NOTE TO USERS

uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Michelle Shaw
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Chemistry)
GRADE / DEGREE

Department of Chemistry
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Application of Density Functional Theory to Large Systems: Implementation of the Continuous Fast Multiple Method and Investigation of Models of Pentacoordinate Phosphorus

TITRE DE LA THÈSE / TITLE OF THESIS

Alain St-Amant
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Gerald Buchanan

René Fournier

Javier Giorgi

Natalie Goto

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Application of Density Functional Theory to Large Systems: Implementation of the Continuous Fast Multipole Method and Investigation of Models of Pentacoordinate Phosphorus

Michelle Shaw

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
for the PhD degree in Chemistry

Department of Chemistry
Faculty of Science
University of Ottawa

Your file Votre référence
ISBN: 0-494-11022-8
Our file Notre référence
ISBN: 0-494-11022-8

Abstract

The density functional package DeFT is used for systems with a large number of charge distributions, and the bottleneck of the computation is the evaluation of two-electron integrals. Previously, it scaled quadratically with system size. A particular strategy, called the continuous fast multipole method, has been implemented in DeFT and combined with traditional methods for evaluating the two-electron integrals. The result is a code which scales linearly with system size and is able to treat larger systems.

The Continuous Fast Multipole Method (CFMM) uses the divide-and-conquer strategy to approximate the pairwise potentials between continuous charge distributions. Pairwise potentials are computed using the direct method if they are within a defined distance of one another, and are approximated by a multipolar expansion if they are not. The algorithm involves successive subdivisions of the simulation space until the desired precision is reached, which reduces the number of pairwise potentials evaluated. For large systems, this algorithm scales linearly with the number of charge distributions.

Enzyme-catalyzed phosphoryl transfer reactions have many applications in biological systems. It has been shown in previous investigations that phosphoryl transfer in

enzymes proceeds via a pentacoordinate phosphorus intermediate. Few theoretical investigations have been conducted on this class of enzymes due to the large number of atoms treated quantum mechanically.

In this investigation, models of the pentacoordinate phosphorus intermediate are investigated with density functional methods. Comparison of the results of calculations using the VSXC functional with experiment and perturbation theory are employed to validate its use in investigations of phosphoryl transfer reactions.

Acknowledgements

There are several people I would like to thank for their guidance and support while this work was being completed. First, I would like to thank my supervisor Alain St-Amant for his advice and assistance which was invaluable in the completion of this work. I would also like to acknowledge the generous financial support he gave me during my doctoral studies. I would like to thank current and past group members Delphine Courmier, Etienne Paradis, Wei Quan Tian, Edet Archibong, and Sebastien Osborne for their continued support and encouragement. Thanks to the National Sciences and Engineering Research Council of Canada, and the University of Ottawa for financial support and giving me the opportunity to complete my doctoral work here. Finally, I would like to acknowledge the high performance computing virtual laboratory for the use of their facilities to complete this work.

I would like to extend a special thanks to my family and friends, especially my parents. Their continued support and encouragement throughout the completion of the project will always be greatly appreciated.

Contents

List of Tables

List of Figures

List of Abbreviations

HF	Hartree-Fock
DFT	density functional theory
TF	Thomas-Fermi
TFD	Thomas-Fermi-Dirac
TFDW	Thomas-Fermi-Dirac-Weizsacker
HK	Hohenberg-Kohn
KS	Kohn-Sham
XC	exchange-correlation
LDA	local density approximation
SVWN	Slater's exchange with Vosko, Wilk, and Nusair correlation
LSDA	local spin density approximation
VWN	Vosko, Wilk, and Nusair
GGA	generalized gradient approximation
B88	Becke's 1988 exchange functional
P86	Perdew's 1988 correlation functional
LYP	Lee, Yang, and Parr

VSXC	Van Voorhis and Scuseria exchange-correlation
B3LYP	Becke's 3 parameter functional with LYP correlation
PW91	Perdew and Wang's 1991 functional
GTO	Gaussian-type orbital
OS	Obara-Saika
HGP	Head-Gordon-Pople
GR	Greengard-Rokhlin
RBM	recursive bisection method
QCTC	quantum chemical tree code
MAC	multipole acceptance criterion
GvFMM	Gaussian very fast multipole method
CFMM	continuous fast multipole method
NF	near-field
FF	far-field
WS	well-separated
NMR	nuclear magnetic resonance
CCSD	coupled cluster with singles and doubles
CCSD(T)	coupled cluster with singles, doubles, and approximate triples
MCSCF	multiconfigurational self-consistent field
TMS	tetramethylsilane
GIAO	gauge-including atomic orbital
CSGT	continuous set of gauge transformations

CPHF	coupled perturbed Hartree-Fock
AO	atomic orbital
MO	molecular orbital
HOMO	highest occupied molecular orbital
LUMO	lowest unoccupied molecular orbital
SCF	self-consistent field
CPKS	coupled perturbed Kohn-Sham
RNA	ribonucleic acid

Chapter 1

Density Functional Theory

1.1 Introduction

Many of today's methods for calculating the properties of a system require the use of an N-electron wavefunction. Although accurate, these methods are time consuming and computationally expensive. The equations are complex, requiring one to solve for a system of 4N variables (3N spatial and N spin). As a result, they are applied only to small and moderately sized systems [2].

More recently, another set of methods has become popular, particularly for calculations of large systems. These methods are less computationally demanding and for roughly the cost of a Hartree-Fock (HF) calculation, one can include some electron correlation. Collectively called *Density Functional Methods*, they are based on a theory which uses the electronic density to calculate the properties of a system (called Density Functional Theory, or DFT) [3].

1.2 The Thomas-Fermi and Related Models

The pioneering work in developing a density functional method of practical use was done by Thomas and Fermi in the 1920's. Their main motivation was to develop a theory to express the energy of a system in terms of its electronic density instead of an N-electron wavefunction. Since the electron density is a function of only three variables, this would greatly simplify solving the resulting system of equations.

Thomas and Fermi based their model on the *noninteracting uniform electron gas*, where the electrons were uniformly distributed in space and were assumed not to interact with each other. From this model they applied statistics to the distribution of electrons in space and were able to connect the (probability) density of electrons in space to their kinetic energy. Dividing up the space into small cubes of volume V, they obtained the following expression for the density [4–7]:

$$\rho = \frac{\Delta N}{\Delta V} \tag{1.1}$$

Using this definition for the density, Thomas and Fermi defined the kinetic energy as [4–7]:

$$T_{TF}[\rho(\mathbf{r})] = C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} \text{ where } C_F = \frac{3}{10}(3\pi^2)^{2/3}, \tag{1.2}$$

with the following expression for the energy from the Thomas-Fermi (TF) model [4–7]:

$$E_{TF}[\rho(\mathbf{r})] = C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} - Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r_1})\rho(\mathbf{r_2})}{\mathbf{r_{12}}} d\mathbf{r_1} d\mathbf{r_2}. \tag{1.3}$$

The first term in the TF energy functional is the kinetic energy, the second term is the electron-nuclear attraction energy, and the third term is the classic Coulomb repulsion energy. Note that the above expression does not include contributions due to the nonclassical exchange.

In order to make this method applicable to calculations on atomic and molecular systems, an iterative procedure was needed to obtain the density. To do this, it was assumed that there was a density which would minimize the energy functional given above. Applying the constraint that the integration of the density would give the number of electrons and solving the problem variationally, the electron density was obtained. Thomas and Fermi applied this method to calculate properties of atoms from their electron densities.

The TF model, although simple, did not give good results for atoms and failed for molecules. Several others developed models which tried to correct some of the problems. The Thomas-Fermi-Dirac (TFD) model was one case, where the Dirac approximation to exchange was added to the TF model. This exchange term was also derived using the uniform electron gas approximation. The resulting energy expression is:

$$\begin{aligned} E_{TFD}[\rho(\mathbf{r})] = C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} - Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r_1})\rho(\mathbf{r_2})}{\mathbf{r_{12}}} d\mathbf{r_1} d\mathbf{r_2} - \\ C_x \int \rho^{4/3}(\mathbf{r}) d\mathbf{r} \end{aligned} \tag{1.4}$$

The above expression differs from the TF energy in the last term, which is the Dirac exchange functional.

There were a few serious problems with the TF and TFD models. First, the behavior of the density at small values of $\mathbf{r}$ was incorrect. This was in comparison to the density obtained from the N-electron wavefunction. In the limit of $\mathbf{r} \rightarrow 0$, the density should have had a cusp in it. The density obtained from the TF and TFD theories was infinite in the same limit. Second, the TF and TFD models did not predict binding properly. According to these theories, if a molecule AB dissociated into A and B, then AB was unstable relative to A and B. This indicated that the TF and TFD models were unsuitable for describing bonding, and thus gave incorrect results for molecules. It was possible to improve the results of these models if the electron density from the N-electron wavefunction was used instead. One final issue that cannot simply be solved by using an exact density was the lack of treatment of the self-interaction in the TFD model. In Hartree-Fock theory, this self-interaction cancelled in the expression $J - K$ as $J = K$ in this case. However, for the TFD model this was not the case [3].

One improvement to the TFD model, originally proposed by von Weizsacker was to take the inhomogeneous electron distribution into account. This was done by using the gradient of the density. This was the first example of a *gradient-corrected functional.* The method was called the Thomas-Fermi-Dirac-Weizsacker (TFDW) model. The modification

was made to the kinetic energy functional of Thomas and Fermi:

$$T_W[\rho(\mathbf{r})] = \frac{1}{8}\frac{\hbar^2}{m}\int \frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}d\mathbf{r}$$
$$T_{TFDW}[\rho(\mathbf{r})] = T_{TF}[\rho(\mathbf{r})] + \lambda T_W[\rho(\mathbf{r})] \tag{1.5}$$

The original TFDW model employed a value of 1 for λ, but a few other values have also been used. The TFDW model corrected the behavior of the density at the nucleus. The behavior of the density far from the nucleus was also described better with the TFDW model. Binding was also possible with this model, however, the strengths predicted were questionable. One other shortcoming of this method was that the iterative solution of the equations for obtaining the density must be done numerically [3].

1.3 The Hohenberg-Kohn Model

Thomas and Fermi developed a theory relating the electronic density to the kinetic energy which enabled the theory to be applied to atoms. Extensions of this theory allowed it to be applied to molecules. However, no rigorous connection had been made between the electron density and the Hamiltonian. For a rigorous theory this was needed, since the number of electrons and external potential together completely specified the system [3].

In 1964, Hohenberg and Kohn provided the proof needed that the electron density and the density functional theory could be used to determine the properties of a system. They first proved that there was a one-to-one mapping between the electron density $\rho(\mathbf{r})$ and the external potential $\nu(\mathbf{r})$. This meant that the density uniquely determined the

external potential. They also proved that if given a trial density, the energy determined was an upper bound to the true energy.

The resulting energy expression given by the Hohenberg-Kohn (HK) theory was:

$$E_\nu[\rho] = \int \rho(\mathbf{r})\nu(\mathbf{r})d(\mathbf{r}) + F_{HK}[\rho] \tag{1.6}$$

The last expression is a sum of contributions from the kinetic and electron repulsion energies, respectively. Hohenberg and Kohn [8] showed that the above expression would be exact, but no expression for $F_{HK}[\rho]$ was given. Also not given was a practical method of calculating the energy or density from the above expression.

1.4 Kohn-Sham DFT

From HK theory, it was known that an exact form of the energy as a functional of the density existed. However, its form was not completely known. The TF, TFD, and TFDW models provided an approximate method to calculating the kinetic energy functional, however, these methods were still not suitable for calculations on atoms and molecules. In 1965, Kohn and Sham [9] developed a method to apply density functional theory to atomic and molecular systems. They looked at developing a more accurate form of the kinetic energy functional.

The exact form of the kinetic energy functional is given in Eq. 1.7 in terms of the natural orbitals ψ_i and their occupation numbers n_i. In this expression, the occupation

numbers take a value between zero and one.

$$T = \sum_{i}^{N} n_i \left\langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \right\rangle \tag{1.7}$$

For an interacting system of electrons, this expansion was cumbersome to calculate due to the large number of terms in the expansion. Kohn and Sham greatly simplified the calculation of the kinetic energy by introducing a system of noninteracting electrons which had the same density as the true interacting system. The exact kinetic energy T could then be written in two parts: one for this *reference* system of noninteracting electrons ($T_s[\rho]$) and another for the correction resulting from interactions between the electrons ($T[\rho]$ - $T_s[\rho]$). Consequently, the kinetic energy term representing the noninteracting gas can be written in terms of the wavefunction for the noninteracting system

$$\Psi_i = \frac{1}{\sqrt{N!}} det[\psi_1 \psi_2 \cdots \psi_N], \tag{1.8}$$

where the ψ_i in this case refer to the lowest energy eigenstates obtained from the following one-electron equations

$$[-1/2\nabla^2 + \nu_s(\mathbf{r})]\psi_i = \epsilon_i \psi_i. \tag{1.9}$$

For the noninteracting electron reference system, the density is exact and is given as

$$\rho(\mathbf{r}) = \sum_{i}^{N} \sum_{s} |\psi_i(\mathbf{r}, s)|^2. \tag{1.10}$$

The kinetic energy contribution in terms of the exact density ρ is

$$T_s[\rho] = \sum_{i}^{N} \left\langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \right\rangle. \tag{1.11}$$

This was the main idea behind the Kohn-Sham (KS) theory, and it led to the development of an orbital-based DFT method. In this method, the Kohn-Sham orbitals are those defined in Eq. 1.8.

Using this noninteracting system as a reference state for the real Hamiltonian, the resulting energy expression was:

$$E_{KS}[\rho] = \int \rho(\mathbf{r})\nu(\mathbf{r})d(\mathbf{r}) + T_s[\rho] + J[\rho] + (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - J[\rho]) \tag{1.12}$$

The first term represents the interaction of the density with the potential $\nu(\mathbf{r})$. The next two terms represent the kinetic energy of the reference system and the Coulomb repulsion energy of the electrons, respectively. The last two terms represent the correction to the kinetic energy due to interactions between electrons and the nonclassical electron-electron contributions, respectively [3].

To obtain the KS orbital equations analogous to the HF equations, one starts with

the KS one-electron Hamiltonian:

$$\hat{h}_{KS} = -\frac{1}{2}\nabla^2 + \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{1.13}$$

The above one-electron Hamiltonian is a Hermitian operator, with the electron density given by:

$$\rho(\mathbf{r}) = \sum_{i}^{N} |\psi_i(\mathbf{r})|^2 \tag{1.14}$$

The problem to solve is then to find the set of ψ that give the density which minimizes the energy functional $E_{KS}[\rho]$ subject to the constraints that the orbitals be orthonormal and integration of the total density give the number of electrons:

$$\delta E_{KS}[\rho] - \lambda \sum_{i}^{N} \sum_{j}^{N} \epsilon_{ij} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = 0 \tag{1.15}$$

Solving the above expression leads to the one-electron Kohn-Sham equations:

$$\hat{h}_{KS}\psi_i = \sum_{j}^{N} \epsilon_{ij}\psi_i \tag{1.16}$$

Since the above Hamiltonian is Hermitian, a unitary transformation can be used to diagonalize ϵ_{ij}, giving the canonical Kohn-Sham equations:

$$\hat{h}_{KS}\psi_i = \epsilon_i\psi_i \tag{1.17}$$

As written, the above equations have no known solution. One problem is how to express the orbitals in terms of functions. To do this, the orbitals can be expanded in a set of basis functions chosen to reproduce the properties of the orbital:

$$\psi_i(\mathbf{r}) = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu(\mathbf{r}) \tag{1.18}$$

Ideally the basis set ϕ_μ used in the above expression would be complete, but the calculations would not be feasible. However, one can use a finite set of basis functions to give a solution which is an approximation to the exact one obtained from a complete basis.

Another problem is that the KS equations depend on the orbitals which are the solutions to these equations. Thus, an iterative solution is needed to obtain the KS orbitals. One way to do this is to first recast the problem into a matrix eigenvalue equation and use the techniques of matrix algebra to iteratively solve the resulting equations. Doing this, and expanding the orbitals in terms of basis functions leads to the KS matrix equation:

$$\mathbf{H_{KS}C} = \mathbf{SC}\boldsymbol{\epsilon} \tag{1.19}$$

The **C** in the above equation is the matrix whose columns yield the expansion coefficients. The overlap matrix is denoted by **S** and the ϵ's are the energies of the KS orbitals.

One main advantage to the KS way of calculating the properties of the system was that the major part of the kinetic energy was known. However, the problem had the same dimensions as in HF theory instead of the desired three dimensional problem of the TF, TFD, and TFDW theories. Also, it was still not known how to obtain an exact expression for E_{XC}.

1.5 The Local Density and Local Spin Density Approximations

Kohn and Sham gave a method of calculating the energy from the density using a set of orbitals. However, they did not originally define the exchange-correlation (XC) energy explicitly. The many DFT methods available today differ in the way they define this XC energy. The first class of methods to be developed were called *local density approximation* (LDA) methods. The LDA methods used the uniform-electron-gas approximation for exchange, similar to that seen in the TFD method. The correlation energy was assumed to be locally approximated by the uniform-electron-gas model. The LDA definition for the XC energy is:

$$E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r})\epsilon_{XC}(\rho)d\mathbf{r} \tag{1.20}$$

From the above equation, the ϵ_{XC} term can be expanded into a term involving exchange and one involving correlation:

$$\epsilon_{XC}(\rho) = \epsilon_X(\rho) + \epsilon_C(\rho) \tag{1.21}$$

The exchange term in the above expression is the Dirac exchange as applied to the uniform-electron-gas model. An approximate form of the correlation term is given by Vosko, Wilk, and Nusair (VWN) [10]. Data for the parameterization of the functional was done by interpolation of the correlation energy data obtained from the quantum Monte Carlo calculations of Ceperly and Alder [11].

LDA was found to give reasonably accurate results for systems which closely resembled the uniform electron gas model. However, for atoms and molecules, the density does not vary as slowly as in the uniform electron gas model. Prior to the LDA, Slater developed a method which treated the nonlocal Fock operator of HF theory by a local one

$$\hat{F}_{X\alpha} = -\frac{1}{2}\nabla^2 + \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \nu_{X\alpha}(\mathbf{r}), \tag{1.22}$$

where

$$\nu_{X\alpha}(\mathbf{r}) = -\frac{3}{2}\alpha \left\{ \frac{3}{\pi}\rho(\mathbf{r}) \right\}^{1/3}, \tag{1.23}$$

and α is set to one. Comparison of Eq. 1.23 with Dirac's analogous exchange expression

$$\nu_{x,Dirac}(\mathbf{r}) = -\frac{3}{4}\left\{\frac{3}{\pi}\rho(\mathbf{r})\right\}^{1/3}. \tag{1.24}$$

shows they are the same, except that $\alpha = 2/3$ in the Dirac expression. Calculations using the above exchange and correlation expressions yields an ideal α value of 3/4, and gives more reasonable results for atoms and molecules. One of the most common LDA methods is SVWN, which combines Slater's exchange with VWN correlation [3].

Extension of the LDA to include spin dependency resulted in the *local spin density approximation*, or LSDA. The LDA used the total electron density $\rho(\mathbf{r}) = \rho^{\alpha}(\mathbf{r}) + \rho^{\beta}(\mathbf{r})$, where the $\rho^{\alpha}(\mathbf{r})$ and $\rho^{\beta}(\mathbf{r})$ denoted the densities of spin up and spin down electrons, respectively. The LSDA, based on the *spin-polarized uniform electron gas*, differentiated between the densities for spin-up and spin-down electrons. This polarization of the uniform electron gas became important when magnetic fields were present. It was also required for a more accurate physical description of atoms and molecules, especially for systems with unpaired electrons [3].

In the LSDA, the electronic portion of the energy functional in the absence of magnetic fields was:

$$E_{el}^{LSDA}[\rho] = T_s[\rho^{\alpha}, \rho^{\beta}] + J[\rho] + E_{XC}[\rho^{\alpha}, \rho^{\beta}] \tag{1.25}$$

Minimization of the energy functional by application of the variational principle resulted in two sets of equations:

$$\begin{aligned} \hat{h}^{\sigma}_{eff}\phi_{i\sigma}(\mathbf{r}) &= \epsilon^{\sigma}_{i}\phi^{\sigma}_{i}(\mathbf{r}) \text{ where } \sigma = \alpha, \beta \\ \hat{h}^{\sigma}_{eff} &= -\frac{1}{2}\nabla^2 + \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{XC}[\rho^{\alpha}, \rho^{\beta}]}{\delta\rho^{\sigma}(\mathbf{r})} \end{aligned} \tag{1.26}$$

The exchange-correlation functional, as in the LDA, can be divided into a part for exchange and a part for correlation. One popular form of the exchange functional was Slater's spin-polarized exchange [3]:

$$E_X^{LSDA}[\rho^{\alpha}, \rho^{\beta}] = 2^{1/3}\left(\frac{3}{4}\right)\left(\frac{3}{2}\right)\alpha\left[\frac{3}{\pi}\right]^{1/3}\int[(\rho^{\alpha})^{4/3} + (\rho^{\beta})^{4/3}]d\mathbf{r} \tag{1.27}$$

This form for the exchange functional was also one of the simplest forms and differed from Dirac's exchange only in the value of the parameter α.

The correlation energy functional was less straightforward to define. Since there are like spin and unlike spin electron interactions, there was no obvious way to separate the density according to the electron's spin. One way to define the correlation energy was in terms of the polarization [3]:

$$\zeta = \frac{(\rho^{\alpha} - \rho^{\beta})}{\rho} \tag{1.28}$$

Defining the Seitz radius ($r_s = ((3\pi)/(4\rho))^{1/3}$), the correlation energy was expressed as:

$$E_c^{LSDA}[\rho^\alpha, \rho^\beta] = \int \rho\epsilon_c(r_s, \zeta)d\mathbf{r} \tag{1.29}$$

The term $\epsilon_c(r_s, \zeta)$ represents the correlation energy per electron. There is no exact form for this term. Originally, the most popular approximation to it was that done by Vosko, Wilk, and Nusair, and details of the functional and its derivation can be found elsewhere [10]. Combining this method and Slater's exchange gave the popular LSDA method *SVWN*.

The LSDA gave more accurate results for atoms and molecules than the LDA. However, it was still based on a local approximation to a nonlocal density. Also, it did not give correct behavior of the exchange-correlation energy functional far from the nucleus. Consequently, dissociation energies calculated using LSDA were poor.

1.6 The Generalized Gradient Approximation

It was thought that the problem with the LSDA was in the exchange term of the exchange-correlation functional. In HF theory, exchange is exactly described by a nonlocal functional. However, the LSDA uses a local functional for exchange. The functional contribution due to correlation was also local, however, exchange contributions were larger [3, 11].

In general, one can express the exchange-correlation energy functional as follows:

$$E_{XC}^{GGA}[\rho^{\alpha},\rho^{\beta}] = \int \epsilon_{XC}(\rho^{\alpha},\rho^{\beta},\nabla\rho^{\alpha},\nabla\rho^{\beta})d\mathbf{r} \tag{1.30}$$

The gradient terms in the above equation now allows one to account for inhomogeneity in the electron density. Called the *generalized gradient approximation* (GGA), it was found to give better results for atoms and molecules.

Initial studies of the exchange-correlation contribution focused on the exchange term. The general form for the gradient-corrected exchange functional is:

$$E_{X}^{GGA}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int \rho^{4/3} f(s)d\mathbf{r} \tag{1.31}$$

If $f(s) = 1$ in the above equation, then the LSDA is obtained.

There have been several studies on the form of $E_{X}^{GGA}[\rho]$. One of the more popular forms is given by Becke [12–14]. This functional was derived starting from the LSDA approximation for local exchange. Gradient corrections which are first order with respect to the density were added to simulate some nonlocal behavior in the exchange functional. The functional was called *B88*:

$$E_{X}^{B88}[\rho] = E_{X}^{LDA}[\rho] - \beta\sum_{\sigma}\int (\rho^{\sigma})^{4/3} f(x_{\sigma})d\mathbf{r}, \tag{1.32}$$

where

$$f(x_\sigma) = \frac{x_\sigma^2}{1 + 6\beta x_\sigma sinh^{-1} x_\sigma}$$
$$x_\sigma = \frac{|\nabla \rho^\sigma|}{(\rho^\sigma)^{4/3}} \tag{1.33}$$

The σ in the above equation was defined previously. The parameter β was obtained from a fit to exact atomic HF data. Becke showed that the above functional form accurately reproduced exchange energies from HF calculations of noble gas atoms. It also performed well reproducing the exchange energies of atoms H through Ar and transition metals Sc through Zn.

Gradient corrections have also been applied to the correlation energy functional. One popular form is the *P86* functional [15]:

$$E_C^{P86}[\rho^\alpha, \rho^\beta] = \int \rho \epsilon_C(\rho^\alpha, \rho^\beta) d\mathbf{r} + \int d^{-1} e^{-\phi} C(\rho) \frac{|\nabla \rho|^2}{\rho^{4/3}} d\mathbf{r}$$
$$\phi = 1.745 f \frac{C(\infty)}{C(\rho)} \frac{|\nabla \rho|}{\rho^{7/6}} \tag{1.34}$$

ϵ_C in the above equation was determined through a parameterization of the results of Ceperly and Alder [11] and f was chosen so the correlation energy was the same as that of the neon atom. Perdew showed that the correlation energies of a small selection of atoms and ions was improved as compared to the LSDA.

Another popular functional for the correlation energy is that of Lee, Yang, and Parr (LYP) [16]. This functional is built from the second-order gradient expansion of the

density:

$$
\begin{aligned}
E_C^{LYP}[\rho(\mathbf{r})] &= -a \int \frac{\gamma(\mathbf{r})}{1 + d\rho(\mathbf{r})^{-1/3}} \left[\rho(\mathbf{r}) + 2b\rho(\mathbf{r})^{-5/3} f(\rho(\mathbf{r})) e^{-c\rho(\mathbf{r})^{-1/3}}\right] d\mathbf{r} \\
f(\rho(\mathbf{r})) &= 2^{2/3} C_F (\rho(\mathbf{r})^{\alpha})^{8/3} + 2^{2/3} C_F (\rho(\mathbf{r})\beta)^{8/3} - \rho(\mathbf{r}) t_W(\mathbf{r}) + \\
&\quad (1/9)(\rho(\mathbf{r})^{\alpha} t_W(\mathbf{r})^{\alpha} + \rho(\mathbf{r})^{\beta} t_W(\mathbf{r})^{\beta}) + \\
&\quad (1/18)(\rho(\mathbf{r})^{\alpha} \nabla^2 \rho(\mathbf{r})^{\alpha} + \rho(\mathbf{r})^{\beta} \nabla^2 \rho(\mathbf{r})^{\beta}) \\
t_W(\mathbf{r}) &= \frac{1}{8} \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} - \frac{1}{8} \nabla^2 \rho
\end{aligned}
\tag{1.35}
$$

The constants a, b, c, and d were obtained via a fit to the correlation energy for a helium atom with a HF orbital. Lee, Yang, and Parr showed that for a select group of atoms, molecules, and ions agreement with experiment was reasonable. The worst cases were Be^{2+}, H_2O, and Li in the ^2S excited state. Here, the errors in the correlation energy were greater than 10 %.

Miehlich et al. [17] compared the VWN functional with a series of GGA functionals. They tested the ionization energies, electron affinities, dissociation energies, and atomization energies of a series of atoms and molecules. Among the GGA functionals tested were B88, P86, and LYP. In general, the authors found that the gradient corrections were a marked improvement over the LSDA. Also, trends with all functionals were the same, although some performed better than others. For all the functionals, the largest errors occurred with the ionization energies. The performance of the LYP functional was similar to P86 and little or no improvement was obtained in using one functional over the other.

1.7 Meta-GGA Methods

Current GGA's introduce inhomogeneity into the density through the use of density gradients. This inhomogeneity improves the accuracy of these functionals over the LSDA. However, some improvement to their accuracy is still needed. One possibility is to include the kinetic energy density in the exchange-correlation energy functional. In two previous works, Becke showed that the kinetic energy density plays a role in the description of the exchange and correlation energies [18, 19].

There are several new *meta-GGA* functionals currently available which are based on the kinetic energy density dependency in the exchange-correlation functional [20–23]. They all have the general form

$$E_{XC}^{meta-GGA}[\rho^{\alpha},\rho^{\beta}] = \int \rho\epsilon_{XC}(\rho^{\alpha},\rho^{\beta},\nabla\rho^{\alpha},\nabla\rho^{\beta},\tau^{\alpha},\tau^{\beta})d\mathbf{r} \tag{1.36}$$

In the above expression, the τ^{α} and τ^{β} are the kinetic energy densities for α and β spin electrons. The general expression for τ is

$$\tau^{\sigma}(\mathbf{r}) = \frac{1}{2}\sum_{i}^{occ}|\nabla^2\psi_{i,\sigma}(\mathbf{r})|^2, \tag{1.37}$$

where $\sigma = \alpha$ or β. These functionals are relatively new and their accuracy for a variety of different systems is still being explored.

An example of a meta-GGA is the functional of Van Voorhis and Scuseria (VSXC)

[20]:

$$
\begin{aligned}
E_X^{VSXC}[\rho^\alpha, \rho^\beta] &= \sum_\sigma \int \rho^\sigma \epsilon_x^{\sigma,LDA} f(x^\sigma, z^\sigma) d\mathbf{r} \\
E_C^{VSXC} &= \int f^{\alpha,\beta}(x,z) e_C^{\alpha,\beta}, LDA d\mathbf{r} + \int f^{\alpha,\alpha}(x^\alpha, z^\alpha) D^\alpha e_C^{\alpha,\alpha}, LDA d\mathbf{r} + \\
&\quad \int f^{\beta,\beta}(x^\beta, z^\beta) D^\beta e_C^{\beta,\beta}, LDA d\mathbf{r} \\
f(x^\sigma, z^\sigma) &= \left(\frac{a}{\gamma^\sigma(x,y)} + \frac{b(x^\sigma)^2 + cz^\sigma}{(\gamma^\sigma(x,y))^2} + \frac{d(x^\sigma)^4 + e(x^\sigma)^2 z^\sigma + f(z^\sigma)^2}{(\gamma^\sigma(x,y))^3} \right) \qquad (1.38) \\
\gamma(x,z) &= 1 + \alpha(x^2 + z) \\
D^\sigma &= 1 - \frac{(x^\sigma)^2}{4(z^\sigma + C_F)}, x^2 = (x^\alpha)^2 + (x^\beta)^2, z = z^\alpha + z^\beta \\
x &= \frac{|\nabla\rho|}{\rho^{4/3}}, z = \frac{\tau}{\rho^{5/3}} - C_F
\end{aligned}
$$

C_F is the Fermi constant and the parameters in the above equations are obtained by fitting to a large dataset of atoms and molecules using a 6-311+g(3df,2p) basis set.

Performance of the meta-GGA's is promising, in particular the VSXC functional. Results indicate that VSXC gives good geometries and vibrational frequencies, and results are comparable to the popular hybrid functional B3LYP (described in the following section) [20, 24].

1.8 Hybrid Density Functional Methods

The functionals discussed so far do not include exchange exactly, as is done in HF theory. This *exact* exchange could theoretically be added to functionals to improve their description of the exchange energy. This was the original idea of Becke [25]. This added

exchange term was obtained by examination of the formula which connects a noninteracting system to a fully interacting system. Such a formula was called the *adiabatic connection* [11, 25, 26]:

$$E_{XC}[\rho] = \int_0^1 U_{XC}^{\lambda}[\rho] d\lambda \tag{1.39}$$

The λ in the above equation has a different value depending on whether the system is noninteracting ($\lambda = 0$) or fully interacting ($\lambda = 1$), and a connection between them is established by varying λ. In the case of a noninteracting system, there is no dynamic correlation, so the exchange correlation potential represents exact exchange.

Originally, Becke proposed an exchange-correlation functional that used half exact exchange and half exchange for the fully interacting system [25]. However, this functional did not reproduce well the total energies of the G1 dataset, so a new mixing scheme, called *B3*, was developed [26]:

$$E_{XC}[\rho] = E_{XC}^{LSDA}[\rho] + a_0(E_X^{exact}[\rho] - E_X^{LSDA}[\rho]) + a_X \Delta E_X^{B88}[\rho] + a_C \Delta E_C[\rho] \tag{1.40}$$

The ΔE's refer to the gradient corrections. One popular functional is B3LYP, which uses the above definition for the XC functional with LYP correlation corrections. The three a values were determined by fitting the energy functional results to experimental data. Results of the B3 functional with the PW91 correlation correction indicated Becke's functional performed well for total energies and ionization energies for a series of atoms

and molecules from the G1 dataset.

1.9 Conclusion

Through the use of the density, DFT calculates the properties of the system without the use of the N-electron wavefunction. With the advent of the KS formalism, the procedure for obtaining these properties is straightforward. DFT thus offers some advantages to HF and post-HF methods, the most attractive being the time savings while still including some electron correlation. As a result, it has become a popular alternative to the conventional HF and post-HF methods, especially for large systems.

Chapter 2

Linear Scaling DFT

2.1 Introduction

Within DFT, evaluation of the energy requires the solution of the KS equations. To accomplish this, one must evaluate the two-electron Coulomb repulsion integrals. These are four-centered integrals, and traditionally they were computed directly. This operation scaled as $O(N^4)$ for N basis functions. That is, for an increase of two in the number of basis functions, the computational effort to generate these integrals increases by a factor of 16. As a result, this problem has been the bottleneck of calculations in both DFT and HF-based calculations.

In the last two decades, much progress has been made in reducing the scaling for the computation of the two-electron integrals [27–36]. Through the use of both Gaussian and auxiliary basis functions, the number of integrals that must be calculated, and hence the scaling, has been reduced. Also, since the values of some of the integrals will be negligible,

there is no need to evaluate them. By *prescreening* the integrals, those which make a negligible contribution to the Kohn-Sham matrix are not computed [37]. However, even with prescreening, the computation of the two-electron integrals still scales quadratically.

2.2 Definitions

2.2.1 The Coulomb Integral

From classical physics, the Coulomb repulsion between two charge distributions varies as the inverse distance between them [2, 38]:

$$E_{ee}^{Coul} = \iint \rho(\mathbf{r_1}) \frac{1}{|\mathbf{r_1} - \mathbf{r_2}|} \rho(\mathbf{r_2}) d\mathbf{r_1} d\mathbf{r_2} \tag{2.1}$$

where $\rho(\mathbf{r_1})$ and $\rho(\mathbf{r_2})$ represent the two charge distributions [38]. The charge distribution $\rho(\mathbf{r})$ represents the probability of finding an electron in the volume of space about $\mathbf{r}$ and can be expressed in terms of the occupied KS orbitals ψ_i

$$\rho(\mathbf{r}) = \sum_{i}^{N} |\psi_i(\mathbf{r})|^2. \tag{2.2}$$

Substitution of Eq. 2.2 into Eq. 2.1 gives

$$E_{ee}^{Coul} = \sum_{ijkl} \iint \psi_i^*(\mathbf{r_1}) \psi_j(\mathbf{r_1}) \frac{1}{|\mathbf{r_1} - \mathbf{r_2}|} \psi_k^*(\mathbf{r_2}) \psi_l(\mathbf{r_2}) d\mathbf{r_1} d\mathbf{r_2}. \tag{2.3}$$

Eq. 2.3 shows the energy resulting from the Coulomb repulsion between the charge densities at $\mathbf{r}_1$ and $\mathbf{r}_2$ in terms of the occupied KS orbitals. To obtain this energy contribution, the set of integrals

$$(ij|kl) = \iint \psi_i^*(\mathbf{r_1})\psi_j(\mathbf{r_1})\frac{1}{|\mathbf{r_1} - \mathbf{r_2}|}\psi_k^*(\mathbf{r_2})\psi_l(\mathbf{r_2})d\mathbf{r_1}d\mathbf{r_2} \tag{2.4}$$

must be evaluated. These are the *Coulomb repulsion integrals.*

2.2.2 Basis Functions

In DFT (and HF) calculations, the spatial orbitals are expanded in a set of K one-electron functions [2,39]:

$$\psi_i^*(\mathbf{r}) = \sum_a^K C_{ai}\phi_a(\mathbf{r}) \tag{2.5}$$

These *basis functions* are combined to reproduce the behavior of the orbital they represent. Sets of these functions are called *basis sets.*

Any type of function can be chosen for the above expansion, however popular choices are Slater, Gaussian, and Hermite functions. Slater functions give the best results when used, however evaluation of integrals using them is cumbersome. For practical calculations, Gaussian functions are chosen. Although these functions are not as good at reproducing orbital behavior, some of their properties make them ideal for use in calculations. Orbitals

expanded with a set of Gaussian functions are called *Gaussian-type orbitals*, or GTO's [2, 39].

Basis functions are most often represented by *contracted* Gaussians. These are linear combinations of primitive (or uncontracted) Gaussians [2]

$$\phi_a(\mathbf{r}) = \sum_{k=1}^{n_k} D_{ka}\phi_k(\mathbf{r}), \tag{2.6}$$

where the D_{ka} are the contraction coefficients and the $\phi_k(\mathbf{r})$ are the primitive Gaussians.

In DFT, and more recently in HF theory, a set of *auxiliary basis functions* has been used along with the orbital basis functions. Consider an approximate density that is expressed in terms of a set of auxiliary basis functions as

$$\tilde{\rho}(\mathbf{r}) = \sum_{i}^{K} c_i\tilde{\chi}_i(\mathbf{r}). \tag{2.7}$$

Ref. 40 gives the procedure for generating the fitted density. Within this procedure, consider the difference between the fitted and real densities

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \tilde{\rho}(\mathbf{r}). \tag{2.8}$$

The Coulomb repulsion energy can be expressed as

$$E^{coul}[\rho] = 2\iint \frac{\rho(\mathbf{r})\tilde{\rho}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' - \iint \frac{\tilde{\rho}(\mathbf{r})\tilde{\rho}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + \frac{1}{2}\iint \frac{\Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}', \tag{2.9}$$

where the last term is the Coulomb repulsion energy between the error in the fit and itself. The fitting procedure's goal is to minimize this last term. In so doing, the fitting procedure is variational [40].

The auxiliary basis functions are atom-centered and are grouped into sets which share a common exponent. These functions have a few advantages for integral evaluation. First, a small set of them is used to fit the density, and this set is much smaller than the set of products of Gaussian orbital basis functions used to describe the density. As a result, fewer integrals need to be evaluated. Also, because of the shared exponents and the fact that they are atom-centered, groups of them can be evaluated at once and the storage of the auxiliary basis functions for each center requires less memory. This, it turns out, reduces the computational complexity in implementing the integral code, which also reduces computation time.

2.2.3 The Gaussian Product Theorem

An advantage to the use of Gaussians as basis functions is that the product of two Gaussian functions produces a third Gaussian function [2, 41].
This function, shown in Figure 2.1 with center **P**, has a center which lies on the line joining the first two Gaussians, centered on **A** and **B** in the same Figure. Also, the exponent of the third Gaussian is obtained from the first two. As an example, consider two unnormalized

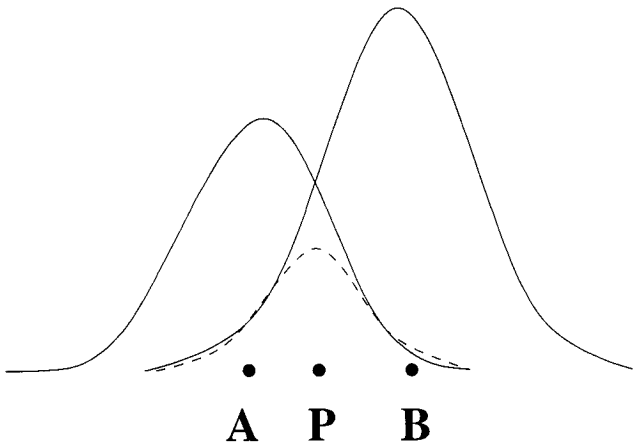


Figure 2.1: Diagram showing the function resulting from the product of two Gaussian functions.

s-type Gaussians:

$$\begin{aligned} \phi_A(\mathbf{r}-\mathbf{R_A}) &= e^{-\alpha r_A^2} \\ \phi_B(\mathbf{r}-\mathbf{R_B}) &= e^{-\beta r_B^2} \end{aligned} \tag{2.10}$$

the Gaussian product can be obtained from the following generating formula [2, 41]:

$$\begin{aligned} \phi_P &= K(\alpha,\beta,\mathbf{A},\mathbf{B})e^{-\gamma \mathbf{r_P}^2} \\ K(\alpha,\beta,\mathbf{A},\mathbf{B}) &= 2^{1/2}\frac{\pi^{5/4}}{\alpha+\beta}e^{-\gamma(\mathbf{A}-\mathbf{B})^2} \\ \gamma &= \frac{\alpha\beta}{\alpha+\beta} \\ \mathbf{P} &= \frac{\alpha\mathbf{A}+\beta\mathbf{B}}{\alpha+\beta} \end{aligned} \tag{2.11}$$

The number of Gaussian products scales as $O(N^2)$, which means the integral evaluation still scales as $O(N^4)$. However, a number of these products are negligible in size and thus do not contribute. Closer inspection of the resulting Gaussian product at $\mathbf{P}$ shows

that its size depends on the distance between the two Gaussians centered at **A** and **B**. As a result, the farther apart the centers **A** and **B** are, the smaller the Gaussian is at **P**. So, this criterion can be used to determine whether an integral contribution is large enough to justify its calculation. Also known as *shellpair screening*, this method reduces the scaling of the integral evaluation to $O(N^2)$.

2.3 Conventional Coulomb Integral Evaluation - Theory

2.3.1 The Obara-Saika Method

A popular method that makes use of the properties of Gaussian functions for two-electron integrals was developed by Obara and Saika [27]. Through prescreening techniques made possible through the use of Gaussian functions, they were able to develop straightforward formulae for the direct computation of the electron repulsion integrals. They were also able to reduce the scaling of the Coulomb integral evaluation from $O(N^4)$ to $O(N^2)$ by exploiting properties of Gaussian-type functions and prescreening techniques. They also demonstrated that integrals of higher angular momentum could be obtained from those of lower angular momentum. As a result, fewer integrals need to be evaluated explicitly.

The method of Obara and Saika (OS) uses a set of recurrence relations to obtain integrals. From these formulae, one need only obtain a small subset of integrals directly. For these integrals, there are simple analytical expressions for their solution. Obara and

Saika showed that to do this, one must introduce a set of integrals which are not kept, but are used to obtain the desired integrals. They called these *auxiliary* integrals. They were able to show that despite the additional computational effort needed to obtain the auxiliary integrals, the overall computational effort was reduced [27].

Obara and Saika also noted that Cartesian Gaussian-type functions obey

$$\frac{\partial}{\partial R_i}\phi(r;\alpha,n,R) = 2\alpha\phi(r;\alpha,n+\mathbf{1_i},R) - n_i\phi(r;\alpha,n-\mathbf{1_i},R)\ i=x,y,z. \tag{2.12}$$

Rearrangement of the above equation gives a relationship between functions of higher and lower angular momentum

$$\phi(r;\alpha,n+\mathbf{1_i},R) = \frac{1}{2\alpha}\left(\frac{\partial}{\partial R_i}\phi(r;\alpha,n,R) + n_i\phi(r;\alpha,n-\mathbf{1_i},R)\right)\ i=x,y,z. \tag{2.13}$$

From the above relationship, the following generating formula for the Coulomb integrals for arbitrary angular momentum was obtained

$$\begin{aligned}
[(\mathbf{a}+\mathbf{1_i})\mathbf{b},\mathbf{cd}]^{(m)} &= (P_i - A_i)(\mathbf{ab},\mathbf{cd})^{(m)} + (W_i - P_i)(\mathbf{ab},\mathbf{cd})^{(m+1)} \\
&+ \frac{1}{2\alpha_{ab}}N_i(\mathbf{a})\left\{[(\mathbf{a}-\mathbf{1_i})\mathbf{b},\mathbf{cd}]^{(m)} - \frac{\rho}{\alpha_{ab}}[(\mathbf{a}-\mathbf{1_i})\mathbf{b},\mathbf{cd}]^{(m+1)}\right\} \\
&+ \frac{1}{2\alpha_{ab}}N_i(\mathbf{b})\left\{[\mathbf{a}(\mathbf{b}-\mathbf{1_i}),\mathbf{cd}]^{(m)} - \frac{\rho}{\alpha_{ab}}[\mathbf{a}(\mathbf{b}-\mathbf{1_i}),\mathbf{cd}]^{(m+1)}\right\} \\
&+ \frac{1}{2(\alpha_{ab}+\alpha_{cd})}N_i(\mathbf{c})[\mathbf{ab},(\mathbf{c}-\mathbf{1_i})\mathbf{d}]^{(m+1)} \\
&+ \frac{1}{2(\alpha_{ab}+\alpha_{cd})}N_i(\mathbf{d})[\mathbf{ab},\mathbf{c}(\mathbf{d}-\mathbf{1_i})]^{(m+1)}\ i=x,y,z,
\end{aligned} \tag{2.14}$$

where the following relationships were used

$$\begin{aligned}
\mathbf{1_i} &= (\delta_{ix}, \delta_{iy}, \delta_{iz}) \\
N_i(\mathbf{a}) &= a_i \\
\alpha_{ab} &= \alpha_a + \alpha_b \\
\rho &= \frac{\alpha_{ab}\alpha_{cd}}{\alpha_{ab} + \alpha_{cd}} \\
\mathbf{P} &= \frac{\alpha_a \mathbf{A} + \alpha_b \mathbf{B}}{\alpha_a + \alpha_b} \\
\mathbf{Q} &= \frac{\alpha_c \mathbf{C} + \alpha_d \mathbf{D}}{\alpha_c + \alpha_d} \\
\mathbf{W} &= \frac{\alpha_{ab}\mathbf{P} + \alpha_{cd}\mathbf{Q}}{\alpha_{ab} + \alpha_{cd}}.
\end{aligned} \tag{2.15}$$

In the above equation, the auxiliary integrals are the set $(\mathbf{ab}, \mathbf{cd})^{(m)}$, for integer values of $m \geq 0$ [27].

