A STUDY OF AB INITIO GENERALIZED VALENCE BOND THEORY: METHODOLOGY, PROGRAMMING AND APPLICATIONS

YOULIANG WANG

A Study of Ab Initio Generalized Valence Bond Theory: Methodology, Programming and Applications

By

©Youliang Wang, B. Eng., M.Sc.

A Thesis

Submitted to the School of Graduate Studies in Partial Fulfillment

of the Requirements for the Degree of Doctor of Philosophy

Department of Chemistry

Memorial University of Newfoundland

April 1995

St. John's, Newfoundland

ISBN 0-612-06157-4

Abstract

The ab initio generalized valence bond (GVB) method has been studied. A procedure for generating starting orbitals for GVB perfect-pairing (GVB-PP) calculations is presented. This is achieved by selecting initial orbitals which correspond to specific bonds or electron pairs. These orbitals can be identified from the localized molecular orbitals (LMOs), for both occupied and virtual orbitals, which are obtained through a unitary transformation of the Hartree-Fock (HF) canonical molecular orbitals (CMOs) using the Boys localization method. Initial GVB-PP orbitals can also be identified, to a limited extent, directly from the CMOs. A scheme has been implemented which achieves optimum convergence of the pairwise orbital optimization. An "object-oriented" GVB program was developed which automatically generates reliable initial GVB orbitals, leading to proper and fast convergence.

A number of properties were calculated to examine the GVB-PP wavefunction. The GVB/6-31G* dipole moments of carbon monoxide and formaldehyde are in very good agreement with the experimental and configuration interaction (CI) results. The special treatment of the ground state allyl radical is also discussed using GVB-PP

approach. The results clearly indicate the π conjugated structure in the C2-C1-C3 three-center bond, and our GVB approach gives the proper description of the π electron resonance structure.

Equilibrium geometries have been obtained at the GVB level for one- and two-heavy-atom molecules and radicals (AH_n and AH_mBH_n) containing first- or second-row elements. The results are compared with the available experimental and theoretical (HF, MP2, and CID) values. The effect of basis set on the GVB equilibrium geometries is also discussed. The results indicate that the addition of polarization or diffuse functions to the split valence basis set is necessary to obtain reasonable GVB geometries. In general, geometrical parameters from the GVB/6-31G* calculations, which treat electron correlation by using the coupling of GVB pairing orbitals, are in better agreement with the experimental data than those obtained from the corresponding HF wavefunction. The GVB/6-31G* geometries are close to the corresponding MP2 or CID geometries. GVB calculations are found to be especially useful for the proper description of multiple bonds.

The GVB wavefunction properly describes homolytic bond dissociation, A-B $\rightarrow$ A· + B·, giving reasonable bond dissociation energies. GVB/6-31G** calculations for A-H and A-B single bond dissociation energies (D_e) are examined systematically in this study. In general, the GVB potential energy curves are consistent with the experimental results. Reaction energies for hydrogenation and reactions converting multiple to single bonds, are also calculated to assess the GVB evaluation of heats of reaction. Overall, the GVB results correlate well with the available gas phase experimental values, demonstrating that GVB can give good estimates of gas-phase thermodynamical data.

GVB/6-31G** has been performed on $A_2X=YB_2$ (A, B=H, F; X, Y=C, Si) to obtain the optimized geometries for planar and twisted singlet structures, and rotational barriers, or π bond energies. The nature of the C-C, Si-Si, and C-Si π bonds has also been investigated. The results show that the C-C π bond energy decreases with increasing fluorine substitution. The Si-Si, C-Si π bond energies are much weaker. The pyramidalization at the carbon or silicon center for the twisted structures decreases the π bond energies in the substituted ethylenes and their silicon counterparts. Fluorine substitution stabilizes both diradical and dipolar twisted singlet structures.

GVB/6-31G** calculations on 1,1-dilithioethylene, 2,2-dilithiosilaethylene and 1,1 dilithiosilaethylene, were carried out to study their singlet, triplet structures, and rotation around the double bond. The singlet-triplet splittings for these systems were obtained along the twisting angle. The twisted triplet is predicted to be the ground state for 1,1-dilithioethylene. The GVB results indicate that there is a strong Li-Li bond in the Li-C-Li three-center group, which stabilizes the perpendicular structure of 1,1-dilithioethylene due to the increase of electron density in the CLi_2 group and the C-C π bond. The GVB dipole moments suggest that the electron deficient substituents withdraw more electron than expected in the triplet or perpendicular structures, where CH_2 acts as an electron donor to the CLi_2 group by retaining the

C=C double bond with a shorter C-C bond distance.

GVB/6-31++G** calculations on S_N2 reactions $Y^- + CH_3X \rightarrow CH_3Y + X^-$ were performed, for Y= F, Cl, OH, NH_2, SH with X=Cl, and for Y=Cl with X= F, OH, NH_2. Optimized structures were reported for reactants, products and transition states. In general, the GVB transition state is found to be looser than that obtained at HF. The calculated GVB frequencies are very close to the experimental harmonic frequencies for reactants and products. The secondary α-deuterium kinetic isotope effects (k_H/k_D) were calculated for these S_N2 reactions with a modified Sim's BEBOVIB-IV program. The results demonstrate that the magnitude of these isotope effects is determined by an inverse stretching vibration contribution and a normal bending vibration to the isotope effect. The out-of-plane bending vibration model for relating the magnitude of secondary α-deuterium kinetic isotope effects to transition state is correct. Larger isotope effects are found in looser S_N2 transition states.

The factors that influence the C=N stretching frequencies and the C=N π bond strength for unprotonated and protonated imines were investigated. The GVB C=N stretching force constants in retinal imine models, show a decrease by 0.47 mdyn/Å upon protonation. However, a higher C=N(H) stretching frequency has been obtained for protonated polyimines, due to a strong coupling between the C=N stretching and C=N-H(D) bending vibrations. On the other hand, protonation causes a strong charge alternation along the chain in all trans polyimines CH_2=(CH-CH=)$_n$NH, and mainly affects the electronic environment of the C=N bond. Our GVB calculations show a large C=N(H) stretching frequency deuterium isotope shift and ^{15}N isotope shift, ~25 and ~ 20 cm^{-1} respectively. The C=N(H) stretching frequency deuterium isotope shift can be regarded as an experimental measure to determine whether the imine nitrogen in visual pigment rhodopsin is protonated.

Overall our GVB/6-31G**//GVB/6-31G** results of dipole moments, equilibrium geometries, bond dissociation energies, heats of reaction, and vibrational frequencies, are comparable to the CISD and MP2 results. GVB calculations can now be widely used as an excellent post-HF method without the need for integral transformation.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Raymond A. Poirier, for his encouragement and guidance during this study. I am also very grateful to Drs. D. Jean Burnell and Peter G. Pickup for many helpful discussions.

It is my pleasure to thank my colleagues, Cory Pye and James Xidos, for their help with computing facilities and discussions.

I thank Memorial University for financial support. I would like to thank Computing and Communications Department for the computer time which made this work possible.

Last but not least, I thank my wife, Xiaomin, and my daughter, Xiao, for their patience, understanding and constant encouragement during the writing of this thesis.

Contents

List of Tables

List of Figures

Chapter One

Theoretical Background

1.1 Hartree-Fock Wavefunction

The objective of the Hartree-Fock molecular orbital method is to find the best possible one-configuration (single-determinant) solution to the Schrodinger equation

$$H\Psi = E\Psi \tag{1}$$

where H is the full non-relativistic electronic Hamiltonian, and Ψ is a determinant of spin-orbitals χ_i whose spatial components are the molecular orbitals (MO) ψ_i. The usual electronic Hamiltonian is given as

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

where N and M are the total number of electrons and nuclei, respectively, r_{iA} is the distance between the ith electron and Ath nucleus, and r_{ij} is the distance between the ith and jth electron. Z_A is the nuclear charge for nucleus A. The simplest case is the closed-shell ground state of a molecule with N electrons, doubly occupying $N/2$ orbitals. Its wavefunction can be written as

$$\Psi = \mathcal{A}|\psi_1\alpha\psi_1\beta\psi_2\alpha\psi_2\beta\cdots\psi_{N/2}\alpha\psi_{N/2}\beta\rangle \tag{3}$$

where $\mathcal{A}$ is the antisymmetrizing operator. This is referred to as a Slater determinant.

The Hartree-Fock (HF) approximation [1] is based on the variation method. The variation principle states that the best wavefunction is the one which gives the lowest possible energy,

$$E_0 = \langle\Psi_0|H|\Psi_0\rangle \tag{4}$$

In practical applications of this theory, the molecular orbitals ψ_i can be expressed as a linear combination of a finite set of basis functions ϕ_μ,

$$\psi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu \tag{5}$$

where $C_{\mu i}$ are the molecular orbital expansion coefficients. If atomic orbitals of constituent atoms are used as basis functions, the treatment is described as linear combination of atomic orbital (LCAO) theory. However, the mathematical treatment is more general, and any set of appropriately defined functions may be used for a basis expansion.

By minimizing E_0 with respect to the coefficients $C_{\mu i}$, one can derive the HF equations to determine the optimum orbitals. HF equations are eigenvalue equations of the form

$$f(1)\psi_i(1) = \epsilon_i\psi_i(1) \tag{6}$$

where $f(1)$ is an effective one-electron operator, called the Fock operator:

$$f(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}} + v^{HF}(1) = h(1) + v^{HF}(1) \qquad (7)$$

The Fock operator $f(1)$ is the sum of a core Hamiltonian operator $h(1)$ and an effective one-electron potential operator called the Hartree-Fock potential $v^{HF}(1)$, which is the average potential experienced by an electron due to the presence of the other electrons. Therefore, $v^{HF}(1)$ depends on the orbitals of the other electrons. Thus, the HF equations are not linear and must be solved in an iterative process called the self-consistent-field (SCF) method.

1.1.1 Closed-Shell Systems

In the closed-shell restricted ground state for a N electron system, each of the occupied spatial molecular orbitals is doubly occupied. As described before, the Fock operator is a sum of a core Hamiltonian operator $h(1)$ and an effective one-electron potential operator $v^{HF}(1)$,

$$v^{HF}(1) = \sum_{a}^{N/2} 2J_a(1) - K_a(1) \qquad (8)$$

where J_a and K_a are the coulomb and exchange operators, respectively, which are defined as,

$$J_a(1)\psi_i(1) = [\int \psi_a^*(2)\frac{1}{r_{12}}\psi_a(2)dr_2]\psi_i(1) \qquad (9)$$

$$K_a(1)\psi_i(1) = [\int \psi_a^*(2)\frac{1}{r_{12}}\psi_i(2)dr_2]\psi_a(1) \qquad (10)$$

The coulomb term represents the average local potential at r_1 arising from an electron in ψ_a. The exchange term arises from the antisymmetric nature of the single determinant and does not have a simple classical interpretation. The closed-shell HF energy is given by

$$E_0 = 2\sum_{i}^{N_{occ}} h_{ii} + \sum_{i,j}^{N_{occ}} (2J_{ij} - K_{ij}) \tag{11}$$

The integrated Hartree-Fock equation can be written as

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i}, \qquad i = 1, 2, \cdots, K \tag{12}$$

where $S_{\mu\nu}$ and $F_{\mu\nu}$ are elements of $K \times K$ Hermitian matrices **S** and **F**, given as

$$S_{\mu\nu} = \int \phi_\mu^*(1)\phi_\nu(1) dr_1 \tag{13}$$

$$F_{\mu\nu} = \int \phi_\mu^*(1) f(1) \phi_\nu(1) dr_1 \tag{14}$$

Eq. (12) is referred to as Roothaan's equation, which can be written as a single matrix equation,

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{15}$$

with the normalization condition,

$$\mathbf{C}^{+}\mathbf{SC} = 1 \tag{16}$$

C is a $K \times K$ square matrix of the expansion coefficients $C_{\mu i}$,

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1K} \\ C_{21} & C_{22} & \cdot\cdot & C_{2K} \\ & & \cdots & \\ C_{K1} & C_{K2} & \cdots & C_{KK} \end{pmatrix} \tag{17}$$

and $\boldsymbol{\epsilon}$ is a diagonal matrix of the orbital energies ϵ_i.

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \cdots & \\ & & & \epsilon_K \end{pmatrix} \tag{18}$$

For an orthonormal basis set, $\mathbf{S} = \mathbf{1}$, the Roothaan equations have the form of the usual matrix eigenvalue problem and the eigenvectors $\mathbf{C}$ and eigenvalues $\boldsymbol{\epsilon}$ are calculated by diagonalizing $\mathbf{F}$. The elements of the one-electron density matrix $\mathbf{P}$ are defined as,

$$P_{\lambda\sigma} = 2\sum_{i=1}^{N_{occ}} C_{\lambda i}^{*} C_{\sigma i} \tag{19}$$

The factor of two indicates that two electrons occupy each molecular orbital. Therefore,

$$F_{\mu\nu} = H_{\mu\nu}^{core} + \sum_{\lambda=1}^{N}\sum_{\sigma=1}^{N} P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma)] \tag{20}$$

In this expression, $H_{\mu\nu}^{core}$ is a matrix representing the energy of a single electron in a field of "bare" nuclei,

$$H_{\mu\nu}^{core} = \int \phi_{\mu}^{*}(1)h(1)\phi_{\nu}(1)dr_1, \tag{21}$$

and the quantities $(\mu\nu|\lambda\sigma)$ are two-electron repulsion integrals,

$$(\mu\nu|\lambda\sigma) = \int\int \phi_{\mu}^{*}(1)\phi_{\nu}(1)\frac{1}{r_{12}}\phi_{\lambda}^{*}(2)\phi_{\sigma}(2)dr_1 dr_2 \tag{22}$$

The electronic energy is now

$$E^{re} = \frac{1}{2}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} P_{\mu\nu}(F_{\mu\nu} + H_{\mu\nu}^{core}) \tag{23}$$

which, when added to the internuclear repulsion,

$$E^{nr} = \sum_{A<}^{M} \sum_{B}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{24}$$

yields an expression for the total energy $E = E^{er} + E^{nr}$.

1.1.2 Open-Shell Systems

Simple molecular orbital theory has been extended to open-shell systems in two ways. The first is described as spin-restricted Hartree-Fock (RHF) theory. In this approach, a single set of molecular orbitals is used. Some orbitals are doubly occupied, whereas some are singly occupied.

The second type of molecular orbital theory in common use for open-shell systems is spin unrestricted Hartree-Fock (UHF) theory. In UHF, different spatial orbitals are assigned to α and β electrons. There are two distinct sets of molecular orbitals ψ_i^α and ψ_i^β ($i = 1, 2, \cdots, N$). For instance, the electron configuration of a five-electron doublet may be written as $(\psi_1^\alpha \alpha)(\psi_1^\beta \beta)(\psi_2^\alpha \alpha)(\psi_2^\beta \beta)(\psi_3^\alpha \alpha)$. According to the variation principle the optimized UHF energy is below the optimized RHF value. The UHF wavefunction, however, is not an eigenfunction of the total spin operator, and is contaminated by functions corresponding to states of higher spin multiplicity.

In UHF theory, the two sets of molecular orbitals are defined by two sets of coefficients,

$$\psi_i^\alpha = \sum_{\mu=1}^{K} C_{\mu i}^\alpha \phi_\mu \tag{25}$$

$$\psi_i^{\beta} = \sum_{\mu=1}^{K} C_{\mu i}^{\beta} \phi_\mu \tag{26}$$

These coefficients are varied independently and results in two sets of equations.

$$\mathbf{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\epsilon^{\alpha} \tag{27}$$

$$\mathbf{F}^{\beta}\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\epsilon^{\beta} \tag{28}$$

ϵ^{α}, ϵ^{β} are diagonal matrices of orbital energies. $\mathbf{C}^{\alpha}$ and $\mathbf{C}^{\beta}$ are $K \times K$ matrices of expansion coefficients.

The above equations are referred to as Pople-Nesbet equations. Since $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ depend on both $\mathbf{C}^{\alpha}$ and $\mathbf{C}^{\beta}$ the two eigenvalue equations must be solved simultaneously. The procedure of solving the Pople-Nesbet equations is essentially identical with the procedure used for solving Roothaan's equations for closed-shell systems.

1.1.3 Localized Molecular Orbitals

As eigenfunctions of the Fock operator, $\{\psi_i, i = 1, 2, \cdots, K\}$ have the special property that they are symmetry adapted, *i.e.*, they have certain symmetry properties characteristic of the symmetry of the molecule, and are referred to as canonical molecular orbitals (CMOs). Consequently, CMOs will form a basis for an irreducible representation of the point group of the molecule, and are generally delocalized over the whole molecule.

Once the CMOs have been obtained, they can be transformed to an infinite number of equivalent sets of MOs by a unitary transformation. [1] In particular,

[1] Such a transformation keeps the SCF results invariant.

there are various criteria [2] for choosing a unitary transformation so that the transformed molecular orbitals are in some sense localized. The localized molecular orbitals (LMOs) are an important conceptual and quantitative link between chemical intuition and quantum theory for molecules, and allow the understanding of many molecular properties in terms of bonds and lone pairs.

The Edminston-Ruedenberg [2(a)] localization method is based on the energetic criterion that the self-repulsion of the orbitals be maximized, *i.e.*,

$$\sum_i \langle \psi_i^2(1) | \frac{1}{r_{12}} | \psi_i^2(2) \rangle = maximum \tag{29}$$

Since it does not depend on the LCAO basis and can be applied in principle even to numerical CMOs, this localization is referred to as an "intrinsic" localization criterion. In this scheme, a "two-by-two" rotation is applied to each pair of MO, until the procedure has converged. In most cases, the Edminston-Ruedenberg LMOs are in very close agreement with the conventional chemical picture and represent clearly the core, lone pairs and bonds of the molecule. Since the two-electron integrals need to be calculated in each step, the Edminston-Ruedenberg procedure is rather time consuming for large molecules.

Boys [2(b)] proposed another localization procedure, where LMOs are defined to maximize the product of the squares of the distances between the centroids of charges. In this case, the unitary transformation is to maximize the distance between the centroids of charges, *i.e.*,

$$\sum_{i<j} |\langle\psi_i(1)|r(1)|\psi_i(1)\rangle - \langle\psi_j(1)|r(1)|\psi_j(1)\rangle| = \text{maximum} \tag{30}$$

The Boys localization is also achieved by "two-by-two" rotation. The Boys LMOs are usually similar to the Edminston-Ruedenberg ones. The main advantage of the Boys localization procedure is that it is faster than the Edminston-Ruedenberg one for large molecules, since only "dipole integrals", $\langle\psi_i(1)|r(1)|\psi_j(1)\rangle$, need to be computed.

1.1.4 Natural Orbitals

Given a normalized wavefunction Φ, then $\Phi(x_1,\cdots,x_N)\Phi^*(x_1,\cdots,x_N)dx_1\cdots dx_N$ is the probability that an electron is in the space-spin volume element dx_1 located at x_1, while simultaneously another electron is in dx_2 at x_2 and so on. The reduced density function, $\rho(x_1)$, for a single electron in an N-electron system is defined as,

$$\rho(x_1) = N\int \Phi(x_1,\cdots,x_N)\Phi^*(x_1,\cdots,x_N)dx_2\cdots dx_N \tag{31}$$

The normalization factor N is included so that the integral of the density equals to the total number of electrons,

$$\int \rho(x_1)dx_1 = N \tag{32}$$

A density matrix, $\gamma(x_1,x_1')$ is generalized from the density function $\rho(x_1)$, defining

$$\gamma(x_1,x_1') = N\int \Phi(x_1,\cdots,x_N)\Phi^*(x_1',\cdots,x_N)dx_2\cdots dx_N \tag{33}$$

The matrix $\gamma(x_1, x_1')$ is called the first-order reduced density matrix or the one-electron reduced density matrix.

As a function of two variables, $\gamma(x_1, x_1')$ can be expanded in the orthonormal basis of Hartree-Fock spin orbitals $\{\chi_i\}$ as,

$$\gamma(x_1, x_1') = \sum_{i,j} \chi_i(x_1)\gamma_{ij}\chi_j^*(x_1') \tag{34}$$

where

$$\gamma_{ij} = \int \chi_i^*(x_1)\gamma(x_1, x_1')\chi_j(x_1')dx_1 dx_1' \tag{35}$$

It is always possible to define an orthonormal basis $\{\eta_i\}$, related to $\{\chi_i\}$ by a unitary transformation, which diagonalizes the matrix $\boldsymbol{\Gamma}$ whose elements are $(\boldsymbol{\Gamma})_{ij} = \gamma_{ij}$.

$$\boldsymbol{\Lambda} = \mathbf{U}^{+}\boldsymbol{\Gamma}\mathbf{U} \tag{36}$$

The elements of the diagonal matrix $\boldsymbol{\Lambda}$ are $\lambda_{ij} = \delta_{ij}\lambda_i$. The elements of the orthonormal set in which $\boldsymbol{\Gamma}$ is diagonal, are called the natural spin orbitals (NO) [3], $\eta_i = \sum_k \chi_k U_{ki}$. λ_i is referred to as the occupation number of the natural orbital η_i in the wavefunction Φ.

The importance of natural orbitals is that they simplify the energy expression for multiconfiguration wavefunctions, and give the most rapidly convergent expansion for configuration interaction (CI) type of calculations.

1.2 Multiple-Determinant Wavefunction

For many chemical systems [4, 5], the single-determinant wavefunction, *i.e.*, restricted Hartree-Fock (RHF) or unrestricted Hartree-Fock (UHF) wavefunction, provides an excellent description of the electronic structures. Configuration interaction studies have shown that the HF wavefunction is the most important configuration in the "exact" wavefunction for molecules near the equilibrium geometry.

The primary deficiency of Hartree-Fock theory is the inadequate treatment of electron correlation. The correlation energy is defined as the difference between the Hartree-Fock limit (E_{HF}) and exact (nonrelativistic) energy ($\mathcal{E}_0$),

$$E_{corr} = \mathcal{E}_0 - E_{HF} \tag{37}$$

This limitation of the closed-shell Hartree-Fock single-determinant wavefunction, therefore, leads to an improper description of bond dissociation. For example, the restricted molecular orbitals of H_2 at the minimal basis set are

$$\psi_1 = [2(1 + S_{12})]^{-\frac{1}{2}}(1s_A + 1s_B) \tag{38}$$

$$\psi_2 = [2(1 - S_{12})]^{-\frac{1}{2}}(1s_A - 1s_B) \tag{39}$$

where $1s_A$ and $1s_B$ are the $1s$ atomic orbitals for the constituent hydrogen atoms. The restricted single determinant wavefunction is

$$|\Psi_0\rangle = |\psi_1\bar{\psi}_1\rangle \tag{40}$$

In this restricted calculation, both electrons are forced to occupy the same spatial molecular orbital ψ_1, independent of the bond length. The spatial part of the wave function is

$$\psi_1(1)\psi_1(2) = \frac{1}{2}[1s_A(1)1s_B(2)+1s_B(1)1s_A(2)]+\frac{1}{2}[1s_A(1)1s_A(2)+1s_B(1)1s_B(2)] \quad (11)$$

The second term on the right-hand side of Eq. (11) can be regarded as a linear combination of two ionic states, one with both electrons on atom A and the other with both electrons on atom B, *i.e.*, $H_A^-H_B^+$ and $H_A^+H_B^-$. According to this wavefunction, the two electrons spend half the time on A and half on B, even when the centers are infinitely separated. Such a description is inappropriate for two separated hydrogen atoms.

A multiple-determinant wavefunction is required for cases which a single-determinant wavefunction is not able to describe properly.

1.2.1 Configuration Interaction

After solving Roothaan's equations in a finite basis set, a set of $2K$ spin orbitals $\{\chi_i\}$ is obtained. The determinant formed from the N lowest energy spin orbitals is the HF determinant $|\Psi_0\rangle$. In addition to $|\Psi_0\rangle$ a large number of other N-electron determinants can be formed from the $2K$ spin orbitals. These determinants are the singly excited determinants $|\Psi_a^r\rangle$, the doubly excited determinants $|\Psi_{ab}^{rs}\rangle$, etc., up to N-tuply excited determinants. These determinants are now used as a basis to expand the exact wavefunction $|\Phi_0\rangle$:

$$|\Phi_0\rangle = C_0|\Psi_0\rangle + \sum_a \sum_r C_a^r|\Psi_a^r\rangle + \sum_{a<b} \sum_{r<s} C_{ab}^{rs}|\Psi_{ab}^{rs}\rangle + \sum_{a<b<c} \sum_{r<s<t} C_{abc}^{rst}|\Psi_{abc}^{rst}\rangle + \cdots \quad (42)$$

This is referred to as the full CI wavefunction [6]. The number of the excited determinants is extremely large even for small molecules with basis sets of medium size. But some of these determinants can be eliminated by utilizing the fact that wavefunctions with different spins do not mix, and linear combination of some determinants can be taken to form proper spin-adapted configurations, which are eigenfunctions of S^2.

Full CI is computationally feasible only for very small molecules. As the number of configurations grows very rapidly, the dimensionality of the full CI matrix becomes computationally impractical. Therefore, the CI expansion for the wavefunction has to be truncated.

Inclusion of single excitations only, termed configuration interaction, singles, or CIS:

$$|\Psi_{CIS}\rangle = C_0|\Psi_0\rangle + \sum_a \sum_r C_a^r|\Psi_a^r\rangle \quad (43)$$

normally leads to no improvement relative to the Hartree-Fock wavefunction or energy. The simplest procedure to have an effect on the calculated energy is the inclusion of double excitations only, which is termed configuration interaction, doubles, or CID:

$$|\Psi_{CID}\rangle = C_0|\Psi_0\rangle + \sum_{a<b} \sum_{r<s} C_{ab}^{rs}|\Psi_{ab}^{rs}\rangle \quad (44)$$

CID is an important practical procedure. At a slightly higher level of theory, both single and double excitations can be included in the configuration interaction treatment.

This model is termed configuration interaction, singles and doubles, or CISD:

$$|\Psi_{CISD}\rangle = C_0|\Psi_0\rangle + \sum_a \sum_r C_a^r|\Psi_a^r\rangle + \sum_{a<b} \sum_{r<s} C_{ab}^{rs}|\Psi_{ab}^{rs}\rangle \tag{45}$$

Here, all coefficients (C_0, C_a^r, C_{ab}^{rs}) are varied to minimize the expectation value of the energy. Since there are nonzero matrix elements of the Hamiltonian between singly and doubly excited determinants, the single excitations contribute to the wavefunction.

However, the most serious deficiency of the CID or CISD is that they fail to satisfy the size-consistency condition [5]. To obtain the corrections to the CID or CISD, a number of efforts have been made. The most commonly used of them is the Davidson's correction [7].

$$\Delta E_{corr} = \Delta E_{CISD} + \Delta E_{CIQ} \tag{46}$$

with

$$\Delta E_{CIQ} \approx (1 - C_0^2)\Delta E_{CISD} \tag{47}$$

where ΔE_{CISD} is the correlation energy at the CISD level, and C_0 is the coefficient of the HF wavefunction $|\Psi_0\rangle$ in the CISD expansion. This corrects a major part of the discrepancy. However, the total energy is still not precisely size-consistent.

1.2.2 Møller-Plesset Perturbation Theory

An important advantage of CI is that it is variational, but as pointed out it is not generally size-consistent. A different systematic procedure for including the correlation energy, which is not variational but is size-consistent, is perturbation theory.

The perturbation theory of Møller-Plesset (MP) [8], closely related to many-body perturbation theory (MBPT) [9], is an approach to the correlation problem. MP models are formulated by introducing a generalized electronic Hamiltonian, H_λ, according to

$$H_\lambda = H_0 + \lambda V \tag{48}$$

where λ is a dimensionless parameter, λV is the perturbation term, and H_0 is taken to be the sum of the one-electron Fock operators. Ψ_λ and E_λ, the exact ground-state wavefunction and energy for a system described by the H_λ, can be expanded in power of λ according to Rayleigh-Schrodinger perturbation theory,

$$\Psi_\lambda = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \cdots \tag{49}$$

$$E_\lambda = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots \tag{50}$$

We call E^n the nth-order energy.

In Møller-Plesset theory, the method is referred to as MP2, MP3 and so forth for second-order, third-order, ..., respectively. The MP energy to first-order is the HF energy. Higher terms are the expansion involving other matrix elements of the operator V. For example, the second-order contribution to the MP energy is

$$E^{(2)} = -\sum_s^D \frac{|V_{s0}|^2}{E_0 - E_s} \tag{51}$$

where $\sum^{D}$ indicates that summation is to be carried out over all double substitutions. If Ψ_s is the double substitution $(ij \rightarrow ab)$, the explicit expression for V_{s0} is

$$V_{s0} = \langle ab||rs\rangle \tag{52}$$

where $\langle ab||rs\rangle$ is an antisymmetrized two-electron integral over spin orbitals, defined as

$$\langle ij||kl\rangle = \langle ij|kl\rangle - \langle ij|lk\rangle \tag{53}$$

In physicists' notation,

$$\langle ij|kl\rangle = \int \chi_i^*(1)\chi_j^*(2)\frac{1}{r_{12}}\chi_k(1)\chi_l(2)dx_1dx_2 \tag{54}$$

The final formula for the second-order contribution to the energy is

$$E^{(2)} = -\sum_{a<b}^{occ}\sum_{r<s}^{virt}\frac{|\langle ab||rs\rangle|^2}{\epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b} \tag{55}$$

An important point to note is that both CI and MP2 require a transformation of the two-electron integrals from the basis functions, ϕ_μ, into molecular orbitals ψ_p.

$$(pq|rs) = \int\int \psi_p^*\psi_q(1)\frac{1}{r_{12}}\psi_r(2)^*\psi_s(2)dr_1dr_2 = \sum_\mu\sum_\nu\sum_\lambda\sum_\sigma C_{\mu p}C_{\nu q}C_{\lambda r}C_{\sigma s}(\mu\nu|\lambda\sigma)$$

where $(\mu\nu|\lambda\sigma)$ are the two-electron integrals over atomic orbitals [Eq. (22)], and molecular orbitals ψ_p are defined in the usual manner as linear combinations of K basis functions ϕ_μ according to Eq. (5). A number of efforts have been made in the development of efficient algorithms for integral transformation.

1.2.3 Multiconfiguration SCF (MCSCF)

A multiconfiguration self-consistent-field (MCSCF) wavefunction is a short CI expansion in which both orbitals and expansion coefficients are optimized.

$$|\Psi_{MCSCF}\rangle = \sum_I C_I |\Psi_I\rangle \tag{56}$$

The MCSCF energy is obtained by minimizing $\langle\Psi_{MCSCF}|H|\Psi_{MCSCF}\rangle$, to determine the optimum C_I coefficients and the optimum form of the orbitals simultaneously. Obviously, the MCSCF approach in principle should show much faster convergence than CI. Details of MCSCF methods can be found in Ref. [10].

Unless restricted by symmetry conditions, MCSCF orbitals tend to be localized. This localization is particularly effective for describing molecular dissociation and bond-breaking process in general, and the removal of symmetry restriction has also been found to produce improved MCSCF results for normal bonding systems.

1.2.4 Valence Bond Approach

The valence bond (VB) approach [11] was first proposed by Heitler and London for H_2, and further developed by Slater and Pauling for extension to more complicated molecules. The VB wavefunction for H_2 is,

$$\Psi_{VB} = (\frac{1}{2})^{1/2}[1s_A(1)1s_B(2) + 1s_B(1)1s_A(2)] \times (\alpha\beta - \beta\alpha) \tag{57}$$

In terms of the VB wavefunction, the two electrons are on different atoms, *i.e.*, one on A and one on B. This is the correct wavefunction for singlet state of H_2 at infinite separation.

In general, for a polyatomic molecule, there are many possible bonding structures. For example, a structure with n atomic orbitals can be considered as a system of $n/2$ chemical bonds, where each bond is described by a Heitler-London type wavefunction.

$$\Psi_{VB} = \mathcal{A}\prod_{i=1}^{n/2}[\phi_i(r_i)\phi_{i+1}(r_{i+1}) + \phi_{i+1}(r_i)\phi_i(r_{i+1})](\alpha\beta - \beta\alpha) \tag{58}$$

This function, when expanded, is equivalent to a linear combination of $2^{n/2}$ determinants.

The classical VB molecular wavefunction is constructed from orbitals that are optimized for the separated atoms. As the atoms approach each other, this wavefunction is relaxed by modifying the orbitals within the valence space of the atoms (*i.e.*, hybridization), including ionic structures and structures with different bonds (resonance). Such relaxation is valuable in qualitative analysis, and the related concepts, such as hybridization and resonance, are very useful in organic chemistry. However, in order to perform ab initio calculations with a quantitative and predictive power, the number of bonding structures required in the VB wavefunction becomes very large even for a small molecule. A possible remedy for this shortcoming is the non-linear, variational optimization of the orbitals.

1.2.5 Generalized Valence Bond Method

The ab initio generalized valence bond (GVB) wavefunction has the same form as the VB wavefunction but allows all orbitals to be solved self-consistently (as in

Hartree-Fock). In the GVB approach, no special hybridization is imposed on the orbitals, and the orbitals are, in principle, permitted to delocalize onto other centers. The GVB wavefunction was first proposed by Goddard [12] in its complete form, which is a special case of multiconfiguration self-consistent-field (MCSCF) method [13]. In GVB wavefunction, each pair of electrons is described by two or more self-consistently optimized molecular orbitals while retaining proper spin symmetry.

The basic idea of the GVB method can be simply illustrated by considering its application to H_2,

$$\Psi_{H_2}^{GVB} = N_1[\psi_{1i}^g\psi_{2i}^g + \psi_{2i}^g\psi_{1i}^g](\alpha\beta - \beta\alpha) \tag{59}$$

where N_1 is a normalizing factor, and the nonorthogonal GVB molecular orbitals, $\psi_{1i}^g, \psi_{2i}^g(\langle\psi_{1i}^g|\psi_{2i}^g\rangle = S_{12})$, are determined variationally. Using the natural orbitals, $\psi_{1i}, \psi_{2i}(\langle\psi_{1i}|\psi_{2i}\rangle = 0)$, the one-pair GVB wavefunction can be expressed as

$$\Psi_{H_2}^{GVB} = N_2[\sigma_{1i}\psi_{1i}\psi_{1i} - \sigma_{2i}\psi_{2i}\psi_{2i}](\alpha\beta - \beta\alpha) \tag{60}$$

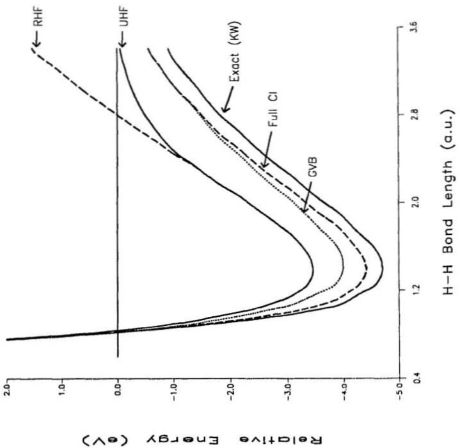
where σ_{1i}, σ_{2i} are referred to as the GVB CI-coefficients. The pair of electrons is now described by two orbitals which are associated with the "bonding" (ψ_{1i}) and "antibonding" (ψ_{2i}) orbitals.

The GVB wavefunction (or equivalently the MCSCF method using two configurations) has the correct form for the singlet state of H_2 at large separation and properly describes the bond dissociation of H_2 into hydrogen atoms. The GVB potential curves of H_2 are illustrated in Figure 1, along with the RHF, UHF, full CI and exact result of Kolos-Wolniewicz [14]. It is clear from Figure 1 that a restricted closed-shell HF calculation fails at large distances, whereas an unrestricted (UHF) wavefunction gives the proper dissociation limit for H_2. However, the unrestricted wavefunction is not a pure singlet. At $R \to \infty$, the wavefunction is contaminated by a triplet.

$$\lim_{R \to \infty} |\Psi_0^{UHF}\rangle = \frac{1}{2}[|\psi_1\bar{\psi}_1\rangle - |\psi_2\bar{\psi}_2\rangle - 2^{\frac{1}{2}}|^3\Psi_1^2\rangle] \quad (61)$$

The GVB potential curve improves with increasing internuclear separation, compared to HF. When R is greater than 3 a.u., the GVB and full CI results are virtually indistinguishable. At R=1.4 a.u., E_{corr}(GVB) =-0.4963 eV, 54% of the full CI correlation energy of -0.9194 eV. The full CI correctly describes the bond dissociation and significantly improves the well depth. The full CI/6-31G** curve nearly parallels the Kolos-Wolniewicz's exact result.

Figure 1: Potential energy curves for the dissociation of H_2.
The RHF, UHF, GVB and full CI calculations are performed with 6-31G** basis sets.
The exact potential curves are from Ref. [14].



RHF
UHF
Exact (KW)
Full CI
GVB
H–H Bond Length (a.u.)
Relative Energy (eV)
0.4
1.2
2.0
2.8
3.6
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
-5.0

An important advantage of the GVB wavefunction is that the GVB orbitals resemble the classical chemically bonded structures, *e.g.*, with "localized" bonds and lone pairs [16]. The emphasis in GVB is on using these orbitals to understand the role of electronic structure in chemical processes. Previous studies [15] have shown that GVB orbitals are highly localized, and very similar to the localized molecular orbitals (LMOs) using the Boys criterion, for example. The GVB orbitals are, however, obtained by a rigorous energetic "localization" technique, whereas the Boys LMOs [2(b)] are obtained by a maximal separation of centroids of charge through a unitary transformation of the HF orbitals. Generally, for a closed-shell system with N electrons, a full GVB perfect-pairing (GVB-PP) wavefunction involves a set of $N/2$ "localized" occupied orbitals and $N/2$ "localized" virtual orbitals which are optimized for electron correlation. The qualitative orbital view of molecules derived from ab initio GVB calculations leads to simple concepts relating to geometries. bond energies, and ordering of electronic states [16].

Another advantage of the GVB wavefunction is that, like SCF calculations, it does not require an integral transformation. GVB orbitals are also a good starting point for post-HF (CI or perturbation) calculations. A further advantage is that the energy gradients can be obtained analytically, allowing for the gradient optimization of geometries. This makes it possible to calculate, for example, dipole moments [17], force constants and fundamental frequencies at the equilibrium geometry with the inclusion of electron correlation.

However, in the GVB iterative process, one difficulty is that, unlike the HF method, it cannot be used as a "blackbox". For GVB calculations, molecular orbitals must be correctly paired using some initial molecular orbitals. The initial perfect-pairing orbitals are either set by some arbitrary default or selected by carefully pairing specific orbitals. If delocalized canonical molecular orbitals (CMOs) are used as the initial GVB orbitals, it is impossible to identify these in terms of specific bonds or lone pairs, and the resulting orbitals may not represent the lowest energy solution.

Even when the orbital pairing is correct, GVB convergence is in general very poor compared to HF. Recently, Goddard et al. [18] have significantly improved GVB convergence using the direct inversion of the iterative subspace (DIIS) technique and a special GVB orbital guess. A reliable initial guess is required to ensure that the GVB calculation is within the DIIS radius of convergence. Although fast convergence is important, a correct initial guess is essential to ensure convergence on the correct state, especially for distorted geometries. Pulay et al. used the natural orbitals of the unrestricted Hartree-Fock wavefunction for generating starting orbitals for GVB and MCSCF calculations [19]. Goddard et al. [18] developed a procedure for constructing initial GVB orbitals by localizing a pseudo-Hartree-Fock (P-HF) molecular orbitals based on the SCF orbitals of the individual atoms. In their approach, the occupied P-HF orbitals are projected on atomic basis functions for GVB first natural orbitals, and the unoccupied HF orbitals are projected for GVB second natural orbitals.

For the GVB-PP wavefunction in Eq. (60), both GVB orbitals ψ_{1i}, ψ_{2i} and GVB CI-coefficients σ_{1i}, σ_{2i} are to be optimized for the minimum energy solution. The normal procedure is to optimize the orbitals first. After the converged orbitals are obtained the GVB CI-coefficients are calculated and the procedure is repeated until a self-consistent solution is obtained. In the iterative process, good starting GVB-PP orbitals will ensure a fast and proper energy convergence. The pairwise orbital optimization procedure [12] is very straightforward.

The main objectives of this study are to develop a scheme for automatically generating the initial GVB-PP orbitals from LMOs and to improve GVB convergence using the pairwise orbital optimization procedure.

