_5-C_6	1.5159	1.5060	1.4846	1.5206
C_5-O	1.4419	1.4830	1.4119	1.4525
C_1-C_2-C_3	121.47	121.18	121.10	121.02
C_2-C_1-C_6	121.87	121.94	121.86	121.67
C_1-C_6-C_5	116.43	116.65	116.91	116.72
C_1-C_6-O	115.74	114.71	115.30	115.25
C_5-C_6-O	58.29	59.49	58.28	58.44
C_5-O-C_6	63.42	61.02	63.44	63.12
$\alpha = C_2$-C_1-C_4-C_5	174.52	174.63	175.89	171.33
$\beta = C_1$-C_6-C_5-O	105.25	104.31	104.46	104.55
$\gamma = C_1$-C_6-C_5-H_5	151.41	154.18	154.35	154.28

The geometric trends with theoretical level are more or less those indicated for the unsubstituted case.

The curvature of *o*-xylene oxide (84.12°) is less than that of 2,7-dimethyloxepin (87.78°) as was the case for the analogous sulfur system, whereas the transition state for their interconversion is more curved (92.68°) for the same reasons. No clear trend emerges in the bonds comprising the reaction coordinate. The planar form shows more bond alternation and wider angles than the boat form.

Table 45: Selected geometric parameters of the *o*-xylene oxide–dimethyloxepin transition state

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_5-CH_3	1.5312	1.5061	1.5053	1.5004
C_2-C_3	1.3800	1.4041	1.3981	1.4045
C_1-C_2	1.3973	1.3651	1.3751	1.3917
C_1-C_6	1.3852	1.4003	1.3990	1.4039
C_5-C_6	1.8960	1.8303	1.8250	1.8473
C_5-O	1.4201	1.4367	1.3767	1.4109
C_1-C_2-C_3	122.59	122.39	122.33	122.33
C_2-C_1-C_6	123.03	123.08	122.91	122.44
C_1-C_6-C_5	110.92	111.71	111.91	111.87
C_1-C_6-O	118.72	116.97	117.62	117.13
C_5-C_6-O	48.12	50.43	48.48	49.11
C_5-O-C_6	83.76	79.14	83.04	81.78
$\alpha = C_2$-C_1-C_4-C_5	159.54	161.49	161.39	159.72
$\beta = C_1$-C_6-C_5-O	110.38	107.71	108.14	107.60
$\gamma = C_1$-C_6-C_5-CH_3	150.42	153.85	154.24	154.02

Table 46: Selected geometric parameters of dimethyloxepin

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_5-CH_3	1.5267	1.4996	1.4974	1.4930
C_2-C_3	1.3243	1.3267	1.3299	1.3643
C_1-C_2	1.4809	1.4654	1.4653	1.4441
C_1-C_6	1.3238	1.3179	1.3235	1.3531
C_5-C_6	2.3260	2.4221	2.3446	2.2703
C_5-O	1.4189	1.4071	1.3768	1.4036
C_1-C_2-C_3	124.80	125.83	124.99	124.39
C_2-C_1-C_6	124.89	125.41	125.10	123.99
C_1-C_6-C_5	105.07	103.60	104.57	105.54
C_1-C_6-O	122.33	122.58	121.94	119.38
C_5-C_6-O	34.95	30.61	31.63	36.03
C_5-O-C_6	110.10	118.78	116.74	107.94
$\alpha = C_2$-C_1-C_4-C_5	155.31	155.51	155.13	153.32
$\beta = C_1$-C_6-C_5-O	125.55	132.77	128.33	118.90
$\gamma = C_1$-C_6-C_5-CH_3	152.47	156.32	155.25	152.01

Table 47: Selected geometric parameters of planar dimethyloxepin

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_5-CH_3	1.5305	1.5022	1.4995	1.4954
C_2-C_3	1.3160	1.3218	1.3223	1.3471
C_1-C_2	1.4863	1.4702	1.4721	1.4619
C_1-C_6	1.3214	1.3179	1.3232	1.3469
C_5-C_6	2.4675	2.5169	2.4651	2.4871
C_5-O	1.4061	1.3919	1.3685	1.3949
C_1-C_2-C_3	126.38	127.62	126.64	126.58
C_2-C_1-C_6	130.24	129.23	129.94	130.49
C_1-C_6-C_5	103.38	103.15	103.42	102.93
C_1-C_6-O	132.05	128.45	129.18	129.86
C_5-C_6-O	28.67	25.29	25.76	26.94
C_5-O-C_6	122.66	129.42	128.48	126.12
$\alpha = C_2$-C_1-C_4-C_5	180.00	180.00	180.00	180.00
$\beta = C_1$-C_6-C_5-O	180.00	180.00	180.00	180.00
$\gamma = C_1$-C_6-C_5-H_5	180.00	180.00	180.00	180.00

Protonated Benzene Oxide

The structural parameters of species involved in the valence tautomerism between protonated benzene oxide and oxepin, at several different theoretical levels, are presented in Tables 48, 49, and 50. Again, the same trends are noticed with basis set with regards to the single and double bonds. The O-H bond shortens upon improving the basis set. The C-O-H angle is overestimated at the 3-21G level relative to HF/6-31G*. The 3-21G basis set is well known to underestimate the pyramidality of sp^3 nitrogen in the isoelectronic RNH_2.

Upon protonation, the C-O bond lengthens considerably at all levels and the C-O-C angle of the oxide form becomes smaller, whereas it increases in the oxepin form. The C-C bonds geminal to the C-O bond shorten slightly. In benzene oxide,

Table 48: Selected geometric parameters of protonated benzene oxide

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4850	1.4695	1.4710	1.4535
C_1-C_2	1.3282	1.3309	1.3313	1.3616
C_1-C_6	1.4917	1.4651	1.4716	1.4544
C_5-C_6	1.5016	1.4791	1.4583	1.4861
C_5-O	1.5163	1.5943	1.5353	1.5766
O-H	0.9986	0.9781	0.9607	0.9937
C_1-C_2-C_3	122.26	121.76	121.85	121.77
C_2-C_1-C_6	119.63	119.84	119.34	119.23
C_1-C_6-C_5	118.02	118.36	118.80	118.79
C_1-C_6-O	115.12	114.65	115.59	115.53
C_5-C_6-O	60.32	62.36	61.65	61.88
C_5-O-C_6	59.36	55.28	56.70	56.24
C_6-O-H	111.45	115.55	113.15	107.78
$\alpha = C_2$-C_1-C_4-C_5	176.47	177.79	178.83	174.68
$\beta = C_1$-C_6-C_5-O	104.49	104.62	105.28	105.30
$\gamma = C_1$-C_6-C_5-H_5	160.22	162.54	161.78	162.24

protonation increases α and γ by about 4° and 9° respectively, whereas in oxepin, α and β are *decreased* by 2° and 8°, respectively.

3.5.2 Energies

Benzene Oxide–Oxepin

The reaction energies and activation barriers pertaining to the valence tautomerism between benzene oxide and oxepin, and the inversion barrier of oxepin, are presented in Table 51. The scatter in the prediction of ΔE at the various levels of theory is quite pronounced. One has to proceed to the correlated levels in order to establish even the correct *sign* of the enthalpy. The inadequacy of the geometry as predicted

Table 49: Selected geometric parameters of the protonated benzene oxide–oxepin transition state

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.3840	1.4008	1.3943	1.4034
C_1-C_2	1.4078	1.3782	1.3891	1.4063
C_1-C_6	1.3738	1.3800	1.3747	1.3766
C_5-C_6	1.9268	1.9032	1.8984	1.9770
C_5-O	1.4565	1.4879	1.4348	1.4628
O-H	0.9993	0.9820	0.9682	1.0033
C_1-C_2-C_3	122.69	122.78	122.65	122.76
C_2-C_1-C_6	122.10	122.69	122.36	121.98
C_1-C_6-C_5	108.50	111.02	111.21	110.15
C_1-C_6-O	117.25	116.32	117.32	116.29
C_5-C_6-O	48.59	50.24	48.58	47.49
C_5-O-C_6	82.82	79.52	82.84	85.02
C_6-O-H	113.46	121.51	115.22	110.01
$\alpha = C_2$-C_1-C_4-C_5	156.95	159.34	158.53	155.09
$\beta = C_1$-C_6-C_5-O	108.50	107.34	108.31	107.68
$\gamma = C_1$-C_6-C_5-H_5	157.34	159.75	158.67	158.65

Table 50: Selected geometric parameters of protonated oxepin

	Basis Set			
Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.3274	1.3314	1.3348	1.3719
C_1-C_2	1.4857	1.4633	1.4645	1.4455
C_1-C_6	1.3180	1.3107	1.3141	1.3420
C_5-C_6	2.4875	2.5312	2.4639	2.3586
C_5-O	1.4565	1.4766	1.4493	1.4655
O-H	0.9909	0.9761	0.9622	0.9988
C_1-C_2-C_3	125.56	126.13	125.45	124.56
C_2-C_1-C_6	126.50	126.05	125.94	124.13
C_1-C_6-C_5	102.44	101.57	102.52	104.09
C_1-C_6-O	119.45	116.95	118.07	116.44
C_5-C_6-O	31.36	31.01	31.78	36.41
C_5-O-C_6	117.28	117.98	116.43	107.16
C_6-O-H	114.17	118.95	113.89	108.48
$\alpha = C_2$-C_1-C_4-C_5	115.04	153.28	153.67	151.09
$\beta = C_1$-C_6-C_5-O	127.26	123.88	123.83	115.66
$\gamma = C_1$-C_6-C_5-H_5	154.73	154.62	154.87	154.87

by the 3-21G basis set is manifested also in the difference in the single point energies. For the enthalpies, the HF/6-31G* geometries are fine, but the activation energies are more sensitive to the use of a correlated geometry. The effect of the various correlated approximations is also seen. For enthalpies, the MP3/6-31G*//MP2/6-31G* seems to be sufficient, but, for valence tautomerization activation barriers, an MP4SDTQ calculation seems to be needed, as the MP3 and MP4 without triple excitations overestimate the barrier relative to the full MP4. The MP3 level seems to be sufficient for inversion barriers. The ZPE/ thermal corrections to 298 K are of opposite sign at the HF/3-21G and HF/6-31G* levels, and given the deficiences of the 3-21G basis set in this system we place more confidence in the HF/6-31G* correction. Similarly, we would trust the entropy term calculated from HF/6-31G* over HF/3-21G.

Our best prediction for the gas-phase ΔH of valence tautomerization (determined by adding the QCISD(T) ΔE to the HF/6-31G* ZPE and thermal corrections) is 0.59 kJ/mol compared with an experimental solution value of 7.1 kJ/mol. The experimental solution value would be expected to favor the species with the higher dipole moment (i.e., benzene oxide), and therefore the experimental gas-phase ΔH should be less than 7.1 kJ/mol. We used self-consistent reaction field theory to quantify the effect of solvation[220]. In this theory, one modifies the Fock operator as

$$F = F_0 - k\mu\langle\mu\rangle,$$

where

$$k = \frac{2(\epsilon - 1)}{2\epsilon + 1}\frac{1}{r^3}.$$

Here, μ is the dipole operator, ϵ the dielectric constant, and r the cavity radius. At the HF/6-31G**//HF/3-21G level we determined that $\frac{\partial \Delta E}{\partial k} = 2.4$MJ bohr^3. For a vacuum, $k = 0$. A reasonable estimate of the radius of the cavity would be 5.0 Å.

The three solvent systems used in the experimental studies are trifluorobromomethane–pentane (2:1) (C), isooctane (A), and water–methanol (85:15) (B). The dielectric constants of water, methanol and pentane at 298 K are 78.5, 32.6 and 1.85 respectively[221]. Isooctane should be nearly identical to n-octane (1.95), given their similarity in refractive index (1.3949, 1.3975) and the excellent relationship $\epsilon = n^2$. For CF_3Br, neither a dielectric constant nor a refractive index could be found. The refractive index was therefore estimated by examining the change in refractive index upon substituting one bromine for iodine (The iodo compounds comprising the test were CH_3I, CH_2ClI, $CHCl_2I$, CCl_3I, CH_2BrI and CH_2I_2.) A linear relationship ($n_I = n_{Br} + 0.10$) was found to apply. The refractive index of CF_3I was 1.3790, giving a predicted refractive index of CF_3Br of 1.28 and an associated dielectric constant of 1.64. From linear interpolation where appropriate, the dielectric constants of the three solutions are estimated to be 1.71, 1.95 and 71.7, respectively, which give k values of 0.000381, 0.000460, and 0.001161. For the system in which ΔH was determined, this suggests a solvent contribution of 0.9 kJ/mol, which improves the agreement to experiment. The other two systems should have solvent contributions of 1.1 and 2.8 kJ/mol, with a difference ($\Delta\Delta G$) of 1.7 kJ/mol.

Vogel has observed the solvent dependence of the valence tautomerism in the latter

two solvent systems from examination of the UV spectra. We may glean a *difference* in the ΔG of the two solvent systems (to compare with the above value for $\Delta\Delta G$) as follows.

The UV spectra give intensities of two bands assigned to benzene oxide (S_{bo}) and oxepin (S_{ox}). Any superscripts denote the solvent system. These may be related to the mole fraction as

$$S_{bo} = \chi_{bo}\epsilon_{bo},$$

where we assume ϵ_{bo} is solvent-independent. We do not know ϵ_{bo}, so we write it in terms of a reference (indane oxide) as

$$\epsilon_{bo} = k_{bo}\epsilon_{io},$$

where k_{bo} is constant. We assume a similar equation for oxepin

$$S_{ox} = \chi_{ox}\epsilon_{ox},$$

with

$$\epsilon_{ox} = k_{ox}\epsilon_{dmo},$$

where our reference is 2,7-dimethyloxepin (Vogel assumed that the extinction coefficients were identical ($k = 1$), while we prefer the more general expression.) The equilibrium constant for the valence tautomerism in solvent A can then be written

$$K^A = \frac{\chi_{ox}}{\chi_{bo}} = \frac{\frac{S^A_{ox}}{k_{ox}\epsilon_{dmo}}}{\frac{S^A_{bo}}{k_{bo}\epsilon_{io}}}.$$

We may write the same expression for solvent system B, and note that

$$\frac{K^B}{K^A} = \frac{\frac{S^B_{gr}}{S^B_{bo}}}{\frac{S^A_{gr}}{S^A_{bo}}}$$

which has no reference to the extinction coefficients. We note that

$$\frac{K^B}{K^A} = \exp\left(-\frac{\Delta G^B - \Delta G^A}{RT}\right).$$

Using Vogel's ratios we compute $\Delta\Delta G = 7.5$ kJ/mol. The sign of the solvent correction is correct, however, there still is some discrepancy. This may be due to a possible significant difference in entropy of the valence tautomerism between the two solvents. In addition, SCRF theory does not explicitly take into account the effects of hydrogen-bonding, which would be very important in the latter solvent (methanol/water).

Benzene Sulfide–Thiepin

The reaction energies and activation barriers pertaining to the valence tautomerism between benzene sulfide and thiepin, and the inversion barrier of thiepin, are presented in Table 52. The scatter in the prediction of the reaction energy is still pronounced. However, all levels predict that the sulfide form is much more stable than thiepin, whereas MNDO predicts them to be nearly isoenergetic[218]. The same trends in enthalpy and activation energy are noted as for benzene oxide–oxepin valence tautomerism. Both the barriers to inversion and to valence tautomerization are much larger than for the oxygen analogue, suggesting that it may be possible to trap thiepin.

Table 51: Benzene oxide system - reaction energies (kJ/mol)

	ΔE^a	$E_{a,+}$	$E_{a,-}$	E_{inv}
STO-3G	59.05	116.23	57.18	13.29
3-21G	-63.45	40.18	103.63	7.69
6-31G*//3-21G	-22.32	57.38	79.70	10.45
6-31G*	-12.05	60.23	72.28	12.23
6-31G**//3-21G	-23.07	56.68	79.75	10.64
MP2/6-31G*//3-21G	19.37	20.95	1.58	18.40
MP2/6-31G*//6-31G*	13.86	19.81	5.95	20.72
MP2/6-31G*	13.90	26.50	12.60	24.74
MP3/6-31G*//MP2	2.82	41.97	39.16	18.13
MP4DQ/6-31G*//MP2	3.74	45.88	42.14	15.51
MP4SDQ/6-31G*//MP2	3.34	46.53	43.18	13.16
MP4SDTQ/6-31G*//MP2	5.41	32.71	27.29	17.42
MP2/6-31G**//MP2	13.55	26.32	12.77	24.75
MP2/6-31G(2d)//MP2	8.37	19.29	10.92	21.71
MP2/6-31+G*//MP2	13.05	23.27	10.22	28.01
MP2/6-311G*//MP2	9.20	19.17	9.96	26.67
QCISD/6-31G*//MP2	1.60	44.25	42.65	13.02
QCISD(T)/6-31G*//MP2	1.43	36.07	34.64	15.60
ZPE corr. (3-21G)	-0.66	-5.35	-4.69	1.12
Thermal corr. (3-21G)	1.15	-0.64	-1.79	-1.79
ΔS (3-21G)	8.55	-3.81	-12.36	-16.75
ZPE corr. (6-31G*)	-2.06	-6.20	-4.14	0.47
Thermal corr. (6-31G*)	1.22	-0.49	-1.71	-1.60
ΔS (6-31G*)	8.34	-3.02	-11.36	-15.09

[a] $E_{a,+}$ = activation barrier (forward), $E_{a,-}$ = activation barrier (reverse), E_{inv} = barrier to inversion.

Table 52: Benzene sulfide system - reaction energies (kJ/mol)

	ΔE	$E_{a,+}$	$E_{a,-}$	E_{inv}
STO-3G	72.68	183.93	111.25	13.24
STO-3G*	74.43	164.30	89.87	13.82
3-21G	51.32	145.84	94.52	33.76
3-21G(*)	12.54	101.49	88.95	34.61
6-31G*	29.79	129.91	100.11	26.76
MP2/6-31G*//6-31G*	47.58	76.36	28.78	40.33
MP2/6-31G*	44.15	79.49	35.34	43.54
MP3/6-31G*//MP2	37.12	106.30	69.18	33.13
MP4DQ/6-31G*//MP2	37.54	107.20	69.66	31.54
MP4SDQ/6-31G*//MP2	35.37	105.77	70.41	30.26
MP4SDTQ/6-31G*//MP2	34.70	85.85	51.14	35.17
MP2/6-31G**//MP2	43.14	78.40	35.26	43.39
MP2/6-31G(2d)//MP2	46.64	76.39	29.74	38.53
MP2/6-31+G*//MP2	38.27	72.51	34.23	47.76
MP2/6-311G*//MP2	44.46	72.55	28.08	47.41
QCISD/6-31G*//MP2	33.60	103.91	70.31	29.78
QCISD(T)/6-31G*//MP2	31.85	92.72	60.87	31.96
ZPE corr. (3-21G)	-2.34	-5.65	-3.31	0.37
Thermal corr. (3-21G)	0.43	-1.35	-1.78	-1.57
ΔS (3-21G)	2.53	-8.81	-11.34	-14.86
ZPE corr. (6-31G*)	-3.56	-6.31	-2.75	0.14
Thermal corr. (6-31G*)	1.03	-0.72	-1.75	-1.59
ΔS (6-31G*)	6.68	-4.66	-11.34	-15.14

The difference between the energies and transition states of the sulfide and oxide systems (which corresponds to an isodesmic process) is shown in Table 53. The difference in the differential between the MP2 and MP4 levels is less than 1 kJ/mol and is reasonable from the HF/6-31G* level upwards. This table presents data which illustrate the accuracy of attempts to predict the value of an isodesmic reaction by using a lower level of theory. Here, MP2 gives essentially the same values as MP4. The utility in such an exercise follows. Suppose we have a large system, A, for which an MP4 calculation cannot be carried out. Assume that the difference between the MP4 and MP2 is primarily a result of the electron correlation difference of the subsystem common to A and a suitably chosen model system (B). Then the MP4 reaction energies can be predicted by calculating the reaction energy at the MP2 level of the system of interest (A) and then adding the MP2 B-A reaction energy difference to the MP4 calculation of the model system B.

Xylene Oxide–Dimethyloxepin

The reaction energies and activation barriers pertaining to the valence tautomerism between *o*-xylene oxide and 2,7-dimethyloxepin, and the inversion barrier of 2,7-dimethyloxepin, are presented in Table 54. The scatter in the prediction of the reaction energy is still pronounced, and the (MP2/6-31G*) result predicts the wrong sign of the enthalpy. The MP4 value is predicted by the method of the previous section, where the differentials are obtained from Table 55 from which we obtain 5.41 - 11.81

Table 53: Benzene oxide - benzene sulfide differential (kJ/mol)

	ΔE	$E_{a,+}$	$E_{a,-}$	E_{inv}
STO-3G	13.64	67.70	54.07	-0.05
STO-3G*	15.38	48.07	32.68	0.53
3-21G	114.77	105.66	-9.11	26.07
3-21G(*)	75.99	44.11	9.24	24.17
6-31G*	41.84	69.67	27.83	14.53
MP2/6-31G*//6-31G*	33.72	56.55	22.84	19.61
MP2/6-31G*	30.25	52.99	22.74	18.80
MP3/6-31G*//MP2	34.30	64.33	30.03	15.00
MP4DQ/6-31G*//MP2	33.80	61.32	27.52	16.03
MP4SDQ/6-31G*//MP2	32.02	59.25	27.22	17.10
MP4SDTQ/6-31G*//MP2	29.29	53.14	23.85	17.75
MP2/6-31G**//MP2	29.59	52.08	22.49	18.64
MP2/6-31G(2d)//MP2	38.27	57.10	18.82	16.82
MP2/6-31+G*//MP2	25.22	49.24	24.01	19.75
MP2/6-311G*//MP2	35.26	53.38	18.12	20.74
QCISD/6-31G*//MP2	32.00	59.66	27.66	16.76
QCISD(T)/6-31G*//MP2	30.42	56.65	26.23	16.36

= -6.40 kJ/mol, which corrects the sign. It is clear that methylation at the 2 and 7 positions slightly increases the inversion barrier. This compares favorably to the actual value of -3.63 kJ/mol.

Protonated Benzene Oxide

The reaction energies and activation barriers pertaining to the valence tautomerism between protonated benzene oxide and oxepin are presented in Table 56. It is immediately clear that protonation stabilizes the oxide form much more than the oxepin form, given the shift in enthalpy to a large positive value. This would suggest also that hydrogen-bonding to benzene oxide would be stronger than to oxepin, which may account for some of the discrepancy in the calculated solvent correction as discused

Table 54: Xylene oxide system - reaction energies (kJ/mol)

	ΔE	$E_{a,+}$	$E_{a,-}$	E_{inv}
STO-3G	40.63	109.18	68.55	20.43
3-21G	-77.41	43.79	121.20	8.51
6-31G*//3-21G	-38.66	58.99	97.64	14.32
6-31G*	-28.85	61.03	89.89	16.19
6-31G**//3-21G	-38.65	58.78	97.44	14.32
MP2/6-31G*//6-31G*	3.30	17.63	14.33	26.14
MP2/6-31G*	2.10	22.37	20.27	30.68
MP3/6-31G*//MP2	-8.54	44.11	52.66	22.76
MP4DQ/6-31G*//MP2	-6.36	47.77	54.12	20.78
MP4SDQ/6-31G*//MP2	-5.78	49.62	55.40	18.81
MP4SDTQ/6-31G*//MP2	-3.63	32.73	36.37	23.71
QCISD/6-31G*//MP2	-7.41	47.71	55.12	18.61
QCISD(T)/6-31G*//MP2	-7.33	38.33	45.66	21.51
ZPE corr. (3-21G)	1.44	-4.74	-6.18	1.20
Thermal corr. (3-21G)	0.62	-0.11	-0.73	-2.04
ΔS (3-21G)	6.67	3.95	-2.72	-20.47
ZPE corr. (6-31G*)	-0.23	-5.34	-5.11	0.07
Thermal corr. (6-31G*)	0.83	-0.11	-0.94	-1.72
ΔS (6-31G*)	8.05	2.25	-5.80	-17.58

Table 55: Benzene oxide - *o*-xylene oxide differential (kJ/mol)

	ΔE	$E_{a,+}$	$E_{a,-}$	E_{inv}
STO-3G	-18.42	-7.05	11.37	7.14
3-21G	-13.96	3.61	17.58	0.81
6-31G*//3-21G	-16.33	1.61	17.94	3.87
6-31G*	-16.81	0.80	17.61	3.96
6-31G**//3-21G	-15.58	2.10	17.68	3.68
MP2/6-31G*//6-31G*	-10.57	-2.18	8.38	5.42
MP2/6-31G*	-11.81	-4.13	7.67	5.94
MP3/6-31G*//MP2	-11.36	2.14	13.50	4.63
MP4DQ/6-31G*//MP2	-10.10	1.89	11.99	5.27
MP4SDQ/6-31G*//MP2	-9.12	3.09	12.21	5.65
MP4SDTQ/6-31G*//MP2	-9.04	0.02	9.06	6.29
QCISD/6-31G*//MP2	-9.01	3.46	12.47	5.58
QCISD(T)/6-31G*//MP2	-8.76	2.26	11.02	5.91

Table 56: Protonated benzene oxide system - reaction energies (kJ/mol)

	ΔE	$E_{a,+}$	$E_{a,-}$
STO-3G	59.29	149.35	90.06
3-21G	-15.73	93.98	109.70
6-31G*	41.46	121.68	80.22
MP2/6-31G*//6-31G*	58.76	65.54	6.78
MP2/6-31G*	54.15	70.56	16.40
ZPE corr. (6-31G*)	-2.66	-4.63	-1.97
Thermal corr. (6-31G*)	0.73	-1.17	-1.90
ΔS (6-31G*)	4.78	-6.85	-11.63

earlier.