As an example, consider the integral $(\mathbf{ps}, \mathbf{ss})$. There are three contributions, depending on whether the angular momentum is in the x, y, or z-direction. From Eq. 2.14, we get the following relationship for the integral $(p_i s, ss)$

$$(p_i s, ss) = (p_i s, ss)^{(0)} = (P_i - A_i)(ss, ss)^{(0)} + (W_i - P_i)(ss, ss)^{(1)} \; i = x, y, z. \tag{2.16}$$

In the above set of three equations, the integrals $(ss, ss)^{(m)}$ can be obtained from the

analytical expression

$$(ss,ss)^{(m)} = \frac{1}{\alpha_{ab}+\alpha_{cd}} K(\alpha_a,\alpha_b,\mathbf{A},\mathbf{B}) K(\alpha_c,\alpha_d,\mathbf{C},\mathbf{D}) F_0(T)$$
$$F_m(T) = \int_0^1 t^{2m} e^{-Tt^2} dt, \tag{2.17}$$

and the expression for the K's are given in Eq.2.11. Details for obtaining $F_m(T)$ can be found in ref.27.

2.3.2 The Head-Gordon-Pople Method

Obara and Saika devised a method to obtain integrals of higher angular momentum from those of lower angular momentum. The expressions used to do this were called *vertical recurrence relations.* Head-Gordon and Pople [28] devised another set of expressions to shift angular momentum from one center to a second in the Gaussian product. They called these expressions *horizontal recurrence relations.*

Head-Gordon and Pople proposed the following relationship between angular momenta on the two centers of a Gaussian product

$$[\mathbf{a}(\mathbf{b}+\mathbf{1_i}),\mathbf{cd}]^{(m)} = [(\mathbf{a}+\mathbf{1_i})\mathbf{b},\mathbf{cd}]^{(m)} + (A_i - B_i)[\mathbf{ab},\mathbf{cd}]^{(m)}. \tag{2.18}$$

In the above expression, the two integrals on the right side can be obtained from the Obara-Saika vertical recurrence relations. Head-Gordon and Pople also showed that the

following expression can be applied to contracted functions

$$(\mathbf{a}(\mathbf{b}+\mathbf{1_i}),\mathbf{cd})^{(m)} = ((\mathbf{a}+\mathbf{1_i})\mathbf{b},\mathbf{cd})^{(m)} + (A_i - B_i)(\mathbf{ab},\mathbf{cd})^{(m)}, \tag{2.19}$$

where the round brackets denote integrals over contracted functions [28].

The Head-Gordon-Pople (HGP) algorithm had three steps. Given an integral class (eg. (pp,ss)), determine the integral of the form (as,cs) closest to the desired integral using the horizontal recurrence relations given in Eq. 2.19. Note that these recurrence relations are evaluated over contracted basis functions, not primitives. Second, relate this integral to the set of uncontracted integrals using

$$(\mathbf{ab},\mathbf{cd})^{(m)} = \sum_i^{N_i}\sum_j^{N_j}\sum_k^{N_k}\sum_l^{N_l} D_{ai}D_{bj}D_{ck}D_{dl}[\mathbf{a_i b_j},\mathbf{c_k d_l}]^{(m)}, \tag{2.20}$$

where the D's are the contraction coefficients. Finally, use the OS vertical recurrence relations in Eq. 2.14 to obtain a set of integrals of the form $[ss,ss]^{(m)}$. These are then solved analytically.

The HGP algorithm just described starts from a desired integral over contracted basis functions and relates this integral to the primitive auxiliary integrals $[ss,ss]^{(m)}$. When the algorithm is executed, it is done in the reverse direction. That is, the primitive auxiliary basis functions are evaluated first, the OS vertical recurrence relations are applied, the resulting integrals are contracted, and the horizontal recurrence relations are applied to obtain the desired integral.

The above algorithm also reduces the number of integrals to be computed directly. Fewer evaluations are done as some of the expressions are evaluated over contracted integrals. Compared to evaluating all integrals over primitives, this amounts to a significant savings in the overall algorithm.

A modification of the HGP algorithm was given by Köster [32] for integrals involving auxiliary basis functions as well as Gaussian products. Essentially, the formulae used are very similar to those of Obara and Saika, and Head-Gordon and Pople. They can be obtained from the relationships of OS and HGP by assigning an exponent of zero to the fourth basis function.

2.4 The Continuous Fast Multipole Method

So far, methods have been discussed which evaluate the Coulomb repulsion integral directly. However, for computations on large systems, these methods are not feasible. Several groups have worked in the last decade on approximations which when applied can yield linear scaling. The most popular methods [42–49] are based on the *divide-and-conquer* approach. Another proposed scheme involves application of a Fourier transform to approximate the Coulomb energy [50]. All of the approaches yield linear (or near-linear) scaling. They are all based on the same idea: one can partition the integrals into far-field and near-field contributions. The far-field contributions can be evaluated using an approximate method derived from classical physics while the near-field interactions must be evaluated using the previous mentioned algorithms.

2.4.1 The Classical Coulomb Interaction

From classical physics, the potential acting on a charge q_2 due to another distant charge q_1 is given by [38]

$$V(\mathbf{r_2}) = k\frac{q_1}{|\mathbf{r_2} - \mathbf{r_1}|}, \tag{2.21}$$

where r_1 is the position of the charge q_1 and r_2 is the position of the distant charge q_2. The definition of k depends on the system of units used. For the cgs system, k=1, whereas for the SI system of units, k= $1/4\pi\epsilon_0$.

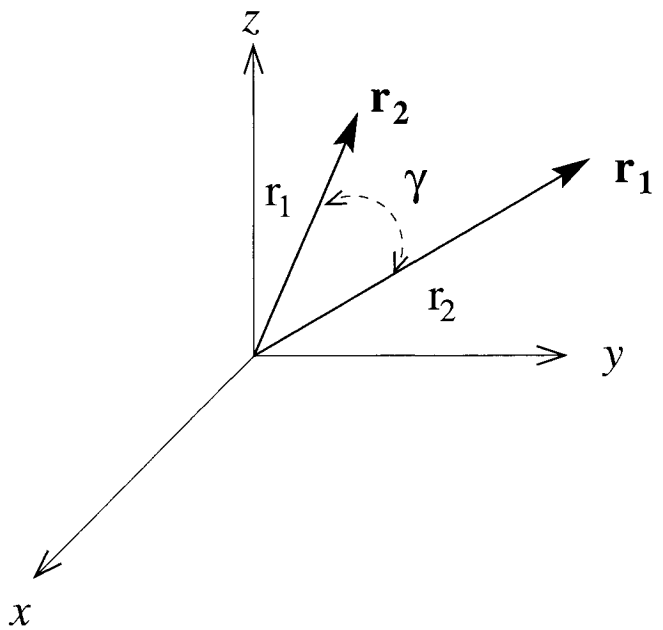


Figure 2.2: A diagram of the interaction between two point charges, one at r_1 and one at r_2.

The interaction described by Eq. 2.21 is displayed in Figure 2.2, and represents a statement of Coulomb's law for point charges. In the case of more than one distant charge, it can be

shown that the potential at a point r_2 due to all charges is a vector sum of the potentials due to each individual charge. The Coulomb energy then becomes an interaction of the charge at r_2 with the potential $V(r_2)$.

If there is no dielectric present, the above potential for the case of many point charges can be obtained from the solutions to the Laplace equation [38]

$$\nabla^2 V = 0. \tag{2.22}$$

The solutions give, in spherical harmonics, an approximation for $|r_{12}|^{-1}$:

$$\begin{aligned} \frac{1}{|r_{12}|} &= 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{1}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}^*(\theta_<, \phi_<) Y_{lm}(\theta_>, \phi_>) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{(l+|m|)!}{(l-|m|)!} \frac{r_<^l}{r_>^{l+1}} P_{lm}(cos\theta_<) P_{lm}(cos\theta_>) e^{-im(\phi_< - \phi_>)}, \end{aligned} \tag{2.23}$$

where the $r_<$ and $r_>$ represent the smaller and larger of the distances r_1 and r_2, respectively. Some of the properties of this expansion are:

- if the two point charges are far from one another, this approximation is accurate
- as the two point charges approach one another, the approximation becomes less accurate
- if the two point charges both lie on the radius of a circle, i.e. $r_1 = r_2$, the series diverges
- the resulting *associated Legendre functions* P_{lm} have a straightforward generating

formula.

Using the approximation given in Eq.2.23 allows one to exploit the application of *divide-and-conquer* techniques to calculate the Coulomb energy. This was the idea originally proposed by Greengard and Rokhlin [51, 52].

The method of Greengard and Rokhlin (GR) was one of a few of the early methods to be developed that combined the *multipole expansion* with a tree data structure. Their algorithm treated the potential energy of two point charges as a multipole expansion. They then used the divide-and-conquer scheme to cluster point charges. They used the idea that multipole expansions can be summed together into one expansion for groups of point charges. As a result, a single expansion could be used to represent the potential felt by a point charge due to everything sufficiently separated from it. This is in contrast to the number of pairwise interactions needed to describe the same set of interactions [51,52].

For example, consider the interaction of two groups of 50 point charges at a large distance from each other. The number of pairwise interactions is 100^2 =10 000. Using GR's algorithm, this interaction has three parts. First, the interaction of the first group of 50 point charges with the expansion of the second group of 50 is 1x50. Second, the interaction of the second group of 50 point charges with the expansion of the first group is also 1x50. Finally, the pairwise interactions within the two groups of 50 point charges is $2(50^2)$ This gives, for the GR algorithm, a total of 5100 steps.

The GR method provided savings for calculating pairwise interactions for large systems. It had also been shown to be an effective linear scaling method. However, the algorithm showed the best scaling for uniform systems. Also, the method was originally

developed for point charge systems and the subdivisions used to build the tree were done in two dimensions. Interest began to develop in how one could apply this method to the Coulomb integrals of molecular systems. The problem was that molecules are not uniform systems and the interactions are between charge distributions, not point charges. As a result, this method was not ideal for use in quantum chemistry calculations [51, 52].

In the last decade, two classes of methods have become increasingly popular for calculating the Coulomb interactions seen in quantum chemistry calculations. The methods have solved the problem of applying the above to charge distributions. They also work well for nonuniform distribution of charges, like those seen in molecules. Both methods use tree data structures to cluster particles. The first class uses the tree to group distributions. The potential energy is then computed using multipoles without the use of additional intermediate steps [45, 46, 49]. The second class uses *translation operators* to share information between groups in the tree [42–44, 47, 48].

The first class has two methods. They are the *Recursive Bisection Method* (RBM) of Pérez-Jordá and Wang [45, 46] and the *Quantum Chemical Tree Code* (QCTC) of Challacombe, Schwegler, and Almlöf [49]. Both use a modified Barnes-Hut tree. A Barnes-Hut tree is formed when space is divided until only one charge distribution remains in the resulting box. Unlike the second class of methods, neither method uses translation operators nor do they pass information through the tree. The boxes are subdivided and tested to see if they are well-separated. If this condition is met, the Coulomb energy is calculated between distributions in one box and the potential of the second box using a multipolar expansion of this energy.

Despite the similarities, they differ in several respects. First, the RBM uses a one-dimensional tree, whereas the QCTC uses a three-dimensional tree. Also, the RBM uses a well-separated index based on the box size and the size of the distributions in it. The QCTC uses a multipole acceptance criterion (MAC) that depends on the multipole expansion elements and the distance between the two boxes. This variable has a similar function to the well-separated criterion of the RBM. Another difference is with the method of evaluating the expansion between well-separated boxes. Although the QCTC does not translate information through the tree, expansions due to boxes whose expansions meet the MAC are translated between one another before the integral contributions are evaluated. The RBM does not use any translation operators. Finally, the QCTC uses Hermite GTFs which are obtained via transformation from Cartesian GTFs. The RBM only uses Cartesian GTFs.

The second class of tree-based methods is bassed on an extension of the fast multipole method of GR to handle charge distributions. The two most popular methods are the *Gaussian very fast multipole method* (GvFMM) of Strain, Scuseria, and Frisch [47,48] and the *Continuous Fast Multipole Method* (CFMM) of White and Head-Gordon [42–44]. These two methods are remarkably similar in every respect. The largest differences between them occur in the expression for the extent of a distribution and in the well-separated criterion.

All of the above mentioned methods have advantages and disadvantages. The RBM is simple to implement and is designed for one-dimensional systems. The scaling is thus much worse for two and three-dimensional systems. The QCTC method is also conceptually simple and is applicable to three-dimensional systems, but requires an extra transformation from Cartesian to Hermite GTFs. The GvFMM and CFMM are more complex

to implement, but use the tree data structure to further group information and can easily be implemented in one, two, and three dimensions. Due to its general applicability, the CFMM was chosen for this work.

The CFMM is a modification of the GR algorithm for a three-dimensional system of distributions which are nonuniformly distributed in space. Also, the translation operators were recast in spherical coordinates. The method also employs a formula for the well-separated criterion that is applied to each distribution, and can therefore have varying values. As with the GR algorithm, the CFMM is applied to reduce the number of integral evaluations done directly. This reduction is made possible through the use of clustering techniques. This clustering is achieved by subdividing the simulation space to group distributions. From these groups, the potential due to each distribution is expanded in multipole moments and these expansions are summed. This resulting expansion is then interacted with individual distributions. This represents the sum of pairwise interactions between the single distribution and all others a minimum distance from it. This is done for all distributions once [42–44].

In the next section, a few terms commonly encountered in the CFMM are explained in some detail.

2.4.2 A Few Definitions

If the expansion in Eq.2.23 is substituted into Eq.2.1, the resulting integral expression is

$$(ij|kl) = \int \rho(\mathbf{r_1}) \left[\sum_{l=0}^{\infty} \sum_{m=-l}^{l} O_{lm}(\mathbf{r_2}) M_{lm}(\mathbf{r_1}) \right] \rho(\mathbf{r_2}) d\mathbf{r_1} d\mathbf{r_2}. \tag{2.24}$$

In the above expression, $r_2 < r_1$. The O_{lm} are called chargeless multipole moments and the M_{lm} are called chargeless Taylor moments. The chargeless multipole moment expansion is also called a *local expansion*, because it represents the potential seen from inside the charge distribution. The chargeless Taylor moment expansion is also called a *distant expansion*, because it represents the potential seen from outside the charge distribution.

The multipole and Taylor moments are obtained from these chargeless expressions using the following relationships

$$\tilde{O}_{lm}(\mathbf{r_2}) = \int \rho(\mathbf{r_2}) O_{lm}(\mathbf{r_2}) d\mathbf{r_2} \tag{2.25}$$

$$\tilde{M}_{lm}(\mathbf{r_1}) = \int \rho(\mathbf{r_1}) M_{lm}(\mathbf{r_1}) d\mathbf{r_1} \tag{2.26}$$

In the CFMM, the interactions must be grouped into near-field (NF) and far-field (FF). The NF interactions occur between distributions with any overlap. The FF interactions occur between distributions with no overlap. The NF interactions are computed with conventional integral evaluation methods like OS and HGP, and the FF interactions are computed with the CFMM. This classification is made based on the size of the distribution and the length of the box it comes in. The size of the distribution can be determined from its *extent*. The extent of a distribution is the distance one must go from its center before

the density falls to zero. For Gaussian products, it is defined from the exponent and the distance between the centers of the GTFs used to generate the product. For this work, the expression for the extent chosen is the one used in the original CFMM work [43, 44]

$$r_{ext} = \frac{1}{2}\sqrt{\frac{2}{\gamma} ln\left(\frac{1}{\epsilon}\right) - \mathbf{AB}^2}. \tag{2.27}$$

In this expression, γ is either the Gaussian product or auxiliary function exponent, depending on the function whose extent is being calculated. The $\mathbf{AB}^2$ in the above equation represents the distance between the centers of the primitive GTFs that were used to generate the Gaussian product. In the case of the auxiliary functions there is only one center, so this value is set to zero. ϵ represents the factor smaller that one can deviate from the true value for the extent.

From this definition of the extent, the *well-separated index* (WS) can be defined as follows

$$WS = max\left(2\left\lceil\frac{r_{ext}}{l}\right\rceil, WS_{ref}\right). \tag{2.28}$$

The WS is taken as the maximum value between two integers, the reference WS (WS_{ref}) and the computed WS. WS_{ref} is a minimum WS value that a distribution can have. The value chosen for this number represents the minimum distance between two distributions before they are treated with the CFMM. The larger it is, the fewer the distributions treated with the CFMM, but the more accurate the approximations are. In previous work, this

value was chosen as that used for the point charge model [42], which is two. l is the length of the simulation space. It is obvious that as the simulation space is subdivided, the box length used in the definition of WS is halved. This doubles the WS. Since the box length is the only variable used in generating WS that changes with subdivision level, the WS at each level can be calculated by simply doubling the value it had in the previous level. Therefore, explicit calculation of the WS is only done once for each distribution. The $\lceil\rceil$ represents the *ceiling* of a number. It is generated by rounding the number within the symbols to the nearest whole integer value.

The WS in the above equation represents the minimum number of boxes that must lie between two boxes before they are considered well-separated. For example, a WS of four for a particular box means only those boxes which lie at least four boxes away will have their interactions computed by the CFMM [42–44].

2.4.3 The CFMM Algorithm

The CFMM algorithm has six parts. First, the tree is built and distributions are placed in the lowest level boxes. Next, the multipole moments about these box centers are generated and translated up the tree. The boxes which are well-separated from one another are then determined and the multipole expansions in one box are converted to Taylor expansions about the second box. These Taylor expansions are then translated down the tree. At the lowest tree level, interactions are computed for each distribution. Finally, the NF contributions are added and the results are added to the Kohn-Sham matrix. These parts will be outlined below.

Build the Tree

The CFMM algorithm begins by surrounding the simulation space with a cube. This cube is just large enough to include all the distributions. The cube is then subdivided in three dimensions.

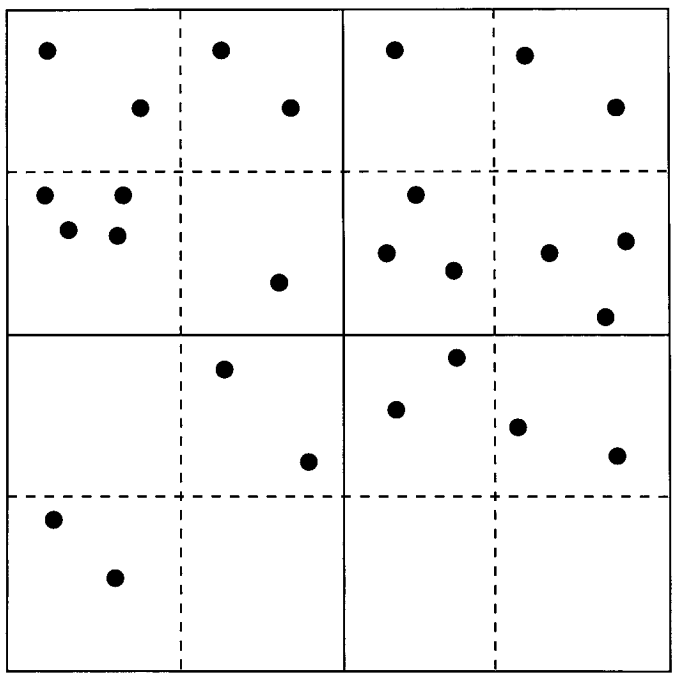


Figure 2.3: Application of three subdivisions to a two-dimensional simulation space.

An example of two subdivisions applied to a two-dimensional system is shown in Figure 2.3. The first subdivision is shown with a solid line. This divides the simulation into four boxes of equal size. The second subdivision is shown with a dashed line. Within each box, this subdivision results in four more boxes of equal size. This set of four boxes that are obtained in each box from this subdivision are called *child* boxes. The box which was subdivided is called the *parent* box. At this subdivision level, there are 16 boxes. Note that the distributions are not uniformly placed in space. As a result, some boxes are empty

and those that are not do not necessarily have the same number of distributions.

The subdivision of the simulation space results in a *computational tree.* In three dimensions, it is called an *oct-tree.* The subdivisions continue until an optimal number of boxes is reached. This optimal number is chosen to balance the savings gained from computing fewer pairwise interactions directly with the overhead of building and passing information through the tree [42–44].

The WS are then determined for each distribution using Eq.2.28. The distributions are then sorted into the lowest level boxes using a *radix sort* [53]. Unlike the original CFMM algorithm, this work uses both auxiliary functions and Gaussian products to represent charge distributions. Since the auxiliary basis functions are used to approximate the density of the system, the potential felt by each Gaussian product is generated using the auxiliary functions. As a result, it is these functions which are grouped together and passed through the tree.

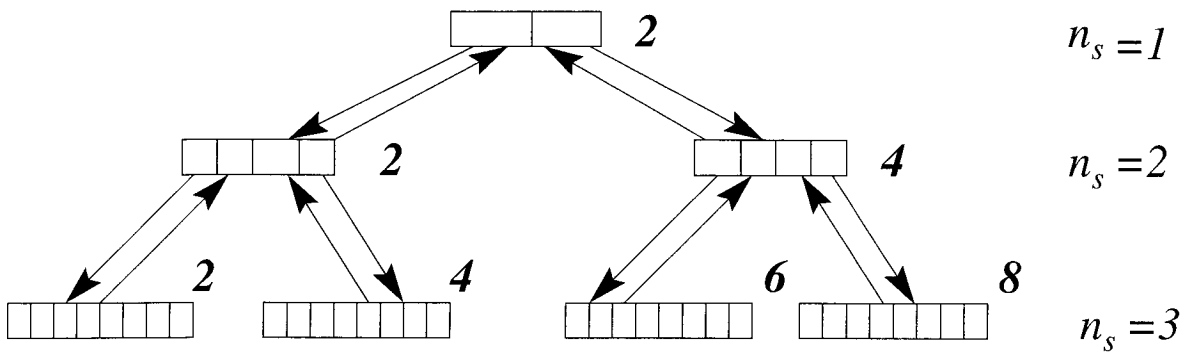


Figure 2.4: A representation of a one-dimensional CFMM tree for three subdivisions in space.

The distributions are sorted first by location in space, then by well-separated index. The resulting tree is shown in Figure 2.4, where the n_s indicate the level of subdivision of

the tree. The leaves of the tree are shown at $n_s = 3$. Each branch of the tree corresponds to a subdivision in space with the same WS for every distribution in it. Within each branch at each level, the boxes represent the subdivided space. At the bottom tree level, each leaf corresponds to eight boxes in one-dimensional space with the WS shown by the number to the right of the box. The arrows show the passage of information through the tree by WS.

Once the boxes are sorted into the leaves of the tree, it is then *pruned.* This removes any boxes that do not have distributions in them.

Form Multipole Moments, Pass Up Tree

The multipole expansions are generated for each charge distribution about its origin using Eq.2.25. To group distributions, they must be gathered into a single expansion about the box center. Since multipole expansions with common centers can be summed, these expansions must be translated to the box center [42–44].

The direction of translations in space is shown by the arrows in Figure 2.5. Also labelled in this diagram are the parent and child boxes, where the child boxes are obtained from the subdivision of the parent boxes. This subdivision is started at the level of the tree where it is first possible to obtain well-separated boxes. In the diagram, this corresponds to $n_s = 3$. Again, this tree was built by subdividing in one dimension in space. Note that information is passed up the tree until the level is reached where well-separated boxes are first encountered. In Fig. 2.5, this level corresponds to the first occurrence of the child boxes. Above this level in the tree, none of the boxes are well-separated and the

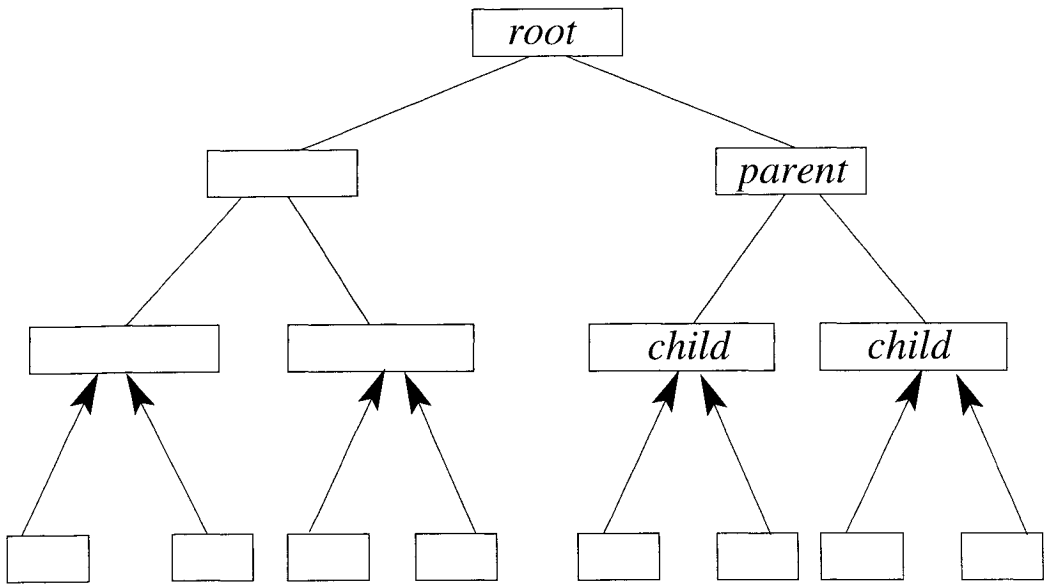


Figure 2.5: Diagram showing direction of translations for the first pass of the CFMM tree.

expansions are not used.

The translations to the box centers in the leaves of the tree and translations up the tree are accomplished with the *local-to-local* translation operator. The general form of this operator is [42–44]

$$O_{lm}(\mathbf{a}+\mathbf{b}) = \sum_{j=0}^{l} \sum_{k=-j}^{j} O_{l-j,m-k}(\mathbf{b}) O_{jk}(\mathbf{a}). \tag{2.29}$$

This corresponds to a translation of origin of the expansion of a charge at **a** from **O** to -**b**.

Convert to Taylor Moments

At each level, there is one multipole expansion centered about each box that is a sum of the multipole expansions generated for each charge distribution located in the box. Now, the boxes whose charge distributions can interact will be determined. To determine

if two boxes will interact, the WS must be generated for the box pair. This step is started at the first subdivision which can have well-separated boxes. If the root of the tree is assigned a level of zero (i.e. $n_s = 0$), this action is started at the second level.

A WS must be generated for each pair of boxes. If the two boxes are in the same WS branch of the tree, the WS for both boxes are the same. The minimum number of boxes that must be between them for them to interact is equal to the WS of either box. If the two boxes are in different WS branches of the tree, then they have different WS. In this case, the minimum number of boxes is taken as the average of the WS of the two boxes [42–44].

Once a WS is assigned to the pair of boxes, the actual number of boxes between them is determined. This number is compared to the WS. If the number of boxes between the pair is less than WS, then no conversions are done at this tree level. If the number of boxes between the pair is at least that of the WS, then the multipole expansions in each of the boxes are converted to Taylor expansions in the other boxes in the pair. The general form of this *local-to-distant* conversion is

$$M_{lm}(\mathbf{a}-\mathbf{b}) = \sum_{j=0}^{\infty} \sum_{k=-j}^{j} M_{j+l,k+m}(\mathbf{b}) O_{jk}(\mathbf{a}) \tag{2.30}$$

The conversion of the multipolar expansion in the first box to a Taylor expansion in the second box in the pair can be obtained from the conversion of the multipolar expansion in the second box to a Taylor expansion in the first box. As a result, these conversions can be done simultaneously, eliminating one function call per pair. It is important to note

that these conversions need only be done at the highest level that two boxes interact. This is true because in the next step of the algorithm, the boxes at lower levels will inherit this information. To make sure this is actually what is done, the algorithm starts at the top of the tree. If the multipole expansions in the two boxes are converted to Taylor expansions, the process is not repeated with their children [42–44].

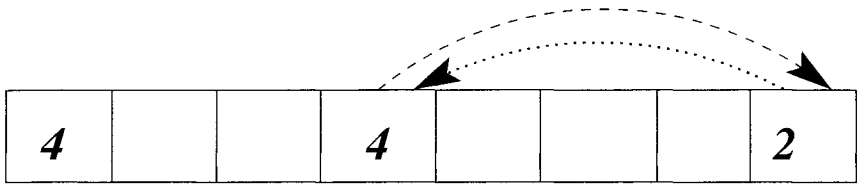


Figure 2.6: Diagram showing direction of translations for the second pass of the CFMM tree.

As an example, consider the set of boxes in Figure 2.6. The numbers in the boxes correspond to their WS. The arrows show the direction of the conversion of expansions. This diagram illustrates three possibilities for interactions at this level:

- Starting from left to right across the boxes in the diagram, consider the first pair of boxes containing numbers. Both boxes have a WS of four. The overall WS is then four. However, there are only two boxes between them, so they cannot interact at this tree level. Also, since they cannot interact at this level, their parent boxes will also not interact. No conversions are done for this box pair at this level.

- Now, consider the two outside boxes, one with WS of four and one with WS of two. The overall WS is three and the number of boxes between them is six. However, the possibility of the parent boxes interacting must be considered. The parent boxes will both have the same WS of two. Pairing boxes from left to right gives the parent

boxes. From this, the number of intervening boxes for the parents is two. Since the children of this box pair will inherit this information, then these two child boxes do not interact. No conversions are done for this box pair at this level.

- Finally, consider the second box with a WS of four and the box with the WS of two. The WS of the pair is three and the number of boxes between them is three. Also, using a similar argument as above, it is found that the parents of these two boxes are not well-separated. Thus, these two boxes interact. The conversions are done for this box pair, as is shown by the arrows in Fig. 2.6.

This process is repeated for each lower level of the tree. Within each box, the Taylor expansions for each well-separated box are summed.

Pass Taylor Moments Down Tree

Once the Taylor expansions for each box are obtained, they must be passed down the tree from parent box to child box. This process is started at the tree level where the boxes can first be well-separated. The general form for this *distant-to-distant* translation is [42]

$$M_{lm}(\mathbf{r}-\mathbf{b}) = \sum_{j=l}^{\infty} \sum_{k=-j}^{j} O_{j-l,k-m}(\mathbf{b}) M_{jk}(\mathbf{a}) \tag{2.31}$$

This corresponds to a translation of the origin O (of a charge at $\mathbf{r}$) to $\mathbf{b}$.

Fig. 2.7 shows the direction of these translations through a one-dimensional binary tree. As in the first translation, the parent and child boxes are labelled and the arrows show the direction of translation. The expansions in each box are summed about its center. Once

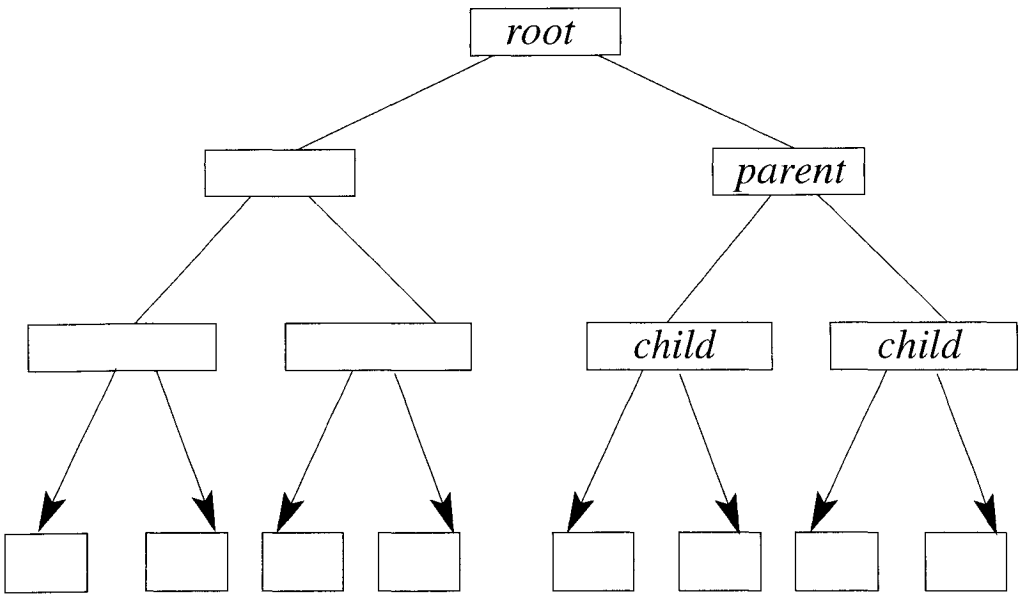


Figure 2.7: Diagram showing direction of translations for the third pass of the CFMM tree.

the translations are done, each leaf box contains a potential expansion due to all boxes well-separated from it [42–44].

Compute FF Interactions

At this point, the FF contributions to the Kohn-Sham matrix are calculated. For each Gaussian product, the multipolar expansions are generated and translated to the box centers. However, in this step, only the expansions for Gaussian products which are part of the same contracted shellpair are summed. These contracted Gaussian products are then interacted with the Taylor expansion in each box.

Generate Kohn-Sham Matrix Elements

The NF contributions to the Kohn-Sham matrix elements are computed and added to the FF contributions. This result is added to Kohn-Sham matrix.

It is worth mentioning that there is a small overhead incurred by generating expansions for the auxiliary GTFs as well as the Gaussian products. However, the algorithm still scales linearly with system size. Also, some savings are obtained in generating the NF contributions as there are far less integrals to calculate than with using Gaussian products alone.

2.4.4 Parallel Extensions to the CFMM

The above algorithm for the CFMM was designed to run on a single processor. Another potential for decreasing the time required to run a calculation is to run on more than one processor. Recently, there has been a lot of work done to design programs that run on multiple processors. One of the main goals for designing parallel programs has been to have calculations take less time to complete and to use less memory.

Before discussing the results of running the CFMM in parallel, a few terms used often in parallel computing will be discussed.

Parallel Programming and Speedup

Parallel programming is the design of programs to run on more than one processor. The goal of programming in parallel is to have as many parts of a calculation as possible running concurrently. This reduces the time and memory needed for the overall calculation.

Parallel programs are designed in one of two ways. One way is to have the same sequence of instructions done on each processor, but with different input values. Parallel programs of this nature are called *data-parallel.* A second way is to run different sets of instructions on each processor. Parallel programs of this nature are called *task-parallel.* Depending on the problem to be solved, programs are either task-parallel or data-parallel [54].

Most scientific programs are data-parallel. An example of a data-parallel program is the calculation of the Coulomb contributions to the Kohn-Sham matrix. This will be discussed later when the parallel CFMM is discussed in more detail. On the other hand, there are a few task-parallel scientific applications. An example of this is the construction of the energy components for a molecular mechanics calculation. Here, different groups of energy components are computed on different processors. Since these groups have different formulae, the set of instructions to generate them is also different.

A metric that is used to compare timings of parallel programs on different computers is *speedup.* The speedup gives an indication of how much faster a parallel program will run if the number of processors is increased. It is calculated using [54]

$$S = \frac{t_s}{t_p}. \tag{2.32}$$

The above relationship is a simple case of calculating speedup S, where the time to run on a single processor is divided by the time to run on multiple processors. This formula is a good indication of the scalability of the algorithm. However, more often, a measure of the

increase in speed of running an algorithm on p processors as compared to one processor is what is actually sought. This can be complicated by part of the algorithm running serially as well as the overhead for interprocessor communication. Taking into account the fraction f of the program that runs sequentially, the time to run in parallel becomes $ft_s + (1-f)t_s/p$. This results in a speedup of

$$S = \frac{t_s}{ft_s + (1-f)t_s/p} = \frac{p}{1+(p-1)f}. \tag{2.33}$$

This is another statement of *Amdahl's law.* This law says that there is a maximum speedup that can be achieved with any parallel algorithm. After the maximum is reached, adding more processors will not decrease the time for the computation. The amount of speedup that is obtained depends heavily on the amount of time the program spends executing instructions sequentially. In fact, in the limit of an infinite number of processors, this maximum value is the inverse of the fraction of the program spent doing sequential tasks on a single processor. Therefore, the best speedups are obtained for codes that spend very little time executing sequential instructions on a single processor [54].

Shared Memory Parallelism

There are two types of parallelism and these depend on the architecture of the computer or computers. They are *shared* and *distributed memory parallelism.* A discussion of distributed memory parallelism will follow.

On today's desktop computer, there is one processor which runs programs from one

main memory. However, one can also have a computer with more than one processor running programs from the same main memory. These types of computers are called shared-memory computers. Physically, the processors are separated from the main memory by some type of interconnection network, such as a bus. Each bit of information in the main memory has an address as part of a single address space. These addresses are used by each processor to access that part of the memory.

Shared memory parallelism exploits the physical layout of shared memory computers. With a single main memory, each processor has access to all of the program and the data required. Parallel code running in a shared memory environment usually has a portion of the code which runs on each processor. Each processor accesses, from main memory, the portion of code it is instructed to run.

Shared memory parallelism has the advantage that everything is stored in the same memory, so little or no communication between processors is necessary. In fact much of the communication is between the processors and the memory. As a result, for systems with a smaller number of processors, this type of parallelism is attractive. However, for large numbers of processors on the same interconnection network, problems can arise where processors are waiting to access the network [54].

Distributed Memory Parallelism and Message Passing

An alternative to the shared memory implementation is to use a distributed architecture. This type of parallel architecture consists of linking entire computers together. Each computer has its own processor and memory and the computers are connected by some

type of network connection, such as an ethernet. In this physical layout, memory is not shared. As a result, some sort of method of sharing data between processors is needed. This is achieved by using a message passing environment.

Distributed memory parallel programs exploit the distributed layout of processors and memory. The problem to be solved is divided up among the processors on the network. Each processor executes a part of the problem. The data is shared between processors by sending it as a message packet on the network.

The most common way to implement message passing is to insert calls from a set of library routines designed for passing messages. This is the idea behind the popular *message passing interface*, or MPI [55]. MPI consists of a set of routines to implement message passing. Some (or all) of these routines are added to existing serial code. There are roughly 120 routines in the MPI library. However, most codes use only a small subset of these frequently. The more commonly used routines will be discussed, and a complete description of the rest can be found in Ref.55. The types of routines discussed here fall into one of three general categories: initialization and termination, point-to-point communication, and global communication.

MPI is started by running the parallel program from the command line using the *mprun* command:

```
mprun -np 4 program
```

The number of processors is specified with the −np switch. Once this is done, mprun grabs the machines that the program is running on from a file with a list of all the machines that can run the program. This file must be defined by the user. The processors that are

specified here are placed in a communication group called the *communication world.* In this group, they will have a rank, which starts at zero and goes to p−1, where p is the number of processors specified by −np.

MPI is initialized in the program using the command MPI_Init. Once this is done, the group rank and size are determined using MPI_Comm_rank and MPI_Comm_size. Once finished, MPI can then be terminated with MPI_finalize.

Point-to-point is one of two types of communication within MPI. It consists of messages being passed between pairs of processors. These messages are passed using one of several variants of *send* and *receive* that exist in the MPI library. The differences between types of sends and receives will not be discussed. To use these operations, a few parameters that they require must be defined. First, some information about the message being sent or received must be specified. The type of data is specified using one of several standard MPI data types such as MPI_Integer and MPI_Real. The number of data elements must also be specified, as does the address of the send or receive buffer. Message tags provide another means of identifying messages. Also, some information about the communicating processes is needed. In a send, the rank of the destination process must be specified. In a receive, the rank of the sending process must be specified. Finally, the communication group of the processor sending or receiving must be given.

The second type of communication is global communication. It is used when messages must be sent or received by all the processors in the communication group. Two common global operations are *broadcast* and *gather.* Broadcast sends information to all processors

and gather collects information from all processors in the communication group. The information required to perform a broadcast or scatter is nearly identical to that required for point-to-point communication. The difference is that in the case of global communication, no message tags are used and only the rank of the process initiating the communication need be specified.

There are a few advantages to message passing. First, it is simpler to implement. One does not have to worry about things like task scheduling and issues concerning concurrent data access and writes. Also, each process has its own sections of data that cannot be overwritten by other processes. Finally, message passing requires little modification to be done on existing serial codes.

When evaluating parallel programs which use message passing, one must include the time for communication between processors in the parallel execution time. Thus, the speedup is modified to reflect this:

$$S = \frac{t_s}{t_{comp} + t_{comm}}. \tag{2.34}$$

Parallel CFMM

The serial implementation of the CFMM described above can be modified to run in parallel in one of two ways. First, the list of integral contributions to the Kohn-Sham matrix can be divided among processors. Second, the list of translations within a step of the CFMM can be divided among processors [56, 57].

In the current implementation of the CFMM, the list of integral contributions (or the

list of Gaussian products) is divided among the processors. These smaller lists are sent to each processor and the entire CFMM algorithm is run concurrently on each processor. This is the simplest way to parallelize the CFMM. However, it is not the most efficient. First, the overhead of the CFMM algorithm resulting from building and translating information through the tree is the same on all processors. Second, when the integrals are subdivided among processors, a point is quickly reached where adding another processor saves no more time. This is because at this point, there are too few integrals on each processor to justify using the CFMM. As a result, dividing up tasks at each step of the CFMM may provide improved speedups for larger number of processors than dividing up the integral contributions.

2.5 Results

To test the CFMM, a series of polyglycine chains are used. These chains consist of n $-NHCH_2CO-$ units capped at one end with $-NH_2$ and at the other end with $-CHO$.

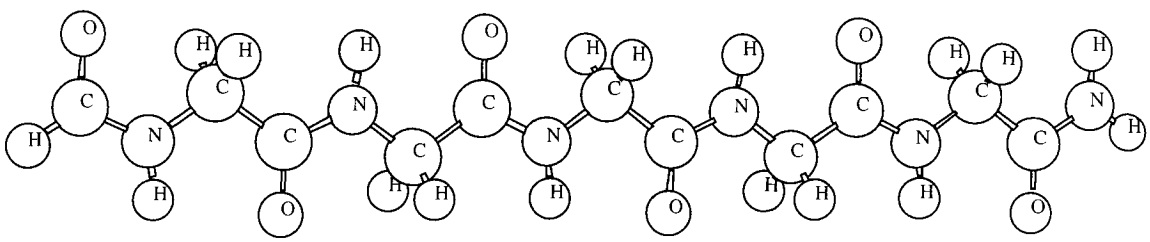


Figure 2.8: A 5 monomer polyglycine chain.

Chains with n=5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 are generated. This corresponds to 41, 76, 111, 146, 181, 216, 251, 286, 321, and 356 atoms, respectively. Figure 2.8 shows an example of a polyglycine chain which has 5 glycine units. These chains provide an excellent test of the method. They are a good approximation to a one-dimensional system and the density is localized within the chain.

For each of the polyglycine chains, single point calculations are performed using the DeFT density functional package [58]. Becke's 1988 gradient-corrected exchange functional [7] is employed together with Perdew's 1986 gradient-corrected correlation functional [8]. The 6−31G* orbital basis set is used. An auxiliary basis set is used to approximate the electron density [59]. This basis set is built by expanding the density in uncontracted Gaussian-type functions of s, p, and d-type. It contains two sets of functions for each atom: one consisting of s-type functions only and one with a set of s, p, and d-type functions. Members in each of these sets share exponents. For example, the carbon atom in DeFT contains four stand-alone s-type functions and four sets of s, p, and d-type functions constrained to having the same exponent. The density matrix convergence criterion is set to 10^{-8}. This corresponds to an energy convergence criterion in the SCF of about 10^{-6} Hartrees. The divide-and-conquer method for fitting the density is used [60, 61]. Cutoffs for the DAC density fitting and subsystem KS calculations were set to 7.0 Å and 15.0 Å,

respectively.

Table 2.1: Errors in the CFMM for each of the polyglycine chains.

number of monomers	Δ E (cal/mol)
5	0.0000
10	0.0013
15	0.0016
20	0.0484
25	0.0377
30	0.0008
35	0.0228
40	0.2462
45	0.0942
50	0.2432

For the CFMM calculations, the multipolar expansion was truncated at 15 terms. It was determined that cutting off the multipolar expansion at 15 terms resulted in an error in the energy of less than 10^{-6} Hartrees. For each system, the error introduced due to the CFMM is shown in Table 2.1. This is less than the error introduced by using a density matrix cutoff of 10^{-8}. The maximum tree depth that can be used is seven. Larger tree depths are not possible as more memory is required than is available.

Three sets of calculations are generated. In all three sets, the near-field contributions are calculated with a combination of the OS and HGP methods. In the first set, the far-field contributions are calculated in the same manner as the near-field contributions. This set of data is labelled the *conventional* method. In the second set, the far-field contributions are obtained using a low-order (but exact) multipole approximation of the fitted density at each atom center. This expansion is then interacted with the Gaussian products which are a sufficient distance from the atom center. This gives the integral

contributions to the Kohn-Sham matrix. This set of calculations is labelled the *multipolar approximation* method. In the third set of calculations, the CFMM is used to generate the far-field contributions.