Chapter Two

Methodology

2.1 GVB Perfect-Pairing (GVB-PP) Approach

2.1.1 GVB-PP Wavefunction

A simple generalization of the HF wavefunction is to allow selected electron pairs to be described by two or more self-consistently optimized orbitals:

$$\phi_i\phi_i(\alpha\beta - \beta\alpha) \rightarrow (\phi^g_{1i}\phi^g_{2i} + \phi^g_{2i}\phi^g_{1i})(\alpha\beta - \beta\alpha) \tag{62}$$

The result is the GVB wavefunction. The closed-shell orbital description of a singlet electron pair is replaced by a GVB pair consisting of two nonorthogonal orbitals coupled into a singlet [12]. In this GVB wavefunction, the strong orthogonality (SO) constraint requires that all orbitals, other than the two within a given singlet pair, be orthogonal. The general form of the GVB wavefunction is

$$\Psi_{GVB} = \mathcal{A}[\Psi_{core}\Psi_{pair}\Psi_{open}] \tag{63}$$

where $\mathcal{A}$ is the antisymmetrizer. For a system with N electrons, $2N_{core}$ electrons are described with doubly occupied orbitals. and N_{open} electrons are high spin coupled ($S = N_{open}/2$):

$$\begin{aligned} \Psi_{core} &= \prod_{i=1}^{N_{core}} (\psi_i^{core}\alpha)(\psi_i^{core}\beta) \qquad \text{and} \\ \Psi_{open} &= \prod_{i=1}^{N_{open}} (\psi_i^{open}\alpha) \end{aligned} \tag{64}$$

where $\psi_i^{core}, \psi_i^{open}$ are orthogonal spatial orbitals. Ψ_{pair}, which is referred to as the restricted active space, can be constructed in a number of ways depending on the choice of the spin-coupling. In the perfect-pairing spin-coupling procedure, the singlet is coupled by pairs of orbitals $(\psi_{1i}^g, \psi_{2i}^g)$ with spin functions $\Theta_{PP} = \prod_{i=1}^{N_{pair}} (\alpha\beta - \beta\alpha)$, where $N_{pair} = (N - N_{open})/2 - N_{core}$. The GVB wavefunction, in which the perfect-pairing and strong orthogonality restrictions are both imposed, is referred to as the GVB-PP wavefunction, and Ψ_{pair} is given as

$$\Psi_{pair} = \prod_{i=1}^{N_{pair}} (\psi_{1i}^g\psi_{2i}^g + \psi_{2i}^g\psi_{1i}^g)(\alpha\beta - \beta\alpha) \tag{65}$$

where ψ_{1i}^g, ψ_{2i}^g are the GVB orbitals in pair i, with $\langle\psi_{1i}^g|\psi_{2i}^g\rangle = S_i$. Computationally, it is more convenient to use orthogonal orbitals. The nonorthogonal GVB orbitals ψ_{1i}^g, ψ_{2i}^g can be replaced by the orthogonal orbitals ψ_{1i}, ψ_{2i} through the following transformation [12],

$$\begin{cases} \psi_{1i}^g = (\sigma_{1i} + \sigma_{2i})^{-\frac{1}{2}}(\sigma_{1i}^{\frac{1}{2}}\psi_{1i} + \sigma_{2i}^{\frac{1}{2}}\psi_{2i}) \\ \psi_{2i}^g = (\sigma_{1i} + \sigma_{2i})^{-\frac{1}{2}}(\sigma_{1i}^{\frac{1}{2}}\psi_{1i} - \sigma_{2i}^{\frac{1}{2}}\psi_{2i}) \\ \langle\psi_{1i}|\psi_{2i}\rangle = 0 \\ \sigma_{1i}^2 + \sigma_{2i}^2 = 1, \qquad \text{and} \qquad \sigma_{1i}, \sigma_{2i} > 0 \end{cases} \tag{66}$$

Under such a transformation each GVB-PP orbital in Eq. (65) is replaced by

$$(\psi^g_{1i}\psi^g_{2i} + \psi^g_{2i}\psi^g_{1i})(\alpha\beta - \beta\alpha) \rightarrow (\sigma_{1i}\psi_{1i}\psi_{1i} - \sigma_{2i}\psi_{2i}\psi_{2i})(\alpha\beta - \beta\alpha) \tag{67}$$

For a two-electron system, this GVB wavefunction is identical with the simple two configuration MCSCF wavefunction

$$\Psi_{MCSCF} = C_A|\psi_{1i}\alpha\psi_{1i}\beta\rangle + C_B|\psi_{2i}\alpha\psi_{2i}\beta\rangle \tag{68}$$

where C_A, C_B are configuration coefficients to be optimized. The natural orbital pairs (ψ_{1i}, ψ_{2i}) are generally well localized and can be easily identified with specific bonds or electron pairs, where ψ_{1i} can be associated with the bonding type orbital and ψ_{2i} with an antibonding type orbital.

Using the orthogonal GVB orbitals, the energy expression for the GVB-PP wavefunction becomes similar to the HF expression,

$$E = 2\sum_{i=1}^{N_{occ}} f_i h_{ii} + \sum_{i,j=1}^{N_{occ}} (a_{ij}J_{ij} + b_{ij}K_{ij}) \tag{69}$$

where N_{occ} is the number of the occupied orbitals, and the standard definitions for the one-electron(h_{ii}), coulomb(J_{ij}), and exchange(K_{ij}) energies are used. The orbital occupation coefficients f_i and the two-electron coupling coefficients a_{ij}, b_{ij}, depend on the GVB wavefunction, which are given by,

$$f_i = \begin{cases} 1 & \text{if } \psi_i \text{ is a core orbital,} \\ \sigma_i^2 & \text{if } \psi_i \text{ is a pair orbital with coefficient } \sigma_i, \\ \frac{1}{2} & \text{if } \psi_i \text{ is an open-shell orbital.} \end{cases} \tag{70}$$

and

$$
\begin{aligned}
a_{ij} &= 2f_i f_j \\
b_{ij} &= -f_i f_j
\end{aligned}
\tag{71}
$$

except that

$$
\begin{aligned}
&a_{ij} = \tfrac{1}{2}, \quad b_{ij} = -\tfrac{1}{2}, && \text{if } \psi_i \text{ and } \psi_j \text{ are both open-shell orbitals.} \\
&a_{ii} = f_i, \quad b_{ii} = 0, && \text{if } \psi_i \text{ is a pair orbital,} \\
&a_{ij} = 0, \quad b_{ij} = \sigma_i \sigma_j, && \text{if } \psi_i \neq \psi_j \text{ are in the same pair.}
\end{aligned}
\tag{72}
$$

The energy expression of Ψ_{GVB-PP} in natural orbitals is simple enough to allow for efficient optimization of both molecular orbitals and GVB CI-coefficients. In general, the GVB-PP calculations are described using the notation, GVB(number of pairs/number of PP orbitals), where the number of PP orbitals is equal to the number of determinants.

2.1.2 Pairwise Orbital Optimization

The general condition for the optimum orbitals, corresponding to the energy expression given by Eq. (69), is that the first-order change in the energy due to changes in the orbitals is zero, leading to

$$
\sum_i \langle \delta\psi_i | F_i | \psi_i \rangle = 0,
\tag{73}
$$

where F_i is the Fock operator for the orbital ψ_i,

$$
F_i = f_i h + \sum_j^{N_{occ}} (a_{ij} J_j + b_{ij} K_j)
\tag{74}
$$

where h is the one-electron operator and J_j, K_j are the coulomb and exchange operators [12]. The energy expression can be written as

$$E = \sum_i [f_i h_{ii} + \langle \psi_i | F_i | \psi_i \rangle] \tag{75}$$

For the optimum orbitals ψ_i and ψ_j, the variational condition is

$$\langle \psi_i | (F_i - F_j) | \psi_j \rangle = 0 \tag{76}$$

Because the Fock operator depends on the orbitals, the variational equations must be solved iteratively.

The independent pairwise optimization allows mixing between only two orbitals at a time,

$$\begin{cases} \psi_i' = (\psi_i + \lambda_{ij}\psi_j)/(1 + \lambda_{ij}^2)^{\frac{1}{2}} \\ \psi_j' = (\psi_j + \lambda_{ji}\psi_i)/(1 + \lambda_{ji}^2)^{\frac{1}{2}} \end{cases} \tag{77}$$

where λ_{ij}, λ_{ji} are the orbital correction factors. Generally, since orbital orthonormality must always be preserved, these corrections cannot be fully independent and orbital orthogonality requires that $\lambda_{ji} = -\lambda_{ij}$. In keeping with the variational condition that the matrix elements $A_{ij} = \langle \psi_i | (F_i - F_j) | \psi_j \rangle$ must be zero, under the second-order approximation [12], the orbital correction factors are given as,

$$\lambda_{ij} = \frac{\langle \psi_i | F_j - F_i | \psi_j \rangle}{\langle \psi_i | (F_j - F_i | \psi_i \rangle - \langle \psi_j | (F_j - F_i) | \psi_j \rangle + B_{ij}} \tag{78}$$

where $B_{ij} = 2(a_{ii} + a_{jj} - 2a_{ij})K_{ij} + (b_{ii} + b_{jj} - 2b_{ij})(J_{ij} + K_{ij})$.

Pairwise orbital optimization deals with the variation of only one variable at a time, and is expected to converge to a solution directly (linearly) accessible through individual pairwise changes. In the pairwise orbital mixing scheme, once λ_{ij} has been determined, orbitals φ_i, φ_j are redefined before proceeding to the next pair. Because each mixing coefficient is determined in a completely independent way, this approach tends to have some oscillatory behaviour. This simple procedure does, however, converge successfully.

2.1.3 GVB CI-Coefficients

In addition to the above orbital optimization, the GVB CI-coefficients σ_{1i}, σ_{2i} in Eq. (67) are optimized in terms of the GVB-PP energy contribution (E_i) from the ith pairing orbitals ψ_{1i}, ψ_{2i}.

$$E_i = \frac{1}{2}(\sigma_{1i}^2 T_{11} + \sigma_{2i}^2 T_{22} + \sigma_{1i}\sigma_{2i} K_{1i,2i}) \tag{79}$$

where

$$\begin{cases} T_{11} = h_{1i,1i} + \frac{1}{2} J_{1i,1i} + \sum\limits_{j \neq 1i,2i} f_j (2J_{j,1i} - K_{j,1i}) \\ T_{22} = h_{2i,2i} + \frac{1}{2} J_{2i,2i} + \sum\limits_{j \neq 1i,2i} f_j (2J_{j,2i} - K_{j,2i}) \end{cases} \tag{80}$$

By defining

$$\alpha_i = \sigma_{2i}/\sigma_{1i} \quad \text{and} \quad T_i = T_{22} - T_{11} \tag{81}$$

and requiring $dE_i/d\alpha_i = 0$,[2] one can get

$$\alpha_i = \frac{T_i}{K_{1i,2i}} - [1 + (\frac{T_i}{K_{1i,2i}})^2]^{\frac{1}{2}} \tag{82}$$

then

$$\sigma_{1i} = \frac{1}{(1+\alpha_i^2)^{\frac{1}{2}}} \quad \text{and} \quad \sigma_{2i} = \frac{\alpha_i}{(1+\alpha_i^2)^{\frac{1}{2}}} \tag{83}$$

[2]Assuming $K_{1i,2i} \neq 0$, $K_{1i,2i} = 0$ results in collapse of the GVB pair.

In the case of more than one pair, the GVB CI-coefficients will be obtained in an iterative procedure, Eqs. (70) and (80).

2.1.4 Basis Sets

Molecular orbitals are expressed as a linear combination of a set of basis functions. $\psi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu$. In the procedure of using contracted Gaussian functions, each basis function is a fixed linear combination (contraction [20]) of Gaussian functions (primitives).

Using g for a normalized Gaussian function, a contracted function consisting of L Gaussians has the form,

$$\phi_\mu^{CGF}(r - R_A) = \sum_{p=1}^{L} d_{p\mu} g_p(\alpha_{p\mu}, r - R_A) \tag{81}$$

where $\alpha_{p\mu}$ are the Gaussian exponents, and $d_{p\mu}$ are the contraction coefficients. For example, in STO-3G, $L = 3$. In 3-21G, $L = 3$ for core, $L = 2$ and 1 for valence orbitals (split-valence).

The 6-31G* and 6-31G** basis sets [21] are formed by adding polarization functions to a 6-31G basis. In the 6-31G contraction, the inner shell functions (core) consist of six primitive Gaussians and the valence orbitals are split into a split-valence shell with L=3 and L=1. The d-type functions that are added to a 6-31G basis to form a 6-31G* basis are a single set of uncontracted d-type primitive Gaussians. The 6-31G** basis differs from the 6-31G* basis by the addition of one set of uncontracted p-type Gaussian primitives for hydrogen.

2.1.5 Energy Gradient and Geometry Optimization

Denoting all the variable geometrical parameters $q_1, \cdots, q_k$, which can be represented as a vector $\mathbf{q} = (q_1, \cdots, q_k)$, the total energy E will depend on $\mathbf{q}$, *i.e.*, $E = E(\mathbf{q})$. The energy optimization (minimization in most cases) is to find the optimum (minimum) value of E with respect to all possible allowed changes in $\mathbf{q}$.

Analytical gradients, or first derivatives of the energy with respect to q_i, can improve the efficiency of optimizing equilibrium structures and searching for transition states [22]. Second derivatives can be calculated by numerical differentiation of the analytical gradients to determine force constants and fundamental frequencies. Hartree-Fock energy gradients have been used extensively to optimize equilibrium geometries and to locate transition states. However, energy surfaces for some reactions cannot be described properly at the HF level. If correlation energy is included via CI or MP theory, energy gradients are more difficult to calculate, since the derivatives of molecular orbital coefficients must be computed by solving the coupled perturbed HF (CPHF) equations [23]. Alternatively, it is possible to describe the energy surfaces for many reactions with sufficient accuracy by using a GVB-PP approach. Since the GVB-PP calculations minimize the energy with respect to the MO and CI-coefficients simultaneously, the CPHF equations can be avoided.

Morokuma et al. [24] first presented the analytical evaluation of GVB energy gradients. Different formalisms and implementations of GVB and MCSCF gradients have been given in the literature [25]. In this study, we calculate the GVB analytical energy gradients based on the procedure proposed by Morokuma et al. [24].

With the total GVB electronic energy given by Eq. (69), the GVB gradients can be cast into a form similar to those for SCF and MCSCF.

$$\frac{\partial E_{GVB}}{\partial q_i} = \sum_{\mu\nu} P_{\mu\nu}(\frac{\partial h_{\mu\nu}}{\partial q_i}) + \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}(\frac{\partial}{\partial q_i})(\mu\nu|\lambda\sigma) - \sum_{\mu\nu} W_{\mu\nu}(\frac{\partial S_{\mu\nu}}{\partial q_i}) + (\frac{\partial E^{nr}}{\partial q_i}) \quad (85)$$

where $P_{\mu\nu}$ is the one-particle density matrix, $\Gamma_{\mu\nu\lambda\sigma}$ is the two-particle density matrix, $h_{\mu\nu}$ are the one-electron integrals, $(\mu\nu|\lambda\sigma)$ are the two-electron integrals, $S_{\mu\nu}$ are the overlap matrix, E^{nr} is the nuclear repulsion energy, and $W_{\mu\nu}$ are the energy-weighted density matrix.

The one-particle density matrix $\mathbf{P}$ is defined as,

$$\mathbf{P} = \mathbf{COC}^{+} \quad (86)$$

where $\mathbf{O}$ is a diagonal matrix ($O_{ij} = \delta_{ij}O_i$) representing the orbital occupancy,

$$O_i = \begin{cases} 2 & \text{if } i \text{ is a core orbital} \\ 2\sigma_i^2 & \text{if } i \text{ is a GVB-PP orbital with coefficient } \sigma_i \\ 1 & \text{if } i \text{ is an open-shell orbital} \end{cases} \quad (87)$$

The two-particle density is defined as,

$$\begin{aligned} \Gamma_{\mu\nu\lambda\sigma} = \; & [2D^{core}_{\mu\nu}D^{core}_{\lambda\sigma} - D^{core}_{\mu\lambda}D^{core}_{\nu\sigma}] \\ & +\sum_{i}^{N_{val}}[(a_{i0} + a_{0i})D^{core}_{\mu\nu}D^{i}_{\lambda\sigma} + (b_{i0} + b_{0i})D^{core}_{\mu\lambda}D^{i}_{\nu\sigma}] \\ & +\sum_{\mu\nu}(a_{ij}D^{i}_{\mu\nu}D^{j}_{\lambda\sigma} + b_{ij}D^{j}_{\mu\lambda}D^{i}_{\nu\sigma}) \end{aligned} \quad (88)$$

where the N_{val} is the number of valence orbitals with variable occupation number, including open-shell orbitals. The occupation number for core orbitals is 2. The density elements for the core, $D^{core}_{\mu\nu}$, and the valence orbitals, $D^{i}_{\mu\nu}$, are given by,

$$D^{core}_{\mu\nu} = \sum_{i}^{N_{core}} C_{\mu i} C_{\nu i}, \qquad D^{i}_{\mu\nu} = C_{\mu i} C_{\nu i} \tag{89}$$

The configurations for the GVB calculation involve the set of valence orbitals, or GVB-PP orbitals, while the core orbitals remain doubly occupied. This separation into core and valence orbitals simplifies the structure of the two-particle density matrix. A smaller $\Gamma^{val}_{\mu\nu\lambda\sigma}$ refers only to the valence space, while for core orbitals it can be obtained in terms of the one-particle density matrix.

Similarly to the closed-shell SCF, the GVB energy-weighted density is defined as,

$$W_{\mu\nu} = \sum_{kl} \epsilon_{kl} C_{\mu k} C_{\nu l} \tag{90}$$

where ϵ_{kl} are elements of the Lagrange matrix.

Our GVB geometry optimizations are based on the analytical gradients and the numerical second derivatives. Usually, the Davidon's optimally conditioned optimization method with gradients, *i.e.*, OC method [26], is used for the ground state geometries. The Powell's minimization of sum of squares of gradients, *i.e.*, VA method [27], is used for the transition state geometries.

2.2 Recent Developments in GVB Methodology

2.2.1 Pseudospectral GVB

The pseudospectral (PS) numerical method for SCF calculations has recently been extended for use in GVB calculations by Goddard et al. [28]. In the PS integration method which was developed by Friesner et al. [29], the two-electron integrals are calculated using both a Gaussian basis set and a numerical grid. In the conventional all-integral GVB calculations, a set of integrals $(\mu\nu|\sigma\eta)$ is evaluated, which are then used to calculate the coulomb and exchange integrals. In PS-GVB calculations, however, a numerical grid is used to directly evaluate the coulomb and exchange integrals. This involves evaluating the quantities,

$$A_{\sigma\eta}(r_g) = \int \frac{\chi_\sigma(2)\chi_\eta(2)}{|r_g - r_2|} d\tau_2 \tag{91}$$

directly for a set of grid points r_g. With Gaussian-type basis functions, $\chi_\sigma = r^l e^{-\alpha r^2} Y_{lm}(\theta,\phi)$, on various centers, the potentials $A_{\sigma\eta}(r_g)$ can be evaluated analytically for each grid point. The coulomb and exchange integrals can be calculated numerically from $A_{\sigma\eta}(r_g)$.

For K basis functions, this procedure scales as K^3 as compared to the conventional K^4 scaling of evaluating integrals over basis functions. For HF and GVB-PP wavefunctions, the PS numerical integration method, as an approximation to the conventional analytical integration, has been shown to give total electronic energies with a discrepancy of < 0.1 kcal/mol relative to the conventional all-integral methods [28].

The substantial reductions in CPU time and disk storage makes it possible to carry out GVB calculations on large systems.

2.2.2 GVB/R

In the GVB-PP wavefunction, the bonding orbitals ψ_{1i}^g, ψ_{2i}^g, which are coupled in the same pair, are nonorthogonal, while orthogonality is enforced between nonbonded spatial orbitals. A novel GVB wavefunction expression in terms of nonorthogonal orbitals has been proposed recently by Messmer et al. [13]. The GVB wavefunction is transformed to a multiconfiguration wavefunction expressed in terms of an orthogonal orbital basis. This simple transformation provides insight for extending the GVB description to the complete-active-space self-consistent-field (CASSCF) limit. An important effect is to allow the GVB orbitals to respond to the different spin couplings. The result is a new "orbital relaxed" GVB (GVB/R) wavefunction.

$$\Psi_{GVB/R} = \sum_i C_i \mathcal{A}[\Psi_{core}\psi_{1i}\psi_{2i}\cdots\psi_{Ni}\theta_i] \tag{92}$$

Each GVB/R configuration i is composed of different nonorthogonal orbitals $\psi_{ji}, j = 1, \cdots, N$, with a total of N nonorthogonal orbitals.

For example, the GVB/R wavefunction for C_2H_4 is,

$$\Psi_{C_2H_4}^{GVB/R} = C_1\mathcal{A}[\Psi_{core}\psi_{1\sigma}\psi_{2\sigma}\psi_{1\pi}\psi_{2\pi}\theta_1] + C_2\mathcal{A}[\Psi_{core}\psi'_{1\sigma}\psi'_{2\sigma}\psi'_{1\pi}\psi'_{2\pi}\theta_2] \tag{93}$$

where θ_1 is the perfect-pairing spin singlet coupling, $\theta_1 = (\alpha\beta - \beta\alpha)(\alpha\beta - \beta\alpha)$, and θ_2 can be considered as two C-C triplet pairs antiferromagnetically coupled to an overall singlet. The GVB orbitals are expanded in terms of the four orthogonal active orbital space. The first term in Eq. (93) is the normal GVB-PP wavefunction.

Nearly all of the correlation energy contained in a corresponding CASSCF expansion can be represented by a GVB/R wavefunction with far fewer variational parameters.

2.2.3 GVB-DIIS

In the early 1980's, Pulay [30] developed the direct inversion of the iterative subspace (DIIS) method for improving the convergence of HF wavefunctions. Instead of extrapolating orbitals from iteration to iteration, DIIS methods average the Fock matrices from different iterations to obtain a new trial Fock matrix for the orbitals of the next iteration. DIIS has been successfully used to closed-shell HF and restricted open-shell HF (ROHF) [31-32], and also applied for geometry optimization.

The DIIS procedure uses the errors (e) for different iterations (n) to find a "best" combination,

$$\sum_{n}^{n_{iter}} q_n e^n \approx 0 \tag{94}$$

where the errors for iterations n are grouped into a supervector e^n. The next iteration then uses the best Fock operator

$$F^{opt} = \sum_{n}^{n_{iter}} q_n F^n \tag{95}$$

where F^n is the general Fock operator at iteration n. Using this predicted Fock operator F^{opt}, rather than the newest Fock operator $F^{n_{iter}}$, leads to accelerated convergence.

The convergency condition is that the error vector ϵ_i^n be zero for all occupied orbitals. A major advantage to DIIS is that it can provide fast convergence generally only seen in second-order methods.

Muller et al. [18] generated a GVB-DIIS multishell Fock operator $\mathcal{F}^g$, and defined the error vector ϵ^n for iteration n. Once the calculation is within the DIIS radius of convergence, the key steps of a GVB-DIIS iteration are (1) the formation of the general Fock operator and error vector, (2) the determination of iteration coefficients from DIIS equations, and (3) to obtain new orbitals from the Fock operator.

The GVB-DIIS method, in general, allows fast and reliable convergence for a wide variety of wavefunctions having arbitrary numbers of core, open, and GVB orbitals. The reliable convergence requires a reliable procedure for the generation of the initial GVB orbitals to ensure that the starting point of a calculation is within the GVB-DIIS radius of convergence.

2.2.4 GVB-CI

The GVB wavefunction provides a clear orbital description of the electronic structure of molecules, but the emphasis of GVB is not on getting 100% of the correlation energy. To obtain more electron correlation, the GVB-CI wavefunction is usually constructed by including all single and double excitations which are within the space spanned by the GVB orbitals.

Carter and Goddard [33] developed the GVB correlation-consistent CI (GVB-CCCI) approach, in which the CI expansion systematically includes all correlations likely to change appreciably in the bond cleavage process. The truncation of this CI expansion is more rapid than the conventional singles and doubles CI approach, while giving more accurate results. Starting with the GVB wavefunction which allows the electrons in the bond that is breaking to occupy their own orbitals, GVB-CCCI includes full correlation of the electrons in the bond that is breaking and all single excitations from all valence orbitals. For example, GVB-CCCI calculations on ethylene lead to a C=C double bond energy of D_e = 174.1 kcal/mol, which is just 4.9 ± 2.5 kcal/mol weaker than the experimental value [33].

In order to mimic the full GVB wavefunction in which various spin couplings are included, a "restricted" CI (GVB-RCI) method [34-35] was developed to include all configurations with two electrons distributed among the two orbitals of each correlated pair (as in GVB-PP). In the natural orbital basis, the GVB-PP wavefunction describing N_{pair} GVB pairs is an expansion in $2^{N_{pair}}$ closed-shell determinants. The GVB-CI coefficients define the transformation from the overlapping GVB orbitals to the orthogonal natural orbitals. Using the structure of the multiconfigurational GVB-PP reference wavefunction makes the GVB-RCI expansion tractable for large systems.

Recently, Murphy et al. [34] developed a GVB-RCI contraction procedure, using more general MRCI expansion from a reference expressed in terms of GVB pairs.

The GVB-RCI contraction procedure consists of two parts: (1) the RCI expansion is defined by making excitations from the contracted, multiconfigurational GVB-PP expansion, rather than by making excitations of the individual $2^{N_{pair}}$ configurations contained in the GVB expansion; (2) the GVB-RCI CI coefficients are functions of the GVB-PP pair coefficients, *i.e.*, GVB CI-coefficients σ_{1i}, σ_{2i}. Furthermore, a fast two-index transformation of PS integrals gives the necessary integrals over orbitals for the CI calculation. In contrast, the conventional methodology requires an expensive (K^5 process) four-index transformation from the basis set to the orbital space. The combination of the novel GVB-RCI contraction and PS integration is shown to allow for accurate GVB-CI calculations on large molecules.

Chapter Three

Computational Developments and Programming Considerations

Both GVB orbitals (ψ_i) and GVB CI-coefficients (σ_i) in Eq. (67) are to be optimized self-consistently. As discussed in Section 1.2.5, GVB calculations cannot be used as a "blackbox", but the molecular orbitals must be correctly paired using some initial molecular orbitals. The initial perfect-pairing orbitals are essential to ensure a fast and proper energy convergence. The objectives of this study are (1) the automatic generation of initial GVB-PP orbitals for specific bonds or electron pairs, (2) the fast and reliable convergence of GVB calculations, and (3) creating an "object-oriented" GVB program.

3.1 Generation of Initial GVB Orbitals

3.1.1 CMO Initial Orbitals

One obvious choice for initial GVB orbitals is the use of canonical molecular orbitals which, for example, can be generated from extended Huckel or HF calculations. With CMOs it is possible, for a few cases, to select automatically orbital pairs which correspond to specific chemical bonds, using an identification matrix ($\mathcal{I}$). The identification matrix element $\mathcal{I}_{iA}^{CMO}$ for CMO ψ_i centered on atom A is given by,

$$\mathcal{I}_{iA}^{CMO} = \frac{\sum\limits_{\mu \in A} (C_{\mu i})^2}{\sum\limits_{\mu = 1} (C_{\mu i})^2} \tag{96}$$

where the $C_{\mu i}$ are the MO coefficients corresponding to ψ_i. According to Eq. (96), $0.00 \leq \mathcal{I}_{iA} \leq 1.00$ and $\sum\limits_{A} \mathcal{I}_{iA} = 1.00$. $\mathcal{I}_{iA}$ can be regarded as a partition function to represent the contribution from atom A to the molecular orbital ψ_i.

From the identification matrix a given ψ_i can be identified as corresponding to a specific bond or electron pair. For instance, at STO-3G the identification matrix elements for $C_2H_4(D_{2h})$ are

$$\mathcal{I}_{8C_1} = \mathcal{I}_{8C_2} = 0.50, \quad \mathcal{I}_{8H_1} = \mathcal{I}_{8H_2} = \mathcal{I}_{8H_3} = \mathcal{I}_{8H_4} = 0.00$$

$$\mathcal{I}_{9C_1} = \mathcal{I}_{9C_2} = 0.50, \quad \mathcal{I}_{9H_1} = \mathcal{I}_{9H_2} = \mathcal{I}_{9H_3} = \mathcal{I}_{9H_4} = 0.00$$

These results indicate that ψ_8 of ethylene is a (π) bonding orbital between the two carbon atoms, and ψ_9 is the corresponding (π) virtual (or antibonding) orbital.

Generally when using CMOs, perfect-pairing orbitals can only be automatically constructed for π orbitals (or some conjugated bonds). It is very difficult to choose the correct pairing for the σ orbitals from CMOs. On the other hand, LMOs are highly localized to atoms, and can generally be characterized as specific bonds between two atoms.

3.1.2 LMO Initial Orbitals

Since GVB orbitals are highly localized, a logical choice for initial GVB orbitals is the use of LMOs. However, if all the occupied orbitals are allowed to mix during the localization procedure, the resulting localized multiple bonds become equivalent. For example, the σ and π bonds in ethylene will transform into two equivalent bent bonds.

This problem can be easily overcome by classifying the orbitals into different localization groups which are defined by the MO selection vector $\mathcal{G}$. The default for $\mathcal{G}$ is two sets of orbitals (one for occupied, and another for virtual),

$$\mathcal{G}_i = \begin{cases} 1 & \text{if } i \text{ is an occupied orbital} \\ 2 & \text{if } i \text{ is a virtual orbital} \end{cases} \tag{97}$$

For example, the default for C_2H_4 at STO-3G is $\mathcal{G}$=(11111111222222). In this case, the first 8 occupied CMOs are localized as one group, and the last 6 virtual CMOs are localized as a second group. When symmetry is used to separate the σ and π bonds, the MO selection vector $\mathcal{G}$ becomes (11111113422222), where the π, π^* orbitals (ψ_8

and ψ_9) are not transformed.

In general, with the MO selection vector $\mathcal{G}$, it is possible to modify the default using a simple menu command [36] and thus to select any number of localization groups,

$$\mathcal{G}_i = \begin{cases} 2 + j & \text{if } i \text{ is a selected occupied orbital} \\ 2 + N_g + j & \text{if } i \text{ is a selected virtual orbital} \end{cases} \tag{98}$$

with $j = 1$ for the first group of selected orbitals and $j = N_g$ for last one, where N_g is the total number of selected groups. In this way, the CMOs can be divided into a total of $2 + 2 \times N_g$ groups for localization. By default the σ and π orbitals are transformed separately resulting in $\sigma\pi$ bonds as opposed to banana bonds.

The two bent bond description of ethylene is more suitable for describing the dissociation of the C=C bond of ethylene into two methylenes with two equivalent bent bonds. However, it is not suitable for describing rotation about the C=C double bond, where twisting of the π bond results in two nonequivalent bonds. Schultz and Messmer [37] have shown that the two bent bond description is energetically favoured over the $\sigma\pi$ description at the GVB level for a number of molecules including C_2H_4. However, our GVB-PP results demonstrate that the $\sigma\pi$ description of the C=C double bond is energetically favoured over the two bent bond description at the equilibrium geometry, which is in agreement with Carter and Goddard's results [38] for $F_2C{=}CF_2$.

3.1.3 Characterizing Bonding Orbitals

Initial orbitals corresponding to specific bonds can be generated from the LMOs, which are obtained through a unitary transformation of the CMOs using the Boys localization method [2(b)] for both occupied and virtual orbitals. The localization procedure is carried out in terms of the MO selection vector $\mathcal{G}$. The LMO coefficient matrix (**L**) and CMO coefficient matrix (**C**) are related by the transformation matrix (**T**) as $\mathbf{L} = \mathbf{CT}$.

Similar to Eq. (96), the LMO identification matrix ($\mathcal{I}^{LMO}$) is defined as,

$$\mathcal{I}_{iA}^{LMO} = \frac{\sum\limits_{\mu \in A} (L_{\mu i})^2}{\sum\limits_{\mu=1} (L_{\mu i})^2} \tag{99}$$

where $0.00 \le \mathcal{I}_{iA} \le 1.00$ and $\sum\limits_{A} \mathcal{I}_{iA} = 1.00$. Using the LMO identification matrix, one can construct a binary type matrix $\mathcal{B}$ with matrix elements defined as,

$$\begin{cases} \mathcal{B}_{iA} = 0, & \text{if } \mathcal{I}_{iA} \le \epsilon \\ \mathcal{B}_{iA} = 1, & \text{if } \mathcal{I}_{iA} > \epsilon \end{cases} \tag{100}$$

where ϵ is a threshold. The threshold is initially set at $\frac{1}{3}$, then decreased by increments (e.g. 0.03) until two bonding atoms are found or the minimum threshold (e.g. 0.10) is reached. Using this scheme, the matrix elements $\mathcal{B}_{iA}$ are determined for all doubly occupied and virtual orbitals. $\mathcal{B}_{iA} = \mathcal{B}_{iB} = 1$ indicates that LMO ψ_i^{LMO} corresponds to a bond between atoms A and B.

3.1.4 Perfect-Pairing

With the binary matrix $\mathcal{B}$, the selection of GVB perfect-pairing orbitals becomes straightforward and automatic. For the occupied ψ_i^{LMO}, which has been identified as a bond between atoms A and B, the corresponding ψ_j^{LMO} can be found using the following criterion:

$$\mathcal{B}_{iA} = \mathcal{B}_{jA}, \qquad \text{for all atoms } A \tag{101}$$

If more than one virtual orbital satisfies this criterion, the following additional condition is used,

$$|\sum_{\mu \in A} L_{\mu i} L_{\mu j} - \sum_{\mu \in B} L_{\mu i} L_{\mu j}| = \text{Maximum} \tag{102}$$

The result is the perfect-pairing orbitals (ψ_i^{LMO} and ψ_j^{LMO}) associated with the A-B bond. The initial GVB pair function ψ_{A-B} corresponding to the A-B bond can be constructed as in Eq. (67), using the pairing orbitals.

$$\psi_{A-B} = (\sigma_i \psi_i^{LMO} \psi_i^{LMO} - \sigma_j \psi_j^{LMO} \psi_j^{LMO})(\alpha\beta - \beta\alpha) \tag{103}$$

With this scheme, our GVB program [39] can (1) automatically generate LMO PP orbitals for all bonds (with frozen cores), (2) generate LMO PP orbitals for selected bonds, (3) read the orbitals which can be obtained from previous GVB calculations or created by other methods, (4) use an undistorted geometry to create GVB orbitals for a distorted geometry, (5) generate GVB orbitals for open shell cases, and (6) generate CMO PP orbitals automatically (for some cases) or from selected orbital pairs. For (1) and (2) the LMOs may, in our current version, be generated from Huckel MOs or from converged HF MOs.

3.2 Orbital Optimization and Convergence

Our GVB procedure is carried out as illustrated in Figure 2. Before STEP 0, the OSIPE (Open Structured Interfaceable Programming Environment) structure of the program [36] ensures that all "*objects*" required for the GVB calculation have been built. For example, the initial orbitals are built by a call to the function *getobj*('MATSQG_COEFF_CMO_AO_GUESS_GVB') (see Appendix A).

3.2.1 Orbital Optimization

Using the basis expansion $\psi_i = \sum_{\mu} C_{\mu i}\phi_\mu$, the energy expression of Eq. (75) can be rewritten as,

$$E = \sum_{\mu}\sum_{\nu}\sum_{i} C_{\mu i} * C_{\nu i} E_{i0} \tag{104}$$

where $E_{i0} = f_i h_{\mu\nu} + \langle\phi_\mu|F_i|\phi_\nu\rangle$.

The orbital coefficients $C_{\mu i}$ are redefined by Eq. (77), which can be expressed as an orbital rotation,

$$\begin{cases} \psi_i' = \psi_i \cos\theta + \psi_j \sin\theta \\ \psi_j' = \psi_j \cos\theta - \psi_i \sin\theta \end{cases} \tag{105}$$

with

$$\cos\theta = \left(\frac{1}{1+(\xi\lambda_{ij})^2}\right)^{\frac{1}{2}} \tag{106}$$

STEP 0 - Optimize CI-coefficients σ_i using the initial orbitals.

STEP 1 - Construct the Fock matrices over AOs, $\langle\psi_i|F_i|\psi_j\rangle$ and
$A_{ij} = \langle\psi_i|(F_i - F_j)|\psi_j\rangle$, where

$$F_i = f_i h + \sum_{j=1}^{N_{orr}} (a_{ij}J_j + b_{ij}K_j)$$

STEP 2 - Pairwise optimization of the GVB-PP orbitals to get the orbital correction factors λ_{ij}.

$$\lambda_{ij} = \frac{\langle\psi_i|F_j - F_i|\psi_j\rangle}{\langle\psi_i|(F_j - F_i)|\psi_i\rangle - \langle\psi_j|(F_j - F_i)|\psi_j\rangle + B_{ij}} \text{ where}$$
$$B_{ij} = 2(a_{ii} + a_{jj} - 2a_{ij})K_{ij} + (b_{ii} + b_{jj} - 2b_{ij})(J_{ij} + K_{ij})$$

STEP 3 - Construct new orbitals.

$$\begin{cases} \psi_i' = \psi_i \cos\theta + \psi_j \sin\theta \\ \psi_j' = \psi_j \cos\theta - \psi_i \sin\theta \end{cases}$$
where $\cos\theta = 1/(1 + (\xi\lambda_{ij})^2)^{\frac{1}{2}}$

STEP 4 - Minimize E with respect to ξ with E_{i0} fixed, go to STEP 5. If no minimum is found or if the orbital corrections are large, go to STEP 2.

STEP 5 - Extrapolate/Interpolate $\mathbf{C}$.

STEP 6 - Recalculate coulomb and exchange integrals (J, K).

STEP 7 - Reoptimize CI-coefficients **only if** ρ_c **is small enough** (e.g. ≤ 0.01).

STEP 8 - Calculate the energy.

$$E = 2\sum_{i=1}^{N_{orr}} f_i h_{ii} + \sum_{i,j=1}^{N_{orr}} (a_{ij}J_{ij} + b_{ij}K_{ij})$$

STEP 9 - Calculate the density matrix and check for convergence $(\rho_{orb}, \rho_c, \rho_d)$.

- GO TO STEP 1

Figure 2: The main steps in the GVB iterative process

The pairwise orbital correction factor Λ_{ij} is calculated using Eq. (78), and a scaling factor (ξ) is introduced in order to accelerate convergence. The scaling factor ξ is optimized to minimize E with E_{i0} fixed (*c.f.* STEP 4 in Figure 2). The default ξ is set to 1.00, and generally the optimized value is about 0.8. The rotation procedure is carried out simultaneously and with a single scaling factor for all orbitals. After the rotation the orbitals are orthogonalized by the Schmidt procedure.

3.2.2 GVB CI-Coefficients and Extrapolation

The GVB CI-coefficients in Eq. (103) are optimized for the initial orbitals, and are only reoptimized when the change in the orbital coefficients is small enough. Pople's 3/4-point [5] and Dewar's [40] extrapolation methods, which have been well studied for HF-SCF in MUNGAUSS [41], are available to extrapolate the orbital coefficients $C_{\mu i}$. In the default procedure the 3/4-point method is used until the orbital correction is small enough, at which point we switch to Dewar's extrapolation method.

3.2.3 Convergence of GVB calculations

Three different convergence criteria are used to determine the optimization accuracy and ensure proper convergence of the GVB energy.

1) The orbital correction convergence criteria, ρ_{orb},

$$\rho_{orb} = \left(\frac{\sum_{i,j=1}^{N_{basis}} A_{ij}^2}{N_{basis} \times (N_{basis}+1)/2 - N_{basis}} \right)^{\frac{1}{2}} \qquad (107)$$

2) The orbital coefficients convergence criteria, ρ_c,

$$\rho_c = \left(\frac{\sum_{\mu=1}^{N_{basis}} \sum_{i=1}^{N_{occ}} (C'_{\mu i} - C_{\mu i})^2}{N_{basis} \times N_{occ}}\right)^{\frac{1}{2}} \tag{108}$$

3) The density matrix convergence criteria, ρ_d,

$$\rho_d = \frac{\left(\sum_{i,j=1}^{N_{basis}} (\mathbf{P}'_{ij} - \mathbf{P}_{ij})^2\right)^{\frac{1}{2}}}{N_{basis}} \tag{109}$$

where $\mathbf{P}$ is the one-particle density matrix as defined in Eq. (86). The density matrix convergence criterion is important since it incorporates convergence in both the orbital coefficients and the orbital occupancy or GVB CI-coefficients.