Protonated benzene oxide is a *minimum* energy structure, not a transition state as suggested by Glusker and coworkers[217]. They found at the HF/6-31G level an *exo* structure whose Hessian contained a negative eigenvalue corresponding to the formation of the carbocation resulting from rupture of the C-O bond. We found an exo C_s structure as well, which was not characterized by a frequency analysis, because an *endo* structure was 10.91 kJ/mol more stable at the HF/6-31G* level. It is well known that sp-basis sets without polarization functions cannot properly predict the pyramidality of heteroatoms.

Oxygen versus Sulfur Systems

In order to probe more fully the difference between the oxygen and sulfur systems, we analyzed a series of isodesmic reactions. The results are shown in Table 57. Reaction 1 is a probe into the effects that arise when a three membered ring is fused to a diene systems. One might expect some stabilization as a result of some π-σ_{CX}

Table 57: Isodesmic reaction energies (kJ/mol)

	ΔE_1[a]	ΔE_2[b]	ΔE_3[c]
STO-3G	7.63	9.37	23.00
STO-3G*	10.43	-5.29	10.10
3-21G	-17.23	-72.49	42.27
3-21G(*)	2.51	-44.37	31.62
6-31G*	7.87	-15.41	26.43
MP2/6-31G*//6-31G*	0.67	-21.97	11.74
MP2/6-31G*	2.84	-18.14	12.10
MP3/6-31G*//MP2	4.80	-16.10	18.21
MP4DQ/6-31G*//MP2	3.20	-16.82	16.98
MP4SDQ/6-31G*//MP2	3.96	-15.18	16.85
MP4SDTQ/6-31G*//MP2	4.33	-15.11	14.18
MP2/6-31G**//MP2	2.58	-18.79	10.80
MP2/6-31G(2d)//MP2	1.21	-25.14	13.14
MP2/6-31+G*//MP2	0.18	-22.27	2.95
MP2/6-311G*//MP2	2.19	-23.26	12.00
QCISD/6-31G*//MP2	4.00	-15.26	16.74
QCISD(T)/6-31G*//MP2	3.93	-15.75	14.67

[a] ΔE_1 = benzene oxide + C_2H_4S → benzene sulfide + C_2H_4O
[b] ΔE_2 = benzene oxide + C_2H_6S → benzene sulfide + C_2H_6O
[c] ΔE_3 = oxepin + C_2H_6S → thiepin + C_2H_6O

or π-n_X which may be different in the oxygen and sulfur, but the small reaction energy suggests that this does not play any role. Reaction 2 is a probe of the relative stability of the three membered rings with oxygen and sulfur. It indicates that sulfur can be accomodated in a three-membered ring much better than oxygen (15 kJ/mol). Reaction 3 is a measure of the difference in ring strain in the 7-membered ring, which shows that thiepin is destabilized (14 kJ/mol) relative to oxepin. These two effects make up the majority of the difference between the oxepin and thiepin systems.

3.5.3 Ionization Energies

The highest occupied molecular orbitals of benzene oxide are predicted to have energies of 8.80, 10.78, 11.88 and 13.33 eV at the HF/6-31G* level and can be described as π_-, $\pi_+ - \sigma_{CO}$, n_O and $\pi_+ + \sigma_{CO}$. We note that the π_+ and σ_{CO} mix. The first three bands actually observed in the photoelectron spectrum[216] are at 8.43, 10.20 and 11.45 eV, and assuming the same assignment, give an excellent linear correlation (slope = 1.026, intercept=0.195, $R^2 = 0.996$). Modelli and coworkers[216] assign π_+ and n_O in the opposite manner and assumes that there is no mixing. Their HF/3-21G calculation gives the same ordering and nearly identical results as our more sophisticated HF/6-31G* calculation, but it is claimed that the calculation overestimates the stability of the σ_{CO} bond, resulting in apparent mixing. Their assignment is based upon comparisons with other experimental spectra. The persistence of the ordering even at HF/6-31G* suggests that our assignment is the correct one, if the Koopman approximation is valid. The sulfide ionization energies are predicted to occur at 8.36, 9.16, 10.89 and 12.15 eV and are assigned as $\pi_- - \sigma_{CS,-}$, n_S, $\pi_+ - \sigma_{CS,+}$, and $\sigma_{CS,-}$.

Oxepin bands do not appear in Modelli's PE spectrum, but these are predicted to lie at 8.10, 10.42, 11.78 and 13.06 eV (π', π'', π',n_O). The bands of thiepin itself are predicted to appear at 8.19, 10.08, 10.32 and 11.64 eV and are similar to those of thiepin with an inversion between the second and third orbitals. The basic distribution of these is in good agreement with several substituted thiepins[219]. For example, a good linear correlation is found with 2,7-di-*tert*-butylthiepin (slope = 1.36, intercept

= -2.33, R^2 = 0.988, where the inductive effect at the 2 and 7 positions shifts the slope considerably away from unity).

3.5.4 Diels-Alder Reaction of Benzene Oxide

In a Diels-Alder reaction with benzene oxide, a dienophile can attack either from the same side of the diene moiety as the oxygen (*syn*) or on the opposite face (*anti*). Experimentally, with *N*-phenylmaleimide and dimethyl acetylenedicarboxylate the reaction proceeds in an exclusively anti fashion[222]. Ethylenic dienophiles usually react syn to oxygen on both cyclopentadiene[130, 223, 224, 225] and cyclohexadiene systems[226].

We modelled this reaction using ethene and ethyne as simple dienophiles. The isodesmic nature of the syn–anti comparison obviate computational deficiencies and the experimental results are predicted accurately. For ethylene the activation barrier for syn addition is much higher than that for anti addition. The same is true for ethyne. The product of syn addition is slightly more stable than the product of anti addition which rules out product-development control in the facial selectivity. Comparing with the parent cyclohexadiene in the critical bond lengths, the syn and anti transition states are earlier and later than cyclohexadiene, respectively, and the activation energies bracket that of cyclohexadiene. The trend is similar for the acetylene transition states but not as pronounced.

In our preliminary investigation of 5-substituted cyclopentadienes, the activation

Table 58: Diels-Alder reaction energies (kJ/mol)

	E^{syn}_{a,C_2H_4}	E^{anti}_{a,C_2H_4}	$\Delta E^{syn}_{C_2H_4}$	$\Delta E^{anti}_{C_2H_4}$	E^{syn}_{a,C_2H_2}	E^{anti}_{a,C_2H_2}
STO-3G	195.20	147.69	-340.29	-342.07	224.39	173.14
3-21G	161.34	125.65	-154.55	-143.02	203.80	146.88
6-31G*//3-21G	212.05	173.53	-108.63	-104.90	241.17	187.31
6-31G*	212.48	174.15	-107.29	-103.10	241.43	188.04
MP2/6-31G*//HF/6-31G*	78.13	38.14			106.52	52.41
MP2/6-31G*	88.16	44.85			114.83	59.44

barrier was partitioned into three terms

$$E_{\text{act}} = E_{\text{def}}^{\text{diene}} + E_{\text{def}}^{\text{dienophile}} + E_{\text{int}},$$

where the deformation energies are the energies required to distort the appropriate molecule into its transition state geometry. The difference in activation barrier between syn and anti is largely manifested in the $E_{\text{def}}^{\text{diene}}$ term. We carried out this analysis for the benzene oxide + ethylene system, obtaining the syn and anti diene deformation values of 145.2 and 101.3 kJ/mol, respectively, consistent with this term being responsible for the facial selectivity.

Why does the diene portion of benzene oxide deform so much in the syn transition state? The deformation energy is reflected in the geometry changes (HF/6-31G*) of the ethylene transition states. The C-C bond of the oxirane portion shortens by 0.033 Å in the syn transition state but only 0.012Å in the anti. Furthermore, α decreases by 42.9° in the syn versus 28.4° in anti, and β increases by 7.4° for syn but only 2.5° for anti. These geometry changes are consistent with a strong steric interaction leading to deformation of the oxirane system in the syn transition state. This effect is large since the oxirane part of the molecule is nearly perpendicular to the diene plane,

Table 59: Structure of benzene oxide + ethylene syn transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4153	1.3991	1.3947	1.4105
C_1-C_2	1.3672	1.3723	1.3823	1.3871
C_1-C_6	1.5228	1.5011	1.5073	1.4968
C_5-C_6	1.4800	1.4609	1.4438	1.4668
C_5-O	1.4364	1.4823	1.4094	1.4532
C_1-C_7	2.2393	2.2460	2.2283	2.3448
C_7-C_8	1.3594	1.3676	1.3797	1.3762
C_1-C_2-C_3	118.76	118.77	118.26	119.03
C_3-C_4-C_5	117.73	116.72	115.97	116.73
C_4-C_5-C_6	114.25	114.80	114.71	115.52
C_4-C_5-O	121.16	121.58	121.69	122.44
C_5-O-C_6	62.02	59.04	61.62	60.62
C_7-C_1-C_2	97.19	98.50	99.21	97.49
C_7-C_1-C_6	95.57	95.22	96.96	94.97
C_1-C_7-C_8	107.84	107.52	107.28	107.12
C_2-C_1-C_4-C_5	146.06	144.96	142.62	146.60
C_4-C_5-C_6-O	113.04	113.66	113.51	114.17
C_2-C_1-C_4-C_8	109.07	110.50	110.95	109.07

whereas the hydrogens of the opposite face are nearly coplanar, as would be the two methyl groups or a three carbon bridge in the derivatives studied in Reference [222].

3.6 cis-5,6-Disubstituted 1,3-Cyclohexadienes

The study of cis-5,6-disubstituted 1,3-cyclohexadienes (shown in Figure 8) is of fundamental importance in understanding the wide range of facial selectivity in their Diels-Alder reactions[226, 227]. Unlike in the more rigid 1,3-cyclopentadiene, there exists a ring inversion process for 1,3-cyclohexadienes. This inversion can be modified by the presence of substituents even to the point where it no longer exists (as in benzene oxide, discussed previously).

Table 60: Structure of benzene oxide + ethylene anti transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4036	1.3817	1.3792	1.3953
C_1-C_2	1.3760	1.3835	1.3922	1.3986
C_1-C_6	1.4951	1.4711	1.4808	1.4861
C_5-C_6	1.5016	1.4896	1.4649	1.4935
C_5-O	1.4377	1.4751	1.4046	1.4438
C_1-C_7	2.2227	2.2095	2.2010	2.2600
C_7-C_8	1.3598	1.3743	1.3846	1.3839
C_1-C_2-C_3	118.83	118.70	118.26	118.63
C_3-C_4-C_5	118.30	118.28	117.91	117.59
C_4-C_5-C_6	114.27	114.52	114.59	115.02
C_4-C_5-O	117.03	116.62	117.70	117.70
C_5-O-C_6	62.96	60.65	62.86	62.29
C_7-C_1-C_2	98.70	99.76	100.18	100.16
C_7-C_1-C_6	92.71	91.75	92.92	91.97
C_1-C_7-C_8	107.96	107.60	107.35	107.40
C_2-C_1-C_4-C_5	147.32	147.52	146.14	146.85
C_4-C_5-C_6-O	107.96	107.68	108.63	108.50
C_2-C_1-C_4-C_8	111.10	112.11	112.22	112.43

Table 61: Structure of benzene oxide + acetylene syn transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4136	1.3987	1.3969	1.4123
C_1-C_2	1.3694	1.3730	1.3798	1.3855
C_1-C_6	1.5251	1.5045	1.5102	1.4999
C_5-C_6	1.4787	1.4601	1.4421	1.4646
C_5-O	1.4365	1.4776	1.4056	1.4496
C_1-C_7	2.1969	2.2182	2.2135	2.2988
C_7-C_8	1.2011	1.2189	1.2219	1.2456
C_1-C_2-C_3	118.35	118.49	118.01	118.76
C_3-C_4-C_5	117.68	116.54	115.93	116.77
C_4-C_5-C_6	113.90	114.51	114.46	115.27
C_4-C_5-O	121.64	122.44	122.56	123.01
C_5-O-C_6	61.95	59.22	61.73	60.69
C_7-C_1-C_2	96.44	96.82	97.19	95.75
C_7-C_1-C_6	95.39	96.03	97.72	95.61
C_1-C_7-C_8	110.14	109.62	109.41	109.04
C_2-C_1-C_4-C_5	144.73	143.74	141.81	145.83
C_4-C_5-C_6-O	113.78	114.77	114.64	114.95
C_2-C_1-C_4-C_8	109.49	109.69	109.73	108.07

Table 62: Structure of benzene oxide + acetylene anti transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4010	1.3820	1.3810	1.3974
C_1-C_2	1.3786	1.3827	1.3893	1.3958
C_1-C_6	1.4969	1.4741	1.4827	1.4708
C_5-C_6	1.4999	1.4840	1.4597	1.4879
C_5-O	1.4381	1.4763	1.4060	1.4458
C_1-C_7	2.1855	2.2019	2.2009	2.2369
C_7-C_8	1.2016	1.2242	1.2267	1.2512
C_1-C_2-C_3	118.42	118.49	118.05	118.44
C_3-C_4-C_5	118.39	118.30	118.10	117.81
C_4-C_5-C_6	113.91	114.38	114.46	114.91
C_4-C_5-O	116.97	116.38	117.42	117.45
C_5-O-C_6	62.86	60.34	62.54	61.94
C_7-C_1-C_2	98.22	98.89	99.02	99.31
C_7-C_1-C_6	92.10	91.22	92.31	91.08
C_1-C_7-C_8	110.23	109.59	109.39	109.26
C_2-C_1-C_4-C_5	146.22	146.95	145.97	146.78
C_4-C_5-C_6-O	108.09	107.51	108.40	108.31
C_2-C_1-C_4-C_8	111.88	112.33	112.08	112.60

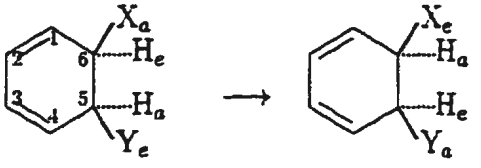
Table 63: Structure of benzene oxide + ethylene syn product

Parameter	STO-3G	3-21G	6-31G*
C_2-C_3	1.3115	1.3196	1.3223
C_1-C_2	1.5328	1.5238	1.5225
C_1-C_6	1.5496	1.5290	1.5334
C_5-C_6	1.4793	1.4618	1.4420
C_5-O	1.4350	1.4788	1.4110
C_1-C_7	1.5577	1.5540	1.5481
C_7-C_8	1.5573	1.5582	1.5498
C_1-C_2-C_3	114.70	114.58	114.23
C_3-C_4-C_5	106.50	106.29	106.04
C_4-C_5-C_6	111.05	111.59	111.62
C_4-C_5-O	116.96	115.87	117.14
C_5-O-C_6	62.06	59.24	61.46
C_7-C_1-C_2	107.17	107.18	107.07
C_7-C_1-C_6	107.71	107.54	108.35
C_1-C_7-C_8	109.41	109.34	109.27
C_2-C_1-C_4-C_5	120.80	120.84	120.29
C_4-C_5-C_6-O	109.60	108.35	109.59
C_2-C_1-C_4-C_8	120.44	120.31	119.89

Table 64: Structure of benzene oxide + ethylene anti product

Parameter	STO-3G	3-21G	6-31G*
C_2-C_3	1.3114	1.3164	1.3188
C_1-C_2	1.5258	1.5101	1.5103
C_1-C_6	1.5491	1.5291	1.5344
C_5-C_6	1.4795	1.4667	1.4465
C_5-O	1.4335	1.4701	1.4036
C_1-C_7	1.5620	1.5634	1.5550
C_7-C_8	1.5589	1.5632	1.5543
C_1-C_2-C_3	114.90	114.96	114.58
C_3-C_4-C_5	108.16	109.25	109.51
C_4-C_5-C_6	111.12	111.56	111.58
C_4-C_5-O	117.52	117.64	118.74
C_5-O-C_6	62.14	59.85	62.03
C_7-C_1-C_2	107.32	107.21	107.31
C_7-C_1-C_6	105.47	104.02	104.08
C_1-C_7-C_8	109.39	109.19	109.16
C_2-C_1-C_4-C_5	123.21	125.09	125.16
C_4-C_5-C_6-O	110.21	111.37	111.44
C_2-C_1-C_4-C_8	120.74	120.50	120.36

Figure 8: cis-5,6-Disubstituted 1,3-cyclohexadiene numbering



3.6.1 1,3-Cyclohexadiene

Geometries

Experimental probes into the structure of 1,3-cyclohexadiene began with the microwave investigation by Butcher[228], who found that the molecule possessed a C_2 point group and had a dipole moment of 0.437 ± 0.014 D. All bond lengths and angles were fixed, and the only varying parameter was the torsion flanked by the two double bonds and the skeletal angles. An important deduction was the nonplanarity of the ring. A subsequent refinement did not change the structural conclusions[229].

An electron diffraction structure derived using the moments of inertia obtained from the microwave spectra of reference [228] was determined, with the only assumption being that C-H bonds of the same type are identical[230]. However, it has been pointed out by Traetteberg[231] that some of these values were inconsistent and another set of values was recommended. Apparently unaware of Traetteberg's work, another determination of the structure of 1,3-cyclohexadiene was undertaken which agreed very well with reference [231]. All predict a twisted skeleton, consistent with the data in Table 65, as does a ^{1}H NMR coupling constant analysis[233].

Our calculated C_2 geometries are shown in Table 66. All theoretical geometries are reasonably similar. Lengths of single bond and valence angles are similar. A modest lengthening of the double bonds is apparent upon proceeding to the correlated level, as was noted for 1,3-cyclopentadiene, also. The ring torsions are greatest at the MP2/6-31G*. The MP2/6-31G* structure is very close to the electron diffraction

Table 65: Experimental 1,3-cyclohexadiene structures

Parameter	Ref. [228][a]	Ref. [230][b]	Ref. [231][c]	Ref. [232]
C_2-C_3	1.47*	1.468(8)	1.465(2)	1.468(14)
C_1-C_2	1.34*	1.339(1)	1.348(1)	1.350(4)
C_1-C_6	1.50*	1.494(17)	1.519(1)	1.523(16)
C_5-C_6	1.50*	1.510(32)	1.538(3)	1.534(20)
C_2-H	1.086*	1.07(1)		1.082(10)
C_1-H	1.086*	*	1.099(4)	
C_5-H_a[d]	1.10*	1.14(1)	1.111(3)	1.096(10)
C_5-H_e	1.10*	*		
C_1-C_2-C_3	120.2	121.6(10)	120.26(23)	120.13(60)
C_3-C_4-C_5	*	118.2(7)	120.25	120.14(50)
C_4-C_5-C_6	110.5	111.5(5)	110.88	110.7
H-C_2-C_3	116*	117(10)	118(4)	118.0(14)
H-C_1-C_2	122*	127(14)	*	122.0(14)
H_a-C_5-H_e	109.47*	99(6)		114.1
C_1-C_2-C_3-C_4	17.5(20)	17	17.96(12)	18.34
C_1-C_6-C_5-C_4	45			

[a] * - constant ; C_3-C_4-C_5 = C_1-C_2-C_3 ; planar double bonds

[b] C_2-H = C_1-H ; C_5-H_a = C_5-H_e

[c] H-C_1-C_2 = H-C_2-C_3

[d] a - axial ; e - equatorial

Table 66: Theoretical 1,3-cyclohexadiene structures

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4885	1.4755	1.4750	1.4635
C_1-C_2	1.3146	1.3221	1.3242	1.3489
C_1-C_6	1.5267	1.5189	1.5111	1.5067
C_5-C_6	1.5481	1.5440	1.5331	1.5287
C_2-H	1.0825	1.0729	1.0762	1.0876
C_1-H	1.0834	1.0733	1.0765	1.0878
C_5-H_a	1.0883	1.0885	1.0905	1.1016
C_5-H_e	1.0916	1.0834	1.0855	1.0952
C_1-C_2-C_3	120.85	120.66	120.60	120.24
C_3-C_4-C_5	121.72	120.54	120.71	119.66
C_4-C_5-C_6	112.37	110.99	111.75	110.61
H_a-C_5-H_e	106.67	107.50	106.36	106.69
C_1-C_2-C_3-C_4	11.67	15.24	14.13	16.04
C_1-C_6-C_5-C_4	35.83	44.17	41.96	49.93

structures of Traetteberg[231] and Oberhammer and Bauer[232].

The geometry of the planar C_{2v} form of cyclohexadiene is given in Table 67. The only significant differences with the C_2 structure besides the ring torsions are a significant lengthening of the C_5-C_6 bond length by about 0.02 Å and a widening of the angles, especially C_4-C_5-C_6.

We may also compare our theoretical structures to those previously determined. Our STO-3G structure is similar to that of Ref. [234] but with a slightly lower energy. Our 3-21G structure is identical with that of Ref. [235] and similar to the split valence structure of Ref. [236]. The MIDI-4 structure of Ref. [237] is similar to the Hartree-Fock structures derived here. Our work is the first optimization of this species with polarization functions or with a correlated wavefunction.

Table 67: Theoretical C_{2v} planar 1,3-cyclohexadiene structures

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4859	1.4733	1.4742	1.4638
C_1-C_2	1.3133	1.3185	1.3211	1.3438
C_1-C_6	1.5253	1.5162	1.5094	1.5066
C_5-C_6	1.5580	1.5648	1.5525	1.5530
C_2-H	1.0824	1.0728	1.0761	1.0875
C_1-H	1.0838	1.0743	1.0773	1.0888
C_5-H	1.0900	1.0857	1.0873	1.0975
C_1-C_2-C_3	121.27	121.50	121.26	121.13
C_3-C_4-C_5	123.69	123.39	123.39	123.30
C_4-C_5-C_6	115.04	115.10	115.35	115.57
H-C_5-H	105.92	106.23	105.16	104.91

Vibrational Frequencies

The calculated vibrational frequencies (unscaled) are shown in Table 68. There are a couple of inversions upon going from 3-21G to 6-31G*. The only theoretical frequencies published are those in Ref. [237], in which the assignment order agrees with ours. Reasonable agreement also exists between our scaled frequencies and the experimental values of Di Lauro and coworkers[238] and of Carriera and coworkers[239], although assigning complicated spectra such as these is seldom error-free. A noteworthy accomplishment is the estimate of the classical barrier to inversion to be 13.1 $\pm$ 0.6 kJ/mol based on the observation of several overtones of the lowest vibrational mode (corresponding to the reaction coordinate).

Table 68: Theoretical 1,3-cyclohexadiene frequencies

Frequency		Symmetry	Mode
6-31G*	3-21G		
199.6	198.7	A	Ring oop[a] def.
332.7	329.9	B	Ring oop def.
518.6	527.3	B	Ring ip def.
559.1	571.2	A	Ring oop def.
609.4	623.5	A	Ring ip def.
745.6	764.9	B	+ + ++ H-C-C oop bend + CH_2 rock
836.6	844.1	B	+ + ++ H-C-C oop bend + CH_2 rock
865.9	868.5	A	+ + −− H-C-C oop bend
908.8	898.0	A	+ + ++ C-C str.
997.3	966.0	B	+0 − 0 C-C str.
1020.2	988.6	A	0 + 0− C-C str.
1081.8	1105.4	B	+ − + − +− ip def.
1105.9	1116.7	A	+ − +− H-C-C oop bend
1110.7	1148.8	B	+ − −+ H-C-C oop bend
1143.7	1133.7	A	+ − −+ C str.
1160.0	1180.9	A	CH_2 rock
1277.6	1311.0	A	H-C-C ip bend
1305.0	1333.9	B	H-C-C ip bend + CH_2 twist
1307.4	1341.7	B	H-C-C ip bend + CH_2 twist
1387.4	1400.7	A	H-C-C ip bend + CH_2 twist
1484.9	1511.9	B	H-C-C ip bend + CH_2 wag
1508.4	1482.8	A	CH_2 wag
1541.3	1537.8	B	H-C-C ip bend + CH_2 wag
1570.9	1572.4	A	H-C-C ip bend
1621.8	1637.8	B	CH_2 scissor
1636.9	1646.8	A	CH_2 scissor
1828.6	1803.5	A	C=C str.
1880.2	1861.3	B	C=C str.
3164.4	3171.9	A	C-H_a str.
3171.7	3180.3	B	C-H_a + C-H_e str.
3236.5	3244.1	A	C-H_e str.
3237.6	3245.8	B	C-H_a + C-H_e str.
3341.3	3335.1	B	+ − +− =C-H str.
3348.8	3342.0	A	+ − −+ =C-H str.
3365.9	3360.7	B	+ + −− =C-H str.
3376.7	3372.8	A	+ + ++ =C-H str.

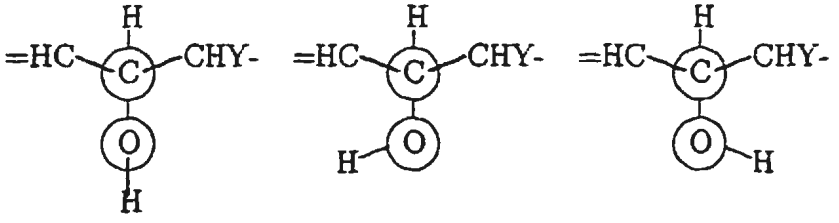
[a]oop refers to out-of-plane, ip refers to in-plane

Table 69: Theoretical 1,3-cyclohexadiene inversion barrier (kJ/mol)

STO-3G	4.07
3-21G	9.34
6-31G*//3-21G	8.34
6-31G*	8.47
6-31G**//3-21G	8.32
MP2/6-31G*//3-21G	14.79
MP2/6-31G*//6-31G*	14.40
MP2/6-31G*	15.67
MP3/6-31G*//MP2	12.30
MP4DQ/6-31G*//MP2	11.95
MP4SDQ/6-31G*//MP2	11.98
MP4SDTQ/6-31G*//MP2	13.49
Experiment.	13.1

Conformation Energies

The electronic barriers to inversion are shown in Table 69. The value reported by Carriera and coworkers[239], because of the way it was determined, is an **electronic barrier without ZPE or entropic terms** present, which is not always appreciated[235]. Fortunately, the ZPE and entropy corrections in this case were found to be very small, so an experimental estimate based upon a rate measurement should agree with that determined by spectroscopy. We obtain excellent agreement with experiment from our best calculation. Sygula[235] is correct in stating that the Hartree-Fock calculations at the 3-21G and 6-311G**//3-21G levels agree very well. However, in general, at the post Hartree-Fock level, polarization functions are needed for an accurate treatment of geometry and energy.



staggered gauche out gauche in

Figure 9: Conformation of 5,6-disubstituted cyclohexadienes

3.6.2 2,4-Cyclohexadien-1-ol

In order to understand the conformational behavior of the cis-5,6-disubstituted 1,3-cyclohexadienes, we first proceed with a calculation of 2,4-cyclohexadien-1-ol. The hydroxyl substituent can be in either an axial or an equatorial position, and in one of three conformations (with respect to the vicinal hydrogen, we have staggered, gauche-in and gauche-out, see Figure 9). Hyperconjugation arguments would suggest that the staggered form would be preferred, with the two gauche structures being close in energy. However, the 'gauche-in' form could be destabilized by a small steric interaction between the hydrogens of the OH and the syn CH of the CH_2, which would not exist for the other conformers. In addition, because of the absence of 1,3-diaxial interactions, there should be no overwhelming preference for an equatorial substituent position.