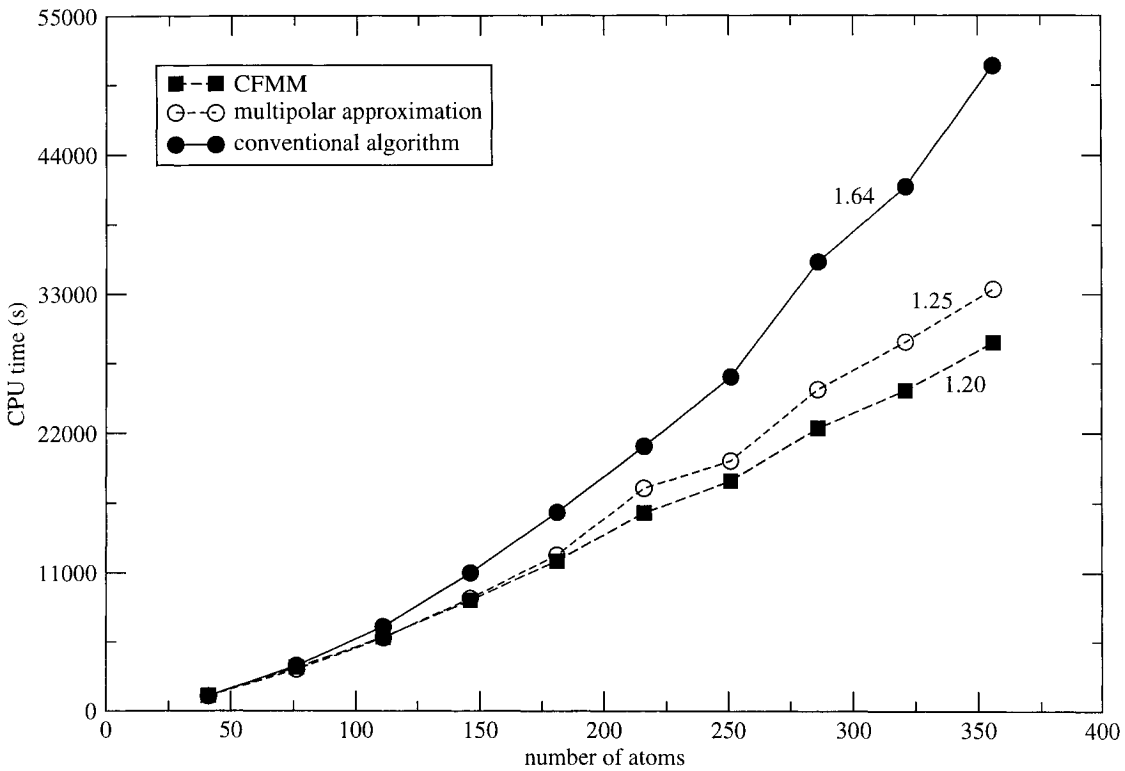


Figure 2.9: Change in total CPU time with system size.

Figure 2.9 shows the change in total CPU time with the number of atoms in the polyglycine chain. Also shown are the scalings for the various methods employed for calculating the far field contributions to the two-electron integrals. These scalings are generated using the last three data points on each curve.

The scalings given by the three methods are as follows. Using only the conventional

method, the overall CPU times scale as $N^{1.64}$. Using instead the multipolar approximation for the far-field interactions, the CPU times scale as $N^{1.25}$. Using the CFMM for the far-field interactions, the CPU times scale as $N^{1.20}$. For the chain with 50 glycine units, the

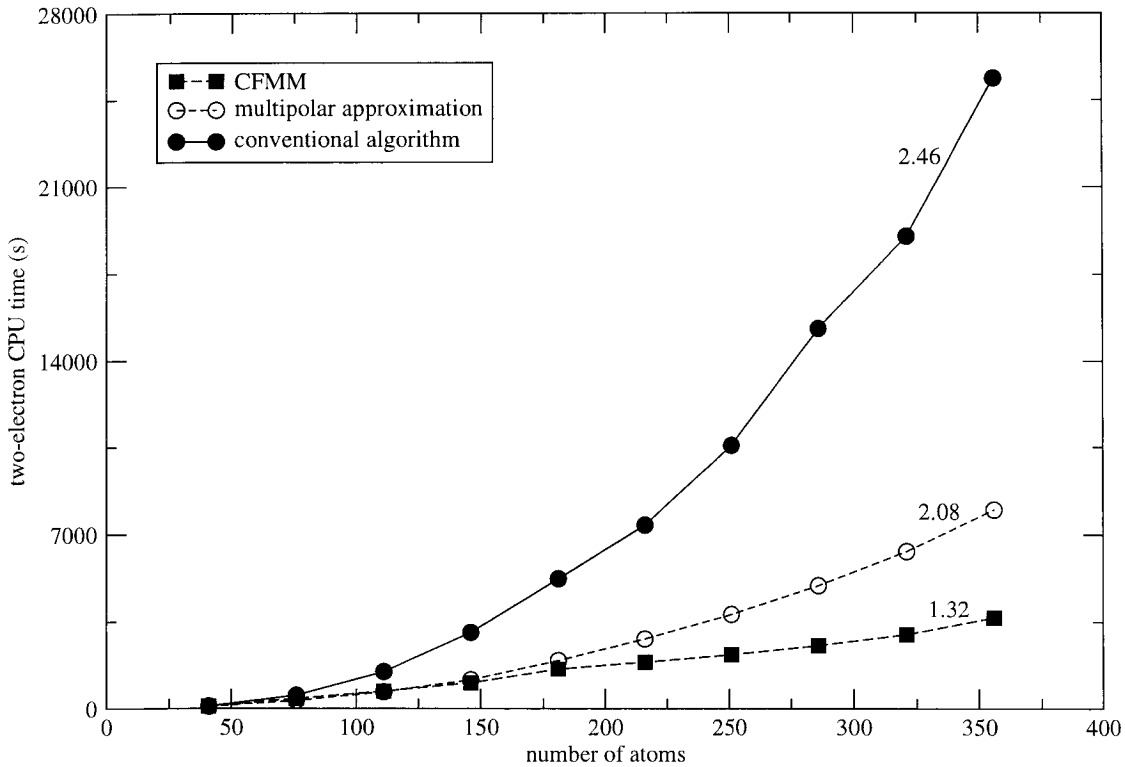


Figure 2.10: Change in two-electron CPU time with system size.

Therefore, for the total CPU times, the CFMM shows significant timing improvements over both the conventional method and the multipolar approximation. Also, use of the CFMM to compute far-field interactions results in effective linear scaling for the total CPU times.

Figure 2.10 shows the change in the CPU times for the calculation of the two-electron

integrals with the number of atoms. Again, the scalings are shown for the three methods. They are calculated using the last five data points. The changes in the scaling are much more pronounced in this Figure than in Figure 2.9. For smaller systems, the calculation of the two-electron integrals dominates the CPU time and changes in it are more noticeable. However, for larger systems, the diagonalization of the subsystem Kohn-Sham matrices with the DAC calculation becomes the dominating contribution to the total CPU time. This effect is large enough that the changes in the timings due to the two-electron integrals are less pronounced.

The scalings for the three methods are as follows. The conventional method scales as $N^{2.46}$, while using the multipolar approximation for the far-field contributions scales as $N^{2.08}$. If the CFMM is used instead for the far-field contributions, this scaling becomes $N^{1.32}$.

The above results show that the scalings for the conventional method and multipolar approximation are quadratic, while for the CFMM, this scaling is nearly linear. From the curve for the CFMM results, an interesting observation is made. The curve shows the overall behavior as effectively linear, but there is locally quadratic behavior. For the first five points and for the last five points, some curvature is seen. At the fifth point, corresponding to a chain of 25 glycine units, there is a peak in the graph separating the two curves.

To explain this locally quadratic behavior, consider the time required to compute the near-field and far-field interactions. The tree depth is chosen so the total time to calculate the two-electron interactions is a balance between the time needed to calculate

the near-field and far-field interactions. Ideally, the near-field interactions should scale as O(Mn), where M represents the average number of charge distributions in the lowest level boxes. To achieve linear scaling, M must remain independent of the overall number of particles. In three dimensions, increasing the depth of the tree by one increases the number of lowest level boxes by eight times. This means in order to keep M independent of the overall number of particles we must increase the number of particles by eight times as well. This is, however, not possible. In the case where an optimal tree depth is achieved and constant over a series of systems, as the number of particles increases, M does as well. This corresponds to an increase in the quadratic behavior. This is what is seen in the above curve. One way to reduce this locally quadratic behavior is to use *adaptive boxing* [62,63]. In the adaptive boxing method, subdivisions are allowed to be done independent of an overall subdivision scheme. This is especially useful for highly nonuniform layouts of charge distributions.

Figure 2.11 shows the change in the total CPU time with the number of atoms for the polypeptide chains. Unlike the previous two graphs, the largest calculation done in parallel is a peptide chain with 45 glycine units. The data for the 50 glycine chain could not be generated due to memory restraints when running in parallel.

The calculations were run on one, two, four, and eight processors. From the graph, speedup is obtained overall when the number of processors is increased. For the polypeptide chain with 45 glycine units, speedups from one to two, two to four, and four to eight processors are 1.3, 1.1, and 1.1, respectively. From these numbers, the largest speedup is seen in going from one to two processors. From two to four and four to eight processors,

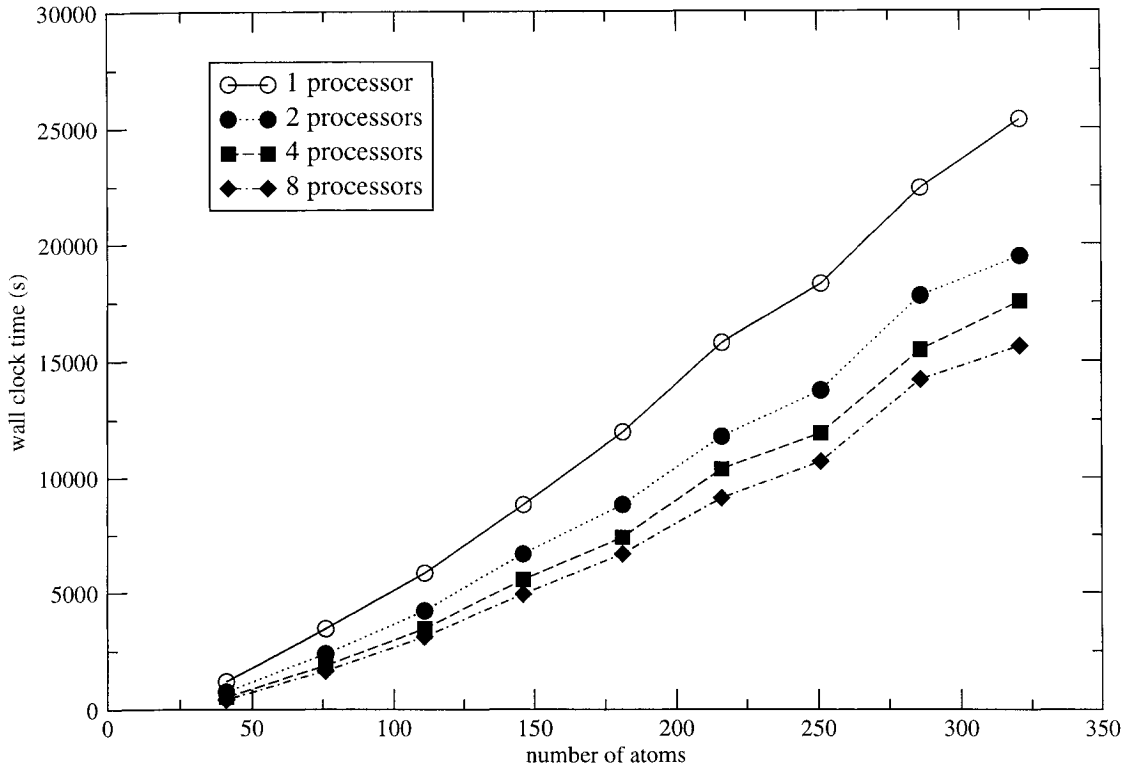


Figure 2.11: Change in total CPU time with system size for 1, 2, 4, and 8 processors.

this speedup is reduced.

Figure 2.12 shows the change in the two-electron CPU time with the number of atoms for the polypeptide chains. These calculations were run on one, two, four, and eight processors. Once again, the 50 glycine chain could not be run due to memory constraints. From the graph, speedups for the evaluation of the Coulomb integrals is more pronounced. For large systems, diagonalization of the Kohn-Sham subsystem matrices dominates the CPU time, so speedups in the two-electron code are less pronounced overall. To evaluate speedups, the timings from the largest calculation are used. In going from one to two processors, a speedup of 2.0 is obtained. In going from two to four processors, a speedup of 1.4 is obtained. In going from four to eight processors, a speedup of 2.2 is obtained. A

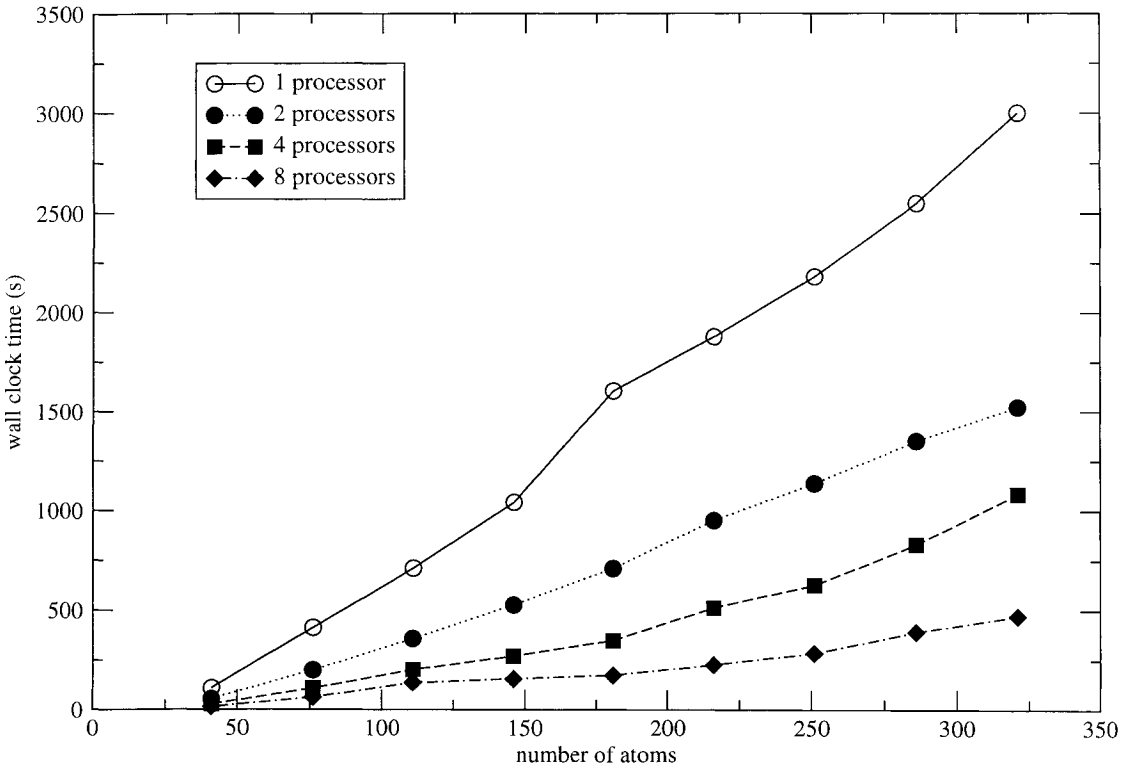


Figure 2.12: Change in two-electron CPU time with system size for 1, 2, 4, and 8 processors.

slight drop in speedup is seen when increasing the number of processors from two to four. Overall, good speedups are seen when the number of processors is increased from one to eight.

It should be mentioned that the machine where the parallel jobs were run is not a dedicated machine. That is, when the jobs were run, several other parallel jobs were running at the same time. Load balancing and increased or decreased communication between processors can affect timings from day to day. These effects are especially pronounced for parallel jobs. As a result, the scalings reported are not quantitative. Essentially, if the same jobs were run again, they may vary slightly. A good example is seen in the case of the two-electron integral calculations on the 45 glycine polypeptide. A speedup of more

than two is seen when the number of processors is increased from four to eight. This is called *superlinear speedup* and can occur because of an inefficiently parallelized algorithm or load balancing problems on the computers. This can cause the timing for either the four or eight processors to be an inaccurate representation of the actual timing for the algorithm to run on these number of processors. A more quantitative gauge of the scalings would be obtained if the jobs were run on a dedicated machine.

2.6 Conclusion

Over the last decade, several methods of reducing the time to compute the two-electron Coulomb integrals have emerged. One method, the CFMM, allows one to generate these integrals by grouping them into far-field and near-field interactions. Calculation of the near-field interactions uses a direct approach to generating the integrals analytically. Calculation of the far-field interactions uses a divide-and-conquer approach resulting in a computational tree which is used to share information within groups. The resulting algorithm is effectively linearly scaling. As a result, the algorithm can be used for calculations which were previously not feasible. Extension of the CFMM algorithm to run on more than one processor has also proven effective in reducing the time needed to perform the calculations on the polyglycine chains.

Chapter 3

Theoretical Nuclear Magnetic Resonance

3.1 Introduction

Nuclear magnetic resonance (NMR) has widely become a useful probe into the electronic structure of a system. The chemical shieldings obtained from NMR are highly sensitive to small changes in this structure. Experimental NMR has been instrumental in the structural determination of many molecules. It has also been used to monitor the progress of a reaction. Despite the widespread use of experimental NMR, theoretical applications in this area have only appeared in the last 20 years.

The origin of the chemical shielding is from an interaction of the nuclei with the moving electrons in the atom or molecule [64, 65]. To study these effects by theoretical methods, some modification on the effects of the field and electrons on the nuclear moments

must be made to the current field-independent theory. A thorough description of these changes can be found in Refs. 66 and 67. Excellent reviews on the method of calculating chemical shieldings with modern theoretical methods can be found in Refs. 68 and 69.

Chemical shieldings can now be calculated with semiempirical, HF, MCSCF, DFT, MP2, and CCSD/CCSD(T) levels of theory with various sizes of basis set, depending on the system size. However, before these shieldings can be compared with experiment, some considerations between theory and experiment must be taken into account.

3.2 Basic Principles

The NMR signal is a result of a combined effect of nuclear and electronic motion within atoms and molecules. The origin of these effects is derived from classical electromagnetic theories [38]. For a better understanding of theoretical NMR, it is first necessary to review some of the properties of moving charges and the magnetic effects that result from their motion.

3.2.1 Physical Principles of Magnetism

The Vector Field

Consider the case where the current density is placed in an applied magnetic field $\mathbf{B}^{app}$. Then the total field felt by the charges is a sum of the applied and induced fields

$$\mathbf{B} = \mathbf{B}^{app} + \mathbf{B}^{ind}. \tag{3.1}$$

The total magnetic field is defined in terms of a vector potential $\mathbf{A}$

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{3.2}$$

where the vector potential, written as a function of the position $\mathbf{r}$, is

$$\mathbf{A}(\mathbf{r}) = \frac{1}{c} \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + |\nabla f|, \tag{3.3}$$

and f is an arbitrary scalar function. In this expression, the term ∇f represents the *gauge transformation*, which is necessary for the definition of $\mathbf{A}$. Since $\mathbf{A}$ uniquely determines $\mathbf{B}$, care must be taken in the form of its definition. The reason for this is that the magnetic field must be invariant to this transformation. If the magnetic field is not invariant to this transformation, properties calculated from it will depend on the choice of *gauge*, or origin of the vector potential. The choice of f is given in terms of the choice of the divergence relation $(\nabla \cdot \mathbf{A})$, and different choices for this value give different definitions of the vector potential. One choice of gauge origin is the Coulomb gauge, where $\nabla \cdot \mathbf{A} = 0$ [38, 67].

The scalar function transforms the vector field, and this transformation is called a *gauge transformation*. Gauge transformations are often chosen to simplify the mathematical relationships specific to the problem studied. A commonly used gauge transformation is the *Coulomb gauge*, which is defined by the divergence relation $\nabla \cdot \mathbf{A} = 0$.

Magnetic Moments

Consider a current density for a small volume of space. This localized current density results in an induced magnetic field at a point distant to the current. If the vector potential is expanded in a series of multipoles, it is found that the largest contributing term in the expansion is the dipolar term. This indicates that the induced magnetic field is dipolar in nature. From this, the *magnetic dipole moment* (also called the magnetic moment) and *magnetization* can be defined in terms of the current density as [38, 67]

$$\boldsymbol{\mu} = \frac{1}{2c}\int \mathbf{r}' \times \mathbf{J}(\mathbf{r}')d\mathbf{r}' \tag{3.4}$$

$$\mathbf{M}(\mathbf{r}) = \frac{1}{2c}(\mathbf{r} \times \mathbf{J}(\mathbf{r})). \tag{3.5}$$

In atomic and molecular systems, the magnetic moment results from the orbital motion of the electrons as well as their spin. For an electron with orbital angular momentum given by $\mathbf{L}$, the magnetic moment due to orbital motion is

$$\boldsymbol{\mu} = \frac{e}{2m_e c}\mathbf{L}, \tag{3.6}$$

and the energy of the magnetic moment in an applied magnetic field $\mathbf{B}^{app}$ is

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}^{app}. \tag{3.7}$$

This energy of interaction between the magnetic moment and the applied field is dependent

on the orientation of the magnet with respect to the magnetic field.

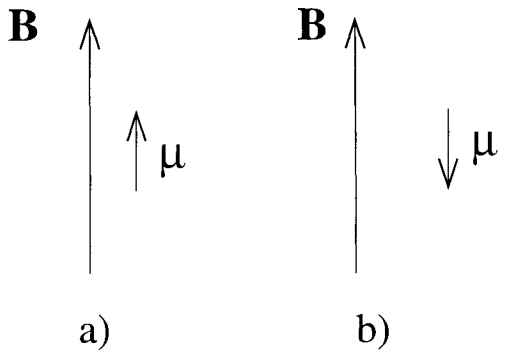


Figure 3.1: Two possible orientations of $\boldsymbol{\mu}$ with respect to the applied magnetic field.

In Fig. 3.1, two possible orientations of $\boldsymbol{\mu}$ with respect to $\mathbf{B}$ are given. In a), the magnetic moment is aligned parallel to the field. In this orientation, the interaction energy is a minimum. In b), the magnetic moment is aligned antiparallel to the field. In this orientation, the interaction energy is a maximum. If the magnetic moments are free to move, they will try to orient themselves as in a) [64, 66].

A second magnetic moment results from the spin angular momentum $\mathbf{S}$ of the electron. For the spin angular momentum, there is no classical description. However, it is related to the classical description $\boldsymbol{\mu} = (e/2m_e c)\mathbf{S}$ by the *g-factor*. This factor, derived from studies on the interaction of electromagnetic radiation with matter, accounts for the nonclassical behavior of spin.

It is clear that when accounting for the interaction of an electron with a magnetic field $\mathbf{B}$, the spin and orbital contributions to the magnetic moment and their interactions with the field must be considered [64, 66, 67].

So far, only magnetic moments from electrons have been discussed. However, nuclei also possess magnetic moments. This is due to the spins of the nuclei and is analogous

to the spins on electrons. However, unlike electrons, this spin can have integral and half-integral values. The magnetic moment due to the spin angular momentum $\mathbf{I}$ of a nucleus is given by

$$\boldsymbol{\mu}_I = \frac{+eg_e}{2m_{nuc}c}\mathbf{I}. \tag{3.8}$$

3.2.2 The Hamiltonian for a Molecule in a Magnetic Field

In the presence of a static applied magnetic field $\mathbf{B}$, electronic motion in this field results in local electric currents. These electric currents induce local magnetic fields in the molecule, usually at the nuclei. This gives rise to two contributions to the magnetic field. One contribution arises from the applied magnetic field. The other arises from induced field due to these electric currents. The induced magnetic field is related to the applied field by the *nuclear magnetic shielding*

$$\mathbf{B}^{ind} = -\boldsymbol{\sigma}\mathbf{B}^{app}. \tag{3.9}$$

The nuclear magnetic shielding is a symmetric tensor, with components in the x, y, and z

directions. It is defined as

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}, \tag{3.10}$$

where the σ_{ij} are the components of the magnetic shielding in the directions i and j, where i,j=x,y, or z.

Prior to application of a magnetic field, the nuclear magnetic moments exist in degenerate spin states. When the field is applied, the degeneracy is removed. The energy difference between the previously degenerate states with respect to the applied magnetic field is

$$\Delta E = \boldsymbol{\mu}_J(1 - \boldsymbol{\sigma})\mathbf{B}^{app}. \tag{3.11}$$

The idea is that if the transitions between different spin states for the nuclei J are investigated, some insight into the molecule's structure can be obtained. This information is obtained through the magnetic shielding. This value is related to the chemical shift of the nucleus by comparing the shielding in the molecule with the shielding from the same nucleus in a reference compound. The chemical shift is then

$$\delta = \sigma_{ref} - \sigma, \tag{3.12}$$

where $\sigma = (1/3)Tr(\boldsymbol{\sigma})$, and Tr refers to the trace of a matrix. The shielding can be obtained from the second derivative of the electronic energy with respect to the applied field and the nuclear moment as

$$\boldsymbol{\sigma} = \left(\frac{\partial^2 E_{el}}{\partial B_j^{app} \partial \mu_{J,i}} \right)_{\mathbf{B}^{app}, \boldsymbol{\mu}_J = 0}. \tag{3.13}$$

where J refers to the nucleus in question and i, j are the components of $\boldsymbol{\mu}_J$ and $\mathbf{B}^{app}$, respectively. The above second derivative expressions can be evaluated explicitly, or the perturbation theory approach can be applied. In this chapter, it is the perturbation theory approach that will be discussed. Further information on the explicit derivative expression evaluation is given in Ref. 69.

The Modified Hamiltonian and the Chemical Shielding

In order to obtain the perturbations due to the magnetic field and the nuclear moments, the electronic Hamiltonian must be modified. From the modified Hamiltonian, terms linear in both the applied magnetic field and the nuclear magnetic moments will be required to evaluate the chemical shielding.

The effects due to the magnetic field and the nuclear moments must be accounted for. For a molecule with no applied magnetic field, the behavior of the system is described by the non-relativistic time-independent Schrödinger equation [66, 67]

$$\hat{H}\Psi(\mathbf{x}, \mathbf{X}) = E\Psi(\mathbf{x}, \mathbf{X}), \tag{3.14}$$

where $\mathbf{x}$ and $\mathbf{X}$ are the electronic and nuclear coordinates, respectively. Applying the Born-Oppenheimer approximation and focusing on the electronic part of the Hamiltonian, the expression for this Hamiltonian in the absence of the magnetic field is

$$\hat{H}_{el} = \frac{\mathbf{p}^2}{2m} + V. \tag{3.15}$$

The two terms in the electronic Hamiltonian are the kinetic energy of the electrons and the potential energy due to electron-electron and electron-nuclear interactions, respectively. For the remainder of the chapter, the subscript *el* will be dropped and the electronic Hamiltonian will be assumed.

To determine the effect of the magnetic field on the nuclear moments, the momentum $\mathbf{p}$ in the above electronic Hamiltonian is replaced by the momentum in an applied vector field $p + e\,\mathbf{A}$. The vector field $\mathbf{A}$ has two contributions

$$\mathbf{A} = \mathbf{A}^{app} + \mathbf{A}^{ind}, \tag{3.16}$$

where the first term is the contribution from the applied field and the second contribution is from the field induced by the electronic currents. The expressions for these two contributions are [66, 67]

$$\mathbf{A}^{app} = \frac{1}{2}\mathbf{B}^{app} \times (\mathbf{r} - \mathbf{R}_G) \tag{3.17}$$

$$\mathbf{A}^{ind} = \frac{\boldsymbol{\mu}_J \times (\mathbf{r} - \mathbf{R}_J)}{|\mathbf{r} - \mathbf{R}_J|^3}. \tag{3.18}$$

In the above expressions, $\mathbf{R}_J$ represents the position of the nucleus, $\mathbf{R}_G$ is the gauge origin, and $\mathbf{r}$ is the spatial coordinate.

Replacing the momentum $\mathbf{p}$ with the expression for the momentum in an applied field, $\boldsymbol{\pi} = p + e\,\mathbf{A}$, in the field-free Hamiltonian and accounting for the interaction of the nuclear spin magnetic moment with the magnetic field gives

$$\begin{aligned}\hat{H} &= \frac{(\mathbf{p}+e\mathbf{A})^2}{2m_e} + V - \boldsymbol{\mu}_J \cdot \mathbf{B}^{app} \\ &= \frac{(\mathbf{p}+e\mathbf{A}^{app}+e\mathbf{A}^{ind})^2}{2m_e} + V - \boldsymbol{\mu}_J \cdot \mathbf{B}^{app} \\ &= \frac{\boldsymbol{\pi}^2}{2m_e} + V - \boldsymbol{\mu}_J \cdot \mathbf{B}^{app} \\ &+ \left(\frac{e}{2m_e}(\boldsymbol{\pi}\cdot\mathbf{A}^{ind} + \mathbf{A}^{ind}\cdot\boldsymbol{\pi}) + \frac{e^2}{2m_e}\mathbf{A}^{ind}\cdot\mathbf{A}^{ind}\right).\end{aligned} \tag{3.19}$$

To calculate the magnetic shieldings, it was previously noted that only terms first order in both the applied field and nuclear magnetic moments are needed. Therefore, from the above equation, the second-order term in the induced vector field can be dropped. This results in the following expression for the Hamiltonian in the presence of an applied magnetic field [66,67]

$$\hat{H} = \frac{\boldsymbol{\pi}^2}{2m_e} + V - \boldsymbol{\mu}_J \cdot \mathbf{B}^{app} + \left(\frac{e}{2m_e}\right)\left(\boldsymbol{\pi}\cdot\mathbf{A}^{ind} + \mathbf{A}^{ind}\cdot\boldsymbol{\pi}\right). \tag{3.20}$$

The first-order contributions in the induced magnetic field are given by

$$\hat{H}^{(1)} = \frac{e}{2m_e}(\boldsymbol{\pi}\cdot\mathbf{A}^{ind} + \mathbf{A}^{ind}\cdot\boldsymbol{\pi}). \tag{3.21}$$

Substituting in the expression for the induced magnetic field and noting that $\boldsymbol{\pi} \cdot \mathbf{A}^{ind} = \mathbf{A}^{ind} \cdot \boldsymbol{\pi}$ gives

$$\hat{H}^{(1)} = \boldsymbol{\mu}_J \frac{e}{m_e} \frac{(\mathbf{r} - \mathbf{R}_J) \times \boldsymbol{\pi}}{|\mathbf{r} - \mathbf{R}_J|^3}. \tag{3.22}$$

Integrating this term over the coordinates of the electrons gives

$$< \psi_e | \hat{H}^{(1)} | \psi_e >= -\boldsymbol{\mu}_J \cdot \mathbf{B}^{ind}, \tag{3.23}$$

where the ψ_e represent the ground state electronic wavefunctions.

In order to obtain the magnetic shielding, the perturbations on the ground state wavefunction must also be obtained with respect to the magnetic field. From the perturbation in the ground state wavefunction due to the applied field, one obtains

$$\psi_e(\mathbf{B}^{app}) = \psi_e^{(0)} + \sum_{i=x,y,z} B_i^{app} \psi_e^{(1)} + \cdots \tag{3.24}$$

Substituting this equation into Eq.3.23 and retaining only the terms linear in the applied field gives

$$\boldsymbol{\mu}_J \cdot \mathbf{B}_D^{ind} = \boldsymbol{\mu}_J \cdot \left\langle \psi_e^{(0)} \left| \sum_k \frac{e^2}{m_e} \frac{(\mathbf{r}_k - \mathbf{R}_J) \times \mathbf{B}^{app} \times (\mathbf{r}_k - \mathbf{R}_{gauge})}{|r_k - R_J|^3} \right| \psi_e^{(0)} \right\rangle \tag{3.25}$$

$$\boldsymbol{\mu}_J \cdot \mathbf{B}_P^{ind} = \boldsymbol{\mu}_J \cdot \sum_{i=x,y,z} B_i \left\langle \psi_{e,i}^{(1)} \left| \sum_k \frac{e}{m_e} \frac{(\mathbf{r}_k - \mathbf{R}_J) \times \mathbf{p}}{|r_k - R_J|^3} \right| \psi_e^{(0)} \right\rangle + CC, \tag{3.26}$$

where the D and P represent the diamagnetic and paramagnetic contributions to the induced magnetic field, respectively, and CC is the complex conjugate of the paramagnetic term. The diamagnetic contribution is usually the larger of the two contributions. This contribution is common to most atomic centers and results in a shift of the signal in the upfield direction. It is related to the ground state wavefunction. The paramagnetic contribution is the smaller of the two contributions. It results in a downfield shift of the signal.

Finally, expressing the Hamiltonian for what the nuclear spins see with respect to the rest of the system gives

$$\hat{H}_{spin} = -\boldsymbol{\mu}_J \cdot \mathbf{B}^{app} + \boldsymbol{\mu}_J \cdot \boldsymbol{\sigma} \cdot \mathbf{B}^{app} \tag{3.27}$$

From this equation, the magnetic shielding contributions can be obtained.

3.2.3 The Gauge Problem

From the relationship between the applied magnetic field and the vector field, it is obvious that the resulting magnetic shieldings depend on the choice of gauge transformation. In other words, they are origin-dependent. However, this choice of origin cannot affect the resulting shieldings. Exact solutions to the Schrödinger equation will give shieldings that are independent of the choice of gauge. However, these solutions require complete basis sets and are not attainable, so one must make use of various approximate methods. For these approximations, the gauge invariance requirement is not met. The only contributor

to this problem is the use of truncated basis sets [68, 69].

If a larger basis set is used, it is expected that better shieldings will be obtained. However, the convergence of approximate methods for calculating shieldings is slow with the size of the basis set, except in the case of perturbation-adapted basis sets (see Ref. 70). It would then be expected that a significant size increase in the basis set would be needed to see any improvement in the shieldings.

For atomic systems, the gauge invariance problem can be overcome by using a gauge centered on the nucleus. However, for molecules, which contain many nuclei, a single gauge origin will not give reasonable results. However, one approach which is popular is to choose many gauge origins. These methods are collectively called *local gauge origin* methods. There are several possible approaches for a local gauge origin. Two of them are at each nuclear center and at points throughout the molecule. Both approaches will be discussed as two of the few popular methods for calculating shieldings are based on these approaches. They are the *gauge-including atomic orbital* [71–74], or GIAO, and the *continuous set of gauge transformations* [74, 75], or CSGT, approaches.

3.3 Methods for Calculating Chemical Shifts

In the case of the application of an external field, whether electric or magnetic, the first-order wavefunction is often used to calculate these same properties. This is the case for calculating NMR chemical shifts. This section will then begin with a discussion of these equations which give the first-order coefficients. They are called the *coupled perturbed*

equations.

In the case of HF or MP2 theory, these coupled equations are a modification of the HF or MP2 equations to account for the perturbation in the wavefunction due to the applied field. These equations are derived from general perturbation theory. The discussion of their derivation and implementation can be found elsewhere [76–78], and will not be presented here.

3.3.1 The Gauge-Including Atomic Orbital Approach

So far, no mention has been made of the gauge invariance problem discussed in the previous section. It turns out that the magnetic shieldings calculated above are dependent on the chosen gauge origin. As a result, different shieldings are found for different gauge origins.

One way to circumvent the gauge invariance problem is to use a different gauge origin for each atom. This can be done by making the molecular orbitals linear combinations of field-dependent atomic orbitals, or London orbitals [71, 72]:

$$\phi_j(\mathbf{B}, \boldsymbol{\mu_J}) = \sum_{\nu} c_{\nu j} e^{-(i/c)\mathbf{A}_\nu \mathbf{r}} \phi_\nu, \tag{3.28}$$

where $\mathbf{A}_\nu = (1/2)\mathbf{B}$ x $\mathbf{R}_\nu$ is the vector potential for atomic orbital ϕ_ν for the nucleus $\mathbf{R}_\nu$.

Using conventional HF theory, the expression to evaluate which contains the field-dependent orbitals is

$$\mathbf{H}(\mathbf{B}, \boldsymbol{\mu}_J)\Psi(\mathbf{B}, \boldsymbol{\mu}_J) = E(\mathbf{B}, \boldsymbol{\mu}_J)\Psi(\mathbf{B}, \boldsymbol{\mu}_J). \tag{3.29}$$

The $\Psi(\mathbf{B}, \boldsymbol{\mu}_J)$ can then be expanded in a single determinant of one-electron orbitals. Following this, and applying the variational principle gives

$$\mathbf{F}(\mathbf{B}, \boldsymbol{\mu}_J)\mathbf{C}(\mathbf{B}, \boldsymbol{\mu}_J) = E(\mathbf{B}, \boldsymbol{\mu}_J)\mathbf{S}(\mathbf{B}, \boldsymbol{\mu}_J)\mathbf{C}(\mathbf{B}, \boldsymbol{\mu}_J). \tag{3.30}$$

The above equation is a modification of the Roothaan equations in the presence of a magnetic field using a linear combination of field-dependent atomic orbitals [71–74].

Evaluation of the magnetic shieldings will require the derivatives of the energy given in the modified Roothaan equations with respect to the applied field and the nuclear moments. Again, the required expressions can be obtained using a perturbation theory approach. Details of the derivation are left to the paper by Ditchfield [72]. The resulting expression for contribution $\sigma_{\alpha\beta}$ to the magnetic shielding using GIAO's is [73]

$$\sigma_{J\alpha\beta} = \sum_{\nu\lambda} \left\{ P^{(0)}_{\nu\lambda} H^{(1,1)}_{J\nu\lambda\alpha\beta} + P^{(1,0)}_{\nu\lambda\alpha} H^{(0,1)}_{J\nu\lambda\beta} \right\}, \tag{3.31}$$

where the superscripts (0), (1,1), and (0,1) refer to the orders with respect to the external field and nuclear moments, respectively and α and β refer to the x, y, or z contribution.

The greek letter subscripts indicate quantities evaluated over basis functions and P is the density matrix. The first term involves the Hamiltonian that is first order in both the applied field and magnetic moments:

$$\begin{aligned} H^{(1,1)}_{J\nu\lambda\alpha\beta} = \left(\frac{1}{2c^2}\right) \{ & \langle \phi_\nu | \mathbf{r}_\lambda \cdot \mathbf{r}_\mathbf{J} \delta_{\alpha\beta} - \mathbf{r}_{\lambda\alpha} \mathbf{r}_{\mathbf{J}\beta} | \phi_\lambda \rangle \\ & + (\mathbf{R}_\nu \times \mathbf{R}_\lambda)_\alpha \left\langle \phi_\nu | L_{J\beta} r_J^{-3} | \phi_\lambda \right\rangle \\ & + \left\langle (\mathbf{R}_{\nu\lambda} \times \mathbf{r}_\nu)_\alpha \phi_\nu | L_{J\beta} r_J^{-3} | \phi_\lambda \right\rangle \}, \end{aligned} \tag{3.32}$$

where $\mathbf{r}_\lambda = \mathbf{r} - \mathbf{R}_\lambda$. The first term in $H^{(1,1)}_{J\nu\lambda\alpha\beta}$ is the diamagnetic contribution to the shielding. The second and third term are the paramagnetic contributions to the shielding. The density matrix in the first term of Eq. 3.31 is obtained from the zeroth-order coefficients.

The second contribution to the shielding involves both the Hamiltonian first order in the magnetic moments and the density matrix first order in the applied field. This contribution is also paramagnetic. The expressions for these are as follows [71–74]:

$$P^{(1,0)}_{\nu\lambda\alpha} = 2 \sum_{j}^{occ} [c^{(0)}_{\nu j} c^{(1,0)}_{\lambda j\alpha} - c^{(1,0)}_{\nu j\alpha} c^{(0)}_{\lambda j}] \tag{3.33}$$

$$H^{(0,1)}_{J\nu\lambda\beta} = -\left(\frac{1}{c}\right) < \phi_\nu | L_{J\beta} r_J^{-3} | \phi_\lambda > . \tag{3.34}$$

To evaluate the density matrix contributions $P^{(1,0)}_{\nu\lambda\alpha}$, the first-order coefficients are needed. These can be obtained by solving the coupled perturbed HF equations that are first order in the applied field [71–74].

One further point is worth mentioning. The gauge invariance is no longer a problem

when using GIAOs as in the evaluation of the integrals required for both the overlap and Fock matrix elements as the gauge origin dependence disappears.

3.3.2 The Continuous Set of Gauge Transformations Approach

The *continuous set of gauge transformations* [75,79], or CSGT, is similar to the GIAO method for obtaining nuclear magnetic shieldings. It also employs a set of gauge origins to remove the dependency of the calculation on the gauge origin. Whereas the GIAO adds a gauge origin to the basis functions, the CSGT uses a separate gauge origin for each point in the simulation space.

The CSGT method begins by expressing the shielding in terms of the first-order current density [75, 79]

$$\begin{aligned}\mathbf{J}^{(1)}(\mathbf{r}) = -\frac{eN}{m}\int d\tau'(\psi^{(0)*}\mathbf{p}\psi^{(1)} + \psi^{(1)*}\mathbf{p}\psi^{(0)}) \\ - \frac{e^2}{mc}\mathbf{A}(\mathbf{r})\rho^0(\mathbf{r}).\end{aligned} \tag{3.35}$$

In Eq.3.35, the $\psi^{(0)}$ represents the ground-state wavefunction and the $\psi^{(1)}$ is the first-order correction to the wavefunction. $\rho^{(0)}(\mathbf{r})$ represents the unperturbed electron density. e, m, N, and c represent the negative value of the charge on the electron, the mass of the electron, the number of electrons, and the speed of light, respectively. If the wavefunction $\psi^{(0)}$ and its first-order correction $\psi^{(1)}$ are written as a determinants of doubly-occupied

molecular orbitals, the first-order current density is given by

$$\begin{aligned}\mathbf{J}^{(1)}(\mathbf{r}) = &-\frac{2e}{m}\sum_{i=1}^{N/2}(\phi_i^{(0)*}\mathbf{p}\phi_i^{(1)} + \phi_i^{(1)*}\mathbf{p}\phi_i^{(0)}) \\ &- \frac{e^2}{mc}\mathbf{A}(\mathbf{r})\rho^0(\mathbf{r}).\end{aligned} \tag{3.36}$$

The first term in Eq.3.36 is the paramagnetic contribution to the first-order current density and the second term is the diamagnetic contribution. The first-order corrections to the wavefunction are written as an expansion in the unperturbed virtual molecular orbitals as [75, 79]

$$\phi_i^{(1)} = \sum_{p=N/2+1} c_{pi}^{(1)}\phi_p^{(0)}. \tag{3.37}$$

The first-order coefficients are obtained as solutions to the coupled perturbed equations. When HF is the method of choice, these correspond to the CPHF equations.

The vector potential $\mathbf{A}(\mathbf{r})$ can be subjected to a gauge transformation given by the expression $\mathbf{A}'(\mathbf{r}) = \mathbf{A}(\mathbf{r}) + \nabla f(\mathbf{r})$. This transformation corresponds to an application of a unitary transformation to the wavefunction. This unitary transformation is given as a phase change. HF calculations are invariant to this transformation as long as a complete basis set is used. However, this is almost always not the case, so the calculations are gauge-dependent. In the GIAO method, this is solved by applying a phase factor to the orbitals which is field-dependent. Another possible way of insuring gauge-invariance is to apply the phase change as a perturbation to the CPHF equations and obtain corrections to the

first-order wavefunctions. These corrections represent a perturbation in the system due to the gauge transformation. Thus, solutions are obtained for the first-order coefficients first for the angular momentum operator for a single gauge origin. The solutions are then obtained for a shift in the gauge origin by solving the CPHF for the components of the linear momentum operator. This corresponds to the solutions to the CPHF equations for any gauge origin [75, 79].

Application of the gauge transformation to the current density and applying a magnetic field along the x-axis gives for the current density

$$\begin{aligned}\mathbf{J}^{(1)}(\mathbf{r}) = \frac{2e}{m}\sum_{i=1}^{n/2}(Be/2mc)[(\phi_i^{(0)*}\mathbf{p}\phi_i^{(1)L_x} + \phi_i^{(1)*L_x}\mathbf{p}\phi_i^{(0)}) \\ - d_y\{\mathbf{r}\}(\phi_i^{(0)*}\mathbf{p}\phi_i^{(1)p_z} + \phi_i^{(1)*p_z}\mathbf{p}\phi_i^{(0)}) \\ + d_z\{\mathbf{r}\}(\phi_i^{(0)*}\mathbf{p}\phi_i^{(1)p_y} + \phi_i^{(1)*p_y}\mathbf{p}\phi_i^{(0)})] \\ - \frac{e^2}{mc}\mathbf{B}\times(\mathbf{r}-\mathbf{dr})\rho^0(\mathbf{r}).\end{aligned} \tag{3.38}$$

The superscripts L_i and p_i, where i is x, y, or z, refer to components of quantities over the angular and linear momentum operators, respectively. The quantities $d_y\{\mathbf{r}\}$ and $d_z\{\mathbf{r}\}$ are components of the function $\mathbf{d}\{\mathrm{r}\}$, which is used to specify the set of gauge transformations.