In order to demonstrate our GVB convergence, some systems were selected to represent the different GVB models, from GVB(1/2) (1 pair) to GVB(10/20) (10 pairs). GVB calculations[3] were performed with various standard Pople basis sets [51] (STO-3G, 3-21G, 6-31G*, 6-31G**, 6-31++G**, and 6-311G**), using HF LMOs as the initial orbitals. GVB/6-31G** calculations were also performed using extended Huckel LMOs as the initial orbitals. All the calculations were carried out using experimental geometries [42] except for glycine, for which the HF/STO-3G geometry is used.

The default accuracy for all convergence criteria (ρ_{orb}, ρ_c and ρ_d) is 10^{-5}. In Ref. [18], Goddard et al. used the SQCDF ($\leq 10^{-9}$) as the convergence criteria, which is

[3]From now on, our GVB calculations in this study are based on the LMO initial orbitals and the automatic generation procedure, without specification.

related to our orbital coefficients convergence accuracy ρ_c.

$$\text{SQCDF} = \rho_c^2 \times N_{basis} \times N_{occ} \tag{110}$$

For moderate size molecules, $N_{basis} \times N_{occ} \approx 10^3$, so that $\rho_c \leq 10^{-5}$ is equivalent to SQCDF $\leq 10^{-7}$. As pointed out the density convergence accuracy, ρ_d, is a better criteria, since it incorporates all GVB variational parameters.

The total energies are summarized in Table 1, and the numbers of iterations for converging GVB calculations are reported in Table 2. As in GVB-DIIS [18], reliable convergence requires a reliable initial guess to ensure that the starting point is within the radius of convergence. Our GVB procedure gives a similar convergence, once a reliable guess is obtained. The GVB(2/4)/6-31G** calculation of C_2H_4 converges at 9 iterations, for instance, compared to 10, 15 and 26 iterations for the corresponding GVB-DIIS, GVB2P5 and G90 respectively [18]. For glycine, our GVB(10/20)/6-31G** converges at 16 iterations, compared to 20, 90 and 75 iterations for the corresponding GVB-DIIS, GVB2P5 and G90 respectively. The GVB/6-31G** calculations using the extended Huckel LMO guess generally converge as well as for the HF LMO guess. Overall, our GVB converges within 10-20 iterations (Table 2), demonstrating that the GVB method as described above provides excellent convergence for a wide variety of wavefunctions based on reliable initial orbitals which are generated automatically. Further improvements in convergence should be possible by incorporating the DIIS method in our GVB.

Table 1: Total RHF and GVB energies (hartrees).

Molecule	Method*	STO-3G	3-21G	6-31G*	6-31G**	6-31++G**	6-311G**
F_2	RHF	-195.96743	-197.64415	-198.67382	-198.67382	-198.68095	-198.72708
	GVB(1/2)	-196.04519	-197.71926	-198.74833	-198.74833	-198.75575	-198.80378
H_2O	RHF	-74.96305	-75.58542	-76.01050	-76.02312	-76.03073	-76.04639
	GVB(2/4)	-75.00315	-75.62810	-76.05287	-76.06411	-76.07155	-76.08826
HNC	RHF	-91.64410	-92.33941	-92.85473	-92.85920	-92.86492	-92.88159
	GVB(2/4)	-91.68875	-92.38253	-92.89744	-92.90157	-92.90765	-92.92486
	GVB(3/6)	-91.69514	-92.39153	-92.90641	-92.91055	-92.91670	-92.93454
	GVB(4/8)	-91.70762	-92.40630	-92.92127	-92.92502	-92.93112	-92.94888
CH_3	RHF	-38.79604	-39.22481	-39.44614	-39.45288	-39.50648	-39.48053
	GVB(3/6)	-39.11445	-39.38276	-39.59914	-39.60375	-39.60639	-39.61251
HCCH	RHF	-75.85287	-76.39528	-76.81710	-76.82116	-76.82682	-76.84025
	GVB(2/4)	-75.91726	-76.44554	-76.86478	-76.86888	-76.87344	-76.88733
	GVB(3/6)	-75.92289	-76.45256	-76.87197	-76.87606	-76.88066	-76.89483
	GVB(5/10)	-75.94894	-76.48010	-76.89980	-76.90326	-76.90787	-76.92230
H_2CCH_2	RHF	-77.07207	-77.59979	-78.03078	-78.03792	-78.04249	-78.05385
	GVB(1/2)	-77.11659	-77.63054	-78.05989	-78.06704	-78.07042	-78.08214
	GVB(2/4)	-77.12496	-77.64065	-78.07017	-78.07743	-78.08070	-78.09275
	GVB(6/12)	-77.18165	-77.69900	-78.13020	-78.13620	-78.13958	-78.15188
Glycine	RHF	-279.10589	-281.22876	-282.80827	-282.82461	-282.83455	-282.89268
	GVB(10/20)	-279.29109	-281.40577	-282.99178	-283.00605	-283.01653	-283.07487

* F_2: GVB(1/2) for the F-F bond. H_2O: GVB(2/4) for two O-H bonds. HNC: GVB(2/4) for two N-C π bonds, GVB(3/6) adding a N-C σ bond, GVB(4/8) adding a H-N bond. CH_3: GVB(3/6) for three C-H bonds. HCCH: GVB(2/4) for two C-C π bonds, GVB(3/6) adding a C-C σ bond, GVB(5/10) adding two C-H bonds. H_2CCH_2: GVB(1/2) for the C-C π-bond, GVB(2/4) adding a C-C σ bond, GVB(6/12) adding four C-H bonds. Glycine: GVB(10/20) for all 10 bonds. The experimental geometries are used [42], except for glycine (HF/STO-3G geometry).

Table 2: Numbers of iterations for converging GVB calculations.

Molecule	Method*	STO-3G	3-21G	6-31G*	6-31G**	6-31G** HMO[a]	6-31++G**	6-311G**	reference
F_2	GVB(1/2)	10	16	9	9	8	9	15	
H_2O	GVB(2/4)	8	14	10	10	15	11	14	$(14)^b$ $(24)^c(89)^d$
HNC	GVB(2/4)	8	9	8	9	11	9	12	
	GVB(3/6)	8	12	8	9	12	9	14	
	GVB(4/8)	8	11	8	9	13	9	20	
CH_3	GVB(3/6)	11	12	12	11	14	13	16	$(14)^b$ $(17)^c(28)^d$
HCCH	GVB(2/4)	6	8	7	7	9	8	8	
	GVB(3/6)	7	8	8	7		7	10	
	GVB(5/10)	6	9	8	8		8	9	
H_2CCH_2	GVB(1/2)	4	11	7	7	10	8	7	
	GVB(2/4)	5	13	9	9	12	9	8	$(10)^b$ $(15)^c(26)^d$
	GVB(6/12)	4	13	9	9	9	10	10	
Glycine	GVB(10/20)	12	20	14	16		19	18	$(20)^b$ $(90)^c(75)^d$

* see notes for Table 1.

a. The initial orbitals are from extended Huckel LMOs, whereas the initial orbitals are from HF LMOs for all other cases.

b. GVB-DIIS [18]; *b*. GVB2P5 [18]; and *d*. G90 [18].

3.3 Program Structure

The program for GVB calculations is a part of MUNGAUSS [41], which has been developed using the OSIPE tools [36]. All *objects*, required for GVB calculations, are handled by the object management tools[4] provided by OSIPE (see Appendix A). The GVB calculations have been tested for a wide variety of wavefunctions having arbitrary numbers of core, open, and GVB-PP orbitals, on VAX/VMS, SGI/UNIX, and DEC-Alpha/UNIX computers.

The following is a brief description of the GVB program, which can be used as a user's and programmer's reference.

3.3.1 GVB Program (GVBCLC)

The GVB computation is carried out in its algorithmic routine GVBCL8. The main steps in GVBCL8 are given in Figure 2, as discussed in Section 3.2. The allocation routine[5] GVBCLC gets the address of *objects* for GVBCL8, and is shown on page 56.

For example, the *scalar SCF_SCI_NUM_OF_CONFIG* is built and assigned to *NCONF* by calling the function, *i.e.*,

```
call getsci ('SCF_SCI_NUM_OF_CONFIG', NCONF)
```

[4] Such as **getobj**, **bldobj**, **putobj** .... Other OSIPE facilities are also used.

[5] In terms of OSIPE basic principles [36], an *object* can only be created by its allocation-routine, which creates nothing else. The OSIPE structure for the allocation routine and its algorithmic routine is illustrated in Appendix A.

```
allocation-routine:
      SUBROUTINE GVBCLC
***********************************************************************
* Includes:
      include 'osipe_stack'
*
* stack addresses:
      ......
* Local scalars:
      ......
*
* Begin:
      lbuild = .false.
* Scalars:
      call getsci ('SCF_SCI_NUM_OF_CONFIG', NCONF)
      call getsci ('SCF_SCI_NUM_OF_FOCK_MATRICES', NFOCK)
      ......
* Input objects:
      ixICONF=getobj ('CONFIGURATION_LIST')
      if(lbuild)ixICONF=bldobj ('CONFIGURATION_LIST')
      ixCMOG=getobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
      if(lbuild)CMOG=bldobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
      ......
* Output objects:
      ixCMOA= putobj ('MATSQG_COEFF_CMO_AO_GVB',
     .                NBASIS*NBASIS, 'REAL', L8)
      ......
* Gather addresses:
      ICONF=objadd(ixICONF)
      CMOG=objadd(ixCMOG)
      CMOA=objadd(ixCMOA)
      ......
* Computation
      CALL GVBCL8 (stack(CMOA), stack(CMOB), ... , NFOCK)
*
      if(pgmchk)call CHKBND ('GVBCL8')
      if(locdbg) then
        call prtobj ('MATSQG_COEFF_CMO_AO_GVB')
      end if
* End of routine GVBCLC
      RETURN
      END
```

The *object* '*MATSQG_COEFF_CMO_AO_GUESS_GVB*' is built as a result of the call to *getobj*. If it is not an existing *object*, it will be built by its allocation routine GVBMOG.

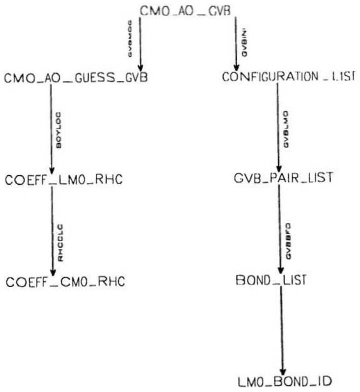
```
CMOG=getobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
if(lbuild)CMOG=bldobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
```

The GVB calculation can be executed by the OSIPE command *output object*,[6] *i.e.*,

```
OUTPUT OBJECT=MATSQG_COEFF_CMO_AO_GVB END
```

Since the *CMO_AO_GVB* is built from other *objects*, such as *CMO_GUESS_GVB* and *CONFIGURATION_LIST*, a dynamic structure defined by the "created-by" relationship between *objects* is followed by OSIPE, as shown in following flow chart.

[6]The OSIPE facility [36] for printing *object* in stack, such as *prtobj* and *show stack*, has proven powerful tool for interpretation the results of a calculation.



Tree structure defined by the "created-by" routine relationship between *objects*. This tree represents the dynamic route followed by OSIPE.

```
OUTPUT OBJECT=MATSQG_COEFF_CMO_AO_GVB END
```

3.3.2 GVB menu (SETGVB)

The GVB calculation can also be executed by,

GVB RUN END

here *GVB* acts as a command, which can be used to change the GVB options (SETGVB). The default *scalars* and the GVB commands are listed in Tables 3 and 4. For example,

GVB acc=*1.0D-7 END*

can be used to reset the convergence accuracy to be 1.0×10^{-7}.

The manipulation of some GVB commands, which are very useful in carrying out the GVB calculations, is described in the following.

- Specific bond. For full valence bond GVB-PP calculations, the default is to correlate all bonds into active space automatically. Specific bonds can be identified by using *sbond=(SI1 SI2)* for the σ bond and *pbond=(SI1 SI2)* for the π bond, in disilene Si_2H_4, for example.

- Lone pair. The specific lone pairs on some atoms can be added if required. For example, in carbon monoxide, the lone pairs at carbon and oxygen can be added by using *llp=(C1)* and *llp=(O2)*. In this case, including the lone pairs into the active space improves the GVB wavefunction and results in the dipole moment for carbon monoxide which is in excellent agreement with experiment

(Section 4.6). The three different lone-pairs of fluorine in hydrogen fluoride can be added by using *1lp=(F1)*, *2lp=(F1)* and *3lp=(F1)*.

- Core. The default active space is for all valence electrons. If necessary, the core is also included. For example, in LiH, using *fc=0* (no frozen cores) to include the Li $1s$ core into active space.

- Three-center bond. In some cases, two-center bonds cannot be correctly identified due to conjugation. In those cases, for example, *tcbond=(C1 C2 C3)* can be used to identify the three-center bond in the allyl radical C_3H_5.

- Perfect-pairing. Perfect-pairing can be generated automatically or selected using *pair= (8 9)* where 8 and 9 refer to the bonding orbital and antibonding orbital, for example.

- Choice of MO. *mo=lmo* or *mo=cmo* can be used to set the type of molecular orbitals.

- Scaling factor. The default scaling factor is 1.0. Using *sca=0.60*, it is changed to 0.60, for example. In general case, *gvbr=incr* is to set the scaling factor in the iterative process [*c.f.*, Eq. (106)] by a increment procedure. While *gvbr=no* is to keep the updated value of the scaling factor, and *gvbr=reset* to reset the scaling factor to the initial value.

- Using undistorted geometry to create the initial GVB-PP orbitals for the distorted geometry with *ppt=guess*.

The data input (menu) of GVB(2/4)/6-31G** calculations for σ and π bonds in disilene Si_2H_4 (it is run on VAX/VMS).

```
set
    run name = "TESTgvb" END
    dir = "DISK$USER4:[SCRATCH.qcgroup]"
END

MOLECULE
       UNITS=ANGSTROM
       CHARGE=0
       MULTIPLICITY=1
ADD atom=SI1                                                      END
ADD atom=SI2    FROM=( SI1 )          BOND=CC                     END
ADD atom=H3     FROM=( SI1 SI2 )      BOND=CH ANGLE=ALPHA         END
ADD atom=H4     FROM=( SI1 SI2 H3 )   BOND=CH ANGLE=ALPHA TORSION=DA1 END
ADD atom=H5     FROM=( SI2 SI1 H3 )   BOND=CH ANGLE=ALPHA TORSION=DA2 END
ADD atom=H6     FROM=( SI2 SI1 H5 )   BOND=CH ANGLE=ALPHA TORSION=DA1 END
DEFINE
CC       =     2.240
CH       =     1.498
ALPHA    =   116.33
DA1      =   131.50
DA2      =    48.50
END
end !molecule

BASIS name=6-31G** end

GVB
 sbond=( SI1 SI2 )
 pbond=( SI1 SI2 )
end ! gvb commands

OUTPUT OBJECT=MATSQG_COEFF_CMO_AO_GVB END

STOP
```

Table 3: The default scalars and the corresponding GVB sub-commands.

Scalar:object name	default	command*
GVB_SCC_GUESS_MO	'LMO'	MOguess=READ
GVB_SCC_GUESS_TYPE	'BOND'	LMOGtype=PAIR
GVB_SCC_GUESS_SSP	'NO'	SSP=YES
GVB_SCR_BOND_THRESHOLD	0.20D0	BTHreshold=0.13
GVB_SCR_PBOND_THRESHOLD	0.45D0	PBThreshold=0.35
GVB_SCR_SCA_FACTOR	0.0D0	SCAle=0.70
GVB_SCR_CI_COEFF	0.7071070D00	CICoeff=0.0
GVB_SCI_NUM_CORES	1	FCOre=2
GVB_SCC_COEF_PRT	'NOPPT'	PPT=PUNC
GVB_SCI_MAX_ITERATION	50	GVBIteration=100
GVB_SCI_MAX_PAIRWISE	0	MXPairwise=2
GVB_SCC_EXT_METHOD	'3/4 POINT'	GVBExtrapolation=DEWAR
GVB_SCC_ROT_METHOD	'RESET'	GVBRotation=NO
GVB_SCB_ROTATION_METHOD	.false.	ROTMIX=.true.
GVB_SCR_DEN_ACCURACY	1.0D-5	ACCuracy=1.0D-6
GVB_SCR_ORB_ACCURACY	1.0D-5	ORBaccuracy=1.0D-6

* The values given here are just some examples.

Table 4: The objects and the corresponding GVB sub-commands.

Object name	command*
GVB_PAIR_LIST	PAIR=(1 2)
GVB_PAIR_LIST_SBOND	SBond=(1 2) or (C1 H2)
GVB_PAIR_LIST_PBOND	PBond=(1 2) or (C1 N2)
GVB_PAIR_LIST_TCBOND	TCbond=(1 2 3) or (C1 C2 C3)
GVB_PAIR_LIST_TBOND	PPbond=(1 2) or TBond=(1 2) or (C1 C2)
GVB_OPEN_LIST	OP=(9)

* The values given here are just some examples.

Chapter Four

GVB Results of Some Simple Molecules

In order to examine the performance of our GVB approach, a number of calculations have been carried out for some simple molecules, *e.g.*, CH_2, H_2O, C_2H_4, and C_2H_6. The GVB treatment for the allyl radical C_3H_5 is also presented and discussed. The results of GVB correlation energies and molecular properties, such as the O-H σ bond dissociation energy in H_2O, π bond energy in C_2H_4, rotational barrier in C_2H_6, and dipole moments of CO and H_2CO, are discussed and compared with experimental and other available GVB data.

4.1 CH_2

GVB(3/6)/3-21G calculations of singlet CH_2 were performed by using our GVB program and Gaussian 92 (G92). Generally, it is very difficult to select the GVB orbital pairing by using G92.[6] The G92 data input for CH_2 is given by the G92/DFT

[6] The special command *guess=(local, lowsym, alter)* is required to alter the localized molecular orbitals into the "**correct but unclear**" order when using G92.

programmer's reference book [13]. In our approach, the GVB(3/6) calculation, which includes two C-H σ bonds and one carbon lone pair, is carried out by using the following OSIPE commands.

```
GVB  sbond=( C H1 ) sbond=( C H2 ) 1lonepair=( C ) END
OUTPUT OBJECT=MATSQG_COEFF_CMO_AO_GVB END
```

The GVB(3/6) wavefunction is

$$\Psi_{CH_2}^{GVB} = \mathcal{A}[\Psi_{core}\psi_{C-H1}\psi_{C-H2}\psi_{C-lp}] \tag{111}$$

where ψ_{C-H}, ψ_{C-lp} are GVB pair functions corresponding to the C-H bonds and the carbon lone pair respectively [*c.f.* Eq. (103)].

The comparison in Table 5 shows that our GVB procedure gives faster convergence than G92, while the GVB energies and the equilibrium geometries with the energy gradient optimization are very close to each other. For example, the optimized C-H bond length is 1.1370 Å (this study) versus 1.1368 Å (G92), and the H-C-H bond angle is 101.23^0 (this study) versus 101.18^0 (G92). [7]

[7] The small discrepancies are mainly due to different integral and convergence accuracies in the different programs.

Table 5: Comparison of our GVB(3/6)/3-21G results of singlet CH_2 with G92

	this study	G92
Energy[a] (a.u.)	-38.691302	-38.691310
N_{iter}	10	24
Equilibrium geometry		
R_{C-C} (Å)	1.1370	1.1368
Θ_{H-C-H}	101.23°	101.18°
Energy after optimization (a.u.)		
	-38.701886	-38.701886

[a] this energy is for the initial geometry, $R_{C-C} = 1.08$ Å, $\Theta_{H-C-H} = 120.°$.

4.2 H_2O

In order to make comparisons with previous ab initio calculations [13]. GVB calculations on H_2O were performed by using Dunning's double-zeta basis [44] and Schaefer's CISD equilibrium geometry [45]. The GVB(2/4) wavefunction of H_2O was constructed to correlate the two O-H bonds, while the O_{1s} and the two oxygen lone pairs were treated as closed shell cores. The four H_2O orbitals GVB(4/8) consist of two equivalent O-H bonds and two equivalent nonbonding oxygen lone pairs. The GVB(4/8) wavefunction is

$$\Psi^{GVB}_{H_2O} = \mathcal{A}[\Psi_{core}\psi_{O-H1}\psi_{O-H2}\psi_{O-lp1}\psi_{O-lp2}] \tag{112}$$

where ψ_{O-H}, ψ_{O-lp} are GVB pair functions corresponding to the O-H bonds and the oxygen lone pairs respectively. The results for the GVB calculations on H_2O are summarized in Table 6. For GVB(4/8), the total correlation energy (ΔE^{corr}_{total}) can be written as,

$$\Delta E^{corr}_{total} = 2\Delta E_{O-H} + 2\Delta E_{O-lp} \tag{113}$$

where ΔE_{O-H} represents the correlation energy contributed from a O-H bond pair, and ΔE_{O-lp} represents the correlation energy contributed from an oxygen lone pair. Our GVB(4/8) result for ΔE^{corr}_{total} is -0.0632 a.u., with ΔE_{O-H}=-0.022 a.u. and ΔE_{O-lp}=-0.009 a.u., which is in good agreement with Goddard's values of -0.0642 a.u., -0.021 a.u. and -0.011 a.u. respectively [12(b)].

The GVB(2/4)/6-31G** potential curve for $H_2O \rightarrow HO\cdot + H\cdot$ dissociation is shown in Figure 3. The GVB(2/4) calculation was performed with the optimized minimum at 6-31G**, then keeping C_s symmetry to dissociate the O-H1 bond, with the O-H2 bond length and H1-O-H2 bond angle fixed. The resulting bond dissociation energy (BDE, D_e) is 100.5 kcal/mol, compared to the available experimental result of D_e=126 kcal/mol [46]. The reported HF limit is 89 kcal/mol at 6-31G**, and 116 kcal/mol at MP4/6-31G** [5]. The GVB(4/8) result of 101.3 kcal/mol shows that the oxygen lone pairs have little effect on the O-H BDE. The bond-breaking process can be described reasonably well even with such a simple wavefunction consisting of independent GVB bond pairs.

Table 6: GVB energies of H_2O.

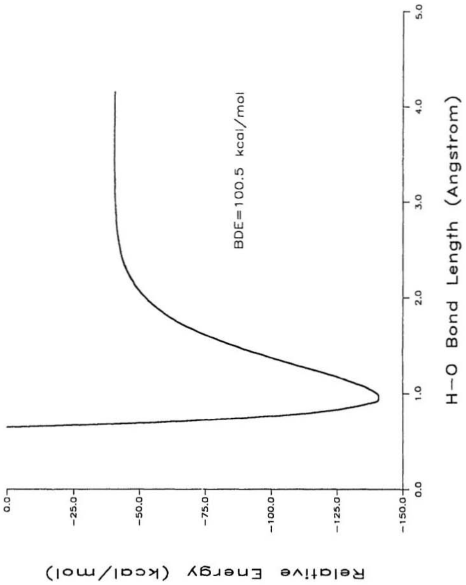
Method[a]	Energy (hartrees)		Correlation energy	
	this work	reference[b]	this work	reference[c]
HF	-76.00984	-76.00984	-	-
GVB(1/2)	-76.03216	-	-0.02232	-0.0209
GVB(2/4)	-76.05415	-76.05415	-0.04421	-0.0418
GVB(4/8)	-76.07299	-	-0.06315	-0.0642

[a] GVB(1/2): for a O-H1 bond. GVB(2/4): adding a O-H2 bond. GVB(4/8): adding two oxygen lone pairs.

[b] Messmer et al. SOPP approach with the same basis set and geometry, see Ref. [13].

[c] Goddard et al. results with a different basis set and geometry, see Ref. [12(b)].

Figure 3: GVB(2/4)/6-31G** potential energy curve for $H_2O \rightarrow HO + H$.



BDE=100.5 kcal/mol
H—O Bond Length (Angstrom)
Relative Energy (kcal/mol)
0.0
1.0
2.0
3.0
4.0
5.0
0.0
-25.0
-50.0
-75.0
-100.0
-125.0
-150.0

4.3 C_2H_4

For C_2H_4, GVB calculations were carried out with the 6-31G** basis set and the D_{2h} experimental geometry [28] with R_{CC} = 1.338 Å, R_{CH} = 1.085 Å, Θ_{HCH} 117.8°. For the 90° twisted conformation all the geometrical parameters were kept constant. The GVB(2/4) wavefunction was constructed from the C=C double bond (σ, π) pairs, and has the form:

$$\Psi^{GVB}_{C_2H_4} = \mathcal{A}[\Psi_{core}\psi_{C1-C2(\sigma)}\psi_{C1-C2(\pi)}] \tag{114}$$

and the total correlation energy can be expressed as

$$\Delta E_{total} = \Delta E_{C1-C2(\sigma)} + \Delta E_{C1-C2(\pi)} \tag{115}$$

The GVB results for C_2H_4 are summarized in Table 7. Our GVB(2/4) correlation energy is slightly lower than recent PS-GVB(2/4) results [28] for both the planar (D_{2h}) and the twisted (D_{2d}) geometry, using the same geometry but a different basis set. However, both of the GVB(2/4) calculations give the same result (74.2 kcal/mol) for the double bond singlet rotational barrier ($D_r = E_{D_{2d}} - E_{D_{2h}}$). The GVB(6/12) calculation, which is comprised of two C1-C2 bond pairs and four C-H bond pairs, gives a singlet rotational barrier of 76.9 kcal/mol, which is higher by 2.7 kcal/mol than the GVB(2/4) result with the same geometry. It is noted that in a recent paper of Goddard et al. [18], the one-pair GVB description of the π bond in C_2H_4 predicts a singlet rotational barrier of 199 kcal/mol.[8] This result is presumably due to a

[8]From Ref. (18), the GVB energy of the twisted structure is higher than that of the planar structure by 0.317619 a.u.

misprint or due to an incorrect convergence and not a failure of the one-pair GVB.

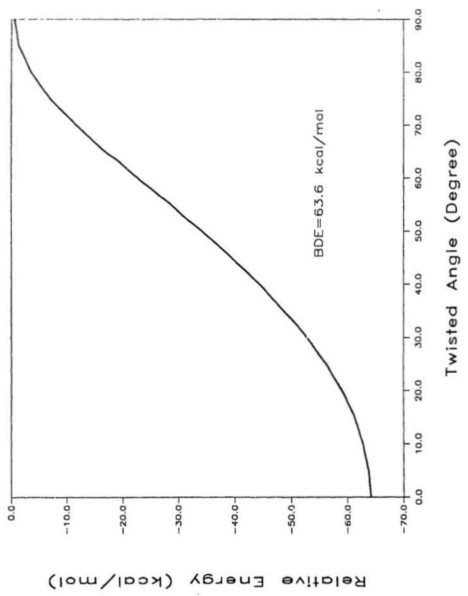
Since the optimized C-C bond in the twisted structure is 1.49 Å, longer than in the planar structure (1.35 Å), the rotational barrier using optimized geometries is much lower than that obtained with a constant C-C bond length. Figure 4 illustrates the GVB(2/4)/6-31G** potential curve for rotation about the C=C double bond of C_2H_4, with the optimized geometries. The 63.6 kcal/mol barrier to rotation is in excellent agreement with the previous CI result of 61 kcal/mol [47], the GVB result of 65.6 kcal/mol [48] and the experimental result of 64 kcal/mol [49].

Table 7: Total energies (hartrees) and rotational barriers (kcal/mol) for C_2H_4.

	HF	GVB(2/4)			GVB(6/12)	
	total	total	correlation energy		total	correlation
	energy	energy	this work	reference[a]	energy	energy
Planar(D_{2h})	-78.03797	-78.07741	-0.03944[b]	-0.03852	-78.13618	-0.09921
Twisted(D_{2d})	-77.86080	-77.95908	-0.09828	0.09273	-78.01352	0.15272
Rotational Barrier[c]	111.2	74.2	63.6[d]	74.2	76.9	

[a] Goddard's PS-GVB result from Ref. [28].
[b] Murphy's GVB-SOPP(2/4) gives a correlation energy of 0.03513 a.u. using the two bent bond description [13].
[c] The experimental rotational barrier of ethylene is 64 kcal/mol [49].
[d] Using optimized geometries.

Figure 4: GVB(2/4)/6-31G** potential energy curve for rotation around the C–C double bond in C_2H_4.


BDE=63.6 kcal/mol
Twisted Angle (Degree)
Relative Energy (kcal/mol)
0.0 10.0 20.0 30.0 40.0 50.0 60.0 70.0 80.0 90.0
0.0 -10.0 -20.0 -30.0 -40.0 -50.0 -60.0 -70.0

4.4 C_2H_6

GVB(7/14) calculations on C_2H_6 were carried out using the 6-31G** basis set and the experimental geometry [50]. The GVB(7/14) wavefunction, in which seven bond pairs are included (1 C-C and 6 equivalent C-H bonds), was constructed as

$$\Psi^{GVB}_{C_2H_6} = \mathcal{A}[\Psi_{core}\psi_{C1-C2}\psi_{C1-H1}\psi_{C1-H2}\psi_{C1-H3}\psi_{C2-H4}\psi_{C2-H5}\psi_{C2-H6}] \qquad (116)$$

and the total correlation energy can be expressed as

$$\Delta E_{total} = 6\Delta E_{C-H} + \Delta E_{C-C} \qquad (117)$$

The same geometrical parameters were used for both the staggered and eclipsed conformations. Table 8 summarizes the GVB results. The rotational barrier of 3.3 kcal/mol is in good agreement with the experimental value of 2.93 kcal/mol [51] and the previous GVB result of 3.1 kcal/mol [12(b)], with the staggered conformation being more stable in every case.

Table 8: Total energies (hartrees) and rotational barrier (kcal/mol) for C_2H_6.

	HF	GVB(7/14)		
	total energy	total energy	correlation energy	
			this work	reference[a]
Staggered	-79.23631	-79.34370	-0.10739	-0.1083
Eclipsed	-79.23074	-79.33839	-0.10765	-0.1086
Rotational Barrier[b]	3.5	3.3	-	3.1

[a] Goddard's results from Ref. [12(b)].
[b] the experimental barrier is 2.93 kcal/mol, see Ref. [51].

4.5 Allyl Radical C_3H_5

The electronic structure of the allyl radical C_3H_5 (C_{2v}) is a challenging case, for which RHF fails to give a proper description. The π electron system of the allyl radical is the smallest π system displaying what is known in valence bond (VB) as resonance.

$$C{=}C{-}\dot{C} \leftrightarrow \dot{C}{-}C{=}C$$

Levin and Goddard [52] reported GVB and CI calculations for the ground state and some excited states of the allyl radical. Tantardini and Simonetta [53] performed π-electron ab initio valence bond calculations.

The GVB wavefunction for the π electrons of the allyl radical is,

$$\Psi^{GVB}_{C_3H_5} = \mathcal{A}[\Psi_{core}\Psi^{g}\Theta] \tag{118}$$

There are many possible ways to construct the GVB orbitals (Ψ^{g}) and the spin functions (Θ) for the doublet ground state. One simple way is to use a three-center bonding orbital (ψ_{C-C-C}) coupled with another three-center antibonding orbital ($\psi^{\bullet}_{C-C-C}$) into a singlet, along with a delocalized open-shell orbital (ψ^{o}). The resulting GVB wavefunction is

$$\Psi^{GVB}_{C_3H_5} = \mathcal{A}[\Psi_{core}\psi_{C-C-C}\psi^{\bullet}_{C-C-C}\psi^{o}(\alpha\beta - \beta\alpha)\alpha] \tag{119}$$

The basis set used was the Huzinaga $(9s5p)$ basis set on each carbon and $(4s)$ set on each hydrogen [20(b)]. The resulting $(9s5p/4s)$ basis was contracted to $[4s2p/2s]$, as suggested by Dunning [20(c)]. The above basis was supplemented by two additional p-type (π) Gaussian functions [52] on each carbon with orbital exponents of 0.0382 and 0.0127. The GVB calculations were performed on the same geometry (C_{2v}) as used by other VB calculations [54], with R_{CC} = 1.40 Å, R_{CH} = 1.08 Å, and all bond angles = 120^0. Table 9 summarizes the energies obtained for the ground state allyl radical from various methods.

The results show that our simple GVB(1/3) description based on Eq. (119) gives the same energy as Levin and Goddard [62], who used different GVB wavefunctions. Our GVB(1/3) calculation gives a correlation energy of -24.5 kcal/mol, only 2.7 kcal/mol less than the previous CI result for the π electrons.

In our approach, two C-C σ bonds can also be included to perform a GVB(3/7) calculation, and five C-H σ bonds included to perform a GVB(8/17) calculation. The calculated total correlation energies are -38.6 kcal/mol and -84.5 kcal/mol for the GVB(3/7) and GVB(8/17) wavefunctions respectively.

Table 9: Total energies and relative energies of the ground state allyl radical.

Method[a]	Total energy hartrees	Relative energy kcal/mol
HF[b]	-116.377697	27.2
HF(this work)	-116.377817	27.2
GVB-GF[b]	-116.414974	3.8
GVB(1/3)[b]	-116.416438	2.9
GVB(1/3)(this work)	-116.416668	2.7
CI[b]	-116.421089	0.00

[a] GVB(1/3): using one pair, three orbitals for π electrons. GVB(3/7): adding two C-C σ bonds. GVB(8/17): adding five C-H σ bonds.
[b] see Ref. [52].

4.6 Dipole Moments of CO and H_2CO

The electronic dipole moment is a vector quantity, and hence is characterized not only by a scalar magnitude, but also by a direction. The Hartree-Fock wavefunction fails to predict the correct dipole moment for carbon monoxide. The experimental dipole moment is 0.11 D in the direction $^-CO^+$, whereas the near Hartree-Fock limit [5] value is 0.27 D in the reverse direction, *i.e.*, $^+CO^-$. The calculation of dipole moment can be utilized as a good measure of the quality of wavefunction.

In this study, GVB dipole moments have been calculated at the optimized geometries for carbon monoxide and formaldehyde, and compared with those obtained from HF, CID, CISD and experiment [55] in Table 10. There is very good agreement between the calculated and experimental dipole moments for formaldehyde, even at the HF level. However, the CID calculation with 200 double excitations still predicts the wrong sign for the dipole moment of carbon monoxide. A CISD calculation with over 200 single and double excitations is required to obtain the correct sign. On the other hand, the simple GVB(3/6) wavefunction already gives the proper sign for the dipole moment of carbon monoxide. After adding the lone pairs on carbon and oxygen, the GVB(5/10)/6-31G* dipole moment (0.10 D) is very close to the experimental (0.11 D) and the CISD value (0.12 D).

Table 10: The calculated dipole moments for CO and H_2CO.

Method	CO^a	H_2CO
HF/6-31G** (this study)	-0.328	2.78
GVB(3/6)/6-31G** (this study)	0.021	
GVB(5/10)/6-31G* (this study)	0.102	-
GVB(6/12)/6-31G** (this study)	-	2.42
CID/6-311G** [55]	-0.20	2.71
CISD/6-311G** [55]	0.12	2.43
Experiment [55]	0.11	2.43

[a] The positive sign of CO dipole moment corresponds to $^-CO^+$.

Chapter Five

Applications

5.1 Molecular Equilibrium Geometries

Theoretical studies of equilibrium geometries have been widely used for two main purposes: (a) developing theoretical methods and assessing their likely accuracy for structure predictions; and (b) investigating molecular geometries for those molecules whose structures are not easily amenable to experimental investigation. Based on systematic ab initio calculations of molecular geometries, Pople et al. [5, 56-59] have concluded that SCF calculations using relatively modest basis sets agree reasonably well with experiment and can therefore be used predictively within "chemical" accuracy [5]. However, there are exceptions where in order to get accurate molecular geometries larger basis sets and inclusion of electron correlation are important [5, 56-59]. Many methods exist for treating electron correlation but few are practical for determining molecular geometries. Systematic studies of equilibrium geometries have been performed with MP2 [58] and CID [59] treatment of electron correlation.

Another method which can be used to determine equilibrium geometries is the generalized valence bond (GVB) approach [12]. However, no systematic studies on optimized GVB geometries have been performed before. In this study GVB equilibrium geometries are examined for a set of simple molecules containing single or multiple bonds, and for a set of simple radicals.

5.1.1 Computational Method

The GVB geometry optimizations were carried out with appropriate symmetry constraints. C_2H_6, for example, was fixed at the D_{3d} symmetry in the optimization. GVB equilibrium geometries were obtained using the standard STO-3G, 3-21G, 6-31G*, 6-31G**, 6-31++G**, and 6-311G** basis sets [5, 60]. In the minimal basis set, GVB calculations were performed to correlate only bonds. In the split valence or larger basis sets, only the cores were frozen and both bonds and the lone pairs were correlated. For example, for CH_3F which has four bonds and a total of 14 valence electrons, the minimal basis set GVB-PP calculation is represented as GVB(4/8), while in the split valence or larger basis sets GVB-PP calculation is represented as GVB(7/14). Energy gradient geometry optimizations were carried out with the OC method [26] for the molecular ground states. GVB optimized geometries were obtained for a set of AH_n molecules, AH_n radicals, and AH_mBH_n molecules, where A, B are the first- or second-row elements.

The convergence accuracy of the GVB coefficients and density [39] was set to 10^{-5}, and the optimization convergence accuracy was set to a gradient length of 5×10^{-4}. The numerical deviations for the geometrical parameters are therefore in the third decimal place for bond lengths (in Å) and in the first decimal place for angles (in degrees). Whenever possible the theoretical equilibrium structures are compared to gas-phase experimental structural data [5, 61]. The deviations of our calculated GVB geometrical parameters from the experimental data are examined for different basis sets and different types of bonding. The mean absolute deviation (d) of the calculated parameters (x_i) from the experimental parameters (x_i^0) is defined as

$$d = \sum_{i=1}^{n} |x_i - x_i^0|/n \tag{120}$$

where n is the total number of parameters.

Tables 11 and 12 summarize the optimized GVB bond lengths for the AH_n and AH_mBH_n species at the various basis sets. The calculated GVB/STO-3G, GVB/3-21G, GVB/6-31G*, and GVB/6-31G** HAH and HAB bond angles for all molecules considered are listed in Table 13. The comparison of GVB/6-31G* results with experimental structural data is illustrated in Figures 5-7 and the line drawn in all plots is of unit slope. The GVB/6-31G* geometrical data are also compared with HF/6-31G*, MP2/6-31G* and CID/6-31G* quantities taken from the literature [5, 56-59], whereas UHF and UMP2 results are given for the radicals. Table 14 gives the values of mean absolute deviation (d), and Table 15 summarizes the correlation coefficients for our GVB geometries related to the experimental and theoretical results.

5.1.2 A-H Bond Lengths

The mean absolute deviation from experiment (66 comparisons) for the GVB/STO-3G A-H bond length is 0.039 Å, compared to 0.028 Å for HF/STO-3G. Compared to GVB/STO-3G results, GVB/3-21G gives shorter A-H bond lengths for the elements of greater electronegativity than hydrogen, and longer A-H bond lengths for the elements of less electronegativity than hydrogen, with most (85%) of A-H bond lengths systematically longer than experimental values. For BH^+, however, GVB/3-21G underestimates the experimental bond length by ~0.02 Å. Overall GVB/3-21G gives an improved description of A-H bond lengths, with the mean absolute deviation (66 comparisons) reduced to 0.022 Å. But GVB/3-21G offers no improvement over HF/3-21G ($\bar{d}$ is 0.019 Å). In general, the minimal STO-3G and split-valence 3-21G basis sets are not suitable for use in a GVB treatment of molecular geometries. Similar observations have been made for MP2 [58] and CI [59].

The addition of polarization or diffuse functions to a basis usually makes a noticeable difference in calculated structures [56-59]. In general, the GVB A-H bond distances obtained with the 6-31G*, 6-31G**, 6-31++G**, and 6-311G** basis sets are shorter and significantly better than those derived from the smaller basis sets (STO-3G and 3-21G). Most (85%) of the GVB/6-31G* A-H bond distances are systematically longer than the experimental parameters as illustrated in Figure 5, with a correlation coefficient of 0.999. Differences among GVB/6-31G*, GVB/6-31G**, GVB/6-31++G**, and GVB/6-311G** are very small (within 0.005 Å). The largest deviations are for LiH and NaH, where GVB/6-31G* overestimates the experimental bond length by ~ 0.06 Å. The GVB/6-311G** LiH bond distance is improved, but still overestimated by 0.04 Å. In H_2O and H_2S, the H-O and H-S bond lengths are 0.961 Å and 1.339 Å from GVB/6-31++G**, which are very close to the experimental values of 0.958 Å and 1.336 Å, respectively. The deviation of GVB/6-31G* from experiment, excluding LiH and NaH, is within ~0.02 Å. The mean absolute deviations are 0.011, 0.010, 0.013 and 0.010 Å for GVB/6-31G* (66 comparisons), GVB/6-31G** (66 comparisons), GVB/6-31++G** (24 comparisons), and GVB/6-311G** (19 comparisons), respectively.