Conformational Energies

The relative energies of the conformers are found in Table 70. We find, at all levels examined, that the axial staggered conformer is preferred. At the 3-21G level

Table 70: Relative energies - 2,4-cyclohexadiene-1-ol (kJ/mol)

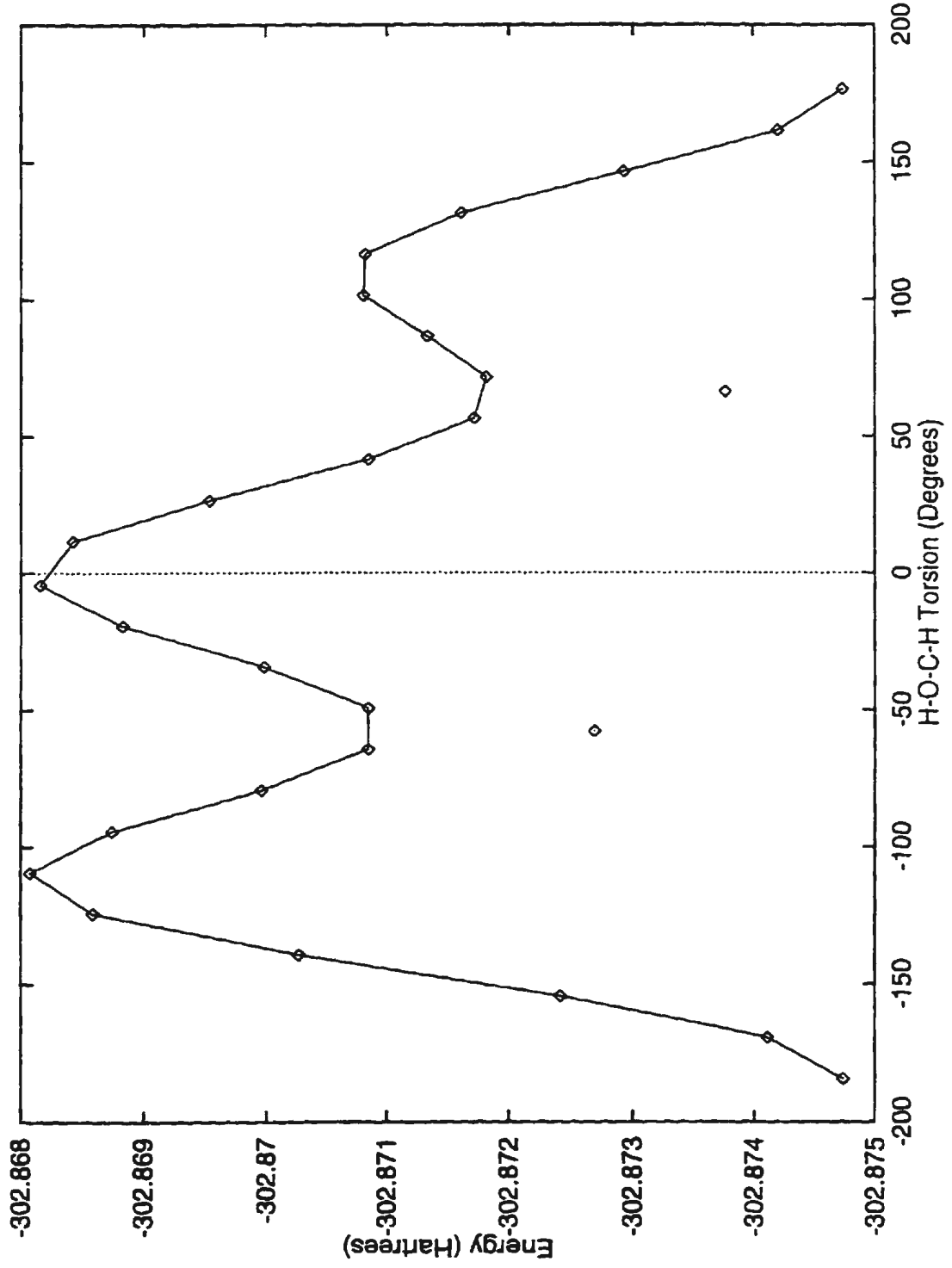
OH position	OH conf.	STO-3G	3-21G	6-31G*//3-21G	6-31G*
axial	staggered	0.00	0.00	0.00	0.00
	gauche-out	2.57			7.31
	gauche-in	5.36	15.80	10.60	10.11
equatorial	staggered	3.38	1.43	2.89	2.67
	gauche-out	4.63	6.72	6.59	6.15
	gauche-in	4.90	5.84	5.63	5.09

the axial gauche-out conformer could not be located, and it is believed that this level of theory predicts no minimum attributable to this 'intuitive' conformer, as rigid rotor scans performed starting at the staggered geometry give no minimum at this position (Figure 11), unlike for STO-3G (Figure 10). The equatorial conformers all gave distinct minima, as shown in Figures 13 and 14. The 6-31G* rigid rotor curves mimic the 3-21G curves (as shown in Figures 12 and 15), however the axial curve near the hidden minimum is a broad inflection region. The points off the curve represent the *optimized* gauche conformers and illustrate the accuracy (or lack thereof) of the rigid rotor approximation.

Geometries

In Tables 71 and 72 the geometry of the various conformers of 2,4-cyclohexadien-1-ol are presented. The changes in the length of the C-C bond geminal to the C-O bond are consistent with strong $n_O \rightarrow \sigma^*_{CC}$ hyperconjugation. The most unusual aspect of the geometries is the value of the H-O-C-H torsion in the axial gauche out conformer (80.50°). In this particular conformer a 1,3-diaxial interaction (repulsive)

Figure 10: 2,4-Cyclohexadien-1-ol axial rigid rotor scan (STO-3G)



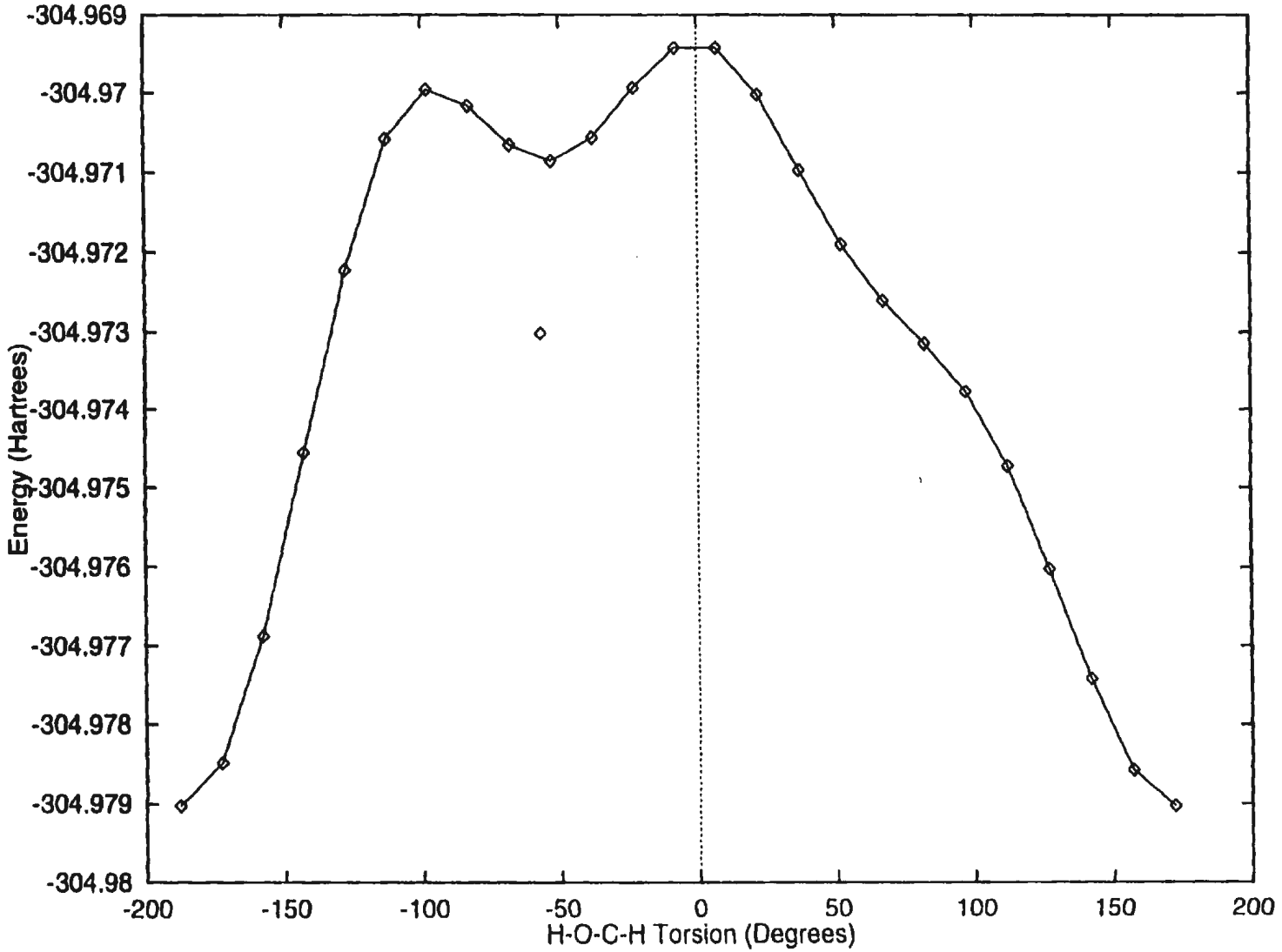


Figure 11: 2,4-Cyclohexadien-1-ol axial rigid rotor scan (3-21G)

Figure 12: 2,4-Cyclohexadien-1-ol axial rigid rotor scan (6-31G*)

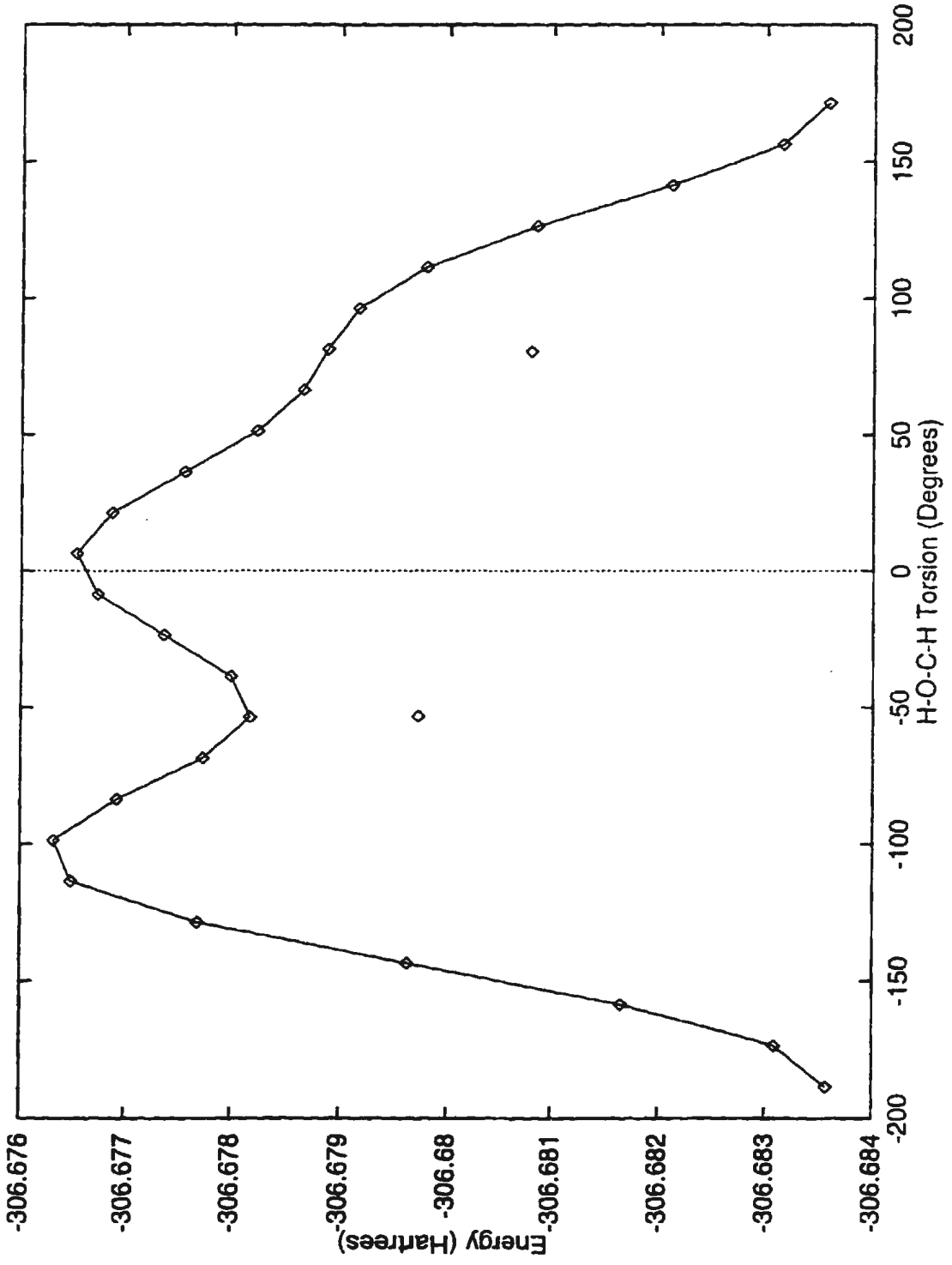


Figure 13: 2,4-Cyclohexadien-1-ol equatorial rigid rotor scan (STO-3G)

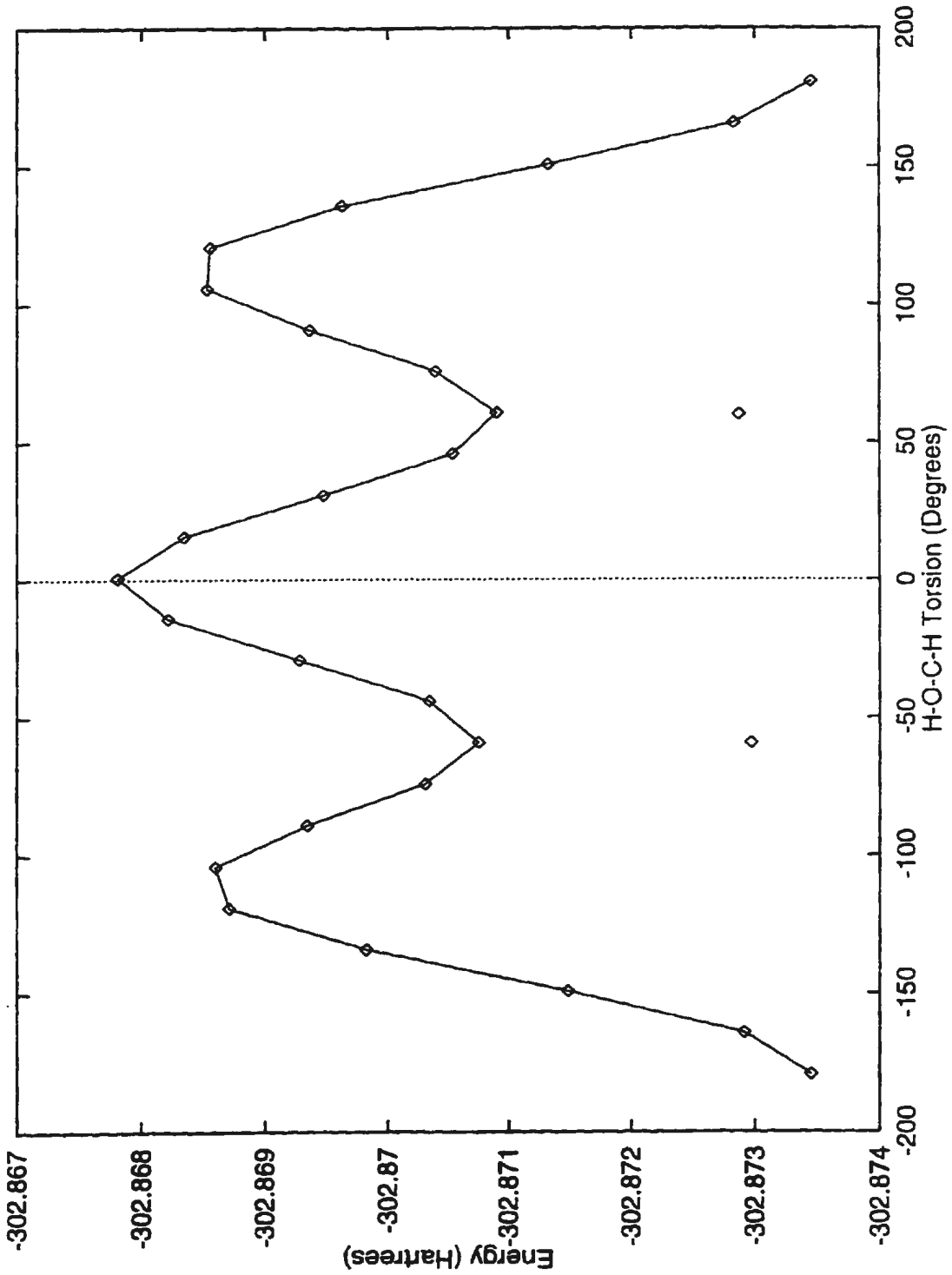


Figure 14: 2,4-Cyclohexadien-1-ol equatorial rigid rotor scan (3-21G)

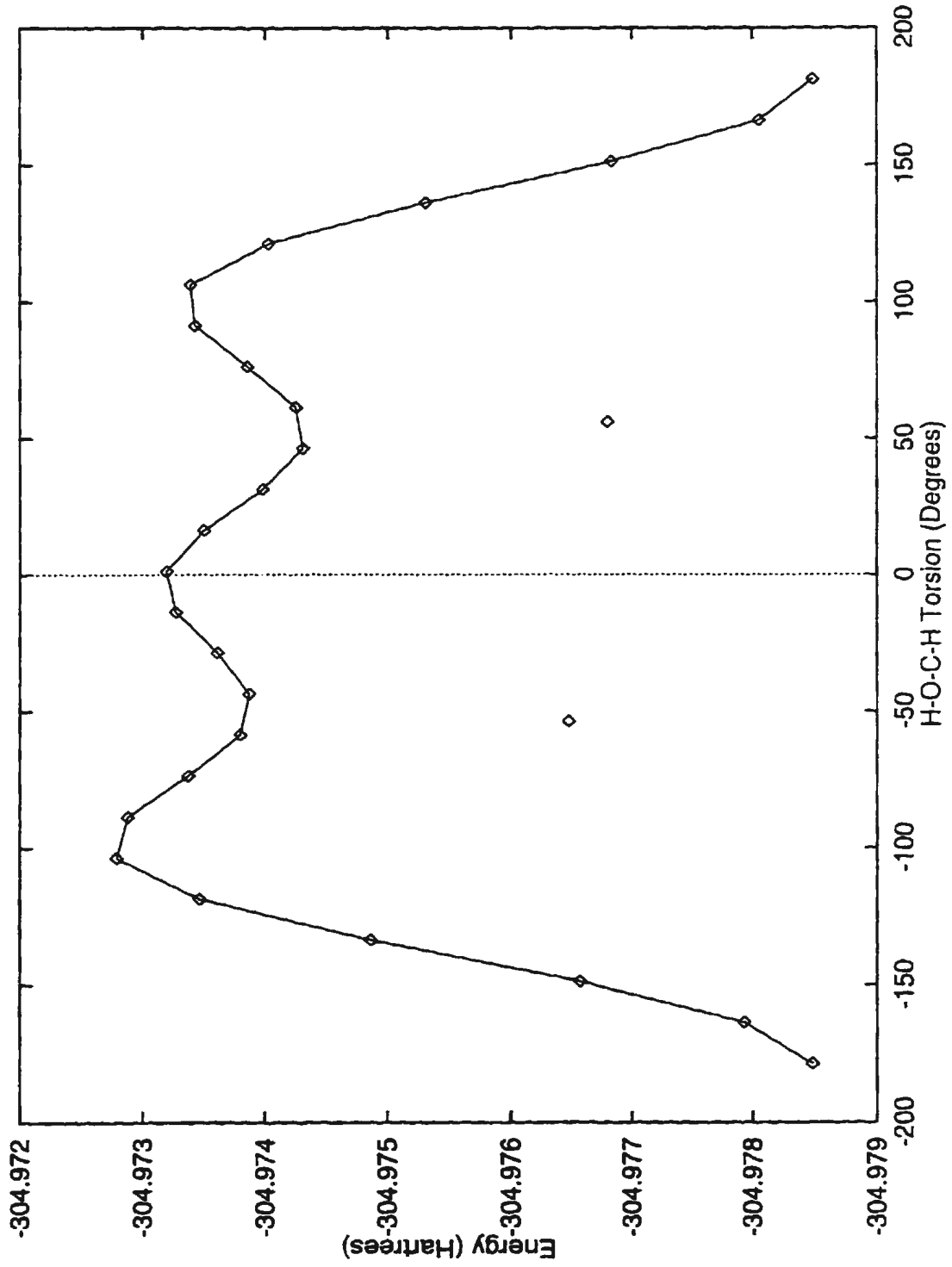
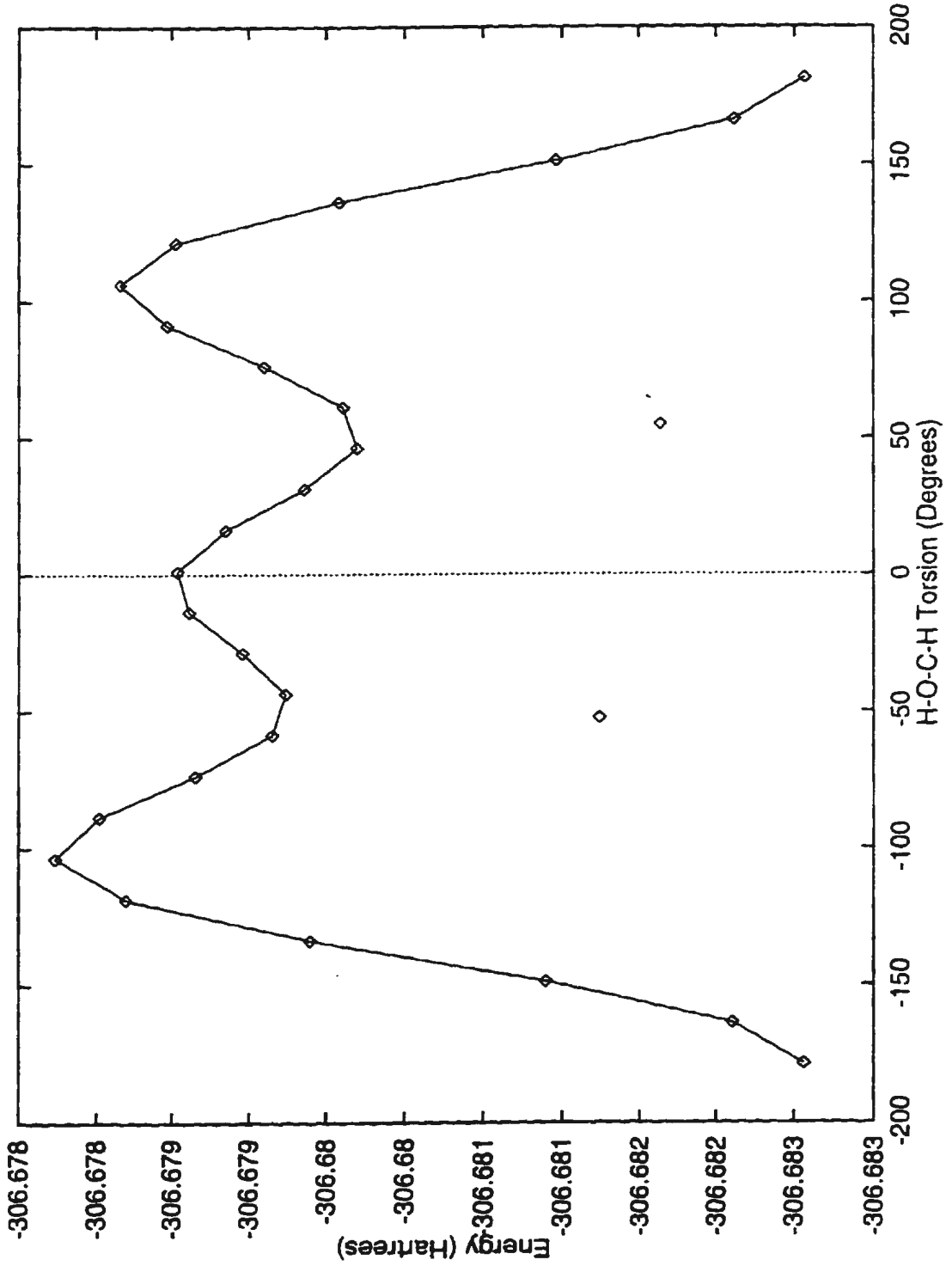


Figure 15: 2,4-Cyclohexadien-1-ol equatorial rigid rotor scan (6-31G*)



between an oxygen lone pair and the syn hydrogen on C_6 is apparent. Another noteworthy geometrical parameters is the C_1-C_6-C_5-C_4 torsion, which because of ring strain, cannot achieve the 'ideal' 60° rotation, but is between 37° and 45 °, with axial substitution preferring a smaller torsion. Similarly, the C_1-C_2-C_3-C_4 torsion is between 11° and 15°, with axial substitution preferring a smaller torsion. The values of the two torsions for cyclohexadiene itself is closer to those of the equatorial conformers, suggesting that axial substitution tends to flatten out the ring a little.