The magnetic shielding tensor for a nucleus N_J is related to the first-order current density by

$$\boldsymbol{\sigma}_{N_J} = -\frac{1}{Bc}\int(\mathbf{r}_{N_J}\times\mathbf{J}^{(1)}(\mathbf{r})/r_{N_J}^3). \tag{3.39}$$

In the GIAO method, the gauge origins are the nuclear centers and these are included in the basis functions. However, in CSGT, the basis functions are written in terms of the field-independent functions. The gauge origins are placed at points throughout the simulation space. This is done through the use of a function $\mathbf{d}\{\mathrm{r}\}$ to represent a continuous set of gauge transofrmations. This function was chosen to give the most accurate first-order current density distribution for the molecule. In the original CSGT implementation, this function was defined as an exponential in r

$$d\{\mathbf{r}\} = \mathbf{r} - \sum_{\sigma}(\mathbf{r} - \mathbf{R}_{\sigma})exp[-\alpha_{\sigma}(\mathbf{r} - \mathbf{R}_{\sigma})^4], \tag{3.40}$$

where the σ represent the nuclei near $\mathbf{r}$, and the α_{σ} values are empirical parameters chosen to guarantee the current distribution's symmetry properties are obtained. Three reasons for this choice of $\mathbf{d}$ $\{\mathrm{r}\}$ were given. First, description of the current density with the gauge origin near the nucleus is accurate. Second, the current density in the nonnuclear regions are also well described with the choice of gauge origin above. Finally, the simplicity of the function allows for the shielding and magnetic susceptibility to be calculated in a straightforward manner [75].

3.4 Improving on the HF Approach

It is often the case that inclusion of electron correlation in the prediction of energies and geometries improves their results over HF theory [80]. As was previously shown, this is also true for NMR chemical shifts and shielding values [69,81,82]. In these works, ^{13}C,

^{15}N, ^{17}O, and ^{31}P chemical shifts for a variety of systems calculated with GIAO-HF and GIAO-MP2 are discussed. These authors show that in most of the cases studied, the shieldings improve significantly when electron correlation is included at the MP2 level of theory. It is suggested that the origins of the improvements are due to the sensitivity of the paramagnetic portion of the shielding to the molecular geometry. These studies also suggest that in general, systems with multiple bonds or lone pairs benefit most from the inclusion of electron correlation. This may be due to the decrease in the HOMO-LUMO gap in these cases where mixing with lower lying excited states is possible.

An excellent review on electron-correlated approaches to the calculation of chemical shieldings is given in Ref.69. Methods for including correlation currently available for NMR chemical shifts are MP2, CCSD/CCSD(T), and DFT. The MP2 and DFT methods will be discussed at some length in this section.

3.4.1 MP2-NMR

As previously stated, the magnetic shieldings are generated by differentiating the electronic energy with respect to the magnetic field and the nuclear magnetic moments. Shieldings obtained from MP2 are derived from differentiation of the MP2 energy as [69, 81, 83]

$$\begin{aligned} \sigma_{J,ji} &= \frac{\partial^2 E_{MP2}}{\partial B_{J,i} \partial m_{J,j}} \\ &= \sum_{\mu\nu} P_{\mu\nu} \frac{\partial^2 h_{\mu\nu}}{\partial B_i \partial m_{J,j}} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial B_i} \frac{\partial h_{\mu\nu}}{\partial m_{J,j}} \end{aligned} \tag{3.41}$$

where the second expression gives the MP2 shieldings in the AO basis whereas the third expression gives the MP2 shieldings in the MO basis. In the above equation for the MO basis, the $h_{pq}^{\mu_{\alpha j} B_{\beta J}}$ represent the second derivatives of the one-electron integrals and the $h_{pq}^{\mu_J}$ represent the first derivatives with respect to the nuclear magnetic moment μ_J. The P_{pq} are blocks of the density matrix where the p, q, r, ... can represent orbitals that are either occupied or virtual. The $U_{pr}^{B_i *}$ and $U_{pr}^{B_i}$ are the first-order coefficients of the wavefunction and are obtained as solutions to the CPHF equations. The details of the derivation of the shielding expression and the origins of the matrix elements within it are described in the literature [69, 81, 83].

Figure 3.2 shows the steps taken in the calculation of the shieldings within the GIAO-MP2 approach. The first step is to generate the one- and two-electron integrals within the GIAO approach. These integrals are required for the solution of the SCF equations. This step is done in an analogous manner to the Roothaan equations in HF theory.

The integrals obtained with GIAO differ from those obtained with the usual field-independent basis functions only in the treatment of the field-dependent phase factor $exp(-(i/c)\mathbf{A}_\nu\mathbf{r})$ [84].

The second step involves a transformation of the two-electron integrals from the AO to the MO basis. The two-electron integrals are needed to evaluate the MP2 second-order energy correction. The integrals calculated for use in the modified Roothaan equations are expressed in the AO basis. This transformation is a necessary step as the two-electron integrals needed in the MP2 second-order energy correction are expressed in the MO basis.

From these transformed integrals, the contributions to the MP2 energy are evaluated. This evaluation is followed by the computation of the MP2 density matrix.

The next step is to evaluate the first and second derivatives of the one- and two-electron GIAO integrals with respect to the magnetic field and nuclear moments. These integrals are required to evaluate the magnetic shieldings. From Eq.3.41, the diamagnetic contributions to the SCF and MP2 shieldings are evaluated by multiplying the second derivatives of the GIAO integrals by their respective SCF and MP2 density contributions. This gives, respectively, the SCF and MP2 diamagnetic shielding contributions. The remaining integrals are kept as they are needed to evaluate paramagnetic contributions to the shielding [69, 81, 83].

To evaluate the paramagnetic contributions to the SCF and MP2 shieldings, the derivative of the density matrix with respect to the magnetic field is needed. For the MP2 shieldings, this corresponds to the second term in the third expression of Eq.3.41. To obtain these contributions, the first-order coefficients are also needed, and these can

be obtained from the solutions to the CPHF equations. This is thus the next step in the algorithm. For the MP2 shieldings, the two-electron integrals are left in the MO basis and the CPHF equations are also solved in the MO basis. This is done as the integral contributions needed for MP2 shieldings are represented in the MO basis. It is after the CPHF equations are solved that the paramagnetic contributions to the SCF shieldings are generated. For the SCF shieldings, the CPHF equations are solved in the AO basis.

The final step in the calculation of the MP2 magnetic shieldings is to obtain the paramagnetic contributions to the MP2 shieldings. There are several stages to this last step of the computation of the paramagnetic contribution to the shielding. These will not be covered here, but can be found in ref. [83]. The overall result is to transform the AO representation of the GIAO two-electron integral derivatives to the MO basis and calculate, in stages, the various contributions to the MP2 paramagnetic shielding. These contributions are then added to the shielding contributions previously calculated to yield the final shieldings for the nuclei in the system [69, 81, 83].

3.4.2 DFT-NMR

DFT provides a cheaper way to include electron correlation in NMR without resorting to perturbation theory. As such, it is the method of choice when calculating chemical shifts on large systems such as transition metal complexes and proteins. There are a few methods of obtaining NMR shieldings using DFT, however, only the GIAO and CSGT methods will be discussed.

DFT-GIAO

The computation of NMR shieldings with DFT using GIAO's is similar in spirit to the HF GIAO-NMR approach previously discussed. Therefore, the complete derivation of the shielding will not be discussed. Also, no introduction into the underlying theory of DFT will be given as it was discussed in Chapter 1.

From Chapter 1, properties of interest are obtained once the solutions to the Kohn-Sham equations

$$h^{KS}\Psi_i^{KS} = \epsilon_i^{KS}\Psi_i^{KS} \tag{3.42}$$

are known. In Eq.3.42, the system is assumed to be field-free. When a magnetic field is applied, the above set of equations must be expressed in terms of the dependence on the magnetic field. This is done by modifying the KS equations (and thus the energy) to include this field dependence. If one starts with the energy expression, the total electronic energy is expressed in terms of the applied field as [79, 85, 86]

$$E_{el,\mathbf{B}}^{KS} = E_{kin,\mathbf{B}} + E_{ne,\mathbf{B}} + E_{ee,\mathbf{B}} + E_{xc,\mathbf{B}}[\rho, \mathbf{J}], \tag{3.43}$$

where the first term is the kinetic energy of the electrons in the magnetic field, the second term is the field-dependent nuclear-electron attraction, the third is the field-dependent Hartree term (density interacting with itself), and the fourth is the exchange-correlation term. The exchange-correlation energy functional in the presence of an applied magnetic

field is now dependent on the current density $\mathbf{J}$ as well as the electronic density ρ. However, it is possible to make the approximation that the current density contribution to the exchange-correlation functional can be neglected. Use of this approximation leads to the *uncoupled* density functional theory of NMR [79, 85, 86].

Inclusion of the approximation for the current density-independent exchange-correlation functional and application of a magnetic field gives the modified KS Hamiltonian

$$h^{KS}(\mathbf{B}) = h_0^{KS} + h_1^{KS} + h_2^{KS}, \tag{3.44}$$

where the first term represents the field-free Hamiltonian, the second term represents the first-order correction to the Hamiltonian due to the applied field, and the third term represents the second-order correction to the Hamiltonian. A discussion of the origins of the first and second-order terms is given earlier in this chapter.

From Eqs. 3.42 and 3.43, and noting that the wavefunction Ψ_i^{KS} now depends on the magnetic field gives the modified KS equations [79, 85, 86]

$$h^{KS}(\mathbf{B})\Psi_i^{KS}(\mathbf{B}) = \epsilon_i^{KS}(\mathbf{B})\Psi_i^{KS}(\mathbf{B}), \tag{3.45}$$

where the $\Psi_i^{KS}(\mathbf{B})$ are the field-dependent solutions to these modified KS equations. They

can be expressed as a linear combination of field-dependent basis functions

$$\Psi_i^{KS}(\mathbf{B}) = \sum_a c_{ai}(\mathbf{B})\phi_i^{KS}(\mathbf{B}), \tag{3.46}$$

where the $\phi_i^{KS}(\mathbf{B})$ are given by Eq.3.28.

Following the method used to obtain the shieldings for HF with GIAO's, the expression for the shielding is

$$\begin{aligned} \sigma_{N_J} &= \sigma_{N_J}^d + \sigma_{N_J}^p \\ &= \left\langle \mathbf{h}^{(1,1)}\mathbf{P}^{(0,0)} \right\rangle + \left\langle \mathbf{h}^{(0,1)}\mathbf{P}^{(1,0)} \right\rangle, \end{aligned} \tag{3.47}$$

where the superscripts (0,0), (0,1), and (1,1) denote the orders of the derivatives with respect to the applied magnetic field and nuclear moments, respectively. Also, the d and p refer to the diamagnetic and paramagnetic shielding contributions, respectively. The $\mathbf{P}$ in the above equation is the density matrix, whose elements are given by [79,85,86]

$$\mathbf{P}_{\mu\nu} = \sum_a^{N/2} c_{\mu a} c_{\nu a}^*, \tag{3.48}$$

and the c's are the Kohn-Sham orbital coefficients. The derivatives of the Hamiltonian are

given by

$$\mathbf{h}_{\mu\nu}^{(1,1)} = \left\langle \chi_\mu \left| \frac{1}{2c} \frac{\mathbf{r}\cdot(\mathbf{r}-\mathbf{R}_{N_J})\delta_{ij} - \mathbf{r}_i(\mathbf{r}-\mathbf{R}_{N_J})_j}{|\mathbf{r}-\mathbf{R}_{N_J}|^3} \right| \chi_\nu \right\rangle \tag{3.49}$$

$$\mathbf{h}_{\mu\nu}^{(0,1)} = \left\langle \chi_\mu \left| -\frac{i}{c} \frac{[(\mathbf{r}-\mathbf{R}_{N_J})\times\boldsymbol{\nabla}]_j}{|\mathbf{r}-\mathbf{R}_{N_J}|^3} \right| \chi_\nu \right\rangle . \tag{3.50}$$

Similar to the HF GIAO-NMR method, the first step is to solve the uncoupled modified KS equations to obtain the zeroth-order KS orbitals. From these, the diamagnetic part of the shielding can then be calculated. In order to calculate the paramagnetic contribution to the shielding, the first-order corrections to the KS orbitals are needed. These are obtained by solving the coupled perturbed KS (CPKS) equations

$$\mathbf{F}\mathbf{P}_{ov}^{(1,0)} - \mathbf{P}_{ov}^{(1,0)}\mathbf{F} - \mathbf{G}(\mathbf{P}_{ov}^{(1,0)} + \mathbf{P}_{vo}^{(1,0)})_{ov} = \mathbf{h}_{ov}^{(1,0)} + \mathbf{G}_{ov}^{(1,0)}(\mathbf{P})_{ov} - \mathbf{F}\mathbf{S}_{ov}^{(1,0)} + \mathbf{G}(\mathbf{S}_{ov}^{(1,0)})_{ov} \tag{3.51}$$

where the subscript ov refers to the occupied-virtual block of the density matrix $\mathbf{P}$. The solutions to the CPKS give the first-order density matrix elements for the occupied-virtual block. The occupied-occupied and virtual-virtual blocks of the first-order density matrices are computed in a straightforward manner by [79, 85, 86]

$$\mathbf{P}_{oo}^{(1,0)} = -\mathbf{S}_{oo}^{(1,0)} \tag{3.52}$$

$$\mathbf{P}_{vv}^{(1,0)} = 0. \tag{3.53}$$

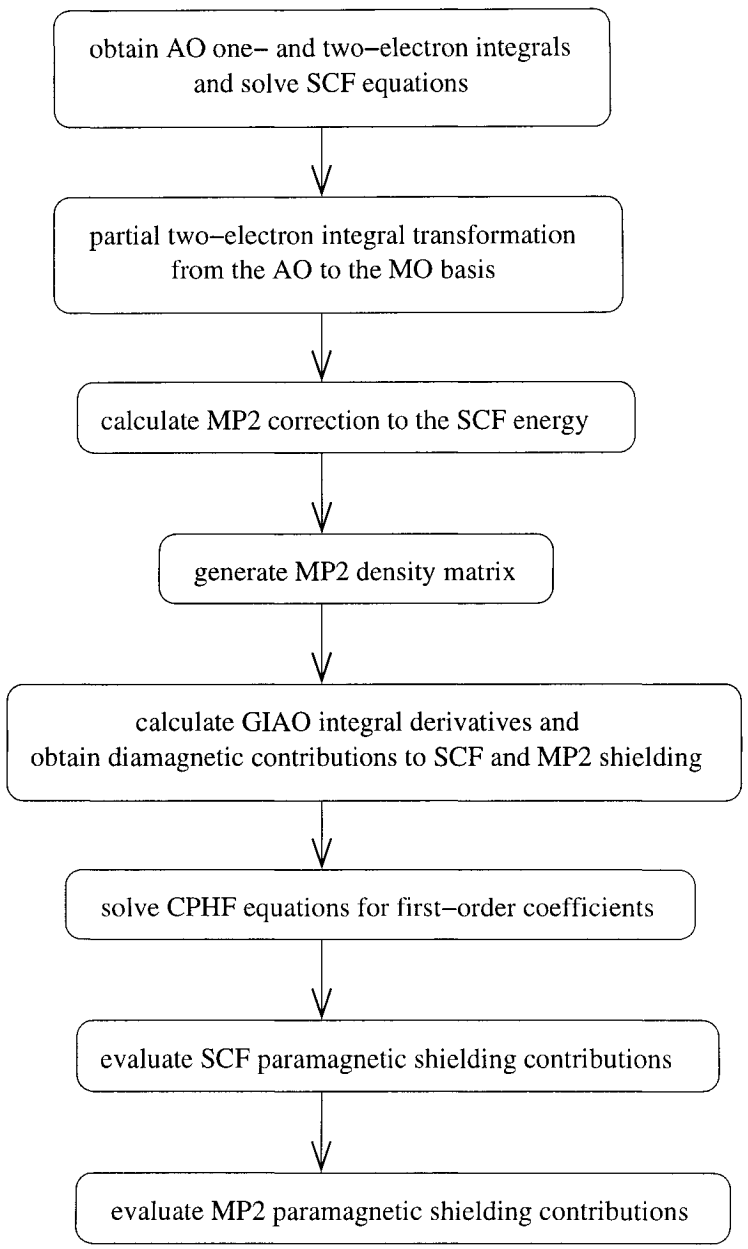


Figure 3.2: Steps taken in the computation of the MP2-GIAO shieldings.

Eq.3.51 depends on the derivatives of the overlap matrix, one-electron Hamiltonian, and two-electron integrals with respect to the magnetic field. These derivatives over GIAO's are given in Ref.79. It should be noted that although the exchange-correlation term of the Kohn-Sham matrix **F** is assumed to be dependent only on the electronic density of the system, there is a field dependence through the use of the GIAO's. As a result, the gradients of the exchange-correlation term of the KS matrix are needed [79,85,86].

The procedure for evaluating the NMR shieldings is analogous to the procedure for the HF shielding tensors, and is summarized in Fig. 3.3.

DFT-CSGT

As for the DFT-GIAO approach, application of DFT with CSGT to compute magnetic shielding tensors is analogous to the application of HF with CSGT. In the case of DFT, the current density is computed from the Kohn-Sham orbitals obtained as solutions of the Kohn-Sham equations. The first-order components of the density matrix are obtained by solving the CPKS equations for the three components of the perturbation in the angular momentum, giving the first-order coefficients. The equations are also solved for the three components of the perturbation in the linear momentum. These are used to obtain the first-order current density. The magnetic shieldings are then obtained via numerical integration of the right hand side of Eq.3.39 [75,79]. These steps are summarized in Fig. 3.4.

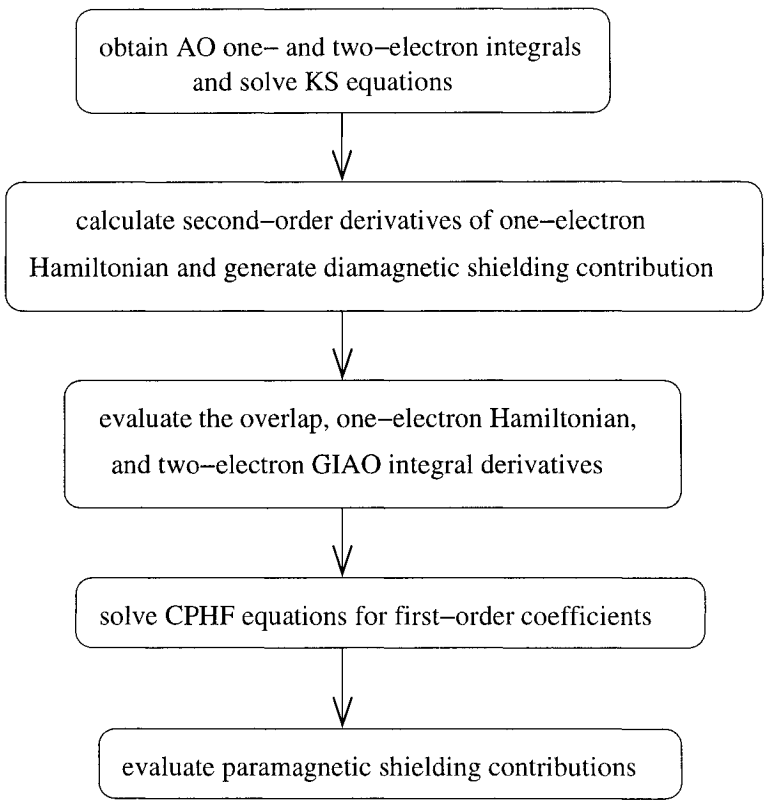


Figure 3.3: Diagram for the steps of a DFT-GIAO calculation of the NMR shieldings.

3.5 Comparisons With Experiment - Practical Considerations

Theoretical NMR calculations correspond to chemical shieldings at a fixed geometry. For structures calculated by geometry optimization, this is the equilibrium structure. This structure can differ significantly from the experimental one, which is averaged over the rotations and vibrations within the molecule. Also, theoretical values of shielding are computed for the individual molecule at the zero-point vibrational level. Unless explicitly included, intermolecular and solvent effects are also not included [64, 87–90].

Experimental NMR chemical shifts are not computed for individual molecules. Experiments are run at temperatures well above zero Kelvin in solution, solid state, neat liquid, or gas phase. To be able to compare theoretical values to experiment, rovibrational and intermolecular effects must be taken into account. Intermolecular effects are expected to be small for most neat liquids, however solvent effects are expected to be significant. Corrections due to rotation and vibration as well as solvent and intermolecular effects must be added to theoretical shieldings or the experimental results must be corrected to correspond to the theoretical values.

3.5.1 Rovibrational Contributions - Adding Temperature Effects

Theoretical shieldings are computed for molecules at a fixed geometry. The molecule's geometry is either input or corresponds to the equilibrium value obtained from an optimization. However, in an experiment, the average structure at a given temperature is

what corresponds to the chemical shifts observed. This average structure can be close to its equilibrium counterpart, but often this is not the case. The experimental structure is an average over all the possible structures which result from rotations and vibrations within the molecule. These effects cause the electronic structure around the nucleus to change. Since the shielding is a sensitive probe of the electronic structure at a given nucleus, it also changes. This results in experimental shielding values (and thus chemical shifts) which can be considerably different from those obtained theoretically. Corrections due to rotations and vibrations can now be computed using a variety of theoretical methods. They can be added to a calculated absolute shielding value and compared to experiment [87–89].

Corrections due to rovibrational effects must be averaged over the rotational and vibrational states available to the molecule at a given temperature. This requires both the vibrational wavefunction and the variations in the shielding with geometry. If a vibrational wavefunction is available or can be generated, corrections to the shielding can be calculated with the following steps:

- using the Born-Oppenheimer potential, calculate rovibrational energy and wavefunction for the state
- calculate shielding constant for equilibrium geometry
- average shielding over rovibrational wavefunction
- repeat the first three steps for many states
- take a Boltzmann-weighted average of the shielding over these states

Table 3.1 lists a few molecules and the rovibrational corrections for the given nuclei. The larger corrections are seen for the heavier nuclei ^{19}F and ^{17}O.

Table 3.1: Rovibrational corrections for a few nuclei in selected molecules.

molecule	nucleus	(σ (300K) - σ (r_e)) [a] (ppm)
F_2	^{19}F	-27.22
N_2	^{14}N	-3.47
N_2	^{15}N	-3.48
CO	^{13}C	-1.80
CO	^{17}O	-4.69

[a]Data from Ref. 88.

3.5.2 Solvent Effects

In an experiment, the molecule of interest is not alone in a vacuum. When the molecule of interest is in solution or part of a larger sample, then the surrounding molecules will have an influence on its electronic structure. This can occur directly through some type of weak bonding interaction. It can also occur through distant interactions. If the neighbouring molecules give rise to a weak electric field, then this will alter the electron density at the nucleus of interest slightly, resulting in an increase or decrease in shielding about the nucleus. This change in the nucleus' shielding results in the chemical shift being different than for the isolated molecule.

As a result of the intermolecular effects, a nucleus in a molecule can have different shielding values depending on the solvent it is placed in. This shielding, again, will differ from that of the isolated atom. Comparisons to experiment must take into account this solvent effect. Theoretical calculations often employ some type of continuum model to

account for the effects of the solvent. The continuum models do not include solvent structure, but actually mimic the effect of the solvent using a point charge model. Since this point charge model does not explicitly include solvent structure, the resulting shieldings calculated using it may not be accurate. However, calculations using explicit solvent molecules are expensive, so continuum models provide an estimate of the solvent effects on the calculated shielding values [90].

Table 3.2 lists some shielding values calculated at the B3LYP/6-311++G** level of theory using the polarizable continuum model available in Gaussian 03 [91]. For all the nuclei listed, a small solvent effect is seen in the absolute shieldings depending on the solvent used.

Table 3.2: Calculated shieldings, in ppm, of a few nuclei in selected molecules in solution.

molecule	nucleus	$\sigma_{isolated}$ (ppm) [a]	σ_{water} (ppm)	$\sigma_{cyclohexane}$ (ppm)
F_2	^{19}F	-248.2	-240.1	-244.0
N_2	^{14}N	-79.8	-75.1	-78.7
CO	^{13}C	-11.3	-10.2	-11.0
CO	^{17}O	-72.3	-66.8	-70.7
PH_3	^{31}P	562.9	571.7	564.2

[a] Shaw, D. M., unpublished data (for all shieldings in this table). No experimental results for the nuclear shieldings in water and cyclohexane could be found.

3.6 Conclusion

Experimental NMR is a useful tool for structural determination in many areas of chemistry. Due to its sensitivity in probing the electronic structure around a nucleus, it has become a widely used technique. Chemical shieldings are available for a variety of nuclei using methods such as HF, MP2, and DFT. For light nuclei, the theoretical results

are comparable to experiment. However, for heavier nuclei, some work still needs to be done to include corrections for relativistic and rovibrational effects as well as solvent. As these corrections improve, theoretical chemical shieldings for the heavier nuclei will also compare well with experiment.

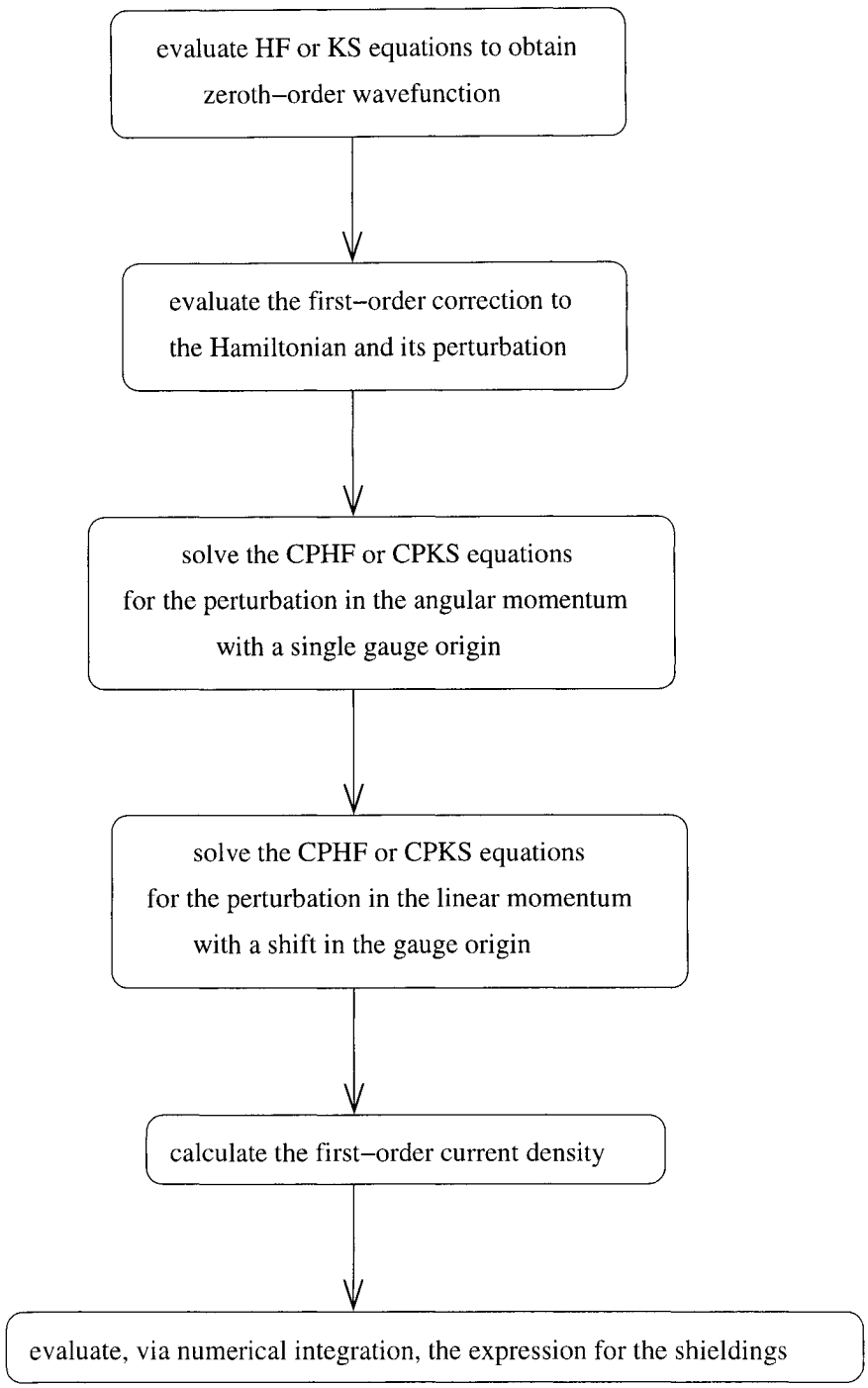


Figure 3.4: Diagram for the steps of a DFT-CSGT calculation of the NMR shieldings.

Chapter 4

Pentacoordinate Phosphorus Models: Background

4.1 Introduction

Phosphoryl transfer enzymes catalyze the transfer of a phosphate group to either a water or another group. These enzymes are responsible for a variety of important functions in living organisms. Some of these include production of second messengers [92, 93] and RNA cleavage [94]. Understanding the mechanism of phosphoryl transfer is essential to understanding how these enzymes work.

There is a significant amount of literature available which is based on experimental studies of phosphoryl transfer reactions. From these studies, three reaction pathways have been proposed. Although the experimental results are numerous, there is still some question over the exact mechanism of phosphoryl transfer in the systems studied.

Theoretical investigations have been conducted for phosphoryl transfer of some phosphate monoesters and diesters. They have been useful in providing insight into the mechanisms of phosphoryl transfer both in solution and in the gas phase. However, most of these studies concentrate mainly on reproducing geometries and energetics of these reactions. Little work has been done recently on investigating the quality of the geometries and spectroscopic results for the transition states and intermediates involved in these reactions. This is especially true for theoretical vibrational frequencies and ^{31}P NMR absolute shielding results.

To be successful in describing the mechanisms of phosphoryl transfer in enzymes, theoretical studies must be able to describe phosphoryl transfer mechanisms in general. This involves successfully describing any intermediates and transition states along the reaction pathway. Therefore, a proper description of pentacoordinate phosphorus is essential to the study of phosphoryl transfer in enzymes.

4.2 Pentacoordinate Phosphorus

4.2.1 General Structural Characteristics

Previous studies on pentacoordinate phosphorus compounds indicate it adopts a trigonal bipyramidal structure, as shown in Fig. 4.1. In this structure, the axial bonds are longer than the equatorial bonds. Ideal angles between axial ligands, equatorial ligands, and those spanning axial and equatorial ligands are 180°, 120°, and 90°, respectively. There is some variation in these angles for various groups attached to phosphorus. In the

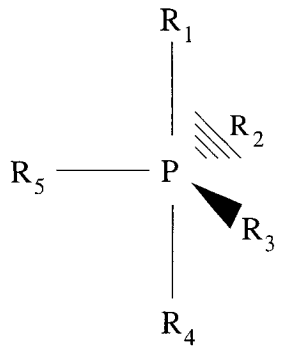


Figure 4.1: Diagram of an ideal trigonal bipyramid.

case of phosphorus participating in a ring, the ring usually spans the axial and equatorial positions [1, 95].

The trigonal bipyramid is the structure adopted by acyclic phosphorus compounds and those in which phosphorus participates in a single ring. However, some pentacoordinate phosphorus compounds adopt a square pyramidal structure, shown in Fig. 4.2, where four of the five ligands form a square plane and the fifth ligand lies directly above the center of the plane. In this structure, the bond to the fifth ligand is slightly longer than the bonds to the other four ligands (nomenclature refers to these ligands as axial and basal, respectively). The axial-basal and basal-basal angles in a square pyramid are roughly 104° and 87°, respectively [1, 95].

The trigonal bipyramid and square pyramid are the only two geometrical structures observed for pentacoordinate phosphorus. A previous investigation of energy contributions to both structures indicates that energetically the two are similar. However, in most cases the trigonal bipyramid is more favored [96].

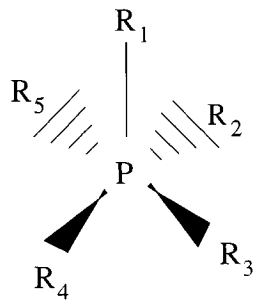


Figure 4.2: Diagram of a square pyramidal geometry adopted in pentacoordinate phosphorus compounds.

Trends in Bonding Environments at Pentacoordinate Phosphorus

Considerable work has been done on experimental investigations of structures of pentacoordinate phosphorus molecules [1, 97]. Studies included both cyclic and acyclic systems. Among the experimental work are x-ray structures, microwave and vibrational spectroscopy, and NMR results.

The experimental results show a variety of trends. Bond lengths and angles from x-ray diffraction or microwave studies conducted on acyclic pentacoordinate phosphorus compounds indicate they adopt a trigonal bipyramidal structure. For identical ligands occupying both the axial and equatorial positions, the axial bonds are longer than the equatorial bonds. The more electronegative atoms and groups reside in the axial positions.

In the fluorophosphorane series, substitution of bulkier and less electronegative groups into the equatorial positions of the trigonal bipyramid causes the structure to distort. Both bond lengths and angles are altered. Distortion in the F_{ax}-P-F_{ax} angle can be as much as 4°. In the case of monosubstituted fluorophosphoranes, the angle between the two equatorial fluorines is compressed by about 4 degrees. However, as a second methyl group

replaces an equatorial fluorine atom, the angle between the two methyl groups increases by 5°. This distortion can be explained by applying the VSEPR theory. Overall repulsive forces are smaller if the methyl groups are placed equatorially instead of axially. However, repulsion between the bonding electrons in the methyl groups and the lone pairs on the axial fluorines causes the axial fluorines to be pushed away from the methyl groups. This results in the observed decrease in the angle between the equatorial methyl groups as well as the decrease in the angle between the axial fluorines [1].

Structural experiments show that P-X and P-F bond lengths in a series PX_nF_{5-n} increase upon substitution of a less electronegative group.

Table 4.1: Bond lengths for a series of fluorophosphoranes (from Ref. 1).

molecule	r_{ax}(P-F, Å)	r_{eq}(P-F, Å)	r_{eq}(P-C, Å)
PF_5	1.577	1.534	-
PF_4CH_3	1.612	1.543	1.780
$PF_3(CH_3)_2$	1.643	1.553	1.798
$PF_2(CH_3)_3$	1.685	-	1.813

Table 4.1 shows the bond lengths for a series of methyl substitutions on pentafluorophosphorane. What is observed is both the P-F and P-C bond lengths increase as methyl groups replace fluorines. Since the methyl group is less electronegative, this would imply that substitution of less electronegative groups increases axial and equatorial bond lengths. In most cases, this is what is observed. Starting from the bottom of the table and moving up, the methyl groups are replaced by fluorines. In this case, bond lengths decrease. Therefore, one would expect that bond lengths would decrease upon substitution of more electronegative groups.

Spectroscopic analysis of substituted fluorophosphoranes and chlorophosphoranes indicates the symmetry group to which they belong. This information indicates the positioning of the substituted groups. It can also be used to determine if there is free rotation about bonds to substituents.

Table 4.2: Space groups for a series of fluorophosphoranes.

molecule	space group
PF_5	D_{3h}
PF_4CH_3	C_{2v}
PF_4CF_3	C_{2v}
$PF_3(CH_3)_2$	C_{2v}
$PF_2(CH_3)_3$	D_{3h}
PCl_5	D_{3h}
PCl_4CF_3	C_{3v}
$PCl_3(CF_3)_2$	D_{3h}

Table 4.2 gives the space groups for a series of fluorophosphoranes and chlorophosphoranes. For the fluorophosphoranes, the symmetry given can only occur if the CF_3 and CH_3 groups are placed equatorially and the hydrogens and fluorines of these groups rotate freely (they are therefore equivalent). This group placement is consistent with the electronegativities of the groups relative to fluorine. Since they are less electronegative, they will occupy equatorial positions. For the chlorophosphoranes, the CF_3 groups are free to rotate, but they now occupy axial positions, in accord with the observed C_{3v} symmetry. This is also in agreement with the placement of the groups according to their electronegativity [1].

In pentacoordinate phosphorus molecules containing one or more rings, geometrical parameters indicate a trigonal bipyramidal structure. The more electronegative substituents occupy the axial positions, even when these groups occur within the rings. The

atoms in the ring that are attached to phosphorus usually occupy the axial and equatorial positions. In a select number of cases where the molecule contains two rings including phosphorus, the preferred geometry is a square pyramid.

For trigonal bipyramids containing cycles, distortions are evident from the observed geometrical parameters. The resulting structure lies somewhere between trigonal bipyramidal and square pyramidal. As a result, the coordinate that gives the distortion follows a path of conversion from the trigonal bipyramid to the square pyramid. Where along this path the distortion occurs depends on steric and electronic effects as well as ring strain. Some electronic effects that influence the distortion path include substituent electronegativity, electron delocalization, and repulsion between pairs of electrons. The molecule will try to reduce any energetic penalty due to these effects. However, in the majority of cases with small rings (4-7 atoms), the positioning of ring atoms in an axial and equatorial position is more significant than the electronegativities of the substituents. Larger rings can have two atoms in them occupy two equatorial positions [1, 97].

The observation that some groups prefer the axial positions over equatorial positions in pentacoordinate phosphorus was investigated by Holmes [98]. He focused on a variety of cyclic and acyclic systems, many of them containing highly electronegative groups attached to phosphorus. He developed an empirical model which could be used to predict the preference of an axial position by a group, also known as *apicophilicity*. These empirical results were compared with both x-ray structures and NMR exchange spectra. For the acyclic phosphoranes, the model was based on electronegativities of the substituents and steric effects. For the cyclic phosphoranes, ring strain effects were also investigated and

taken into account in the empirical model. For the experimental results given, the model gave accurate predictions. However, to obtain a better model, experimental results for a variety of systems containing less electronegative or bulkier substituents would need to be added. However, especially for acyclic systems, this is not possible as these systems are too unstable to study experimentally. As a result, more current predictions and explanations for apicophilicity are given by theoretical investigations.

The experimental results indicate trends in bonding at phosphorus. However, these trends depend highly on the systems studied, and there are a number of interesting pentacoordinate phosphorus molecules that are too unstable to study experimentally. This is especially true for some models of reaction intermediates and transition states in enzyme-catalyzed phosphoryl transfer reactions. Theoretical models, capable of reproducing experimental trends, have thus become an attractive method of explaining results of experiment. They have also provided a method of investigating systems too unstable to study with experiment.

4.2.2 Theoretical Models of Trigonal Bipyramidal Phosphorus Compounds

Investigations of bonding at trigonal bipyramidal phosphorus compounds have been conducted for over 40 years. The more studied of these compounds include the halophosphoranes PF_5 and PCl_5, although treatment has been extended to a variety of systems. Of these, studies involving halogen-phosphorus bonding as well as systems containing P-C, P-N, P-O, and P-S bonds have been reported. These studies also include cyclic

systems with phosphorus participating in one or more rings of varying size. Studies include investigations of geometrical parameters and various spectra in order to explain several well-known trends. These trends include longer apical bonds compared to equatorial bonds, apicophilicities and correlation to size and electronegativity, and stability due to stereoelectronic effects, conjugation, and donor interactions.

Several models have been proposed to explain the origins of these effects. They include a simple model of repulsion of bonding and lone pair electrons [99,100], inclusion of d orbitals in σ bonds through hybridization [95,96,99], donation of electron density from the ligands to the empty d orbitals on phosphorus [95,96,99,101,102], and construction of the molecular orbitals without the use of d orbitals [95,99,103–105]. Investigations into the origins of bonding in hypervalent compounds focused on answering the debate over whether the d orbitals were the reason behind bonding in these compounds. There is no debate that, in theoretical calculations, including d orbitals provides more accurate results. However, the exact nature of the d orbital participation is still being debated.

Valence Shell Electron Pair Repulsion

Some investigations into the bonding at pentacoordinate phosphorus looked at the problem as one of electron pair repulsion. The main advantage of this model is that it explained why these compounds had a trigonal bipyramidal shape. It also gave insight into why this shape was more stable than the alternative square pyramidal model for many of these compounds.

To give support to the theory of repulsion of electron pairs, a few studies applied the

valence shell electron pair repulsion (VSEPR) model. In the VSEPR model, electron pairs are classified according to whether they are lone pairs or they participate in a bond. The main idea of this model is that electron pairs will try to minimize repulsive interactions by moving as far from each other in space as possible. This idea follows from the Pauli exclusion principle that electrons of like spin will avoid each other. Since electrons in bonds or lone pairs are paired, then the model seeks to minimize the repulsion energy between all pairs of electrons.

Application of VSEPR theory to pentacoordinate phosphorus models gives a trigonal bipyramidal structure. In this structure, lone pairs and more electronegative groups occupy the axial positions preferentially. In the axial positions, repulsion between the lone pairs, or between bonding electrons closer to the more electronegative atoms is reduced. The model also explains the origin of longer axial bonds in these compounds. Not only does axial positioning reduce repulsions, lengthening of bonds does as well. In fact, a greater reduction in repulsions is seen for a greater lengthening of axial bonds as compared to equatorial bonds [99].

Recently, the electron localization function (ELF) has been applied to some hypervalent systems including PF_5 and PCl_5 to validate the results of VSEPR theory [100]. In this study, the electron populations of the bonding and lone pair regions of these and other systems were generated. In the case of PF_5 and PCl_5, a diagrammatic description of these populations was used to support the trigonal bipyramidal structure. An increase in population at phosphorus and in the bonding region between phosphorus and the ligand as one moves from F to Cl was observed. This supported the increase in the P-X bond

length seen in P-Cl as compared to P-F. Also, the electron population in axial bonding regions is lower than in equatorial bonding regions. The population of the axial ligands is greater than the population of equatorial ligands. These two results support the observed trend of longer axial bond lengths as compared to equatorial bond lengths.

Although simple, the VSEPR model is accurate in describing trends in the bonding and shape of pentacoordinate phosphorus molecules. The application of VSEPR theory, however, is not sufficient to determine orbital involvement in bonding.

d Orbital Participation Through Hybridization

The hybridization model for a species with five bonds suggests that there are five hybrid orbitals formed from the combination of an s, three p, and one d orbital. In the case of phosphorus, the hybrids are formed from mixing the 3s, three 3p, and a 3d orbital. For symmetry reasons, it was proposed that the hybrid orbital configuration was $sp^3d_{z^2}$.

The hybrids in the $sp^3d_{z^2}$ configuration at phosphorus are not equivalent. The two axial hybrids are equivalent, as are the three equatorial hybrids. However, the axial and equatorial hybrids differ in the orbitals mixed to form them. The three equatorial hybrids are formed from mixing of an s orbital and the p_x and p_y orbitals. The two axial hybrids are formed from mixing of the p_z and d_{z^2} orbitals. This would imply that the two bonding environments, and thus the corresponding bond lengths, are quite different. This result supports the differing bond lengths observed in the axial and equatorial positions at pentacoordinate phosphorus [95, 96, 99].

d Orbital Participation Via (d-p)$_\pi$ Donation

In the hybridization scheme, the d orbitals participate in σ bonds at phosphorus. This would imply a strong contribution of the d orbitals to bonding at phosphorus. However, this strong contribution is generally not observed in calculations on pentacoordinate phosphorus. This would imply that the d orbitals do not participate in the σ bonds to phosphorus. Instead, they may be available to accept donation of electron density from the attached ligands. In this case, electron density from lone pairs in p orbitals on the ligands would be shifted into empty d orbitals on the phosphorus atom. This type of interaction would explain both the evidence of d orbital participation and the fact that this participation is not as large as it should be if it participated in the σ bonds [95,96,99].

Evidence of this type of electron donation comes from ^{19}F NMR spectra of singly substituted PF_4SR. In this case, the R groups are alkyl groups. Donation of electron density from a p orbital on S to a d orbital on P would introduce π character into the P-S bond. It was argued that if this effect were not present, then the two axial fluorines should be in the same environment, and a single NMR peak would be observed for both. However, the NMR spectrum gave two different peaks for the axial fluorines, indicating that the two fluorines were in different environments. This observation can only be explained by restricted rotation of the alkyl group. This was thought to occur as a result of interaction of the sulfur with select d orbitals on phosphorus. This would restrict the alkyl group rotation, resulting in the two different NMR peaks [95].

Theoretical studies involving d orbitals participating in back bonding assume that some contraction of these orbitals occurs due to attached electronegative ligands. Upon

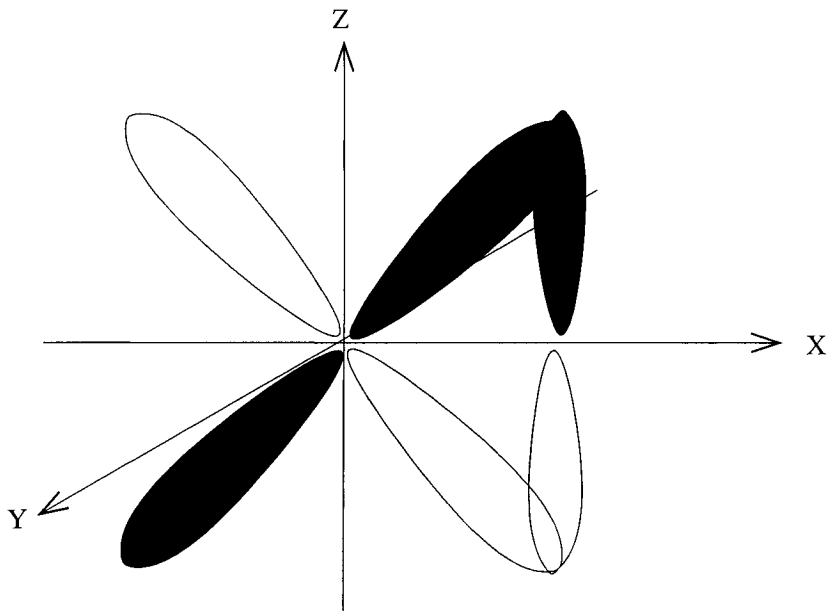


Figure 4.3: Diagram of overlap of a ligand p_z orbital with a phosphorus d_{xz} orbital (not to scale).

contraction, they are then available to accept electron density from the ligand p orbitals. It is then a matter of symmetry and the placement of electronegative ligands or those with lone pairs around phosphorus that determines the particpating d orbital. For example, in the PF_4SR system mentioned above, electron density from the lone pairs on sulfur would be donated into either the $d_{x^2-y^2}$, the d_{xz}, or the d_{xy} orbital. This type of overlap for the d_{xz} orbital is shown in Fig. 4.3 [95, 101, 102].