The HF/6-31G*, MP2/6-31G* and CID/6-31G* results are also given in Tables 11 and 12 for comparison. The deviations in A-H bond lengths at these levels from experiment are very close to each other, in general. The Li-H and Na-H bond lengths are both overestimated at HF/6-31G*. GVB/6-31G* predicts even longer bond lengths in those cases, as do MP2 and CID. The HF/6-31G* O-H and S-H bond length in H_2O and H_2S, which are both underestimated, are overcorrected at GVB/6-31G*. In the $OH(^2\Pi)$, $OH^+(^3\Sigma^-)$ and $OH_2^+(^2B_1)$ radicals, the O-H bond distances, which are underestimated at UHF/6-31G*, are all within 0.01 Å at GVB/6-31G* and are significantly better than the UMP2/6-31G* results. Even though most of the GVB/6-31G* bond lengths are longer than the corresponding HF/6-31G* bond lengths, this discrepancy is systematic as reflected in the excellent linear correlation (correlation coefficient of 1.000). Overall, GVB/6-31G* (d is 0.011 Å) shows only a small improvement in the mean absolute deviation over HF/6-31G* (d is 0.013 Å).

On the other hand, GVB/6-31G* results (d is 0.011 Å for 66 comparisons) are close to MP2/6-31G* (d is 0.007 Å for 56 comparisons) and CID/6-31G* (d is 0.007 Å for 36 comparisons) results for A-H bond lengths. The correlation coefficients are 0.999 for GVB/6-31G* versus MP2/6-31G* and 0.999 for GVB/6-31G* versus CID/6-31G*, respectively, indicating that deviation in GVB equilibrium geometries is systematic when compared with other ab initio calculations and suggesting that GVB equilibrium geometries are similar to those of MP2 or CID.

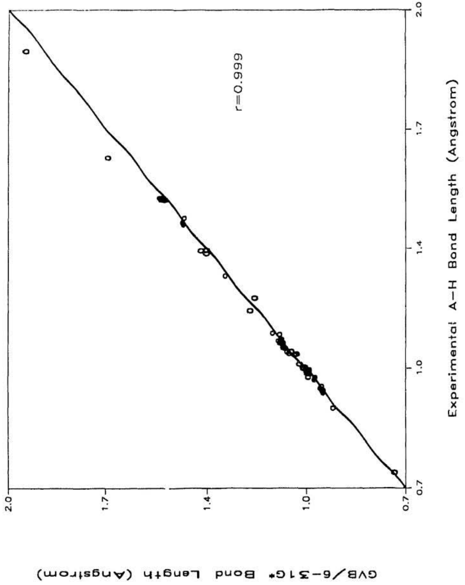
Figure 5 demonstrates the comparison of GVB/6-31G* bond lengths with the experimental results for all A-H bond lengths calculated, which clearly indicates that there is a good linear correlation with the correlation coefficient of 0.999. Among all standard basis sets examined, the 6-31G* basis set is the simplest to offer reasonable GVB equilibrium geometries.

Table 11: Equilibrium bond lengths (Å) for AH_n species

	Bond	GVB/ STO-3G	GVB/ 3-21G	HF/ 6-31G*	MP2/ 6-31G*	CID/ 6-31G*	GVB/ 6-31G*	GVB/ 6-31G**	GVB/ 6-31++G**	GVB/ 6-311G**	Expt.
AH_n Molecules											
$H_2(D_{\infty h})$	HH	0.734	0.758	0.730	0.753	0.746	0.753	0.753	0.754	0.757	0.742
$LiH(C_{\infty v})$	LiH	1.550	1.677	1.636	1.640	1.649	1.667	1.661	1.662	1.637	1.596
$NaH(C_{\infty v})$	NaH	1.661	1.972	1.914	-	-	1.952	1.950	1.957	-	1.887
$HF(C_{\infty v})$	FH	0.994	0.962	0.911	0.934	0.931	0.929	0.918	0.920	0.914	0.917
$HCl(C_{\infty v})$	ClH	1.343	1.324	1.266	1.280	-	1.288	1.286	1.287	-	1.275
$H_2O(C_{2v})$	OH	1.026	0.992	0.947	0.969	0.966	0.967	0.961	0.961	0.959	0.958
$H_2S(C_{2v})$	SH	1.361	1.385	1.326	1.340	-	1.350	1.350	1.339	-	1.336
$NH_3(C_{3v})$	NH	1.066	1.025	1.002	1.017	1.016	1.022	1.020	1.018	1.020	1.012
$PH_3(C_{3v})$	PH	1.411	1.456	1.403	1.415	-	1.427	1.428	1.428	-	1.420
$CH_4(T_d)$	CH	1.103	1.101	1.084	1.090	1.091	1.103	1.102	1.103	1.104	1.092
$SiH_4(T_d)$	SiH	1.442	1.507	1.475	1.484	-	1.497	1.497	1.497	-	1.481
AH_n Radicals											
$BeH(^2\Sigma^+)$	BeH	1.321	1.372	1.346	1.348	1.355	1.365	1.364	1.362	1.360	1.343
$BH^+(^2\Sigma^+)$	BH	1.217	1.193	1.179	1.194	-	1.191	1.192	1.192	1.195	1.215
$CH(^2\Pi)$	CH	1.180	1.152	1.107	1.120	1.128	1.133	1.135	1.131	1.133	1.120
$NH(^3\Sigma^-)$	NH	1.118	1.074	1.022	1.040	1.045	1.042	1.041	1.040	1.040	1.036
$NH^+(^2\Pi^+)$	NH	1.168	1.092	1.045	1.064	-	1.065	1.069	1.071	1.072	1.070
$OH(^2\Pi)$	OH	1.051	1.013	0.958	0.979	0.980	0.978	0.973	0.973	0.970	0.970
$OH^+(^3\Sigma^-)$	OH	1.103	1.067	1.011	1.035	-	1.029	1.025	1.026	1.026	1.028
$FH^+(^2\Pi)$	FH	1.050	1.053	1.005	1.030	-	1.019	1.000	1.001	0.993	1.001
$BH_2(^2A_1)$	BH	1.180	1.202	1.186	1.188	1.193	1.204	1.203	1.202	1.203	1.181
$CH_2(^3B_1)$	CH	1.098	1.084	1.072	1.077	1.081	1.085	1.085	1.085	1.086	1.078
$NH_2(^2B_1)$	NH	1.093	1.050	1.012	1.029	1.030	1.031	1.030	1.028	1.029	1.024
$OH_2^+(^2B_1)$	OH	1.058	1.027	0.988	1.011	-	0.998	0.998	0.998	0.996	0.999
$CH_3(^2A_2'')$	CH	1.099	1.087	1.072	1.079	1.081	1.089	1.088	1.088	1.089	1.079

UHF and UMP2 results are given for radicals.

Figure 5: Comparison of the GVB/6-31G* A-H bond lengths with the experimental data.
The line drawn in this plot is of unit slope.


r=0.999
0.7
1.0
1.4
1.7
2.0
Experimental A–H Bond Length (Angstrom)
GVB/6-31G* Bond Length (Angstrom)

5.1.3 A-B Bond Lengths

For 39 comparisons of bonds connecting non-hydrogen atoms (*i.e.*, A-B bonds), the mean absolute deviation of GVB/STO-3G is 0.054 Å, which is larger than for the corresponding A-H bonds ($\bar{d}$ is 0.039 Å). For H_2O_2, F_2 and HOF molecules, GVB/STO-3G parameters are fortuitously in good agreement with experiment. GVB/3-21G gives a poorer (overestimating) description for most of A-B bond lengths and an even worse description of A-B bond distances in the molecules containing second-row elements such as H_2S_2, with a mean absolute deviation of 0.095 Å. These results suggest that the minimal and double-zeta valence basis sets are not good enough to give the correct description of virtual orbitals in GVB-PP especially for second-row elements for which the addition of *d*-polarization or diffuse functions becomes necessary. The mean absolute deviation is 0.052 Å for GVB/3-21G A-B multiple bond lengths (17 comparisons).

As expected, the addition of *d*-polarization functions significantly reduces the deviations in GVB A-B bond lengths. For instance, the GVB/6-31G* O-O distance (1.499 Å) for H_2O_2 is longer than the experimental value (1.452 Å), but significantly reduced from the GVB/3-21G value (1.590 Å) and slightly improved over the HF/6-31G* value (1.393 Å). The GVB/6-31G* S-S distance (2.108 Å) for H_2S_2 is significantly improved over the GVB/3-21G result (2.357 Å), but still longer than the HF/6-31G* (2.064 Å) and experimental (2.058 Å) values. Most (82%) of GVB/6-31G* and GVB/6-31G** equilibrium A-B bond distances are longer than the experimental values. Differences between GVB/6-31G* and GVB/6-31G** results are small, suggesting that the addition of p-polarization functions on hydrogen does not significantly affect the equilibrium structures. Our GVB/6-31G* geometries of CO(1.131 Å), N_2(1.103 Å), and F_2(1.492 Å) are in good agreement (±0.01 Å) with those of Goddard et al [34] (1.119, 1.092, and 1.485 Å, respectively), using a different basis set. For 39 comparisons of A-B bond lengths in Table 12, the mean absolute deviation is 0.025 Å at both the GVB/6-31G* and GVB/6-31G** levels. The largest deviation is for bonds involving highly electronegative atoms (O and F), where GVB/6-31G* overestimates those bond lengths. For instance, in F_2, H_2O_2, and HOF molecules, GVB/6-31G* overestimates A-B bond lengths from experiment by ~0.05 Å. For others, the deviation of GVB/6-31G* from experiment is within ~0.02 Å. The mean absolute deviation is 0.018 Å for all single bonds (64 comparisons) at both GVB/6-31G* and GVB/6-31G**, and 0.009 and 0.010 Å for all multiple bonds (17 comparisons) respectively. The comparison of GVB/6-31G* A-B bond lengths with experimental results is illustrated in Figure 6, demonstrating a good linear correlation with a correlation coefficient of 0.998 (37 comparisons).

Table 12 also gives HF/6-31G*, MP2/6-31G* and CID/6-31G* results. The deviations of those theoretical quantities from experiment for all bond lengths are close to each other. Overall, the GVB/6-31G* geometrical parameters (d is 0.025 Å) are only slightly better than HF/6-31G* (d is 0.027 Å) for A-B bond lengths. In general, the

GVB/6-31G* A-B single bonds are longer than the HF/6-31G* values by 0.01 – 0.03 Å, with a small improvement in the deviation from experiment. For instance, the C-F bond length in CH_3F is 1.365 Å at HF/6-31G*, whereas the GVB/6-31G* value is 1.393 Å, which is longer by 0.01 Å compared to the experimental value of 1.383 Å.

GVB/6-31G* A-B multiple bonds, however, are significantly improved (*d* is 0.010 Å) over the corresponding HF/6-31G* (*d* is 0.025 Å). It is interesting to note that all HF/6-31G* A-B multiple bond lengths are shorter than experiment, while most of the MP2/6-31G*, GVB/6-31G* and CID/6-31G* values are longer than experiment. For instance, the C-N bond distances in HCN and HNC are 1.133 Å, and 1.154 Å at HF/6-31G*, while the GVB/6-31G* values are 1.157 Å, and 1.173 Å, respectively, within 0.004 Å of the experimental values. In HCP the HF/6-31G* C-P bond is underestimated by 0.025 Å whereas the GVB/6-31G* bond is overestimated by 0.013 Å, again representing a significant improvement. The GVB/6-31G* multiple bonds in C_2H_2 and C_2H_4 are also closer to experiment and both within 0.01 Å. For the N-N bond in N_2, N-P bond in NP, P-P bond in P_2, and N-N bond in N_2H_4, the GVB/6-31G* bond distances are all significantly improved and within 0.01 Å with the exception of the P-P bond.

GVB/6-31G* A-B bond lengths are close to MP2/6-31G* (*d* is 0.016 Å for 32 comparisons) and CID/6-31G* (*d* is 0.027 Å for 17 comparisons) values. The comparisons of GVB/6-31G* with HF/6-31G*, MP2/6-31G*, CID/6-31G* quantities demonstrate a linear correlation for A-B bond lengths, with their correlation coefficients (see Table

15) being 0.995 (39 comparisons), 0.997 (32 comparisons) and 0.997 (17 comparisons), respectively, suggesting that the GVB equilibrium geometries are uniformly consistent with other post Hartree-Fock calculations.

The GVB/6-31G* Li-Li bond length, although still too long (0.044 Å), is closer to experiment than the HF/6-31G*, MP2/6-31G* and CID/6-31G* values. The GVB/6-31G* F-F bond in F_2 (1.492 Å) is overcorrected from HF/6-31G* (1.345 Å), and longer than experiment (1.412 Å), while the MP2/ 6-31G* and CID/6-31G* are within the deviation of ~0.01 Å. For molecules containing bonds involving highly electronegative atoms, *e.g.*, the F-F bond in F_2, more configurations than those which are included in the GVB-PP wavefunction, are required to get accurate equilibrium geometries. For instance, GVB-RCI [34] gives a F-F bond distance of 1.422 Å, corrected from the corresponding GVB-PP result of 1.485 Å. However, since GVB-RCI has no gradient capacity, only a limited number of geometries can be obtained at that level.

Table 12: Equilibrium bond lengths (Å) for AH_mBH_n molecules, where A and B are first- and second-row elements.

Molecule	Bond	GVB/ STO-3G	GVB/ 3-21G	HF/ 6-31G*	MP2/ 6-31G*	CID/ 6-31G*	GVB/ 6-31G*	GVB/ 6-31G**	Expt.
$Li_2(D_{\infty h})$	LiLi	2.844	2.732	2.812	2.782	2.724	2.717	2.717	2.673
$N_2(D_{\infty h})$	NN	1.184	1.114	1.078	1.131	1.103	1.103	1.103	1.098
$F_2(D_{\infty h})$	FF	1.391	1.536	1.345	1.421	1.399	1.492	1.492	1.412
$CO(C_{\infty v})$	CO	1.168	1.148	1.114	1.151	1.133	1.131	1.131	1.128
$CS(C_{\infty v})$	CS	1.546	1.589	1.520	1.546	-	1.537	1.537	1.535
$NP(C_{\infty v})$	NP	1.538	1.583	1.455	-	-	1.501	1.501	1.491
$P_2(D_{\infty h})$	PP	1.893	2.029	1.859	1.936	-	1.916	1.916	1.893
$NaF(C_{\infty v})$	NaF	1.767	1.864	1.885	1.920	-	1.894	1.894	1.926
$ClF(C_{\infty v})$	ClF	1.772	1.827	1.613	-	-	1.693	1.693	1.628
$Cl_2(D_{\infty h})$	ClCl	2.133	2.323	1.990	2.015	-	2.071	2.071	1.988
$HCN(C_{\infty v})$	CN	1.192	1.163	1.133	1.177	1.154	1.157	1.158	1.153
	CH	1.086	1.064	1.059	1.070	1.067	1.073	1.073	1.065
$HNC(C_{\infty v})$	NC	1.191	1.177	1.154	1.187	1.171	1.173	1.173	1.169
	NH	1.033	0.998	0.985	1.002	0.997	0.999	0.998	0.994
$HCP(C_{\infty v})$	CP	1.524	1.594	1.515	1.562	-	1.553	1.553	1.540
	CH	1.088	1.073	1.063	1.076	-	1.079	1.079	1.069
$HNO(C_s)$	NO	1.298	1.316	1.175	1.237	1.206	1.220	1.220	1.212
	NH	1.117	1.062	1.032	1.058	1.052	1.050	1.050	1.063
$HPO(C_s)$	PO	1.623	1.632	1.460	1.519	-	1.509	1.497	1.512
	PH	1.436	1.477	1.431	1.453	-	1.452	1.453	-
$HOF(C_s)$	OF	1.432	1.555	1.376	1.444	1.420	1.491	1.491	1.442
	OH	1.045	1.007	0.973	0.952	0.979	0.971	0.969	0.966
$HOCl(C_s)$	OCl	1.823	1.907	1.670	-	-	1.761	1.759	1.690
	OH	1.040	1.005	0.951	-	-	0.972	0.967	0.975
$H_2CO(C_{2v})$	CO	1.273	1.248	1.184	1.221	1.205	1.215	1.214	1.208
	CH	1.114	1.079	1.092	1.104	1.101	1.106	1.105	1.116
								...	

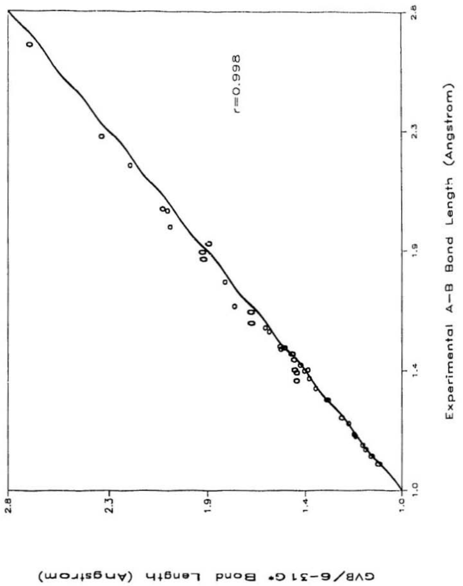
Table 12. (continued)

Molecule	Bond	GVB/ STO-3G	GVB/ 3-21G	HF/ 6-31G*	MP2/ 6-31G*	CID/ 6-31G*	GVB/ 6-31G*	GVB/ 6-31G**	Expt.
$H_2CS(C_{2v})$	CS	1.635	1.685	1.597	1.617	-	1.634	1.633	1.611
	CH	1.106	1.088	1.078	1.090	-	1.094	1.095	1.093
$H_2O_2(C_2)$	OO	1.480	1.590	1.393	1.467	-	1.499	1.499	1.452
	OH	1.037	0.998	0.949	0.976	-	0.969	0.965	0.965
$H_2S_2(C_2)$	SS	2.117	2.375	2.064	2.069	-	2.108	2.109	2.058
	SH	1.367	1.390	1.327	1.344	-	1.350	1.350	1.345
$N_2H_2(C_{2h})$	NN	1.350	1.310	1.216	1.267	1.242	1.249	1.257	1.252
	NH	1.093	1.046	1.014	1.036	1.031	1.032	1.037	1.028
$CH_2NH(C_s)$	CN	1.329	1.296	1.250	1.282	1.268	1.282	1.281	1.273
	CH	1.106	1.094	1.084	1.096	1.092	1.099	1.100	1.103
	CH	1.105	1.090	1.080	1.090	1.087	1.095	1.096	1.081
	NH	1.080	1.040	1.006	1.027	1.021	1.026	1.025	1.023
$CH_3F(C_{3v})$	CF	1.427	1.437	1.365	1.392	1.382	1.393	1.393	1.383
	CH	1.114	1.097	1.082	1.092	1.090	1.100	1.100	1.100
$CH_3OH(C_{3v})$	CO	1.479	1.480	1.400	1.424	1.415	1.429	1.427	1.421
	CH	1.109	1.095	1.081	1.090	1.088	1.098	1.099	1.094
	CH	1.113	1.102	1.087	1.097	1.095	1.105	1.106	1.094
	OH	1.027	0.990	0.946	0.970	0.963	0.965	0.960	0.963
$C_2H_2(D_{\infty h})$	CC	1.198	1.208	1.185	1.218	1.202	1.211	1.210	1.203
	CH	1.082	1.059	1.057	1.066	1.065	1.056	1.072	1.061
$C_2H_4(D_{2h})$	CC	1.347	1.345	1.317	1.336	1.328	1.338	1.346	1.339
	CH	1.100	1.091	1.076	1.085	1.084	1.094	1.093	1.085
$C_2H_6(D_{3d})$	CC	1.560	1.566	1.527	1.527	-	1.549	1.548	1.531
	CH	1.106	1.102	1.086	1.094	-	1.105	1.105	1.096
$CH_3NH_2(C_s)$	CN	1.508	1.510	1.453	1.465	1.460	1.475	1.474	1.471
	CH	1.114	1.108	1.091	1.100	1.098	1.111	1.111	1.099
	CH	1.108	1.099	1.084	1.092	1.091	1.102	1.102	1.099
	NH	1.053	1.027	1.001	1.018	1.014	1.018	1.017	1.010

Table 12.(continued)

Molecule	Bond	GVB/ STO-3G	GVB/ 3-21G	HF/ 6-31G*	MP2/ 6-31G*	CID/ 6-31G*	GVB/ 6-31G*	GVB/ 6-31G**	Expt.
...									
$CH_3SiH_3(C_{3v})$	CSi	1.882	1.939	1.888	-	-	1.912	1.911	1.867
	CH	1.103	1.103	1.086	-	-	1.106	1.105	1.093
	SiH	1.443	1.510	1.478	-	-	1.500	1.500	1.485
$CH_2PH(C_s)$	CP	1.683	1.737	1.652	-	-	1.695	1.695	1.670
	CH	1.099	1.089	1.075	-	-	1.093	1.093	-
	CH	1.099	1.090	1.076	-	-	1.094	1.094	-
	PH	1.420	1.469	1.409	-	-	1.432	1.433	-
$CH_3PH_2(C_s)$	CP	1.863	1.937	1.861	-	-	1.878	1.877	1.862
	CH	1.106	1.100	1.082	-	-	1.104	1.104	1.094
	CH	1.103	1.098	1.084	-	-	1.102	1.102	1.094
	PH	1.412	1.446	1.404	-	-	1.421	1.421	1.432
$CH_3SH(C_s)$	CS	1.833	1.947	1.817	1.817	-	1.849	1.849	1.819
	CH	1.106	1.094	1.082	1.091	-	1.099	1.099	1.091
	CH	1.107	1.095	1.081	1.090	-	1.100	1.100	1.091
	SH	1.362	1.388	1.327	1.341	-	1.350	1.349	1.336
$CH_3Cl(C_{3v})$	CCl	1.837	1.937	1.785	1.778	-	1.814	1.813	1.781
	CH	1.107	1.091	1.078	1.088	-	1.097	1.097	1.096
$N_2H_4(C_2)$	NN	1.536	1.537	1.413	1.439	1.430	1.455	1.451	1.449
	NH	1.070	1.030	0.999	1.016	1.012	1.018	1.017	1.021
	NH	1.073	1.035	1.003	1.021	1.016	1.025	1.021	1.021
$NH_2OH(C_s)$	NO	1.476	1.533	1.403	1.453	1.433	1.437	1.437	1.453
	NH	1.064	1.002	1.002	1.021	1.016	1.018	1.017	1.016
	OH	1.042	1.002	0.946	0.971	0.964	0.970	0.965	0.962
$SiH_3F(C_{3v})$	SiF	1.666	1.653	1.594	1.619	-	1.612	1.611	1.596
	SiH	1.442	1.497	1.470	1.481	-	1.491	1.492	1.480
$SiH_3Cl(C_{3v})$	SiCl	2.105	2.215	2.067	-	-	2.085	2.086	2.048
	SiH	1.444	1.495	1.468	-	-	1.489	1.490	1.481
$Si_2H_6(D_{3d})$	SiSi	2.257	2.404	2.353	2.338	-	2.375	2.376	2.327
	SiH	1.445	1.509	1.478	1.487	-	1.501	1.501	1.486
$P_2H_4(C_2)$	PP	2.229	2.478	2.214	-	-	2.256	2.257	2.219
	PH	1.422	1.457	1.401	-	-	1.424	1.425	1.417
	PH	1.412	1.461	1.402	-	-	1.424	1.425	1.414

Figure 6: Comparison of the GVB/6-31G* A-B bond lengths with the experimental data for AH_mBH_n molecules.
The line drawn in this plot is of unit slope.


r=0.998
Experimental A–B Bond Length (Angstrom)
GVB/6-31G* Bond Length (Angstrom)
1.0
1.4
1.9
2.3
2.8

5.1.4 Bond Angles

Generally, GVB bond angles are in good agreement with the experimental values. GVB/STO-3G equilibrium geometries indicate that HAH and HAB bond angles are in reasonable accord with the experimental data. However, the equilibrium dihedral angle in H_2O_2 at GVB/STO-3G (180.0^0)[9] is far larger than experiment (119.1^0), while there is good agreement between GVB/STO-3G (93.6^0) and experiment (90.8^0) for the dihedral angle in H_2S_2. For 26 comparisons of GVB/STO-3G bond angles, excluding dihedral angles, the mean absolute deviation is 3.0^0. The mean absolute deviation of GVB/3-21G is reduced to 1.9^0, giving an improved description of bond angles over GVB/STO-3G and HF/3-21G (d is 2.1^0). However there is still no change in the dihedral angle of H_2O_2.

After adding polarization functions, GVB/6-31G* and GVB/6-31G** equilibrium bond angles are better than those derived from the smaller basis sets (STO-3G, 3-21G). The dihedral in H_2O_2 is changed to be 124.5^0 at GVB/6-31G* and 124.3^0 at GVB/6-31G**, and closer to the experimental value of 119.1^0. As for bond lengths, bond angles also show very small differences between GVB/6-31G* and GVB/6-31G**. The addition of p-polarization functions on hydrogen again does not significantly affect the GVB equilibrium structures. For 26 comparisons of GVB bond angles with experiment, the mean absolute deviation is 1.1^0 for both GVB/6-31G* and GVB/6-31G**. At this level, the bond angles in $CH_2(^3B_1)$ and $BH_2(^2A_1)$ still

[9] This is a "well known" basis set problem in ab initio calculations [5], not a failure of the GVB approach.

deviate from experiment by 5^0. Bond angles are much better described, relative to bond lengths in molecules containing highly electronegative atoms (such as HOF and H_2O_2), with the addition of d-polarization functions (such as in the 6-31G* basis set). The HOF bond angle at GVB/6-31G*, for example, is 96.6^0, very close to experiment ($96.8\ ^0$). The HOO bond angle in H_2O_2 at GVB/6-31G* is 98.5^0, compared with the experimental value of 100.0^0. Figure 7 illustrates a comparison of GVB/6-31G* bond angles with experimental results for all bond angles in Table 13, which shows a linear correlation with a correlation coefficient of 0.987.

Table 13 also gives the HF/6-31G*, MP2/6-31G* and CID/6-31G* bond angles. GVB/6-31G* (d is 1.1^0) results are slightly better than HF/6-31G* (d is 1.5^0 for 26 comparisons), and very close to the MP2/6-31G* (d is 1.0^0 for 22 comparisons) and CID/6-31G* (d is 1.2^0 for 11 comparisons) results. For instance, the HF/6-31G* HOH bond angle and dihedral angle in H_2O_2 are 102.2^0 and 115.2^0, while the GVB/6-31G* values are 98.5^0 and 124.5^0, which are closer to the experimental values of 100.0^0 and 119.1^0, respectively. The HF/6-31G* HSH bond angle and dihedral angle in H_2S_2 are 99.1^0 and 87.9^0, while the GVB/6-31G* angles are 98.2^0 and 90.7^0, which are in excellent agreement with the experimental values of 98.1^0 and 90.8^0, respectively. The GVB/6-31G* bond angles show good linear correlation with those from HF/6-31G*, MP2/6-31G* and CID/6-31G*, with correlation coefficients of 0.958, 0.995 and 0.997, respectively (see Table 15).

Table 13: Equilibrium bond angles (degree) for molecules considered.

System	Angle	GVB/ STO-3G	GVB/ 3-21G	HF/ 6-31G*	MP2/ 6-31G*	CID/ 6-31G*	GVB/ 6-31G*	GVB/ 6-31G**	Expt.
HNO	HNO	104.9	105.6	108.8	107.3	108.0	107.5	107.6	108.6
HPO	HPO	96.7	100.4	105.4	105.6	-	103.5	101.5	104.7
HOF	HOF	98.7	95.7	99.8	97.2	98.3	96.6	96.8	96.8
HOCl	HOCl	98.1	99.5	105.1	-	-	102.0	102.0	102.5
H_2O	HOH	97.5	104.8	105.5	104.0	104.3	104.2	104.5	104.5
H_2S	HSH	91.6	94.2	94.4	93.3	-	93.7	93.7	92.1
BH_2	HBH	123.3	127.6	126.1	127.6	127.9	126.2	126.4	131
CH_2	HCH	123.5	128.8	128.4	131.6	132.0	128.2	128.6	136
NH_2	HNH	97.7	104.1	104.4	103.3	103.1	103.7	103.6	103.3
OH_2^+	HOH	105.9	114.7	111.7	109.9	-	110.1	111.0	110.5
NH_3	HNH	100.8	109.3	107.2	106.3	106.3	106.2	106.2	106.7
PH_3	HPH	92.7	94.6	95.4	94.6	-	94.8	95.0	93.3
CH_3	HCH	120.0	120.0	120.0	120.0	120.0	120.0	120.0	120
CH_3F	HCH	108.9	110.0	109.8	109.8	109.8	110.2	110.2	110.6
CH_3Cl	HCH	110.4	112.3	110.5	110.1	-	110.7	110.7	110.0
H_2CO	HCH	116.1	116.4	115.7	115.6	115.8	116.7	116.9	116.5
H_2CS	HCH	113.3	117.7	115.5	116.0	-	116.3	116.4	116.9
H_2O_2	HOO	97.4	95.6	102.2	98.7	-	98.5	98.7	100.0
	HOOH	180.0	180.0	115.2	121.3	-	124.5	124.3	119.1
H_2S_2	HSS	95.7	94.4	99.1	99.0	-	98.2	98.1	98.1
	HSSH	93.6	96.5	87.9	90.3	-	90.7	90.8	90.8
C_2H_4	HCH	116.1	116.4	116.4	116.6	116.3	116.4	116.6	117.8
C_2H_6	HCH	108.3	108.3	107.7	107.7	-	108.8	108.8	107.8
SiH_3F	HSiH	109.7	109.7	110.2	110.0	-	110.3	110.3	110.6
SiH_3Cl	HSiH	111.2	111.7	108.3	-	-	110.6	110.7	110.9
CH_3SiH_3	HCH	107.5	108.2	107.8	-	-	107.7	107.8	107.7
	HSiH	108.8	108.3	108.3	-	-	108.4	108.4	108.3
Si_2H_6	HSiH	107.8	108.8	108.5	108.6	-	108.4	108.4	107.8

Table 14: The mean absolute deviations (d) of theoretical geometries from experiment.

Method	r_{A-H}[(a)]	r_{A-B}[(b)]	r_{Single}[(c)]	$r_{Multiple}$[(d)]	r_{all}[(e)]	θ_{all}[(f)]
HF/STO-3G	0.028	0.045	0.028	0.025	0.031	2.5
GVB/STO-3G	0.039	0.054	0.042	0.043	0.044	3.0
HF/3-21G	0.019	0.035	0.037	0.017	0.029	2.1
GVB/3-21G	0.022	0.095	0.055	0.052	0.050	1.9
HF/6-31G*	0.013	0.027	0.019	0.025	0.018	1.5
MP2/6-31G*	0.007	0.016	0.008	0.018	0.010	1.0
CID/6-31G*	0.007	0.011	0.008	0.005	0.008	1.2
GVB/6-31G*	0.011	0.025	0.018	0.010	0.016	1.1
GVB/6-31G**	0.010	0.025	0.018	0.010	0.015	1.1
GVB/6-31++G**	0.013	-	-	-	0.014	-
GVB/6-311G**	0.010	-	-	-	0.011	-

Bond lengths (r) are in Å, and bond angles (θ) in degrees.

[a] A-H bond lengths. 66 comparisons, except 24 for GVB/6-31++G**, 19 for GVB/6-311G**, 56 for MP2/6-31G*, and 36 for CID/6-31G*.

[b] A-B bond lengths. 39 comparisons, except 32 for MP2/6-31G*, and 17 for CID/6-31G*.

[c] single bond lengths in AH_mBH_n molecules. 65 comparisons, except 50 for MP2/6-31G*, and 30 for CID/6-31G*.

[d] multiple bond lengths in AH_mBH_n molecules. 17 comparisons, except 16 for MP2/6-31G*, and 10 for CID/6-31G*.

[e] all bond lengths considered. 105 comparisons, except 24 for GVB/6-31++G**, 19 for GVB/6-311G**, 88 for MP2/6-31G*, and 54 for CID/6-31G*.

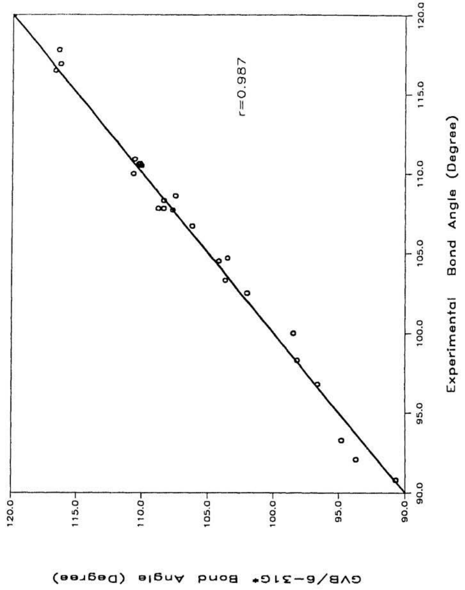
[f] all bond angles considered. 26 comparisons, except 22 for MP2/6-31G*, and 11 for CID/6-31G*.

Table 15: The correlation coefficients for our GVB geometries related to the experimental, and other theoretical results.

	r_{A-H}	r_{A-B}	bond angles
GVB/6-31G* *vs* expt.	0.999(66)	0.998(37)	0.987(26)
GVB/6-31G* *vs* HF/6-31G*	1.000(66)	0.995(39)	0.958(26)
GVB/6-31G* *vs* MP2/6-31G*	0.999(56)	0.997(32)	0.995(22)
GVB/6-31G* *vs* CID/6-31G*	0.999(36)	0.997(17)	0.997(11)

The numbers in parentheses are the number of comparisons.

Figure 7: Comparison of the GVB/6-31G* bond angles with the experimental data for molecules considered.
The line drawn in this plot is of unit slope.


r=0.987
Experimental Bond Angle (Degree)
GVB/6-31G* Bond Angle (Degree)
90.0
95.0
100.0
105.0
110.0
115.0
120.0

5.2 Bond Dissociation Energies and Heats of Reaction

Bond dissociation energies (BDE) are fundamental quantities in describing the changing of chemical bonds. The accurate calculation of BDE is usually required to study photochemical processes and the mechanisms of catalysis. Hartree-Fock theory with any basis set gives poor results for direct calculation of the energy of a homolytic dissociation process A-B $\rightarrow$ A$\cdot$ + B$\cdot$, as illustrated in Figure 1 for the dissociation of H_2. High level perturbation methods or other post-Hartree-Fock methods are required to compute BDE. Møller-Plesset theory (MP2, MP3, MP4) has been shown to give good results for BDE of many small molecules [5]. With some further improvements, ab initio calculations are capable of reproducing the experimental BDE to high accuracy ($\pm$2 kcal/mol) [5].

Bauschlicher et al. [62-63] have shown that BDE for C-C, C-H single bonds and C-C, N-N multiple bonds, calculated by using MRCI and CASSCF with atomic natural orbital (ANO) basis sets, are in agreement with the experimental results to very high accuracy. Neuheuser et al. [64] recently reported BDE for some radical reactions and charge transfer reactions, from CI calculations by employing localized molecular orbitals (LMO).

As is widely recognized, the GVB wavefunction [12] properly describes the homolytic bond dissociation process. For example, the GVB potential energy curve for $H_2O \rightarrow HO\cdot + \cdot H$, illustrated in Figure 3, has been discussed in Section 4.2. For the ground state ($X^1\Sigma^+$) and some excited states ($^3\Sigma^+, ^3\Pi, ^1\Pi$) of hydrogen fluoride, the calculated GVB and GVB-CI potential energy curves have been studied by Dunning [65]. The spectroscopic constants obtained in these calculations are given in Table 16, and compared with Hartree-Fock and experimental results. There is a significant change in nearly all of the fundamental constants in going from HF to GVB. The calculated GVB constants are in good agreement with the experimental data, with the exception of the absolute bond dissociation energies (D_e).

The GVB wavefunction provides a clear orbital description of the electronic structure of molecules and properly describes bond dissociation, but the emphasis of GVB is not on getting 100% of the correlation energy. Therefore, it is very important systematically to examine the consistency of GVB bond dissociation energies. De Maré et al. [17] has presented the behaviour of GVB potential energy surfaces (PES) for primary α-CH bond rupture in propene, ethylene, methane, water and hydrogen fluoride, showing a similar PES for A-H bond rupture with different values of D_e. The results are useful for studying the dynamics of unimolecular decomposition via primary C-H bond dissociation in olefins [66].

In this work, GVB D_e values were obtained for a series of A-H and A-B single bond dissociation reactions, where A, B are first- or second-row elements. The reaction energies, or heats of reaction, for hydrogenation and for reactions converting multiple to single bonds, were further studied to examine the GVB evaluation of reaction energies (ΔE).

Table 16: The calculated and experimental spectroscopic constants for the ground state ($X^1\Sigma^+$) of hydrogen fluoride.

	HF[a]	GVB[a]	GVB-CI[a]	Expt.[b]
E_e, hartree	-100.06247	-100.08523	-100.08618	-100.530
R_e,Å	0.899	0.9165	0.9170	0.91680
D_e, eV	4.27	4.89	4.92	6.122
k_e, $10^5 dyn/cm$	11.43	9.65	9.63	9.654
ω_e, cm^{-1}	4503	4139	4134	4138.73
$\omega_e\chi_e$	112.6	126	125	90.05
$B_e cm^{-1}$	21.82	20.97	20.95	20.9555
α_e	0.785	0.921	0.918	0.7958

[a] The calculated results are from T. H. Dunning, Jr., J. Chem. Phys. 65, 3854 (1976).
[b] The experimental results are from B. Rosen, Spectroscopic Data Relative to Diatomic Molecules , Pergamon, New York, (1970).

5.2.1 Computational Method

The GVB A-B bond D_e is calculated by the formula,

$$D_e(A-B) = E(R_0 + a) - E(R_0)$$

where R_0 is the equilibrium A-B bond length, and a is the stretching distance.[10] At $(R_0 + a)$, the A-B antibonding orbital GVB CI-coefficient is 0.707 $(\frac{1}{\sqrt{2}}) \pm 0.005$ which is a criterion for bond rupture. The full correlation of electrons in the bond that is breaking is consistent in the GVB description as suggested by Carter and Goddard [67].

For all calculations in this section, the GVB/6-31G** energies were calculated using the full valence GVB-PP wavefunction which included all bonds and valence lone pairs. The geometries were optimized at the same level [42] and have been reported in Section 5.1.

As an example, the symmetrical stretching dissociation of the two C-H bonds in H_2CO was calculated using GVB(6/12)/6-31G**. The GVB(6/12) wavefunction consists of two C-H bonds, one C-O σ bond, one C-O π bond, and two oxygen lone pairs, *i.e.*,

$$\Psi^{GVB}_{H_2CO} = \mathcal{A}[\Psi_{core}\psi_{C-H1}\psi_{C-H2}\psi_{C-O(\sigma)}\psi_{C-O(\pi)}\psi_{O-lp1}\psi_{O-lp2}] \tag{121}$$

The GVB(6/12)/6-31G** equilibrium geometry for H_2CO is R_{C-H}=1.214 Å, R_{C-O}=1.106 Å, and $\Theta_{H-C-O} = 121.56^0$. The total energy, dipole moment, and GVB

[10] Generally, a is taken to be 5.0-6.0 Å. At $(R_0 + a)$, the A-B bond is essentially broken.

CI-coefficient for C-H antibonding orbital, are listed in Table 17 as a function of the C-H bond distance. The relative energies and the dipole moments are illustrated in Figure 8. As suggested by De Maré et al [17], the GVB dipole moments clearly indicate the dissociation process. At R_0+5.0 Å, the potential energy reaches the point of dissociation and the GVB CI-coefficient is 0.707 ($\frac{1}{\sqrt{2}}$), where R_0 is the equilibrium C-H bond distance.