3.6.3 cis-3,5-Cyclohexadiene-1,2-diol

Given an understanding of the rigid rotor profiles of the axial and equatorial alcohols, we may predict, in principle, a rigid rotor profile of the diol by creating a 'direct sum' of the domain to obtain a two dimensional grid, the range of which would be the sum of the appropriate components. Any deviation from the true profile should be rationaled by steric interactions and/or internal hydrogen-bond formation.

The actual potential of the diol as a function of the two H-O-C-H torsions, is represented in a three-dimensional and contour fashion in Figures 16 and 17, respectively. The energy is evaluated on the direct sum grid points, but the optimized structure corresponding to the lowest point is used for the remaining parameters. At

Table 71: Theoretical 2,4-cyclohexadien-1-ol axial structures

	Parameter	STO-3G	3-21G	6-31G*
	C_1-C_6	1.5244	1.5145	1.5078
	C_4-C_5	1.5348	1.5163	1.5129
	C_5-C_6	1.5580	1.5359	1.5321
	O-C	1.4417	1.4486	1.4107
	H-O	0.9915	0.9675	0.9484
	C_1-C_2-C_3	121.31	120.55	120.46
Axial	C_2-C_3-C_4	120.78	121.23	121.35
stag.[a]	C_4-C_5-C_6	111.54	111.20	111.51
	C_5-C_6-C_1	113.07	112.63	113.25
	O-C_5-C_6	111.80	111.18	112.00
	H-O-C	103.91	109.02	108.75
	C_1-C_2-C_3-C_4	10.79	12.12	11.87
	C_1-C_6-C_5-C_4	35.52	41.05	40.51
	H-O-C-H	176.61	172.13	171.41
	C_1-C_6	1.5239		1.5068
	C_4-C_5	1.5347		1.5135
	C_5-C_6	1.5526		1.5267
	O-C	1.4432		1.4136
	H-O	0.9915		0.9472
Axial	C_1-C_2-C_3	120.73		120.22
gauche	C_2-C_3-C_4	121.29		121.26
out	C_4-C_5-C_6	112.08		111.73
	C_5-C_6-C_1	113.27		113.35
	O-C_5-C_6	106.97		107.20
	H-O-C	104.10		109.20
	C_1-C_2-C_3-C_4	10.43		11.57
	C_1-C_6-C_5-C_4	33.39		39.23
	H-O-C-H	66.38		80.50
	C_1-C_6	1.5235	1.5126	1.5064
	C_4-C_5	1.5322	1.5113	1.5084
	C_5-C_6	1.5585	1.5382	1.5340
	O-C	1.4419	1.4484	1.4111
	H-O	0.9917	0.9678	0.9479
Axial	C_1-C_2-C_3	120.79	120.54	120.45
gauche	C_2-C_3-C_4	121.33	121.37	121.43
in	C_4-C_5-C_6	111.68	111.25	111.76
	C_5-C_6-C_1	113.22	112.83	113.74
	O-C_5-C_6	112.29	112.01	112.32
	H-O-C	103.94	110.12	109.27
	C_1-C_2-C_3-C_4	10.82	12.14	11.48
	C_1-C_6-C_5-C_4	34.64	40.29	37.31
	H-O-C-H	152-57.93	-57.22	-53.34

[a]With respect to vicinal hydrogen.

Table 72: Theoretical 2,4-cyclohexadien-1-ol equatorial structures

	Parameter	STO-3G	3-21G	6-31G*
	C_1-C_6	1.5257	1.5172	1.5106
	C_4-C_5	1.5371	1.5147	1.5120
	C_5-C_6	1.5584	1.5348	1.5299
	O-C	1.4377	1.4429	1.4052
	H-O	0.9913	0.9675	0.9483
	C_1-C_2-C_3	120.86	120.42	120.42
Equat.	C_2-C_3-C_4	121.03	120.54	120.64
stag.	C_4-C_5-C_6	111.78	110.74	111.37
	C_5-C_6-C_1	113.01	111.27	112.36
	O-C_5-C_6	111.57	111.16	111.41
	H-O-C	104.07	109.69	109.12
	C_1-C_2-C_3-C_4	10.95	14.58	13.37
	C_1-C_6-C_5-C_4	34.00	44.53	41.26
	H-O-C-H	180.76	181.35	181.26
	C_1-C_6	1.5253	1.5174	1.5108
	C_4-C_5	1.5377	1.5165	1.5133
	C_5-C_6	1.5529	1.5272	1.5232
	O-C	1.4388	1.4433	1.4063
	H-O	0.9911	0.9665	0.9471
Equat.	C_1-C_2-C_3	120.75	120.21	120.22
gauche	C_2-C_3-C_4	121.05	120.39	120.43
out	C_4-C_5-C_6	111.86	110.40	111.04
	C_5-C_6-C_1	113.13	110.73	111.75
	O-C_5-C_6	106.60	106.11	107.25
	H-O-C	104.13	110.78	109.70
	C_1-C_2-C_3-C_4	10.89	15.06	14.06
	C_1-C_6-C_5-C_4	33.76	46.70	43.56
	H-O-C-H	-59.35	-53.65	-52.51
	C_1-C_6	1.5257	1.5187	1.5117
	C_4-C_5	1.5328	1.5102	1.5074
	C_5-C_6	1.5576	1.5324	1.5280
	O-C	1.4383	1.4435	1.4061
	H-O	0.9912	0.9661	0.9471
Equat.	C_1-C_2-C_3	120.78	120.30	120.34
gauche	C_2-C_3-C_4	120.89	120.33	120.49
in	C_4-C_5-C_6	111.53	110.26	110.94
	C_5-C_6-C_1	112.51	110.47	111.52
	O-C_5-C_6	111.80	111.50	111.84
	H-O-C	103.97	110.54	109.45
	C_1-C_2-C_3-C_4	12.04	15.69	14.51
	C_1-C_6-C_5-C_4	36.42	47.63	44.68
	H-O-C-H	153 59.99	56.11	54.95

about (-100°,135°) there exists a strong steric interaction between the two hydrogens resulting in a strong maximum, overwhelming the predicted (gauche in, staggered) minimum at about (-60°,180°). The (staggered, staggered) conformer corresponds to a likely hydrogen-bonding scenario, and using the isodesmic reaction, 2,4-cyclohexadien-1-ol axial staggered + 2,4-cyclohexadien-1-ol equatorial staggered → cis-3,5-cyclohexadiene-1,2-diol + cyclohexadiene, at STO-3G, the energy is predicted to be about -376.704350 Hartrees, compared with the obtained value of -376.705770 Hartrees. This suggests that hydrogen-bond formation only results in a net stabilization of 3.7 kJ/mol. This is an order of magnitude smaller than the corresponding water dimer result[43] and may be related to the inability to achieve a linear configuration. One difficulty in this interpretation, however, is the lack of accounting for destabilizing factors such as dipole-dipole interactions (C-O with C-O) and lone-pair repulsion. These factors are difficult to assess in general, but the latter cannot be present in the H-bonded structure. Perhaps a better method is to use the difference between the point predicted by the mono alcohols and the actual value, which gives 12.2 kJ/mol.

The predicted minima are represented in Table 73. The position of the (staggered,staggered) minimum is in reasonable agreement with the prediction from the scan and in moderate agreement with those predicted from the mono forms, i.e. (176.61, 180.76). The equatorial hydrogen 'swings' in, presumably to H-bond more effectively with the axial oxygen's lone-pair. The four structures which can H-bond

Figure 16: cis-3,5-Cyclohexadiene-1,2-diol rigid rotor scan (STO-3G)

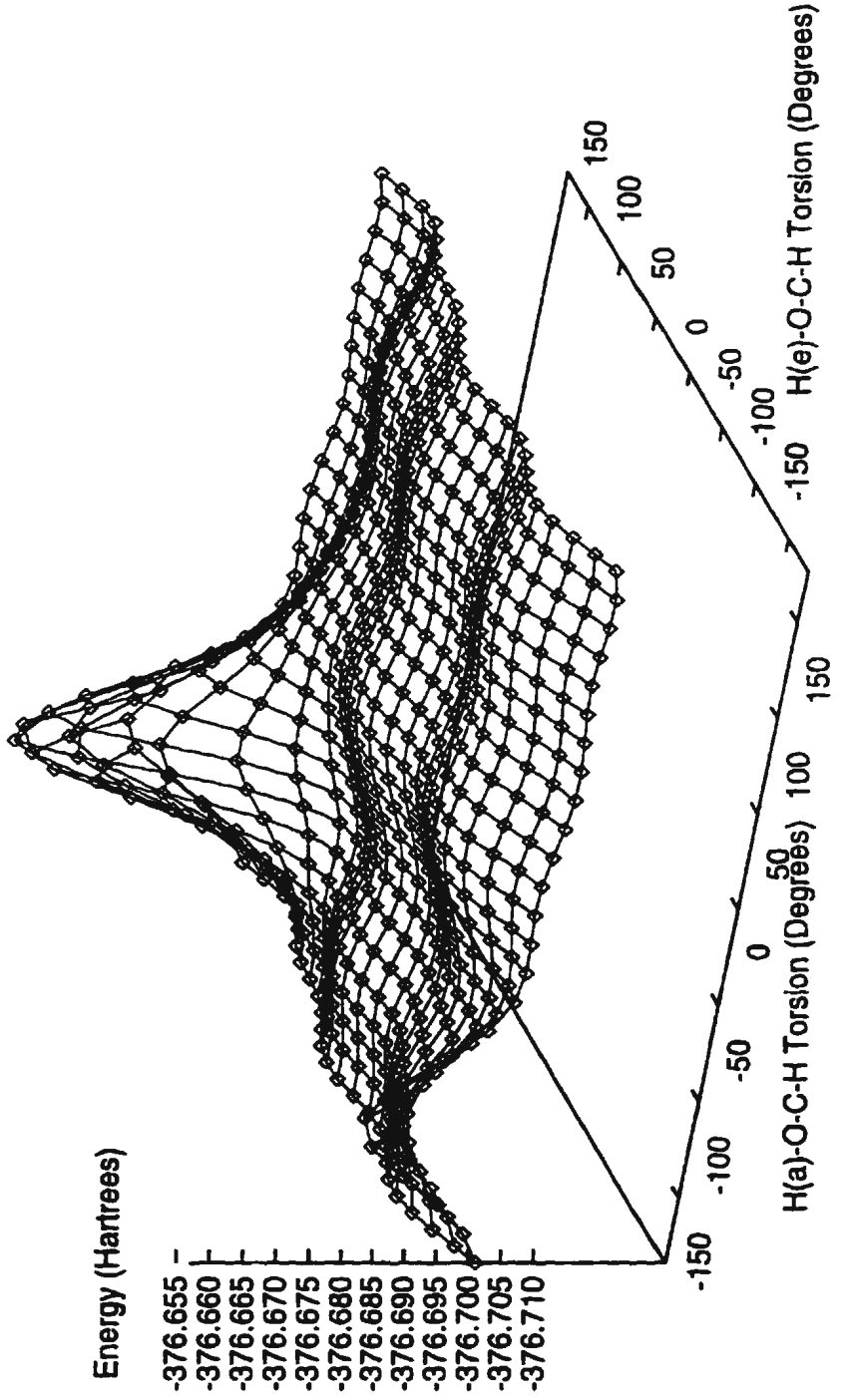


Figure 17: cis-3,5-Cyclohexadiene-1,2-diol rigid rotor scan (STO-3G)

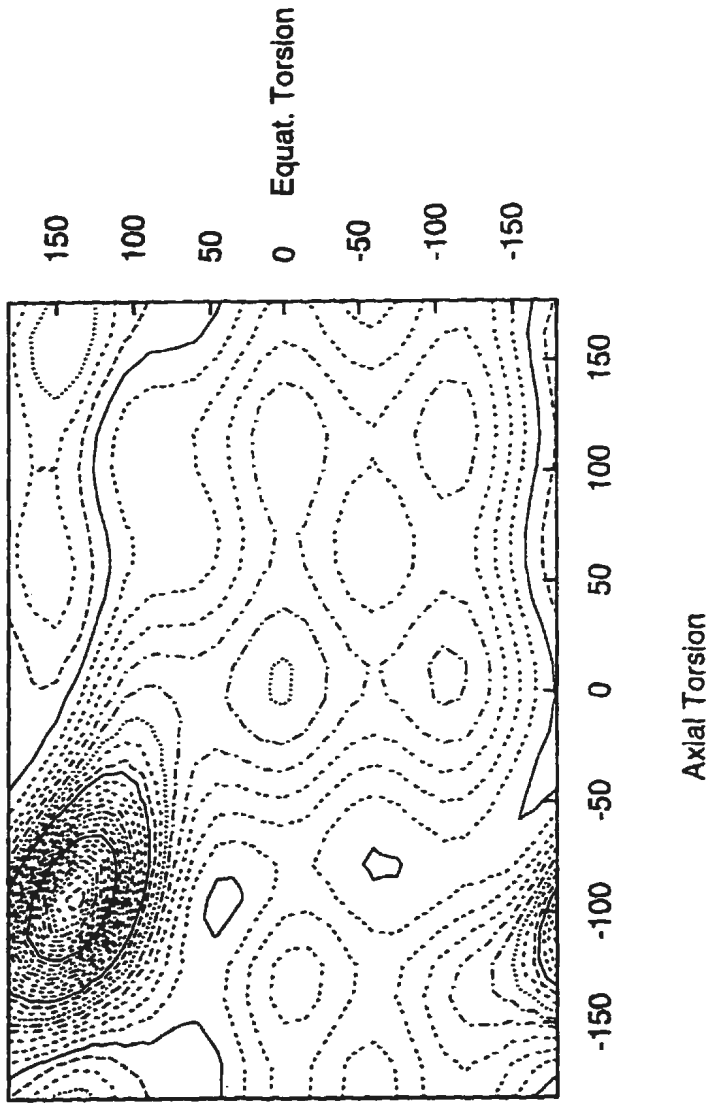


Table 73: Predicted and actual STO-3G minima (torsions in degrees)

	Predicted		Actual		Energy
	Axial	Equat.	Axial	Equat.	(kJ/mol)
(stag,stag)	160	150	162.38	144.51	0.00
(stag, gau out)	-170	-50	-175.52	-59.36	14.42
(stag, gau in)	-170	70	-178.41	70.95	7.43
(gau out, stag)	50	160	58.79	158.39	2.26
(gau out, gau out)	65	-60	62.77	-62.01	20.19
(gau out, gau in)	70	70	66.89	73.22	15.84
(gau in, stag)			-59.06	188.59	7.85
(gau in, gau out)	-80	-60	-78.44	-68.21	6.07
(gau in, gau in)	-95	50	-94.66	41.33	7.77

are, with one exception, all lower in energy than those which cannot. The three highest energy structures can all exhibit lone pair–lone pair repulsion. The optimized structures are close to those predicted from the scan. In addition, the mininum with a strong steric interaction actually distorts to relieve the strain, giving a reasonably stable minimum at STO-3G.

It would be very difficult to obtain all of the various transition states connecting the minima through both rotation about C-O bonds and through the inversion process. However, we can estimate, from the (gau out, gau out) structure and its (necessarily C_s) inversion transition state, the inversion barrier to be 9.20 kJ/mol at STO-3G. This is more than double the value for cyclohexadiene itself. In the C_s structure, the dipole-dipole interaction of the 2 C-O bonds would be maximized, which may account for the difference.

At the 3-21G level only four minima could be located, as suggested in Table 74. Three of the disappearing minima are those with the axial gauche out conformation,

Table 74: Conformers of cis-3,5-cyclohexadiene-1,2-diol (torsions in degrees) and relative energies (kJ/mol)

	3-21G			6-31G*		
	Axial	Equat.	ΔE	Axial	Equat.	ΔE
(stag,stag)	158.0221	151.9548	0.00	161.4653	154.4530	0.00
(stag, gau out)	-170.3523	-52.9102	33.45	-178.5063	-51.4812	21.98
(stag, gau in)						
(gau out, stag)				73.0985	162.3478	7.66
(gau out, gau out)				74.7724	-55.2519	33.72
(gau out, gau in)						
(gau in, stag)				-37.7082	184.5494	10.34
(gau in, gau out)	-80.3923	-65.7674	18.79	-70.6971	-60.7750	12.99
(gau in, gau in)	-93.6682	38.2564	19.97	-82.1969	37.9770	13.50

which was suggested by the scans and optimization of the simpler alcohol. In addition, the (stag, gau in) and (gau in, stag) minima disappear, and this may be related to the fact that these conformers cannot hydrogen-bond and also have strong steric or 1,3-diaxial interactions. Since 3-21G overestimates the strength of the hydrogen-bond[43] it may alter the surface so as to destroy intermediate minima. At the 6-31G* level only seven minima could be located, the two missing minima having an equatorial gauche in conformation and thus a 1,3-diaxial interaction, and also possibly a lone pair–lone pair repulsion. A preference for the (stag, stag) at all theoretical levels considered is clearly seen.

In Table 75 the 6-31G* C-O and O-H bond lengths are given. If the conformer cannot hydrogen-bond, then the bond lengths would resemble those of the various conformations of the mono alcohol (i.e., C-O_a = 1.411-4, C-O_e = 1.405-6, H-O = 0.947-8), with the C-O bonds tending to be a little shorter. C-O bonds corresponding to H-bond acceptors are longer than average, whereas those corresponding donors are

Table 75: 6-31G* O-H, C-O and H...O bond lengths (Å), and O-H...O angle (degrees) of diol

	C-O		H-O		H-bond?	H...O	O-H...O
	Axial	Equat.	Axial	Equat.	Axial is		
(stag,stag)	1.4165	1.3969	0.9488	0.9501	Acceptor	2.1799	114.40
(stag, gau out)	1.4011	1.3993	0.9484	0.9470	No	3.0027	
(stag, gau in)					Donor		
(gau out, stag)	1.4159	1.3971	0.9468	0.9497	Acceptor	2.1601	110.73
(gau out, gau out)	1.4051	1.3983	0.9470	0.9470	No	3.6196	
(gau out, gau in)					Acceptor		
(gau in, stag)	1.4124	1.4041	0.9473	0.9488	No	2.4290	
(gau in, gau out)	1.4011	1.4080	0.9494	0.9469	Donor	2.2029	109.09
(gau in, gau in)	1.4013	1.4117	0.9495	0.9473	Donor	2.1940	113.99

shorter than average. The H-O bond of a donor is longer than average.

3.6.4 Cyclic Derivatives of the Diol

Derivatization of the oxygen functionalities of cis-3,5-cyclohexadiene-1,2-diol (in our numbering, the 5,6-dihydroxy derivative of 1,3-cyclohexadiene) may be expected to alter both the rate of Diels-Alder cycloaddition and its selectivity. Presumably, the ideal cycloaddition of a dienophile to a 1,3-cyclohexadiene moiety would occur in which the double bonds form a plane, as this would ensure maximum overlap between the reacting orbitals of the diene and dienophile. However, 1,3-cyclohexadiene and the diol derivative are inherently nonplanar. Cyclic derivatization of the diol may force the diene to adopt a planar or nearly planar configuration, which should increase the rate. Benzene oxide, an extreme example of derivatization, adopts a C_s structure as mentioned previously, and illustrates the point quite clearly.

Table 76: Ring twisting frequency (cm^{-1}) of C_s diol derivatives

Substituent	Frequency	
	3-21G	6-31G*
-CH_2- endo	10	41
-CH_2- exo	61i	32i
-CMe_2- endo	30	31
-CMe_2- exo	57i	26i
-SiH_2-	60i	66i
-$SiMe_2$-	50i	56i
-CO-	60i	50i
-BH-	39	35

Frequencies of Twisting Mode

The cyclic substituents that we consider, formed by replacing the two hydrogens attached to the oxygens, are –CH_2–, –CMe_2–, –SiH_2–, –$SiMe_2$–, –CO–, and –BH–. A measure of the floppiness of the C_s structures obtained would be the vibrational frequency corresponding to the ring twist. An imaginary frequency suggests that the structure is unstable with respect to twisting and would be distorted away from C_s in its preferred conformation. These frequencies are given in Table 76. The only two direct comparisons that can be made with the literature are for the substituents –CMe_2– endo and –$SiMe_2$–, whose diene portion are planar and non-planar (3-21G). This correlates with their rate of reaction relative to cyclohexadiene (> 100 and 2.7) respectively.

Geometries

The geometries of the pentacyclic portion of our substituted dienes are given in Tables 77 and 78, under the assumption of C_s symmetry. The choice of basis set

affects the C-O, C=O, B-O, Si-O and B-H bond lengths significantly, especially upon the addition of polarization functions on oxygen. This, in turn, affects the bond angles and torsions of most of the species, except for the two endo structures. For the nonplanar rings (-OCR_2O-) which can exhibit exo and endo forms, the choice of conformation does not affect the bond length.

Based on the 6-31G* structures and the corresponding frequency, an attempt was made to find a correlation between structural parameters and the twisting frequency. A clear correlation between the C-O-X angle and the square of the twisting frequency was identified (see Figure 18). The angle at which the frequency becomes zero is predicted to be about 111°. Larger angles result in imaginary frequencies. A rationalization of this is that the ideal ring angles of 108° (or a tetrahedral 109.5°) is desired, and that twisting would result in these angles being closer to this ideal.

3.6.5 Ionization Energies

In an effort to understand the facial selectivity of various substituted cyclohexadienes, the photoelectron spectra were measured by Klapstein[240]. Below, we compare the orbital energies thus obtained (Table 79) with our predicted values (Table 80).