One theoretical study suggests that the presence of non-negligible density contribution to the d orbital on phosphorus in PF_5 is due to donation of electron density from the fluorines. Further proof is given by the contribution of the d_{z^2} orbital. The authors suggest that the d orbital is not just polarizing the phosphorus, but is accepting electron density from the axial fluorine atoms. They also suggest that previous studies which

simply remove the d functions from the basis set and compare results with and without d functions are in fact introducing some basis set superposition error. Calculations done excluding the d_{z^2} orbital while keeping the other four orbitals and accounting for these errors show that in fact this orbital contribution is significant [102].

A second theoretical study investigated the effect of the ligands on the d functions in the basis set. The authors found that the 3d orbitals do contract in the field of electronegative ligands. Orbital energy levels, populations, and atomic charges supported the d-p bonding in some hypervalent phosphorus molecules. It was found that the 3d energy levels dropped and the orbital populations increased. These results support some participation of the d orbitals in bonding. Since the contributions to σ bonding are too small, the authors suggest that some $(d\text{-}p)_{\pi}$ bonding must be occurring in the hypervalent systems studied [101].

From the above work, it seems likely that the back bonding picture of d orbital participation is valid. However, in terms of the theoretical studies, basis sets are small or their functions are highly localized. It is possible that this may introduce artefacts into the simulations that would otherwise be interpreted as evidence of back bonding.

Molecular Orbitals Without d Orbital Participation

As proposed above, d functions likely act as polarization functions, essentially making the basis set for phosphorus more flexible and providing a more accurate description of bonding at phosphorus. In this case, the contributions would also be small. However, whether this results from some type of hyperconjugation is debatable. In order to show

that the bonds about phosphorus do not need the d orbitals, some theoretical studies rely on forming the molecular orbitals without the use of d orbitals. This is the idea behind the molecular orbital model of bonding at pentacoordinate phosphorus. Support for this idea has been given by analysis of the molecular orbitals of PH_5, shown in Fig. 4.4.

In PH_5, the 3s and 3p orbitals on phosphorus combine with the five 1s orbitals on the hydrogens. These have been shown to give four bonding orbitals, four antibonding orbitals, and one nonbonding orbital. The three equatorial bonds are then described by three of the bonding molecular orbitals. These are the two 1e' degenerate orbitals and the $1a_1$' orbital in Fig. 4.4. The two axial bonds are described by a combination of a bonding and a nonbonding molecular orbital. These are the $1a_2$' and the $2a_1$' orbitals in Fig. 4.4. The effect of the nonbonding molecular orbital is to destabilize slightly the bonding molecular orbital, weakening the axial bonds. This results in longer axial bonds compared to the equatorial bonds [95, 99, 103–105].

Inclusion of d orbitals can be shown to only affect the nonbonding orbital, not the remaining bonding orbitals. Also, the energy stabilization resulting from including d orbitals is small. The only effect of this stabilization is a slight shift of electron density towards phosphorus. Thus, bonding at phosphorus can be effectively described without the use of d orbitals.

An interesting study based on the molecular orbital model was conducted by Magnusson [104]. In this work, he not only studied the effect of basis sets with and without d orbitals on phosphorus, he also investigated these same basis sets on the ligands. While

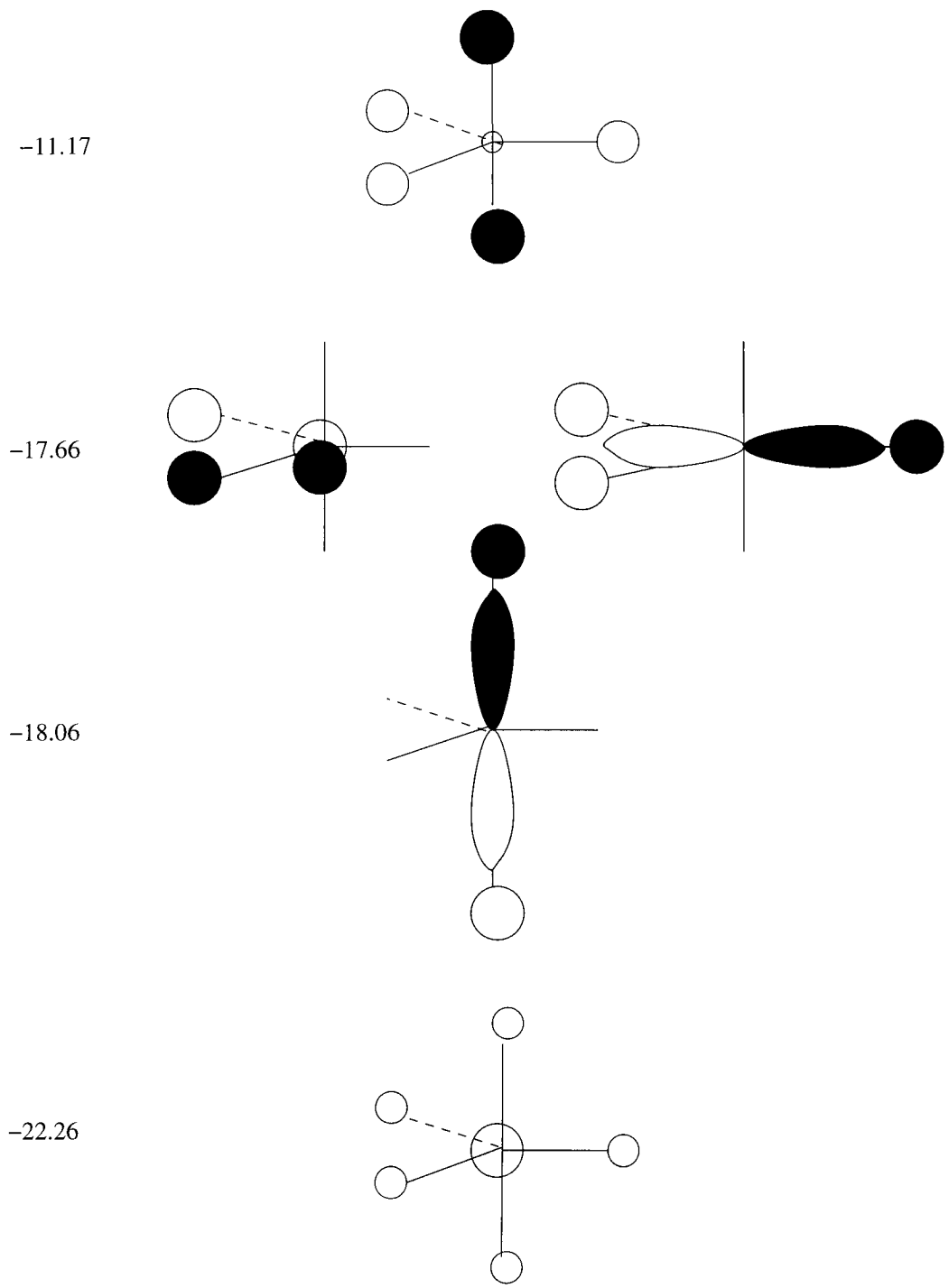


Figure 4.4: Diagram of the molecular orbitals of PH_5 and their symmetries.

he does not dispute that d functions are required to properly reproduce geometries of pentacoordinate phosphorus, his findings do not support d orbital participation in bonding at phosphorus. For studies employing small basis sets, the effect of d functions on energetics is larger than if a larger basis set is used. Another argument in support of d orbital participation was that the energy of pentacoordinate phosphorus molecules drops as d functions are added to phosphorus. However, Magnusson also found the same sort of energy drop if d functions are added to the ligands. Also, where d orbital participation was largest, charge transfer to highly electronegative ligands was also high. In this case, d functions (or polarization functions in general) provide a better description of the electronic structure of the molecule. The presence of basis functions of higher angular momentum has been shown to shorten bonds in these systems. This bond shortening has been attributed to augmentation of the basis set with higher angular momentum functions like d functions. Addition of higher angular momentum functions results in polarization of charge along the bonds. As previously mentioned, these d functions also contribute to higher-lying molecular orbitals and do not greatly influence the lower-lying bonding molecular orbitals. All of the above evidence points to the d orbitals being polarization functions. They therefore do not participate in bonding in pentacoordinate phosphorus [104].

d Orbitals or d Functions?

From the studies conducted on hypervalent second-row species, the results suggest the model has a significant effect on the interpretations of d orbital participation. More recent investigations focus on the participation of the d functions in the basis set as polarization

functions. Their presence allows the charge in the bonding regions to be shifted away from the nucleus. Earlier investigations focus on $(d\text{-}p)_\pi$ bonding and the d orbital participation in bonding of hypervalent second-row species. One general theme is present, however. These molecules are not well described by basis sets which do not include d orbitals.

Theoretical Investigations of Observed Experimental Trends

Theoretical investigations have been done on several of the halophosphoranes [106–108], as well as a few oxyphosphoranes [109,110]. Structural information for the halophosphoranes has been compared to experiment. Electronic structure effects have been used to try to explain geometries, spectral data, and trends in bonding in these compounds. Calculations on geometries and spectra are done Hartree-Fock level with a double zeta quality basis set incorporating polarization functions [106–109]. One investigation of a series of substituted oxyphosphoranes included calculations done at the second-order Möller-Plesset level of theory employing Pople's 6-31+G* basis set. Core orbitals were not included in the calculations [110].

For the PX_5 series, where X is either F or Cl, the Hartree-Fock calculations correctly reproduce observed geometries [108]. Bond lengths are in agreement with those observed

experimentally. It was found that the stable minimum structure is a trigonal bipyramid.

Table 4.3: Energy differences for the trigonal bipyramid and square planar geometries of the PX_5 series, where X=F, Cl, Br, I.

molecule	r_{ax}(Å)	r_{eq}(Å)	$\Delta E_{sqp-tbp}$ (kcal/mol)
PF_5	1.566	1.531	4.87
PCl_5	2.127	2.013	4.78
PBr_5	2.374	2.227	3.85
PI_5	2.697	2.466	3.44

For PF_5, the square planar geometry lies 4.87 kcal/mol higher in energy relative to the trigonal bipyramid. These results are in agreement with a similar study for PF_5 which gives an energy difference of 4.24 kcal/mol [107]. For PCl_5, the square planar geometry lies 4.78 kcal/mol higher in energy. These results and those for PBr_5 and PI_5 are included in Table 4.3. For PBr_5 and PI_5, no experimental results are available. It is likely that these two molecules are too unstable to isolate for long enough to investigate experimentally.

A few overall trends are worth noting. First, as the size of the halogen increases, the ratio between the axial and equatorial bond lenths also increases. This observation parallels the increase in size and electronegativity of the halogen atoms. Therefore, as the electronegativity of the halogen on phosphorus decreases and its size increases, both bond lengths increase. However, the axial bond lengths increase in greater proportion as compared to the equatorial bond lengths. For experimental studies on PF_5 and PCl_5, this trend is also observed [108].

Theoretical work on PF_4CH_3 shows the trigonal bipyramid with the methyl group in the equatorial position is the minimum energy geometry. The square pyramidal geometry

with the methyl group in the axial position is 3.95 kcal/mol higher in energy. This indicates the molecule is able to invert and place the methyl group in any of the three equatorial positions. For the $PF_3(CH_3)_2$ system, the molecule adopts a trigonal bipyramidal structure with the two methyl groups in the equatorial positions. In this case, inversion about phosphorus will give a product with one methyl group in the axial position and one in the equatorial position. For inversion to take place, the barrier to inversion must be very small. However, an inspection of the energy difference between the two isomers shows that it also must be very small. If the energy difference is large, then the energy required to surpass the activation barrier to inversion is too large and inversion does not occur. However, it was found that the axial-equatorial placement of the methyls results in a structure 14.01 kcal/mol higher in energy than if they were both placed equatorially. This suggests that inversion does not occur in this molecule. All of the above results support experimental observations [107].

Studies on substituted oxyphosphoranes had no experimental counterparts for the systems studied. Therefore, assessment of the quality of the results must be based on observed trends. Again, the calculated geometries are trigonal bipyramidal. The axial bond lengths are longer than the equatorial bond lengths in the case where the substituents are the same. A comparison of bond lengths of remaining substituents with the electronegativity of the substituted group shows that a more electronegative element produces a shortening of the remaining bonds [109, 110].

Apicophilicities have been calculated for both the fluorophosphoranes, hydrogen phosphoranes, and a select group of oxyphosphoranes [106, 107, 110–113]. All the studies reported geometries at the HF level of theory. Refs. 106,107 report energies at the MP4 level of theory, while Refs. 110–113 report energies at the HF level of theory. Refs. 106,107 still report on the participation of d orbitals in bonding. Although these studies are all based on the molecular orbital model described above, some investigations [106, 107] still indicate d orbital involvement in bonding. Regardless of the model, the general observation is that a variety of factors contribute to apicophilicity. All the above studies have shown that electronegativity, in many cases, correlates directly to the observed apicophilicities. Also important are steric effects, solvent effects, and, in the case of cyclic phosphoranes, ring strain. However, theoretical studies have indicated that the σ donating/withdrawing (inductive) and π donating/withdrawing (π interactive) abilities of the attached ligands must also be considered.

McDowell and Streitwieser Jr. [106], as well as Wang and coworkers [112], studied substituent effects on the apicophilicities of the PH_4X series. They attributed the apicophilicity to both inductive effects and π interactions. The origin of the inductive effects is in the nonbonding or slightly antibonding character of the highest occupied molecular orbital (HOMO). This nonbonding character results in much of the electron density being placed on the ligand. For ligands of low electronegativity placed in the apical position, a large amount of electron density remains in this nonbonding orbital and the ligand participates very little in this bond. This results in longer and weaker bonds. If these ligands are placed in the equatorial position, the electron density remains in a bonding orbital,

and the less electronegative ligand can participate more effectively in bonding. For more electronegative ligands placed apically, much of the electron density in the HOMO is transferred to the ligand. This creates a stronger bond with an increased ionic character. If these ligands are placed in the equatorial position, the ligand would remove electron density from the bonding region, weakening the bond. This accounts in part for the preference of more electronegative ligands for the axial sites.

π interactions are described by considering the π donating or accepting capability of the ligands. π accepting ligands prefer the axial positions. It is believed that some stability is gained by interaction of the apical electron-withdrawing group with a degenerate pair of π-type orbitals on the equatorial plane. The electron-withdrawing nature of the ligand in the axial position moves electron density onto the ligand itself from the two degenerate equatorial orbitals. These orbitals experience a decrease in the already present antibonding character, resulting in a stabilizing interaction. On the contrary, if an electron-donating ligand is placed axially, it donates electron density into the same pair of equatorial orbitals, increasing their antibonding character, resulting in a destabilizing interaction. This increases the amount of electron density shared between the three P-Y equatorial substituents [106, 112].

Initial work by Holmes [98] indicated an empirical model was effective for explaining the experimentally observed apicophilicities of some pentacoordinate phosporus molecules. The model was based on electronegativity, steric, and ring strain effects. However, more recent theoretical models described above challenge this empirical model. These newer apicophilicity comparisons have shown that in some cases discrepancies in the observed

apicophilicities can be explained in terms of inductive effects and π interactions. The important conclusion from all the studies is that the above competing effects must all be taken into account to explain both the experimentally and theoretically observed apicophilicities.

Both the theoretical and experimental studies, no matter the model used, show that properties of the ligand play a role in the site they occupy, the stability of the given isomer relative to other possible isomers, and the geometry of pentacoordinate phosphorus. These effects are important in the mechanistic interpretation of phosphoryl transfer reactions. Also important is the ability of the pentacoordinate phosphorus molecule to exchange ligand positions. An example of this is the PF_4CH_3 system, where the methyl group can occupy one of three equatorial positions. Experimentally, it has been shown that the methyl group interchanges these positions. This is supported with theoretical calculations showing a relatively low barrier to inversion. There are two models for this interconversion process. They are the turnstile and Berry pseudorotation mechanisms. It has been shown that the intermediate states in the Berry pseudorotation mechanism are of lower energy than in the turnstile mechanism, so this is the preferred path for ligand exchange. However, there may be cases where this mechanism of exchange is blocked and the turnstile mechanism is the preferred route [95].

4.2.3 The Berry Pseudorotation Mechanism

The Berry pseudorotation mechanism was first proposed to explain observed ^{19}F NMR signals in PF_5. Only one signal was found to represent the fluorines. However, two were expected, one each for the axial and equatorial environments. Some sort of intermolecular

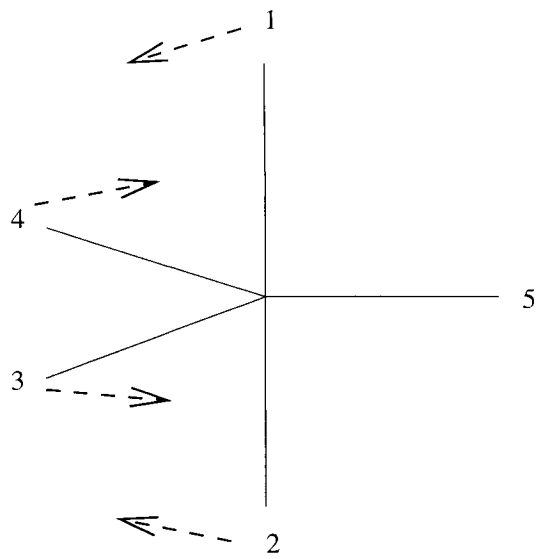


Figure 4.5: Diagram of a vibrational mode which distorts a trigonal bipyramid geometry to give a square pyramid.

exchange was occurring to make all five fluorines equivalent.

An examination of the vibrational modes of PF_5 showed that one mode involving the simultaneous bend of the axial and equatorial fluorines would distort the geometry to square pyramidal. An example of such a bending mode is shown in Fig. 4.5.

If this distortion continues, the square pyramid will be converted into a new trigonal bipyramid with four of the five fluorine atoms exchanged. This exchange is equivalent to selecting the fifth (and unmoved) fluorine atom as a pivot point and rotating the remaining four by 90° [1, 95].

Fig. 4.6 shows the intermolecular exchange process for the Berry pseudorotation mechanism. The ligands are labeled to uniquely identify them. The vibrational distortion shown in Fig. 4.5 causes the axial ligands 1 and 2 to bend towards the equatorial ligands 3 and 4. As this occurs, the equatorial ligands 3 and 4 bend toward ligand 5. This distortion

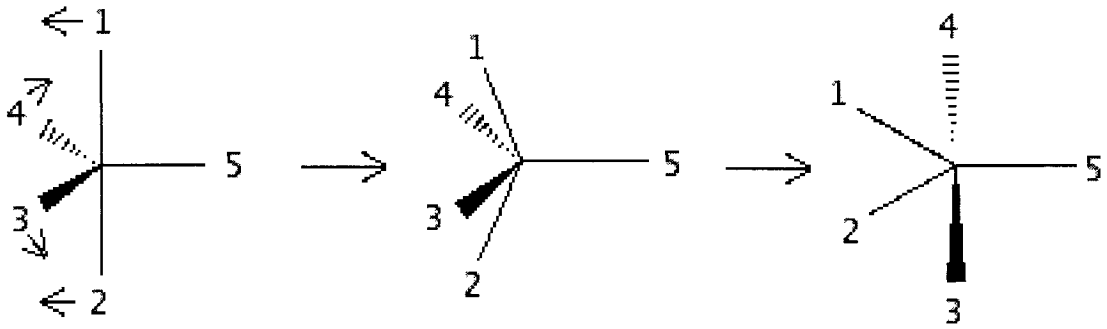


Figure 4.6: Diagram of the exchange coordinate for the Berry pseudorotation.

results in the square pyramidal transition state. Further distortion along this coordinate places ligands 3 and 4 in the axial positions and ligands 1 and 2 in the equatorial positions.

A single interconversion, however, did not account for all five fluorines being equivalent. However, a second interconversion, done about another fluorine atom, would be sufficient to make the fluorines indistinguishable from one another. It was proposed that this type of exchange occurred in PF_5 faster than the time scale for an NMR experiment. This was the reasoning for the observed spectrum of ^{19}F NMR [1, 95].

Structural characteristics and ligand exchange have been extensively studied for many pentacoordinate phosphorus compounds. These studies have proven useful for explaining various parts of the reaction mechanisms observed for phosphoryl transfer reactions.

4.3 The Phosphoryl Transfer Reaction

4.3.1 Introduction

Little work has been done on investigating phosphoryl transfer reactions in enzymes. As many of these enzymes rely on metal atoms in the active site to achieve their function, their size alone makes them a challenge to study theoretically. As a result, much of the theoretical work of enzyme-catalyzed phosphoryl transfer has been done on small model reactions. These reactions are conducted in solution and the mechanism is tailored to mimic the corresponding enzyme mechanism. Experiment has revealed that only two of the three possible mechanisms for phosphoryl transfer actually occur in enzyme systems. As a result, the theoretical models focus on these two pathways. However, to obtain a thorough understanding of phosphoryl transfer reactions, it is instructive to review these reactions in general.

4.3.2 General Considerations

Phosphoryl transfer reactions in enzymes proceed via nucleophilic substitution at phosphorus [95]. The substrates in these reactions are alkyl phosphate esters. The general equation for these reactions is given in Fig. 4.7.

The mechanisms for these substitution reactions are analogous to the S_N1, S_N2, and addition-elimination mechanisms which occur at tetrahedral carbon [114].

Before considering the nucleophilic substitution at phosphorus (as opposed to carbon), several important factors concerning the reactivity of the nucleophile should be

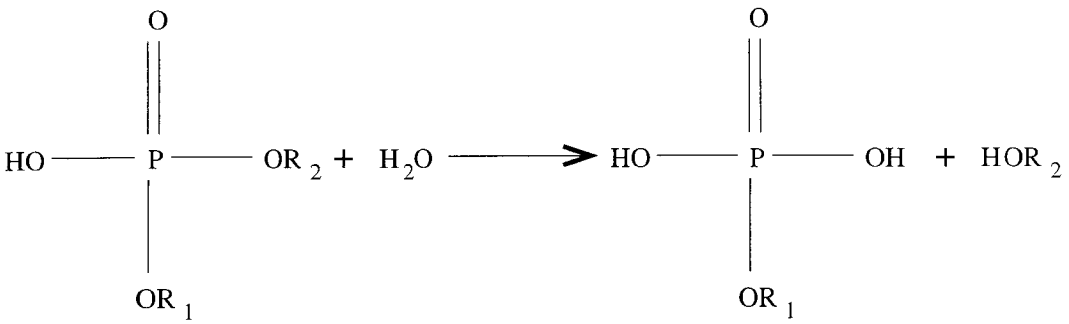


Figure 4.7: Equation for a general phosphoryl transfer reaction.

discussed. The strength of the bonds to both the nucleophile and leaving group is an important factor in determining reactivity in the S_N2 and addition-elimination mechanisms. It may be a determining factor in the preference for nucleophilic entry and electrophilic exit at the apical positions. This is due to the apical bonds in the pentacoordinate intermediate or transition state being weaker than the equatorial bonds. Similar to nucleophilic substitution at carbon, basicity of the nucleophile is also an important factor in determining reactivity. In the case of phosphate esters, it has been suggested that both the size and apicophilicity of the nucleophile may also affect its reactivity. Another effect, which is more of a factor in the reactivity at phosphorus versus carbon centers, is the ease at which the bond to the leaving group is weakened. A final point to consider is the position of the nucleophilic atom relative to atoms bearing lone pairs [95]. The alpha effect [96] also affects the reactivity of a nucleophile. This effect describes an improvement in the observed nucleophilicity of a group containing an atom with a lone pair next to the nucleophilic atom. From the effects discussed above, it is clear that, as in nucleophilic substitution at carbon centers, reactivity of a nucleophile depends on several contributing factors.

Before considering each of the mechanisms and some of the evidence to support them, some important points are worth noting. First, the evidence collected for the three mechanisms does not give unquestionable proof of one mechanism over the other. It is often the case that intermediates postulated for a mechanism are too reactive to isolate. Also, stereochemical evidence can be somewhat misleading since reorganization of the groups attached to phosphorus can occur throughout the course of the reaction. This is especially true in the case of pentacoordinate intermediates.

4.3.3 The Dissociative Mechanism

The dissociative mechanism of phosphate esters has been thought to be analogous to the S_N1 mechanism of carbon centers [95, 114]. In this case, the reaction rate should follow first order kinetics and thus depend on the concentration of the phosphate ester only. Much of the evidence collected shows this is true mainly for phosphate monoester monoanions. For phosphate diesters and triesters, the reaction proceeds via a concerted or addition-elimination mechanism. These two mechanisms will be discussed in the next two subsections.

The simplest mechanism to support the first-order rate observations is that of a two step process analogous to the S_N1 mechanism of carbon compounds and is shown in Fig. 4.8.

As with carbon compounds, the first step of the reaction is the exit of the leaving group, forming a trigonal planar, tricoordinate phosphorus intermediate. The charge of the phosphorus in this intermediate is formally positive, analogous to the carbocations

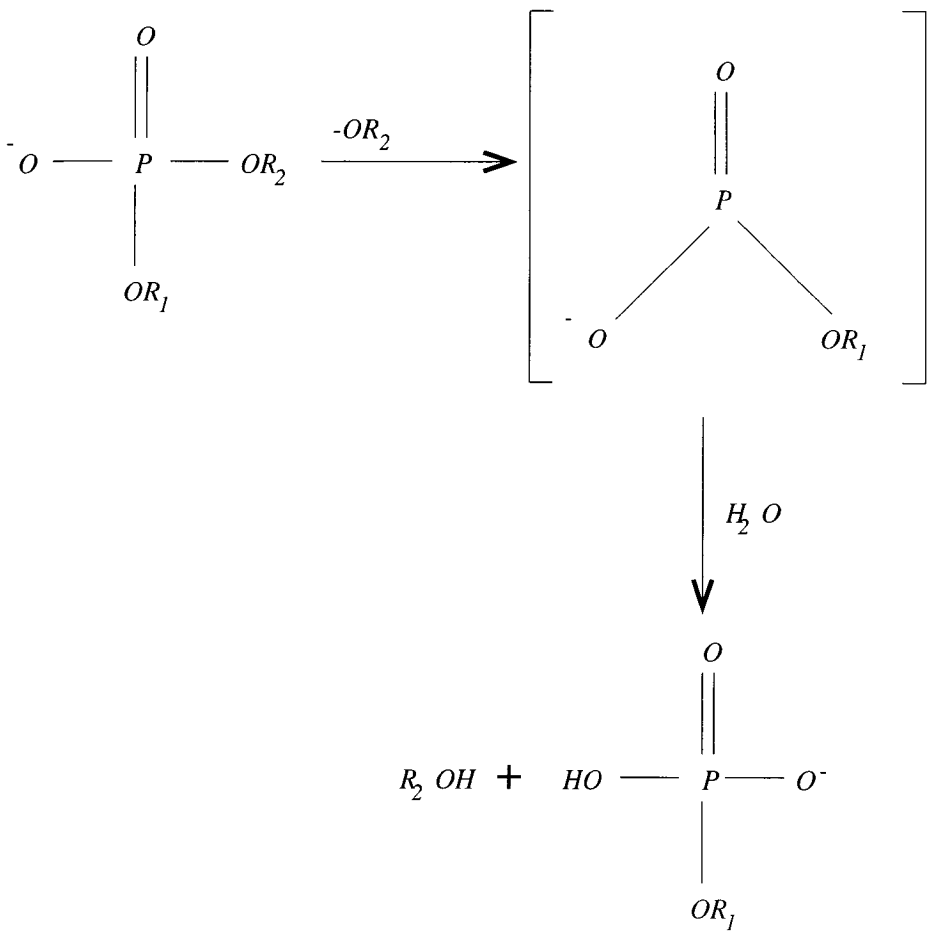


Figure 4.8: Diagram showing the dissociative mechanism for a phosphate monoester.

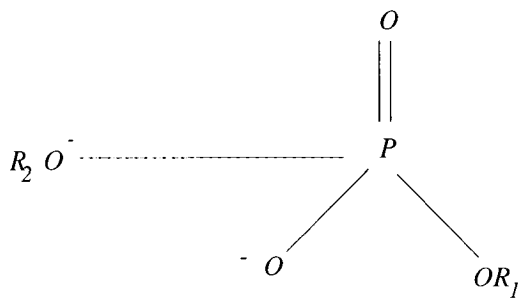


Figure 4.9: Diagram of a possible example of a metaphosphate intermediate.

formed in S_N1 reactions at carbon centers. Since the rate depends on the concentration of the starting phosphate ester only, this is the rate-determining step. In the second step, nucleophilic attack occurs. This can occur from one of two positions: above and below the plane formed from the three groups attached to phosphorus. Due to the reactivity of the intermediate, this step is expected to be rapid [95, 114].

Little evidence has been given to prove the existence of a free monomeric metaphosphate intermediate. The phosphorus atom of this intermediate is highly electrophilic, and as a result the intermediate is extremely reactive. It has been suggested that in many cases it is not free metaphosphate that is observed. Instead, the metaphosphate may form a complex with the leaving group, as shown in Fig. 4.9. This would stabilize the metaphosphate and also assist in the reaction with the nucleophile and exit of the leaving group. This is especially true in cases where the rate of decomposition of the metaphosphate intermediate is faster than the time for the nucleophile to diffuse through the solution to react with it. For example, for a monomeric phosphate containing an acidic proton, the rate-determining step is the concerted proton transfer and phosphate-leaving group bond

cleavage. This results in the formation of an intermediate where the leaving group and metaphosphate are associated [115–118].

Stereochemical evidence can be provided to support the dissociative mechanism. Due to the presence of a trigonal planar intermediate, nucleophilic attack can occur from either side of the plane produced by the three attached groups. If the starting phosphate ester is chiral, a racemic mixture should be produced [95, 114, 116].

Along with stereochemical studies, solvent effects have also been studied for the dissociative mechanism. It has been found that the nature of the solvent can be a deciding factor on whether the mechanism for phosphoryl transfer proceeds via a dissociative or associative mechanism. The reason for this is twofold. First, it was found that in experiments conducted in aqueous solution, no evidence could be found for the presence of a metaphosphate intermediate. It is likely that the intermediate reacts so rapidly with water that it is too short-lived to be detected. Experimental results conducted in the gas-phase indicate the possibility of a metaphosphate intermediate. This has also been shown for reactions in nonpolar solvents. Second, the reactant, a phosphate ester monoanion, is polar and contains a single negative charge. The metaphosphate anion has a charge of -2 and is less polar. As a result, a more polar solvent would better solvate the reactants than the intermediate. This would destabilize the intermediate relative to the reactant. Thus, better evidence for a dissociative mechanism is expected in a less polar medium [95, 116].

A significant amount of additional evidence has also been proposed to support the metaphosphate intermediate, both in the gas phase and in theoretical calculations [115–122]. Among this evidence are the following observations: the presence of small entropies

of activation, little or no dependence of the rate on the nature of the nucleophile, linear free energy relationships relating the log of the rate of the reaction to the pK_a of the leaving group, and the presence of inverse isotope effects. More information on these results can be found in two reviews on nucleophilic substitution of phosphate esters [115, 116] and in a metaphosphate paper by Westheimer [117].

4.3.4 The Concerted Mechanism

Recently, a paper was published which questioned the validity of the evidence which suggested that hydrolysis of phosphate esters in solution proceeded by a dissociative mechanism [123]. The authors suggested that the evidence collected could be interpreted as evidence for both a dissociative-type mechanism as well as an associative mechanism. Much of the data they examined was based on monoanions and dianions of monoesters, and monoanions of diesters. The data included comparisons of kinetics of reactions of phosphoesters, the linear free-energy relationships with varying pK_a values for both nucleophile and leaving group, kinetic isotope effects on the bridging oxygen of the phosphate with the leaving group, and entropies of activation. What the authors found was that there was no clear evidence to support either mechanism. In fact, several factors seem to be important when considering the type of mechanism postulated, including the pH of the solution, pK_a of the leaving group and nucleophile, and the solvent used. It seemed that the distinction between a metaphosphate interacting with the leaving group and a transition state involving significant bond breakage to the leaving group but a small amount of bond formation to a nucleophile was not clear.

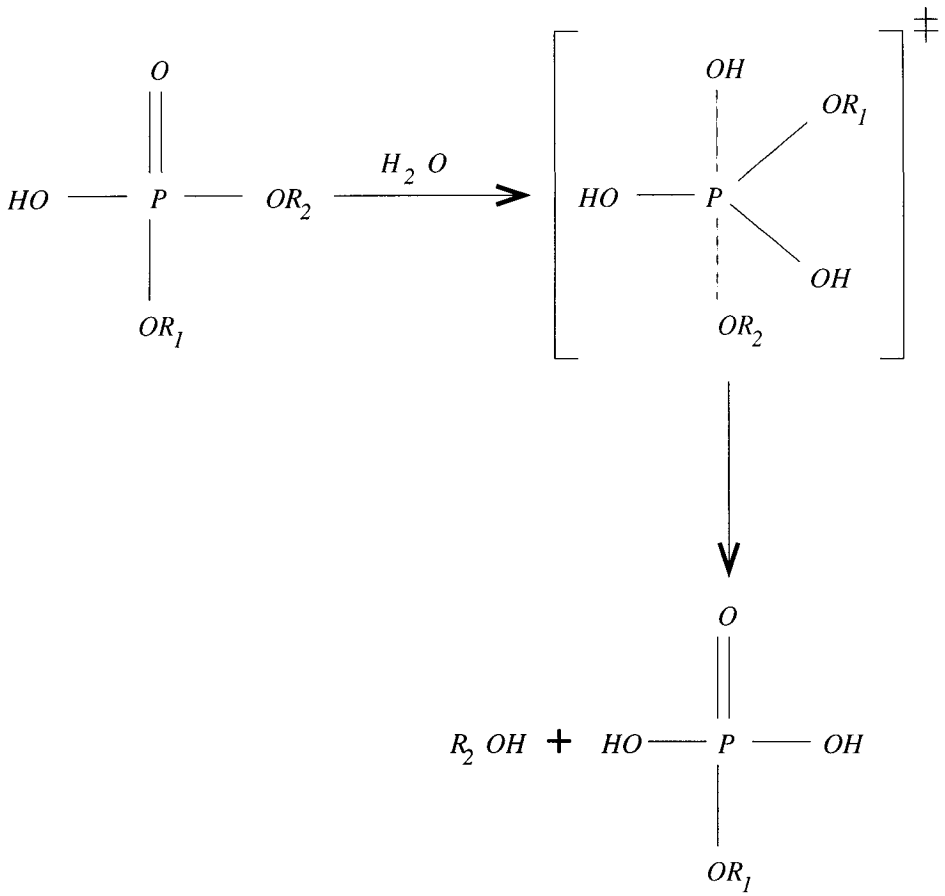


Figure 4.10: An example of the S_N2 mechanism of phosphate esters.

So, it seems that in aqueous solution, depending on the reaction conditions, the more plausible mechanism is one which mirrors the S_N2 mechanism of carbon compounds. For phosphoryl transfer reactions believed to follow this type of mechanism, the rate follows second-order kinetics. It is dependent both on the initial concentrations of the phosphate ester and the nucleophile. The simplest mechanism to support such observations is a single step process where both the nucleophile and the starting ester are present in the transition state [95, 114]. An example of such a mechanism is shown in Fig. 4.10.

It is found that for bimolecular substitution mechanisms involving the above type of transition state, activation entropies are negative and usually quite large. This is due to the large increase in order seen as one goes from the less ordered reactants to the more ordered transition state. Evidence from kinetic isotope effects indicates the presence of a transition state involving both the nucleophile and the starting phosphate ester. Both these results seem to indicate the mechanism follows the one shown in Fig. 4.10. Also, in reactions which follow the above mechanism, no intermediates were isolated, indicating that the reaction proceeds through a species that is too reactive to isolate [95, 124, 125].

In S_N2 reactions at chiral carbon, the mechanism shows that the reactions are expected to proceed with complete inversion of configuration. In the case of primary alkyl halides, this is in fact what happens. In the case of phosphate esters, this is also seen [95, 114].

In the case of substitution at carbon, the transition state is fairly nonpolar in nature. As a result, these reactions usually occur in nonpolar solvents. However, the transition state in the case of phosphoryl transfer is often more polar than the reactants. As a result, more polar solvents, in particular protic solvents, favor the S_N2 type mechanism.

4.3.5 The Addition-Elimination Mechanism

In many enzymes which catalyze phosphoryl transfer, the reaction is believed to occur via a concerted mechanism through a pentacoordinate transition state. There has, however, been considerable debate whether this transition state is actually an intermediate. Studies including kinetics of disappearance of such an intermediate as well as the

stereochemistry of the products relative to the reactants are quoted as proof of one type of pathway over the other. It seems the difficulty of determining the difference between the presence of a pentacoordinate transition state and an intermediate is that the intermediates are often too short-lived to be isolated [95, 99, 115, 116, 126].

The mechanism of phosphoryl transfer reactions which proceed via a pentacoordinate intermediate is analogous to the addition-elimination mechanism of carbonyl compounds. The kinetics of these reactions is similar to a bimolecular nucleophilic substitution. The rate is dependent on both the nucleophile and the starting ester. The simplest mechanism to explain the above results is a two step process, shown in Fig. 4.11.
In the addition-elimination mechanism, the first step is addition of the nucleophile to the phosphorus to form a pentacoordinate intermediate. This is the rate-determining step of the reaction. This step is followed by the exit of the leaving group [95, 114].

Some of the early work on associative mechanisms at phosphates was done by Westheimer [99, 116, 126, 127]. These studies focused on reactions of cyclic phosphate esters, most containing a five-membered ring. However, some of the resulting guidelines may also be applied to other phosphates. The Westheimer guidelines can be summarized as follows:

- Nucleophilic attack on phosphate esters occurs via a trigonal bipyramidal pentacoordinate species. This is either a transition state or an intermediate.

- In the case of a pentacoordinate phosphorus intermediate, ligand reorganization can occur via ligand exchange processes.

- The nucleophile enters in the apical position. The leaving group leaves from the

apical position.

- site preference rules:
  - small rings containing phosphorus span apical-equatorial sites
  - more electronegative ligands prefer apical sites
  - π donating ligands prefer equatorial sites
  - bulky ligands prefer equatorial sites as this reduces steric repulsion.

Evidence for a reaction mechanism which proceeds via a pentacoordinate species is based on kinetics, linear free-energy relationships, thermodynamics, and spectroscopic studies. The results from kinetics, thermodynamics, and linear free-energy relationships are similar to those observed for S_N2 mechanisms and will not be discussed. The most interesting result of these studies is many of the results do not distinguish between a pentacoordinate intermediate and transition state. Since the postulated intermediates for many of the reactions studied have not been isolated, much of the evidence for these reactive species has been based on stable pentacoordinate models. It was believed that if stable pentacoordinate species exist, then it was possible the intermediates also exist, but are too short-lived to study spectroscopically. Combined data on stable pentacoordinate species and stereochemical investigations for a variety of reactions have been used to support the pentacoordinate intermediate [99, 115, 116, 126].

Stereochemistry has provided much insight into the possibility of a pentacoordinate phosphorus intermediate. It was believed that reactions on chiral phosphates proceeding via the dissociative mechanism resulted in complete racemization of the product mixture.

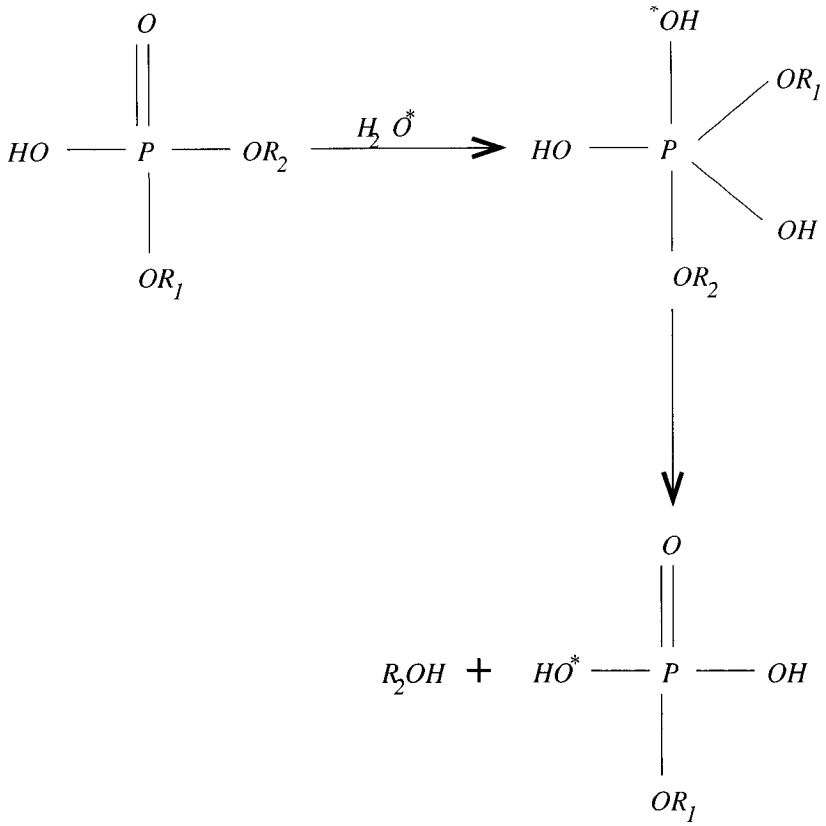


Figure 4.11: Diagram for inline nucleophilic attack at phosphate esters.

Therefore, reactions on chiral phosphates which showed either complete retention or inversion in the products must occur via the S_N2 or addition-elimination mechanism. In other words, the pentacoordinate species was a transition state or an intermediate in these reactions. Formation of a pentacoordinate intermediate was thought to occur by one of two modes of attack.

Fig. 4.11 shows one of two modes of attack. In this diagram, the nucleophile attacks the tetrahedral face formed from the ligands other than the leaving group. The resulting pentacoordinate species has the nucleophile and leaving group spanning the axial positions.

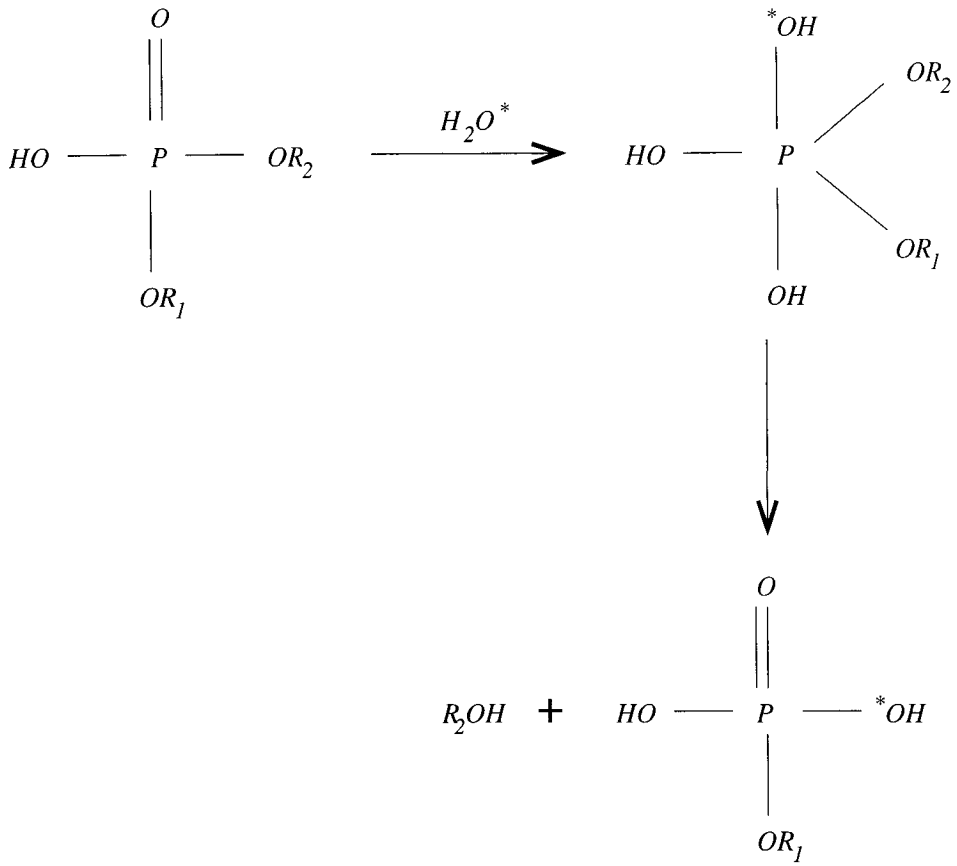


Figure 4.12: Diagram for adjacent nucleophilic attack at phosphate esters.

They both occur along the same line, hence the mechanism is called inline substitution. According to the Westheimer rules, the leaving group should also leave from the apical site it occupies. Inversion of configuration at chiral phosphorus is expected for this mode of attack [99, 116, 126].

The second mode of attack is shown in Fig. 4.12. In this diagram, the nucleophile attacks from a position other than the tetrahedral face. The nucleophile occupies the apical position while the leaving group occupies the equatorial position. However, according to the Westheimer rules, the leaving group departs from the apical position. This makes sense

since the apical bonds are the weakest and thus the easiest to break. Since the leaving group is positioned equatorially, ligand exchange, presumably by Berry pseudorotation, must occur to position the leaving group apically. If this occurs, reactions at chiral phosphates that occur by adjacent attack result in retention of configuration in the products.