5.2.2 A-H Bond Dissociation Energies

Potential energy curves for hydrogen fluoride were calculated by using the GVB(4/8) wavefunction with 6-31G**, 6-31++G** and 6-311G** basis sets, in order to examine the performance of our GVB approach and to compare with the previous results. The GVB(4/8) wavefunction includes the H-F σ bond and three fluorine lone pairs. The relative energies are illustrated in Figure 9. Our GVB potential energy curve (6-311G**) with R_e=0.9170 Å and D_e=114 kcal/mol is in excellent agreement with Dunning's of R_e=0.9165 Å and D_e=114 kcal/mol (Table 16) [65].

The calculated GVB/6-31G** A-H bond dissociation energies (D_e) are listed in Table 18, and the available values for HF/6-31G**//HF/6-31G*, MP2/6-31G**//HF/6-31G* and for experiment [5] are also given for comparison. The values of ZPE [62-63] and D_0 for some dissociation reactions are given in Table 19. In general, our GVB D_e values are in good agreement with previous calculations. For instance, our GVB D_e for LiH is 43 kcal/mol, compared to 44 kcal/mol given by Goddard [56].

Table 17: GVB(6/12)/6-31G** total energy, dipole moment and GVB CI-coefficient for the C-H antibonding orbital as a function of the change in the C-H bond distance for H_2CO.

Δr_{C-H} (Å)	E (hartrees)	Dipole moment (Debye)	σ^*
0.0	-113.969233	2.423	0.0882
0.1	-113.959460	2.336	0.1027
0.2	-113.936395	2.244	0.1186
0.3	-113.906674	2.149	0.1358
0.4	-113.874305	2.056	0.1543
0.5	-113.841736	1.967	0.1738
0.6	-113.810435	1.883	0.1941
0.7	-113.781212	1.806	0.2145
0.8	-113.754417	1.739	0.2350
0.9	-113.730116	1.680	0.2555
1.0	-113.708209	1.634	0.2760
1.5	-113.628533	1.575	0.3820
2.0	-113.588909	1.890	0.5411
3.0	-113.574969	2.110	0.6818
4.0	-113.574423	2.119	0.7043
5.0	-113.574413	2.119	0.7068
6.0	-113.574413	2.119	0.7070
7.0	-113.574413	2.119	0.7071
8.0	-113.574413	2.119	0.7071
9.0	-113.574413	2.119	0.7071

Figure 8: GVB(6/12)/6-31G** relative energies and dipole moments for the symmetrical dissociation of the C-H bonds in H_2CO (C_{2v}).

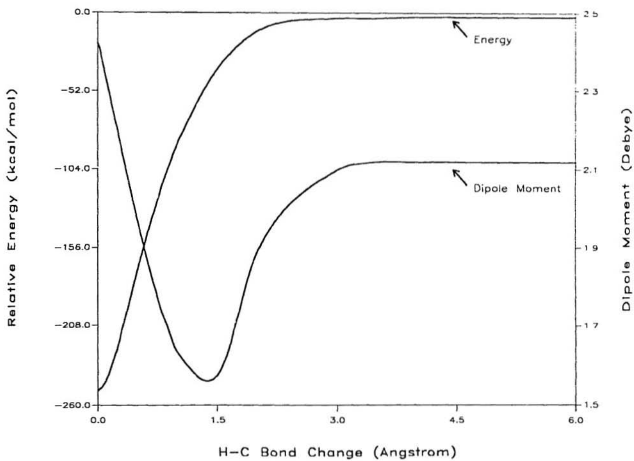
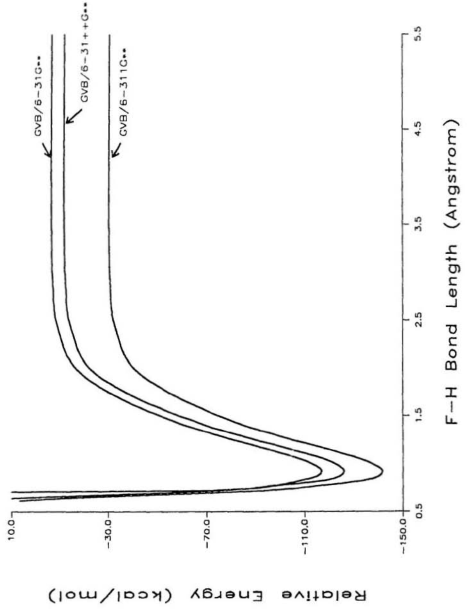

Relative Energy (kcal/mol)
0.0
-52.0
-104.0
-156.0
-208.0
-260.0
Dipole Moment (Debye)
2.5
2.3
2.1
1.9
1.7
1.5
0.0
1.5
3.0
4.5
6.0
H-C Bond Change (Angstrom)
Energy
Dipole Moment

Figure 9: GVB(4/8) potential energy curve for the dissociation of HF.


GVB/6-31G**
GVB/6-31++G**
GVB/6-311G**
10.0
-30.0
-70.0
-110.0
-150.0
0.5
1.5
2.5
3.5
4.5
5.5
Relative Energy (kcal/mol)
F—H Bond Length (Angstrom)

It is obvious from Tbales 18 and 20 that HF calculations exhibit large errors (20–40 kcal/mol) in the evaluation of D_e values for A-H bonds. Compared to HF calculations, the GVB D_e values are improved by 10-20 kcal/mol. However, they are still about 10-20 kcal/mol lower than the experimental values. Overall, the mean absolute deviation of our GVB D_e values is 15 kcal/mol (9 comparisons), compared to 31 kcal/mol for HF (7 comparisons) and 7 kcal/mol for MP2 (7 comparisons) respectively.

The largest deviation is for hydrogen fluoride, where GVB predicts a D_e of 110 kcal/mol, significantly better than the HF value of 87 kcal/mol, but still far from the experimental value of 141 kcal/mol [68] and worse than the MP2 result of 131 kcal/mol [5]. Since the GVB potential energy curve is similar to experiment as shown by Dunning [65], the discrepancy is only due to insufficient treatment of the correlation energy for the equilibrium structure.[11]

GVB D_e values for C-H bonds in hydrocarbons are predicted to within 10 kcal/mol of the experimental values, and in good agreement with previous post-HF results. For instance, our GVB C-H D_e value for C_2H_4 is 111 kcal/mol, compared to 116 kcal/mol (GVB-CCCI [67]), 117 kcal/mol (MRCI [62]) and 118 kcal/mol (experimental value [5]). The C-H D_e values for CH_4 and C_2H_6 are 106 and 104 kcal/mol, respectively, compared to 109 and 110 kcal/mol from a recent CI calculation [64]. MRCI produced a C-H D_e of 110.9 kcal/mol for CH_4 [62]. For the C-H D_0 in acetylene, our GVB result

[11] As discussed in the dissociation of H_2, GVB(1/2)/6-31G** obtains 54% of the correlation energy of full CI at the equilibrium structure, and it is almost 100% when the bond distance is greater than 3.0 a.u.

is 121 kcal/mol, 10 kcal/mol lower than the experimental value of 131.3±0.7 kcal/mol [63]. Many post-HF calculations of D_0 are available for comparison, such as 130.1 (MRCI) [69], 129.7 (GVB-CCCI) [70], 133.5±3 (G1) [71], and 131.5 kcal/mol (CIQ) [72]. The comparison shows that much simpler GVB wavefunction gives reasonably good C-H D_e which are comparable to CI results.

5.2.3 A-B Single Bond Dissociation Energies

A selection of calculated A-B single bond dissociation energies (D_e) are compared with the experimental values in Table 20. The HF/6-31G**//HF/6-31G* D_e values are in very poor agreement with the experimental values with errors are larger than 50 kcal/mol in some cases. Both HOF and F_2 molecules are calculated to be thermodynamically unstable with respect to OF and FF bond rupture. A zero bond energy is calculated for the central cleavage in hydrogen peroxide H_2O_2.

GVB significantly improves the D_e over HF. For instance, the GVB D_e is improved by 50 kcal/mol over the HF D_e for F_2. In general, our GVB/6-31G** A-B single bond D_e is in error by 10-25 kcal/mol, similar to the A-H D_e. While the MP2/6-31G** D_e is in error by only a few kcal/mol.

For hydrocarbons, GVB C-X single bond D_e values are in good agreement with experimental values, and comparable to those from previous post-HF calculations. For example, the GVB C-C D_e in C_2H_6 is 85.3 kcal/mol, compared to 80.8 (CASSCF), 93.0 (MRCI), 94.4 (MRCI+Q), and 94.2 kcal/mol (MCPF) [63], whereas D_0 [73] is

88.7 (G1) and 87.8 kcal/mol (experiment). Our GVB result for the C-F D_e in CH_3F is 96 kcal/mol, compared to the CI value [64] of 102 kcal/mol and the experimental value of 114 kcal/mol [64].

For the homolytic dissociation of a polar bond A^+-B^-, there is a charge transfer [64] from B to A since the products are two radicals A· and B·. In CH_3X, higher C-X D_e are obtained for more polar C-X bonds, for example, D_e is highest for X=F. The N-N D_e in NH_2NH_2 is lower than the C-N D_e in CH_3NH_2, because there is a larger charge transfer for homolytic breaking of the C-N bond in CH_3NH_2. The GVB N-N D_e in N_2H_4 is 54.3 kcal/mol, compared to 51.8 (2×2 CASSCF), 64.8 (CASSCF), 67.2 (MRCI), 68.3 (MRCI+Q), and 66.8 kcal/mol (MCPF) [63], while D_0 [73] is 64.6 (G1) and 65.4 kcal/mol (experiment).

Overall, our simple GVB wavefunction gives a proper description of the bond dissociation process with reasonable D_e values, that are much improved over HF and are comparable to other post-HF (MP2 and CI) results. Figure 10 compares GVB and experimental single bond D_e values. The correlation coefficient is 0.976 and a regression equation is obtained with $D_e(expt) = 0.893 D_e(GVB) + 25.1$ kcal/mol, which can be used to correct GVB D_e results. Such a good correlation indicates that the errors in GVB D_e values are systematic.

Table 18: A-H bond dissociation energies (kcal/mol).

Bond dissociation reaction	HF/6-31G** //HF/6-31G* [5]	MP2/6-31G** //HF/6-31G* [5]	GVB/6-31G** //GVB/6-31G** this study	expt. [5]
$H_2 \rightarrow H\cdot + H\cdot$	85	101	96.1	109
$LiH \rightarrow Li\cdot + H\cdot$	32	45	42.8	58
$FH \rightarrow F\cdot + H\cdot$	93	131	110	141
$BH_3 \rightarrow BH_2^{\cdot} + H\cdot$	90	106	100	107
$CH_4 \rightarrow CH_3^{\cdot} + H\cdot$	87	109	106	113
$CH_3F \rightarrow \cdot CH_2F + H\cdot$			104	
$CH_3OH \rightarrow \cdot CH_2OH + H\cdot$			106	
$CH_3CH_3 \rightarrow \cdot CH_2CH_3 + H\cdot$			104	
$CH_3NH_2 \rightarrow \cdot CH_2NH_2 + H\cdot$			98.8	
$H_2CO \rightarrow \cdot HCO + H\cdot$			99.7	
$HCCH \rightarrow \cdot CCH + H\cdot$		138.2[a]	128	138
$H_2CCH_2 \rightarrow \cdot HCCH_2 + H\cdot$		116.6[a]	111	118
$H_2O \rightarrow HO\cdot + H\cdot$	86	119	102	126
$CH_3OH \rightarrow CH_3O\cdot + H\cdot$			94.6	
$HOF \rightarrow FO\cdot + H\cdot$			93.5	
$HOOH \rightarrow HOO\cdot + H\cdot$			91.3	
$NH_3 \rightarrow NH_2^{\cdot} + H\cdot$	83	110	97.3	116
$CH_3NH_2 \rightarrow CH_3NH\cdot + H\cdot$			93.4	
$NH_2NH_2 \rightarrow NH_2NH\cdot + H\cdot$			89.7	

[a] MRCI result [62].

Table 19: Results of ZPE and D_0 (kcal/mol) for some dissociation reactions.

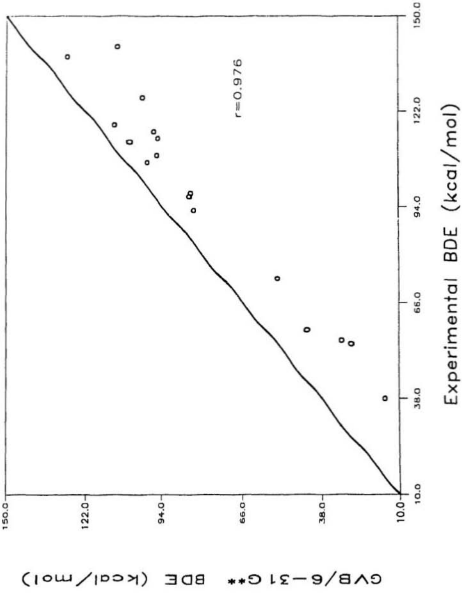
reaction[a]	ZPE	D_0(cal.)	D_0(expt.)
$CH_4 \rightarrow CH_3 + H$	9.2	97	104
$C_2H_4 \rightarrow C_2H_3 + H$	8.2	103	110
$C_2H_2 \rightarrow CCH + H$	6.8	121	131
$C_2H_6 \rightarrow CH_3 + CH_3$	9.0	76.3	88
$N_2H_4 \rightarrow NH_2 + NH_2$	7.2	47	65

[a] D_0(cal.) is our GVB/6-31G** results. ZPE and D_0 (expt.) are from [62-63].

Table 20: A-B bond dissociation energies (kcal/mol).

Bond dissociation reaction	HF/6-31G** //HF/6-31G* [5]	MP2/6-31G** //HF/6-31G* [5]	GVB/6-31G** //GVB/6-31G** this study	expt. [5]
$CH_3CH_3 \rightarrow {}^{\cdot}CH_3 + {}^{\cdot}CH_3$	69	99	85.3	97
$CH_3NH_2 \rightarrow {}^{\cdot}CH_3 + {}^{\cdot}NH_2$	58	93	83.6	93
$CH_3OH \rightarrow {}^{\cdot}CH_3 + {}^{\cdot}OH$	58	98	84.8	98
$CH_3F \rightarrow {}^{\cdot}CH_3 + {}^{\cdot}F$	69	113	95.7	114
$NH_2NH_2 \rightarrow {}^{\cdot}NH_2 + {}^{\cdot}NH_2$	34	73	54.3	73
$HOOH \rightarrow HO^{\cdot} + HO^{\cdot}$	0	53	29.9	55
$HOF \rightarrow HO^{\cdot} + F^{\cdot}$	-11	48	26.6	54
$F_2 \rightarrow F^{\cdot} + F^{\cdot}$	-33	35	14.8	38

Figure 10: Comparison of the GVB/6-31G** D_e with the experimental data. The line drawn in this plot is of unit slope.


r=0.976
Experimental BDE (kcal/mol)
GVB/6-31G** BDE (kcal/mol)
10.0
38.0
66.0
94.0
122.0
150.0

5.2.4 Energies of Hydrogenation Reactions

To examine further how effective GVB is in describing correlation energies and reaction energies, heats of reaction (ΔE) were calculated for a series of hydrogenation reactions. For the hydrogenation reactions studied, H_2 is added to the extent required for complete reduction of a given molecule to simple hydrides. For example, the complete hydrogenation of formaldehyde is

$$CH_2{=}O + 2H_2 \rightarrow CH_4 + H_2O$$

A number of ab initio calculations for hydrogenation energies have been previously reported. HF and MP2 results are systematically discussed in Ref. [5]. MP2 results are generally in very good agreement with experiment. We have calculated, for the first time, hydrogenation energies at the GVB/6-31G**//GVB/6-31G** level, including all bonds and valence lone pairs for a selection of molecules comprising first- and second-row elements. The GVB results are listed in Table 21 along with the available HF, MP2 and experimental values [5].

In general, the GVB hydrogenation energies are in good agreement with the experimental values. In most cases, the deviations from experiment are small, and very close to those of HF and MP2. For first-row molecules, the GVB results are about 0-10 kcal/mol in error, with a mean absolute deviation of 6 kcal/mol. The errors in GVB ΔE values are 0-5 kcal/mol for systems with only single bonds with the exception of F_2, while they are 0-10 kcal/mol for systems with multiple bonds with the exception of HCCH. The worst cases are for F_2 and HCCH, where the ΔE values are 20 and 18 kcal/mol in error, respectively. For second-row molecules, the results of GVB ΔE calculations are about 0-20 kcal/mol in error, with a mean absolute deviation of 11 kcal/mol, which is larger than for the first-row systems. For example, the GVB hydrogenation energy for HCP is in error by 28 kcal/mol, compared to 7 kcal/mol for HCN. GVB ΔE values are very close to HF and MP2 results for single bond systems. The errors relative to the experimental values are 0-10 kcal/mol with the exception of SiH_3F and SiH_3Cl where the errors are about 20 kcal/mol. GVB ΔE values are improved over HF results for systems with second-row multiple bonds. For example, the GVB hydrogenation energy for P_2 is only 6 kcal/mol in error, whereas 23 kcal/mol for HF and 24 kcal/mol for MP2. For PN and HCP, the GVB hydrogenation energies are in error by 27 and 28 kcal/mol, while 42 and 33 kcal/mol for HF, respectively.

The error in the GVB heats of reaction is dependent on the GVB evaluation of the bond dissociation energies. The error can be expressed as,

$$\Delta E_{GVB} - \Delta E_{expt.} = \sum \nu_i [D_e(GVB) - D_e(expt.)]_i \tag{122}$$

where ν_i is the stoichiometry coefficients, and $[D_e(GVB) - D_e(expt.)]_i$ is the discrepancy between the GVB D_e and the experimental D_e for a specific bond which is broken in the reaction. For example, $[D_e(GVB) - D_e(expt.)]$ for H_2, F_2, and HF are 17, 23, and 31 kcal/mol, respectively. The error in the GVB evaluation of hydrogenation energy for F_2 is 20 kcal/mol, which is very close to the result of 22 kcal/mol obtained from Eq. (122).

Table 21: The calculated energies of hydrogenation reactions (kcal/mol).

Reaction	HF/6-31G** //HF/6-31G* [5]	GVB/6-31G** //GVB/6-31G** this study	MP2/6-31G** //HF/6-31G* [5]	Expt. [5]
$CH_3\text{-}CH_3 + H_2 \rightarrow 2\,CH_4$	-18	-19	-15	16(-19)
$CH_3\text{-}NH_2 + H_2 \rightarrow CH_4 + NH_3$	-25	-26	-22	-23(-26)
$CH_3\text{-}OH + H_2 \rightarrow CH_4 + H_2O$	-27	-27	-25	27(-30)
$CH_3\text{-}F + H_2 \rightarrow CH_4 + HF$	-25	-24	-24	-27(-29)
$NH_2\text{-}NH_2 + H_2 \rightarrow 2NH_3$	-45	-45	-43	-45(-48)
$HO\text{-}OH + H_2 \rightarrow 2H_2O$	-84	-76	-80	83(-86)
$F\text{-}F + H_2 \rightarrow 2HF$	-131	-110	-123	-130(-133)
$CH_2{=}CH_2 + 2H_2 \rightarrow 2CH_4$	-55	-56	-52	48(-57)
$CH_2{=}NH + 2H_2 \rightarrow CH_4 + NH_3$	-53	-52	-49	-55(-64)
$CH_2{=}0 + 2H_2 \rightarrow CH_4 + H_2O$	-49	-49	-46	-50(-50)
$HN{=}NH + 2H_2 \rightarrow 2NH_3$	-70	-67	-66	-58(-68)
$HN0 + 2H_2 \rightarrow H_2O + NH_3$	-95	-101	-90	-93(-103)
$HCCH + 3H_2 \rightarrow 2CH_4$	-103	-108	-96	-90(-105)
$CO + 3H_2 \rightarrow CH_4 + H_20$	-45	-55	-44	-49(-63)
$N_2 + 3H_2 \rightarrow 2NH_3$	-18	-23	-13	-22(-37)

Table 21. (continued)

Reaction	HF/6-31G** //HF/6-31G* [5]	GVB/6-31G** //GVB/6-31G** this study	MP2/6-31G** //HF/6-31G* [5]	Expt. [5]
CH_3-SiH_3 + H_2 → CH_4 + SiH_4	-10	-11	-3	-3(-6)
CH_3-Cl + H_2 → CH_4 + HCl	-20	24	-14	-20(-22)
HOCl + H_2 → H_2O + HCl	-67	-62	-59	-62(-59)
SiH_3-F + H_2 → SiH_4 + HF	31	26	36	49(48)
ClF + H_2 → HCl + HF	-75	-71	-67	-75(-73)
Na_2 + H_2 → 2NaH	37	37	39	26(29)
SiH_3-SiH_3 + H_2 → 2SiH_4	-10	-9	-1	-3(-5)
SiH_3-Cl + H_2 → SiH_4 + HCl	17	13	22	34(33)
PH_2-PH_2 + H_2 → 2PH_3	-9	-11	-1	-2(-4)
HSSH + H_2 → 2H_2S	-20	-22	-12	-13(-14)
Cl_2 + H_2 → 2HCl	-53	-51	-46	-44(-41)
CH_2=S + 2H_2 → CH_4 + H_2S	-57	-55	-42	-47(-54)
HCP + 3H_2 → CH_4 + PH_3	-98	-83	-57	-56(-67)
PN + 3H_2 → PH_3 + NH_3	-67	-62	-29	-34(-45)
P_2 + 3H_2 → 2PH_3	-51	-44	-14	-38(-44)

The experimental data after ZPE correction are in parentheses.

The errors in the GVB hydrogenation energies for systems with multiple bonds are larger than for systems with only single bonds. GVB does not describe bonds in a consistent manner. In general, the large differential correlation effect for an isogyric[12] reaction, such as a hydrogenation reaction, leads to larger error in the calculated heats of reaction. For an isodesmic reaction,[13] the differential correlation effect is smaller and the reaction energies are generally better described.

5.2.5 Energies of Reactions Converting Multiple to Single Bonds

The energies of reactions converting multiple to single bonds are given in Table 22, for compounds containing first-row elements only. In this type of reaction, the unsaturated systems are converted into molecules containing single bonds between the same pairs of atoms. For example,

$$H_2C{=}O + CH_4 + H_2O \rightarrow 2\ CH_3\text{-}OH$$

is a reaction converting a C=O double bond to a C-O single bond.

Since the process of converting multiple to single bonds is isodesmic, the differential correlation effect is smaller as compared to hydrogenation reactions. For the direct evaluation of heats of reaction (ΔE) in this type of reaction, the agreement between the GVB and experimental data is very good. These reactions are better

[12] In an isogyric reaction, the number of electron pairs is conserved.
[13] In an isodesmic reaction, not only is the number of electron pairs held constant but also the chemical bond types are conserved.

described than the hydrogenation reactions. GVB ΔE values are 0-10 kcal/mol in error with a mean absolute deviation of only 4 kcal/mol. For example,

$$HCCH + 4CH_4 \rightarrow 3\ CH_3\text{-}CH_3$$

the GVB ΔE is 5 kcal/mol in error. The differential correlation effect in this reaction is much smaller than for the hydrogenation of HCCH,

$$HCCH + 3H_2 \rightarrow 2CH_4$$

where three H-H bonds and a C-C triple bond are totally broken to form six new C-H bonds.

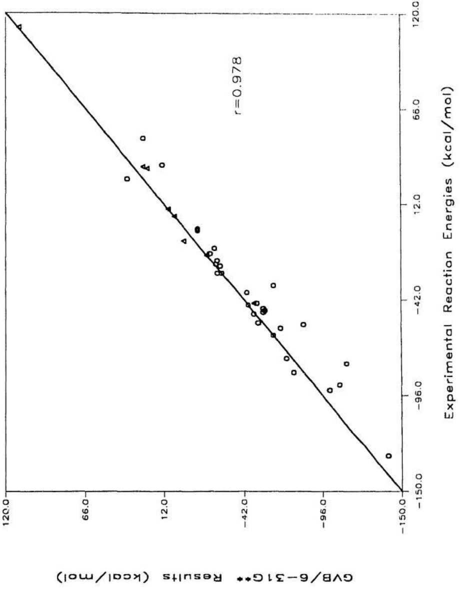
Figure 11 illustrates the correlation of GVB results and experimental data, for all the heats of reaction calculated. The correlation coefficient is 0.978, showing that the GVB results are in reasonably good agreement with the experimental data. A regression equation is obtained as $\Delta E(expt) = 0.936\Delta E(GVB) + 4.08$ kcal/mol, which can be used to correct GVB ΔE results.

Table 22: The calculated energies (kcal/mol) for reactions converting multiple to single bonds.

Reaction	HF/6-31G** //HF/6-31G* [5]	GVB/6-31G** //GVB/6-31G** this study	MP2/6-31G** //HF/6-31G* [5]	Expt. [5]
$H_2C{=}CH_2 + 2\ CH_4 \rightarrow 2\ CH_3\text{-}CH_3$	-19	-17	-23	-17(-20)
$H_2C{=}NH + CH_4 + NH_3 \rightarrow 2\ CH_3\text{-}NH_2$	-4	-1	-5	-9(-12)
$H_2C{=}O + CH_4 + H_2O \rightarrow 2\ CH_3\text{-}OH$	5	5	6	5(1)
$HN{=}NH + 2\ NH_3 \rightarrow 2\ NH_2\text{-}NH_2$	20	23	20	32(28)
$HCCH + 4\ CH_4 \rightarrow 3\ CH_3\text{-}CH_3$	-49	-49	-54	-44(-49)
$HCN + 2\ CH_4 + 2NH_3 \rightarrow 3\ CH_3\text{-}NH_2$	9	9	9	9(4)
$CO + 2\ CH_4 + 2H_2O \rightarrow 3\ CH_3\text{-}OH$	36	26	33	33(27)
$N_2 + 4\ NH_3 \rightarrow 3\ NH_2\text{-}NH_2$	115	111	114	112(107)

The experimental data after ZPE correction are in parentheses.

Figure 11: Comparison of the GVB/6-31G** reaction energies (ΔE) with the experimental data.
The line drawn in this plot is of unit slope.
Triangles are used for reactions converting multiple to single bonds; circles are used for hydrogenation reactions.


r=0.978
Experimental Reaction Energies (kcal/mol)
GVB/6-31G** Results (kcal/mol)
120.0
66.0
12.0
-42.0
-96.0
-150.0

5.3 π Bond Energies

Double bond energies are often divided into σ and π contributions, but these are not directly observable quantities. Ab initio molecular orbital theory affords an excellent way to assess these components. There are two ways to obtain the π bond energies.

The first procedure is to compare the bond dissociation energies of double bonds, D_e(X=Y), and single bond, D_e(X-Y). However, the energy difference, D_e(X=Y)-D_e(X-Y), does not correspond to the barrier to rotation about the X=Y bond. For example, the experimental data [74] for ethylene are 173.3 and 86.1 kcal/mol for D_e(X=Y) and D_e(X-Y), respectively. The difference, 87.2 kcal/mol, is far greater than the rotational barrier (65 kcal/mol) [75]. Obviously, the directly deduced quantity from D_e(X=Y)-D_e(X-Y) is not the π bond energy in ethylene. Schleyer et al. [76] used isogyric reactions,

$$H_2X{=}YH_n + XH_4 + YH_{n+2} \rightarrow 2\ H_3X\text{-}YH_{n+1}$$

to compare the energies of X=Y double bonds with those of X-Y single bonds. For ethylene, the π bond energy (E_π) of 69.6 kcal/mol was obtained at MP4/6-31G*//6-31G*+ZPE. Meanwhile, Schmidt et al. [77] used hydrogenation thermochemical cycles to give similar results, *e.g.*, E_π(ethylene)=62 kcal/mol. However, there are some discrepancies between the calculated π bond energies of Schmidt et al. and those of Schleyer et al. For instance, the calculated C=N π bond energy by Schmidt et al. is 65 kcal/mol, which is smaller than 81 kcal/mol obtained by Schleyer et al. The discrepancies are due to the use of different sets of reference species in the thermochemical cycles.

Another commonly used procedure is to equate the π bond energy directly to the rotational barrier [75, 77-80], provided that the transition state has no π bond character while retaining all σ bonds. The transition state, for singlet rotation around a double bond, is the 90^0 twisted singlet diradical. For this diradical structure, a single determinant wavefunction cannot afford a proper description, and a multiconfiguration calculation must be used. Due to the difficulty in characterizing the twisted singlet transition state which is very sensitive to the basis set and correlation energy, few efforts have been made to obtain the optimized geometries with extensive basis sets. Schmidt et al. [77] reported a MCSCF/6-31G*//3-21G* rotational barrier, or π bond energy, for ethylene of 65.4 kcal/mol. Borden et al. [48,79] carried out GVB-PP(1/2)/6-31G* calculations of the rotational barrier in ethylene (65.6 kcal/mol), and TCSCF (two configuration SCF) calculations of the rotational barrier in disilene (22.1 kcal/mol). Our preliminary GVB(2/4)/6-31G** calculations [39] give a proper description of the rotation about the π bond in C_2H_4, and give a π bond energy of 63.6 kcal/mol. The results are illustrated in Figure 4, and have been discussed in Section 4.5.

The objective of this part of this work is to characterize the transition state for the singlet rotation around C=C, C=Si and Si=Si double bonds, by employing GVB calculations. The π bond energies and the fluorine substitution effects are further discussed.

5.3.1 Computational Method

In this study, GVB(6/12)/6-31G** calculations were carried out for $A_2X{=}YB_2$ (A, B=H, F; X, Y=C, Si). The GVB(6/12) wavefunction includes two X-A bonds, two B-Y bonds, one X-Y σ and one X-Y π bond, *i.e.*,

$$\Psi^{GVB}_{A_2X=YB_2} = \mathcal{A}[\Psi_{core}\psi_{X-Y(\sigma)}\psi_{X-Y(\pi)}\psi_{X-A1}\psi_{X-A2}\psi_{Y-B1}\psi_{Y-B2}] \qquad (123)$$

The planar or trans-bent structures of $A_2X{=}YB_2$ were optimized with the OC method [26]. The singlet transition state structures were optimized with the VA method [27] (*c.f.* Section 2.1.5). All equilibrium structures were characterized to have no imaginary frequency, while the transition states have only one imaginary frequency. In the transition state, the GVB CI-coefficients for π orbitals are $0.707(\frac{1}{\sqrt{2}}) \pm 0.005$.

When a doubly-bonded $A_2X{=}YB_2$ molecule is rotated by 90^0 to the diradical transition state, the X-Y π bond is broken while retaining all σ bonds. The barrier to X=Y double bond rotation can be defined as the π bond energy, E_π. π bond energies reported here are not corrected for zero-point energy (ZPE).[14]

GVB total energies for ground states and transition states, and rotational barriers are summarized in Table 23. Optimized geometrical parameters are given in Tables 24 and 25 for ground states and transition states, respectively, where r_1, r_2 and r_3 are

[14] The MCSCF/3-21G* [77] ZPE correction for the π bond energies in $CH_2{=}CH_2$, $CH_2{=}SiH_2$, and $SiH_2{=}SiH_2$, are 4.2, 1.4, and 0.6 kcal/mol, respectively.

the X-Y, A_1-X and B_1-Y bond lengths, θ_1 and θ_2 are the A_1-X-Y and B_1-Y-X bond angles. ϕ_1, ϕ_2 and ϕ_3 are the A_2-X-Y-A_1, B_1-Y-X-A_1 and B_2-Y-X-B_1 torsion angles. The X-Y bond orders [120], and Mulliken charges on X and Y, are also listed in Table 26.

5.3.2 The C=C Bond

The C-C π bond in ethylene is no doubt the best characterized π bond, both experimentally and theoretically. A number of theoretical papers [77, 81-82] have indicated that the triplet state and rotated singlet state represent a violation of Hund's rule. In accordance with Hund's rule, the two unpaired π electrons should give the triplet ground state. GVB treatments [81a], along with MCSCF [77] and CI [82], predict that the lowest energy rotation path on the singlet potential energy surface is entirely below the triplet surface for ethylene. In other $A_2X=YB_2$ molecules, (A, B=H, F; X, Y=C, Si), the 90^0 twisted triplet is about 1-3 kcal/mol lower than the rotated singlets, with geometries similar to their singlets.

A. Optimized Geometries

For planar $H_2C=CH_2$, $H_2C=CF_2$ and $F_2C=CF_2$, the GVB geometries are very close to the CISD [83(d)] and MCSCF [77] results. The C=C and C-F bond lengths decrease with increasing fluorine substitution, as observed from early HF calculations [83(d)].

The twisted structure of ethylene is illustrated in Figure 12(a). For twisted

$H_2C{=}CH_2$, $H_2C{=}CF_2$ and $F_2C{=}CF_2$, the GVB C-C bond lengths are 1.492, 1.490, and 1.501 Å, respectively, which are very close to Schaefer's CI C-C bond length of 1.494 Å for $H_2C{=}CH_2$ [84(a)]. The C-C bond length in the twisted structure is close to the single C-C bond length of 1.531 Å in C_2H_6 obtained at full valence GVB/6-31G** [42].

The F-C-C bond angle (θ_2 in Tables 24 and 25) clearly shows that CF_2 is changed from a planar (sp^2 hybridization, $\sim 120^0$) conformation to a tetrahedral (sp^3 hybridization, $\sim 113^0$) conformation, indicating the pyramidalization of CF_2 in the twisted structure.

B. π Bond Energy

Our GVB/6-31G** π bond energies (E_π) are in very good agreement with the previous ab initio calculations and the available experimental results. For instance, E_π(ethylene) is 65.4 kcal/mol, compared to 67.9 (Schaefer's CI [84(a)]), 65.4 (MCSCF [77]), 65.6 (GVB(1/2) [48]), and 64 kcal/mol (experimental [75]). For fluoroethylene, E_π contains a contribution from the pyramidalization of CF_2. The intrinsic π bond energy is the calculated result without pyramidalization. If the CF_2 radical center is constrained to be planar in the diradical 1,1-difluoroethylene, E_π is higher by 11.5 kcal/mol than for ethylene, indicating that the π bond in 1,1-difluoroethylene is intrinsically stronger. In general, the pyramidalization of the carbon or silicon center from $sp^2(120^0)$ to $sp^3(109.5^0)$ stabilizes the diradical transition state. The value of E_π is in fact related to the B-Y-X bond angle in the transition state, *i.e.*, θ_2 in Table 25.

For example, θ_2 are 121.1^0, 115.2^0, 113.9^0 for $H_2{=}CH_2$, $CH_2{=}CF_2$, and $CF_2{=}CF_2$, respectively. Their deviations from 109.5^0 are 10.6^0, 5.7^0, and 4.4^0, which correlate well with the calculated π bond energies of 65.1, 63.2, and 51.7 kcal/mol. With two carbon centers pyramidalized, the rotational barrier in tetrafluoroethylene is much lower. The C-C π bond energy decreases with increasing fluorine substitution.

For the isodesmic reaction,

$$CH_2{=}CH_2 + CF_2{=}CF_2 \rightarrow 2\ CH_2{=}CF_2$$

the GVB reaction energy (ΔE) is -16.5 kcal/mol, which is in excellent agreement with the experimental value of -14.6 kcal/mol [48]. The change from E_π contribution is -9.3 kcal/mol, consistent with the reaction energy. The exothermicity of this reaction can be attributed to the replacement of the homopolar σ and π C-C bonds in ethylene and the tetrafluoroethylene with the heteropolar σ and π C-C bonds in 1,1-difluoroethylene.

The C-C bond orders (Table 26) in the twisted structures are ~0.9, demonstrating that the C-C π bond is totally broken. The C=C double bond orders are 1.778, 1.738 and 1.670 for planar $CH_2{=}CH_2$, $CH_2{=}CF_2$ and $CF_2{=}CF_2$, respectively. The differences between the C=C double bond orders in the planar ground states and the C-C single bond orders in the transition states are 0.812, 0.788 and 0.767, respectively, which are measures of π bond strengths.

Table 23: GVB(6/12)/6-31G** total energies (hartrees) and π bond energies (kcal/mol).

Molecule	Total energy		E_π	E_π
	Planar	Twisted	(this work)	reference
$H_2C=CH_2$	-78.13651	-78.03224	65.4	65.4[77], 69.6[76]
$H_2C=SiH_2$	-329.13391	-329.08352	31.6	35.6[76], 36.1[77]
$H_2Si=SiH_2$	-580.16608	-580.12900	23.2	22.7[77], 24.2[76]
$H_2C=CF_2$	-275.85207	-275.75140	63.2	63.3[48]
$F_2C=CF_2$	-473.54140	-473.45909	51.7	50.3[48]
$F_2C=SiH_2$	-526.83074	-526.79163	24.6	-
$H_2C=SiF_2$	-526.96988	-526.93590	21.3	-

5.3.3 The Si=Si Bond

Because disilene (Si_2H_4) is the simplest species with a Si-Si double bond and its ground state shows an unusual preference for a trans-bent C_{2h} structure rather than a planar D_{2h} structure, it has been the subject of much interest experimentally and theoretically since its isolation in 1981 [77, 79, 81(b), 83(a-b), 85(a)].

A. Optimized Geometries

The trans-bent singlet disilene is more stable than the planar structure. Our GVB equilibrium geometry is very close to the TCSCF/6-31G* result [79]. For the $H_2Si{=}SiH_2$ trans-bent structure (C_{2h}) illustrated in Figure 12(c), the Si-Si distance is 2.240 Å, compared to 2.160 Å (experimental [86]), 2.217 Å (MCSCF/3-21G* [77]), 2.242 Å (TCSCF/6-31G* [79]), 2.248 Å (GVB(2/4)/6-31G** [83(b)]), and 2.236 Å (CASSCF/6-31G** [83(b)]). Such a long bond length reflects the diradical character of the Si-Si bond, even though it is still shorter than that of a single Si-Si bond. However, some double bond character still exists in disilene (C_{2h}) as indicated by the Si-Si bond order of 1.380. For the twisted (C_2) structure illustrated in Figure 12(d), our GVB Si-Si single bond is 2.358 Å, compared to 2.351 Å (TCSCF/6-31G* [79]) and 2.355 Å (MCSCF/3-21G* [77]).

B. π Bond Energy

The energetics of disilene states are significantly affected by the incorporation of electron correlation corrections. The calculated GVB barrier to rotation is 23.2

kcal/mol, which is close to 22.1 (TCSCF/6-31G* [79]), 22.7 (MCSCF/6-31G*//3-21G* [77]), and 25.0 kcal/mol (CISD/DZP [79]). The small Si-Si π bond energy is consistent with the small Si-Si bond order. It is noted from Table 26 that rotation slightly increases the positive charge at silicon in disilene, while it slightly increases the negative charge at carbon in ethylene.

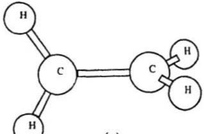
5.3.4 The C=Si Bond

Silaethylenes have long been implicated as reactive intermediates. 1,1-Dimethylsilaethylene (DMSE) is the simplest silaethylene detected in the gas-phase. Silaethylene differs from ethylene in that the rotated singlet will be pyramidalized at the silicon.

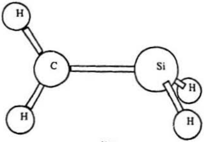
A. Optimized Geometries

For silaethylene $H_2C{=}SiH_2$, the C-Si distance is 1.734 Å, compared to the experimental value of 1.70 - 1.80 Å [101] and 1.743 Å from a MCSCF/3-21G* calculation [77]. For planar (C_{2v}) $F_2C{=}SiH_2$ and $H_2C{=}SiF_2$, the C-Si bond lengths are 1.749 and 1.688 Å, respectively. For the twisted conformation (C_s), the GVB C-Si bond lengths are 1.884, 1.912, and 1.858 Å, for $H_2C{=}SiH_2$, $F_2C{=}SiH_2$, and $H_2C{=}SiF_2$, respectively. Due to the strong dipolar (ionic) interaction between CH_2 and SiF_2 in $H_2C{=}SiF_2$, its C=Si double bond is very short. The C-Si bond in the twisted structure is also a short single bond.

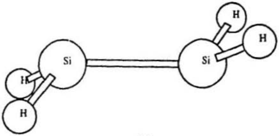
Figure 12: The structure of planar and perpendicular conformations of $A_2X=YB_2$ (a) twisted C_2H_4; (b) twisted $CH_2{=}SiH_2$; (c) trans-bent Si_2H_4; (d) twisted Si_2H_4.



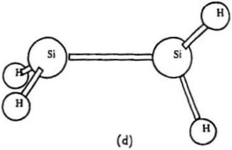
(a)



(b)



(c)



(d)

Table 24: GVB(6/12)/6-31G** ground state geometries for A_2X YB_2.