The energy of the $\pi^{-}_{C=C}$ orbital is predicted to within 0.11 eV of the measured experimental value. For the alcohols, the most stable conformer predicted at HF/6-31G* was used to determine orbital energies. The endo isomer of the isopropylidene derivative was assumed. The orbital energies of the most likely populated conformers

Table 77: Cyclic cis-5,6-disubstituted 1,3-cyclohexadiene structures (-OCR_2O-)

Parameter	STO-3G	3-21G	6-31G*
-OCH_2O- endo			
C_5-C_6	1.5681	1.5655	1.5547
C_5-O	1.4498	1.4581	1.4126
O-C	1.4329	1.4296	1.3880
C_5-C_6-O	104.89	104.11	103.61
C_6-O-C	104.75	107.26	107.17
O-C-O	107.64	105.52	106.17
C_5-C_6-O-C	21.00	20.34	20.64
C_6-O-C-O	-35.92	-33.85	-34.77
-OCH_2O- exo			
C_5-C_6	1.5686	1.5625	1.5540
C_5-O	1.4489	1.4568	1.4112
O-C	1.4326	1.4288	1.3861
C_5-C_6-O	104.90	104.44	103.67
C_6-O-C	104.47	108.65	107.10
O-C-O	107.72	106.47	106.49
C_5-C_6-O-C	-21.29	-16.26	-20.35
C_6-O-C-O	36.50	26.95	34.38
-$OCMe_2O$- endo			
C_5-C_6	1.5642	1.5613	1.5487
C_5-O	1.4458	1.4517	1.4078
OC	1.4403	1.4404	1.4031
C_5-C_6-O	105.25	104.47	104.12
C_6-O-C	106.95	110.03	110.52
O-C-O	107.61	105.10	105.64
C_5-C_6-O-C	16.77	14.81	13.74
C_6-O-C-O	-28.12	-24.02	-22.41
-$OCMe_2O$- exo			
C_5-C_6	1.5661	1.5601	1.5502
C_5-O	1.4461	1.4524	1.4082
OC	1.4398	1.4386	1.4007
C_5-C_6-O	104.99	104.59	103.93
C_6-O-C	106.07	110.42	109.50
O-C-O	106.97	105.58	105.38
C_5-C_6-O-C	-19.44	-12.83	-16.83
C_6-O-C-O	32.70	20.82	27.63

Table 78: Cyclic cis-5,6-disubstituted 1,3-cyclohexadiene structures (-OXO-)

Parameter	STO-3G	3-21G	6-31G*
	-OCOO-		
C_5-C_6	1.5655	1.5625	1.5466
C_5-O	1.4457	1.4596	1.4193
OC	1.3971	1.3614	1.3289
C=O	1.2099	1.1841	1.1739
C_5-C_6-O	105.13	103.03	102.93
C_6-O-C	108.72	112.29	111.89
O-C-O	112.28	109.27	110.36
C_5-C_6-O-C	0.34	1.95	0.49
C_6-O-C-O	-0.59	-3.28	-0.84
	-OBHO- exo		
C_5-C_6	1.5708	1.5677	1.5567
C_5-O	1.4472	1.4687	1.4212
OB	1.3589	1.3842	1.3536
BH	1.1577	1.1720	1.1820
C_5-C_6-O	104.43	104.21	104.38
C_6-O-B	108.08	110.01	108.92
O-B-O	114.98	111.53	113.39
C_5-C_6-O-B	0.03	1.08	0.36
C_6-O-B-O	-0.05	-1.86	-0.64
	-$OSiH_2O$-		
C_5-C_6	1.5875	1.5831	1.5708
C_5-O	1.4377	1.4557	1.4188
O-Si	1.6801	1.6867	1.6451
C_5-C_6-O	110.16	108.09	108.41
C_6-O-Si	109.44	114.06	112.97
O-Si-O	100.23	94.99	97.14
C_5-C_6-O-Si	4.85	5.59	2.09
C_6-O-Si-O	-7.10	-7.87	-3.00
	-$OSiMe_2O$-		
C_5-C_6	1.5873	1.5825	1.5707
C_5-O	1.4353	1.4531	1.4148
O-Si	1.6786	1.6908	1.6545
C_5-C_6-O	109.98	107.90	108.38
C_6-O-Si	110.00	114.56	113.44
O-Si-O	99.80	94.11	96.19
C_5-C_6-O-Si	3.27	6.61	2.67
C_6-O-Si-O	-4.77	-9.26	-3.80

Figure 18: C_s cis-5,6-Disubstituted-1,3-cyclohexadiene ν^2_{twist} vs. C-O-X

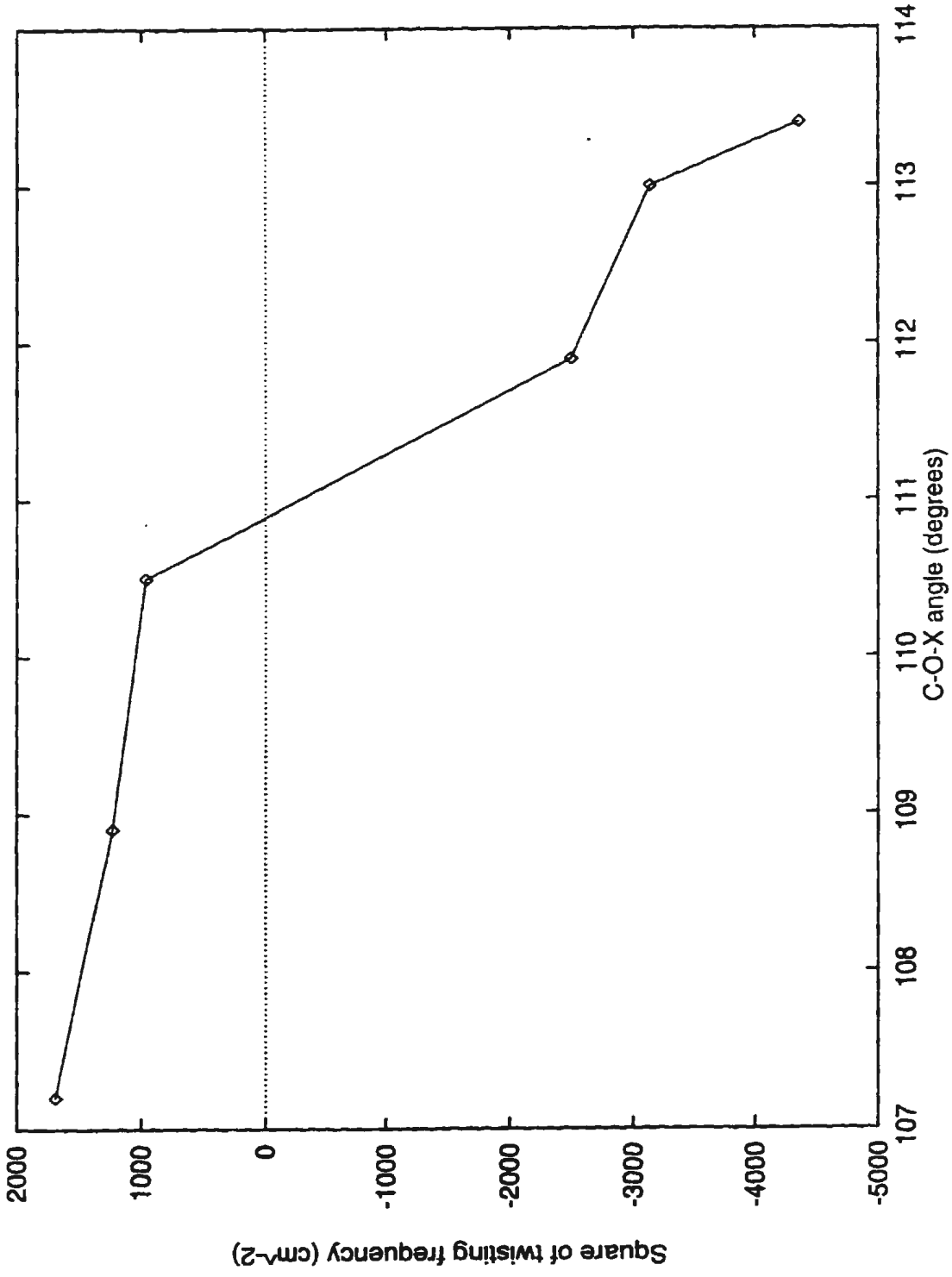


Table 79: Experimental ionization energies (eV)

	Chd[241]	Chd ol[216]	Chd diol[240]	Chd diOAc[240]	Chd $OCMe_2O$[240]
$\pi^{-}_{C=C}$	8.25	8.67	8.88	8.85	8.65
n_O		10.25	10.21	10.50	9.5
n_O			10.97	11.39	10.1
$\pi^{+}_{C=C}$	10.7	11.15	11.6?	11.8?	11.2?

Table 80: Theoretical (HF/6-31G*) ionization energies (eV)

	Chd[241]	Chd ol[216]	Chd diol[240]	Chd OCMe2O[240]
$\pi^{-}_{C=C}$	8.14	8.60	8.85	8.73
n_O		11.36	11.45	10.94
n_O			11.89	11.59
$\pi^{+}_{C=C}$	11.37	11.90	12.64	11.89

differ by at most about 0.3 eV, except for the $\pi^{+}_{C=C}$ which exhibits a larger range. There is no obvious trend between the rate or selectivity of the Diels-Alder reaction and the HOMO energy of the diene, which agrees with the analogous cyclopentadiene results[140]. The only noticeable trend is the lowering of the $\pi_{C=C}$ orbital energies upon hydroxy substitution. For inner valence shell or core electrons, the Koopman approximation would break down, necessitating the use of techniques such as many-body Green's Function[242]. In our case, the orbitals (with the HOMO an exception) are slightly lower in energy than experiment (i.e., these electrons are predicted to be more tightly bound).

3.6.6 Diels-Alder Reaction

The structures of the transition states for the Diels-Alder reaction of cyclohexadiene with ethene and ethyne are given in Tables 81 and 82. One remarkable feature

of both transition states is their incredible similarity. The bond lengths and angles of the two cyclohexadiene fragments are nearly identical with each other, with maximum differences being 0.003 Å and 1°. The geometry of the transition state would be consistent with a strongly exothermic reaction, according to the well-known Hammond's postulate. The C_5-C_6 bond length is very similar to that in the planar form of 1,3-cyclohexadiene. The forming C-C bonds are shorter in the ethene TS, which is consistent with the generally shorter σ bonds to unsaturated carbons. In spite of the shorter forming C-C bonds, the torsion C_2-C_1-C_4-C_5 is larger. The angle of approach of the dienophile, C_2-C_1-C_4-C_8 is also larger in the ethyne case. This is consistent with a larger steric interaction between the incoming dienophile and the sp^3-hydrogens of the diene in the case of ethene. However, the activation barrier for ethyne is consistently 5-6 kJ/mol larger than that of ethene, suggesting that the electronic difference between the two dienophiles primarily determines the barrier.

Table 81: Structure of 1,3-cyclohexadiene + ethene transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4181	1.3972	1.3932	1.4124
C_1-C_2	1.3604	1.3698	1.3793	1.3814
C_1-C_6	1.5310	1.5125	1.5155	1.5107
C_5-C_6	1.5536	1.5604	1.5485	1.5517
C_1-C_7	2.2312	2.2222	2.2119	2.3112
C_7-C_8	1.3567	1.3705	1.3809	1.3781
C_1-C_2-C_3	118.44	118.69	118.19	118.65
C_3-C_4-C_5	120.20	119.67	119.31	119.65
C_4-C_5-C_6	112.27	112.24	112.26	113.10
C_7-C_1-C_2	98.52	98.95	99.54	98.79
C_7-C_1-C_6	94.24	94.40	95.48	93.63
C_1-C_7-C_8	107.71	107.57	107.30	107.10
C_2-C_1-C_4-C_5	147.13	146.44	145.03	147.78
C_2-C_1-C_4-C_8	110.47	111.06	111.34	110.50

Table 82: Structure of 1,3-cyclohexadiene + ethyne transition state

Parameter	STO-3G	3-21G	6-31G*	MP2/6-31G*
C_2-C_3	1.4158	1.3976	1.3954	1.4147
C_1-C_2	1.3617	1.3679	1.3753	1.3775
C_1-C_6	1.5323	1.5236	1.5165	1.5119
C_5-C_6	1.5530	1.5585	1.5467	1.5498
C_1-C_7	2.1862	2.2027	2.2004	2.2824
C_7-C_8	1.1995	1.2216	1.2239	1.2466
C_1-C_2-C_3	118.14	118.57	118.09	118.63
C_3-C_4-C_5	120.72	120.25	120.10	120.55
C_4-C_5-C_6	111.98	112.12	112.16	113.07
C_7-C_1-C_2	98.69	98.68	98.93	98.52
C_7-C_1-C_6	91.84	91.93	92.98	90.41
C_1-C_7-C_8	110.07	109.69	109.47	109.03
C_2-C_1-C_4-C_5	147.19	147.21	146.22	149.55
C_2-C_1-C_4-C_8	112.19	112.18	112.04	111.52

Table 83: Activation barriers of 1,3-cyclohexadiene + dienophile (kJ/mol)

	C_2H_4	C_2H_2
STO-3G	162.82	180.61
3-21G	153.38	165.81
6-31G*//3-21G	194.75	200.28
6-31G*	194.79	200.17
MP2/6-31G*//6-31G*	64.21	70.91
MP2/6-31G*	72.89	78.15

Chapter 4

Optimization Observations and Improvements

In the following sections we make some observations on the Z-matrix optimization behavior of several systems. Each section is then followed by the selection of suitable test cases and an analysis of the optimization behavior when automatically generated natural internal coordinates are used.

4.1 5-Substituted Cyclopentadienes

4.1.1 Z-Matrix Optimizations

The relatively minor changes in the geometry of the carbon framework upon changing either the basis set or the substituent allow the geometrical parameters to be very transferable, resulting in reasonably fast convergence for most of the molecules with

the OC method. Occasionally the DIIS method needs to be used to improve the accuracy of the geometry when the OC method fails to lower the gradient norm enough. For some systems, the VA method with full Hessian evaluation was used, and these converged in six steps or less, illustrating the effect that the Hessian quality has on the optimization. The line search algorithm in OC occasionally generates a large number of function evaluations with no immediate improvement in function value or gradient norm. In some cases poor convergence resulted from an unfortunate choice of Z-matrix in which the specification of the 5-position (carbon and its substituents) resulted in (nearly) linear bond angles, which was subsequently corrected. The optimization is thus also dependent on a suitable coordinate choice.

A statistical analysis of the optimization data was performed. A total of 218 optimizations were performed, of which 14 were already converged at the first iteration and thus excluded. Of the remaining 204, 93 were OC, 68 were DIIS, and 43 were performed with VA. Of these, there were 5, 0, and 4 blunders, respectively, for the three methods (in which a Z-matrix was used containing either redundant parameters or nearly linear angles). When we consider the 88, 68 and 39 optimizations left in our sample, 8, 11, and 4 failed to converge sufficiently (9, 16 and 10 % respectively). It is clear that DIIS fails the most on average, the reason presumably being that DIIS does not update its Hessian matrix, and thus it exceeds the maximum allowable iteration count (MAXIT). OC usually fails before reaching MAXIT, because it cannot find a way of lowering the energy. VA usually fails because the Hessian matrix update

changes the eigenvalue spectra drastically, and the optimizer either enters or considers the current point as being in a region of different curvature (usually too many negative eigenvalues). The number of iterations correlates very roughly with both the number of parameters and the initial gradient.

4.1.2 MUNGAUSS 1.0 Improvements

There are four test cases considered in this section, namely:

1. 2,4-cyclopentadien-1-ol anticlinal: 3-21G VA optimization from STO-3G, STO-3G initial Hessian (13 iterations/13 function evaluations)

2. 2,4-cyclopentadien-1-ol anticlinal: 6-31G* VA optimization from 3-21G, 3-21G updated Hessian (40/40M + 7/7)

3. 2,4-cyclopentadien-1-amine anticlinal: 3-21G VA optimization from STO-3G, STO-3G initial Hessian (18/18)

4. 2,4-cyclopentadiene-1-phosphine synclinal: 3-21G OC optimization from STO-3G, STO-3G initial Hessian (13/15)

As will be discussed later, Test case 1 is a transition state linking the synclinal and antiperiplanar structures. Using the automatic coordinates generated by MUNGAUSS 1.0, under the same conditions, the geometry converged in just 8 iterations, a saving in CPU time of almost 40 % . Test 2 converged in just 8 iterations, a savings in CPU time of over 80 % . Test 3 took slightly longer at 20 iterations (-11 %). However,

Test 4, using OC, took much longer, taking 27 function evaluations (-80 %). This is most likely due to the poorer Hessian default guess as currently implemented in the PIC routine. An overall savings in function evaluations over the four test cases of 15 % is noted.

4.2 1,2-Heterotropic Shift Transition States in 5-Substituted Cyclopentadienes

4.2.1 Z-Matrix Optimizations

The 78 optimizations proceeded reasonably quickly, usually converging within about 20 iterations of the next-best theoretical level. The number of failures with the OC method (11 out of 54) for these symmetrical transition states is about twice as prevalent (20 %), however, which may indicate that the Hessian guess used is not as good for these transition states. The majority of failures (8) occur in cases where there is relative orientation of the two fragments (the ring and the migrating group). All of the failures (3) in which this possibility is excluded (i.e., the migrating group is a single atom) occur when OC is given a nearly converged geometry with a gradient less than ten times the cutoff tolerance. In both types of failure, many additional function evaluations did not lower the energy. In all of the cases, however, subsequent application of the DIIS method allowed the structures to converge.

4.2.2 MUNGAUSS 1.0 Improvements

Because of the symmetry of these transition states, the number of parameters used in the Z-matrix optimization space is just over half ($0.555 \pm .01$) of the total number of parameters expected. We thus consider any PIC optimization proceeding within 1.8 times as many function evaluations as the corresponding Z-matrix optimization as being of equal performance. (Implicit here is the assumption that the number of function evaluations scales linearly in the number of parameters.) The gradient should lie entirely in the totally symmetric subspace. Seven test cases were considered.

1. CpH: 6-31G** OC optimization from 6-31G*
2. CpF: 6-31G* optimization from 3-21G
3. CpCl: 3-21G(*) optimization from 3-21G
4. CpOH staggered: 3-21G optimization from 3-21G eclipsed parameters and STO-3G HOD Z-matrix angle.
5. CpSH staggered: STO-3G optimization from STO-3G eclipsed parameters and HOD=80.0°.
6. $CpNH_2$ staggered: 6-31G* optimization from 3-21G
7. $CpPH_2$ staggered: 6-31G* optimization from 3-21G(*).

Table 1 gives the number of iterations/function evaluations for the various optimizations performed. The first series of PIC optimizations worked poorly as the

optimization was performed in the full optimization space, which contains the transition vector. If the gradient becomes non-orthogonal to the transition vector then the symmetry becomes destroyed and the geometry would start converging to the minimum structure. In our case, a large step is taken and the iterative procedure that is used to convert the proper internal coordinates to Cartesians fails. One reason for the gradient deviation may have been that the ring coordinates did not have the correct local pseudosymmetry. (The two criteria it uses to choose the starting atom for a ring coordinate, which is especially important in odd-membered rings, failed and a default was chosen.) The pseudosymmetry was explicitly forced in optimization series 2. A slight improvement in many test cases was noted. Symmetrization to C_s of the coordinates and removing coordinates of representation A" gave another slight improvement. It must be noted that no consistent trend regarding the better choice of Z-matrix or proper internal coordinates can be observed for these systems. Of course, for transition states, there is no exact prescription for the 'best' coordinates and thus considerable improvements in this choice may be made.

4.3 Metal-Water Complexes

4.3.1 Z-Matrix Optimizations

Lithium aquo complexes provide us with stringent tests of any optimization procedure. The optimization of the mono- (C_{2v}) and di-aquo (D_{2d} and D_{2h}) species was straightforward, since reasonable guesses for the force constants of the stretching

Table 1: OC Optimizations of transition states of the 1,2-heterotropic shift of 5-substituted 1,3-cyclopentadienes

Optimization	1	2	3	4	5	6	7
0[a]	8/25E[b]	9/14	15/20	14/26E	9/21E	20/26	9/21E
1[c]	4/23E	5/13B[d]	33/50EC[e]	8/14B	1/2B	0/1B	8/12B
2[f]	6/9	11/36B	8/20	9/12B	1/2B	0/1B	14/17B
3[g]	9/14E	18/21	8/18E	1/18B	1/18B	0/1B	15/32E

[a]Z-Matrix: 15 (3),17 (2) or 18 (2) parameters, respectively.

[b]E refers to an optimization by terminated by the energy criterion before the gradient criterion has been reached.

[c]PIC auto: Coordinates relating the orientation of the components were generated by hand.

[d]B refers to the failure of the BLDXYZ routine to generate a suitable set of Cartesian coordinates for the current parameters.

[e]C refers to convergence toward the minimum energy structure.

[f]PIC: As PIC set 1, but with slightly different ring coordinate definition (see text).

[g]PIC: A symmetrized version of PIC set 2.

and bending motions were available. The optimization of the two tri-aquo species of D_{3h} symmetry was also straightforward, since there were no loose librational modes as optimization parameters. However, the preferred D_3 species has 1 mode corresponding to the twist of the water relative to the C_3 axis. The DIIS method, if started very near the optimal geometry, would sometimes take many optimization steps before converging to the optimal geometry, especially with a tight convergence criterion. The same behavior was noted for the tetra-aquo species of S_4 symmetry, but not for the D_{2d} structure, which had no totally symmetric torsional modes. The C_2 penta-aquo species gave much difficulty during the optimization, especially at the 6-31G* level, since at least 16 attempts were made to optimize the geometry. The T_h hexa-aquo structure posed no problems during the optimization, because the high symmetry rendered no loose torsional modes. The dodeca-aquolithium species posed many problems because of the many second-sphere librational modes possessing very

small force constants.

There were a total of 89 recorded optimizations for $LiAq_n, n = 1, \ldots, 6$, of which 10 were pre-converged. Of the 29 OC and 50 DIIS optimizations recorded, there were 9 (29 %) and 0 failures recorded, but this data is skewed because many DIIS failures were not initially recorded. The data clearly indicate the difficulties when loose librational modes are present.

The most difficult case proved to be octa-aquolithium. Initially two models were considered for octaaquolithium. The dodeca-aquolithium calculation suggested the second sphere waters could be distinguished into a set that was hydrogen-bonded twice to the inner-sphere and to a set that was hydrogen-bonded just once. The STO-3G basis set suggested that these two structures were nearly isoenergetic, whereas the 3-21G basis set clearly favored the structure with more hydrogen-bonds. For the structure with 8 hydrogen-bonds, the STO-3G structure needed only 4 DIIS attempts with MUNGAUSS to optimize, the 3-21G structure, 16, and the 6-31G* structure, 41 attempts! At the end of the 6-31G* optimization with MUNGAUSS, the structure was still not converged, even with the use of proper internal coordinates, so we began using Gaussian 92. With a combination of analytic force constant evaluation and (initially) the eigenvector following method, we were able to obtain a structure that contained four additional hydrogen-bonds. Recalculation of the force constant and subsequent optimization gave us the minimum structure in a total of 21 Gaussian optimization steps and two Hessian calculations. With the 6-31G* geometry in hand we optimized

the 6-31+G* geometry with the default Gaussian method in 27 iterations.

Similarly, octaaquoberyllium was difficult to optimize. A structure similar to octaaquolithium was initially imposed. After 103 iterations and five Hessian evaluations in Gaussian 92, the geometry converged, using both the default and the eigenvector following search. The hydrogen-bond arrangement in the final structure was unlike the original as discussed before, and thus the major difference in geometry accounts for the difficulty.

4.3.2 MUNGAUSS 1.0 Improvements

The only Z-matrix optimization which both failed and which was sufficiently documented was for tetra-aquolithium(I) ion (S_4). This 6-31G* OC optimization used the 3-21G angles and torsions, and the 6-31G* bond lengths of a more symmetric D_{2d} species. The old Z-matrix optimization went 7 iterations/ 20 function evaluations before failing to converge (termination by failure to reduce the energy sufficiently - EPS). The proper internal coordinate optimization converged quickly in 3 iterations/ 4 function evaluations, a savings of 80 % . This example demonstrates the power of using PIC's in geometry optimization.

4.4 Bimolecular Complexes

4.4.1 Z-Matrix Optimizations

For these systems, 383 optimizations were performed, of which 42 were preconverged. Of the remainder, there were 271 with OC, 49 with DIIS, 20 with DOC, and 1 with VA. A total of 35 OC failures were recorded (9 % of total). Many of the 'preconverged' structures were not fully converged, and failed on the first OC optimization. DIIS was successful in converging any incompletely converged OC structures.

For the most part, the optimization of these bimolecular complexes was straightforward. Surprisingly, the worst results were for the linear systems, which comprise 28 of the 35 failures. This may be related to the very small number of optimization parameters (3 or 4). Of the remaining 7, 6 started with very small gradient lengths. Of the 28 linear system failures, 21 were not attributable to any conditions known to give failure. In addition, for the cases with f-functions, in which the analytic derivatives were not available, the optimization takes many function evaluations.

4.4.2 MUNGAUSS 1.0 Improvements

It is clear that for the linear systems, proper internal coordinates are essentially the same as the Z-matrix coordinates, and thus the majority of the failures in these cases cannot be attributed to the coordinate system. We found three suitable test

cases:

1. FH ... H_2O: STO-4G OC optimization from STO-3G (13/14)

2. H_2O ... CO_2: STO-4G OC optimization from STO-3G (5/13E)

3. H_2O ... H_2O: 6-21G* DIIS optimization from 3-21G* (11/11)

With the proper internal coordinate optimization, the number of iterations/function evaluations changed to 4/5, 4/15E and 4/4. The first and last optimization was speeded up by about 65 %, whereas the middle optimization was slightly less efficient. A slightly different set of coordinates was tried for Case 2, where the bend and out-of-plane bend about C were replaced by a linear bend pair, but this gave problems during the iterative transformation to Cartesian coordinates.

4.5 Benzene Oxide–Oxepin Valence Tautomerization

4.5.1 Z-Matrix Optimizations

A total of 152 MUNGAUSS optimizations were performed, of which 15 were pre-converged. Of the remainder, there were 59 carried out with OC, 32 with DIIS and 46 with VA. The geometries for each system are generally independent of theoretical level and thus guesses from a lower theoretical level provide good estimates for those of higher levels. Some exceptions to this are the poor behavior of the 3-21G basis set in describing the C-S bond length of the three-membered ring of benzene

sulfide and in the H-O-C angle in the protonated species. The bicyclic species fared slightly worse during optimization compared with the monocyclics, probably because of a combination of a poorer Hessian guess (2 loose bonds instead of 1) and of the problems with 3-membered rings resulting in poorer geometry guesses. In general, the planar oxepin forms optimized more quickly than the boat form, but the initial inclusion of symmetry-redundant parameters hampered initial attempts. The valence tautomerization transition states were difficult to optimize partly because a redundant parameter corresponding to a free rotation was inadvertently introduced, and thus the Hessian was nearly singular. When this was removed the VA optimization proceeded smoothly. The default algorithm in Gaussian 92 (opt=ts) always crashed because of a poor Hessian update changing the eigenvalue structure of the initial 3-21G Hessian. Recalculation of the 3-21G Hessian and continuing with the partially optimized geometry solved the problem. With this system, one can certainly learn a lot from one's blunders!

4.5.2 MUNGAUSS 1.0 Improvements

Three test cases were examined for this system:

1. Benzene oxide : 6-31G* OC optimization from 3-21G (11/17)

2. Oxepin: 6-31G* OC optimization from 3-21G (11/15)

3. Planar oxepin: 6-31G* OC optimization from 3-21G (17/21)

With proper internal coordinates, the optimizations converged in 15/17, 21/23, and 9/10 iterations/function evaluations respectively. Comparing this to the Z-matrix optimization, we see that there is no change, 53 % worsening, and 52 % improvement. Overall, there is a 6 % improvement in the number of function evaluations. We note that no net trend is observed on a case-by-case basis, and this may be related to a poor choice of PIC initial diagonal Hessian.

4.6 1,3-Cyclohexadienes and derivatives

4.6.1 Z-Matrix Optimization

A total of 97 MUNGAUSS optimizations (3 pre-optimized) and 29 Gaussian optimizations were performed, of which 52 were OC, 27 were DIIS and 15 were VA. OC sometimes failed on the second step when the predicted energetic lower bound was too low, resulting in a large step and subsequent SCF failure. 1,3-Cyclohexadiene itself (C_2 and C_{2v}) proved to be relatively easy to optimize, taking around 20 function evaluations. In other cases OC failed to optimize completely, but DIIS cleaned this up. This may be related to the lack of a ring closure coordinate and thus a poorer Hessian guess. The 6 conformers of 2,4-cyclohexadien-1-ol took about the same number of iterations at STO-3G, if started with a suitably modified diol geometry. The 3-21G optimization of the axial gauche out conformer gave a lot of difficulty, initially converging to the staggered conformer. This conformer was never found, and a rigid rotor scan casts doubt on the existence of this conformer as a local minimum at this

level of theory. Even with Gaussian 92, this conformer (at 6-31G*) took twice as many iterations (8) as the other conformers (4), because of its unusual torsion angle.