Both mechanisms of attack occur for phosphate esters. Thatcher and Kluger [116] indicate that direct displacement is most popular in enzyme reactions, although adjacent attack occurs in some cases. Gillespie and coworkers [126] believe it depends on the electronegativities of the leaving group, nucleophile, and remaining ligands. For reactions where the nucleophile and leaving group are both significantly more electronegative than the remaining ligands, then direct inline attack is the favored mode. Any pseudorotations that occur will give inversion of configuration only if there are an even number of them. If the nucleophile is more electronegative than the leaving group and remaining ligands, but the electronegativities of the leaving group and remaining ligands are comparable, then adjacent attack is highly likely. Retention of configuration occurs for an odd number of inversions.

Since the transition state in the S_N2 mechanism is too short-lived for pseudorotation to occur, the direct inline attack is the postulated mechanism. This would suggest that the best evidence for a pentacoordinate intermediate is retention of configuration. However, in many cases, inversion of configuration is often observed. The exact mechanism is still not discernable from the evidence as it is indicative of either mechanism.

In the case of cyclic phosphate esters, the addition-elimination mechanism often leads to ring cleavage. This mechanism occurs to reduce the ring strain and is believed to occur

via pentacoordinate intermediate. This is especially true for phosphate esters containing five-membered rings. In this case, there is a reduction in ring strain when the ring angle is reduced from 99° in the tetrahedral phosphate to 90° in the pentacoordinate intermediate. As compared to the all carbon analog, cyclopentane, the ring angle in the phosphate ring is smaller than the 108° expected for cyclopentane. This reduction in ring angle is due to the much longer P-O bond lengths as compared to the C-C bond lengths. The reduction in ring strain is believed to be one of the driving forces for the formation of the pentacoordinate intermediate [99, 116, 126].

4.3.6 Catalysis in Nucleophilic Substitutions of Phosphate Esters

Catalysis is common in nucleophilic substitutions at phosphate esters. This is especially true for phosphoryl transfer in enzyme systems. The catalyst is often an acid, base, or some type of metal atom.

The most convincing evidence for catalysis is the reduction in the overall rate. That is, the rate of both of the above steps must be less than the rate of nucleophilic attack without the catalyst. Little speculation has been made on the role of the acid or base in reducing the overall rate. However, metal atoms have been thought to play a variety of roles. First, the metal may be used to bring reacting species close enough for a reaction to occur. Second, the metal may act to stabilize charge in the case where the transition state or intermediate is highly charged. Third, the metal may aid in the formation of a ring system at the phosphate. Evidence indicates reactions occurring in cyclic esters are much

faster than their acyclic counterparts. Fourth, the metal may act to aid in the formation of a particular stereoisomer. Finally, the metal may complex water or other molecules in which a hydroxyl is present [95, 116].

4.4 Conclusion

A significant amount of work has been done on structural features of pentacoordinate phosphorus models. In most of these studies, however, the molecules contain highly electronegative ligands or rings. As a result, these models and their predictions are based mainly on these systems. However, theoretical investigations have shown a quality similar to experiment. Theoretical models may be useful in describing pentacoordinate phosphorus compounds that are not stable enough to study experimentally.

Three mechanisms have been proposed for the transfer of a phosphate to water or another group. They are the S_N1, S_N2, and addition-elimination mechanisms. In the S_N2 mechanism, the transition state is a highly reactive pentacoordinate phosphorus system. In the addition-elimination mechanism, this pentacoordinate phosphorus is an intermediate. Therefore, a proper description of pentacoordinate phosphorus compounds is essential to modeling phosphoryl transfer reactions.

Chapter 5

Pentacoordinate Phosphorus Models: Results

5.1 Introduction

In the previous chapter, the importance of a proper description of pentacoordinate phosphorus for the investigation of phosphoryl transfer was discussed. Theoretical studies showed that d functions were important for the proper description of properties of pentacoordinate phosphorus systems.

The phospholipase C family of enzymes are responsible for the catalysis of hydrolysis of phospholipids present in membranes. Bacterial phosphatidylinositol-specific phospholipase C cleaves the inositol end of the phospholipid to produce diacylglycerol and inositol-1-phosphate. The mammalian version of this enzyme also produces diacylglycerol, but the second product is inositol-1,4,5-trisphosphate. In mammalian systems, these

two species are potent second messengers. The inositol-1,4,5-trisphosphate activates a membrane-bound enzyme which is responsible for the release of Ca^{2+}, which activates other calcium-sensitive enzymes. The diacylglycerol activates protein kinase C, which phosphorylates specific Ser and Thr residues on other proteins, thus changing their catalytic activity [128].

Experimental work [92] has indicated that this enzyme reacts via a pentacoordinate transition state, but the exact mechanism and participation of active site residues is still being debated. To date, no theoretical studies have been conducted on any of the PI-PLC enzymes. The bacterial PI-PLC enzyme is a good model for this class of enzymes as it is the smallest one in the group and can be studied theoretically. Often studies of enzyme mechanisms involve a model where the active site and a part or all of the substrate is treated by quantum mechanical methods and the remainder of the protein and any solvent are treated with molecular mechanics. Investigation of bacterial PI-PLC by this method is suitable as little change in the overall structure of the enzyme is expected throughout the course of the reaction. The active site can thus be approximated as rigid.

Before an investigation into the mechanism of bacterial PI-PLC in the enzyme can be conducted, some indication of the quality of the results from the quantum mechanical portion of the calculation is necessary. If this part of the calculation gives poor results, then the overall results will be poor. Therefore, the first part of the study involves a study of the treatment of the active site and a portion of the phosphatidylinositol substrate by quantum mechanics. However, this is difficult to do outside the enzyme, so a study of the reaction of hydrolysis of phosphatidylinositol will be done. The protocol used will be the

one that is expected to be used in the study of the enzyme mechanism, so the results will be used to validate its use in the enzyme study.

Since the mechanism of bPI-PLC is thought to proceed via a pentacoordinate transition state, proper description of this system is essential. Experimentally, a great deal of work has been done on pentacoordinate phosphorus compounds with highly electronegative ligands or multiple ring systems. However useful these results have been, they do not accurately describe the intermediates or transition states of enzyme reactions. These systems usually contain one small ring containing phosphorus, long or bulky alkyl side chains, and less electronegative ligands. As a result, these systems are a problem to isolate experimentally. Theoretically, due to their enormous size, they provide a challenge to study with accurate methods.

In this chapter, a study of models of pentacoordinate analogs of the postulated intermediates and transition states for phosphoryl transfer enzymes is presented. Where possible, experimental results will be provided. However, before such models can be studied, a test of the method and basis set is conducted. The results of this test are compared to some of the better studied pentacoordinate phosphorus systems. The accuracy of the method is compared to the post-HF method, second-order Möller-Plesset perturbation theory.

5.2 Quality of VSXC for Structures and Vibrational Frequencies

Previous investigations of pentacoordinate phosphorus models have shown that HF methods are not able to quantitatively reproduce energies and geometries of pentacoordinate phosphorus systems. Earlier investigations pointed, not to HF, but to the lack of inclusion of d functions on phosphorus. A study by Magnusson [104] showed that d functions are essential for the proper treatment of pentacoordinate phosphorus systems. Another study conducted by Magnusson [129] showed that inclusion of electron correlation is essential for the proper treatment of pentacoordinate phosphorus. In fact, in his study on electron correlation effects, he showed that both correlation and d functions must be present for a proper description of pentacoordinate phosphorus.

There are many methods available that include electron correlation. The bulk of these are based either on multiconfigurational approaches or corrections upon Hartree-Fock theory to include correlation. Although accurate, these methods are computationally expensive. For treating systems on the order of enzyme reaction intermediates, they become prohibitive. In these cases, density functional theory has provided a cheap alternative for the inclusion of electron correlation. DFT methods are on the order of HF methods for computational complexity, making them ideal for use with large systems. Of all the functionals available, B3LYP is the most widely used functionals.

Previous investigations have shown that the non-hybrid VSXC functional gives geometries and vibrational frequencies of similar quality to the hybrid B3LYP functional

[130–132]. Previous studies have shown that the B3LYP functional is adequate for describing structures and energetics of phosphate ester hydrolysis [133]. Thus, it is probable that the VSXC functional will also provide adequate results for the structures and spectra of pentacoordinate phosphorus, although no investigations to date have been performed.

5.3 Computational Methods

The transition state for the hydrolysis of phosphatidylinositol by bPI-PLC is the oxyphosphorane resulting from addition of an OH attached to the inositol ring to the phosphorus. However, this species is too short-lived to be studied experimentally. As a result, model oxyphosphoranes containing a five-membered ring will be used. However, since systems which look similar to the proposed transition state only have NMR results, other models will be used to validate the protocol for geometries and vibrational frequencies. A considerable amount of theoretical and experimental work on the geometries and vibrational frequencies of halophosphoranes is available, so these systems will be used to validate the use of the chosen method for reproducing geometries and vibrational frequencies.

Energies were calculated for isomers of the various halophosphoranes to determine which isomers were the lowest in energy and what the energy difference was. This was done to support the geometries given experimentally, which were then used in the reports on geometries, vibrational frequencies, and NMR shieldings. Geometry optimizations were carried out on these isomers using the B3LYP functional with the 6-311+G* basis set

(6-311++G** when hydrogens were present) in the Gaussian 03 suite of programs [91]. Minima were obtained when the energies and forces reached a default threshold defined in Gaussian 03. These minima were confirmed via vibrational analysis. For calculation of the energy differences between different isomers, it is expected that the MP2 and VSXC results will be similar and that the trend will be reproduced, so these results will not be reported. The energies reported for comparison of apical and equatorial placement of ligands are zero-point corrected.

Geometries reported were optimized at the VSXC level of theory with the 6-311+G* Pople basis set using the Gaussian 03 suite of programs [91]. Where hydrogens were present, the 6-311++G** basis set was used. A larger basis set than the one chosen could have been used, however, this basis set was chosen to validate it for future studies on the phosphoryl transfer reaction of bacterial phosphatidylinositol-specific phospholipase C. A larger basis set for the enzyme study would have been prohibitive. Minimization was carried out until the default cutoffs for the energies and forces in Gaussian 03 were reached. Stationary points were confirmed via vibrational analysis. For all reported structures, no imaginary frequencies were found. For the halophosphoranes, comparisons were made with analogous calculations done at the MP2 and B3LYP level of theory. These were done to compare the VSXC functional with the popular B3LYP functional and the simplest post-HF method, MP2. The purpose was to validate the use of the VSXC functional for use in describing pentacoordinate phosphorane models.

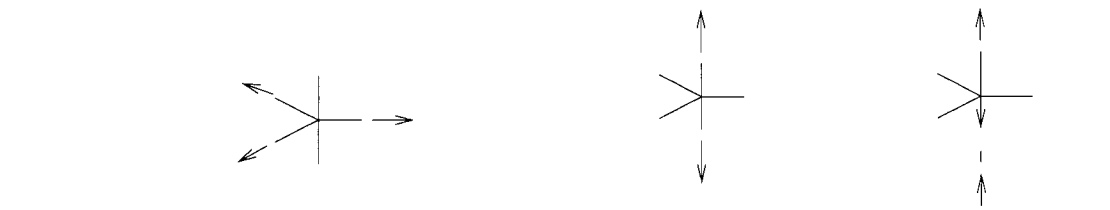
Harmonic frequencies were also calculated with the method and basis sets described above. Scale factors for vibrational frequencies are often applied to generate results more

comparable to experiment. However, in this case no scale factors were available for either the basis set or method used. Therefore, calculated harmonic frequencies are reported as unscaled values. In all cases, only the frequencies of modes which include phosphorus are reported. Since the proper description of pentacoordinate phosphorus was the goal of the work, these frequencies are the only ones of interest.

5.4 Structures and Vibrational Frequencies of Selected Halophosphoranes

Figures 5.1 and 5.2 show the vibrational frequencies for trigonal bipyramids with D_{3h}, C_{2v}, and C_{3v} symmetry. The descriptions of the vibrational modes given in the tables to follow correspond to the diagrams in the figure. The modes shown correspond to the ideal case where the only mode of motion that occurs is shown in the diagram. For each vibrational mode, the symmetry under the corresponding point group is given. A and B refer to nondegenerate modes and E refers to doubly degenerate modes. A modes are symmetric and B modes are asymmetric with respect to rotation about the principle axis. Numerical subscripts denote symmetry of rotation about an axis other than the principle axis. The superscripts ' and " refer to the symmetry with respect to reflection through σ_h.

In many cases, there was more than one minimum geometry found. As described in detail in the previous chapter, ligands can occupy both apical and equatorial positions in a trigonal bipyramid. In the case of the fluorophosphoranes, the investigation of the energetics for all possible cases for apical-equatorial positioning of ligands is reported.



mode		*symm. equatorial stretch*	*symm. axial stretch*	*asymm. axial stretch*
symmetries	D_{3h}	A'_1	A'_1	A''_2
	C_{2v}	A_1	A_1	B_2
	C_{3v}	A_1	A_1	A_1

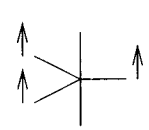
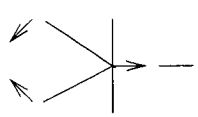
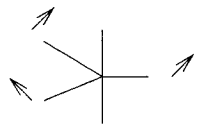
mode		*asymm. equatorial stretch*	*asymm. equatorial stretch*
symmetries	D_{3h}	E'	E'
	C_{2v}	A_1	B_1
	C_{3v}	E	E

Figure 5.1: Diagram of the stretching vibrational modes of a trigonal bipyramid.

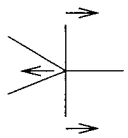
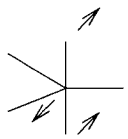
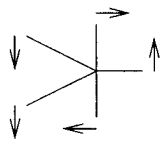
However, the structures and vibrational frequencies reported correspond to the isomer found experimentally.

5.4.1 PF_5 and PCl_5

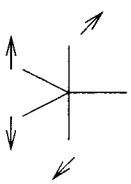
Table 5.1 shows the apical and equatorial P-F and P-Cl bond lengths calculated at the MP2, B3LYP, and VSXC levels of theory. These are compared to experimentally generated structures [1]. Overall, all of the methods reproduce the P-F and P-Cl bond lengths. For all three methods, the minimum energy structure is a trigonal bipyramid.

mode		*out–of–plane bend*	*equatorial in–plane bend*	*equatorial in–plane bend*
symmetries	$\mathbf{D_{3h}}$	A''_2	E'	E'
	$\mathbf{C_{2v}}$	B_2	A_1	B_1
	$\mathbf{C_{3v}}$	A_1	E	E

mode		*axial bend*	*axial bend*	*rock*
symmetries	$\mathbf{D_{3h}}$	E'	E'	E''
	$\mathbf{C_{2v}}$	A_1	B_1	B_2
	$\mathbf{C_{3v}}$	E	E	E



mode		*twist*
symmetries	$\mathbf{D_{3h}}$	E''
	$\mathbf{C_{2v}}$	A_2
	$\mathbf{C_{3v}}$	E

Figure 5.2: Diagram of the bending and torsion vibrational modes of a trigonal bipyramid.

The general trend in PX_5 systems where the apical bonds are longer than the equatorial bonds is reproduced. In all cases, the ratio of P-X_{ax} to P-X_{eq} for chlorine is larger than for fluorine, as is expected from comparison of experimental results.

Table 5.1: Selected Structural Parameters for PF_5 and PCl_5.

coordinate	experiment	MP2	B3LYP	VSXC
PF_5				
$P-F_{eq}$ (Å)	1.534	1.562	1.570	1.569
$P-F_{ax}$ (Å)	1.577	1.596	1.604	1.602
PCl_5				
$P-Cl_{eq}$ (Å)	2.020	2.041	2.072	2.067
$P-Cl_{ax}$ (Å)	2.124	2.151	2.189	2.166

The MP2 method gives the best description of these bonds as compared to experimental values. All of the bond lengths are within 0.030 Å of the experimental value. For both P-F and P-Cl bond lengths, MP2 overestimates their values.

VSXC gives results that agree well with B3LYP. The VSXC P-F bond lengths are within 0.002 Å of the B3LYP results. The situation is worse for the P-Cl bond lengths, where the two functionals differ by as much as 0.023 Å. B3LYP and VSXC bond lengths are both comparable to the MP2 results. For the P-Cl_{ax} bond, the VSXC result is the closest to experiment, followed by B3LYP. Therefore, for PX_5 systems, the VSXC geometries will

agree well with B3LYP and MP2.

Table 5.2: Selected Frequencies for PF_5 and PCl_5.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PF_5				
PF_3 symm. stretch	816	782	756	751
PF_2 symm. stretch	648	625	610	610
PF_2 antisymm. stretch	947	945	916	916
PF_3 antisymm. stretch	1025	996	966	948
PF_2 axial bend	533	518	501	487
PF_3 in-plane bend	174	164	158	149
PF_3 symm. out-of-plane bend	575	558	541	535
rocking	520	498	479	476
PCl_5				
PCl_3 symm. stretch	392	391	358	357
PCl_2 symm. stretch	286	283	257	270
PCl_2 antisymm. stretch	446	458	407	459
PCl_3 antisymm. stretch	586	569	525	547
PCl_2 axial bend	276	271	255	248
PCl_3 in-plane bend	90	88	82	77
PCl_3 symm. out-of-plane bend	299	310	289	289
rocking	261	268	249	247

Experimentally, PF_5 and PCl_5 exhibit D_{3h} symmetry. They are trigonal bipyramids and both have 12 vibrational degrees of freedom. Diagrammatically, these are shown in Figures 5.1 and 5.2 for a generic trigonal bipyramid. Also given are the symmetries of the modes under the D_{3h}, C_{2v}, and C_{3v} space groups. The above theoretical methods also show D_{3h} symmetry for both PF_5 and PCl_5. Experimentally, only eight vibrational modes are observed. These, along with the corresponding theoretically assigned values, are given in Table 5.2.

An analysis of the 12 vibrational modes of PF_5 and PCl_5 shows eight fundamental transitions. Four of these are degenerate. They are the pairs of asymmetric equatorial

stretches, equatorial in-plane bends, axial bends, and the rock and twist pair. From symmetry considerations of the D_{3h} point group, the symmetric equatorial and axial stretches, as well as the rock and twist are expected to be IR inactive, but Raman active. Although no theoretical Raman intensities for these molecules were obtained, the IR inactive modes were confirmed from the calculated intensities.

For each of the three theoretical methods, the harmonic frequencies are reported. These are compared to the anharmonic frequencies of these systems from experiment. The three theoretical methods employed are able to reproduce the experimentally determined fundamental vibrations of PF_5 and PCl_5. Considering that a comparison of the theoretical harmonic frequencies to the experimental anharmonic frequencies is expected to make agreement worse, this is a fortuitous result. In spite of this, all three methods generally underestimate the frequencies of these two molecules. The MP2 method gives the best agreement with experiment as compared to B3LYP and VSXC. The VSXC results are nearly identical to B3LYP for most of the frequencies calculated and are generally less in magnitude.

5.4.2 PF_4CH_3, PF_4CF_3, and PCl_4CF_3

Pentacoordinate PY_4CX_3, where Y is F or Cl and X is F or H, are similar to PF_5 and PCl_5, but they have a lower symmetry than D_{3h}. These molecules have four identical ligands that span apical and equatorial positions. The fifth nonidentical ligand can be placed either apically or equatorially. Apical positioning gives C_{3v} symmetry, while equatorial positioning gives C_{2v} symmetry. Preference for apical or equatorial positioning will

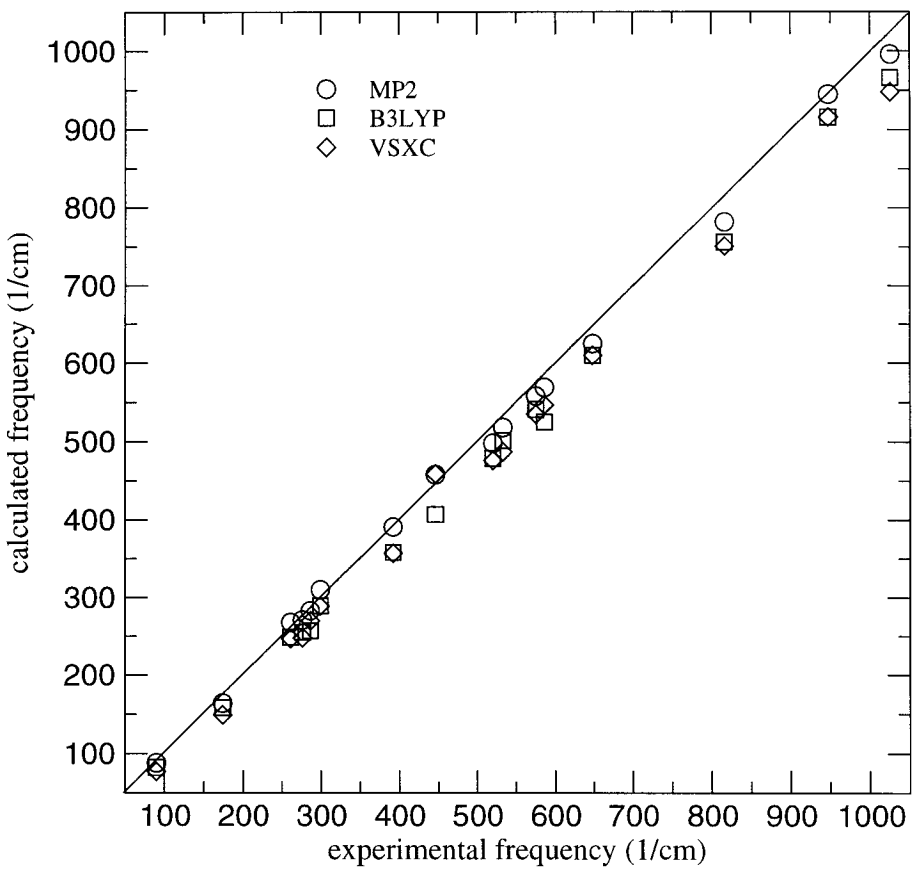


Figure 5.3: Harmonic frequencies for PF_5 and PCl_5.

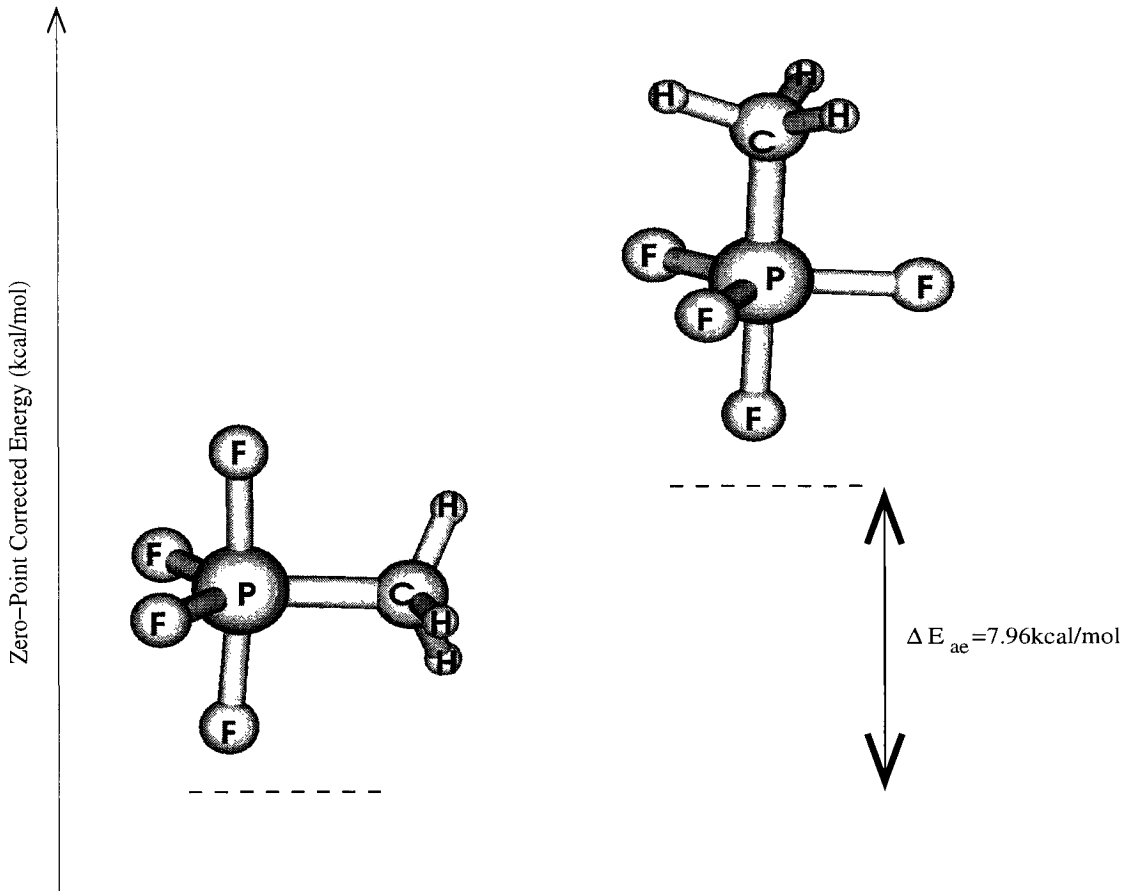


Figure 5.4: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CH_3 group on PF_4CH_3.

depend primarily on the sizes and electronegativities of the two types of ligands. Experimental observations of the symmetries of PF_4CF_3, PF_4CH_3, and PCl_4CF_3 give C_{2v}, C_{2v}, and C_{3v}, respectively [1]. These assignments have been based on the infrared spectra of these species. However, a microwave spectrum of PF_4CF_3 gives C_{3v} symmetry. A more recent investigation into the infrared spectrum of this molecule shows the two isomers are separated by about 2 kcal/mol and may exist in equilibrium with one another [134]. This observation is not seen with either PF_4CH_3 or PCl_4CF_3.

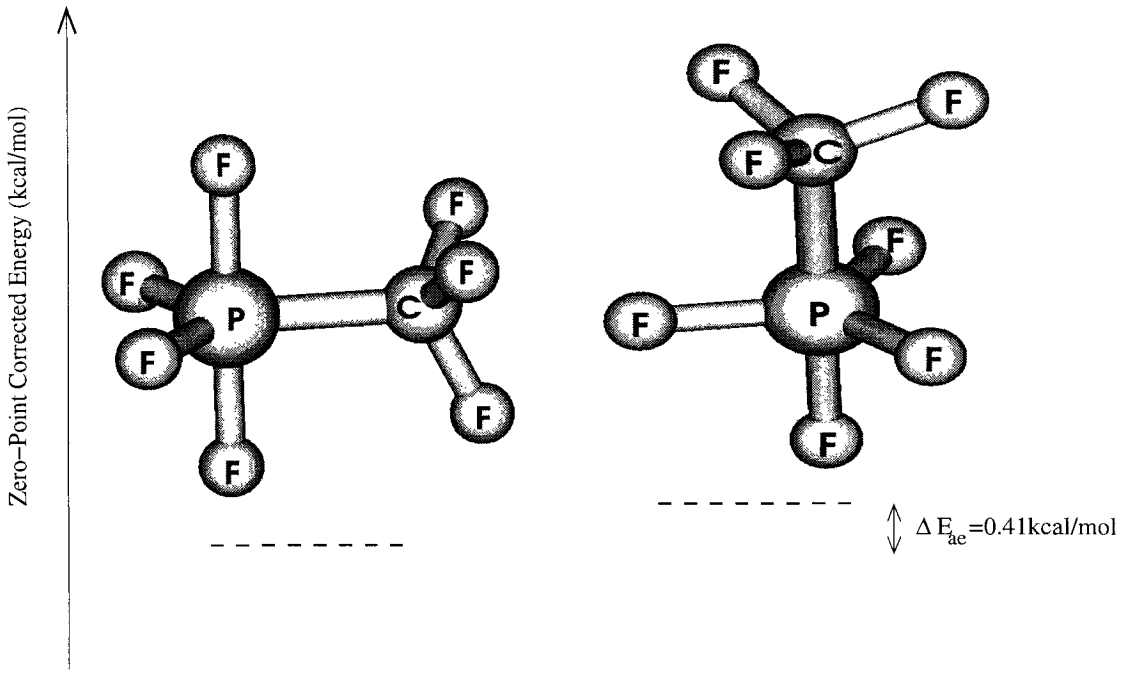


Figure 5.5: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CF_3 group on PF_4CF_3.

Figure 5.4 depicts an energy diagram representing the apical and equatorial positioning of the CH_3 group of PF_4CH_3. Frequency analyses of both structures indicate they are minima. The reported energy difference is $\Delta\ E_{ae} = E_{apical}$ - $E_{equatorial}$. From this result, the isomer with the methyl group placed equatorially is 7.96 kcal/mol lower in energy than if the methyl group was placed apically. This indicates that the equatorial positioning of the methyl group is favored over the apical positioning. On the basis of the sizes and electronegativities of the fluorine and methyl ligands, this is the expected result. The smaller and more electronegative groups prefer the apical position to the equatorial position.

The isomers showing apical and equatorial placement of CF_3 and the energy difference between them for PF_4CF_3 is depicted in Figure 5.5. In the case of the CH_3 placement

in PF_4CH_3, the electronegativity difference is larger. Using group electronegativities for CH_3 given in Reference 135, the difference in electronegativities is 1.8 for the methyl and fluorine ligands. According to the same group electronegativities, the difference for F and CF_3 is 1.3. Therefore, one would expect the CF_3 group to show relatively strong preference for the equatorial position, as is seen with CH_3. However, the calculated energy difference is only 0.41 kcal/mol. The preferred site is the equatorial position. However, this energy difference is small enough that a clear assignment of equatorial positioning cannot be made within the error of the calculation.

A spectroscopic and thermodynamic analysis of the apical-equatorial equilibrium for PF_4CF_3 is given in Ref. 134. Experimental evidence indicates an enthalpy difference of 1.03 kcal/mol in favor of the apical positioning of the CF_3 group. An entropy difference of 3.7 cal mol^{-1} K^{-1} was also determined, with the entropy of the equatorial isomer being larger than for the apical isomer. The equilibrium constant for the apical-equatorial interconversion was found to be between 0.10 and 1.59 for temperatures from 149 K to 379 K. From this data, it was concluded that an equilibrium exists between the two isomers, even at low temperatures. Therefore, there is little difference in energy for apical versus equatorial positioning of the CF_3 group. These results, however, seem to contradict the arguments for strong equatorial preference based on size and electronegativity. The theoretical enthalpy difference calculated in this work is 0.45 kcal/mol in favor of the equatorially substituted isomer, which is similar to the experimental result.

A diagram showing the energy difference between apical and equatorial positioning of the CF_3 group in PCl_4CF_3 is given in Figure 5.6. The difference in energy found in

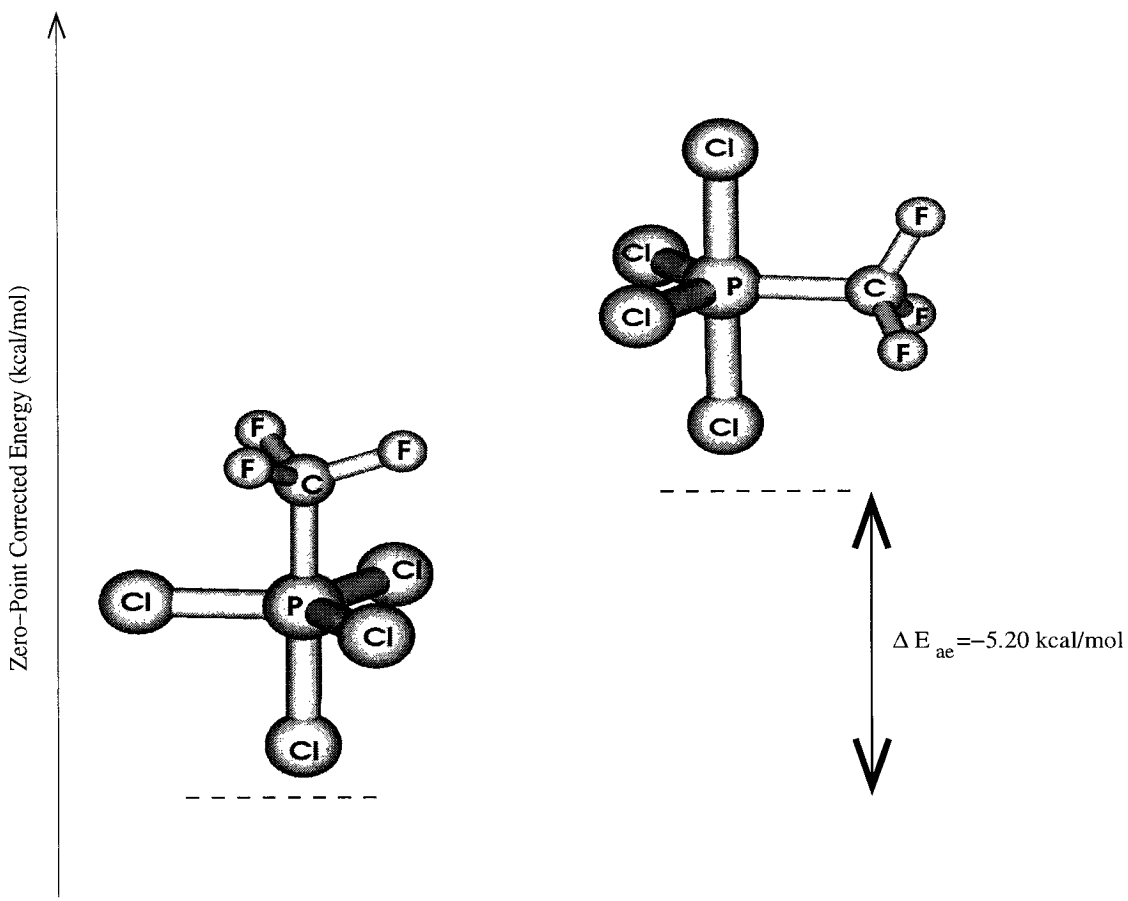


Figure 5.6: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CF_3 group on PCl_4CF_3.

placing the CF_3 group apically versus equatorially was -5.20 kcal/mol. This indicates that in PCl_4CF_3, the CF_3 group prefers to be positioned apically as opposed to equatorially. The energy difference is large enough that it indicates a strong preference for this isomer. This is in contrast with the expected result. First, the electronegativity of CF_3 is given in Reference 135 as 2.71, whereas the Pauling electronegativity for chlorine is 2.9. Therefore, chlorine should prefer to be positioned apically, not equatorially. Also, since the electronegativity difference is small, there should not be a large preference for one site over the other. Since spectroscopic results support the apical positioning of the CF_3 group, then either the CF_3 electronegativity reported in Reference 135 is incorrect or there are other effects which must be accounted for to describe this preference.

The theoretical results reported for the geometries of PF_4CH_3, PF_4CF_3, and PCl_4CF_3 correspond to those with C_{2v}, C_{2v}, and C_{3v} symmetries, respectively. These are consistent

with the experimentally determined symmetries for these systems.

Table 5.3: Selected Structural Parameters for PF_4CH_3 and PF_4CF_3.

coordinate	experiment	MP2	B3LYP	VSXC
PF_4CH_3				
$P - F_{eq}$ (Å)	1.543	1.574	1.586	1.581
$P - C_{eq}$ (Å)	1.780	1.793	1.807	1.797
$P - F_{ax}$ (Å)	1.612	1.631	1.644	1.646
$F_{eq} - P - F_{eq}$ (deg.)	115.6	116.9	117.2	116.1
$F_{eq} - P - F_{ax}$ (deg.)	91.8	89.0	88.4	89.1
PF_4CF_3				
$P - F_{eq}$ (Å)	1.537	1.564	1.574	1.570
$P - C_{eq}$ (Å)	1.881	1.896	1.920	1.888
$P - F_{ax}$ (Å)	1.573	1.607	1.618	1.615
$F_{eq} - P - F_{eq}$ (deg.)	117.4	117.4	117.3	118.4
$F_{eq} - P - F_{ax}$ (deg.)	90.0	90.4	90.2	90.7

Table 5.3 gives the structural parameters for PF_4CH_3 and PF_4CH_3 [1, 136]. The apical P-F bonds in both cases are longer than the equatorial P-F bonds. The P-F bonds in PF_4CH_3 are nearly identical to the analogous P-F bonds in PF_4CF_3. The largest difference is in the P-C bond. In PF_4CH_3, this bond is shorter than in PF_4CF_3. This may be due to repulsion between the lone pairs on the fluorines attached to carbon and the fluorines attached to phosphorus.

Overall, the structural parameters in PF_4CH_3 and PF_4CF_3 are well reproduced with MP2, B3LYP, and VSXC. As with PF_5 and PCl_5, bond lengths calculated using MP2 are all overestimated. MP2 is slightly better at reproducing the geometries of these two molecules as compared to B3LYP and VSXC. However, the results for all three methods are similar. B3LYP and VSXC give very similar results and both overestimate the bond

lengths. VSXC, however, gives slightly better results for these two molecules than B3LYP.

Table 5.4: Selected Structural Parameters for PCl_4CF_3.

coordinate	MP2	B3LYP	VSXC
$P - Cl_{ax}$ (Å)	2.133	2.176	2.146
$P - Cl_{eq}$ (Å)	2.048	2.083	2.072
$P - C$ (Å)	1.936	1.974	1.934
$Cl_{eq} - P - Cl_{eq}$ (deg.)	119.9	120.0	119.9
$Cl_{eq} - P - Cl_{ax}$ (deg.)	91.4	90.8	91.5
$Cl_{eq} - P - C$ (deg.)	88.6	89.2	88.5

Table 5.4 gives the structural parameters for PCl_4CF_3. Although no experimental structural data was found for this molecule, comparisons between MP2, B3LYP, and VSXC can still be made. Overall, VSXC has the best agreement with MP2. B3LYP differs significantly from MP2. Longer apical P-Cl bond lengths are seen with all three methods. This is expected as apical P-X bond lengths are generally longer than equatorial P-X bond lengths. A slight distortion in the angles about phosphorus is also seen. This is expected as a substitution of a chlorine for the larger CF_3 group increases repulsions, distorting the trigonal bipyramid from its ideal geometry seen in PCl_5.

Comparisons of the structures of PF_4CF_3 and PCl_4CF_3 shows the P-C bond in PF_4CF_3 is shorter than in PCl_4CF_3. PCl_4CF_3 angles indicate a slight distortion from the ideal trigonal bipyramid values of 120° and 90°. This is also seen with the angles reported for PF_4CF_3, most notably the F_{eq}-P-F_{eq} angles. These trends are consistently

reproduced with all three theoretical methods.

Table 5.5: Selected Frequencies for PF_4CH_3.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PF_2 symm. stretch [a]	932	927	891	893
PC stretch [b]	725	715	679	686
PF_2 asymm. stretch [c]	1009	997, 856	978, 843	962, 823
PF_2' symm. stretch	596	579	560	561
PF_2' asymm. stretch	843	831	799	792
PF_2' bend	514	502	480	475
PF_2' bend	467	452	436	419
CPF_2 in-plane bend	179	168	166	148
PF_2 in-plane bend	179	171	159	158
CPF_2 out-of-plane bend	538	521	505	493
CPF_2F_2' rock	412	390	366	378
PF_2F_2' twist	397	480	464	453

[a] In Ref. 1, the authors identify this frequency as an equatorial P-F symmetric stretch. The theoretical vibrational analysis shows this frequency is an equatorial P-F symmetric stretch that also contains a significant contribution from the P-C stretch.

[b] In Ref. 1, the authors identify this frequency as a P-C symmetric stretch. The theoretical vibrational analysis shows this frequency is a P-C symmetric stretch which contains a significant contribution from the equatorial P-F symmetric stretch.

[c] In Ref. 1, the authors identify this frequency as an asymmetric equatorial P-F stretch. The theoretical vibrational analysis shows two frequencies, both with significant contributions from the asymmetric equatorial stretch.

Tables 5.5, 5.6, and 5.7 give the experimentally assigned vibrational frequencies for PF_4CH_3, PF_4CF_3, and PCl_4CF_3, respectively. Only those frequencies that include phosphorus have been included. These molecules each have 21 vibrational degrees of freedom, 12 of which are the vibrations for a trigonal bipyramid. However, assignment for a trigonal bipyramid of C_{2v} (or C_{3v} in the case of PCl_4CF_3) symmetry will not exhibit exactly the vibrations shown in Figures 5.1 and 5.2. As a result, the three tables indicate the

experimentally assigned labels for the given vibrations.

Table 5.6: Selected Frequencies for PF_4CF_3.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PF_2 symm. stretch [a]	909	882, 755	853, 738	845, 718
PC stretch	423	315	299	313
PF_2 asymm. stretch	986	987	952	948
PF_2' symm. stretch	674	647	625	622
PF_2' asymm. stretch	892	900	868	859
PF_2' bend	556	551	532	527
PF_2' bend	482	474	457	442
CPF_2 in-plane bend	170	94	92	74
PF_2 in-plane bend	130	172	163	148
CPF_2 out-of-plane bend [b]	579	563, 522	549, 506	535, 494
CPF_2F_2' rock	384	352	338	350

[a] In Ref. 1, the authors attribute this frequency to a stretch of the two apical P-F bonds. The theoretical calculations show this mode to be a P-F_2 stretch as well, but with a significant contribution from the P-C stretch.

[b] In Ref. 1, the authors attribute this frequency to the bend of the CF_3 and the two equatorial fluorines out of the plane they form. The theoretical vibrational analysis shows two frequencies corresponding to this mode. The lower frequency corresponds to the bend of the two equatorial fluorines from the CPF_2 plane. The higher frequency corresponds to the bend of the CF_3 group out of this same plane.

Table 5.7: Selected Frequencies for PCl_4CF_3.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PC axial stretch	493	518	471	513
PCl axial stretch	324	309	282	289
PCl_3 symm. stretch	374	377	341	345
PCl_3 antisymm. stretch	587	581	514	509
$CPCl$ axial bend	304	337	314	323
PCl_3 symm. out-of-plane deformation	260	248	223	241
PCl_3 antisymm. deformation	94	90	81	87
rocking	239	265	246	243

Figure 5.7 shows the MP2, B3LYP, and VSXC harmonic frequencies for PF_4CH_3, PF_4CF_3, and PCl_4CF_3 compared to experiment. As with PF_5 and PCl_5, all the methods underestimate the majority of the frequencies for all three systems. MP2 gives frequencies

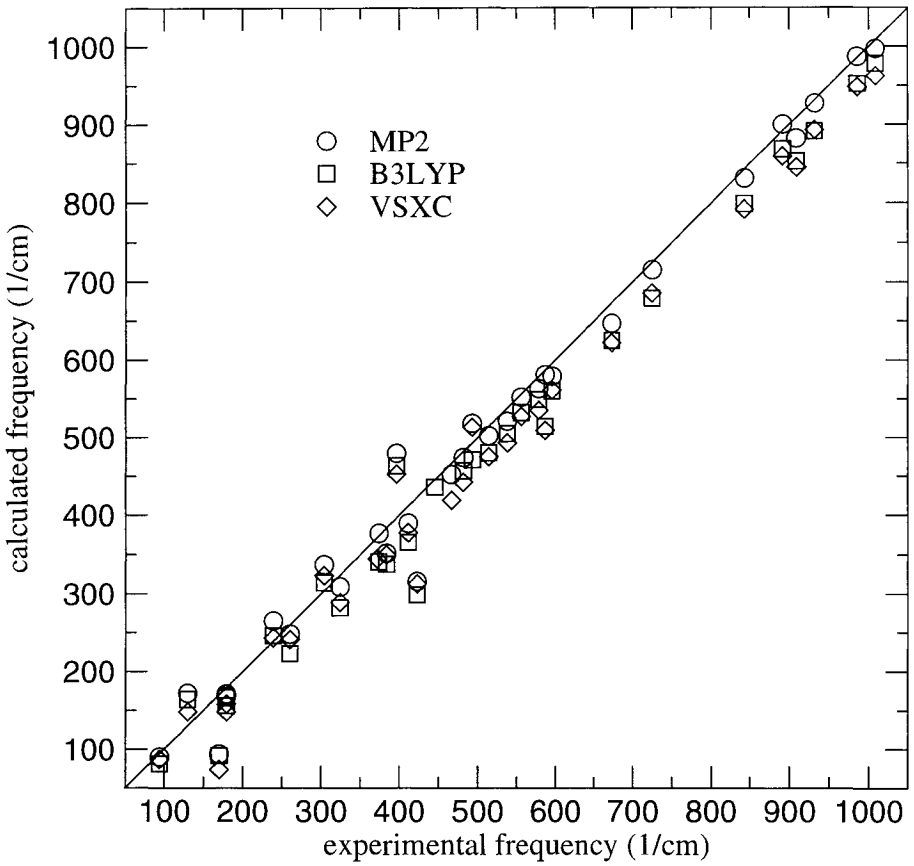


Figure 5.7: Selected harmonic frequencies for PF_4CF_3, PCl_4CF_3, and PF_4CH_3.

which are most accurate as compared to experiment. The B3LYP and VSXC calculated frequencies are similar to one another in quality. In some cases, the B3LYP frequencies are more accurate than VSXC. In other cases, the reverse is true.

Two trends are worth noting. In PF_4CH_3, the frequencies involving apical ligands are lower than those found for PF_4CF_3, and those involving equatorial ligands are higher. In PCl_4CF_3, all of the frequencies are lower than for PF_4CH_3 and PF_4CF_3. This is observed experimentally, and all three theoretical methods reproduce these trends.