Molecule	r_1	r_2	r_3	θ_1	θ_2	ϕ_1	ϕ_2	ϕ_3
$H_2C=CH_2$	1.346	1.093	1.093	121.7	121.7	180.0	0.0	180.0
$H_2C=SiH_2$	1.734	1.093	1.486	122.4	122.1	180.0	0.0	180.0
$H_2Si=SiH_2$	2.240	1.498	1.498	116.3	116.3	131.5	48.5	131.5
$H_2C=CF_2$	1.327	1.087	1.329	119.9	125.1	180.0	0.0	180.0
$F_2C=CF_2$	1.324	1.325	1.325	123.4	123.4	180.0	0.0	180.0
$F_2C=SiH_2$	1.749	1.339	1.477	125.3	118.7	180.0	0.0	180.0
$H_2C=SiF_2$	1.688	1.091	1.586	121.2	126.6	180.0	0.0	180.0

Bond lengths are in Å, and bond angles in degrees.

Table 25: GVB(6/12)/6-31G** transition state geometries for $A_2X=YB_2$.

Molecule	r_1	r_2	r_3	θ_1	θ_2	ϕ_1	ϕ_2	ϕ_3
$H_2C=CH_2$	1.492	1.091	1.091	121.1	121.1	180.0	90.0	180.0
$H_2C=SiH_2$	1.884	1.093	1.496	122.7	112.1	180.0	118.1	123.8
$H_2Si=SiH_2$	2.358	1.496	1.497	112.5	114.7	126.7	161.4	126.8
$H_2C=CF_2$	1.490	1.087	1.341	119.9	115.2	180.0	115.7	128.6
$F_2C=CF_2$	1.501	1.330	1.330	113.9	113.9	129.0	155.4	129.0
$F_2C=SiH_2$	1.912	1.346	1.489	117.8	110.8	134.3	156.8	123.3
$H_2C=SiF_2$	1.858	1.093	1.603	122.5	109.7	180.0	121.8	116.4

(a) Bond lengths are in Å, and bond angles in degrees.
(b) In the transition state, A_1-X and A_2-X bond lengths, and the related bond angles, are not equal but very close. Only the above parameters are listed.

The structure of twisted silaethylene is illustrated in Figure 12(b). Our GVB H-Si-C bond angle (θ_2 in Table 25) in the twisted structure (112.1^0) is different from the previous MCSCF/3-21G* result (118.0^0) [77], whereas other geometrical parameters are very close. The comparison suggests that the H-Si-C bond angle in Ref. [77] is incorrectly reported.

B. π Bond Energy

The intrinsic (without pyramidalization at silicon) π bond energy of H_2C=SiH_2 from our GVB calculations is 40.7 kcal/mol. Pyramidalization of the silicon center from $sp^2(120^0)$ to $sp^3(109.5^0)$ stabilizes the diradical transition state and decreases the π bond energy. Our GVB(6/12)/6-31G** E_π(silaethylene) is 31.6 kcal/mol, which is close to the experimental result of 34±4 kcal/mol [81(c)] and the MCSCF/6-31G*//3-21G* result of 35.6 kcal/mol [77]. E_π(silaethylene) is less than the experimental estimate of 41±5 kcal/mol for 1,1-Dimethyl-silaethylene [87]. However, the Goddard et al. GVB(2/4) result of 59.9 kcal/mol [81(c)] is too high, because the geometries were not optimized at the GVB level and C=SiH_2 was constrained to be planar in the twisted conformation.

Because of silicon pyramidalization, the GVB H(F)-Si-C bond angles (θ_2 in Table 25) in the transition states are 112.1^0, 110.8^0 and 109.7^0 for silaethylene, 2,2-difluorosilaethylene and 1,1-difluorosilaethylene, respectively. Their deviations from 109.5^0 are 2.6^0, 1.3^0, and 0.2^0, which correlate well with the π bond energies of 31.6, 24.6, and 21.3 kcal/mol.

For the following reaction

$$H_2C=CH_2 + H_2Si=SiH_2 \rightarrow 2\ H_2C=SiH_2$$

the GVB reaction energy (ΔE) is 21.8 kcal/mol compared to 25.4 kcal/mol calculated using the difference between the GVB π bond energy. For

$$H_2Si=SiH_2 + F_2C=CF_2 \rightarrow 2\ F_2C=SiH_2$$

the GVB reaction energy (ΔE) is 28.9 kcal/mol compared to 25.7 kcal/mol calculated using the difference between the GVB π bond energy. The π bond change is the dominant contribution to the reaction energy in those two reactions.

According to the above discussion, the pyramidalization at the carbon or silicon center for the twisted structures decreases the π bond energies in the substituted ethylenes and their silicon counterparts. Fluorine substitution stabilizes both diradical and dipolar twisted singlet structures.

Table 26: GVB(6/12)/6-31G** results of X=Y bond orders and Mulliken charges at atoms X and Y for planar and twisted singlets.

Molecule	planar			twisted		
	Q_X	Q_Y	X=Y bond order	Q_X	Q_Y	X=Y bond order
$H_2C=CH_2$	-0.214	-0.214	1.778	-0.227	-0.227	0.966
$H_2C=SiH_2$	-0.446	0.431	1.709	-0.017	0.337	0.899
$H_2Si=SiH_2$	0.241	0.241	1.380	0.250	0.250	0.863
$H_2C=CF_2$	-0.355	0.716	1.738	-0.292	0.679	0.950
$F_2C=CF_2$	0.622	0.622	1.670	0.637	0.637	0.903
$F_2C=SiH_2$	0.486	0.317	1.576	0.423	0.499	0.837
$H_2C=SiF_2$	-0.556	1.192	1.771	-0.494	1.165	0.881

5.4 Singlet and Triplet Structures of 1,1-Dilithioethylene $CH_2{=}CLi_2$, 2,2-Dilithiosilaethylene $SiH_2{=}CLi_2$ and 1,1-Dilithiosilaethylene $CH_2{=}SiLi_2$

The GVB(6/12)/6-31G** rotational barrier around the C=C double bond in ethylene (65.4 kcal/mol [39]) is in excellent agreement with the experimental value (64 kcal/mol [75]). The π bond is totally broken in the twisted transition state, and the rotational barrier can be regarded as the π bond energy (or π bond strength). This barrier can be reduced significantly by π donor or π acceptor substituents at carbon, which will preferentially stabilize either the diradical or especially the dipolar configuration of the twisted form by electron delocalization, as for fluorine in the tetrafluoroethylenes studied in Section 5.3.

Apeloig et. al. [88] first discovered the unexpected properties of 1,1-dilithioethylene, $CH_2{=}CLi_2$, by ab initio calculations. They noted that, not only is the rotational barrier about the C=C double bond low, but the ground state equilibrium geometry is the triplet perpendicular structure. Their results have been confirmed by Nagase and Morokuma [89], and Laidig and Schaefer [90], and more importantly the related molecules have already been prepared in the laboratory [91]. The surprising preference for the perpendicular structure over the planar structure for both singlet and triplet has been interpreted as being due to a combination of strong hyperconjugation involving the C-Li bonds and the formation of a delocalized three-center Li-C-Li bond. Other olefinic systems [92] involving electron deficient substituents, such as $CH_2{=}C(BH)_2$, have also been found to prefer the perpendicular structure while retaining the C=C double bond. Recently, Schaefer et al. [93] have surveyed the potential energy surface for the singlet 1,2-dilithioethylene and acetylenic $C_2H_2Li_2$ using CCSD/DZP.

Since the structural parameters in these systems are very sensitive to basis set and electron correlation, it is important to obtain the optimized geometries with the inclusion of electron correlation using large basis sets. The objective of this study is to examine further whether the special stability in the triplet or perpendicular 1,1-dilithioethylene is due to the strong π bond between CH_2 and the Li-C-Li three-center group, or a complex of Li_2 and the vinylidene.

5.4.1 Computational Method

The full valence GVB/6-31G** calculations were carried out for $H_2X{=}YLi_2$, with three (X, Y) combinations as (C, C), (Si, C) and (C, Si). In this scheme, GVB(6/12) was used for all singlets, which includes X-Y σ and π bonds with two Y-Li and two X-H bonds, *i.e.*,

$$\Psi_S^{GVB} = \mathcal{A}[\Psi_{core}\psi_{X-Y(\sigma)}\psi_{X-Y(\pi)}\psi_{X-H1}\psi_{X-H2}\psi_{Y-Li1}\psi_{Y-Li2}] \qquad (124)$$

and GVB(5/10) was used for all triplets, which includes X-Y σ bond with two Y-Li and two X-H bonds, *i.e*,

$$\Psi_T^{GVB} = \mathcal{A}[\Psi_{core}\psi_{X-Y(\sigma)}\psi_{X-H1}\psi_{X-H2}\psi_{Y-Li1}\psi_{Y-Li2}\Psi_{open}] \qquad (125)$$

The geometries for both singlets and triplets were optimized using the OC method [26] (*c.f.* Section 2.1.5) without any symmetry constraint.[15]

5.4.2 Potential Curves and Relative Energies

The GVB total energies and relative energies are summarized in Table 27. Potential energy curves for rotation around the double bond are shown in Figures 13-15 for both singlets and triplets.

The GVB/6-31G** relative energies for 1,1-dilithioethylene are very close to Schaefer's CISD results [90]. The perpendicular triplet is predicted to be the minimum energy structure with the planar triplet lying only 1.33 kcal/mol higher. The singlets are about 13 kcal/mol higher than the triplets, and the singlet-triplet energy splitting decreases slightly from the planar (13.3 kcal/mol) to the perpendicular (12.6 kcal/mol) structure. However, the planar singlet is the lowest state of $H_2Si{=}CLi_2$, with a rotational barrier of 5.25 kcal/mol which is much smaller than that in $H_2C{=}SiH_2$, showing that a strong π bond still exists in the perpendicular structure. For $H_2C{=}SiLi_2$, the lowest state is also the planar singlet with a rotational barrier of 23.4 kcal/mol which is close to that of $H_2C{=}SiF_2$ (see Section 5.3).

For the C=Si systems, the singlets are all lower in energy than the triplets. In the process of rotation, the energies of singlets increase, whereas those of triplets decrease. The singlet-triplet energy splittings in the planar structures are therefore much higher

[15] For twisted structures, both X and Y centers are allowed to be pyramidalized during the geometry optimization.

than in the perpendicular structures, 13.1 versus 4.98 kcal/mol for $H_2Si{=}CLi_2$ and 35.9 versus 4.40 kcal/mol for $H_2C{=}SiLi_2$. For singlet planar structures, $H_2C{=}SiLi_2$ is more stable than $H_2Si{=}CLi_2$, and their energy difference decreases for the perpendicular structures. However, planar triplet $H_2Si{=}CLi_2$, which is more stable than planar triplet $H_2C{=}SiLi_2$, has strong C-Si π bond character with a bond order of 1.729 (Table 29).

The hyperconjugation of substituents can increase the X-Y π bond strength even in the twisted structures. Since the X-Y π bond still exists in the perpendicular singlet, it is impossible to obtain X-Y π bond energy from the rotational barrier of dilithium substituted X=Y doubly bonded system. The assumption that the π bond energy is equal to the rotational barrier (in Section 5.3) is only suitable when the π bond is totally broken in the perpendicular structure. This can be ascertained from the X-Y π orbital GVB CI-coefficient. For example, the GVB CI-coefficients for the C-C π and π^* orbitals are both $0.707(\frac{1}{\sqrt{2}})$ in the perpendicular singlet of ethylene, indicating that the π bond is totally broken. The GVB CI-coefficients for the C-C π and π^* orbitals of the perpendicular singlet of 1,1-dilithioethylene, however, are 0.979 and 0.205, demonstrating that the π bond still exists.

Table 27: GVB/6-31G** total energies (hartrees) and relative energies (kcal/mol).

Molecule	State	Planar	Perpendicular	$\Delta E(PP-P)$
CH_2CLi_2	S	-91.76803	-91.77129	-2.04
	T	-91.78925	-91.79137	-1.33
	$\Delta E(T-S)$	-13.3	-12.6	
SiH_2CLi_2	S	-342.81313	-342.80477	5.25
	T	-342.79227	-342.79683	-2.86
	$\Delta E(T-S)$	13.1	4.98	
CH_2SiLi_2	S	-342.84686	-342.80951	23.4
	T	-342.78964	-342.80249	-8.06
	$\Delta E(T-S)$	35.9	4.4	

S: for singlet; T: for triplet. $\Delta E(T-S) = E(T) - E(S)$: the energy gap between triplet and singlet; $\Delta E(PP-P) = E(PP) - E(P)$: the energy difference between the perpendicular and planar structures.

Figure 13: Potential energy curves for rotation around the $C=C$ in $CH_2=CLi_2$.

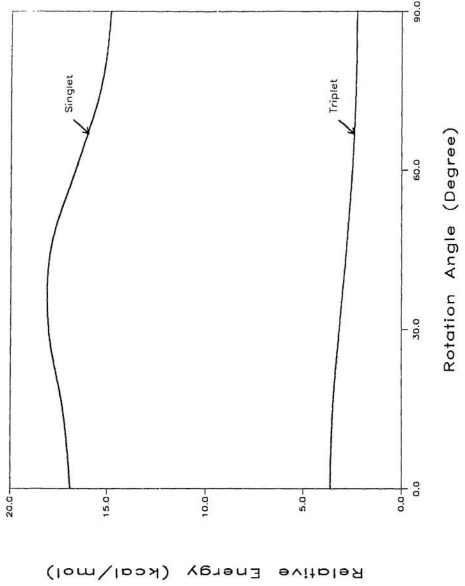

Singlet
Triplet
0.0
30.0
60.0
90.0
Rotation Angle (Degree)
0.0
5.0
10.0
15.0
20.0
Relative Energy (kcal/mol)

Figure 14: Potential energy curves for rotation around the Si=C in $SiH_2=CLi_2$.

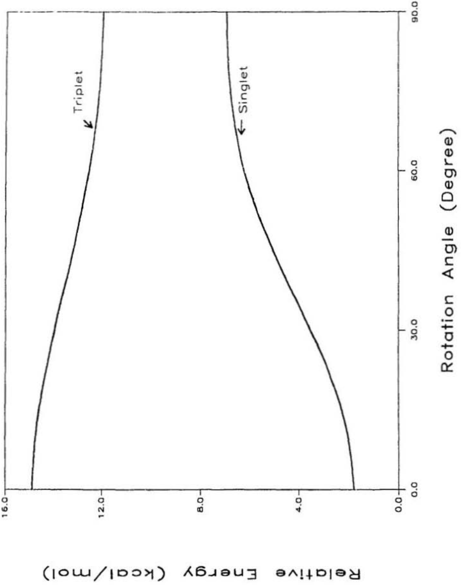
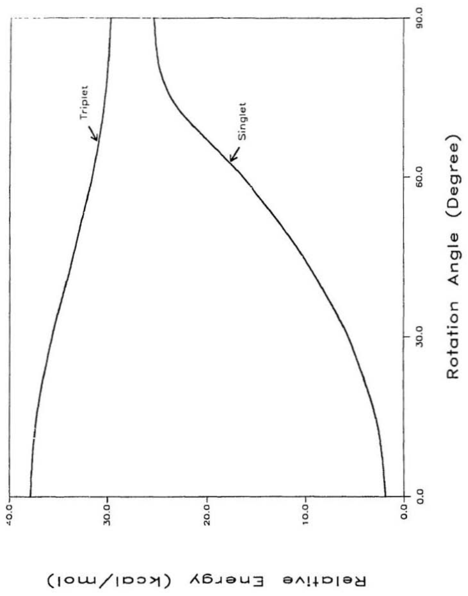

Triplet
Singlet
Rotation Angle (Degree)
Relative Energy (kcal/mol)
0.0
30.0
60.0
90.0
0.0
4.0
8.0
12.0
16.0

Figure 15: Potential energy curves for rotation around the C=Si in $CH_2 = SiLi_2$.


Triplet
Singlet
Rotation Angle (Degree)
Relative Energy (kcal/mol)
0.0
30.0
60.0
90.0
10.0
20.0
30.0
40.0

5.4.3 Optimized Geometries

The GVB/6-31G** optimized geometries for both singlets and triplets are listed in Table 28. The structures of $CH_2{=}CLi_2$ are illustrated in Figure 16.

For the planar singlet 1,1-dilithioethylene, the C-C, C-Li, and Li-Li bond lengths are 1.369, 2.042, and 3.573 Å, respectively, as compared to 1.356, 2.000, and 3.676 Å for HF/DZP[16] [90], showing an effect of electron correlation on the geometry. The GVB C-C, C-Li bond distances are longer, but the Li-Li bond distance shorter than the HF results. The GVB C-C bond length in planar 1,1-dilithioethylene is longer than that in ethylene (1.346 Å at GVB(6/12)/6-31G** [42]). In the perpendicular singlet, the C-C, C-Li, and Li-Li bond distances are shorter than in the planar singlet. The Li-C-Li bond angle is 99^0, 32^0 smaller than in the planar singlet, resulting in a much shorter Li-Li bond.

There are some special characteristics in the GVB geometries of triplet 1,1-dilithioethylene. First, the C-C bond is almost the same in both the planar and perpendicular structures, close to the equilibrium bond length in ethylene, and shorter than in the corresponding singlet, indicating that the C=C double bond becomes stronger in the triplet. Secondly, the Li-C-Li bond angle is about 74^0, with a Li-Li bond distance of ~2.5 Å which is shorter than that in the Li_2 molecule,[17] demonstrating a strong Li-Li bond in the triplet structure.

In 2,2-dilithiosilaethylene $H_2Si{=}CLi_2$, the C-Si bond distance (1.725 Å) is shorter

[16]Schaefer et al. DZP basis set is comparable to 6-31G*. The differences between our GVB and their HF geometry are due to the electron correlation rather than basis set.

[17]The experimental Li-Li bond distance in the Li_2 is 2.67 Å [42].

than in silaethylene (1.734 Å *c.f.* Section 5.3). For the perpendicular singlet, the C-Si bond distance (1.711 Å) is much shorter than in the twisted silaethylene (1.884 Å), and even shorter than in the planar structure. The shorter C-Si bond indicates that the π bond in $H_2Si{=}CLi_2$ is stronger than in $H_2Si{=}CH_2$, and it is even stronger in the perpendicular structure than in the planar structure. The C-Si bond distances in the planar triplets and the perpendicular triplets are 1.745 and 1.740 Å, respectively, which are longer than those in the corresponding singlets. Those distances are between a C-Si single bond and a C=Si double bond, indicating a weaker π bond in the triplet than that in the singlet. The Li-Li bond distance (2.5 Å) and the Li-C-Li bond angle (75^0) are similar to those in 1,1-dilithioethylene.

For 1,1-dilithiosilaethylene $CH_2{=}SiLi_2$, the structure is very different, compared to $H_2Si{=}CLi_2$. The C-Si bond length is about 1.80 Å for the singlet and 1.95 Å for the triplet, much longer than in silaethylene. In the triplet, there is a typical C-Si single bond.

In the perpendicular $Li_2C{=}SiH_2$ and $H_2C{=}SiLi_2$, the planarity (sp^2 hybridization) of $SiLi_2$ or SiH_2 is favoured and no silicon pyramidalization is found.

5.4.4 The Electronic Structures

The bond orders, Mulliken charges and dipole moments are calculated for both singlets and triplets in the planar and perpendicular forms. The results are listed in Table 29.

In $CH_2{=}CLi_2$ and $SiH_2{=}CLi_2$, the Li-Li bond order for the triplets (0.54) is larger than for the singlets (0.10), and the C-C and C-Si bond orders are also slightly larger for the triplets. However, those bond orders are smaller for triplet $H_2C{=}SiLi_2$.

For $H_2X{=}CLi_2$, the singlets are more dipolar-like, while the triplets are more diradical-like. Large dipole moments are obtained for the singlets (4.0-6.0D). However, the dipole moments for the triplets are much smaller (1.0-1.5D) and of the *opposite* sign, for $H_2C{=}CLi_2$ and $H_2Si{=}CLi_2$. The negative dipole moment implies $H_2C^+ ={}^-CLi_2$ polarity. This dipolar direction clearly shows that the π bond in the triplet increases the electron density at the CLi_2 group. Whereas, the dipole moment for triplet $H_2C{=}SiLi_2$ is close to that of the singlet.

The singlet and triplet diradical structures for perpendicular ethylene are very close, with the singlet lower in energy by about 0.6 kcal/mol (CI result [77]). However, the triplet is lower in energy with the singlet-triplet splitting of 12.6 kcal/mol for perpendicular 1,1-dilithioethylene. The energy preference for the triplet in 1,1-dilithioethylene is due to the strong π bond between CH_2 and the Li-C-Li three-center group.

In conclusion, the GVB relative energies are in excellent agreement with previous CISD results for 1,1-dilithioethylene. The perpendicular $H_2X{=}CLi_2$ conformation is only stabilized by a strong π bond between XH_2 and CLi_2. The X=C double bond increases the electron density at the electron deficient group CLi_2, leading to a Li-Li bond.

Table 28: GVB/6-31G** geometries for $H_2X{=}YLi_2$.

Molecule	State	X-Y	H-X	Y-Li	Li-Li	H-X-H	Li-Y-Li
CH_2CLi_2	S (Planar)	1.3691	1.1073	2.0427	3.5728	110.98	121.98
	S (Perpendicular)	1.3554	1.1047	1.8844	2.8688	113.14	99.14
	T (Planar)	1.3459	1.0981	2.0992	2.5126	115.02	73.52
	T (Perpendicular)	1.3473	1.0985	2.0563	2.4939	115.32	74.66
SiH_2CLi_2	S (Planar)	1.7246	1.5140	1.9502	3.5665	105.02	132.24
	S (Perpendicular)	1.7114	1.5098	1.8759	3.1316	105.22	113.40
	T (Planar)	1.7453	1.4919	2.0976	2.4942	113.06	72.96
	T (Perpendicular)	1.7404	1.4923	2.0720	2.5307	112.06	75.28
CH_2SiLi_2	S (Planar)	1.7965	1.0987	2.5367	4.7116	112.30	136.46
	S (Perpendicular)	1.7881	1.0965	2.3561	3.1494	113.64	83.88
	T (Planar)	1.9641	1.0960	2.4856	4.4794	113.72	128.60
	T (Perpendicular)	1.9101	1.0966	2.4957	3.7093	113.92	137.02

S: for singlet; T: for triplet. Bond lengths are in Å, and bond angles in degrees.

Figure 16: The structure of $CH_2{=}CLi_2$.
(a) planar singlet; (b) perpendicular singlet; (c) planar triplet; (d) perpendicular triplet.

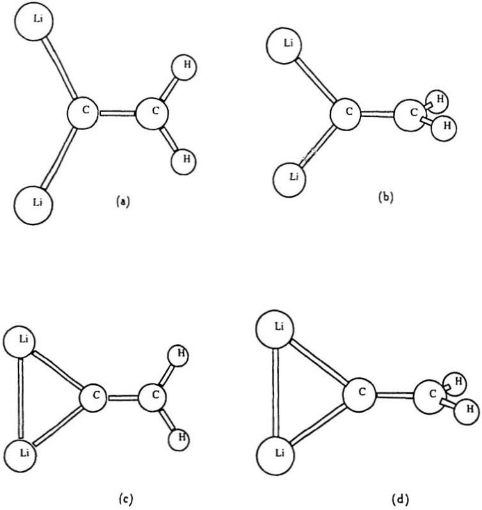

Li
Li
C
C
H
H
(a)
Li
Li
C
C
H
H
(b)
Li
Li
C
C
H
H
(c)
Li
Li
C
C
H
H
(d)

Table 29: Calculated bond orders, Mulliken charges and dipole moments.

$H_2X{=}YLi_2$	State	X-Y	Li-Li	X	Y	μ(Debye)
CH_2CLi_2	S(Planar)	1.668	0.076	-0.199	-0.306	5.004
	S(Perpendicular)	1.622	0.107	-0.216	-0.236	4.515
	T(Planar)	1.802	0.540	-0.288	-0.270	-0.970
	T(Perpendicular)	1.733	0.530	-0.269	-0.260	-1.494
SiH_2CLi_2	S(Planar)	1.620	0.071	0.385	-0.524	6.065
	S(Perpendicular)	1.643	0.143	0.401	-0.480	5.628
	T(Planar)	1.729	0.541	0.378	-0.494	-0.979
	T(Perpendicular)	1.718	0.531	0.420	-0.502	-1.667
CH_2SiLi_2	S(Planar)	1.525	0.103	-0.428	0.076	4.149
	S(Perpendicular)	1.435	0.288	-0.404	0.011	6.053
	T(Planar)	0.812	0.104	-0.450	0.130	4.583
	T(Perpendicular)	0.891	0.104	-0.454	0.071	3.760

S: for singlet; T: for triplet. The negative dipole moment implies $H_2X^+ = {}^-CLi_2$ polarity.

5.5 S_N2 Reaction: Transition State Structures, Frequencies and Kinetic Isotope Effects

Bimolecular nucleophilic substitution reactions (S_N2) at carbon

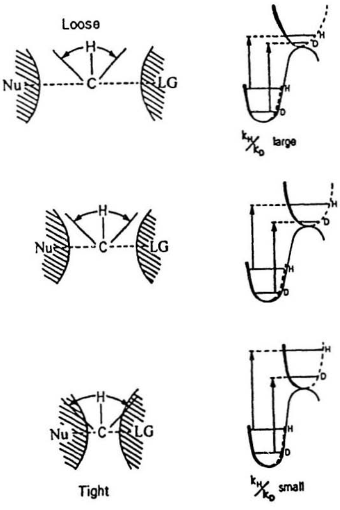
$$Y^- + CH_3X \rightarrow YCH_3 + X^-$$

have been extensively studied. Early experimental and theoretical studies have established an energy profile consisting of a double-well potential energy surface [94]. Since it is difficult to obtain a detailed energy profile from experiment, ab initio calculations afford an excellent source of data. There have been a number of ab initio calculations on gas-phase S_N2 reactions to study energy profiles [95], transition state structures [96], isotope effects [97] and solvation effects [98]. The effect of electron correlation on S_N2 reactions has also been studied by many researchers using different post Hartree-Fock levels, such as CI, MP2, MCSCF, MRCI, MBPT, and DFT methods, with extensive basis sets [99].

Secondary α-deuterium kinetic isotope effects (KIEs) have been widely used to investigate transition state structures and substituent effects [100] in the study of S_N2 reaction mechanisms. For a particular substrate, a tighter transition state, with short nucleophile-α-carbon and/or α-carbon-leaving group bonds, will have high-energy C_α-H(D) out-of-plane bending vibrations and the secondary α-deuterium KIE is predicted to be small. If the S_N2 transition state is looser, with longer nucleophile-α-carbon and/or α-carbon-leaving-group bonds, the C_α-H(D) out-of-plane bending vibrations will be lower in energy and a larger secondary α-deuterium KIE is predicted, Figure 17. Although it has been accepted that secondary α-deuterium KIEs are determined primarily by out-of-plane bending vibrations of the C_α-H(D) bonds in the reactant and transition state, recent theoretical calculations by Truhlar et al. [98], and by Wolfe et al. [97] have suggested that the inverse KIEs in S_N2 reactions are primarily the result of changes in the C_α-H(D) stretching vibrations rather than the C_α-H(D) out-of-plane bending vibrations. This suggests a smaller KIE is associated with a looser, rather than a tighter, S_N2 transition state.

Our HF/6-31+G* study [97(c)], for S_N2 reactions between different nucleophiles and methyl and ethyl chloride and fluoride, demonstrates that the magnitude of the secondary α-deuterium KIE is determined by an inverse stretching vibration contribution and a normal bending vibration contribution to the isotope effect. The out-of-plane bending vibration model that has been used to interpret secondary α-deuterium KIEs is therefore basically correct, *i.e.*, changes in the bending vibrations from the reactants to transition state determine the magnitude of the KIE for a particular substrate.

Figure 17: The relationship between the looseness of the S_N2 transition state and the magnitude of the secondary α-deuterium KIE as determined by the C_α-H(D) out-of-plane bending vibrations.


Loose
H
Nu
C
LG
H
D
H
D
kH/kD large
H
Nu
C
LG
H
D
H
D
H
Nu
C
LG
Tight
H
D
H
D
kH/kD small

Previous studies [95,96,99] have shown the effect of electron correlation on the C-Y and C-X bond distances in the transition state structures. Since GVB properly describes the bond dissociation process in general, it is expected that GVB would also properly describe the energy profile and the transition state structure of an S_N2 reaction, where a C-X bond is breaking and a C-Y bond is forming. On the other hand, early studies [19(b), 65, 102] demonstrated that GVB improves the SCF force constants and frequencies, which are important for calculating accurate secondary α-deuterium KIEs.

In this work, S_N2 reactions with the nucleophiles (Y)=F, Cl, OH, NH_2, SH, and the leaving group (X)=Cl, and with Y=Cl and X=F, OH, NH_2, have been studied using GVB/6-31++G** calculations. The GVB-PP wavefunction includes all bond pairs. The reactants and transition states were optimized using the OC [26] and VA [27] optimization methods (*c.f.* Section 2.1.5), respectively. All transition states were verified to ensure that they have only one imaginary frequency. Sims's BEBOVIB-IV program [103] modified to accept ab initio geometries and force constants [104], was used to carry out a vibrational analysis and to calculate secondary α-deuterium KIEs at 298 K. The corresponding HF/6-31+G* results from our previous studies, and other available results on HF, MP2 and DFT, are also given for comparison.

5.5.1 Geometries of Transition States

The optimized geometries are summarized in Table 30 for reactants and products, and in Table 31 for transition states.

For reactants and products, our HF/6-31+G* geometries are very close to the previous HF/6-31++G** results of Boyd et al. [96(a)]. For instance, the C-X bond lengths are 1.372, 1.786, 1.401, and 1.454 Å, compared to 1.371, 1.786, 1.401, and 1.453 Å for X=F, Cl, OH, and NH_2, respectively. The GVB geometries are close to the MP2 [96(b)] and DFT [99] results, with longer (by 0.02 Å) C-H bonds and C-X bonds when compared to HF.

For transition states, our HF geometries are almost the same as the Boyd's HF results for X=Cl, F. For X=OH, both C-X and C-Cl bond distances are slightly different due to the different basis sets on the OH hydrogen. Both HF and GVB results show that the C-H bonds are shorter in transition states than in the reactants by ~0.02 Å.

Our GVB results show that the effect of electron correlation on the geometries of S_N2 transition states is mainly in the bond distances. For instance, the GVB C-H bonds are longer by 0.02 Å than HF C-H bonds. The GVB C-X and C-Y bonds are either longer or shorter, with the total distance $[R_{TS} = (C-Y)_{TS} + (C-X)_{TS}]$ larger than HF R_{TS}. The GVB transition state structures are therefore looser than those obtained from HF and MP2.

Table 30: Optimized geometries for CH_3X.

X		F	Cl	OH	NH_2
r_{C-X}					
	HF	1.371	1.786	1.401	1.453
	HF(this study)	1.372	1.786	1.401	1.454
	GVB(this study)	1.398	1.814	1.431	1.481
	DFT	1.400	1.778	-	-
	MP2	1.406	1.779	1.429	1.464
r_{C-H_1}					
	HF	1.081	1.078	1.081	1.090
	HF(this study)	1.081	1.078	1.080	1.090
	GVB(this study)	1.100	1.097	1.100	1.108
	DFT	1.099	1.095	-	-
	MP2	1.090	1.089	1.090	1.099
r_{C-H_2}					
	HF	-	-	1.087	1.084
	HF(this study)	-	-	1.087	1.084
	GVB(this study)	-	-	1.105	1.103
	MP2	-	-	1.096	1.092
$\angle H_1CX$					
	HF	108.5	108.4	107.1	114.5
	HF(this study)	108.5	108.4	106.9	114.5
	GVB(this study)	108.5	108.4	106.7	114.4
	DFT	108.7	108.2	-	-
	MP2	108.1	108.9	106.1	115.0
$\angle H_2CX$					
	HF	-	-	111.7	109.2
	HF(this study)	-	-	111.7	109.2
	GVB(this study)	-	-	111.5	108.9
	MP2	-	-	111.7	108.7

HF results from Ref. [96(a)], MP2 results from Ref. [96(b)], and DFT from Ref. [99]. Bond lengths are in Å, and bond angles in degrees.

Table 31: S_N2 transition state geometries.

Y		r_{C-Y}	r_{C-X}	r_{C-H_1}	r_{C-H_2}	∠ H_1CY	∠ H_2CY
Cl							
	HF	2.393	2.394	1.062	-	90.0	-
	HF(this study)	2.394	2.394	1.061	-	90.0	-
	GVB(this study)	2.405	2.405	1.079	-	90.0	-
	DFT	2.342	2.342	1.080	-	90.0	-
	MP2	2.316	2.316	1.072	-	90.0	-
F							
	HF	2.126	2.133	1.062	-	82.7	-
	HF(this study)	2.129	2.132	1.062	-	82.6	-
	GVB(this study)	2.036	2.216	1.079	-	85.2	-
	MP2	2.013	2.142	1.073	-	84.4	-
HO							
	HF	2.267	2.102	1.063	1.062	76.7	83.8
	HF(this study)	2.282	2.093	1.064	1.062	79.7	81.8
	GVB(this study)	2.326	2.075	1.082	1.082	74.5	83.2
NH_2							
	HF(this study)	2.388	2.072	1.062	1.063	84.9	79.1
	GVB(this study)	2.343	2.163	1.080	1.080	87.3	80.6
HS							
	HF(this study)	2.603	2.281	1.063	1.062	85.7	87.3
	GVB(this study)	2.658	2.254	1.081	1.080	84.6	86.3

see notes for Table 30.

5.5.2 Frequencies

It is well known that HF/6-31G* and MP2/6-31G* overestimate the experimental frequencies by ~10% and ~5% [5] respectively, mainly due to the unharmonicity. Previous calculations have shown that GVB stretching force constants are smaller than HF values [65,102] and close to the values obtained at MP4 [19(b)]. In this study, the GVB force constants were calculated numerically. Our results show that GVB stretching frequencies are very close to the experimental values, and that the bending frequencies are overestimated by ~5%. The GVB bending frequencies are therefore scaled[18] by a factor of 0.95.

The calculated GVB vibrational frequencies for the reactants and products are listed in Table 32, along with the corresponding HF and experimental values [5]. The HF and GVB frequencies for the transition states are also given in Tables 33 and 34, respectively.

GVB frequencies for reactants and products are very close to experimental and scaled HF frequencies. In CH_3Cl, for example, the GVB C_α-H stretching frequencies are ± 3 cm^{-1} and C_α-H bending frequencies are $\pm$ 10 cm^{-1} in deviation from the experimental data [5]. For the transition state, the GVB imaginary frequencies are larger, C_α-H stretching frequencies are smaller, and some of the bending frequencies are slightly smaller, compared to HF.

On the other hand, the C_α-H stretching is much stronger, but the bending is

[18]GVB frequencies have not been studied systematically. Only a few results are reported by Pulay [19(b)], in which the different scaling factors are used for each modes. In general, bending frequencies are scaled by a factor of 0.95 in our full valence GVB/6-31G** calculated frequencies.

much weaker, as compared to those in the reactants and products. From reactant to transition state, the C_α-H stretching frequencies increase by $\sim$ 200 cm^{-1}, while the bending frequencies decrease by $\sim$ 100 cm^{-1}.

5.5.3 Secondary α-Deuterium k_H/k_D

The total KIE can be expressed (see Appendix B) as the product of translational, rotational, vibrational, and tunneling terms,

$$k_H/k_D = (k_H/k_D)_{trans}(k_H/k_D)_{rot}(k_H/k_D)_{vib}(k_H/k_D)_{tunneling} \tag{126}$$

where the tunneling term is determined using the Wigner correction [103]. The vibrational contribution to the KIE can be further factored,

$$(k_H/k_D)_{vib} = (k_H/k_D)_{stretching}(k_H/k_D)_{bending}(k_H/k_D)_{other} \tag{127}$$

where the $(k_H/k_D)_{stretching}$ term contains the contribution from the C_α-H(D) stretching vibrations to the KIE and the $(k_H/k_D)_{bending}$ term contains the contribution from the C_α-H(D) bending vibrations. Three stretching vibrations and five bending vibrations are used to calculate the $(k_H/k_D)_{stretching}$ and $(k_H/k_D)_{bending}$ terms, respectively. The contribution from other vibrations is included in $(k_H/k_D)_{other}$ term.

Table 32: The calculated frequencies (cm^{-1}) for reactants and products.

Molecules	Description of Mode	HF/6-31+G*	GVB/6-31++G**	Expt.
CH_3Cl	CH_3 d-stretch (e)	3032	3037	3039
	CH_3 s-stretch (a_1)	2940	2955	2937
	CH_3 d-deform (e)	1462	1462	1452
	CH_3 s-deform (a_1)	1382	1386	1355
	CH_3 rock (e)	1023	1018	1017
	CCl stretch (a_1)	700	700	732
CH_3F	CH_3 d-stretch (e)	2999	3004	3006
	CH_3 s-stretch (a_1)	2920	2934	2930
	CH_3 d-deform (e)	1479	1474	1467
	CH_3 s-deform (a_1)	1467	1463	1464
	CH_3 rock (e)	1174	1163	1182
	CF stretch (a_1)	1037	1027	1049
CH_3OH	OH stretch (a')	3703	3709	3681
	CH_3 d-stretch (a')	2981	2992	3000
	CH_3 d-stretch (a'')	2915	2929	2960
	CH_3 s-stretch (a')	2873	2893	2844
	CH_3 d-deform (a')	1490	1487	1477
	CH_3 d-deform (a'')	1482	1479	1477
	CH_3 s-deform (a')	1466	1464	1455
	OH (a')	1337	1358	1345
	CH_3 rock (a'')	1157	1150	1165
	CH_3 rock (a')	1060	1046	1060
	CO stretch (a')	1038	1027	1033
	torsion (a'')	300	347	295

Table 32. (continued).

Molecules	Description of Mode	HF/6-31+G*	GVB/6-31++G**	Expt.
CH_3NH_2	NH_2 a-stretch (a'')	3441	3425	3427
	NH_2 s-stretch (a')	3365	3354	3361
	CH_3 d-stretch (a'')	2953	2967	2985
	CH_3 d-stretch (a')	2920	2936	2961
	CH_3 s-stretch (a')	2847	2871	2820
	NH_2 scis. (a')	1660	1662	1623
	CH_3 d-deform (a'')	1494	1490	1485
	CH_3 d-deform (a')	1479	1479	1473
	CH_3 s-deform (a')	1442	1443	1430
	NH_2 twist (a'')	1325	1336	1419
	CH_3 rock (a'')	1153	1167	1195
	CN stretch (a')	1031	1015	1044
	CH_3 rock (a')	945	951	1130
	NH_2 wag (a')	837	903	780
	torsion (a'')	295	315	268

The experimental data from Ref. [5]. HF frequencies are scaled by 0.90, and our GVB bending frequencies are scaled by 0.95.

Table 33: HF/6-31+G* frequencies (cm^{-1}) for S_N2 transition states.

Y	Cl	F	HO	NH2	HS
imaginary	-384	-395	-394	-432	-420
H-C bending	901	955	944	966	905
	901	955	956	973	925
	1116	1174	1165	1151	1089
	1394	1396	1398	1391	1391
	1394	1396	1398	1396	1392
H-C stretching	3079	3073	3069	3068	3074
	3271	3249	3234	3235	3258
	3271	3249	3245	3242	3267
H-X stretching	-	-	3617	3192	2539
	-	-	-	3264	-
other stretching	179	208	175	162	173
	179	208	191	187	181
torsion	186	243	221	248	70
	-	-	243	309	157
	-	-	386	372	385

see notes for Table 32.

Table 34: GVB/6-31++G** frequencies (cm^{-1}) for S_N2 transition states.

Y	Cl	F	HO	NH2	HS
imaginary	-460	-469	-401	-475	-441
H-C bending	908	961	963	961	917
	938	976	993	989	947
	1082	1133	1167	1132	1083
	1388	1385	1390	1376	1382
	1389	1389	1395	1387	1392
H-C stretching	3080	3067	3055	3064	3070
	3255	3237	3214	3232	3239
	3261	3239	3217	3238	3245
H-X stretching	-	-	3541	3197	2482
	-	-	-	3122	-
other stretching	185	234	172	181	172
	188	238	193	197	178
torsion	192	257	223	221	134
	-	-	251	251	186
	-	-	408	439	376

see notes for Table 32.