The cis-diol gave conformational behavior which was highly basis-set dependent. The STO-3G surface scan, using the staggered geometry, suggested 8 minima, but we were able to find a ninth conformer which was near the maximum energy structure in the torsional subspace generated by the rigid staggered geometry. Thus at this level, the intuitive picture of $3 \times 3 = 9$ minima is realized. Starting with these minima and their corresponding Hessians, optimizations were carried out at the 3-21G and 6-31G* levels. The number of minima obtained were 4 and 7 respectively, with the 'missing' minimum collapsing to another minimum as depicted in Figures 1 and 2. The arrows here depict the iterates moving from the starting configuration to the final configuration.

4.6.2 MUNGAUSS 1.0 Improvements

A total of 5 test cases were examined:

1. 1,3-cyclohexadiene: 6-31G* OC optimization from 3-21G (9/16)

2. 2,4-cyclohexadiene-1-ol axial staggered: 3-21G OC optimization from STO-3G (20/25)

3. cis-3,5-cyclohexadiene-1,2-diol axial gauche in, equatorial gauche out: 3-21G OC optimization from STO-3G (18/37E)

Figure 1: cis-3,5-Cyclohexadiene-1,2-diol optimization (3-21G)

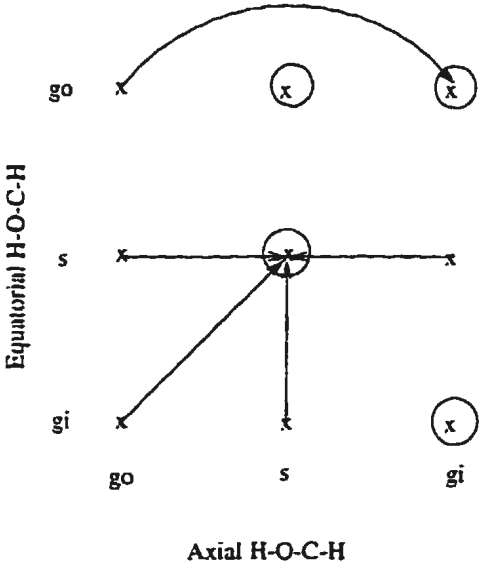
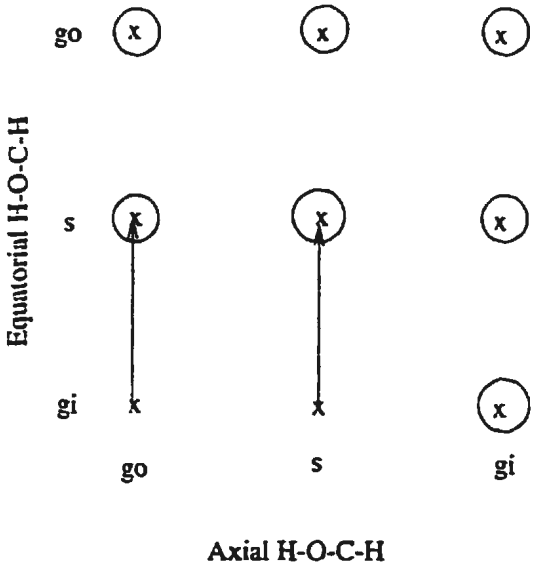


Figure 2: cis-3,5-Cyclohexadiene-1,2-diol optimization (6-31G*)



4. 1,3-cyclohexadiene carbonate: 3-21G OC optimization from STO-3G, lower bound set to -490.4 Hartrees (25/37E)

5. 1,3-cyclohexadiene + ethyne Diels-Alder transition state: 3-21G VA optimization from STO-3G, with STO-3G Hessian (40/40M)

With proper internal coordinates, these converged in 10/19, 15/20, 16/17B, 18/32E, and 7/7 iterations/function evaluations, respectively. Case 3 gave problems while transforming to Cartesians, but the last point was within two times the gradient criterion. These correspond to savings of -19 %, 20 %, 54 %, 14 %, and 83 %. Overall, the proper internal coordinates gave an improvement of 39 % .

4.7 Summary

It is clear that on average, the use of proper internal coordinates speeds up the optimization, especially where large-amplitude motions are involved. Problems sometimes arise during the iterative transformation to Cartesian coordinates, especially where transition states are involved.

Chapter 5

Concluding Remarks

In this study, several significant computational developments in optimization as applied to chemistry were made.

1. The choice of initial Hessian is now much more flexible, without sacrificing ease of input. One can use a unit Hessian, a diagonal Hessian, or any combination of forward, backward or central difference estimations for optimization parameters.

2. The DIIS algorithm has been expanded to include a BFGS Hessian update, a cubic geometry correction and several choices for a throwaway strategy.

3. Three layers of connectivity are defined corresponding to normal, weak, and very weak bonds which can be easily tailored to meet the needs of particular systems.

4. Efficient algorithms from graph theory have been used to analyse the molecular connectivity to determine the number of components, the number of fundamental rings, and the ring assemblies. This is accomplished through the use of pruning and homeomorphic reduction.

5. From the topology, the natural internal coordinates of Pulay can automatically be constructed for all acyclic structures containing atoms of valence 4 or less.

6. The natural coordinates of cyclic substructures containing exactly one ring can be constructed automatically. The relative motion of spiro-fused rings can also be handled.

7. Suggestions are made for dealing with fused ring systems of the type [n.m.0] and [m.n.k] in a non-redundant fashion.

8. Single weak bonds between components and very weak bonds can be handled automatically, and suggestions for coordinates are made for the case where there are 2 or more weak bonds between two components.

During the Z-matrix optimization of several systems of interest, several points are noted:

1. When the OC method fails to converge sufficiently, the DIIS method can usually remedy the problem.

2. The DIIS method can fail because of the lack of a Hessian update.

3. The VA method can fail when the approximate Hessian changes its eigenvalue spectrum.

4. Z-matrix optimizations of systems with loose modes tend to converge more slowly than those systems without.

5. Redundant parameters corresponding to translation or rotation can slow optimization considerably.

The use of natural internal coordinates on the above troublesome cases demonstrates that on average, the use of natural internal coordinates speeds up the optimization, in some cases very drastically.

The structures of 5-substituted cyclopentadienes agree well with most of the available experimental results. A pronounced basis set dependence of the structures is noted in some cases. The conformational trends in geometry and energy can be explained on the basis of hyperconjugation, especially where lone pairs are involved.

The structures of the transition states of the 1,2-shift of the substituent in 5-substituted cyclopentadienes are, with the exception of the migrating group, very similar. A dimensionless stretching parameter is defined and is shown to correlate well with the activation barrier to this migration. The Hartree-Fock method seems to overestimate the barrier by 20 - 50 kJ/mol compared with available experimental estimates.

The metal-water complexes of Li^{+}, Be^{2+}, Mg^{2+}, Al^{3+}, Sc^{3+}, Zn^{2+}, Ga^{3+}, Cd^{2+}, and In^{3+} all give a symmetric stretch frequency much lower than that measured in

aqueous solution. Including a second hydration sphere in the calculation brings the frequency to much closer agreement with experiment as demonstrated for Li^+, Be^{2+}, and Mg^{2+}.

Complexation to the hydrogen of HF results in a lengthening of the HF bond in accordance with the interaction energy. The hydrogen-bond lengths correlate roughly with the interaction energy, with shorter bonds indicating a stronger interaction, especially when comparing with the same acceptor atom.

The enthalpy of the valence tautomerization between benzene oxide and oxepin is close to zero and thus this system is a sensitive probe of the accuracy of a particular theoretical level. Methyl substitution at the 2 and 7 positions destabilizes the oxide form relative to oxepin. Both replacement of the oxygen by sulfur, and protonation have the opposite effect. The transition states are more curved than either reactant or product. Self-consistent reaction field theory can give the sign of the solvent correction. A literature photoelectron spectrum is reassigned. The source of the facial selectivity in benzene oxide is due to steric interaction between the dienophile and the oxygen.

The geometry of 1,3-cyclohexadiene, optimized here for the first time with polarization functions and/or correlation, agrees very well with the available experimental information, predicting a twisted skeleton. The barrier to ring inversion is in excellent agreement to the experimental value as determined by overtone measurements of the ring-twisting mode. The potential surface in the H-O-C-H torsional subspace of the

5-hydroxylated and cis-5,6-dihydroxylated species is very basis set dependent. Diol species capable of forming internal hydrogen-bonds are generally the most stable. Cyclization of the diol can give either a stable or unstable C_s species, and a correlation between the square of the 6-ring inversion frequency of the C_s structure and the value of the C-O-X angle is apparent. Where this angle is smaller than about 111°, the C_s structure is stable. The available photoelectron spectra compares favorably with our calculations but shows no trend with the facial selectivity of the Diels-Alder reaction.

Possible further work is outlined:

1. the extension of automatic natural internal coordinates to atoms of valence 5 or more;

2. the extension of automatic natural internal coordinates to multiple close contacts;

3. the extension of automatic natural internal coordinates to ring systems of the type [n.m.0] or [n.m.k.0];

4. the derivation and extension of automatic natural internal coordinates to complicated ring systems such as [m.n.k] systems;

5. the incorporation of a better default Hessian for natural internal coordinates;

6. a correlated treatment of the 1,2-shift of 5-substituted 1,3-cyclopentadienes.

Bibliography

[1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1st ed. (revised), McGraw-Hill, New York (1989), and references therein.

[2] See, for example, L. I. Schiff, *Quantum Mechanics*, McGraw-Hill, New York (1968).

[3] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[4] J. J. Moré and D. C. Sorensen, *Math. Prog.*, **16**, 1 (1979).

[5] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press, London (1981).

[6] M. J. D. Powell, in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (eds.), Academic Press, New York, pp 31-65 (1970).

[7] W. C. Davidon, Argonne National Labs Report ANL-5990 (1959).

[8] R. Fletcher and M. J. D. Powell, *Comp. J.*, **6**, 163 (1963).

[9] C. G. Broyden, *J. Inst. Maths. Appls.*, **6**, 76 (1970).

[10] C. G. Broyden, *J. Inst. Maths. Appls.*, **6**, 222 (1970).

[11] R. Fletcher, *Comp. J.*, **13**, 317 (1970).

[12] D. Goldfarb, *Math. Comp.*, **24**, 23 (1970).

[13] D. F. Shanno, *Math. Comp.*, **24**, 647 (1970).

[14] D. F. Shanno and K.-H. Phua, *Math. Prog.*, **14**, 149 (1978).

[15] M. Contreras and R. A. Tapia, *J. Optim. Theor. Appl.*, **78**, 93 (1993).

[16] W. C. Davidon, *Math. Prog.*, **9**, 1 (1975).

[17] M. J. D. Powell, *Math. Prog.*, **12**, 141 (1977).

[18] L. Lukšan, *J. Optim. Th. Appl.*, **83**, 27 (1994).

[19] Y. F. Hu and C. Storey, *J. Opt. Th. Appl.*, **83**, 421 (1994).

[20] R. B. Schnabel and T.-T. Chow, *Tech. Rep. U. Colorado*, CU-CS-439-89, (1989).

[21] M. J. D. Powell, *Comp. J.*, **7**, 303 (1965).

[22] K. Levenberg, *Quart. Appl. Math.*, **2**, 164 (1944).

[23] D. W. Marquardt, *J. Soc. Indust. Appl. Math.*, **11**, 431 (1963).

[24] J. E. Dennis, Jr., D. M. Gay and R. E. Welsch, *ACM Trans. Math. Soft.*, **7**, 348, 369 (1981).

[25] M. J. D. Powell, in *Numerical Methods for Nonlinear Algebraic Equations*, P. Rabinowitz (ed.), Gordon & Breach, London p 87, 115 (1970).

[26] C. G. Broyden, *Math. Comp.*, **19**, 577 (1965).

[27] M. J. D. Powell, Subroutine VA05AD, AERE Subroutine Library, Harwell, Didcot, Berkshire, UK.

[28] R. A. Poirier, Y. Wang and C. C. Pye, MUNGAUSS 1.0 (OSIPE Version), Chemistry Department, Memorial University of Newfoundland, St. John's, Newfoundland, Canada (1996).

[29] R. A. Poirier, M. R. Peterson and A. Yadav, MUNGAUSS, Chemistry Department, Memorial University of Newfoundland, St. John's, Newfoundland, Canada (1989).

[30] M. R. Peterson and R. A. Poirier, MONSTERGAUSS, Department of Chemistry, University of Toronto, Toronto, Ontario, Canada and Department of Chemistry, Memorial University of Newfoundland, St. John's, Newfoundland, Canada (several versions)

[31] D. Poppinger, *Chem. Phys. Letters*, **34**, 332 (1975).

[32] R. Fletcher, *Comp. J.*, **13**, 317 (1972).

[33] R. Fletcher, Atomic Energy Research Establishment report AERE-R7125 (1972).

[34] B. A. Murtagh and R. W. H. Sargent, *Comp. J.*, **13**, 185 (1970).

[35] P. W. Payne, *J. Chem. Phys.*, **65**, 1920 (1976).

[36] A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield, and M. Conrad, *Chem. Phys. Letters*, **45**, 595 (1977).

[37] J. D. Head and M. C. Zerner, *Chem. Phys. Letters*, **122**, 264 (1985).

[38] J. D. Head and M. C. Zerner, *Adv. Quant. Chem.*, **20**, 239 (1989).

[39] P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.*, **101**, 2550 (1979).

[40] P. Császár and P. Pulay, *J. Mol. Struct.*, **114**, 31 (1984).

[41] P. L. Cummins and J. E. Gready, *J. Comput. Chem.*, **10**, 939 (1989).

[42] T. H. Fischer and J. Almlöf, *J. Phys. Chem.*, **96**, 9768 (1992).

[43] C. C. Pye, R. A. Poirier, D. Yu, and P. Surján, *J. Mol. Struct (Theochem)*, **307**, 239 (1994).

[44] W. Rudolph, M. H. Brooker and C. C. Pye, *J. Phys. Chem.*, **99**, 3793 (1995).

[45] C. C. Pye, W. Rudolph and R. A. Poirier, *J. Phys. Chem.*, **100**, 601 (1996).

[46] H. B. Schlegel, *J. Comput. Chem.*, **3**, 214 (1982).

[47] J. S. Binkley, R. A. Whiteside, R. Krishnan, R. Seeger, D. J. DeFrees, H. B. Schlegel, S. Topiol, L. R. Kahn, and J. A. Pople, GAUSSIAN 80, Carnegie Mellon University, Pittsburgh, PA 1980.

[48] H. B. Schlegel, *Theor. Chim. Acta (Berl.)*, **66**, 333 (1984).

[49] R. Lindh, A. Bernhardsson, G. Karlström, and P-Å Malmqvist, *Chem. Phys. Letters*, **241**, 423 (1995).

[50] A. Banerjee, N. Adams, J. Simons and R. Shepard, *J. Phys. Chem.*, **89**, 52 (1985).

[51] J. F. Stanton and D. E. Bernholdt, *J. Comput. Chem.*, **11**, 58 (1990).

[52] S. Vogel, T. H. Fischer, J. Hutter and H. P. Lüthi, *Int. J. Quant. Chem.*, **45**, 679 (1993).

[53] P. Pulay, G. Fogarasi, X. Zhou, and P. W. Taylor, *Vib. Spect.*,**1**, 159 (1990).

[54] G. Fogarazi, X. Zhou, P. W. Taylor and P. Pulay, *J. Am. Chem. Soc.*, **114**, 8191 (1992).

[55] J. Baker, *J. Comput. Chem.*, **14**, 1085 (1993).

[56] P. Pulay and G. Fogarasi, *J. Chem. Phys.*, **96**, 2856 (1992).

[57] J. Baker and W. J. Hehre, *J. Comput. Chem.*, **12**, 606 (1991).

[58] J. Baker, *J. Comput. Chem.*, **13**, 240 (1992).

[59] J. Baker and D. Bergeron, *J. Comput. Chem.*, **14**, 1339 (1993).

[60] H. Nakano, T. Nakajima and S. Obara, *Chem. Phys. Letters*, **177**, 458 (1991).

[61] H. B. Schlegel, *Int. J. Quant. Chem., Quant. Chem. Symp.*, **26**, 243 (1992).

[62] C. F. Bender, P. K. Pearson, S. V. O'Neil and H. F. Schaefer III, *J. Chem. Phys.*, **56**, 4626 (1972).

[63] S. V. O'Neil, P. K. Pearson, H. F. Schaefer III and C. F. Bender, *J. Chem. Phys.*, **58**, 1126 (1973).

[64] F. Bernardi and M. A. Robb, in *Adv. Chem. Phys.*, **27**, Ab Initio Methods in Quantum Chemistry 1, K. P. Lawley, (ed.), 155 (1987).

[65] H. B. Schlegel, in *Adv. Chem. Phys.*, **27**, Ab Initio Methods in Quantum Chemistry 1, K. P. Lawley, (ed.), 249 (1987).

[66] H. B. Schlegel, in *Modern Electronic Structure Theory. Part 1*, Advanced Series in Physical Chemistry, Vol. 2, D. R. Yarkony (ed.), World Scientific, Singapore (1995).

[67] J. W. McIver, Jr., *Acc. Chem. Res.*, **7**, 72 (1974).

[68] R. E. Stanton and J. W. McIver, Jr., *J. Am. Chem. Soc.*, **97**, 3632 (1975).

[69] P. Pechukas, *J. Chem. Phys.*, **64**, 1516 (1976).

[70] J. N. Murrell and K. J. Laidler, *Trans. Faraday Soc.*, **64**, 371 (1968).

[71] J. N. Murrell and G. L. Pratt, *Trans. Faraday Soc.*, **66**, 1680 (1970).

[72] J. Brocas and R. Willem, *J. Am. Chem. Soc.*, **105**, 2217 (1983).

[73] J. W. McIver, Jr. and A. Komornicki, *J. Am. Chem. Soc.*, **94**, 2625 (1972).

[74] A. Komornicki and J. W. McIver, Jr., *J. Am. Chem. Soc.*, **95**, 4512 (1973).

[75] A. Komornicki and J. W. McIver, Jr., *J. Am. Chem. Soc.*, **96**, (1974).

[76] D. Poppinger, *Chem. Phys. Letters*, **35**, 550 (1975).

[77] W. Kliesch, K. Schenk, D. Heidrich and H. Dachsel, *J. Comput. Chem.*, **9**, 810 (1988).

[78] S. Bell, J. S. Crighton and R. Fletcher, *Chem. Phys. Letters*, **82**, 122 (1981).

[79] M. J. Rothman and L. L. Lohr, Jr., *Chem. Phys. Letters*, **70**, 405 (1980).

[80] J. E. Sinclair and R. Fletcher, *J. Phys. C: Solid State Phys.*, **7**, 864 (1974).

[81] S. Fischer and M. Karplus, *Chem. Phys. Letters*, **194**, 252 (1992).

[82] S. Bell and J. S. Crighton, *J. Chem. Phys.*, **80**, 2464 (1984).

[83] P. Scharfenberg, *J. Comput. Chem.*, **3**, 277 (1982).

[84] O. Tapia and J. Andrés, *Chem. Phys. Letters*, **109**, 471 (1984).

[85] P. Culot, G. Dive, V. H. Nguyen, and J. M Ghuysen, *Theor. Chim. Acta*, **82**, 189 (1992).

[86] J. M. Bofill, *J. Comput. Chem.*, **15**, 1 (1994).

[87] P. G. Mezey, in *Prog. Theor. Org. Chem. 2: Applications of MO Theory in Organic Chemistry*, I. G. Csizmadia (ed.), Elsevier, Amsterdam p 127 (1977); M. R. Peterson and I. G. Csizmadia, *Ibid.*, p 117; P. G. Mezey, M. R. Peterson and I. G. Csizmadia, *Can. J. Chem.*, **55**, 2941 (1977)

[88] G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.*, **144**, 462 (1971). This method is designed to find minima adjacent to that of the initial minima by passing through the transition states linking them.

[89] J. Pancíř, *Coll. Czech. Chem. Commun.*, **40**, 1112 (1975).

[90] M. V. Basilevsky and A. G. Shamov, *Chem. Phys.*, **60**, 347 (1981).

[91] C. J. Cerjan and W. H. Miller, *J. Chem. Phys.*, **75**, 2800 (1981).

[92] J. Simons, P. Jørgensen, H. Taylor and J. Ozment, *J. Phys. Chem.*, **87**, 2745 (1983).

[93] J. Simons and J. Nichols, *Int. J. Quant. Chem., Quant. Chem. Symp.*, **24**, 263 (1990).

[94] D. T. Nguyen and D. A. Case, *J. Phys. Chem.*, **89**, 4020 (1985).

[95] J. D. Head, B. Weiner, and M. C. Zerner, *Int. J. Quant. Chem.*, **33**, 177 (1988).

[96] T. A. Halgren and W. N. Lipscomb, *Chem. Phys. Letters*, **49**, 225 (1977).

[97] K. Müller and L. D. Brown, *Theor. Chim. Acta (Berl.)*, **53**, 75 (1979).

[98] K. Müller, *Angew. Chem. Int. Ed. Engl.*, **19**, 1 (1980).

[99] A. Jensen, *Theor. Chim. Acta (Berl.)*, **63**, 269 (1983).

[100] M. J. S. Dewar, E. F. Healy and J. J. P. Stewart, *J. Chem. Soc., Faraday Trans. 2*, **80**, 227 (1984).

[101] C. Cárdenas-Lailhacar and M. C. Zerner, *Int. J. Quant. Chem.*, **55**, 429 (1995).

[102] C. Peng and H. B. Schlegel, *Israel J. Chem.*, **33**, 449 (1993).

[103] C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch, *J. Comput. Chem.*, **17**, 49 (1996).

[104] I. V. Ionova and E. A. Carter, *J. Chem. Phys.*, **98**, 6377 (1993).

[105] I. V. Ionova and E. A. Carter, *J. Chem. Phys.*, **103**, 5437 (1995).

[106] J. Baker, *J. Comput. Chem.*, **7**, 385 (1986) .

[107] J. S. Binkley, M. J. Frisch, D. J. DeFrees, K. Raghavachari, R. A. Whiteside, H. B. Schlegel, E. M. Fluder and J. A. Pople, GAUSSIAN 82, Carnegie Mellon University, Pittsburgh, PA 15213 .

[108] P. Pulay, *Theor. Chim. Acta (Berl.)*, **50**, 299 (1979).

[109] R. M. Badger, *J. Chem. Phys.*, **2**, 128 (1934); **3**, 227 (1935).

[110] F. Colonna, L.-H. Jolly, R. A. Poirier, J. G. Ángyán, and G. Jansen, *Comput. Phys. Commum.*, **81**, 293 (1994).

[111] Y. Wang, *A Study of Ab Initio Generalized Valence Bond Theory: Methodology, Programming and Applications*, Ph. D. Thesis, Memorial University of Newfoundland, St. John's, Newfoundland Canada (1995).

[112] Y. Wang and R. A. Poirier, *J. Comput. Chem.*, **7**, 313 (1996).

[113] Y. Wang and R. A. Poirier, *J. Mol. Struct. (Theochem)*, **340**, 1 (1995).

[114] W. J. Hehre, *Practical Strategies for Electronic Structure Calculations*, Wavefunction, Irvine, CA (1995).

[115] Spartan V 4.1, Wavefunction, Inc., Irvine, CA (1996).

[116] N. Deo, *Graph Theory with Applications to Engineering and Computer Science*, Prentice-Hall, Englewood Cliffs, NJ (1974).

[117] L. Matyska, *J. Comput. Chem.*, **9**, 455 (1988).

[118] J. Baker, A. Kessi and B. Delley, *J. Chem. Phys.*, **105**, 192 (1996).

[119] J. Xidos, Memorial University of Newfoundland, St. John's, Newfoundland, Canada, unpublished observations.

[120] S. Saebø, F. R. Cordell and J. E. Boggs, *J. Mol. Struct. (Theochem)*, **104**. 221 (1983).

[121] J. Kao and L. Radom, *J. Am. Chem. Soc.*, **100**, 379 (1978).

[122] L. H. Scharpen and V. W. Laurie, *J. Chem. Phys.*, **43**, 2765 (1965).

[123] R. C. Benson and W. H. Flygare, *J. Am. Chem. Soc.*, **92**, 7523 (1970).

[124] D. Damiani, L. Ferretti, E. Gallinella, *Chem. Phys. Letters*, **37**, 265 (1976).

[125] G. Liebling and R. E. Marsh, *Acta Cryst.*, **19**, 202 (1965).

[126] V. Schomaker and L. Pauling, *J. Am. Chem. Soc.*, **61**, 1769 (1939).

[127] R. Breslow, J. M. Hoffman, Jr., *J. Am. Chem. Soc.*, **94**, 2110 (1972).

[128] M. Saunders, R. Berger, A. Jaffe, J. M. McBride, J. O'Neill, R. Breslow, J. M. Hoffmann, Jr., C. Perchonock, W. Wasserman, R. S. Hutton and V. J. Kuck, *J. Am. Chem. Soc.*, **95**, 3017 (1973).

[129] M. A. McClinton and V. Sik, *J. Chem. Soc., Perkin Trans. 1*, 1891 (1992).

[130] S. Winstein, M. Shatavsky, C. Norton, R. B. Woodward, *J. Am. Chem. Soc.*, **77**, 4183 (1955).

[131] M. Rosenblum, *J. Am. Chem. Soc.*, **79**, 3179 (1957).

[132] F. Tureček and Z. Havlas, *J. Org. Chem.*, **51**, 4066 (1986).

[133] V. A. Mironov, M. E. Dolgaya, V. T. Luk'yanov and S. A. Yankovskii, *J. Org. Chem. USSR*, **12**, 1422 (1976) (Zh. Org. Khim. 1436).

[134] S. B. King and B. Ganem, *J. Am. Chem. Soc.*, **113**, 5089 (1991).

[135] M. Ishida, T. Aoyama and S. Kato, *Chem. Letters*, 663 (1989).

[136] M. Ishida, Y. Beniya, S. Inagaki, S. Kato, *J. Am. Chem. Soc.*, **112**, 8980 (1990).

[137] M. Ishida, T. Aoyama, Y. Beniya, S. Yamabe, S. Kato, S. Inagaki, *Bull. Chem. Soc. Jpn.*, **66**, 3430 (1993).