5.4.3 $PF_3(CH_3)_2$ and $PF_3(CF_3)_2$

$PF_3(CH_3)_2$ and $PF_3(CF_3)_2$ are trigonal pyramids that have three fluorines and two CX_3 groups, where X is either F or H. The positioning of these groups depends on their preference for apical or equatorial placement. Apical placements of both CX_3 groups results in D_{3h} symmetry. Equatorial placements of both CX_3 groups results in C_{2v} symmetry. Apical placement of one CX_3 group and equatorial placement of the other CX_3 group results in C_s symmetry. Experimental results indicate C_{2v} symmetry, indicating diequatorial placement of the CX_3 groups. This result is expected due to the higher electronegativity and smaller size of fluorine relative to CF_3 and CH_3.

A diagram of the analysis of the energies of the isomers resulting from the apical and equatorial positioning of the CH_3 groups in $PF_3(CH_3)_2$ is given in Figure 5.8. Diequatorial positioning is more energetically favorable than diapical positioning by 12.64 kcal/mol. Diapical positioning is more energetically favorable than apical-equatorial positioning by 2.07 kcal/mol. This is expected since placement of the methyl groups as apical-equatorial places them closest together, as compared to apical-apical or equatorial-equatorial positioning. With both methyl groups placed apically, they are 180° apart. With both methyl groups placed equatorially, they are 120° apart. With one methyl placed apically and one methyl placed equatorially, they are placed 90° apart. These results also explain the slight distortion seen in the Y_{ax}-P-F_{ax} angle when one or both the methyl groups are placed equatorially.

Figure 5.9 shows an energy diagram analogous to the one in Figure 5.8. In this case, only one energy difference was calculated. The difference in energy between apical-apical

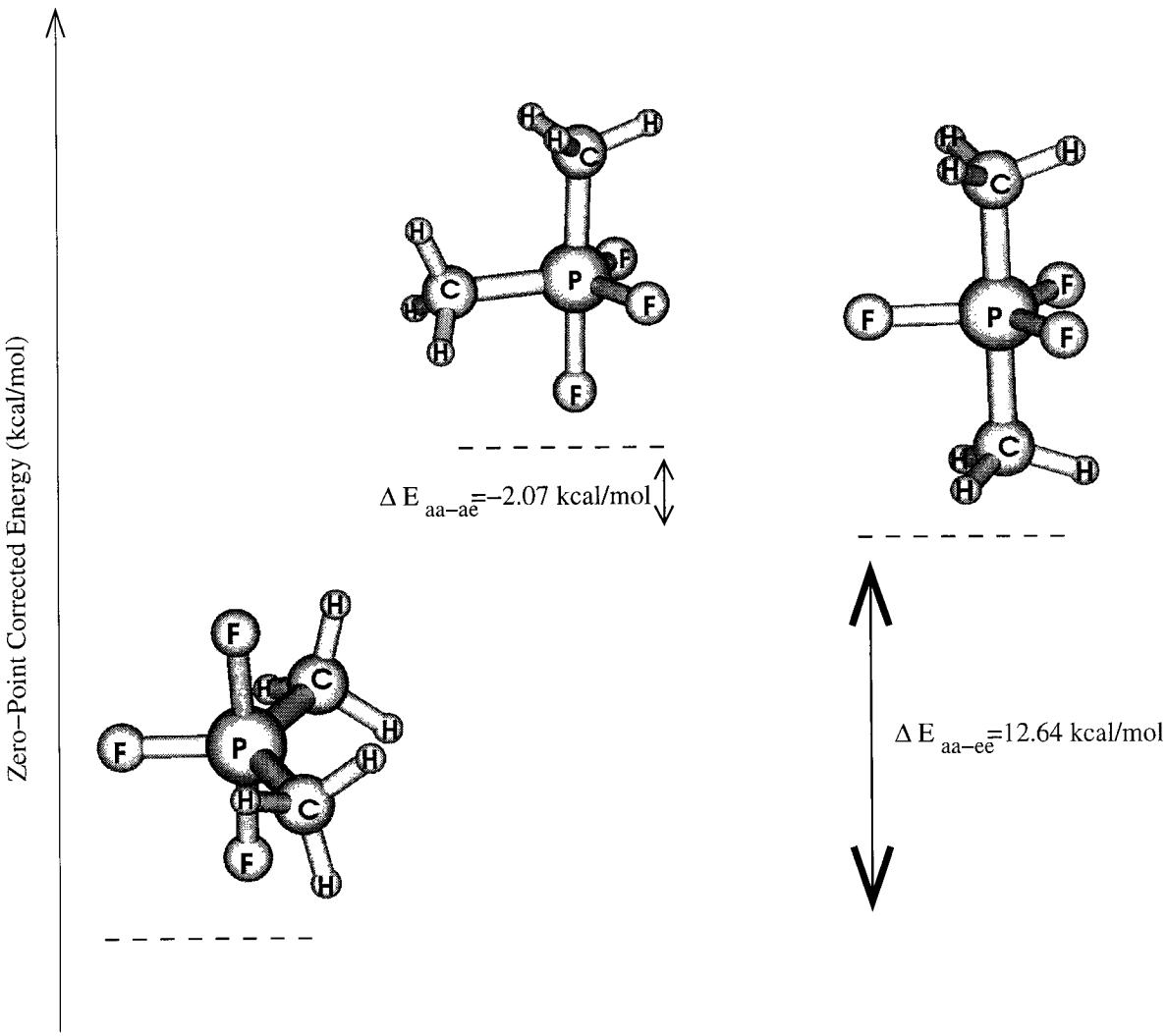


Figure 5.8: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CH_3 groups on $PF_3(CH_3)_2$.

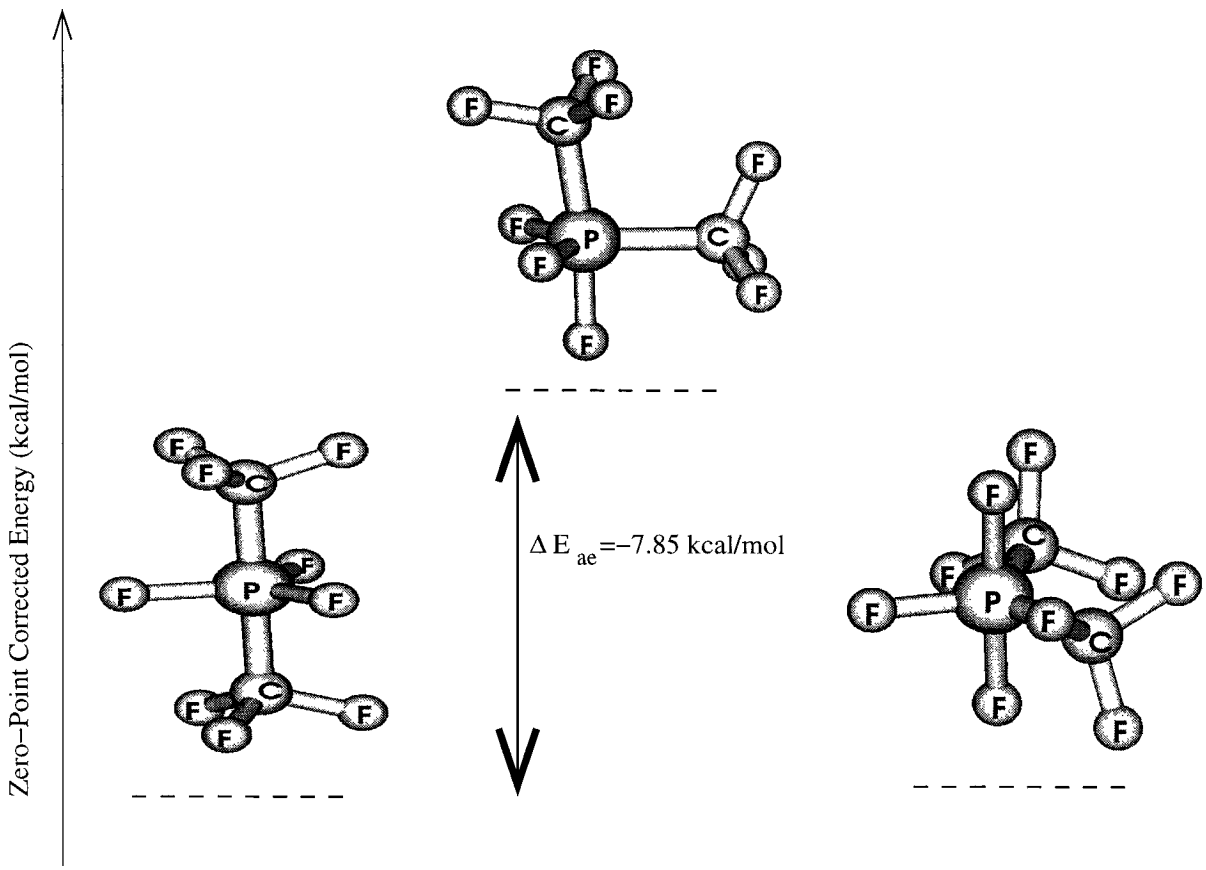


Figure 5.9: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CF_3 groups on $PF_3(CF_3)_2$.

and equatorial-equatorial positioning of the CF_3 groups is roughly 0.1 kcal/mol, with the diapical positioning of the CF_3 groups being the more stable isomer. This energy difference is much too small to distinguish which isomer is more stable. The diapical positioning of CF_3 groups is 7.85 kcal/mol more favorable than the apical-equatorial positioning. As in the case of $PF_3(CH_3)_2$, this energy difference results from a combination of factors. Equatorial positioning of a CF_3 group would be expected to be favored due to the preference of fluorine for the apical position. However, a much stronger effect is the repulsion between the fluorines in the CF_3 groups. It is larger for apical-equatorial positioning than either apical-apical or equatorial-equatorial positioning. It is believed that relief of this repulsion by diapical or diequatorial positioning of the CF_3 groups is the largest effect contributing to the observed energy difference.

Table 5.8: Selected Structural Parameters for $PF_3(CH_3)_2$.

coordinate	experiment	MP2	B3LYP	VSXC
$P - F_{eq}$ (Å)	1.553	1.589	1.599	1.594
$P - C_{eq}$ (Å)	1.798	1.801	1.815	1.804
$P - F_{ax}$ (Å)	1.643	1.675	1.689	1.690
$C_{eq} - P - C_{eq}$ (deg.)	124	125	126	127
$F_{eq} - P - F_{ax}$ (deg.)	88.9	88.3	88.0	88.5

Table 5.8 shows the experimentally determined structural parameters and those determined at the MP2, B3LYP, and VSXC levels of theory for $PF_3(CH_3)_2$. These parameters are reported for the C_{2v} symmetry found by experiment and supported by the given energy analysis diagrams. The apical P-F bonds are longer than the equatorial P-F bonds. The C-P-C and F_{eq}-P-F_{ax} angles are also given, and are distorted from the ideal values in

the trigonal bipyramid. This distortion results from repulsion between the hydrogens of the methyl groups placed equatorially. Distortion also occurs between the larger methyl groups and the apically placed fluorines.

The theoretically determined parameters also indicate longer P-F apical bonds and distorted angles as indicated by experiment. The MP2 method gives the best reproduction of the experimentally determined geometries. All three methods overestimate the bond lengths. VSXC gives results that are more accurate than B3LYP.

Table 5.9: Selected Structural Parameters for $PF_3(CF_3)_2$.

coordinate	MP2	B3LYP	VSXC
$P - F_{eq}$ (Å)	1.569	1.578	1.582
$P - C_{eq}$ (Å)	1.896	1.922	1.888
$P - F_{ax}$ (Å)	1.625	1.636	1.638
$C_{eq} - P - C_{eq}$ (deg.)	123.7	125.8	102.7
$C_{eq} - P - F_{eq}$ (deg.)	118.2	117.1	128.5
$C_{eq} - P - F_{ax}$ (deg.)	87.7	87.9	93.4
$F_{eq} - P - F_{ax}$ (deg.)	90.8	90.8	88.7

No experimentally determined spectra were found for $PF_3(CF_3)_2$. Ref. 136 gives geometric parameters for apical-apical positioning of the CF_3 groups, but no experimental structures for diequatorial positioning. However, the equatorial P-F and the P-C bond lengths for the diequatorial isomer should be similar to those found for the diapical isomer. Hence, the theoretical results for these two bonds will be compared to the analogous bonds in the experimentally determined structure. However, the accuracy can only be fully determined by comparing the experimental diequatorial isomer with the theoretical calculations. The equatorial P-F bond length is 1.559 Å, while the P-C bond length is

1.884 Å. The theoretical results are shown in Table 5.9. The theoretical values for the equatorial P-F bond lengths are comparable to experiment for all three methods. The P-C bond length calculated with all three methods is also comparable to the experimental value. It is then expected that the results for the diapical positioning of the CF_3 groups will be accurately reproduced by the above three theoretical methods.

MP2 bond lengths are correctly reproduced with B3LYP and VSXC. For B3LYP and MP2, the equatorial-equatorial angles are similar, however some discrepancy in the VSXC determined angles is seen. It is likely that there is a problem with the VSXC description of this system.

P-C bond lengths in $PF_3(CF_3)_2$ are longer than their counterparts in $PF_3(CH_3)_2$, whereas the apical and equatorial P-F bond lengths are shorter. This is reasonable since it is expected that the P-C bond lengths would be longer for the CF_3 ligand than the methyl ligand. This bond elongation results from repulsion between fluorines on the CF_3 groups and causes the equatorial and apical P-F bonds to shorten. This trend is seen for all three

theoretical methods.

Table 5.10: Selected Frequencies for $PF_3(CH_3)_2$.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PF stretch	836	836	813	807
PC_2 symm. stretch	675	676	637	651
PC_2 asymm. stretch	779	817	789	790
PF_2' symm. stretch	540	539	518	509
PF_2' asymm. stretch	755	731	698	669
PF_2' axial bend	440	424	409	390
PF_2' axial bend	459	456	436	430
PC_2 in-plane bend	184	165	163	144
C_2PF in-plane bend	184	186	180	170
C_2PF out-of-plane bend	496	371	352	354
C_2PFF_2' rock	471	477	461	449
C_2PF_2' twist	401	390	370	382

Table 5.10 gives the experimentally assigned vibrational frequencies for $PF_3(CH_3)_2$. As for the PY_4CX_3 series, only those frequencies that include phosphorus have been included. $PF_3(CH_3)_2$ has 30 vibrational degrees of freedom. 12 degrees of freedom result from the vibrations for a trigonal bipyramid. The table indicates the experimentally assigned labels for the given vibrations. The highest frequency vibrations given in the table result from the symmetric and asymmetric stretching modes. This is followed by the bending modes.

Also shown are the frequencies determined using the MP2, B3LYP, and VSXC methods. Overall, the stretching frequencies have the largest value, followed by the bending frequencies. This is also what is observed experimentally. Figure 5.10 is a plot of the experimentally determined frequencies versus the theoretical values for the three methods. This diagram indicates that overall the majority of the frequencies are underestimated

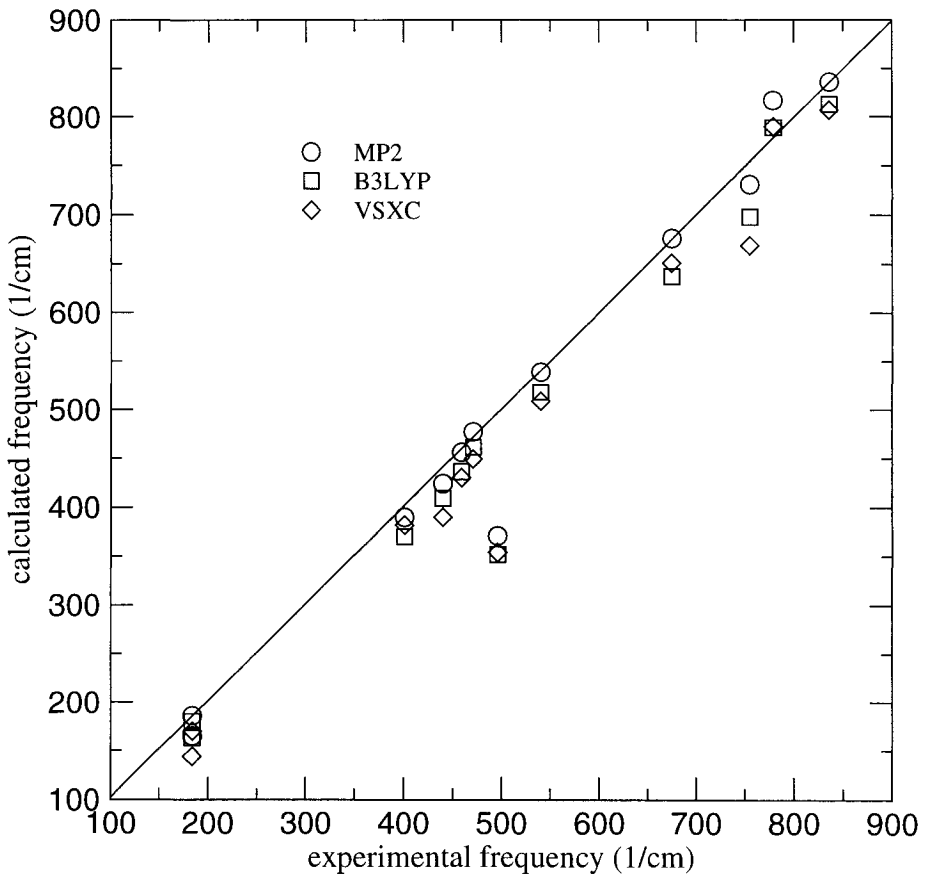


Figure 5.10: Selected harmonic frequencies for $PF_3(CH_3)_2$.

by all three methods. Best agreement with experiment is seen with the MP2 method, followed by B3LYP and VSXC. The agreement with experiment and B3LYP and VSXC is very similar and both methods produce similar values for the frequencies.

5.4.4 $PF_2(CH_3)_3$ and $PF_2(CF_3)_3$

$PF_2(CH_3)_3$ and $PF_2(CF_3)_3$ are trigonal bipyramids with two fluorines and three CX_3 groups, where X is either H or F. The three CX_3 groups can either occupy all three equatorial sites, two equatorial sites and one apical site, or two apical sites and one equatorial site. Experimental results for the vibrational spectra exist only for $PF_2(CH_3)_3$, so only these vibrations will be compared to theory. Experimental spectra indicate $PF_2(CH_3)_3$ adopts D_{3h} symmetry. The resulting geometrical and spectral analyses will closely follow those of PF_5 and PCl_5.

Figure 5.11 shows the energy difference between the isomers of $PF_2(CF_3)_3$ with triequatorial and diequatorial-apical positioning of the CF_3 groups. The isomer with diapical-equatorial placement of the CF_3 groups was not included as a minimum energy structure was not found. A minimization in which the diapical C-P-C angle was constrained to 180° was performed and when no minimum was found, the resulting structure was reoptimized without constraints. The resulting structure was the triequatorial isomer. It is believed that the diapical isomer is not a minimum on the potential energy surface for $PF_2(CF_3)_3$.

The triequatorial isomer was favored by 10.25 kcal/mol. This is reasonable considering the size and electronegativity of CF_3 as opposed to fluorine. Also, triequatorial

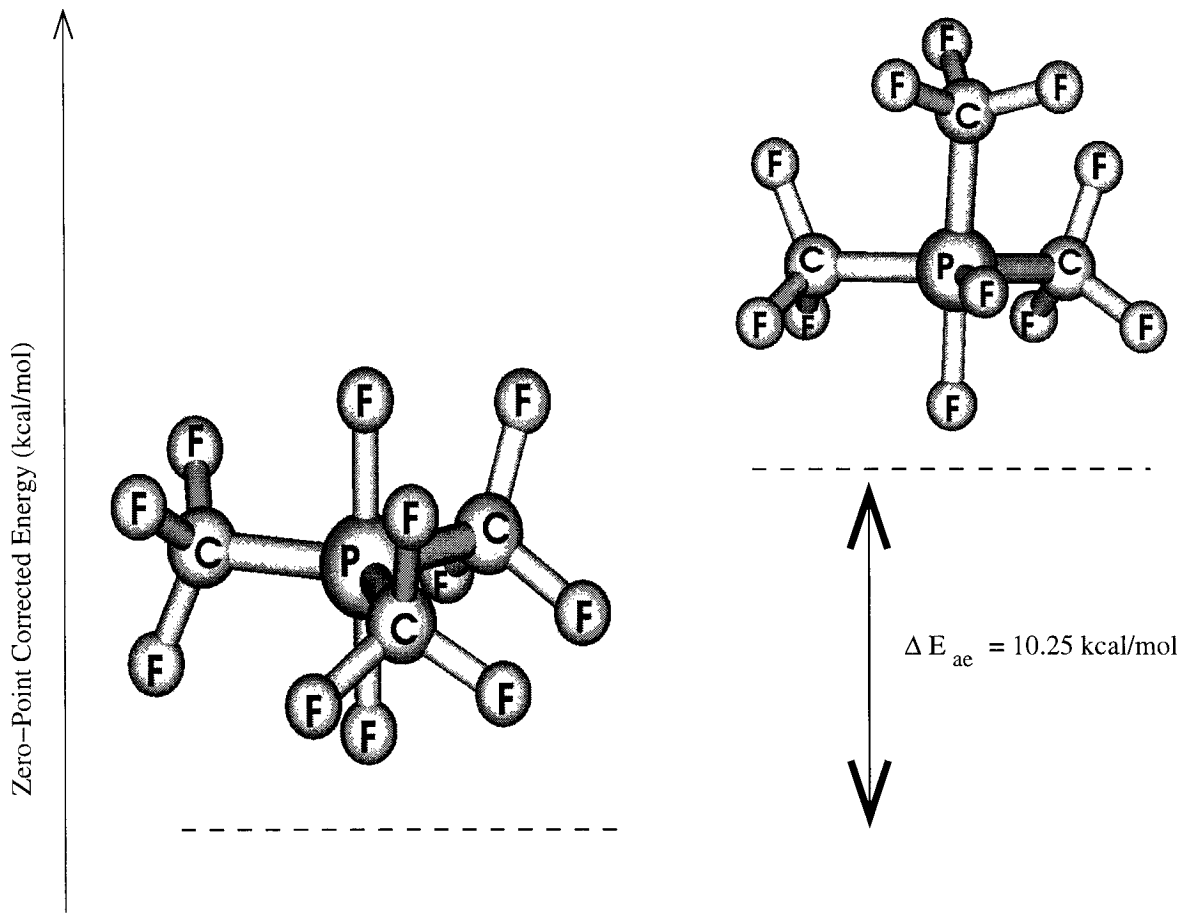


Figure 5.11: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CF_3 groups on $PF_2(CF_3)_3$.

positioning of these groups results in a minimum repulsion between the fluorines of the CF_3 groups and between the CF_3 and apical fluorines.

$PF_2(CH_3)_3$ has three possible isomers based on the placement of the methyl groups. The methyl groups can occupy both apical positions and one equatorial position. They can also occupy all three equatorial positions or two equatorial positions and one apical position. A diagram showing the energy differences between these isomers is shown in Figure 5.12. The diapical-equatorial isomer was obtained by constraining the C_{ap}-P-C_{ap} angle to a value of 180° and optimizing the geometry. It was confirmed to be a minimum. However, when the geometry was relaxed, the triequatorial isomer resulted. This isomer was included to illustrate the large preference of the methyl groups for the equatorial sites. The diequatorial-apical isomer was favored by 9.89 kcal/mol. The triequatorial isomer was favored by 18.00 kcal/mol. Due to its smaller size and larger electronegativity, the fluorines will prefer the apical sites. This rationalization can be used to explain the energy difference between the preference of the methyl groups for the equatorial sites.

Table 5.11: Selected Structural Parameters for $PF_2(CF_3)_3$.

coordinate	experiment	MP2	B3LYP	VSXC
$P - C_{eq}$ (Å)	1.888	1.900	1.926	1.880
$P - F_{ax}$ (Å)	1.600	1.642	1.652	1.658
$F - C - F$ (deg.)	108.3	108.9	108.7	109.5
$C_{eq} - P - C_{eq}$ (deg.)	-	118.1	118.8	123.9
$C_{eq} - P - F_{ax}$ (deg.)	-	91.8	88.7	94.0
$F_{ax} - P - F_{ax}$ (deg.)	-	178.8	179.3	176.5

Tables 5.11 and 5.12 give the structural parameters for $PF_2(CF_3)_3$ and $PF_2(CH_3)_3$,

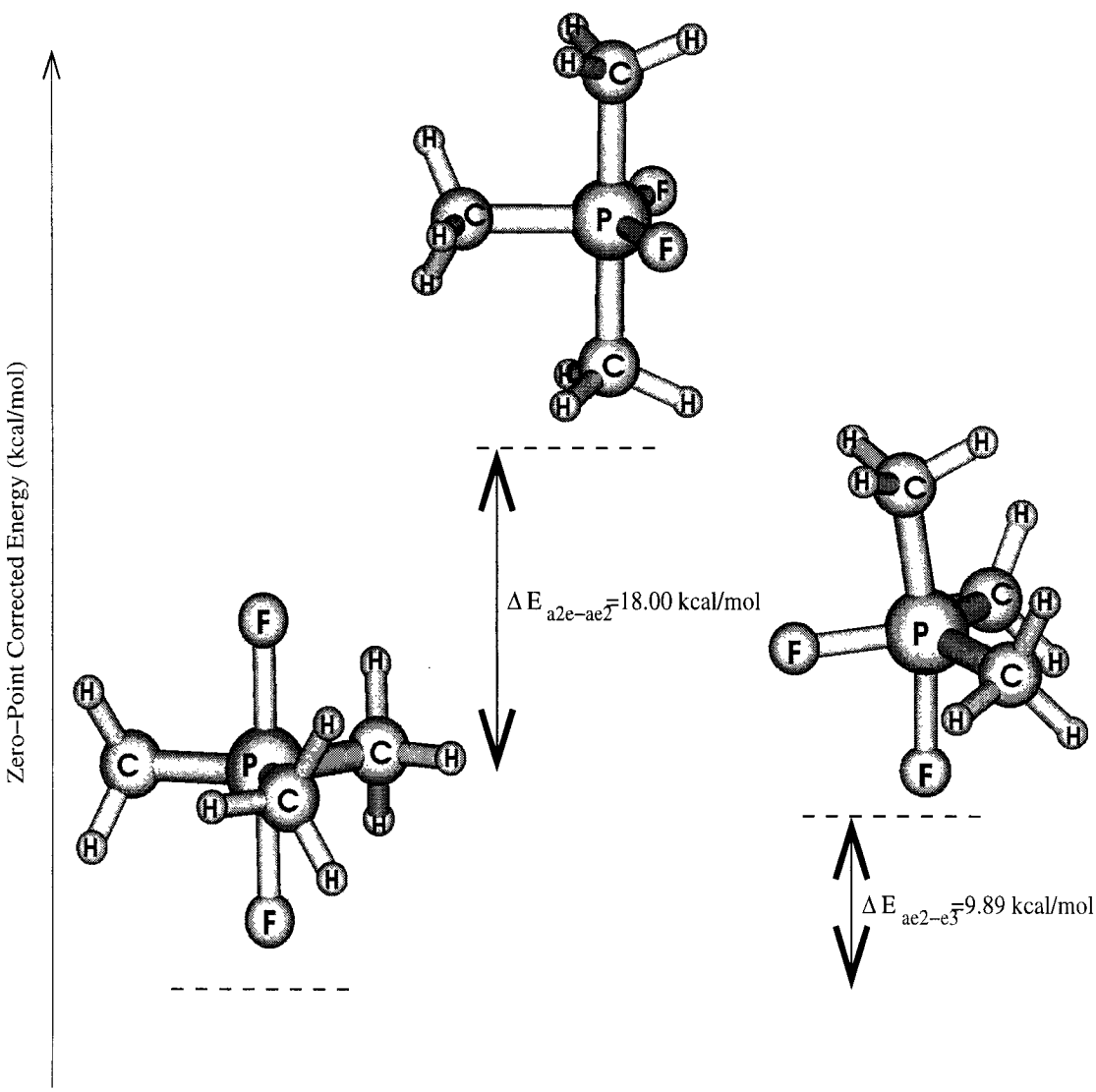


Figure 5.12: Diagram showing the B3LYP energy difference between apical and equatorial positioning of the CH_3 groups on $PF_2(CH_3)_3$.

respectively. Experimental geometries for these molecules were obtained from x-ray diffraction studies [136]. Not all reported angles were included in the experimental results. For those angles with no experimental results, only the theoretical results are reported.

Table 5.12: Selected Structural Parameters for $PF_2(CH_3)_3$.

coordinate	experiment	MP2	B3LYP	VSXC
$P - C_{eq}$ (Å)	1.813	1.811	1.826	1.813
$P - F_{ax}$ (Å)	1.685	1.725	1.739	1.741
$C_{eq} - P - C_{eq}$ (deg.)	-	119.8	120.0	119.9
$C_{eq} - P - F_{ax}$ (deg.)	-	88.7	90.0	89.9
$F_{ax} - P - F_{ax}$ (deg.)	-	180.0	180.0	179.2

The experimental results reported both $PF_2(CF_3)_3$ and $PF_2(CH_3)_3$ as having D_{3h} symmetry, with the CF_3 and CH_3 groups all placed equatorially. Since the only angle reported was the F-C-F angle (no H-C-H angle reported), the distortion from the ideal trigonal bipyramidal geometry could not be determined from the experimental results.

The MP2 method most accurately describes the geometries of $PF_2(CF_3)_3$ and $PF_2(CH_3)_3$, as shown in Tables 5.11 and 5.12, respectively. The B3LYP and VSXC methods also accurately describe these systems. Comparison of the two DFT functionals indicate comparable results. For the angles where no experimental results were reported, all three methods give similar values. With the theoretical results, a small distortion in the angles

from the ideal trigonal biypramidal geometry is seen.

Table 5.13: Selected Frequencies for $PF_2(CH_3)_3$.

frequency (cm^{-1})	experiment	MP2	B3LYP	VSXC
PC_3 symm. stretch	647	657	618	643
PF_2 symm. stretch	501	501	480	477
PF_2 antisymm. stretch	670	639	607	594
PC_3 antisymm. stretch	775	805	769	784
PF_2 axial bend	422	409	391	381
PC_3 in-plane bend	190	182	174	151
PC_3 symm. out-of-plane bend	367	349	344	324
rocking	392	371	367	365

Table 5.13 gives the fundamental frequencies observed experimentally for $PF_2(CH_3)_3$. As in the previous analyses, only those frequencies that include phosphorus are included. These include the twelve fundamental modes shown in Figures 5.1 and 5.2. As indicated above, the molecule has D_{3h} symmetry, so eight fundamental frequencies are expected, four of which are degenerate modes. In the table, eight frequencies are reported from the experimental spectrum. Also included in the table are the theoretically calculated frequencies for the experimentally assigned spectrum.

Figure 5.13 is a plot of the experimental frequencies for $PF_2(CH_3)_3$ versus those calculated using the three theoretical methods. All three methods underestimate the frequencies for this molecule. The best approximation to the experimental spectrum is obtained with the MP2 method. VSXC and B3LYP are comparable to one another in accuracy and generally underestimate the frequencies more than does MP2.

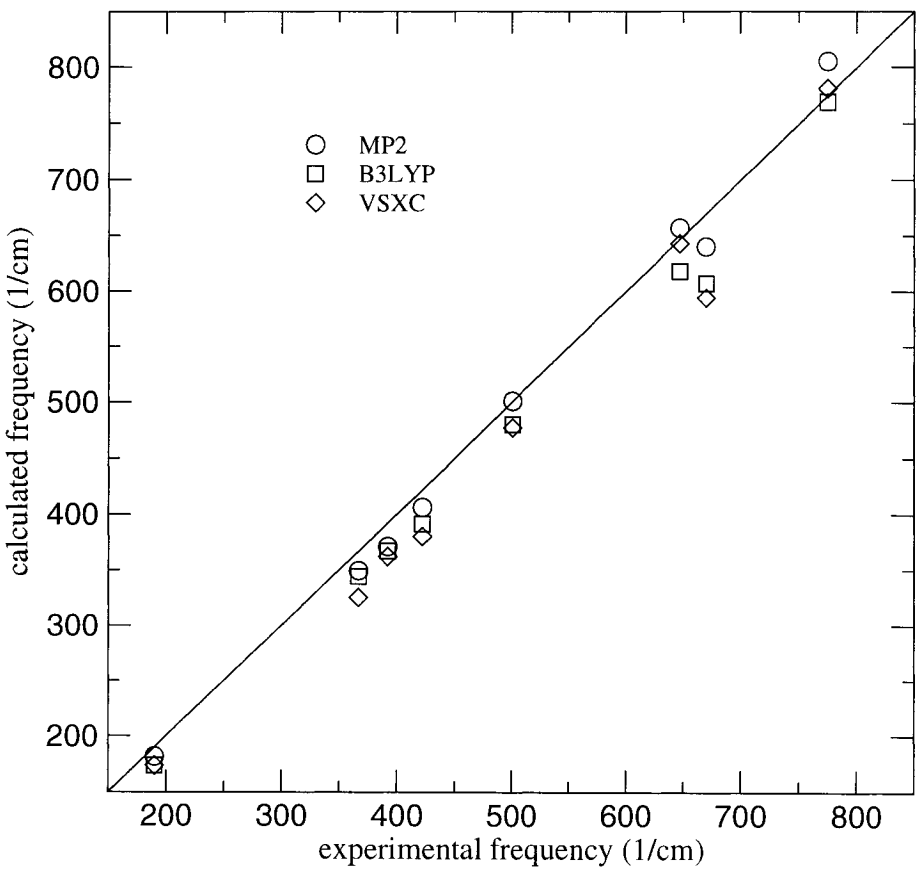


Figure 5.13: Selected harmonic frequencies for $PF_2(CH_3)_3$.

5.4.5 Overall Performance of VSXC

Overall, the geometries and vibrational frequencies are reproduced by the VSXC method. It is comparable to B3LYP in accuracy, improving on it in some cases. Based on the analyses of the given systems, it is expected that VSXC is a feasible method for studying pentacoordinate phosphorus compounds. For geometries and vibrational frequencies, it gives results comparable to MP2 in quality.

One final note is worth mentioning. MP2 still gives the most accurate results for the halophosphorane geometries and vibrational frequencies. However, it is much more computationally expensive than the two DFT methods. For example, the optimization and vibrational analysis of PCl_4CF_3 with MP2 took roughly three days and six hours on a Sun Fire cluster machine. B3LYP calculations of the same system on the same machine took roughly 2 hours and 55 minutes, while VSXC took 3 hours and 15 minutes. Since the difference in accuracy between MP2 and the two DFT functionals is small for the halophosphoranes, DFT is an alternative for studying these systems.

5.5 ^{31}P NMR of Fluorophosphoranes and Oxyphosphoranes

5.5.1 ^{31}P NMR Shieldings: Theoretical Considerations

Investigations into the NMR chemical shifts of the phosphorus nucleus have important applications in investigations of structures and reactions in inorganic and biological

systems. In general, application of a magnetic field to a sample induces motion in the charge density at a nucleus. This motion gives rise to internal currents. These currents, depending on their origin, can oppose or augment the external field. The net result is a local field experienced by the nucleus. Since the magnitude of this local field depends on the induced currents, probing it can give information about the electronic environment about the nucleus in question.

There are two main contributions to the shielding of the nucleus. A small local field is produced due to rotational motion of the charge density about the nucleus. This field is opposed to the applied field, causing increased shielding of the nucleus. This is the diamagnetic part of the shielding. It is not significantly affected by the nuclear environment. Interactions between electronic currents within various orbitals give rise to the paramagnetic contribution to the shielding. These currents are perturbed by the applied field. The interaction between them depends on the energy difference between the orbitals in question. Often interactions occur between higher-lying occupied and lower-lying unoccupied molecular orbitals. The paramagnetic contribution to the shielding is very sensitive to the nuclear environment within the molecule [137].

Several studies have been done on the ^{31}P shielding in tricoordinate and tetracoordinate phosphorus compounds [137–140]. The results indicate that the paramagnetic part contributes significantly to the overall shielding at phosphorus. Since a more spherical charge distribution leads to a smaller paramagnetic contribution, this indicates the charge distribution at phosphorus is not spherical. It also indicates a significant contribution due to mixing of lower-lying excited states into the ground-state wavefunction as this is the

origin of the paramagnetic contribution. Results of these studies also indicate it is for this reason that inclusion of correlation beyond what is described in the Hartree-Fock approach is needed to properly reproduce ^{31}P shieldings from theoretical calculations. This is especially true for systems which have multiple bonds or lone pairs as these lower the HOMO-LUMO gap and allow more mixing of excited states into the ground state.

The phosphorus shieldings are sensitive to the electronic environment of the nucleus. Changes in this electronic environment can occur from geometrical changes, addition of substituents, either directly to phosphorus or to species attached to phosphorus, and changes in the charges on various atoms or groups in the molecule. Chesnut and coworkers have done investigations into the change in the bond lengths or angles involving the phosphorus atom and the effects on the shieldings [137,139]. They have indicated that the shieldings are quite sensitive to changes in the bond lengths and angles at phosphorus, with the sensitivity being largest for changes in bond lengths. In general, an increase in the bond lengths and angles has a deshielding effect. This results from a larger paramagnetic contribution to the shielding which results from the elongation of bonds and widening of angles.

From the above information, it is clear that correlation effects are important in describing the phosphorus chemical shifts. These correlation effects can be included either through the use of a post-Hartree-Fock method or DFT. The cheapest post-HF method is MP2. However, even for moderately sized systems, this method is quite expensive. DFT provides a cheap way of including correlation effects and is more attractive for application of ^{31}P NMR shielding calculations to biological or larger inorganic systems. In the previous

section of this chapter, both B3LYP and VSXC were shown to give accurate geometries and frequencies for a set of substituted halophosphoranes. It is expected that the NMR shielding calculations should also give reasonable results when calculated with B3LYP and VSXC.

Previous investigations into the use of DFT for calculating ^{31}P NMR shieldings has focused on tricoordinate and tetracoordinate systems [141, 142]. However, these provide some insight into the quality of DFT shieldings for pentacoordinate phosphorus. They noted that in general, the LDA and GGA functionals all give ^{31}P shieldings that are too deshielded compared to experiment. Also, this effect seems to be largest for phosphorus systems that have large paramagnetic contributions, such as the phosphorus halides. Also, these shieldings are smaller than those predicted with either HF or MP2. It is interesting to note that DFT methods are in general successful at reproducing chemical shifts in systems with small paramagnetic contributions. For systems with larger paramagnetic contributions, they do not perform as well. This may be a result of their inability to properly describe the paramagnetic part of the shielding. To quantify this, shieldings of various nuclei with a range of sizes of paramagnetic contributions would need to be calculated with DFT at the experimental geometries. Also, in general, methods which do not produce a good approximation to the ground-state wavefunction (or density) are expected to be poor at reproducing the paramagnetic contribution to the shielding. This may result from deviations from the true wavefunction being amplified in the perturbed wavefunction, which is needed only in the paramagnetic part of the shielding.

Investigation into the shieldings of some halophosphoranes and oxyphosphoranes was

done in two parts. First, a protocol was needed to study some model systems of enzyme intermediates and transition states. This investigation was done for a set of substituted halophosphoranes. Next, the VSXC functional was further tested against experiment for a series of oxyphosphoranes whose environment at phosphorus closely resembled the model systems investigated in the last part of this study.

5.5.2 ^{31}P NMR Shieldings of Substituted Halophosphoranes

Protocol

The VSXC functional was again compared to MP2 and B3LYP results for the halophosphoranes to determine the accuracy of the calculated NMR shieldings. Since only the CSGT method was implemented with VSXC, VSXC-CSGT results were compared with B3LYP-CSGT methods. Due to the above concerns with using DFT to calculate NMR shieldings of phosphorus, comparisons needed to be made with MP2. However, since CSGT is not implemented with MP2, MP2-GIAO was compared to B3LYP-GIAO to first validate the use of B3LYP for calculating ^{31}P NMR shieldings. From this result, VSXC results could be indirectly compared to MP2 results.

For MP2, B3LYP, and VSXC, the geometries of the halophosphoranes were obtained using the 6-311+G* basis set (6-311++G** for systems with hydrogens). Although NMR shielding quality converges slowly with basis set, the purpose of the investigation was to eventually apply the same protocol to larger model systems. As a result, a compromise was made in the size of the basis set. Optimizations were performed for each method so the shieldings were calculated on geometries obtained at the same level of theory. Minima

were confirmed from an analysis of the resulting frequencies.

Before a comparison of the theoretical shieldings could be made with experiment, experimental chemical shifts were converted to absolute shieldings. It is believed that a comparison to absolute shieldings is more accurate. Errors present in the shielding values for the reference system and the sample are not masked by a relative comparison. This comparison may cancel them out partially or entirely by subtraction of the two shielding values to obtain chemical shifts. The experimental chemical shifts were converted to shielding values using

$$\sigma_{sample} = \delta + \sigma_{ref}, \tag{5.1}$$

where the reference shielding used was that of 85 % phosphoric acid given in Ref. 143 as 328.35 ppm.

Results

^{31}P NMR absolute shieldings were cacluated for PF_5, PF_4CF_3, PF_4CH_3, $PF_3(CF_3)_2$, $PF_3(CH_3)_2$, $PF_2(CF_3)_3$, $PF_2(CH_3)_3$, PCl_5, and PCl_4CF_3. The absolute theoretical and

experimental shieldings are given in Table 5.14.

Table 5.14: ^{31}P NMR Shieldings (in ppm) for Substituted Halophosphoranes.

molecule	experiment	MP2-GIAO	B3LYP-GIAO	B3LYP-CSGT	VSXC-CSGT
PF_5	408.7	400.3	344.6	342.2	385.6
PF_4CF_3	394.8	388.5	333.6	334.3	379.8
PF_4CH_3	358.3	356.7	298.0	300.5	346.7
$PF_3(CF_3)_2$	379.3	381.3	326.8	331.7	377.4
$PF_3(CH_3)_2$	320.4	329.0	267.4	273.5	320.2
$PF_2(CF_3)_3$	388.2	400.2	346.8	354.9	401.6
$PF_2(CH_3)_3$	344.2	365.5	306.3	313.2	356.2
PCl_5	248.4	325.7	221.0	225.9	296.1
PCl_4CF_3	-	342.3	241.7	247.1	306.1

No obvious trend in substitution is apparent from the halophosphorane ^{31}P shieldings. However, as with PF_3 and PCl_3, PCl_5 is more deshielded than PF_5. This suggests a larger paramagnetic contribution for PCl_5 than for PF_5. This trend is reproduced for all four methods. PF_4CH_3 is more deshielded than PF_4CF_3, suggesting a larger paramagnetic contribution for this system. This effect is seen for the $PF_3(CH_3)_2$ - $PF_3(CF_3)_2$ and $PF_2(CH_3)_3$ - $PF_2(CF_3)_3$ pairs as well. The trend is also reproduced for all the four theoretical methods.

Figure 5.14 shows the correlation between experimental and theoretical shieldings for a series of substituted halophosphoranes. Overall, the shieldings are well reproduced for these systems. The largest errors are seen for PCl_5, $PF_2(CH_3)_3$, and $PF_2(CF_3)_3$. MP2-GIAO overestimates most of the shieldings, while it underestimates the rest. Both B3LYP-GIAO and B3LYP-CSGT underestimate all of the shieldings. A few of the shieldings are overestimated with VSXC-CSGT, but the majority are underestimated. The DFT results

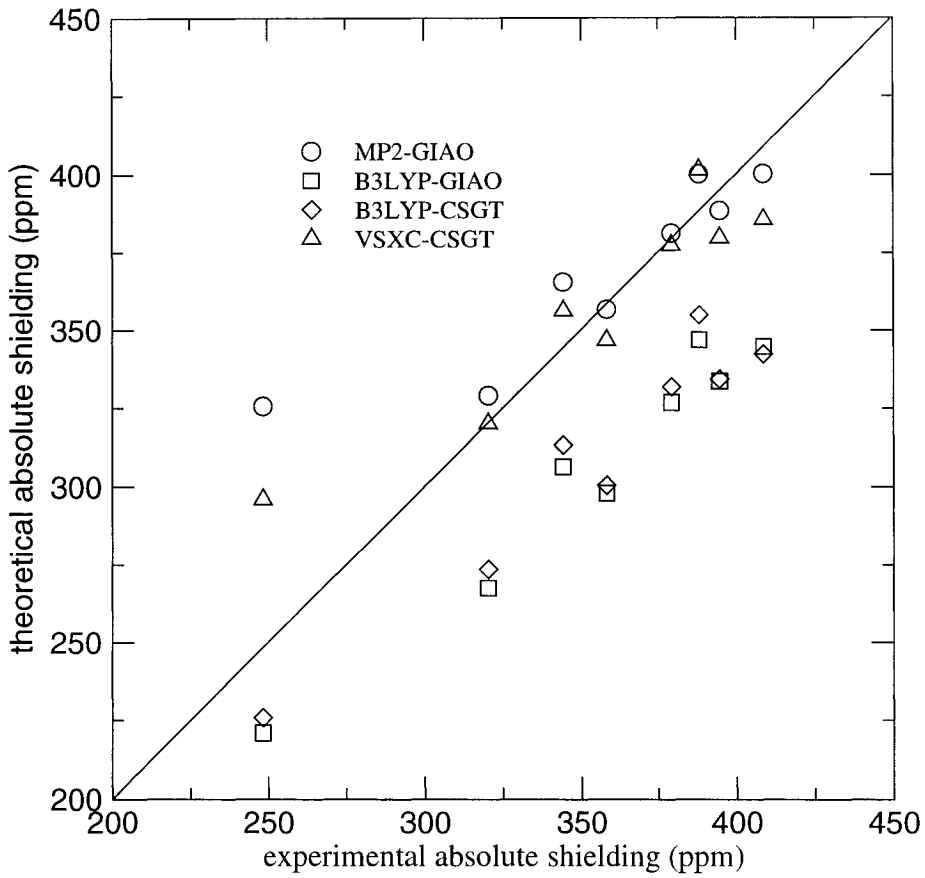


Figure 5.14: ^{31}P NMR absolute shieldings for selected fluorophosphoranes.

are not unexpected as it was previously mentioned that DFT underestimates shieldings for phosphorus.