Table 35 summarizes the α-deuterium secondary k_H/k_D for the S_N2 reactions considered. The GVB KIEs are slightly larger than our previously reported HF values [97(c)]. Where they can be compared, the KIEs in this study are in excellent agreement with those reported by Wolfe et al. [97(a)]. For the methyl chloride identity reactions, for example, Wolfe et al. reported k_H/k_D of 0.94 (HF) and 0.96 (MP2), compared to 0.935 (HF) and 0.957 (GVB) obtained in this study.[19]

The relationship between the magnitude of the secondary α-deuterium KIE and the structure of the transition state is investigated in an effort to analyze the factors that determine the magnitude of these KIEs in S_N2 reactions [97(c)]. This study will examine whether these KIEs are primarily determined by changes in the stretching vibrations as proposed by Wolfe et al. [97(a-b)] or whether they are the result of changes in the out-of-plane bending vibrations as had been previously believed [100]. In particular, the objective is to further investigate the electron correlation effect at the post-HF (GVB) level. In terms of the results in Table 35, all stretching contributions to the KIE are large and inverse at both the HF and GVB levels, and the bending contributions to the KIE are close to or greater than unity.

[19]It should be pointed out that in Wolfe et al. recent paper [97(b)], k_H/k_D for the $Cl^- + CH_3F$ system is incorrectly reported to be 1.30. The correct value of k_H/k_D is 1.002 (HF) and 1.031 (GVB) as given in Table 35.

Table 35: Calculated k_H/k_D for S_N2 reactions at 298 K.

X	Y		$\frac{k_H}{k_D}$	$(\frac{k_H}{k_D})_s$	$(\frac{k_H}{k_D})_b$	$(\frac{k_H}{k_D})_{vib}$	L_{TS}	R_{TS}
F	Cl	HF[a]	0.882	0.705	1.098	0.711	74.5	4.261
		GVB[a]	0.921	0.733	1.091	0.743	-	4.252
		HF	0.88	0.73	1.20			
		MP2	0.91					
OH	Cl	HF[a]	0.907	0.714	1.128	0.704	79.7	4.371
		GVB[a]	0.926	0.757	1.084	0.718	-	4.401
		HF	0.92					
NH_2	Cl	HF[a]	0.897	0.716	1.112	0.665	80.2	4.460
		GVB[a]	0.927	0.757	1.114	0.685	-	4.506
Cl	Cl	HF[a]	0.935	0.680	1.178	0.739	68.0	4.788
		GVB[a]	0.957	0.704	1.157	0.756	-	4.810
		HF	0.94	0.71	1.33			
		MP2	0.96	0.75	1.29			
		MP2	0.97					
SH	Cl	HF[a]	0.953	0.690	1.197	0.638	70.9	4.884
		GVB[a]	0.978	0.722	1.167	0.706	-	4.912
Cl	F	HF[a]	1.002	0.665	1.157	0.762	74.5	4.261
		GVB[a]	1.031	0.692	1.325	0.784	-	4.252
		HF	1.00	0.73	1.37			
Cl	OH	HF[a]	0.980	0.612	1.274	0.753	79.7	4.371
		GVB[a]	0.960	0.655	1.322	0.738	-	4.401
		MP2	0.960					
Cl	NH_2	HF[a]	0.840	0.594	1.048	0.651	80.2	4.460
		GVB[a]	0.879	0.686	1.273	0.681	-	4.506

[a] This study, see also [97(c)]. Other results are from Wolfe et al. [97(a-b)].

The relationship between the total KIE and $(k_H/k_D)_{stretching}$ is illustrated in Figure 18. The poor correlation (r=-0.748 at HF and -0.694 at GVB) indicates that the change in the total KIE is not determined by the change in $(k_H/k_D)_{stretching}$. A plot of the total KIE versus $(k_H/k_D)_{bending}$ as illustrated in Figure 19, shows an excellent correlation (r=0.995 at HF and 0.943 at GVB). The trends in KIEs are, therefore, in agreement with the commonly accepted view that the KIE is determined by changes in the bending vibrations and not by changes in the stretching vibrations as suggested by Wolfe et al. [97(a)]. If changes in the bending vibrations are responsible for the trends in the KIE, a S_N2 reaction with a looser transition state would be expected to have a larger KIE. For example, the $(k_H/k_D)_{bending}$ contribution is larger for Y=Cl (1.157 at GVB) than for Y=F (1.091 at GVB), suggesting that the transition state is looser and that there is less "steric hindrance" exerted on the out-of-plane bending vibration in the transition state. The larger bending contribution results in a larger KIE, 0.957 for Y=Cl versus 0.921 for Y=F, although the stretching contribution is smaller, 0.704 versus 0.733.

Figure 18: The total KIE versus the stretching vibration contribution to the KIE for $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$.
Circles are for HF results (r=-0.748), and triangles are for GVB results (r=-0.694).

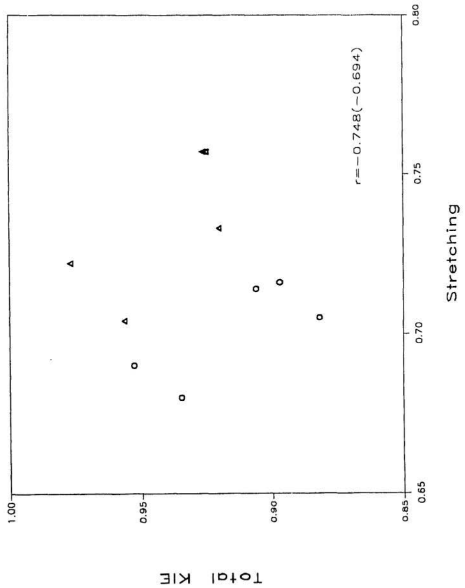
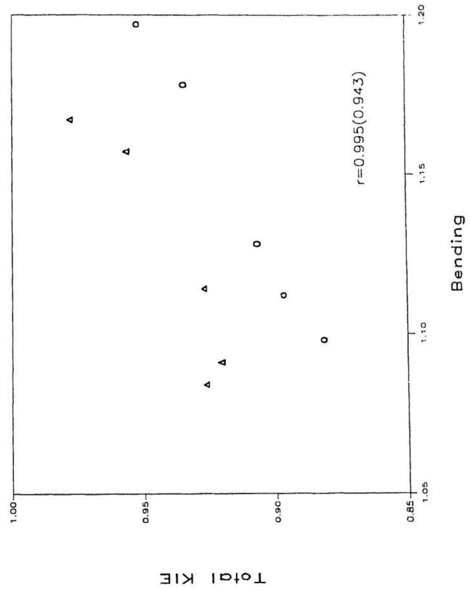

Total KIE
Stretching
r=-0.748(-0.694)
1.00
0.95
0.90
0.85
0.65
0.70
0.75
0.80

Figure 19: The total KIE versus the bending vibration contribution to the KIE for $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$.
Circles are for HF results (r=0.995), and triangles are for GVB results (r=0.943).



r=0.995(0.943)
1.05
1.10
1.15
1.20
Bending
1.00
0.95
0.90
0.85
Total KIE

The inverse relationship between the KIE and the looseness parameter [97(a)], L_{TS}, found by Wolfe et al.[20] suggested that a reaction with a looser transition state will have a smaller KIE. Our previous study [97(c)] showed that there is no correlation between the KIE and L_{TS} for identity reactions. The lack of a correlation between the KIE and L_{TS} is more clearly seen when the KIEs for the non-identity methyl fluoride and methyl chloride reactions are considered. We plotted Wolfe et al. [97(b)] $(k_H/k_D)_{total}$ versus L_{TS} for methyl fluoride and methyl chloride with different nucleophiles, Figures 20 and 21, respectively. The correlation coefficients for the total KIE versus L_{TS} are small (0.153 and -0.558 respectively). Our HF/6-31+G* results for methyl chloride also confirm that L_{TS} is not a good measure of the looseness of the transition state, Figure 22.

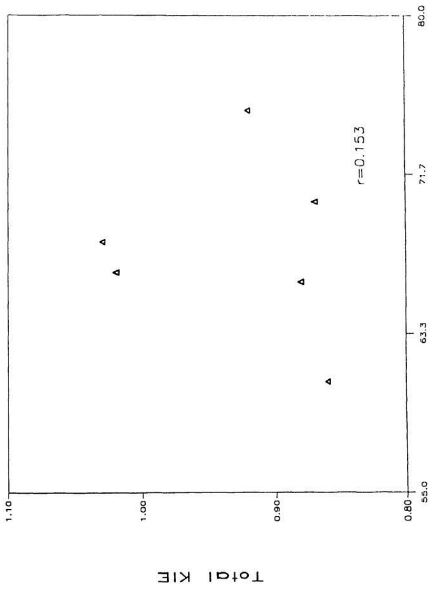
The magnitude of the KIE is, however, related to the "space" available for the out-of-plane bending vibrations in the transition state, since the trends in the KIE are the result of changes in the bending vibrations. Such "space" can be represented by R_{TS}, *i.e.*, the distance between the nucleophile and the leaving group in the S_N2 transition state. The relationship between the total KIE and R_{TS} which is illustrated in Figure 23 for the reaction $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$, shows a very good correlation (r=0.965 at HF and 0.957 at GVB). The correlation coefficient for the $(k_H/k_D)_{bending}$ versus R_{TS} in Figure 24, is 0.968 at HF and 0.966 at GVB. Thus, as R_{TS} increases (the transition state becomes looser), both $(k_H/K_D)_{bending}$ and the total KIE increase as

[20]Wolfe et. al. defined, $L_{TS} = (\%C - Y + \%C - X)$, where the percent increase in the C-X bond length in going from the ion-complex to the transition state, %C-X, can be expressed as $\%C - X = \frac{(C-X)_{TS} - (C-X)_{complex}}{(C-X)_{complex}} \times 100\%$.

the out-of-plane bending vibration model predicts. The plot of Wolfe et al. total KIE [97(b)] versus their R_{TS} [96(a)], Figure 25, gives a correlation coefficient of 0.951, for methyl fluoride, as compared to 0.153 for correlation between KIE and L_{TS} in Figure 20. It is clearly shown that KIE is well correlated to R_{TS}, and not correlated to L_{TS}.

Overall both HF and GVB results of KIE support that $(k_H/k_D)_{bending}$ and the trend in the total KIE is determined by the looseness of the S_N2 transition state, R_{TS}, for S_N2 reactions of a substrate with a particular leaving group. Although the stretching vibration plays a significant role in determining the magnitude of the KIE, $(k_H/k_D)_{bending}$ determines the trends in these KIEs.

Figure 20: The total KIE versus the looseness parameter L_{TS} for $Y^- + CH_3F \rightarrow YCH_3 + F^-$.
The data in this plot are from Ref. [97(b)].



Total KIE
L
r=0.153
1.10
1.00
0.90
0.80
55.0
63.3
71.7
80.0

Figure 21: The total KIE versus the looseness parameter L_{TS} for $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$.
The data in this plot are from Ref. [97(b)].

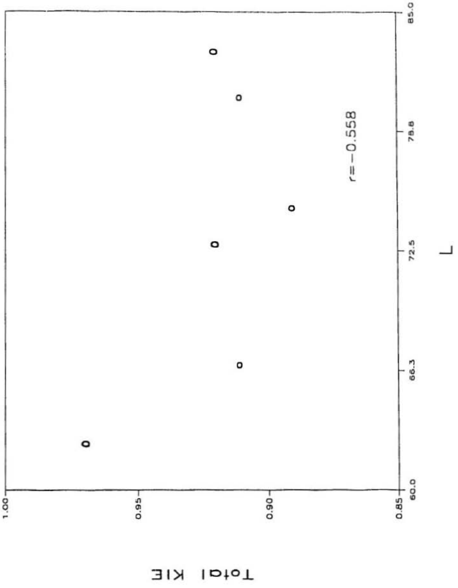

r=-0.558
1.00
0.95
0.90
0.85
60.0
66.3
72.5
78.8
85.0
L
Total KIE

Figure 22: The total KIE versus the looseness parameter L_{TS} for $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$.
The data in this plot are from our study, see Ref. [97(c)].

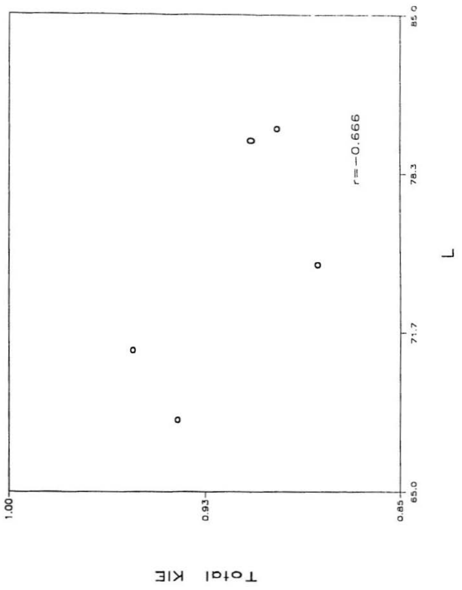
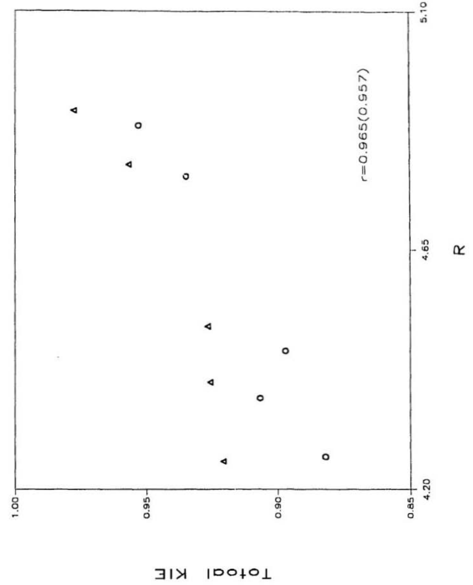

Total KIE
1.00
0.93
0.85
65.0
71.7
78.3
85.0
L
r=−0.666

Figure 23: The total KIE versus R_{TS} for $Y^- + CH_3Cl \rightarrow YCH_3 + Cl^-$.
Circles are for HF results (r=0.965), and triangles are for GVB results (r=0.957).



Total KIE
1.00
0.95
0.90
0.85
4.20
4.65
5.10
R
r=0.965(0.957)

Figure 24: The bending vibration contribution to the KIE versus R_{TS} (in Å) for Y^- + $CH_3Cl \rightarrow YCH_3 + Cl^-$.
Circles are for HF results (r=0.968), and triangles are for GVB results (r=0.966).

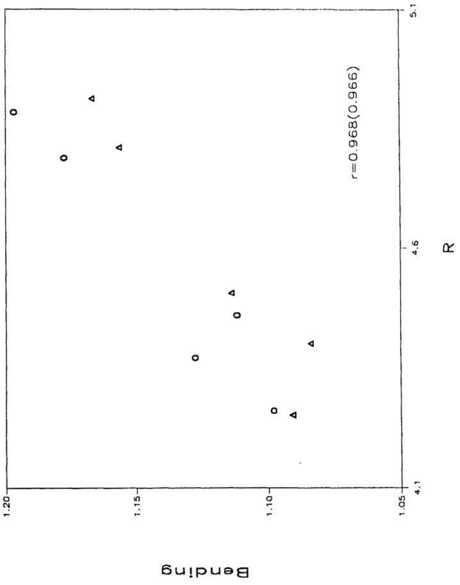
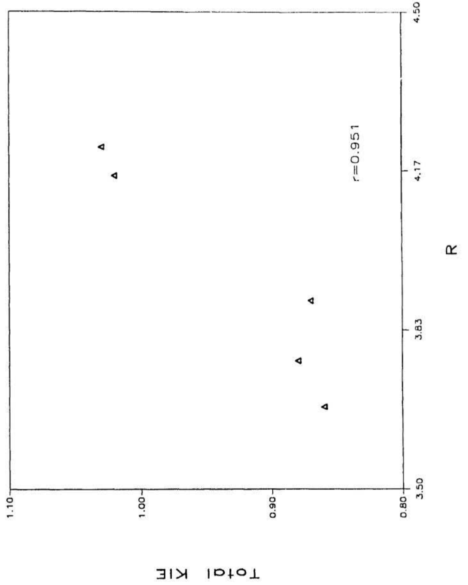

Bending
1.20
1.15
1.10
1.05
4.1
4.6
5.1
R
r=0.968(0.966)

Figure 25: The total KIE versus R_{TS} (in Å) for $Y^- + CH_3F \rightarrow YCH_3 + F^-$. The data in this plot are from Boyd et al. [97(b)], [96(a)] with the correction of (Cl, F) KIE to 1.03.


Total KIE
1.10
1.00
0.90
0.80
R
3.50
3.83
4.17
4.50
r=0.951

5.6 Protonation of Imines: Geometries, Force Constants, Frequencies and C=N Rotational Barrier

The imine (C=N) linkage has attracted much interest because of its occurrence in the visual pigment rhodopsin and bacteriorhodopsin (bR) [106]. Imines play an important role in biological catalysis since the C=N linkages are versatile in their physical and chemical properties. The C=N bond, for example, is fairly labile and can be hydrolyzed and re-formed readily.

Protonation of the C=N nitrogen of an imine containing chromophore generally leads to a marked red shift in the absorption spectrum. This reaction is important in controlling the optical properties of retina in the visual pigment rhodopsin. Many spectroscopic studies and theoretical calculations have been carried out in an effort to understand the absorption spectrum of rhodopsin [107].

On the other hand, the imine protonation reaction, or reaction with Lewis acids in general, results in an increase in the C=N stretching frequency (blue shift) [108]. Changes in the C=N stretching frequency in deuterated solutions have been regarded as the most reliable measure of whether the imine is protonated [109]. For instance, rhodopsin has a C=N stretching frequency ($\nu_{C=N(H)}$) of 1660 cm^{-1} and an unusually large deuterium shift of $\sim$ 25 cm^{-1} [110]. The large $\nu_{C=N(H)}$ deuterium shift suggests that the imine nitrogen in rhodopsin interacts strongly with a hydrogen bonded group. The C=N stretching frequency is used as a measure of π-electron distribution in the imine region of the molecule, since this is the most reactive part of the chromophore and chemical modification of the protein and chromophore makes it possible to study the C=N bond in considerable detail. Therefore, an accurate understanding of the factors that determine the C=N stretching frequency is clearly needed.

Methylimine (CH_2=NH), the simplest imine, and its protonated derivative, provide models which could be used to study the electronic charge distribution in the C=N bonding region where the nitrogen lone pair is bonded to a proton. Ab initio calculations of the force field for methylimine have been carried out. Eades et al. [111] reported that the C=N bond length increases by 0.019 Å upon protonation, and they also obtained the vibrational frequencies for unprotonated and protonated methylimine. In contrast to Eades et al. HF calculations, Harrison's GVB calculations [102] showed a slight decrease in the C=N bond length, and an increase of 0.51 mdyn/Å in the C=N force constant. Instead of calculating the $\nu_{C=N}$ in methylimine directly, Harrison et al. examined the frequency changes for CH_3HC=NCH_3 using an empirical force field which is constructed based on their GVB methylimine geometry and force constants [108]. The study showed that an increase in the C=N stretching force constant (by 0.51 mdyn/Å) for the protonated species translates into an increase in the C=N stretching frequency of $\sim 30\ cm^{-1}$.

However, in the protonation of a polyimine, the π bond order decreases resulting in a corresponding decrease in the C=N force constant. Honig et al. [112,113], using an empirical force field, found that coupling of the C=N stretching with the C=N-H bending makes an important contribution to the C=N stretching and pushes it to the higher frequencies that are observed experimentally. A normal coordinate analysis of the model aromatic imine trans-N-benzylideneaniline (tBA) using AM1 method, obtained a decrease in the C=N stretching force constant but an increase in the C=N stretching frequency by 45 cm^{-1} [117]. Protonation causes many changes in conformation, which is responsible for the changes in the vibrational spectrum. Recently, based on HF and MP2 calculations for the force field of N-ethylidene methylamine ($CH_3HC=NCH_3$) and its protonated species, Tsuda et al. [118] stated that the C=N stretching force constants cannot be evaluated properly without taking into account the effect of electron correlation.

In order to obtain a clearer description of the electronic environment changes that occur in the protonation of imines, GVB calculations of the force field for imines have been carried out. In this study, accurate force fields for the unprotonated and protonated $CH_2=NH$ and trans-$CH_3HC=NCH_3$ were calculated using GVB(6/12)/6-31G*, and frequencies were then calculated to investigate whether the C=N-H bending motion makes an important contribution to the C=N stretching frequency shift. The change of C=N π bond energy upon protonation was also studied by calculating the barrier to rotation around C=N double bond. The equilibrium geometries for all trans polyimines, taken as retinal analogs, were obtained at the GVB level to study the conjugation effect on the C=N bond length, force constants, rotation barrier and charge distribution.

In this study, GVB(6/12) calculations were carried out. The GVB(6/12) wavefunction includes the C-N σ and π bonds, two C-H bonds, one N-H bond, adding one nitrogen lone pair for the unprotonated and one N-H bond for the protonated methylimine $H_2C{=}NH$ respectively, *i.e.*,

$$\Psi^{GVB}_{(unprotonated)} = \mathcal{A}[\Psi_{core}\psi_{C-N(\sigma)}\psi_{C-N(\pi)}\psi_{C-H1}\psi_{C-H2}\psi_{N-H}\psi_{N-lp}] \tag{128}$$

$$\Psi^{GVB}_{(protonated)} = \mathcal{A}[\Psi_{core}\psi_{C-N(\sigma)}\psi_{C-N(\pi)}\psi_{C-H1}\psi_{C-H2}\psi_{N-H1}\psi_{N-H2}] \tag{129}$$

Similar GVB(6/12) wavefunctions were constructed for trans-$CH_3HC{=}NCH_3$ and polyimines. The equilibrium structures were optimized using the OC method [26]. The transition state structures were optimized using the VA method [27] and were characterized to have only one imaginary frequency. The GVB force constants and frequencies were calculated as described in Section 5.5 (see Appendix B).

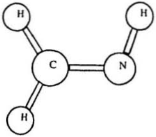
5.6.1 Geometries

The structures of the unprotonated and protonated $H_2C{=}NH$ are shown in Figure 26. The GVB(6/12) geometries at 6-31G*, 6-31G** and 6-311G** basis sets for the neutral and protonated methylimine are listed in Tables 36 and 37, along with the twisted structures at 6-31G* basis set. For methylimine, the GVB geometries at the three different basis sets are very close to each other, and there is good agreement between the GVB and the experimental geometrical parameters, as discussed in Section 5.1. Our GVB(6/12)/6-31G* geometries are much improved over the previous ab initio calculations. For instance, the GVB/6-31G* C=N bond length (1.282 Å) is closer to the experimental value (1.273 Å) [5], compared to the previous results, *e.g.*, 1.262 Å (HF) [111] and 1.289 Å (GVB) [102]. When protonated, however, HF results indicate that the $C{=}NH_2^+$ bond becomes longer by 0.02 Å, while previous GVB results [102] indicated that it shortened by 0.008 Å. Our GVB/6-31G* $C{=}NH_2^+$ bond length (1.281 Å) is only shorter by 0.001 Å than the C=N bond length. The H-N-C bond angle changes from 110.5^0 (sp^3) to 121.9^0 (sp^2) upon protonation.

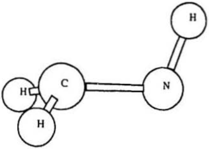
The GVB(6/12)/6-31G* optimized geometries for unprotonated and protonated trans-$CH_3CH{=}NCH_3$ are listed in Table 38. The structure of protonated trans-$CH_3CH{=}NCH_3$ is shown in Figure 27. With methyl substituents on carbon and nitrogen, the C=N bond is a little shorter than that in methylimine, but slightly longer upon protonation.

In general, an all trans polyimine can be represented by the formula for $H_2C{=}(CH{-}CH{=})_nX$ ($X{=}NH$, NH_2^+), where n is the number of conjugated π bonds. The structure for the n=6 species (polyimine **6**) is illustrated in Figure 27, and it was taken as a retinal analog in Poirier et al. calculations [114]. Table 39 lists some geometrical parameters for unprotonated and protonated polyimine **0**, **1**, **2**, **3** optimized at GVB(6/12)/6-31G* level. The C=N bond distance increases from n=0 to n=1, especially in the protonated imines. The changes in the C=N bond length upon protonation are -0.001, 0.011, 0.019, 0.022 and 0.022 Å, for n=0, 1, 2, 3, and 6. Overall, for retinal analogs, the GVB C=N bond distance is longer by ~0.02 Å upon protonation (see Table 40), which is in agreement with previous HF results [114].

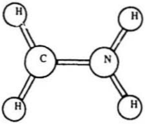
Figure 26: The structure of $H_2C{=}NH$ and $H_2C{=}NH_2^+$.
(a) planar $H_2C{=}NH$; (b) twisted $H_2C{=}NH$; (c) planar $H_2C{=}NH_2^+$; (d) twisted $H_2C{=}NH_2^+$.



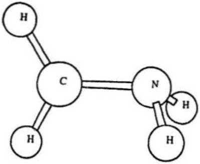
(a)



(b)



(c)



(d)

Table 36: GVB(6/12)/6-31G* geometries for $H_2C=NH$ and $H_2C=NH_2^+$.

Parameters	planar		twisted	
	C=N	$C=N^+$	C-N	$C-N^+$
C=N	1.282	1.281	1.446	1.459
C-H1	1.099	1.091	1.092	1.088
C-H2	1.095	1.091	1.092	1.087
N-H3	1.026	1.021	1.029	1.026
N-H4	-	1.021	-	1.026
H1-C-N	124.4	119.5	115.3	114.7
H2-C-N	118.9	119.5	115.3	114.7
H3-N-C	110.5	121.9	108.1	121.1
H4-N-C	-	121.9	-	122.0
H3-N-C-H1				153.2
H1-C-N-H3			109.0	103.6
H4-N-C-H1				180.0
H2-N-C-H3			-109.0	

Bond lengths are in Å, and bond angles in degrees.

Table 37: GVB(6/12) geometries for $H_2C{=}NH$ and $H_2C{=}NH_2^+$ at 6-31G** and 6-311G** basis sets.

Parameters	GVB(6/12)/6-31G**		GVB(6/12)/6-311G**	
	C=N	C=N$^+$	C=N	C=N$^+$
C=N	1.281	1.281	1.279	1.280
C-H1	1.100	1.092	1.101	1.093
C-H2	1.096	1.092	1.097	1.093
N-H3	1.025	1.020	1.025	1.021
N-H4	-	1.020	-	1.021
H1-C-N	124.3	119.6	124.2	119.6
H2-C-N	118.9	119.6	118.9	119.6
H3-N-C	110.4	121.7	110.3	121.7
H4-N-C	-	121.7	-	121.7

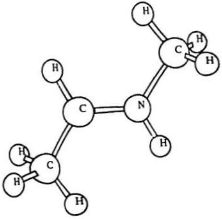
Bond lengths are in Å, and bond angles in degrees.

Table 38: GVB(6/12)/6-31G* geometries for unprotonated and protonated trans-$CH_3CH{=}NCH_3$.

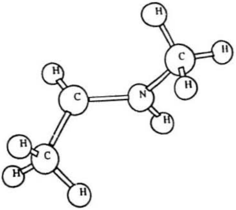
Parameters	planar		twisted	
	C=N	$C{=}NH^+$	C=N	$C{=}NH^+$
C15=N	1.278	1.285	1.446	1.461
C15-H	1.104	1.094	1.093	1.089
C15-C	1.525	1.515	1.523	1.521
N-C	1.477	1.499	1.470	1.481
N-H	-	1.021	-	1.025
H-C15-N	121.7	116.5	113.2	112.7
C-C15-N	121.5	123.2	116.2	117.5
C-N-C15	118.0	125.8	112.0	123.2
H-N-C15	-	117.9	-	118.6
C-N-C15-H	0.0	0.0		
C-C15-N-C	180.0	180.0	-106.0	155.9
H-N-C15-H	-	180.0		
H-C15-N-C	0.0	0.0	111.4	105.7

Bond lengths are in Å, and bond angles in degrees.

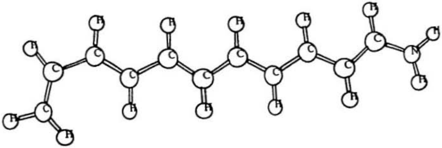
Figure 27: The structure of retinal analogs.
(a) trans-$CH_3HC{=}NH^+CH_3$; (b) twisted trans-$CH_3HC{=}NH^+CH_3$; (c) protonated polyimine **6**.



(a)



(b)



(c)

Table 39: GVB(6/12)/6-31G* geometrical parameters for all trans polyimines **n** [$H_2C{=}(CH\text{-}CH{=})_nX$ ($X{=}NH$, NH_2^+)].

n	C15=N	N-H1	N-H2	C15-C14	C14-C13	C13-C12	C12-C11	C11-C10	C10-C9
0	1.282	1.026							
	1.281	1.021	1.021						
1	1.284	1.027	-	1.493	1.320				
	1.295	1.018	1.017	1.455	1.331				
2	1.284	1.027	-	1.489	1.327	1.464	1.323		
	1.303	1.016	1.015	1.432	1.350	1.446	1.330		
3	1.285	1.027	-	1.488	1.328	1.459	1.330	1.463	1.324
	1.307	1.015	1.014	1.421	1.360	1.427	1.346	1.451	1.328

The bond lengths in Å. For each n, the first line is for X=NH, and the second line is for $X{=}NH_2^+$.

Table 40: GVB(6/12)/6-31G geometry for planar polyimine **6**.

	X=NH	$X=NH_2^+$		X=NH	$X=NH_2^+$
		Bond lengths (Å)			
C15-X	1.302	1.324	N-H	1.034	1.010
C14-C15	1.481	1.402	N-H		1.009
C13-C14	1.333	1.377	C15-H	1.093	1.092
C12-C13	1.454	1.405	C14-H	1.077	1.073
C11-C12	1.337	1.366	C13-H	1.077	1.078
C10-C11	1.452	1.421	C12-H	1.077	1.073
C9-C10	1.337	1.353	C11-H	1.077	1.078
C8-C9	1.453	1.437	C10-H	1.077	1.074
C7-C8	1.336	1.344	C9-H	1.077	1.078
C6-C7	1.468	1.462	C8-H	1.076	1.074
C5-C6	1.329	1.330	C7-H	1.077	1.077
			C6-H	1.076	1.074
			C5-H	1.073	1.072
			C5-H	1.074	1.073
		Bond angles (degrees)			
H-C15-X	116.4	115.4	C14-C15-X	126.7	125.3
H-C14-C15	116.9	119.3	C13-C14-C15	122.8	120.5
H-C13-C14	119.0	117.8	C12-C13-C14	124.6	125.4
H-C12-C13	116.7	118.6	C11-C12-C13	123.8	121.7
H-C11-C12	119.1	118.2	C10-C11-C12	124.4	125.2
H-C10-C11	116.6	118.1	C9-C10-C11	124.0	122.0
H-C9-C10	119.2	118.6	C8-C9-C10	124.3	124.9
H-C8-C9	116.2	117.1	C7-C8-C9	123.6	121.9
H-C7-C8	118.0	117.7	C6-C7-C8	127.1	127.5
H-C6-C7	114.6	115.1	C5-C6-C7	127.5	126.3
H-C5-C6	121.1	120.9			
H-N-C15	113.0	121.7			
H-N-C15		121.7			

5.6.2 C=N Stretching Force Constants and Frequencies

A normal coordinate analysis of unprotonated, protonated and deuterated imines is necessary to assign the C=N stretching mode and to understand the origin of the various frequency shifts in greater detail. Honig et al. [112(b)] carried out normal coordinate vibrational frequency calculations on empirical force fields, to study the factors that influence the C=N stretching frequencies in protonated imines, visual pigments, and bacteriorhodopsin (bR).

For methylimine, the C=N stretching force constant from GVB is much smaller than that from HF, as shown in Table 41. The GVB(6/12)/6-31G* calculated frequencies for methylimine are listed in Table 42, compared with the experimental, HF and MP2 results. The GVB frequencies are in good agreement with the experimental data [5]. For example, the GVB C=N stretching frequency is 1628 cm^{-1} which is very close to 1638 cm^{-1} (MP2) and 1640 cm^{-1} (experiment). While the HF result is 1719 cm^{-1}, which is too large due to the too short HF C=N bond distance.

For unprotonated and protonated methylimine, our GVB C=N stretching force constants are smaller than Harrison's [102]. The C=N stretching force constant increases upon protonation by 0.70 mdyn/Å which is slightly larger than the 0.51 mdyn/Å obtained by Harrison. For trans-$CH_3HC=NCH_3$, the C=N stretching force constant is 10.61 mdyn/Å, which is a little larger than that in $CH_2=NH$ due to a shorter bond distance. Tsuda et al. [118] stated that the MP2 C=N stretching force constants of the unprotonated and protonated $CH_3HC=NH^+CH_3$ are nearly equal. However, no details were reported. Our GVB C=N stretching force constant in trans-CH_3HC=NH^+CH_3 is smaller than that in CH_2=NH_2^+ due to a longer bond distance. The increase in the C=N stretching force constant upon protonation of trans-CH_3HC=NCH_3 is 0.24 mdyn/Å, which is smaller than the 0.51 mdyn/Å used by Harrison in their model calculations [108]. Our GVB results show a strong coupling between the C=N stretching and C=N-H bending modes in protonated H_2C=NH (0.33 mdyn/rad) and protonated trans-CH_3HC=NCH_3 (0.31 mdyn/rad). Based on the GVB force constants, the calculated $\nu_{C=N}$ is 1688 cm^{-1} in trans-CH_3HC=NCH_3, and $\nu_{C=N(H)}$ is 1746 cm^{-1} upon protonation. The C=N(H) stretching frequency deuterium isotope shift is 20 cm^{-1}.

As mentioned, since the C=N bond in protonated polyimines is longer by 0.02 Å than in unprotonated polyimines (Tables 39 and 40) [114], it is incorrect to use the methylimine force constants for other polyimines as Harrison did. It is proper to take trans-CH_3HC=NCH_3 as a model and to calculate the force constants directly at the ab initio level, to obtain C=N stretching frequency isotope shifts. In terms of the GVB calculated force constants and normal coordinate vibrational frequencies $\nu_{C=N}$ with the different isotope effects, the effect of protonation on the C=N stretching and the contribution of C=N-H(D) bending to the C=N stretching can be investigated.

In this work, trans-CH_3HC=NCH_3 was taken as a model to calculate the accurate force constants and frequencies at GVB level. The geometries are optimized with the constraint of the C=N bond distances being 1.284 Å and 1.303 Å respectively, which are the equilibrium C=N bond distances for polyimine **3** [$H_2C=(CH\text{-}CH=)_3X$ ($X=NH$, NH_2^+)]. The calculated force constants, C=N stretching frequencies and isotope shifts are given in Table 39. The C=N stretching force constant for protonated imine is 9.70 mdyn/Å, a decrease of 0.47 mdyn/Å from unprotonated imine, instead of the increase as suggested by Harrison. The GVB force constants from this study are very close to the empirical values used before [112,113], providing a theoretical basis for those empirical values.

Our GVB calculated $\nu_{C=N}$ is 1659 cm^{-1} in the unprotonated, 1681 cm^{-1} in the protonated imines. Our GVB calculated ^{15}N isotope shift is 18 cm^{-1}, as observed in experiment of 15 cm^{-1} [116]. The C=N stretching frequencies are close to the previous calculations [112(b)] of 1630 cm^{-1} in unprotonated imine and 1659 cm^{-1} in protonated imine, whereas experimental data [111] are 1632 and 1650-1662 cm^{-1} in retinal imine and retinal protonated imine respectively. It is surprising that in protonated imine, the C=N force constant is lower by 0.47 mdyn/Å, while the C=N stretching frequency is higher by 22 cm^{-1} than in unprotonated imine. This increase is obviously not contributed by C=N stretching force constant itself, rather it is contributed by the C=N-H bending force constant and the C=N/C=N-H coupling force constant (0.31 mdyn/rad). The coupling of C=N stretching with C=N-H bending vibrations makes an important contribution to the C=N stretching mode and pushes it to the higher frequencies that are observed experimentally, as suggested by previous studies [112,113]. Therefore, the C=N stretching frequency depends on both the C=N stretching and C=N-H bending force constants. Due to the contributions of these two modes, the C=N stretching frequency is particularly sensitive to the hydrogen bonding environment at nitrogen.

The GVB $\nu_{C=N(H)}$ deuterium isotope shift is 26 cm^{-1}, which is in good agreement with the experimental observation of $\sim$ 25 cm^{-1} [111, 115] and Honig et al. calculation of 28 cm^{-1} [112,113]. Since the C=N stretching frequency and its deuterium isotope shift are significantly dependent on the C=N-H bending and the C=N/C=N-H coupling force constants, the C=N(H) stretching frequency deuterium isotope shift can be used as a measure of whether the nitrogen in imines is protonated. It is generally proposed that a large $\nu_{C=N(H)}$ deuterium isotope shift, $\sim$ 25 cm^{-1} in rhodopsin and bacteriorhodopsin, is characteristic of protonation. The large deuterium shift can only be obtained when the proton is covalently bounded to the nitrogen resulting in a large C=N-H bending force constant which is responsible for shifting the C=N stretching frequency of protonated imines. If the proton were not tightly bounded to the nitrogen, the observed frequency shift would have been much lower.

On the other hand, counterion dependencies in the C=N spectral region of protonated imines have been observed using resonance Raman and Fourier transform infrared (FTIR) spectroscopy [119]. The counterion can affect the $C{=}NH^{+}$ environment by means of hydrogen bonding. Since bonding of the proton with the counterion changes the C=N-H bending force constant, perturbation by the hydrogen bonding of the counterion affects the C=N(H) frequency more than the C=N(D) frequency, whereas the non-hydrogen-bonded counter ion leads to near-equal shifts of both frequencies. The deuteration-induced frequency shift of the C=N vibration is, therefore, the most explicit indicator of the strength of hydrogen bonding interaction of the imine group with its surroundings. Recently, FTIR spectra [119] show that the C=N stretching vibrational frequency is sensitive to which halide (Cl^-, Br^-, I^-) is present as a counterion for the protonated imine. The results indicate that the halide ions form direct hydrogen-bonded salt bridges with the protonated imine. The $C{=}NH^+$ stretch frequency correlates with the strength of the hydrogen bond formed by halide. However, as Birge et al. pointed out [107], there is no clear quantitative relationship between the deuterium isotope shift and the strength of the hydrogen bond.

Table 41: The calculated C=N bond lengths, stretching force constants, frequencies, and rotational barriers in unprotonated and protonated imines.

	$H_2C=NH$		$CH_3HC=NCH_3$		Model	
	C=N	$C=N^+$	C=N	$C=N^+$	C=N	$C=N^+$
bond length (Å)						
HF [111]	1.262	1.282	-	-	-	-
GVB [102]	1.289	1.281	1.289	1.281	-	-
GVB (this work)	1.282	1.281	1.278	1.285	1.284	1.303
force constants						
C=N stretching (mdyn/Å)						
k_{HF} [111]	13.77	13.25	-	-	-	-
k_{GVB} [102]	11.14	11.65	10.60	11.10	-	-
k_{GVB} [a]	10.48	11.19	10.61	10.85	10.17	9.70
k_{GVB}(this work, 6-31G**)	10.49	11.17				
C=N-H bending[a] (mdyn Å/rad^2)	1.12	1.06	-	1.20	-	1.20
C=N/C=N-H coupling[a] (mdyn /rad)	0.43	0.33	-	0.31	-	0.31
stretching freq.[a] (cm^{-1})						
C=N	1628		1688		1659	
$C=N^{15}$	1614		1668		1641	
$^{13}C=N$	1597		1653		1625	
DC=N	1585		1669		1640	
$C=NH^+$		1738		1746		1681
$C=ND^+$		1685		1726		1655
$\Delta\nu_{C=N(H/D)}$		53		20		26
$^{13}C=NH^+$		1723		1718		1655
$DC=NH^+$		1721		1732		1665
rotational barrier[a] (kcal/mol)	70.3	83.8	69.2	81.8	-	-

[a] this is our GVB(6/12)/6-31G* results.

Table 42: The calculated frequencies (cm^{-1}) for $H_2C=NH$.

Modes	HF	MP2	GVB	expt.
NH stretch (a')	3347	3291	3254	3297
CH_2 s-stretch (a')	3012	3091	3018	3036
CH_2 a-stretch (a')	2929	2960	2935	2924
CN stretch (a')	1719	1638	1628	1640
in-plane bend (a')	1465	1448	1453	1453
in-plane bend (a')	1346	1341	1372	1347
torsion (a'')	1143	1101	1133	1123
out-of-plane bend (a'')	1048	1052	1050	1069
in-plane bend (a')	1011	1045	979.4	1059

The GVB results are GVB(6/12)/6-31G* in this study. HF, MP2 and experimental data are from Ref. [5].