[138] M. Ishida, S. Kakita and S. Inagaki, *Chem. Letters*, 469, (1995).

[139] M. Ishida, S. Tomohiro, M. Shimizu and S. Inagaki, *Chem. Letters*, 739 (1995).

[140] N. H. Werstiuk, J. Ma, J. B. Macaulay, A. G. Fallis, *Can. J. Chem.*, **70**, 2798 (1992).

[141] J. B. Macaulay and A. G. Fallis, *J. Am. Chem. Soc.*, **110**, 4074 (1988).

[142] J. B. Macaulay and A. G. Fallis, *J. Am. Chem. Soc.*, **112**, 1136 (1990).

[143] A. Yadav, P. R. Surján and R. A. Poirier, *J. Mol. Struct. (Theochem)*, **165** 297 (1988).

[144] J. G. Angyan, I. G. Csizmadia, R. Daudel and R. A. Poirier, *Chem. Phys. Letters*, **131**, 247 (1986).

[145] W. J. Hehre, L. Radom, P. v. R. Schleyer, J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York (1986).

[146] R. A. Poirier and I. G. Csizmadia, in *The Chemistry of Organic Selenium and Tellurium Compounds*, Vol. 1, S. Patai and Z. Rappoport (eds.), Wiley, New York (1986).

[147] V. K. Yadav, A. Yadav and R. A. Poirier, *J. Mol. Struct. (Theochem)*, **186**, 101 (1989).

[148] S. D. Kahn, J. Korppi-Tommola, R. Y. N. Leung, B. M. Pinto, *J. Mol. Struct. (Theochem)*, **303**, 163 (1994).

[149] U. Salzner and P. v. R. Schleyer, *J. Am. Chem. Soc.*, **115**, 10231 (1993).

[150] G. Boche, M. Bernheim, W. Schrott, *Tetrahedron Lett.*, **23**, 5399 (1982).

[151] P. Jutzi and H. Saleske, *Chem. Ber.*, **117**, 222 (1984).

[152] R. A. Bartlett, A. Cowley, P. Jutzi, M. M. Olmstead, and H.-G. Stammler, *Organometallics*, **11**, 2837 (1992).

[153] S. Cradock and D. H. W. Rankin, *J. Chem. Soc., Faraday Trans. 2*, 940 (1972).

[154] J. E. Bentham, E. A. V. Ebsworth, H. Moretto and D. W. H. Rankin, *Angew. Chem. Int. Ed. Engl.*, **11**, 640 (1972).

[155] S. Cradock, E. A. V. Ebsworth, H. Moretto and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 390 (1975).

[156] W. W. Schoeller, *Z. Naturforsch.*, **38b**, 1635 (1983).

[157] W. W. Schoeller, *Z. Naturforsch.*, **39b**, 1767 (1984).

[158] S. M. Csicsery, *J. Org. Chem.*, **25**, 518 (1960).

[159] V. A. Mironov, E. V. Sobolev and A. N. Elizarova, *Tetrahedron*, **19**, 1939 (1963).

[160] S. McLean and P. Haynes, *Tetrahedron Lett.*, 2385 (1964).

[161] S. McLean, C. J. Webster and R. J. D. Rutherford, *Can. J. Chem.*, **47**, 1555 (1969).

[162] J. W. De Haan and H. Kloosterziel, *Rec. Trav. Chim. Pays-Bas*, **87**, 298 (1968).

[163] M. J. S. Dewar, *Chem. Brit.*, **11**, 97 (1975).

[164] A. P. Hagen and P. J. Russo, *Inorg. Nucl. Chem. Letters*, **6**, 507 (1970); *Inorg. Synth.*, **7**, 172 (1977).

[165] J. E. Bentham and D. W. H. Rankin, *J. Organometal. Chem.*, **30**, C54 (1971).

[166] A. P. Hagen and P. J. Russo, *J. Organometal. Chem.*, **51**, 125 (1973).

[167] S. Cradock, R. H. Findlay and M. H. Palmer, *J. Chem. Soc., Dalton Trans.*, 1650 (1974).

[168] A. Bonny and S. R. Stobart, *J. Chem. Soc., Dalton Trans.*, 224 (1980).

[169] A. F. Cuthbertson and C. Glidewell, *J. Organometal. Chem.*, **221**, 19 (1981).

[170] S. R. Stobart, *J. Organometal. Chem.*, **33**, C11 (1971); P. C. Angus and S. R. Stobart, *J. Chem. Soc., Dalton Trans.*, 2374 (1973).

[171] M. J. Barrow, E. A. V. Ebsworth, M. M. Harding and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.*, 603 (1980).

[172] A. V. Kisin, V. A. Korenevsky, N. M. Sergeyev, Yu. A. Ustynyuk, *J. Organometal. Chem.*, **34**, 93 (1972).

[173] T. K. Brunck and F. Weinhold, *J. Am. Chem. Soc.*, **101**, 1700 (1979).

[174] A. E. Reed and F. Weinhold, *Israel J. Chem.*, **31**, 277 (1991).

[175] L. Radom, W. J. Hehre and J. A. Pople, *J. Am. Chem. Soc.*, **94**, 2371 (1972).

[176] W. R. Roth, *Tetrahedron Lett.*, 1009 (1964).

[177] N. M. Sergeyev, G. I. Avramenko, A. V. Kisin, V. A. Korenevsky and Yu. A. Ustynyuk, *J. Organometal. Chem.*, **32**, 55 (1971).

[178] R. B. Larrabee, *J. Organometal. Chem.*, **74**, 313 (1974).

[179] N. G. Rondan and K. N. Houk, *Tetrahedron Lett.*, **25**, 2519 (1984).

[180] S. M. Bachrach, *J. Org. Chem*, **58**, 5414 (1993).

[181] H. Jiao and P. v. R. Schleyer, *J. Chem. Soc., Faraday Trans.*, **90**, 1559 (1994).

[182] M. H. Brooker, in *The Chemical Physics of Solvation*, R. R. Dogonadze, E. Kalman, A. A. Kornyshev, J. Ulstrup (eds.), Elsevier, Amsterdam (1986); Part 3; Chapter 4, p 119.

[183] U. Olsher, R. M. Izatt, J. S. Bradshaw, and N. K. Dalley, *Chem. Rev.*, **91**, 137 (1991).

[184] Y. Kameda, H. Ebatu and O. Uemura, *Bull. Chem. Soc. Jpn.*, **67**, 929 (1994).

[185] F. Rull, *Z. Naturforsch.*, **50 a**, 292 (1995).

[186] Y. Kameda and O. Uemura, *Bull. Chem. Soc. Jpn.*, **66**, 384 (1993).

[187] V. M. Tret'yak, V. I. Baranovskii, O. V. Sizova, and G. V. Kozhevnikova, *Zh. Strukt. Chim.*, **19**, 594 (1978) (Engl. *J. Struct. Chem.*, 519).

[188] J. Sadlej, *Adv. Mol. Relax. Proc.*, **1976**, 8, 117.

[189] C. W. David, *Comp. Chem.*, **2**, 65 (1978).

[190] B. M. Rode, G. J. Reibnegger, S. Fujiwara, *J. Chem. Soc., Faraday Trans. 2*, **76**, 1268 (1980).

[191] E. Kaufmann, J. Gose, P. v. R. Schleyer, *Organometallics*, **1989**, 8, 2577.

[192] V. I. Baranovskii and O. V. Sizova, *Zh. Strukt. Khim.*,**32**, 158 (1991) (Engl.: J. Struct. Chem., p 593).

[193] D. A. Zhogolev, B. Kh. Bunyatan, Yu. A. Kruglyak, *Chem. Phys. Letters*, **18**, 135 (1973).

[194] F. Ramondo, L. Bencivenni, V. Rossi, and R. Caminiti, *J. Mol. Struct. (Theochem)*, **277**, 185 (1992).

[195] D. Feller, E. D. Glendening, R. A. Kendall, and K. A. Peterson, *J. Chem. Phys.*, **100**, 4981 (1994).

[196] E. D. Glendening and D. Feller, *J. Phys. Chem.*, **99**, 3060 (1995).

[197] D. E. Woon and T. H. Dunning, Jr., *J. Am. Chem. Soc.*, **117**, 1090 (1995).

[198] C. W. Bock and J. P. Glusker, *Inorg. Chem.*, **32**, 1242 (1993).

[199] G. D. Markham, J. P. Glusker, C. L. Bock, M. Trachtman and C. W. Bock, *J. Phys. Chem.*, **100**, 3488 (1996).

[200] C. W. Bock, A. Kaufman, and J. P. Glusker, *Inorg. Chem.*, **33**, 419 (1994).

[201] M. M. Probst and K. Hermansson, *J. Chem. Phys.*, **96**, 8995 (1992).

[202] R. Åkesson, L. G. M. Pettersson, M. Sandström, U. Wahlgren, *J. Phys. Chem.*, **96**, 150 (1992).

[203] R. Åkesson, L. G. M. Pettersson, M. Sandström, P. E. M. Siegbahn, U. Wahlgren, *J. Phys. Chem.*, **96**, 10773 (1992).

[204] R. Åkesson, L. G. M. Pettersson, M. Sandström, U. Wahlgren, *J. Am. Chem. Soc.*, **116**, 8691 (1994).

[205] R. Åkesson, L. G. M. Pettersson, M. Sandström, U. Wahlgren, *J. Am. Chem. Soc.*, **116**, 8705 (1994).

[206] C. W. Bock, A. K. Katz, and J. P. Glusker, *J. Am. Chem. Soc.*, **117**, 3754 (1995).

[207] B. J. Mhin, S. Lee, S. J. Cho, K. Lee and K. S. Kim, *Chem. Phys. Letters*, **197**, 77 (1992).

[208] For a review, see E. Vogel and H. Günther, *Angew. Chem. Int. Ed. Engl.*, **6**, 385 (1967).

[209] G. Maier, *Angew. Chem. Int. Ed. Engl.*, **6**, 402 (1967).

[210] D. R. Boyd and M. E. Stubbs *J. Am. Chem. Soc.* **105**, 2554 (1983).

[211] P. Politzer and K. C. Daiker, in *Excited States in Organic Chemistry and Biochemistry*, B. Pullman and N. Goldblum (eds.), Reidel, Dordrecht, Holland, 1977, p 331.

[212] D. M. Hayes, S. D. Nelson, W. A. Garland, and P. A. Kollman *J. Am. Chem. Soc.* **102**, 1255 (1980).

[213] D. Cremer, B. Dick, and D. Christeu *J. Mol. Struct. (Theochem)* **110**, 277 (1984).

[214] J. M. Schulman, R. L. Disch, and M. L. Sabio *J. Am. Chem. Soc.* **106**. 7696 (1984).

[215] C. W. Bock, P. George, J. J. Stezowski, and J. P. Glusker, *Struct. Chem.*, **1**. 33 (1989).

[216] F. Scagnolari, A. Modelli, A. Bottoni, D. Jones and D. Lazzari, *J. Chem. Soc., Faraday Trans.*, **92**, 1447 (1996).

[217] P. George, C. W. Bock and J. P. Glusker, *J. Phys. Chem.*, **94**, 8161 (1990).

[218] R. Thieme and C. Weiss, *Stud. Biophys.*, **93**, 273 (1983).

[219] R. Gleiter, G. Krennrich, D. Cremer, K. Yamamoto, and I. Murata *J. Am. Chem. Soc.* **107**, 6874 (1985).

[220] O. Tapia and G. Johannin *J. Chem. Phys.* **75**, 3624 (1981).

[221] D. R. Lide (ed.), *Handbook of Chemistry and Physics*, 71 ed., CRC Press, Boca Raton, FL (1991).

[222] J. R. Gillard, M. J. Newlands, J. N. Bridson and D. J. Burnell, *Can. J. Chem.*, **69**, 1337 (1991).

[223] D. W. Jones, *J. Chem. Soc., Chem. Commun.*, 739 (1980).

[224] L. C. Burry, J. N. Bridson, and D. J. Burnell, *J. Org. Chem.*, **60**, 5931 (1995).

[225] D. F. Harvey and E. M. Grezner, *J. Org. Chem.*, **61**, 159 (1996).

[226] J. R. Gillard and D. J. Burnell, *Can. J. Chem.*, **70**, 1296 (1992).

[227] J. R. Gillard and D. J. Burnell, *J. Chem. Soc., Chem. Commun.*, 1439 (1989).

[228] S. S. Butcher, *J. Chem. Phys.*, **42**, 1830 (1965).

[229] G. Luss and M. D. Harmony, *J. Chem. Phys.*, **43**, 3768 (1965).

[230] G. Dallinga and L. H. Toneman, *J. Mol. Struct.*, **1**, 11 (1967).

[231] M. Trætteberg, *Acta Chim. Scand.*, **22**, 2305 (1968).

[232] H. Oberhammer and S. H. Bauer, *J. Am. Chem. Soc.*, **91**, 10 (1969).

[233] W. Auf der Heyde and W. Lüttke, *Chem. Ber.*, **111**, 2384 (1978).

[234] A. J. Birch, A. L. Hinde and L. Radom, *J. Am. Chem. Soc.*, **103**, 284 (1981)

[235] A. Sygula and P. W. Rabideau, *J. Mol. Struct. (Theochem)*, **262**, 117 (1992).

[236] S. Saebø and J. E. Boggs, *J. Mol. Struct.*, **73**, 137 (1981).

[237] C. Ehrendorfer and A. Karpfen, *J. Mol. Struct. (Theochem)*, **306**, 123 (1994).

[238] C. Dilauro, N. Neto and S. Califano, *J. Mol. Struct.*, **3**, 219 (1969).

[239] L. A. Carreira, R. O. Carter and J. R. Durig, *J. Chem. Phys.*, **59**, 812 (1973).

[240] D. Klapstein, personal communication to D. J. Burnell.

[241] K. Kimura (ed.), *Handbook of HeI Photoelectron Spectra of Fundamental Organic Molecules: Ionization Energies, ab initio Assigmnents, and Valence Electron Structure for 200 Molecules*, Halsted, New York (1981) p 68.

[242] M. Deleuze, P. Horeczky, J. Delhalle, and B. T. Pickup, *Int. J. Quant. Chem., Quant. Chem. Symp.*, **26**, 31 (1992).

Appendix A

Finite Difference Approximations to Higher Derivatives

Most finite-difference schemes to approximate higher order derivatives are based on a Taylor-series approximation about a point of interest, i.e.

$$f(x) = f(x^*) + \sum_{i=1}^{N} \left.\frac{\partial f}{\partial x_i}\right|_{x^*} (x_i - x_i^*) + \sum_{i=1}^{N}\sum_{j=1}^{N} \left.\frac{1}{2!}\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{x^*} (x_i - x_i^*)(x_j - x_j^*)$$
$$+ \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \left.\frac{1}{3!}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\right|_{x^*} (x_i - x_i^*)(x_j - x_j^*)(x_k - x_k^*) + \mathcal{O}(x^4), \quad (39)$$

where $x = (x_i)$. We may define $\lambda = x - x^*$. This equation can be written in the more compact tensor notation

$$f(\lambda) = f^* + f_i^* \lambda_i + \frac{1}{2!} H_{ij}^* \lambda_i \lambda_j + \frac{1}{3!} T_{ijk}^* \lambda_i \lambda_j \lambda_k + \mathcal{O}(\lambda^4). \quad (40)$$

where the summation over i, j, k, is understood. Some formulas involving only function values are

$$f_i = \frac{f(\lambda_i e_i) - f(0)}{\lambda_i} + \mathcal{O}(\lambda) \quad (41)$$

$$f_i = \frac{f(-\lambda_i e_i) - f(0)}{-\lambda_i} + \mathcal{O}(\lambda) \tag{42}$$

$$f_i = \frac{f(\lambda_i e_i) - f(-\lambda_i e_i)}{2\lambda_i} + \mathcal{O}(\lambda^2) \tag{43}$$

$$H_{ii} = \frac{f(\lambda_i e_i) + f(-\lambda_i e_i) - 2f(0)}{\lambda_i^2} + \mathcal{O}(\lambda) \tag{44}$$

$$H_{ij} = \frac{f(\lambda_i e_i + \lambda_j e_j) + f(-\lambda_i e_i - \lambda_j e_j) - f(-\lambda_i e_i + \lambda_j e_j) - f(\lambda_i e_i - \lambda_j e_j)}{4\lambda_i \lambda_j} + \mathcal{O}(\lambda) \tag{45}$$

where e_i is the i-th unit vector. Some analogous formulas involving only function derivative values are

$$H_{jk} = \frac{f_k(\lambda_j e_j) - f_k(0)}{\lambda_j} + \mathcal{O}(\lambda) \tag{46}$$

$$H_{jk} = \frac{f_k(-\lambda_j e_j) - f_k(0)}{-\lambda_j} + \mathcal{O}(\lambda) \tag{47}$$

$$H_{jk} = \frac{f_k(\lambda_j e_j) - f_k(-\lambda_j e_j)}{2\lambda_j} + \mathcal{O}(\lambda^2) \tag{48}$$

$$T_{jjk} = \frac{f_k(\lambda_j e_j) + f_k(-\lambda_j e_j) - 2f_k(0)}{\lambda_j^2} + \mathcal{O}(\lambda) \tag{49}$$

It was observed that the Hessian, when calculated with forward-difference formula 46, was not symmetric in general, so normally the Hessian was symmetrized by averaging H_{ij} and H_{ji}. Others had taken the difference between H_{ij} and H_{ji} to be a measure of the error in the finite differentiation[73], but there is useful information here, so a relationship was derived,

$$H_{kl} - H_{lk} = \frac{1}{2}(T_{kkl}\lambda_k - T_{llk}\lambda_l) + \mathcal{O}(\lambda^2). \tag{50}$$

From a forward-difference Hessian calculation, one can therefore obtain information about the difference between semidiagonal third derivatives. By itself, this is not useful, but if one of T_{kkl} and T_{llk} is known or can be estimated easily, then the other

can be determined. This is of use in the empirical estimation of the third derivative tensor where k and l correspond to certain classes of parameters.

Appendix B

OSIPE and the Object Concept

B.1 Basic Concepts

The MUNGAUSS program[28] is written according to the Open Structured Interfaceable Programming Environment (OSIPE)[110]. These tools treat scientific data as objects which can be accessed anywhere in the program, and enable the creation of new objects by manipulation of other objects, without affecting other objects. This flexibility is guaranteed by obeying the paradigm *one object* $\Leftrightarrow$ *one routine*, and global accessibility is enabled by storing all objects in the *stack*, a large unique common block. Object manipulation is handled by the the three basic functions of OSIPE, *putobj*, *getobj*, and *bldobj*:

- *Putobj* returns the index of an object to be created, and is usually called in the creation routine associated with the object. Existing objects are destroyed.[1]

[1] *ObjectSize* is the number of elements of the object. *Type* refers to REAL, INTEger, BOOLean, or CHARacter. *WordSize* is the number of bytes each element takes up.

```
indexObject = putobj ('ObjectName',
                      ObjectSize, 'Type', WordSize)
```

- *Getobj* returns the index of an existing object, which may be needed to create another object or to be printed as output. If the object does not exist in memory or on disk, then *bldobj* should be called.

```
indexObject = getobj ('ObjectName')
```

- *Bldobj* returns the index of an object by calling the appropriate creation routine, which then builds the object. *Bldobj* usually follows a *getobj* and is called when the object does not exist.[2]

```
if(lbuild) indexObject = bldobj ('ObjectName')
```

For scalars, which are of much simpler construction than a general object, two sets of routines are used: *putscX* and *getscX*, with *X* either Boolean, Character, Integer or Real. These *routines* bypass the need for addresses and work directly with the scalar value,[3] e.g.

```
call getscI ('Scalar-Name', IntegerVariableName )
call putscR ('Another-Scalar-Name', RealVariableName )
```

[2] *lbuild* is a global scalar which is set to .TRUE. if the object is neither in memory nor on disk.
[3] Unlike *getobj*, the getscX routines call a function *bldscl* (analogous to *bldobj*) directly.

B.2 Example: Hessian Matrix Construction

To illustrate these concepts, we will show how to build the Hessian matrix from a combination of default and finite difference methods (see Appendix A), as carried out in the routine HESCLC. In order to build a Hessian from finite differences, we need to address several issues.

- Will the final Hessian be made symmetric?

```
call getscb('OPT_SCB_HESSIAN_SYMM', HESSYM)
```

- How will a particular row of the Hessian be evaluated? The *defobj* command is similar to *bldobj*, but is used for the default specification of non-existent objects.

```
ixHESTYP=getobj ('HESSIAN_TYPE_BY_PARAM')
if(lbuild)ixHESTYP=defobj ('HESSIAN_TYPE_BY_PARAM')
```

- What are the dimensions of the Hessian? The *objelm* function gives the number of elements of an object.

```
NOPTPR=objelm(ixHESTYP)
HESLEN=NOPTPR*NOPTPR
```

- If we are doing forward differences, we would need the gradients evaluated at the forward-stepped parameters. The gradient is defined as zero if that particular gradient is not needed.

```
ixGRDFOR=getobj ('MATSQG_GRAD_SET_FORWARD')
if(lbuild)ixGRDFOR=bldobj ('MATSQG_GRAD_SET_FORWARD')
```

- We also need the gradients corresponding to backward differences.

```
ixGRDBAC=getobj ('MATSQG_GRAD_SET_BACKWARD')
if(lbuild)ixGRDBAC=bldobj ('MATSQG_GRAD_SET_BACKWARD')
```

- We need the gradients at the current point.

```
ixPARGRD=getobj ('OPT_PARAM_GRADIENTS')
if(lbuild)ixPARGRD=bldobj ('OPT_PARAM_GRADIENTS')
```

- We need the step sizes to be taken.

```
ixDSTEP=getobj ('FORCE_CON_PAR_STEP')
if(lbuild)ixDSTEP=defobj ('FORCE_CON_PAR_STEP')
```

- If we do not calculate by finite differences, we must provide a suitable default.

```
ixHESDEF=getobj ('MATSQG_HESSIAN_DEFAULT')
if(lbuild)ixHESDEF=defobj ('MATSQG_HESSIAN_DEFAULT')
```

- Now let us reserve memory in the stack for the Hessian.

```
ixHSGUES=putobj ('MATSQG_HESSIAN_GUESS',
.                HESLEN , 'REAL', L8)
```

- Convert the indices to stack addresses.

```
HESTYP=objadd(ixHESTYP)

PARGRD=objadd(ixPARGRD)

GRDFOR=objadd(ixGRDFOR)

GRDBAC=objadd(ixGRDBAC)

DSTEP =objadd(ixDSTEP)

HESDEF=objadd(ixHESDEF)

HSGUES=objadd(ixHSGUES)
```

- Copy the default Hessian into the Hessian.

```
call coparr (stack(HESDEF),stack(HSGUES),HESLEN)
```

- Finally, call the scientific code which computes the Hessian from finite differences.

```
CALL HESCL8 (cstack(HESTYP),stack(PARGRD),stack(GRDFOR),
.             stack(GRDBAC),stack(DSTEP),stack(HSGUES),
.             HESSYM,NOPTPR)
```

The scientific code looks like most scientific code in which the parameters are passed in as subroutine arguments.

```
SUBROUTINE HESCL8(HESTYP,! Hessian Type by parameter
```

```
     .                    PARGRD,! Parameter Gradients
     .                    GRDFOR,! Forward Gradient Set
     .                    GRDBAC,! Backward Gradient Set
     .                    DSTEP ,! Step vector from pivot
     .                    HSGUES,! Hessian Guess
     .                    HESSYM,! Symmetrize ?
     .                    NOPTPR)! Number of optimizable parameters
```

Appendix C

Total Energies

Table 1: Total energies (Hartrees) - halo and chalco Cp

	Basis Set			
Subst.	STO-3G	3-21G	6-31G*//3-21G	6-31G*[a]
H[b]	-190.457105	-191.717084	-192.791380	-192.791723
F[c]	-287.902965	-290.026480	-291.633000	-291.634395
Cl[d]	-644.456461	-648.432484	-651.686956	-651.689894
Br[e]	-2734.546596	-2751.193639	-2762.353669	-2762.354209
I	-7040.573021	-7078.980063	-7105.257246	-7105.257602
OH stag[f]	-264.284414	-266.145761	-267.635161	-267.636724
OH eg	-264.279682	-266.139062	-267.631855	-267.631636
OH gau	-264.282058	-266.139485	-267.631177	-267.63293
OH ecli[g]	-264.279141	-266.137576	-267.629759	-267.631426
SH stag[h]	-583.630103	-587.272754	-590.296663	-590.298721
SH eg[i]	-583.626000	-587.267490	-590.291616	-590.293715
SH gau[j]	-583.628277	-587.269839	-590.294630	-590.296796
SH ecli[k]	-583.626265	-587.267959	-590.292842	-590.294791
SeH stag[l]	-2564.011225	-2580.199631	-2590.471634	-2590.471889
SeH eg	-2564.008029	-2580.194794	-2590.467560	-2590.467832
SeH gau	-2564.009613	-2580.197518	-2590.470494	-2590.470799
SeH ecli[m]	-2564.008062	-2580.195854	-2590.468745	-2590.469048
TeH stag	-6737.587268	-6774.897344	-6799.790044	-6799.790395
TeH eg	-6737.585194	-6774.894325	-6799.787268	-6799.787625
TeH gau	-6737.586210	-6774.896842	-6799.790059	-6799.790430
TeH ecli	-6737.585046	-6774.895031	-6799.788477	-6799.788922

[a]For atoms with $Z > 18$, Huzinaga's (4333/433/4) or (43333/4333/43) minimal basis sets, decontracted in all valence shells to an (N1) representation (the filled d-shell is treated as core), with an additional d-polarization function, were used in conjunction with the proper 6-31G* basis set for atoms of $Z \leq 18$.