The MP2-GIAO method gives the most accurate shieldings out of all of the methods. Comparison of B3LYP-GIAO with MP2-GIAO shows B3LYP results are significantly less accurate than MP2. Many of the B3LYP-GIAO shieldings differ from the MP2-GIAO results by around 50 ppm. Comparison of B3LYP-CSGT with B3LYP-GIAO indicates little difference between GIAO and CSGT with B3LYP. This suggests the problem in the shieldings is with the B3LYP functional, not GIAO. Also, if there was a problem with GIAO reproducing the phosphorus shieldings, the MP2 results would show larger deviations from experiment. Comparison of VSXC-CSGT calculated shieldings shows much better agreement with both MP2-GIAO and experiment than does both B3LYP-GIAO and B3LYP-CSGT. This result is not unexpected as in Section 5.4 VSXC gave slightly better geometries than B3LYP. Since the phosphorus shieldings are sensitive to geometry, a better approximation of the experimental geometry gives more accurate theoretical shieldings.

5.5.3 ^{31}P NMR Shieldings of Substituted Oxyphosphoranes

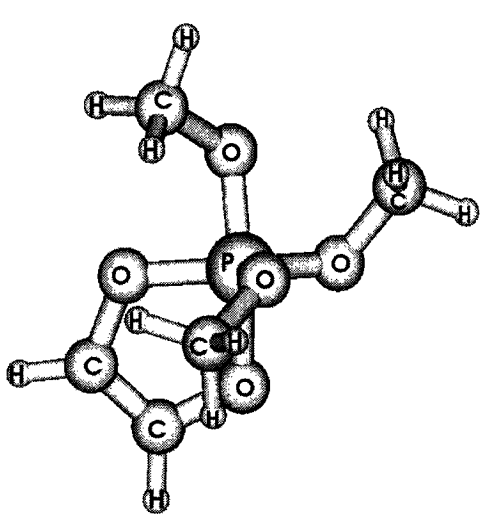
Protocol

The geometries of the oxyphosphoranes were obtained at the VSXC/6-311++G** level of theory. Minima were confirmed from an analysis of the resulting frequencies. Experimental chemical shifts were converted to absolute shieldings in the same manner as was done for the halophosphoranes.

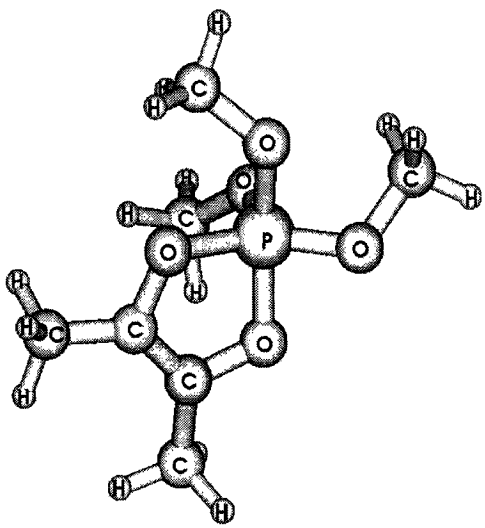
Four model systems were investigated. For each model, a minimum energy structure was obtained for a series of five conformers. These conformers were generated by changing the position of the carbon atom attached to the oxygen side of the P-O bonds. Altering the positions of groups attached to these carbons would have produced a wider variety of conformers. This analysis was not pursued as previous calculations on the effects of bond deformations on phosphorus compounds indicated these effects were small [139]. Also, with the large amount of conformers that would need to be investigated, a thorough investigation would have been prohibitive. So a small sample of conformers was generated which tried to incorporate the largest effect on the shielding. This was done to try to get a better estimate of the shielding for a collection of molecules for each model system. The resulting shieldings for each model system were calculated from an average over the shieldings for the five conformers.

Results

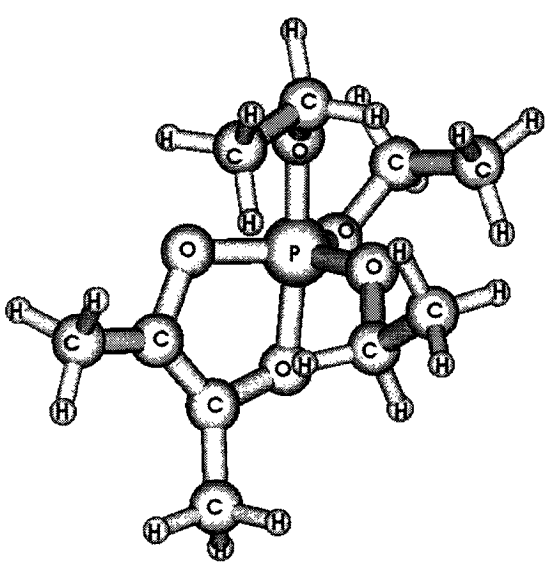
Figure 5.15 shows the lowest energy conformer for each model system. Table 5.15 shows the average shieldings for the model systems.



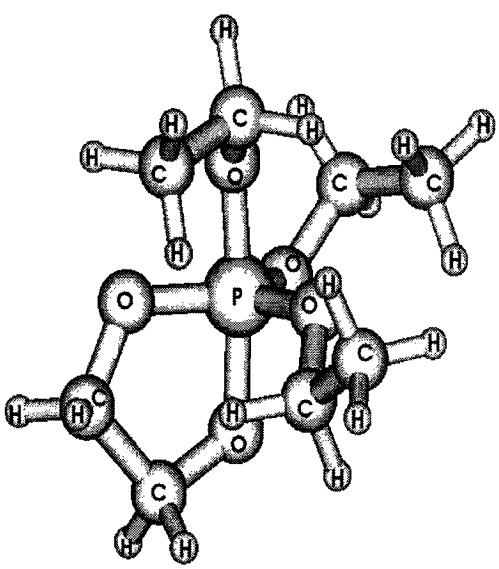
model 1



model 2



model 3



model 4

Figure 5.15: Diagram showing the minimum energy conformers for each of the model oxyphosphoranes.

Table 5.15: ^{31}P NMR Shieldings (in ppm) for Model Oxyphosphoranes.

model system	experiment [a]	VSXC-CSGT
model 1	372.6	353.7
model 2	377.9	364.1
model 3	380.5	363.0
model 4	380.4	366.3

[a]Experimental values from Ref. 1.

The range in the above shieldings is small at roughly 8 ppm in the experimental values. In the theoretical values, the range is roughly 13 ppm. This indicates that addition of methyl groups to the side chains or ring attached to phosphorus has little effect on the shielding values. Since phosphoryl transfer enzymes catalyze phosphate transfer from phospholipids, models which contain small side chains should be sufficient to model the entire phospholipid. Between model 1 and model 2, the difference in the shieldings is largest. This corresponds to addition of methyl groups to the ring. Between model 2 and model 3, the shieldings change very little. This indicates that addition of methyl groups directly to the ring has a larger effect than adding them to the methoxy side chains. Between model 1 and model 2, the theoretical method reproduces this trend, but this is not the case for model 2 and model 3. Since the difference is small in the experimental shielding values, it is likely that it is too small to distinguish theoretically.

Figure 5.16 shows the correlation of the experimental and theoretical ^{31}P NMR shieldings for models 1 to 4. From the diagram, it is apparent that VSXC is able to reproduce experimental shielding values for these four model systems. Since they closely represent

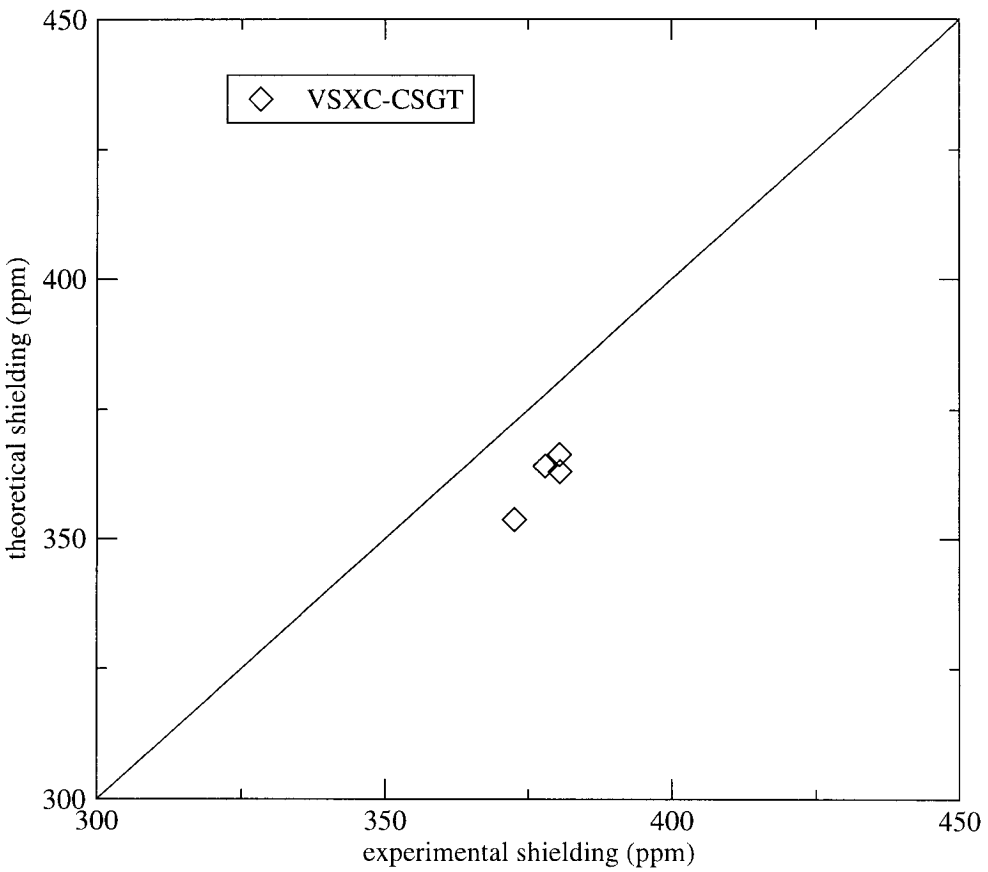


Figure 5.16: ^{31}P NMR absolute shieldings for model oxyphosphoranes.

the part of the enzyme transition state or intermediate that undergoes the largest changes in electronic structure, it is reasonable to assume that ^{31}P NMR shieldings for the enzyme intermediates or transition states can accurately be reproduced using VSXC-CSGT.

A further note should be made concerning acceptable error in ^{31}P NMR shielding calculations. It is generally accepted [137, 140] that as much as 20 ppm of the difference between theory and experiment is considered to be *noise* in the theoretical calculation. This noise may result from a couple of factors. Basis set convergence is often slow for NMR shielding calculations, so the basis set may be too small to obtain accurate shielding values. Also, the accuracy of the paramagnetic portion of the shielding may not be good enough to accurately describe the shielding of nuclei which have a significant contribution to the paramagnetic portion of the shielding. As a result, any shieldings within this value are considered accurate theoretical representations of the experimental shielding. In the case where the NMR shieldings are used to detect the presence of a pentacoordinate phosphorus intermediate in phosphoryl transfer reactions, the reactant-intermediate difference in shieldings is usually around 50 ppm. As long as the accuracy of the theoretical calculations can come within this noise value, they can be used to qualitatively show the presence of an intermediate in phosphoryl transfer reactions.

5.6 Structures and ^{31}P NMR Shieldings of Enzyme Intermediate Models

Phosphatidylinositol-specific phospholipase C catalyses the hydrolysis of phosphatidylinositol to produce inositol-1-phosphate and diacylglycerol. It proceeds via a transition state that involves a pentacoordinate phosphorus species [92]. In order to study the reaction pathway properly, some information on the ability to study pentacoordinate phosphorus transition states by theoretical methods is necessary. However, to do this, one must compare to experiment. This is not possible for transition states. Some insight into the utility of theoretical methods for studying stable pentacoordinate phosphorus species was gained from the previous sections of this chapter. However, it would be useful to study a pentacoordinate analog of the transition state for this enzyme. A short investigation is thus presented for an analog of the transition state of this enzyme. Since it was determined in Section 5.5.3 that the NMR shieldings are not affected by addition of methyl groups to the side chain methoxy groups, the diacylglycerol portion of the molecule is replaced with a methyl group. This smaller analog was also chosen as the DFT calculation for the entire molecule was computationally too demanding to study. It is expected that the shielding values for this system will be close to the values of the oxyphosphoranes studied. A diagram of a phosphatidylinositol without the long fatty acid side chains is presented in Figure 5.17.

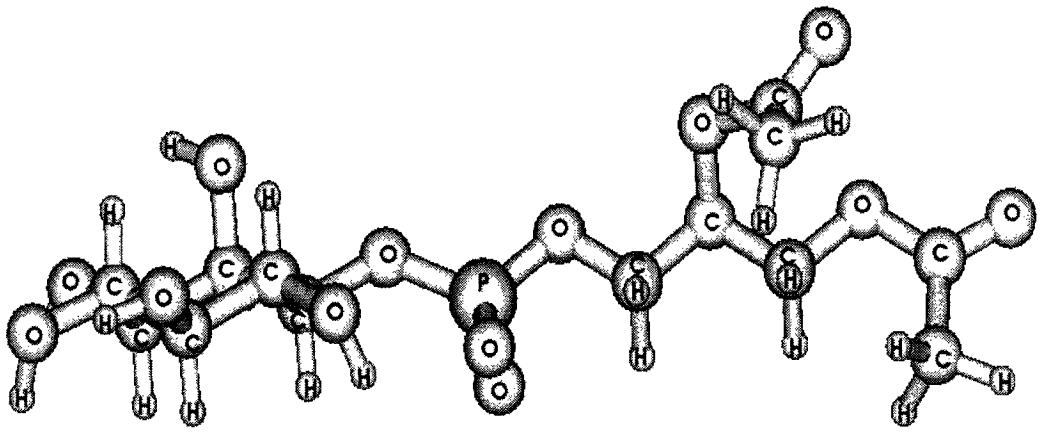


Figure 5.17: Diagram showing a PI-PLC substrate with the fatty acid side chains shortened.

5.6.1 Protocol

Two models were considered for this study. They differ in the position of the ring in the trigonal bipyramid. The first model has both P-O bonds in the five-membered ring diequatorial and the second model has them apical-equatorial. From previous work on cyclic oxyphosphoranes, it is expected that the latter model will be the more stable one. Results will be presented for the lower energy model.

For each model, four conformers were generated. Each conformer differed in the placement of the O^-, OH, and methoxy groups relative to the ring system. A minimum was sought for each conformer using the VSXC/6-311++G** level of theory. Frequency analyses were done to confirm the resulting geometries were minima. ^{31}P NMR shieldings were generated with the VSXC-CSGT method.

5.6.2 Results

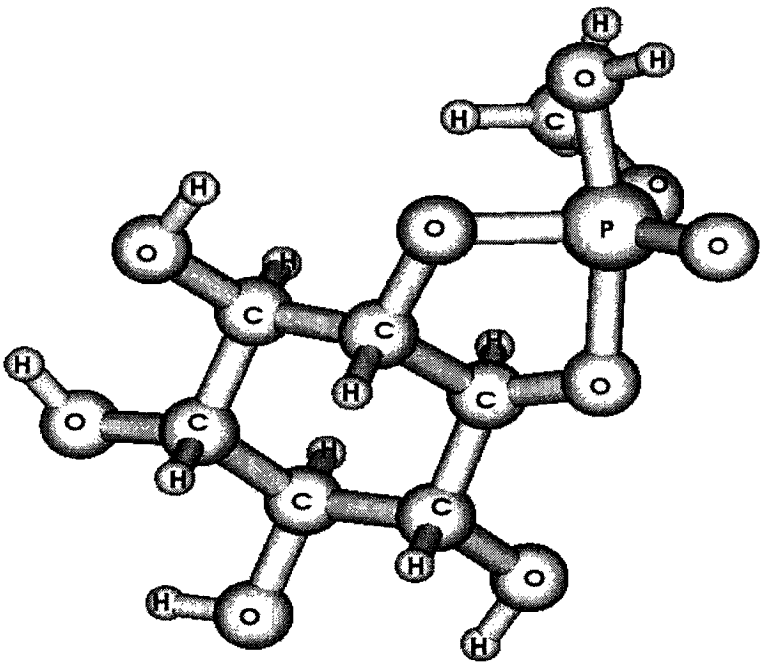
Since no experimental results were found for the two models, only the results of the theoretical analysis are presented. A minimum search for each of the eight conformers produced the two minimum geometries in Figure 5.18.

Table 5.16 shows the energies and ^{31}P NMR shielding values for the two models studied. The models differ in energy by 11.03 kcal/mol in favor of the apical-equatorial placement of the ring P-O bonds. This is consistent with the expected ring placement of five-membered rings in oxyphosphoranes. Ring strain is larger for five-membered rings in a diequatorial placement compared to apical-equatorial placement. Also, a significant change in the deformation from the ideal trigonal bipyramid is seen with both models. However, this distortion is larger for the diequatorial placement of the five-membered ring.

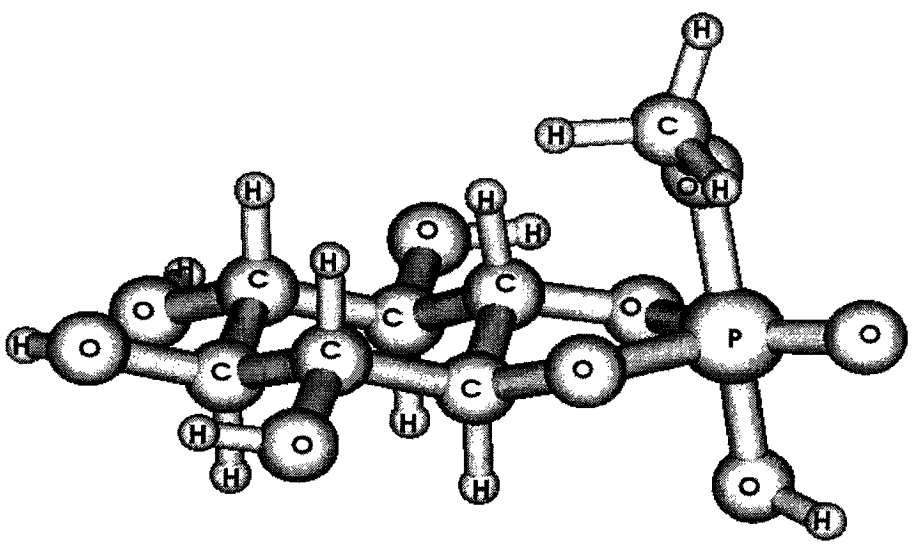
Table 5.16: Energies and ^{31}P NMR shielding values of the lowest energy conformers representing apical-equatorial and equatorial-equatorial ring span in the enzyme models.

enzyme model	E_{zpc} (H)	^{31}P Shielding (ppm)
apical-equatorial ring span	-1294.189415	369.8
equatorial-equatorial ring span	-1294.171832	339.8

There are a number of effects seen in going from the diequatorial to apical-equatorial placement of the ring. Some of these include ring strain, bond and angle deformations, and steric crowding. All of these influence the shielding value at phosphorus. From the ^{31}P shielding values for the two models, the combined effect results in a significant deshielding effect when the ring is placed diequatorially as compared to an apical-equatorial placement.



apical–equatorial ring span



equatorial–equatorial ring span

Figure 5.18: Diagram showing the minimum energy conformers for the models of the enzyme transition state in the phosphoryl transfer reaction catalyzed by PI-PLC.

5.7 Conclusion

VSXC has proven to give accurate structures, vibrational frequencies, and ^{31}P NMR shielding values for a select number of pentacoordinate halophosphoranes. The results are very similar to those of B3LYP and are comparable to MP2. Phosphorus shielding results indicate VSXC yields more accurate shielding values. Although they are underestimated overall by VSXC, they are grossly underestimated by B3LYP. Thus, for studies of pentacoordinate transition states or intermediates using DFT, VSXC is a viable option.

Investigation of an analog for the transition state pentacoordinate oxyphosphorane in the phosphoryl transfer reaction catalyzed by PI-PLC indicate the apical-equatorial isomer is more stable. Also, significant distortion from trigonal bipyramid is seen. These results are common to oxyphosphoranes that contain five-membered rings.

Claims to Original Research

1. Coding of the CFMM and integration into DeFT.

2. Testing of DeFT with the CFMM using polyglycine chains of differing length.

3. Analysis of structures for halophosphoranes for the MP2, B3LYP, and VSXC methods and comparison to experiment.

4. Characterization of halophosphorane frequencies for the MP2, B3LYP, and VSXC methods and comparison to experiment.

5. Comparison of ^{31}P NMR shieldings for the MP2-GIAO, B3LYP-GIAO, B3LYP-CSGT, and VSXC-CSGT methods as well as experimental values for halophosphoranes.

6. Comparison of ^{31}P NMR shieldings for four model oxyphosphoranes for which experimental results were available.

7. Analysis of structures and ^{31}P NMR shieldings for two models of phosphatidylinositol, the substrate for PI-PLC.

Publications

1. Tian W.Q., Shaw, M., St-Amant, A. Quantum Mechanical Investigations on Some Nanoparticles: Icosahedral (I_h) Fullerenes C_{180}, C_{240}, C_{560}, and C_{960}., manuscript in preparation.

2. Shaw, M., St-Amant, A. Investigation of Pentacoordinate Phosphorus Models with Density Functional Methods., manuscript in preparation.

3. Shaw, M., St-Amant, A. Linear Scaling for Density Functional Calculations on Large Molecules With the DeFT Software Package., *J. Theor. Comp. Chem.* 2004, **3**, 419

4. Shaw, M., Arteca, G. Conformationally-Averaged Molecular Surface Areas for Chain Molecules., *Theochem* 2001, **537**, 181

Appendix A

CFMM Code and Input Examples

This appendix gives examples of skeleton code for the CFMM. The CFMM is executed during the computation of the KS matrix elements. This is done from the subroutine which generates the two-electron integrals as follows:

```
      if(natoms.le.1000) then
        if(iter.eq.1) write(6,*)'twoemm> System is too small -> ',
     &            'evaluating integrals directly.'
        timef = zero
      else
        ngp = iwkgp(1)+iwkgp(2)+iwkgp(3)+iwkgp(4)+iwkgp(5)+iwkgp(6)
        ncd = ncds+ncdspd
        icord=ipb+3*ngp

c Parallel stuff...
```

```
ka = 1

kb = igma

kc = icoef

kd = iab2

ke = ipa

kf = ipb

kg = icord


ja=iwkgp(1)+7

jb=3*iwkgp(2)+ja

jc=10*iwkgp(3)+jb

jd=6*iwkgp(4)+jc

je=18*iwkgp(5)+jd

jf=je+(37*iwkgp(6))

jg=jf+5*(ngp+ncd)

jh=jg+(ngp+ncd)


if(jh.gt.5000000) stop 'twoemm> out of bounds in iwkgp '


ti_ff = etime(tarray)

ti_ff = tarray(1)+tarray(2)
```

```
      call setcfmm (ncd, ngp, nsubdiv, xdist, xmincoor, ymincoor,
     &                zmincoor, nlgws, nboxes, noccbox, coord, iboxes,
     &                icdcfunc, icdlfunc, alphacd, wkgp(ka), wkgp(kb),
     &                wkgp(kg), iwkgp(jf), wkgp(kd), iwkvec(1))

      maxws = 2**(nsubdiv-1)
      lldim = (nlevels+1)*(nlevels+2)/2
      ia = 1
      ib = 2*noccbox*lldim+1
      ic = nsubdiv*maxws
      id = nsubdiv*3+1

      nss=iwkgp(1)
      nps=iwkgp(2)
      npp=iwkgp(3)
      nds=iwkgp(4)
      ndp=iwkgp(5)
      ndd=iwkgp(6)

      call main_cfmm (nlevels, natoms, ncds, ncdspd, ncd, nlgws,
     &                 nconts, ncontp, ncontd, ngp, noccbox,
     &                 nsubdiv, nboxes, maxws, nss, nps, npp,
```

```
&                    nds, ndp, ndd, xmincoor, ymincoor, zmincoor,
&                    lldim, ic, id, xdist, iboxes, icdcfunc,
&                    icdlfunc, alphacd, coeff, coefscd, coefpcd,
&                    coefdcd, alpha, cdfitc, wkvec(ia), wkvec(ib),
&                    iwkvec(1), iwkvec(id), wkgp(kb), wkgp(kc),
&                    iwkgp(7), iwkgp(ja), iwkgp(jb), iwkgp(jc),
&                    iwkgp(jd), iwkgp(je), wkgp(kd), wkgp(ke),
&                    wkgp(kf), igpcont, wkgp(kg), iwkgp(jf),
&                    iwkgp(jg), fock)
```

In the above code fragment, two subroutines for the CFMM are called. The first subroutine call sets up arrays and boxing information for use in the main part of the CFMM. Also generated is the well-separated information and box labels for the centers. The second subroutine call executes the CFMM. The code for this subroutine is

```
      subroutine main_cfmm (ll, ncenters, nauxs, nauxspd, naux, nlgws,
&                           nconts, ncontp, ncontd, ngp, noccbox,
&                           nsubdiv, nboxes, maxws, ncontss,
&                           ncontps, ncontpp, ncontds, ncontdp, ncontdd,
&                           xmincoor, ymincoor, zmincoor, lldim, icdim,
&                           iddim, xdist, iboxes, icaux, ilaux,
&                           alphaaux, coeff, coefscd, coefpcd, coefdcd,
&                           alpha, cdfitc, c_boxexp, s_boxexp, imaxdim,
&                           icage, gamma, coeffgp, iptrss, iptrps,
```

```
&                       iptrpp, iptrds, iptrdp, iptrdd, ab2, pa,
&                       pb, igpcont, coord, npartbox, ni_box, fuv)

 implicit real*8(a-h,o-z)

 logical lcage(nboxes)
 real ti_box, tf_box, ti_p1, tf_p1, ti_p2, tf_p2, ti_p3, tf_p3,
&     ti_rest, tf_rest, ti_pl, tf_pl, ti_p4, tf_p4, t_cfmm,
&     tarray(2)

 dimension icaux(*),ilaux(*),alphaaux(*),fuv(*), icage(icdim, *)
 dimension alpha(*), coeff(*), coefscd(*), coefpcd(*),
&          coefdcd(*), cdfitc(*)
 dimension c_boxexp(noccbox, *), s_boxexp(noccbox, *),gamma(*),
&          coeffgp(*), ab2(*), pa(3,*), pb(3,*)
 dimension iptrss(*), iptrps(3,*), iptrpp(10,*), iptrds(6,*),
&          iptrdp(18,*), iptrdd(37,*), igpcont(*), coord(3, *)
 dimension itobox(10000000), iboxes(*), npartbox(5, *), ni_box(*),
&          imaxdim(3,*)

 call cfmm_box(nboxes, nsubdiv, maxws, nlgws, icdim, noccbox,
&              iboxes, npartbox, itobox, lcage, icage, idiv,
```

```
&              noccbox, ni_box, imaxdim)

 call cfmm_pass1 (ll, nsubdiv, nboxes, nauxs, nauxspd, naux, maxws,
&                 lldim, icdim, noccbox, xdist, xmincoor,
&                 ymincoor, zmincoor, lcage, coord, itobox,
&                 iboxes, icaux, ilaux, icage, ni_box, imaxdim,
&                 alphaaux, coefscd, coefpcd, coefdcd, cdfitc,
&                 c_boxexp, s_boxexp)

 call cfmm_pass2 (ll, nsubdiv, nboxes, lldim, icdim,
&                 noccbox, xdist, xmincoor, ymincoor, zmincoor,
&                 lcage, iboxes, icage, imaxdim, c_boxexp,
&                 s_boxexp)

 call cfmm_pass3 (nsubdiv, nboxes, ll, noccbox, lldim, icdim,
&                 xdist, xmincoor, ymincoor, zmincoor, lcage,
&                 iboxes, icage, imaxdim, c_boxexp, s_boxexp)

 call cfmm_pass4 (nsubdiv, nboxes, ll, maxws, lldim,
&                 icdim, noccbox, naux, nauxs, ncontss,
&                 ncontps, ncontds, ncontpp, ncontdp, ncontdd,
&                 ngp, nconts, ncontp, ncontd, xdist, xmincoor,
```

```
&                ymincoor, zmincoor, lcage, iboxes, gamma,
&                itobox, icage, ni_box, imaxdim, coord, coeffgp,
&                c_boxexp, s_boxexp, iptrss, iptrps, iptrds,
&                iptrpp, iptrdp, iptrdd, igpcont, pa, pb, ab2,
&                fuv)

 RETURN
 END
```

The main subroutine is responsible for executing the various steps of the CFMM. Pass 1 generates multipoles and translates information up the tree. Pass 2 converts multipole expansions in one box into Taylor expansions in another box well-separated from it. Pass 3 translates this information down the tree. Pass 4 interacts Taylor expansions in each box with the densities in the box and adds the contributions to the KS matrix. After execution of the main part of the CFMM, all the integral contributions to the KS matrix for the repulsion between two charge densities has been generated.

References

[1] Holmes, R. R. *Pentacoordinated Phosphorus. Volume 1.;* American Chemical Society: Washington, 1980.

[2] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory;* Macmillan Publishing Co.: New York, 1982.

[3] Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules;* Oxford University Press: New York, 1989.

[4] Thomas, L. H. *Proc. Camb. Phil. Soc.* **1927,** *23,* 542.

[5] Fermi, E. *Rend. Accad., Lincei* **1927,** *6,* 602.

[6] Fermi, E. *Z. Phys.* **1928,** *48,* 73.

[7] Fermi, E. *Rend. Accad., Lincei* **1928,** *7,* 342.

[8] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964,** *136,* B864.

[9] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965,** *140,* A1133.

[10] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980,** *58,* 1200.

[11] Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980,** *45,* 566.

[12] Becke, A. *J. Chem. Phys.* **1992,** *96,* 2155.

[13] Becke, A. D. *J. Chem. Phys.* **1988,** *88,* 1053.

[14] Becke, A. D. *Phys. Rev. A* **1988,** *38,* 3098.

[15] Perdew, J. P. *Phys. Rev. B* **1986,** *33,* 8822.

[16] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37,* 785.

[17] Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989,** *157,* 200.

[18] Becke, A. D. *J. Chem. Phys.* **1998,** *109,* 2092.

[19] Becke, A. D. *Int. J. Quantum Chem.* **1983,** *23,* 1915.

[20] Voorhis, T. V.; Scuseria, G. E. *J. Chem. Phys.* **1998,** *109,* 400.

[21] Perdew, J. P.; Kurth, S.; Zupan, A.; Blaha, P. *Phys. Rev. Lett.* **1999,** *82,* 2544.

[22] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003,** *91,* 146401.

[23] Perdew, J. P.; Tao, J.; Staroverov, V. N.; Scuseria, G. E. *J. Chem. Phys.* **2004,** *120,* 6898.

[24] Jaramillo, J.; Scuseria, G. E. *Chem. Phys. Lett.* **1999,** *312,* 269.

[25] Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 1372.

[26] Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 5648.

[27] Obara, S.; Saika, A. *J. Chem. Phys.* **1986,** *84,* 3963.

[28] Head-Gordon, M.; Pople, J. A. *J. Chem. Phys.* **1988,** *89,* 5777.

[29] Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1989,** *23,* 269.

[30] Hamilton, T. P.; III, H. F. S. *Chem. Phys.* **1991,** *150,* 163.

[31] Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993,** *213,* 514.

[32] Köster, A. M. *J. Chem. Phys.* **1996,** *104,* 4114.

[33] Adamson, R. D.; Dombroski, J. P.; Gill, P. M. W. *J. Comp. Chem.* **1999,** *20,* 921.

[34] Gill, P. M. W.; Gilbert, A. T. B.; Adams, T. R. *J. Comp. Chem.* **2000,** *21,* 1505.

[35] Shao, Y.; Head-Gordon, M. *Chem. Phys. Lett.* **2000,** *323,* 425.

[36] Rico, J. F.; Lòpez, R.; Ramìrez, G. *J. Mol. Struct. (Theochem)* **2001,** *537,* 27.

[37] Häser, M.; Ahlrichs, R. *J. Comp. Chem.* **1989,** *10,* 104.

[38] Jackson, J. D. *Classical Electrodynamics;* John Wiley and Sons, Inc.: New York, 1999.

[39] Hehre, W. J.; Radom, L.; v.R. Schleyer, P.; Pople, J. A. *Ab Initio Molecular Orbital Theory;* John Wiley and Sons Ltd: West Sussex, 1986.

[40] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979,** *71,* 3396.

[41] Ahmadi, G. R.; Almlöf, J. *Chem. Phys. Lett.* **1995,** *246,* 364.

[42] White, C. A.; Head-Gordon, M. *J. Chem. Phys.* **1994,** *101,* 6593.

[43] White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M. *Chem. Phys. Lett.* **1994,** *230,* 8.

[44] White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M. *Chem. Phys. Lett.* **1996,** *253,* 268.

[45] Pérez-Jordá, J. M.; Yang, W. *Chem. Phys. Lett.* **1995,** *247,* 484.

[46] Pérez-Jordá, J. M.; Yang, W. *J. Chem. Phys.* **1997,** *107,* 1218.

[47] Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996,** *271,* 51.

[48] Scuseria, G. E. *J. Phys. Chem. A* **1999,** *103,* 4872.

[49] Challacombe, M.; Schwegler, E.; Almlöf, J. *J. Chem. Phys.* **1996,** *104,* 4685.

[50] Dombroski, J. P.; Taylor, S. W.; Gill, P. M. W. *J. Phys. Chem.* **1996,** *100,* 6272.

[51] Greengard, L.; Rokhlin, V. *J. Comp. Phys.* **1987,** *73,* 325.

[52] Greengard, L.; Rokhlin, V. *Acta Num.* **1997,** *6,* 229.

[53] Sedgewick, R. *Algorithms in C;* Addison-Wesley: New York, 1990.

[54] Wilkinson, B.; Allen, M. *Parallel Programming Techniques and Applications Using Networked Workstations and Parallel Computers, Second Edition;* Prentice Hall: Upper Saddle River, 2005.

[55] Gropp, W.; Lusk, E.; Skjellum, A. *Using MPI Portable Programming with the Message-Passing Interface, Second Edition;* MIT Press: Cambridge, 1999.

[56] Greengard, L.; Gropp, W. D. *Computers Math. Applic.* **1990,** *20,* 63.

[57] Hoi, C. H.; Ruedenberg, K.; Gordon, M. S. *J. Comp. Chem.* **2001,** *22,* 1484.

[58] Shaw, D. M.; St-Amant, A. *J. Theor. Comp. Chem.* **2004,** *3,* 419.

[59] Godbout, N.; Salahub, D.; Andzelm, J.; Wimmer, E. *Chinese J. Chem.* **1992,** *70,* 560.

[60] Gallant, R. T.; St-Amant, A. *Chem. Phys. Lett.* **1996,** *256,* 569.

[61] Goh, S. K.; St-Amant, A. *Chem. Phys. Lett.* **1996,** *264,* 9.

[62] Carrier, J.; Greengard, L.; Rokhlin, V. *SIAM J. Sci. Stat. Comput.* **1988,** *9,* 669.

[63] Nabors, K.; Korsmeyer, F. T.; Leighton, F. T.; White, J. *SIAM J. Sci. Stat. Comput.* **1994,** *15,* 714.

[64] Levitt, M. H. *Spin Dynamics, Basics of Nuclear Magnetic Resonance;* John Wiley and Sons: West Sussex, 2001.

[65] Ramsey, N. F. *Phys. Rev.* **1950,** *77,* 567.

[66] Atkins, P. W.; Friedman, R. S. *Molecular Quantum Mechanics, Third Edition;* Oxford University Press: Oxford, 1997.

[67] Singh, J. *Quantum Mechanics, Fundamentals and Applications to Technology;* John Wiley and Sons: New York, 1997.

[68] Chesnut, D. B. *Rev. Comp. Chem.* **1996,** *8,* 245.

[69] Gauss, J.; Stanton, J. F. *Adv. Chem. Phys.* **2002,** *123,* 355.

[70] Gauss, J.; Ruud, K.; Helgaker, T. *J. Chem. Phys.* **1996,** *105,* 2804.

[71] London, F. *J. Phys. Radium* **1937,** *8,* 397.

[72] Ditchfield, R. *Mol. Phys.* **1974,** *27,* 789.

[73] Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Amer. Chem. Soc.* **1990,** *112,* 8251.

[74] Häser, M.; Ahlrichs, R.; Baron, H. P.; Weis, P.; Horn, H. *Theor. Chim. Acta* **1992,** *83,* 455.

[75] Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1993,** *210,* 223.

[76] Gerratt, J.; Mills, I. M. *J. Chem. Phys.* **1968,** *49,* 1719.

[77] Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1979,** *13,* 225.

[78] Lipscomb, W. N. *Adv. Mag. Res.* **1966,** *2,* 137.

[79] Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996,** *104,* 5497.

[80] Hehre, W. J.; Radom, L.; vR. Schleyer, P.; Pople, J. *Ab Initio Molecular Orbital Theory;* Wiley: West Sussex, 1986.

[81] Gauss, J. *J. Chem. Phys.* **1993,** *99,* 3629.

[82] Chesnut, D. B.; Byrd, E. F. C. *Heteroatom Chem.* **1996,** *7,* 307.

[83] Gauss, J. *Chem. Phys. Lett.* **1992,** *191,* 614.

[84] Ruud, K.; Helgaker, T.; Bak, K. L.; Jorgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1993,** *99,* 3847.

[85] Friedrich, K.; Seifert, G.; Groβmann, G. *Z. Phys. D* **1990,** *17,* 45.

[86] Schreckenback, G.; Ziegler, T. *J. Phys. Chem.* **1995,** *99,* 606.

[87] Sundholm, D.; Gauss, J.; Schäfer, A. *J. Chem. Phys.* **1996,** *105,* 11051.

[88] Cromp, B.; Jr., T. C.; Salahub, D. R.; Malkina, O. L.; Malkin, V. G. *J. Chem. Phys.* **1999,** *110,* 7153.

[89] Ruud, K.; Ästrand, P.-O.; Taylor, P. R. *J. Chem. Phys.* **2000,** *112,* 2668.

[90] Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1999,** *110,* 7627.

[91] Frisch, M. J. *et al.* "Gaussian 03, Revision B.04", Gaussian, Inc., Wallingford, CT, 2004.

[92] Hondal, R. J.; Zhao, Z.; Kravchuk, A. V.; Liao, H.; Riddle, S. R.; Yue, X.; Bruzik, K. S.; Tsai, M.-D. *Biochemistry* **1998,** *37,* 4568.

[93] Gottlin, E. B.; Rudolph, A. E.; Zhao, Y.; Matthews, H. R.; Dixon, J. E. *Proc. Natl. Acad. Sci. USA* **1998,** *95,* 9202.

[94] Raines, R. T. *Chem. Rev.* **1998,** *98,* 1045.

[95] Emsley, J.; Hall, D. *The Chemistry of Phosphorus;* Harper and Row Ltd.: London, 1976.

[96] Hudson, R. F.; Blomquist, A. T. *Organic Chemistry, A Series of Monographs* **1965,** *6,* 106.

[97] Kommana, P.; Kumaraswamy, S.; Vittal, J. J.; Swamy, K. C. K. *Inorg. Chem.* **2002,** *41,* 2356.

[98] Holmes, R. R. *J. Amer. Chem. Soc.* **1978,** *100,* 433.

[99] Holmes, R. R. *Pentacoordinated Phosphorus. Volume 2.;* American Chemical Society: Washington, 1980.

[100] Noury, S.; Silvi, B.; Gillespie, R. J. *Inorg. Chem.* **2002,** *41,* 2164.

[101] Li, L.-M.; Xiao, L.; Liang, Z.-X. *Chinese J. Chem.* **1992,** *10,* 97.

[102] Häser, M. *J. Amer. Chem. Soc.* **1996,** *118,* 7311.

[103] Hoffman, R.; Howell, J. M.; Mutterties, E. M. *J. Amer. Chem. Soc.* **1972,** *94,* 3047.

[104] Magnusson, E. *J. Amer. Chem. Soc.* **1990,** *112,* 7940.

[105] Cooper, D. L.; Cunningham, T. P.; Gerratt, J.; Karadakov, P. B.; Raimondi, M. *J. Amer. Chem. Soc.* **1994,** *116,* 4414.

[106] McDowell, R. S.; Jr., A. S. *J. Amer. Chem. Soc.* **1985,** *107,* 5849.

[107] Wasada, H.; Hirao, K. *J. Amer. Chem. Soc.* **1992,** *114,* 16.

[108] Breidung, J.; Thiel, W. *J. Comp. Chem.* **1992,** *13,* 165.

[109] Uchimaru, T.; Tsuzuki, S.; Storer, J. W.; Tanabe, K.; K, T. *J. Org. Chem.* **1994,** *59,* 1835.

[110] Wladkowski, B. D.; Krauss, M.; Stevens, W. J. *J. Phys. Chem.* **1995,** *99,* 4490.

[111] Dieters, J. A.; Holmes, R. R.; Holmes, J. M. *J. Amer. Chem. Soc.* **1988,** *110,* 7672.

[112] Wang, P.; Zhang, Y.; Glaser, R.; Reed, A. E.; vR. Schleyer, P.; Streitweiser, A. *J. Amer. Chem. Soc.* **1991,** *113,* 55.

[113] Thatcher, G. R. J.; Campbell, A. S. *J. Org. Chem.* **1993,** *58,* 2272.

[114] Vollhardt, K. P. C.; Schore, N. E. *Organic Chemistry, Second Edition;* W. H. Freeman and Company: New York, 1994.

[115] Benkovic, S. J.; Schray, K. J. *Transition States in Biochemical Processes* **1978,** *1,* 493.

[116] Thatcher, G. R. J.; Kluger, R. *Adv. Phys. Org. Chem.* **1989,** *25,* 101.

[117] Westheimer, F. H. *Chem. Rev.* **1981,** *81,* 313.

[118] Kirby, A. J.; Varvoglis, A. G. *J. Amer. Chem. Soc.* **1967,** *89,* 415.

[119] Hengge, A. C.; Edens, W. A.; Elsing, H. *J. Amer. Chem. Soc.* **1994,** *116,* 5045.

[120] Loew, L. M.; MacArthur, W. R. *J. Amer. Chem. Soc.* **1977,** *99,* 1019.

[121] Hu, C.-H.; Brinck, T. *J. Phys. Chem. A* **1999,** *103,* 5379.

[122] Bianciotto, M.; Barthelat, J.-C.; Vigroux, A. *J. Amer. Chem. Soc.* **2002,** *124,* 7573.

[123] Aqvist, J.; Kolmodin, K.; Florian, J.; Warshel, A. *Chem. Biol.* **1999,** *6,* R71.

[124] Skoog, M. T.; Jencks, W. P. *J. Amer. Chem. Soc.* **1984,** *106,* 7597.

[125] Dejaegere, A.; Liang, X.; Karplus, M. *J. Chem. Soc. Faraday Trans.* **1994,** *90,* 1763.

[126] Gillespie, P.; Ramirez, F.; Ugi, I.; Marquarding, D. *Angew. Chem. Intl. Ed. Eng.* **1973,** *12,* 91.

[127] Westheimer, F. H. *Acc. Chem. Res.* **1968,** *1,* 70.

[128] Lehninger, A. L.; Nelson, D. L.; Cox, M. M. *Principles of Biochemistry, Second Edition;* Worth Publishers: New York, 1993.

[129] Magnusson, E. *J. Amer. Chem. Soc.* **1993,** *115,* 1051.

[130] van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998,** *109,* 400.

[131] Jaramillo, J.; Scuseria, G. E. *Chem. Phys. Lett.* **1999,** *312,* 269.

[132] Zhao, Y.; Pu, J.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2004,** *6,* 673.

[133] Mercero, J. M.; Barrett, P.; Lam, C. W.; Fowler, J. E.; Ugalde, J. M.; Pedersen, L. G. *J. Comp. Chem.* **2000,** *21,* 43.

[134] Lösking, O.; Willner, H.; Oberhammer, H.; Grobe, J.; Van, D. L. *Inorg. Chem.* **1992,** *31,* 3423.

[135] Boyd, R. J.; Boyd, S. L. *J. Amer. Chem. Soc.* **1992,** *114,* 1652.

[136] Oberhammer, H.; Grobe, J.; Van, D. L. *Inorg. Chem.* **1982,** *21,* 275.

[137] Chesnut, D. B.; Quin, L. D. *Heteroatom Chem.* **1999,** *10,* 566.

[138] Gorenstein, D. G. *Phosphorus-31 NMR. Principles and Applications;* Academic Press, Inc.: New York, 1984.

[139] Dransfeld, A.; Chesnut, D. B. *Chem. Phys. Lett.* **1998,** *234,* 69.

[140] Chesnut, D. B.; Byrd, E. F. C. *Heteroatom Chem.* **1996,** *7,* 307.

[141] van Wüllen, C. *Phys. Chem. Chem. Phys.* **2000,** *2,* 2137.

[142] Patchkovskii, S.; Ziegler, T. *J. Phys. Chem. A* **2002,** *106,* 1088.

[143] Jameson, C. J.; de Dios, A.; Jameson, A. K. *Chem. Phys. Lett.* **1990,** *167,* 575.