[b]CpH: 6-31G**//3-21G = -192.802038, 6-31G** = -192.802419

[c]CpF: 6-31G**//3-21G = -291.642164

[d]CpCl: 3-21G(*) = -648.446039, 6-31G*//3-21G(*) = -651.688966, 6-31G**//3-21G = −651.695890, 6-31G**//3-21G(*) = -651.697875

[e]CpBr: BC//3-21G -2762.099538. BC = Binning-Curtiss basis set for third row p-block elements

[f]H-X-C-H: stag = 180 stag, eg = 120, gau = 60, ecli = 0

[g]CpOH ecli: 6-31G**//3-21G = -267.645234

[h]CpSH stag: 3-21G(*) = -587.305939, 6-31G*//3-21G(*) = -590.298174

[i]CpSH eg: 3-21G(*) = -587.300102, 6-31G*//3-21G(*) = -590.293273

[j]CpSH gau: 3-21G(*) = -587.303672, 6-31G*//3-21G(*) = -590.296292

[k]CpSH ecli: 3-21G(*) = -587.301538, 6-31G*//3-21G(*) = -590.294319, 6-31G**//3-21G = −590.305732, 6-31G**//3-21G(*) = -590.307168

[l]CpSeH stag: BC//3-21G = -2590.365219

[m]CpSeH ecli: BC//3-21G = -2590.362008

Table 2: Total energies (Hartrees) - pnicto and group IV Cp

	Basis Set			
Subst.	STO-3G	3-21G	6-31G*//3-21G	6-31G*
NH_2 stag[a]	-244.763505	-246.419967	-247.803417	-247.805195
NH_2 ecli[b]	-244.760699	-246.420311	-247.801533	-247.803036
NH_2 gau[c]	-244.766333	-246.426475	-247.807834	-247.809306
NH_2 eg[d]	-244.760351	-246.419491	-247.801281	-247.803093
PH_2 stag[e]	-527.956263	-531.278932	-534.082824	-534.084217
PH_2 ecli[f]	-527.953378	-531.275069	-534.077645	-534.078901
PH_2 gau[g]	-527.957473	-531.280574	-534.083203	-534.084601
PH_2 eg[h]	-527.952981	-531.274560	-534.078161	-534.079392
AsH_2 stag[i]	-2400.334032	-2415.728144	-2425.515715	-2425.516566
AsH_2 ecli[j]	-2400.331950	-2415.724197	-2425.510820	-2425.511270
AsH_2 gau	-2400.335341	-2415.729347	-2425.515801	-2425.516181
AsH_2 eclgau	-2400.331580	-2415.724312	-2425.511623	-2425.512072
SbH_2 stag	-6442.406923	-6478.398028	-6502.212254	-6502.212655
SbH_2 ecli	-6442.405432	-6478.393010	-6502.206775	-6502.207126
SbH_2 gauche	-6442.407612	-6478.397291	-6502.210774	-6502.211166
SbH_2 eg	-6442.405044	-6478.393819	-6502.208163	-6502.208540
CH_3 stag[k]	-229.037614	-230.537736	-231.825964	-231.826583
CH_3 ecli[l]	-229.032072	-230.530892	-231.819349	-231.819873
SiH_3 stag[m]	-477.245364	-480.273934	-482.871879	-482.872605
SiH_3 ecli[n]	-477.242213	-480.268692	-482.866779	-482.867385
GeH_3 stag[o]	-2243.412586	-2258.232905		-2267.374437
GeH_3 ecli[p]	-2243.409996	-2258.228230		-2267.370206
SnH_3 stag	-6154.947524	-6189.542977	-6212.428279	-6212.428784
SnH_3 ecli	-6154.945861	-6189.539254	-6212.425256	-6212.425719

[a] 3-21G(N) = -246.441174, 6-31G*//3-21G(N) = -247.804369, 6-31G**//3-21G = -247.819920

[b] 3-21G(N) = -246.440334, 6-31G*//3-21G(N) = -247.801855, 6-31G**//3-21G = -247.817960

[c] 3-21G(N) = -246.447051, 6-31G*//3-21G(N) = -247.808571

[d] 3-21G(N) = -246.439361, 6-31G*//3-21G(N) = -247.802382

[e] 3-21G(*) = -531.326084, 6-31G(*)//3-21G = -534.011010, 6-31G*//3-21G(*) = -534.083704, 6-31G**//3-21G(*) = -534.096827

[f] 3-21G(*) = -531.320423, 6-31G*//3-21G(*) = -534.078512, 6-31G**//3-21G = -534.090995, 6-31G**//3-21G(*) = -534.091742

[g] 3-21G(*) = -531.326979, 6-31G*//3-21G(*) = -534.084024

[h] 3-21G(*) = -531.320795, 6-31G*//3-21G(*) = -534.078930

[i] BC//3-21G = -2425.350590

[j] BC//3-21G = -2425.343899

[k] 6-31G**//3-21G = -231.839607

[l] 6-31G**//3-21G = -231.833024

[m] 3-21G(*) = -480.322487, 6-31G*//3-21G(*) = -482.872008, 6-31G**//3-21G = -482.885234

[n] 3-21G(*) = -480.316386, 6-31G*//3-21G(*) = -482.866890, 6-31G**//3-21G = -482.880141

[o] BC//3-21G = -2267.195257

[p] BC//STO-3G = -2267.176062, BC//3-21G = -2267.190323

Table 3: Total energies (Hartrees) - 1,2-shift transition states of Cp

	Basis Set			
Subst.	STO-3G	3-21G	6-31G*//3-21G	6-31G*
H[a]	-190.364536	-191.649946	-192.729024	-192.729920
F[b]	-287.787043	-289.929154	-291.530567	-291.531432
Cl[c]	-644.375111	-648.378189		-651.623608
Br[d]	-2734.471706	-2751.140283		-2762.300165
I[e]	-7040.510758	-7078.936035		-7105.214504
OH	-264.185384	-266.060652	-267.536692	-267.538107
SH[f]	-583.551410	-587.213865		-590.229949
SeH	-2563.940822	-2580.143150	-2590.414627	-2590.415165
TeH	-6737.528386	-6774.852609	-6799.745483	-6799.745936
NH_2[g]	-244.626151	-246.314335	-247.690752	-247.691905
PH_2[h]	-527.876060	-531.228286		-534.032113
AsH_2	-2400.265611	-2415.686412	-2425.475570	-2425.476265
SbH_2	-6442.357861	-6478.371503	-6502.187564	-6502.187965
CH_3[i]	-228.918961	-230.453520	-231.735104	-231.735993
SiH_3[j]	-477.183429	-480.237783		-482.836890
GeH_3	-2243.355634	-2258.200768		-2267.341356
SnH_3		-6189.523708	-6212.409898	-6212.410419

[a] 6-31G**//3-21G = -192.741961, 6-31G** = -192.742885
[b] 6-31G**//3-21G = -291.539828
[c] 3-21G(*) = -648.382906, 6-31G*//3-21G(*) = -651.622858, 6-31G**//3-21G = -651.631954
[d] Huz(5d) = -2762.291238
[e] Huz(5d) = -7105.061443
[f] 3-21G(*) = -587.239116, 6-31G*//3-21G(*) = -590.229177
[g] 6-31G**//3-21G = -247.707564
[h] 3-21G(*) = -531.275324, 6-31G*//3-21G(*) = -534.031570, 6-31G**//3-21G = -534.043768
[i] 6-31G**//3-21G = -231.749249
[j] 3-21G(*) = -480.287195, 6-31G*//3-21G(*) = -482.836481, 6-31G**//3-21G = -482.849406

Table 4: Total energies (Hartrees) - lithium-water complexes

		Basis Set			
Subst.	Symmetry	STO-3G	3-21G	6-31G*	6-31+G*
H_2O[a]	C_{2v}	-74.965901	-75.585960	-76.010746	-76.017743
Li^+[b]		-7.135448	-7.187094	-7.235537	-7.235537
$LiAq_1^+$[c]	C_{2v}	-82.226811	-82.863928	-83.309321	-83.311792
$LiAq_2^+$[d]	D_{2d}	-157.307766	-158.531301	-159.375852	-159.381124
$LiAq_2^+$	D_{2h}	-157.306213	-158.530503	-159.375496	
$LiAq_3^+$[e]	D_3	-232.365675	-234.181210	-235.430142	-235.438058
$LiAq_3^+$	D_{3h} (1)	-232.361561	-234.179566	-235.427343	
$LiAq_3^+$	D_{3h} (2)	-232.363471	-234.177060	-235.428346	
$LiAq_4^+$[f]	S_4	-307.402731	-309.812881	-311.472448	-311.483872
$LiAq_4^+$	D_{2d}	-307.401573	-309.811099	-311.470793	
$LiAq_4^+$	C_{2v}	-307.399415			
$LiAq_5^+$	C_2	-382.403629	-385.430199	-387.499013	
$LiAq_5^+$	C_{2v} (1)	-382.402860	-385.418258	-387.496165	
$LiAq_5^+$	C_{2v} (2)	-382.401818			
$LiAq_5^+$	C_{2v} (3)	-382.401989			
$LiAq_5^+$	C_{2v} (4)	-382.402923			
$LiAq_6^+$	T_h	-457.403398	-461.023032	-463.522675	
$LiAq_8^+$	D_{2d}	-607.375526	-612.317894	-615.598715	-615.625420
$LiAq_8^+$	S_4	-607.376058	-612.283858		
$LiAq_{12}^+$	S_4	-907.327995	-914.804356	-919.713806	

[a] 6-31G** = -76.023615, 6-311G* = -76.032400, 6-31+G(5d) = -76.016549

[b] 6-31G = -7.235480, 6-31G* = -7.235537, 6-311G* = -7.235839, 6-31+G(5d) = -7.235485

[c] 6-31G** = -83.322262, 6-311G* = -83.334557, 6-31+G(5d) = -83.310390, 6-31G(5d,O+) = -83.310342, 6-31G(O5d+) = -83.310118

[d] 6-31G** = -159.401675, 6-311G* = -159.425121, 6-31+G(5d) = -159.378559

[e] 6-31G** = -235.468877, 6-311G* = -235.502645, 6-31+G(5d) = -235.434296

[f] 6-31G** = -311.524039, 6-311G* = -311.567379, 6-31+G(5d) = -311.478928

Table 5: Total energies (Hartrees) - metal-water complexes

		Basis Set		
Subst.	Symmetry	STO-3G	3-21G	6-31G* or Huz[a]
$BeAq_4^{2+}$	S_4	-314.250308	-316.659789	-318.299157
$BeAq_8^{2+}$	D_{2d}			-622.516144
$MgAq_6^{2+}$	T_h	-647.056820	-651.916609	-655.423792
$MgAq_{18}^{2+}$	T_h (1)		-1559.358386	-1567.834433
$MgAq_{18}^{2+}$	T_h (2)		-1559.401705	-1567.834331
$AlAq_6^{3+}$	T_h	-688.489827	-693.544359	-697.180781
$ScAq_6^{3+}$	T_h	-1201.864145	-1209.108215	-1214.934705
$ScAq_8^{3+}$	C_4	-1351.960852	-1360.407505	-1367.045533
$FeAq_6^{2+}$	T_h	-1698.452606		
$ZnAq_6^{2+}$	T_h		-2222.524942	-2233.070920
$GaAq_6^{3+}$	T_h	-2350.226269	-2366.751560	-2376.600280
$CdAq_6^{2+}$	T_h	-5861.440983	-5893.514809	-5916.759302
$InAq_6^{3+}$	T_h	-6132.340262	-6167.976073	-6191.107939

[a]Sc,Zn = (53321/5211*/41); Ga = (43321/4321/41*); Cd = (433321/43211*/421); In = (433321/43321/431*)

Table 6: Total energies (Hartrees) - benzene oxide–oxepin

	Benzene Oxide	Trans. State	Oxepin	Planar Oxepin
STO-3G	-301.687829	-301.643559	-301.665339	-301.660278
3-21G	-303.766826	-303.751522	-303.790992	-303.788062
6-31G*//3-21G	-305.479218	-305.457363	-305.487721	-305.483741
6-31G*	-305.485236	-305.462294	-305.489824	-305.485167
6-31G**//3-21G	-305.489734	-305.468144	-305.498521	-305.494468
MP2/6-31G*//3-21G	-306.419080	-306.411101	-306.411703	-306.404694
MP2/6-31G*//6-31G*	-306.418341	-306.410796	-306.413061	-306.405171
MP2/6-31G*	-306.423343	-306.413249	-306.418048	-306.408625
MP3/6-31G*//MP2	-306.450567	-306.434580	-306.449494	-306.442590
MP4DQ/6-31G*//MP2	-306.451963	-306.434488	-306.450538	-306.444631
MP4SDQ/6-31G*//MP2	-306.462965	-306.445244	-306.461691	-306.456677
MP4SDTQ/6-31G*//MP2	-306.506256	-306.493798	-306.504194	-306.497558
MP2/6-31G**//MP2	-306.470559	-306.460535	-306.465397	-306.455971
MP2/6-31G(2d)//MP2	-306.496471	-306.489122	-306.493283	-306.485013
MP2/6-31+G*//MP2	-306.444001	-306.435138	-306.439029	-306.428361
MP2/6-311G*//MP2	-306.535333	-306.528032	-306.531827	-306.521668
QCISD/6-31G*//MP2	-306.465187	-306.448332	-306.464577	-306.459619
QCISD(T)/6-31G*//MP2	-306.504742	-306.491004	-306.504196	-306.498256
3-21G: ZPE (kJ/mol)	294.92	289.57	294.26	295.38
3-21G: Therm (kJ/mol)	12.71	12.07	13.86	12.07
3-21G: S (J/mol K)	301.84	298.03	310.39	293.64
6-31G*: ZPE (kJ/mol)	296.16	289.96	294.10	294.57
6-31G*: Therm (kJ/mol)	12.57	12.08	13.79	12.19
6-31G*: S (J/mol K)	300.89	297.87	309.23	294.14

Table 7: Total energies (Hartrees) - benzene sulfide–thiepin

	Benzene Sulfide	Trans. State	Thiepin	Planar Thiepin
STO-3G	-621.033182	-620.963125	-621.005498	-621.000455
STO-3G*	-621.073645	-621.011068	-621.045297	-621.040034
3-21G	-624.929806	-624.874258	-624.910260	-624.897402
3-21G(*)	-624.951097	-624.912440	-624.946319	-624.933135
6-31G*	-628.161697	-628.112218	-628.150349	-628.140156
MP2/6-31G*//6-31G*	-629.045289	-629.016204	-629.027167	-629.011808
MP2/6-31G*	-629.047907	-629.017630	-629.031091	-629.014509
MP3/6-31G*//MP2	-629.085589	-629.045100	-629.071450	-629.058833
MP4DQ/6-31G*//MP2	-629.086437	-629.045605	-629.072137	-629.060123
MP4SDQ/6-31G*//MP2	-629.095821	-629.055534	-629.082350	-629.070823
MP4SDTQ/6-31G*//MP2	-629.138794	-629.106097	-629.125576	-629.112180
MP2/6-31G**//MP2	-629.095238	-629.065376	-629.078806	-629.062281
MP2/6-31G(2d)//MP2	-629.117289	-629.088194	-629.099523	-629.084847
MP2/6-31+G*//MP2	-629.063801	-629.036184	-629.049225	-629.031033
MP2/6-311G*//MP2	-629.155684	-629.128053	-629.138749	-629.120692
QCISD/6-31G*//MP2	-629.097943	-629.058365	-629.085146	-629.073802
QCISD(T)/6-31G*//MP2	-629.138445	-629.103128	-629.126314	-629.114142
3-21G: ZPE (kJ/mol)	288.34	282.69	286.00	286.37
3-21G: Therm (kJ/mol)	14.42	13.07	14.85	13.28
3-21G: S (J/mol K)	317.88	309.07	320.41	305.55
6-31G*: ZPE (kJ/mol)	289.02	282.71	285.46	285.60
6-31G*: Therm (kJ/mol)	13.80	13.08	14.83	13.24
6-31G*: S (J/mol K)	313.57	308.91	320.25	305.11

Table 8: Total energies (Hartrees) - xylene oxide–dimethyloxepin

	Xylene Oxide	Trans. State	DMO	Planar DMO
STO-3G	-378.857387	-378.815803	-378.841913	-378.834131
3-21G	-381.415759	-381.399079	-381.445243	-381.442003
6-31G*//3-21G	-383.559134	-383.536667	-383.573857	-383.568404
6-31G*	-383.564757	-383.541511	-383.575747	-383.569582
6-31G**//3-21G	-383.576011	-383.553621	-383.590733	-383.585277
MP2/6-31G*//6-31G*	-384.766355	-384.759641	-384.765099	-384.755143
MP2/6-31G*	-384.771357	-384.762837	-384.770559	-384.758873
MP3/6-31G*//MP2	-384.814729	-384.797928	-384.817985	-384.809315
MP4DQ/6-31G*//MP2	-384.818176	-384.799983	-384.820597	-384.812682
MP4SDQ/6-31G*//MP2	-384.831901	-384.813002	-384.834103	-384.826940
MP4SDTQ/6-31G*//MP2	-384.885583	-384.873113	-384.886965	-384.877936
QCISD/6-31G*//MP2	-384.834744	-384.816574	-384.837568	-384.830480
QCISD(T)/6-31G*//MP2	-384.883869	-384.869270	-384.886660	-384.878469
3-21G: ZPE (kJ/mol)	448.90	444.16	450.34	451.54
3-21G: Therm (kJ/mol)	21.02	20.91	21.64	19.60
3-21G: S (J/mol K)	363.01	366.96	369.68	349.21
6-31G*: ZPE (kJ/mol)	450.22	444.88	449.99	450.06
6-31G*: Therm (kJ/mol)	20.46	20.35	21.29	19.57
6-31G*: S (J/mol K)	358.08	360.33	366.13	348.55

Table 9: Total energies (Hartrees) - protonated benzene oxide–oxepin

	Benzene Oxide	Trans. State	Oxepin
STO-3G	-302.082163	-302.025278	-302.059581
3-21G	-304.114655	-304.078861	-304.120645
6-31G*	-305.816604	-305.770258	-305.800813
MP2/6-31G*//6-31G*	-306.740326	-306.715362	-306.717944
MP2/6-31G*	-306.745634	-306.718758	-306.725006
6-31G*: ZPE (kJ/mol)	330.90	326.27	328.24
6-31G*: Therm (kJ/mol)	13.66	12.49	14.39
6-31G*: S (J/mol K)	307.23	300.38	312.01

Table 10: Total energies (Hartrees) - isodesmic reaction comparisons

	C_2H_4O	C_2H_6O	C_2H_4S	C_2H_6S
STO-3G	-150.928501	-152.133869	-470.276760	-471.482790
STO-3G*			-470.318291	-471.517672
3-21G	-152.000703	-153.213209	-473.157121	-474.348578
3-21G(*)			-473.185931	-474.380579
6-31G*//3-21G	-152.861866	-154.062355		
6-31G*	-152.867354	-154.064741	-475.546814	-476.735331
6-31G**//3-21G	-152.868594	-154.071780		
MP2/6-31G*//3-21G	-153.301933	-154.502042		
MP2/6-31G*//6-31G*	-153.301624	-154.502083	-475.928827	-477.120662
MP2/6-31G*	-153.303583	-154.503455	-475.929228	-477.121108
MP3/6-31G*//MP2	-153.319311	-154.528005	-475.956162	-477.156895
MP4DQ/6-31G*//MP2	-153.322890	-154.532930	-475.958581	-477.160997
MP4SDQ/6-31G*//MP2	-153.327376	-154.536848	-475.961739	-477.163923
MP4SDTQ/6-31G*//MP2	-153.341476	-154.548247	-475.975663	-477.175029
MP2/6-31G**//MP2	-153.335365	-154.551756	-475.961027	-477.169277
MP2/6-31G(2d)//MP2	-153.341973	-154.545281	-475.963254	-477.156524
MP2/6-31+G*//MP2	-153.314735	-154.514561	-475.934603	-477.125880
MP2/6-311G*//MP2	-153.367075	-154.570171	-475.988262	-477.181664
QCISD/6-31G*//MP2	-153.328462	-154.538574	-475.962740	-477.165519
QCISD(T)/6-31G*//MP2	-153.341405	-154.549428	-475.976605	-477.177133

Table 11: Total energies (Hartrees) - Diels-Alder - benzene oxide + ethylene

	syn TS	anti TS	syn prod	anti prod
STO-3G	-378.687437	-378.705531	-378.891394	-378.892074
3-21G	-381.306362	-381.319957	-381.426680	-381.422287
6-31G*//3-21G	-383.430146	-383.444819	-383.552286	-383.550866
6-31G*	-383.435957	-383.450555	-383.557749	-383.556156
MP2/6-31G*//HF/6-31G*	-384.672934	-384.688164		
MP2/6-31G*	-384.674792	-384.691288		
3-21G: ZPE (kJ/mol)	448.99	449.16	464.53	463.09
3-21G: Therm (kJ/mol)	16.96	17.10	15.06	15.45
3-21G: S (J/mol K)	333.24	335.13	321.26	324.23
6-31G*: ZPE (kJ/mol)	448.77	448.75		
6-31G*: Therm (kJ/mol)	17.07	17.15		
6-31G*: S (J/mol K)	334.15	335.37		

Table 12: Total energies (Hartrees) - Diels-Alder - benzene oxide + acetylene

	syn TS	anti TS
STO-3G	-377.458613	-377.478131
3-21G	-380.085161	-380.106841
6-31G*//3-21G	-382.205118	-382.225629
6-31G*	-382.211038	-382.231372
MP2/6-31G*//HF/6-31G*	-383.442404	-383.463014
MP2/6-31G*	-383.446401	-383.467497
3-21G: ZPE (kJ/mol)	376.87	378.64
3-21G: Therm (kJ/mol)	17.36	16.81
3-21G: S (J/mol K)	336.22	332.01
6-31G*: ZPE (kJ/mol)	377.12	378.28
6-31G*: Therm (kJ/mol)	17.31	16.88
6-31G*: S (J/mol K)	335.67	332.44

Table 13: Total energies (Hartrees) - dienophiles

	C_2H_4	C_2H_2
STO-3G	-77.073956	-75.856248
3-21G	-77.600989	-76.395958
6-31G*//3-21G	-78.031694	-76.817755
6-31G*	-78.031719	-76.817826
6-31G**//3-21G	-78.038809	
MP2/6-31G*//3-21G	-78.284100	-77.064660
MP2/6-31G*//6-31G*	-78.284350	-77.064634
MP2/6-31G*	-78.285028	-77.066794
MP3/6-31G*//MP2	-78.305966	-77.075828
MP4DQ/6-31G*//MP2	-78.308697	-77.077655
MP4SDQ/6-31G*//MP2	-78.311396	-77.081541
MP4SDTQ/6-31G*//MP2	-78.319826	-77.093574

Table 14: Total energies (Hartrees) - 1,3-cyclohexadienes

	C_2	C_{2v} TS
STO-3G	-229.043836	-229.042284
3-21G	-230.543233	-230.539674
6-31G*//3-21G	-231.831628	-231.828453
6-31G*	-231.831916	-231.828689
6-31G**//3-21G	-231.844926	-231.841756
MP2/6-31G*//3-21G	-232.590754	-232.585119
MP2/6-31G*//6-31G*	-232.591366	-232.585879
MP2/6-31G*	-232.593619	-232.587648
MP3/6-31G*//MP2	-232.636518	-232.631835
MP4DQ/6-31G*//MP2	-232.639114	-232.634564
MP4SDQ/6-31G*//MP2	-232.646712	-232.642148
MP4SDTQ/6-31G*//MP2	-232.678383	-232.673246
ZPE (3-21G)	346.68	346.45
ZPE (6-31G*)	345.66	345.21

Table 15: Total energies (Hartrees) - 2,4-cyclohexadiene-1-ol

OH position	OH conf.	STO-3G	3-21G	6-31G*//3-21G	6-31G*
axial	stag.	-302.874737	-304.979027	-306.682050	-306.683579
	gau. out	-302.873760			-306.680795
	gau. in	-302.872694	-304.973008	-306.678012	-306.679729
equatorial	stag.	-302.873449	-304.978483	-306.680950	-306.682563
	gau. out	-302.872972	-304.976467	-306.679540	-306.681237
	gau. in	-302.872870	-304.976801	-306.679904	-306.681639

Table 16: Total energies (Hartrees) - cis-3,5-cyclohexadiene-1,2-diol

OH conf.	STO-3G	3-21G	6-31G*
stag, stag (1)	-376.705770	-379.419090	-381.536251
stag, gau out (8)	-376.700276	-379.406350	-381.527880
stag, gau in (7)	-376.702939		
gau out, stag (2)	-376.704907		-381.533335
gau out, gau out (4)	-376.698080		-381.523407
gau out, gau in (3)	-376.699735		
gau in, stag (9)	-376.702781		-381.532313
gau in, gau out (5)	-376.703459	-379.411934	-381.531302
gau in, gau in (6)	-376.702811	-379.411482	-381.531109
gau out, gau out C_s	-376.694577	-379.388184	

Table 17: Total energies (Hartrees) - cis-3,5-cyclohexadiene-1,2-diol derivatives

Derivative	STO-3G	3-21G	6-31G*//3-21G	6-31G*
-CH2- endo	-414.159758	-417.072396	-419.399422	-419.404877
-CH2- exo	-414.160160	-417.069524	-419.398407	-419.404414
-BH-	-400.658487	-403.502521	-405.749053	-405.754211
-CO-	-486.845755	-490.384946	-493.140550	-493.145448
-CMe2- endo	-491.331277	-494.729431	-497.483444	-497.488107
-CMe2- exo	-491.332236	-494.727491	-497.483346	-497.488409
-SiH2-	-662.399133	-666.904956	-670.536655	-670.541466
-SiMe2-	-739.598634	-744.604854	-748.640771	-748.645405

Table 18: Total energies (Hartrees) - 1,3-cyclohexadiene + dienophile transition state

	C_2H_4	C_2H_2
STO-3G	-306.055778	-304.831295
3-21G	-308.085800	-306.876037
6-31G*//3-21G	-309.789145	-308.573102
6-31G*	-309.789442	-308.573500
MP2/6-31G*//6-31G*	-310.851008	-309.628991
MP2/6-31G*	-310.850883	-309.630647