5.6.3 C=N Rotational Barrier

Another important factor which can be used to determine the strength of C=N bond is the barrier to rotation around C=N double bond. The twisted structures are considered as the transition state for singlet rotation, as discussed in Section 5.3.

The twisted diradical structures for the unprotonated and protonated methylimine are listed in Table 36, as illustrated in Figure 26. The C-N bond in the twisted structures is about 1.45 Å, which is very close to the MCSCF/3-21G* result [77] of 1.46 Å, showing a typical C-N single bond. The C-NH_2^+ bond in the twisted structure is slightly longer. The twisted structure of protonated trans-$CH_3CH=NCH_3$ is shown in Figure 27. In the twisted structure (Table 38), the C-N bond is a single bond and it becomes longer by 0.015 Å upon protonation.

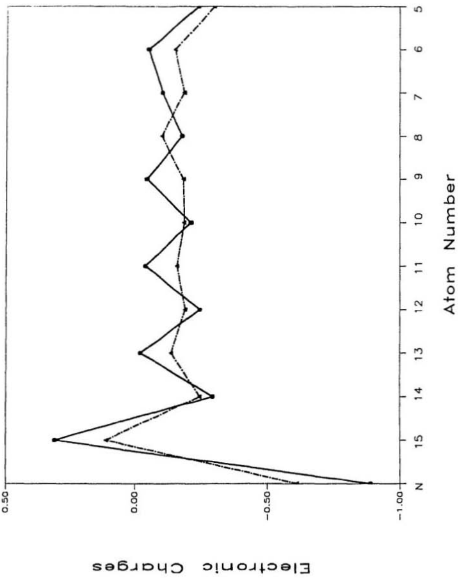
The GVB(6/12)/6-31G* results of the rotational barrier for methylimine and trans-$CH_3HC=NCH_3$ are summarized in Table 41, and their transition state structures are illustrated in Figure 26 and 27, respectively. For methylimine, the barrier is 70.3 kcal/mol, which is in agreement with MCSCF/6-31G* (72.8 kcal/mol) [77]. Upon protonation, the barrier is increased by 13 kcal/mol, indicating an increase in the C=N strength. In trans-$CH_3HC=NCH_3$, the rotational barrier decreases by 1 kcal/mol, while it decreases by 2 kcal/mol upon protonation. The results indicate that the C=N bond energy will decrease with an increase in conjugation.

5.6.4 Electronic Structures

The GVB(6/12)/6-31G geometries for retinal analogs are listed in Table 40. The GVB charge distribution is illustrated in Figure 28. Compared to the early HF/STO-3G results [114], the GVB results give much shorter bond lengths, especially much improved in C=N region. Two main differences between the protonated and unprotonated forms have been observed: (1) the NH_2^+ group causes an increase in conjugation along the chain which is reflected by the equalization of the bond lengths for single and double C-C bonds; and (2) the extra positive charge on NH_2^+ causes a strong charge alternation along the chain. The negative charge increases on nitrogen, but decreases on carbon upon protonation. The increased negative charge on nitrogen of C=N-H will increase the energy of C=N-H bending and thus raise $\nu_{C=N(H)}$ [107,112].

In conclusion, the GVB C=N stretching force constants in unprotonated and protonated imine models, show a decrease by 0.47 mdyn/Å upon protonation. However, a higher C=N(H) stretching frequency has been obtained for protonated imines, due to a strong coupling between the C=N stretching and C=N-H(D) bending modes. On the other hand, protonation causes a strong charge alternation along the chain in all trans polyimines CH_2=(CH-CH=)$_n$NH, and mainly affects the electronic environment of the C=N bond.

Figure 28: The charge distribution for unprotonated (dashed line) and protonated (solid line) polyimine **6**.



Electronic Charges
0.50
0.00
−0.50
−1.00
Atom Number
N
15
14
13
12
11
10
9
8
7
6
5

Chapter Six

Concluding Remarks

In this study, significant computational developments in GVB calculations have been made:

1. The generation of GVB-PP orbitals for specific chemical bonds has obvious computational and conceptual advantages, and results in the automatic generation of reliable initial orbitals.

2. The Boys localized molecular orbitals are good starting orbitals for GVB calculations.

3. The pairwise orbital optimization procedure gives reliable and fast convergence. GVB calculations converged within 10-20 iterations for all molecules examined using different basis sets and GVB pairs.

4. The algorithms described in this work can also be applied to MCSCF calculations.

5. Since the initial orbitals are localized, our approach provides a simple, adequate and consistent description of the potential energy surface.

6. The GVB wavefunction gives a proper description of dipole moments.

The systematic study on GVB equilibrium geometry shows that:

1. As with other calculations (MP2 and CI), small basis sets are not suitable for GVB equilibrium geometries. The addition of d-polarization or diffuse functions significantly reduces the deviation of GVB equilibrium geometry from experiment, but the addition of p-polarization functions on hydrogen does not affect the equilibrium structures significantly. GVB/6-31G* is the simplest calculation to offer reasonable GVB equilibrium geometries, giving a the mean absolute deviation (d) being 0.016 Å for bond lengths (105 comparisons) and 1.1^0 for bond angles (26 comparisons).

2. There is a good linear correlation between GVB/6-31G* and experimental geometrical parameters with a correlation coefficient of 0.999.

3. GVB/6-31G* A-B multiple bond lengths are significantly improved (d is 0.010 Å) over the corresponding HF/6-31G* results (d is 0.025 Å).

4. GVB/6-31G* equilibrium geometries are uniformly in good agreement with other theoretical (HF/6-31G*, MP2/6-31G* and CID/6-31G*) results. Overall, GVB/6-31G* shows improved geometrical parameters over HF/6-31G* with d being 0.018 Å for bond lengths (105 comparisons) and 1.5^0 for bond angles (26

comparisons). GVB/6-31G* results are close to CID/6-31G* and MP2/6-31G* results suggesting that GVB equilibrium geometries are similar to those of MP2 or CID.

The mean absolute deviation of our GVB bond dissociation energies (D_e) for A-H bonds is 15 kcal/mol (9 comparisons), which improves the HF results by 10-20 kcal/mol. GVB D_e values for C-H bonds in hydrocarbons are predicted to within 10 kcal/mol of the experimental values. Overall, our simple GVB wavefunction gives a proper description of the bond dissociation process and reasonably good D_e values, which are much improved over HF and are comparable to other post-HF (MP2 and CI) results. There is a very good correlation between GVB D_e values with the experimental results for all single bonds considered, with a correlation coefficient of 0.976. For first-row molecules, GVB hydrogenation energies are about 0-10 kcal/mol in error, with a mean absolute deviation of 6 kcal/mol. For second-row molecules, the mean absolute deviation is 11 kcal/mol. GVB much improves hydrogenation energies over HF, for molecules containing multiple bonds between second-row elements. For reactions converting multiple to single bonds, GVB heats of reaction are in very good agreement with experimental data, with a mean absolute deviation of 4 kcal/mol. GVB reaction energies correlate well with experimental data, with a correlation coefficient of 0.978.

GVB(6/12)/6-31G** study on the singlet rotation around the double bond of $A_2X{=}YB_2$ (X, Y=C, Si and A, B=H, F), gives the twisted singlet structures and π

bond energies. Our GVB C-C π bond energies are in very good agreement with previous ab initio calculations and available experimental results. The results indicate that the C-C bond energy decreases with increasing fluorine substitution. Pyramidalization at carbon or silicon decrease the energy of the transition state more than that of the reactant, thus decreasing the rotational barrier. In general, pyramidalization at carbon or silicon center from sp^2 (120^0) to sp^3 (109.5^0) stabilizes the diradical and the dipolar transition state.

GVB(6/12)/6-31G** calculations have been carried out to study potential energy curves along the twisting angle and the singlet and triplet structures of $H_2X{=}YLi_2$ (X, Y=C, Si). In triplet 1,1-dilithioethylene, short C-C and Li-Li bonds are observed. In the perpendicular geometries for dilithiosilaethylenes, no silicon pyramidalization is found. GVB relative energies are in excellent agreement with previous CISD results for 1,1-dilithioethylene. The perpendicular triplet is predicted to be the energy minimum structure, and the planar triplet is only 1.33 kcal/mol higher in energy. For silaethylenes, the planar singlet is the ground state, and twisting decreases the singlet-triplet splittings since the triplet becomes lower and the singlet higher during the process of rotation. The negative dipole moment shows $H_2X^+ = ^-CLi_2$ polarity in triplets, indicating that the π bond in the triplet increases the electron density at the CLi_2 group with a strong Li-Li bond. The perpendicular $H_2X{=}CLi_2$ is stabilized by a strong π bonding between XH_2 and CLi_2.

The GVB/6-31++G** calculations on several S_N2 reactions show that the effect of electron correlation on the transition state structures is mainly in the bond lengths. The GVB transition state structures are looser than those from HF and MP2. The calculated GVB frequencies for reactants and products are very close to the experimental results and the scaled HF frequencies. For transition states, GVB gives lower C-H stretching and slightly lower frequencies in some bending modes, as compared to HF. The GVB KIEs are slightly larger than the corresponding HF values. Our GVB results also show that the KIE is determined by changes in the bending vibrations and not by changes in the stretching vibrations. Our GVB results confirm that the magnitude of the KIE is related to the "space" available for the out-of-plane bending vibrations in the transition state, *i.e.*, the distance between the nucleophile and the leaving group in the S_N2 transition state, R_{TS}.

Our GVB(6/12)/6-31G* geometries of methylimine much improve the previous ab initio (HF and GVB) calculations. Generally, the C=N bond in protonated polyimines is longer by 0.02 Å compared to unprotonated polyimines. The calculated GVB force constants for methylimine, and trans-$CH_3HC{=}NCH_3$, and their protonated species, clearly show that the C=N force constant increases upon protonation. The calculated GVB frequencies for methylimine is in very good agreement with the experimental data. GVB results of the C=N stretching force constants for polyimines, show a decrease (by 0.47 mdyn/Å) upon protonation. However, a higher C=N(H) stretching frequency has been obtained, along with a large deuterium isotope shift, due to a strong coupling (0.31 mdyn/rad) between C=N stretching and C=N-H bending. This coupling pushes the C=N stretching to higher frequency. Our GVB calculations provide the theoretical basis for the previous empirical study and support that the C=N(H) stretching frequency deuterium isotope shift can be used as a measure of whether the nitrogen of imines is protonated. This study also shows that protonation increases the C=N π bond energy and changes the electronic environment around C=N bond.

Possible further work is outlined below: (1) developing GVB-CI or MP2-GVB, to obtain accurate potential energy surfaces; (2) implementing this algorithm into MCSCF calculations; (3) including more than two orbitals into one GVB pair, which could be combined with the development of MCSCF calculations; (4) optimizing basis sets which are suitable for GVB calculation, especially for bond dissociation energy; (5) using GVB to study chemical problems, such as the mechanisms in the photochemistry of retinal imines.

Overall, our GVB/6-31G**//GVB/6-31G** calculations demonstrate that the GVB dipole moments, equilibrium geometries, bond dissociation energies, heats of reaction, and vibrational frequencies, are comparable to the CISD and MP2 results. GVB calculations can now be widely used as an excellent post-HF method without the need for integral transformation.

References

1. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Macmillan, New York, 1982.

2. (a) C. Edminston and K. Ruedenberg, *Rev. Mod. Phys.*, **35**, 457 (1963); (b) S. F. Boys, *Rev. Mod. Phys.*, **32**, 232 (1960); (c) S. F. Boys, in *Quantum Theory of Atoms, Molecules and the Solid State*, P. O. Lowdin, Ed., Interscience, New York, 1966, p 253.

3. P. O. Lowdin, *Adv. Chem. Phys.*, **2**, 207 (1959).

4. C. E. Dykstra, *Ab initio Calculation of the Structures and Properties of Molecules*, Elsevier, Amsterdam, 1988.

5. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab initio Molecular Orbital Theory*, Wiley, New York, 1986.

6. I. Shavitt, in *Methods of Electronic Structure Theory*, H. F. Schaefer III, Ed., Plenum, New York, 1977.

7. S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.*, **8**, 61 (1974).

8. C. Møller and M. S. Plesset, *Phys. Rev.*, **46**, 618 (1934).

9. R. J. Bartlett, *Ann. Rev. Phys. Chem.*, **32**, 359 (1981).

10. A. C. Wahl and G. Das, in *Methods of Electronic Structure Theory*, H. F. Schaefer III, Ed., Plenum, New York, 1977.

11. D. B. Cook, *Ab initio Valence Calculations in Chemistry*, Butterworth, London, 1974.

12. (a) P. C. Ladner and W. A. Goddard III, *J. Chem. Phys.*, **51**, 1073 (1969); (b) W. J. Hunt, P.J. Hay, and W. A. Goddard III, *J. Chem. Phys.*, **57**, 738 (1972); (c) P. J. Hay, W. J. Hunt, and W. A. Goddard III, *J. Am. Chem. Soc.*, **94**, 8293 (1972); (d) F. W. Bobrowicz and W. A. Goddard III, in *Methods of Electronic Structure Theory*, H. F. Schaefer III, Ed., Plenum, New York, 1977, p 79.

13. R. B. Murphy and R. P. Messmer, *J. Chem. Phys.*, **98**, 7958 (1993).

14. (a) W. Kolos and L. Wolniewicz, *J. Chem. Phys.*, **46**, 1426 (1967); (b) *J. Chem. Phys.*, **49**, 410 (1968).

15. D. A. Dixon, T. H. Dunning, Jr., R. A. Eades, and P. G. Gassman, *J. Am. Chem. Soc.*, **105**, 7011 (1983).

16. (a) W. A. Goddard III and L. B. Harding, *Ann. Rev. Phys. Chem.*, **29**, 363 (1978); (b) G. Ohanessian and W. A. Goddard III, *Acc. Chem. Res.*, **23**, 386

(1990).

17. (a) G. J. Collin, H. Deslauriers, G. R. De Maré, and R. A. Poirier, *J. Phys. Chem.*, **94**, 134 (1990); (b) G. R. De Maré, E. M. Evleth, R. A. Poirier, and G. J. Collin, *Can J. Chem.*, **72**, 1230 (1994).

18. (a) R. P. Muller, J. M. Langois, M. N. Ringnalda, R. A. Friesner, and W. A. Goddard III, *J. Chem. Phys.*, **100**, 1226 (1994); (b) J. M. Langlois, T. Yamasaki, R. P. Muller, and W. A. Goddard III, *J. Phys. Chem.*, **98**, 13498 (1994).

19. (a) P. Pulay and T. P. Hamilton, *J. Chem. Phys.*, **88**, 4926 (1988); (b) J. M. Coffin and P. Pulay, *J. Phys. Chem.*, **95**, 118 (1991).

20. (a) T. H. Dunning and P. J. Hay, in *Methods of Electronic Structure Theory*, H. F. Schaefer III, Ed., Plenum, New York, 1977; (b) S. Huzinaga, *J. Chem. Phys.*, **42**, 1293 (1965); (c) T. H. Dunning, *J. Chem. Phys.*, **53**, 2823 (1970).

21. P. C. Hariharm and J. A. Pople, *Theoret. Chim. Acta*, **28**, 213 (1973).

22. P. Pulay and G. Fogarsi, *Ann. Rev. Phys. Chem.*, **35**, 191 (1984).

23. R. Krishnan, H. B. Schlegel, and J. A. Pople, *J. Chem. Phys.*, **72**, 4654 (1980).

24. S. Kato and K. Morokuma, *Chem. Phys. Lett.*, **65**, 19 (1979).

25. (a) H. B. Schlegel and M. A. Robb, *Chem. Phys. Lett.*, **93**, 43 (1982); (b) M. Dupuis, *J. Chem. Phys.*, **74**, 578 (1981).

26. W. C. Davidon, *Math. Prog.*, **9**, 1 (1975).

27. M. J. D. Powell, Subroutine VA05AD, AERE Subroutine Library Harwell, Didcot, Berkshire, U. K.

28. J. M. Langlois, R. P. Muller, T. R. Coley, W. A. Goddard III, M. N. Ringnalda, Y. Won, and R. A. Friesner, *J. Am. Chem. Soc.*, **92**, 7488 (1990).

29. R. A. Friesner, *J. Phys. Chem.*, **92**, 3091 (1988).

30. P. Pulay, *J. Comput. Chem.*, **3**, 556 (1982).

31. T. P. Hamilton and P. Pulay, *J. Chem. Phys.*, **84**, 5728 (1986)

32. J. H. Van Lenthe, J. Verbeek, and P. Pulay, *Molec. Phys.*, **73**, 1159 (1991).

33. E. A. Carter and W. A. Goddard III, *J. Chem. Phys.*, **88**, 3132 (1988).

34. R. Murphy, R. A. Friesner, M. N. Ringnalda, and W. A. Goddard III, *J. Chem. Phys.*, **101**, 2986 (1994).

35. L. B. Harding and W. A. Goddard III, *J. Am. Chem. Soc.*, **97**. 6293 (1975);

36. F. Colonna, L.-H. Jolly, R. A. Poirier, J. G. Angyan, and G. Jansen, *Comput. Phys. Commun.*, **81**, 293 (1994).

37. (a) P. A. Schultz and R. P. Messmer, *Phys. Rev. Lett.*, **58**, 2416 (1987); (b) R. P. Messmer and P. A. Schultz, *ibid*, **60**, 860 (1988).

38. E. A. Carter and W. A. Goddard III, *J. Am. Chem. Soc.*, **110**, 4077 (1988).

39. Y. Wang and R. Poirier, *J. Comput. Chem.*, (accepted).

40. M. J. S. Dewar and P. K. Weiner, *Comp. & Chem.*, **2**, 31 (1978), with changes from P. Badziag and D.F. Solms, *Comp. & Chem.*, **12**, 233 (1988).

41. (a) R. A. Poirier, Y. Wang, and C. C. Pye, MUNGAUSS V1.0 (OSIPE Version, see Ref. [36]), Chemistry Department, Memorial University of Newfoundland, St. John's, Newfoundland, Canada (1995); (b) R. A. Poirier and C. D. Keefe, *J. Mol. Struc. (THEOCHEM)*, **234**, 19 (1991).

42. Y. Wang and R. Poirier, *J. Mol. Struc. (THEOCHEM)*, (accepted).

43. Gaussian 92/DFT, Revision F.4, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. W. Wong, J. B. Foresman, M. A. Robb, M. Head-Gordon, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1993.

44. T. H. Dunning, Jr., *J. Chem. Phys.*, **53**, 2823 (1970).

45. P. Saxe, H. F. Schaefer III, and N. C. Handy, *Chem. Phys. Lett.*, **79**, 202 (1981).

46. Y.-R. Luo and J. L. Holmes, *J. Mol. Struc. (THEOCHEM)*, **281**, 123 (1993).

47. H. J. Kohler and H. Lischka, *J. Am. Chem. Soc.*, **104**, 5884 (1982).

48. A. Nicolaides and W. T. Borden, *J. Am. Chem. Soc.*, **114**, 8682 (1992).

49. S. S. Parmar and S. W. Benson, *J. Am. Chem. Soc.*, **111**, 57 (1989).

50. R. M. Pitzer and W. N. Lipscomb, *J. Chem. Phys.*, **39**, 1995 (1963).

51. S. Weiss and G. Leroi, *J. Chem. Phys.*, **48**, 962 (1968).

52. G. Levin and W. A. Goddard III, *J. Am. Chem. Soc.*, **97**, 1649 (1975).

53. G. F. Tantardini and M. Simonetta, *Int. J. Quantum Chem.*, **12**, 131 (1978).

54. S. D. Peyerimhoff and R. J. Buenker, *J. Chem. Phys.*, **51**, 2528 (1969).

55. F. Grimaldi, A. Lecourt, and C. Moser, *Int. J. Quantum Chem.*, **S1**: 153 (1967).

56. J. A. Pople, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III, Plenum, New York, Vol. 4 (1977).

57. (a) W. A. Lathan, W. J. Hehre, L. A. Curtis, and J. A. Pople, *J. Am. Chem. Soc.*, **93**, 6377 (1971); (b) J. A Pople, L. Farnell, and L. Radom, *J. Phys. Chem.*, **87**, 79 (1983).

58. D. J. DeFrees, B. Alevis, S. K. Pollack, W. J. Hehre, J. S. Binkley, and J. A. Pople, *J. Am. Chem. Soc.*, **101**, 4085 (1979).

59. D. J. Defrees, R. Krishnan, H. B. Schlegel, and J. A. Pople, *J. Am. Chem. Soc.*, **104**, 5576 (1982).

60. M. J. Frisch, J. S. Binkley, H. B. Schlegel, K. Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder, and J. A. Pople, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh PA, 1984.

61. (a) M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman, D. A. Ramsay, F. J. Lovas, W. J. Lafferty, and G. A. Maki, *J. Phys. Chem. Ref. Data*, **8**, 619 (1979); (b) K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure*, Van Nostrand Reinhold, New York, (1979).

62. (a) C. W. Bauschlicher and S. R. Langhoff, *Chem. Phys. Lett.*, **177**, 133 (1991); (b) C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, *ibid.*, **171**, 42 (1990).

63. S. R. Langhoff, C. W. Bauschlicher, and P. R. Taylor, *Chem. Phys. Lett.*, **180**, 88 (1991).

64. T. Neuheuser, M. Von Arnim, and S. D. Peyerimhoff, *Theoret. Chim. Acta*, **83**, 123 (1992).

65. T. H. Dunning Jr., *J. Chem. Phys.*, **65**, 3854 (1976).

66. G. R. De Maré and E. M. Evleth, *J. Mol. Struc. (THEOCHEM)*, **259**, 273 (1992).

67. (a) E. A. Carter and W. A. Goddard III, *J. Phys. Chem.*, **91**, 4651 (1987); (b) E. A. Carter and W. A. Goddard III, *J. Chem. Phys.*, **88**, 1752 (1988).

68. B. Rosen, *Spectroscopic Data Relative to Diatomic Molecules*. Pergamon, New York, 1970.

69. C. W. Bauschlicher and S. R. Langhoff, *Chem. Phys. Lett.*, **173**, 367 (1990).

70. C. J. Wu and E. A. Carter, *J. Am. Chem. Soc.*, **112**, 5893 (1990).

71. L. A. Curtiss and J. A. Pople, *J. Chem. Phys.*, **91**, 2420 (1989).

72. J. A. Montgomery and G. A. Detersson, *Chem. Phys. Lett.*, **168**, 75 (1990).

73. J. A. Pople, M. Head-Gordon, and K. Raghavachari, *J. Chem. Phys.*, **87**, 5968 (1987).

74. H. M. Rosenstock, K. Praxl, B. W. Steiner, and J. T. Heron, *J. Phys. Chem. Ref. Data*, 6 (1977), Supplement 1.

75. J. E. Douglas, B. S. Rabinovitch, and F. S. Looney, *J. Chem. Phys.*, **23**, 315 (1955).

76. P. v. R. Schleyer and D. Kost, *J. Am. Chem. Soc.*, **110**, 2105 (1988).

77. M. N. Schmidt, P. N. Truong, and M. S. Gordon, *J. Am. Chem. Soc.*, **109**, 5217 (1987).

78. K. D. Dobbs and W. J. Hehre, *Organometallics*, **5**, 2057 (1986).

79. D. A. Hrovat, H. Sun, and W. T. Bordon, *J. Mol. Struc. (THEOCHEM)*, **163**, 51 (1988).

80. (a) M. Hanamura, S. Nagase, and K. Morokuma, *Tetrahedron Lett.*, **22**, 1813 (1981); (b) M. J. S. Dewar, R. H. Lo, and C. A. Ramsden, *J. Am. Chem. Soc.*, **97**, 1311 (1975).

81. (a) A. F. Voter, M. M. Goodgame, and W. A. Goddard III, *Chem. Phys.*, **98**, 7 (1985); (b) H. Teramae, *J. Am. Chem. Soc.*, **109**, 4140 (1987); (c) S. K. Shin, K. K. Irikura, J. L. Beauchamp, and W. A. Goddard III, *J. Am. Chem. Soc.*, **110**, 24 (1988); (d) E. A. Carter and W. A. Goddard III, *J. Am. Chem. Soc.*, **108**, 2180 (1986); (e) E. A. Carter and W. A. Goddard III, *J. Phys. Chem.*, **88**, 1485 (1984).

82. (a) Y. Yamaguchi, Y. Osamura, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **105**, 7506 (1983); (b) R. J. Buenker and S. D. Peyerimhoff, *Chem. Phys.*, **9**, 75 (1976).

83. (a) R. A. Poirier and J. D. Goddard, *Chem. Phys. Lett.*, **80**, 37 (1981); (b) H. Lishchka and H. J. Kohler, *Chem. Phys. Lett.*, **85**, 467 (1982); (c) Y. Yoshioka, J. D. Goddard, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **103**, 2452 (1981); (d) D. A. Dixon, T. Fukunaga, and B. E. Smart, *J. Am. Chem. Soc.*, **108**, 1585 (1986).

84. (a) B. R. Brook and H. F. Schaefer III, *J. Am. Chem. Soc.*, **101**, 307 (1979); (b) J. D. Goddard, Y. Yoshioka, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **102**, 7644 (1980).

85. (a) H. J. Kohler and H. Lischka, *J. Am. Chem. Soc.*, **104**, 5884 (1982); (b) M. W. Schmidt, M. S. Gordon, and M. Dupuis, *J. Am. Chem. Soc.*, **107**, 2585 (1985).

86. M. J. Fink, M. J. Michalczyk, K. J. Haller, R. West, and J. Michl, *J. Chem. Soc., Chem. Commun.*, 1010 (1983).

87. (a) R. Walsh, *Acc. Chem. Res.*, **14**, 246 (1981); (b) W. J. Pietro and W. J. Hehre, *J. Am. Chem. Soc.*, **104**, 4329 (1982).

88. Y. Apeloig, P. v. R. Schleyer, J. S. Binkley, and J. A. Pople, *J. Am. Chem. Soc.*, **98**, 4332 (1976).

89. S. Nagas and K. Morokuma, *J. Am. Chem. Soc.*, **100**, 1661, (1978).

90. W. D. Laidig and H. F. Schaefer III, *J. Am. Chem. Soc.*, **101**, 7184 (1979).

91. J. A. Morrison, C. Chung, and R. J. Lagow, *J. Am. Chem. Soc.*, **97**, 5015 (1975).

92. (a) K. Krogh-Jespersen, D. Cremer, D. Poppinger, J. A. Pople, P. v. R. Schleyer, and J. Chandrasekhar, *J. Am. Chem. Soc.*, **101**, 4843 (1979); (b) P. H. M. Budzelaar, K. Krogh-Jespersen, T. Clark, and P. v. R. Schleyer, *ibid.*, **107**, 2773 (1985).

93. (a) E. E. Bolton, W. D. Laidig, P. v. R. Schleyer, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **116**, 9602 (1994); (b) Y. Apeloig, T. Clark, A. J. Kos, E.

D. Jennis, and P. v. R. Schleyer, *Isr. J. Chem.*, **20**, 43 (1980); (c) P. v. R. Schleyer, E. Kaufmann, A. J. Kos, T. Clark, and J. A. Pople, *Angew. Chem. Int. Ed. Engl.*, **25**, 169 (1986).

94. A. Dedieu and A. Veillard, in *Quantum Theory of Chemical Reactions*, R. Daudel, A. Pullman, L. Salem, and A. Veillard, Eds., Reidel, Dordrecht, 1979, p 69, and references therein.

95. (a) F. Keil and R. J. Ahlrichs, *J. Am. Chem. Soc.*, **98**, 4787 (1976); (b) Z. Havlas, A. Merkel, J. Kalchev, R. Janoschek, and R. Zahradnik, *Chem. Phys.*, **127**, 53 (1988).

96. (a) Z. Shi and R. J. Boyd, *J. Am. Chem. Soc.*, **111**, 1575 (1989); (b) Z. Shi and R. J. Boyd, *ibid.*, **112**, 6789 (1990); (c) W. Hu and D. G. Truhlar, *J. Phys. Chem.*, **98**, 1049 (1994).

97. (a) S. Wolfe and C. K. Kim, *J. Am. Chem. Soc.*, **113**, 8056 (1991); (b) R. J. Boyd, C. K. Kim, Z. Shi, N. Weinberg, and S. Wolfe, *ibid.*, **115**, 10147 (1993); (c) R. A. Poirier, Y. Wang, and K. C. Westaway, *ibid.*, **116**, 2526 (1994); (d) J. A. Barnes and I. H. Williams, *J. Chem. Soc., Chem. Commun.*, 1286 (1993).

98. (a) S. C. Tucker and D. G. Truhlar, *J. Am. Chem. Soc.*, **112**, 3347 (1990); (b) X. G. Zhao, S. C. Tucker, and D. G. Truhlar, *ibid.*, **113**, 826 (1991).

99. L. Deng, V. Branchadell, and T. Ziegler, *J. Am. Chem. Soc.*, **116**, 10645 (1994), and references therein.

100. K. C. Westaway, in *Isotopes in Organic Chemistry*, E. Buncel and C. C. Lee, Eds., Elsevier, Amsterdam, 1987, Vol. 7, Chapter 5.

101. N. Wiberg, G. Wagner, G. Muller, and J. Riede, *J. Organomet. Chem.*, **271**, 381 (1984).

102. J. J. Lopez-Garriga, S. Hanton, G. T. Babcock, and J. F. Harrison, *J. Am. Chem. Soc.*, **108**, 7251 (1986).

103. L. B. Sims and E. D. Lewis, in *Isotopes in Organic Chemistry*, E. Buncel and C. C. Lee, Eds., Elsevier, Amsterdam, 1985, Vol. 6, Chapter 4.

104. K. C. Westaway and Z. G. Lai, *Can. J. Chem.*, **67**, 345 (1989).

105. Y. Wang and R. A. Poirier, *Can. J. Chem.*, **92**, 1338 (1994).

106. (a) M. Offolenghi, *Adv. Photochem*, **12**, 97 (1980); (b) S. O. Smith, T. A. Pardoen, P. P. J. Mulder, B. Cuny, J. Lugtenburg, and R. Mathies, *Biochemistry*, **22**, 6141 (1983).

107. R. B. Birge, C. M. Einterz, H. M. Knapp, and L. P. Murray, *Biophys. J.*, **53**, 367 (1988), and references therein.

108. J. J. Lopez-Garriga, G. T. Babcock, and J. F. Harrison, *J. Am. Chem. Soc.*, **108**, 7241 (1986).

109. R. H. Callender and B. Honig, *Ann. Rev. Biophys. Bioeng.*, **6**, 33 (1977).

110. (a) D. R. Narva and R. H. Callender, *Photochem. Photobiol.*, **32**, 273 (1980); (b) N. Livnah and M. Sheves, *J. Am. Chem. Soc.*, **155**, 7863 (1993).

111. R. A. Eades, D. A. Weil, M. R. Ellenberger, W. E. Farneth, D. A. Dixon, and C. H. Douglass, Jr., *J. Am. Chem. Soc.*, **103**, 5372 (1981).

112. (a) H. Deng and R. H. Callender, *Biochemistry*, **26**, 7418 (1987); (b) B. Aton, A. G. Doukas, D. Narva, R. H. Callender, V. Dinur, and B. Honig, *Biophys. J.*, **29**, 79 (1980).

113. H. Kakltanl, T. Kakltanl, H. Rodman, B. Honig, R. Callender, *J. Phys. Chem.*, **87**, 3620 (1983).

114. R. A. Poirier, A. Yadav, and P. R. Surjan, *Can. J. Chem.*, **65**, 892 (1987).

115. B. Aton, A. G. Doukas, R. H. Callender, B. Becher, and T. G. Brey, *Biochemistry*, **16**, 2995 (1977).

116. M. A. Marcus, A. T. Lemley, and A. Lewis, *J. Raman Spectrosc.*, **8**, 22 (1979).

117. Z. Meic, G. Baranovic, and T. Suste, *J. Mol. Struc.*, **296**, 163 (1993).

118. S. Masuda, E. H. Msrita, M. Tasumi, T. Iwasa, and M. Tsuda, *J. Mol. Struc.*, **297**, 29, (1993).

119. T. J. Walter and M. S. Braiman, *Biochemistry*, **33**, 1724 (1994).

120. I. Mayer, *Chem. Phys. Lett.*, **97**, 270 (1983).

Appendix A

Objects in OSIPE

The OSIPE structure for the allocation routine and its algorithmic routine is illustrated on page 251. An *object* can only be created by its allocation-routine, which creates nothing else [36]. *Objects* are manipulated in OSIPE mainly by three functions *putobj*, *getobj* and *bldobj*.

- Function *putobj* returns the address of a new *object*, *i.e.*, an *object* to be created and therefore supposedly not existing. Any previously existing version of this Object will be deleted.

  ```
  syntax:  index   = putobj ('Object-name', Cardinal, 'Type', Bytes)
  ```

  For example, the GVB orbital coefficients are created by,

  ```
  CMOA= putobj ('MATSQG_COEFF_CMO_AO_GVB', NBASIS*NBASIS, 'REAL', L8)
  ```

- Function *getobj* returns the address of an existing *object* needed to create a new *object*.

```
syntax: index = getobj ('Object-name')
```

If the existing *object* is not present in memory or on disk, *getobj* activates its creation by calling a third function, *bldobj*.

```
else if (object-name. eq. 'XXX') then
     call ALLOC_O ('object-name')
```

Adding a new feature to a program merely involves adding a new item to *bldobj*. The program is then able to automatically build this new *object* whenever it is required.

For example, the GVB initial orbital coefficients are obtained by,

```
CMOG=getobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
if(lbuild)CMOG=bldobj ('MATSQG_COEFF_CMO_AO_GUESS_GVB')
```

Getobj returns the address of *CMOG* needed to create *CMOA*. If *CMOG* is not existing, it will be created by calling *bldobj*.

```
else if(nameL(1:ln).eq.'MATSQG_COEFF_CMO_AO_GUESS_GVB ')then
     call GVBMOG
```

Two types of routines deal with *scalars*: *putscX* and *getscX*, where *X* stands for *Boolean*, *Character*, *Integer* or *Real* depending on the *scalar* type. These routines are used to create a *scalar object* with a given value or return its value. For example,

```
call getscI ('SCF_SCI_NUM_OF_CONFIG', NCONF)

call getscR ('GVB_SCR_SCA_FACTOR' , SCALST)

call getscC ('GVB_SCC_ROT_METHOD', ROTMET)

call getscB ('GVB_SCB_ROTATION_METHOD', ROTMIX)

call putscR ('ENERGY_SCR_ELECTRONIC_GVB', EELEC)

call putscR ('ENERGY_SCR_TOTAL', ETOT)
```

Allocation of default values and automatic building of *scalars* is done by a function *bldscl* organized similarly to the function *bldobj* used to build *array objects*.

```
allocation-routine:
      subroutine ALLOC_O (Object-name)
      include 'stack.h'          ! contains array stack
      character*(*) Object-name
      character*(maxstr) O, O1, O2
------------------
      lbuild = .false.            ! initialize
* build existing-Object names from modalities of Object-name (if any):
     O1 = function-of (Object-name)
     O2 = function-of (Object-name)
* gather indices:
     index_O1 = getobj (O1)     ! set lbuild=.true. if O1 does not exist
     if(lbuild) index_O1 = bldobj (O1)    ! automatic creation if needed
     index_O2 = getobj (O2)
     if(lbuild) index_O2 = bldobj (O2)
* creation:
     index_O  = putobj (Object-name, dimension_O, 'REAL', 8)
* gather addresses:
     address_O  = objadd(index_O)
     address_O1 = objadd(index_O1)
     address_O2 = objadd(index_O2)]
-------------------------------------------------
* computation:
      call CALC_O (Stack(address_O),
     .             Stack(address_O1),
     .             Stack(address_O2),
     .             dimension_O,
     .             dimension_O1,
     .             dimension_O2)
      end
 algorithmic-routine:
      subroutine CALC_O (O ,
     .                   O1 ,
     .                   O2,
     .                   dimension_O,
     .                   dimension_O1,
     .                   dimension_O2)
      double precision O(dimension_O)
      double precision O1(dimension_O1)
      double precision O2(dimension_O2)
             ... Non-declarative part of the code ...
          O = function-of (O1, O2)
      end
```

Appendix B

The Calculation of Kinetic Isotope Effects

The kinetic isotope effect (KIE) is defined as the ratio of k/k^*.

$$\mathrm{KIE} = k/k^* \tag{130}$$

where k and k^* are the rate constants for reactant R and its isotopic isomer R* respectively. KIEs are very important for studying reaction mechanisms, especially for the transition state structures.

In terms of transition state theory and statistical thermodynamics, the expression of calculating kinetic isotope effects is:

$$\mathrm{KIE} = \mathrm{MMI} \times \mathrm{EXC} \times \mathrm{ZPE} \times Q_t \tag{131}$$

where

- MMI - mass-moment of inertia term, represents a factor based on structural and mass differences of the isotopic isomers; which can be expressed as a product of the energy differences due to isotopic substitution that arises from rotation and translation of the reactant and transition state as expressed:

$$\mathrm{MMI} = \frac{(M/M^*)^{\frac{3}{2}} \prod_i^{n_{rot}} (I_i/I_i^*)^{\frac{1}{2}})_{trans.}}{(M/M^*)^{\frac{3}{2}} \prod_i^{n_{rot}} (I_i/I_i^*)^{\frac{1}{2}})_{reac.}} \tag{132}$$

where M and M^* are the mass, I and I^* are the moment of inertia and n_{rot} is the number of rotational degrees of freedom.

- ZPE – zero-point energy contribution;

$$ZPE = \frac{[exp(-\sum_{i=1}^{n_{vib}} \Delta u_i/2)]_{trans.}}{[exp(-\sum_{i=1}^{n_{vib}} \Delta u_i/2)]_{reac.}} \tag{133}$$

$$\Delta u_i = u_i - u_i^* \tag{134}$$

where $u_i = h\nu_i/kT$, ν_i are frequencies, and n_{vib} is the number of vibrational frequencies.

- EXC – vibrational excitation term;

$$EXC = \frac{[\prod_{i=1}^{n_{vib}}(1-e^{-u_i^*})/(1-e^{-u_i})]_{trans.}}{[\prod_{i=1}^{n_{vib}}(1-e^{-u_i^*})/(1-e^{-u_i})]_{reac.}} \tag{135}$$

- The kinetic tunneling correction factor (Q_t) was determined by the Wigner's correlation equation,

$$Q_t = 1 + u_i^2/24 + 7u_i^4/5760 \tag{136}$$

In this study, a modified version of Sims' BEBOVIB-IV program (QCPE No.337, c.f. [103]) has been developed, so that it can calculate the normal mode vibrational frequencies and KIEs directly using ab initio (SCF, GVB, MP2, etc.) optimized geometries and force constants [104]. The descrption of data input is briefly given in next page. A sample input is also given.

```
! This is a reference to input example for using BEBOVIB to obtain
! the frequencies and KIEs.
! ------- start here --------------
    The calculation of KIE using the BEBOVIB program
                 10 October,1991
!   the first two line is a title!
  5    9    0    2    1    1 298.00000
!  NOAT(I5) -- The number of atoms
!  NOCORD(I5) --  The number of internal coordinates
!  NLIBE(I5) -- The number of linear bending coordinates
!  NISO(I5) -- Th number of different isotopic reactants
!             (including normal, NISO=2)
!  NCPLX(I5) -- The number of reactant models (1)
!  NTEMP(I5) -- The number of temperatures
!  TEMPT(F10) -- temperatures
  1   0.10000
!  NPRINT(I5) -- set as 1
!  CUTOFF(F10) -- below this frequencies ignored
 Z-Matrix from ab initio (cut the output from MUNGAUSS directly)
 Force constants from ab initio (cut the output from MUNGAUSS directly)
  5    1    0
! NA(I5) -- the number of atoms
! IFVECT(I5) -- a control for print ( 0  1  2 )
! NCYCL(I5) -- the number of cyclic moieties in the model
 12.00000  34.96885   1.00783   1.00783   1.00783
 12.00000  34.96885   2.01410   2.01410   2.01410
! Atomic weight for isotopes
                       normal
                       (h3-h4-h5) D
!   title for output
     Isotop          sdddd1

 Cartesian coordinates (cut the output from MUNGAUSS directly)
0
! check frequencies order, 0 means no changes
! otherwise input the new order of frequencies
! use for calculate the contribution.
 Following by the input for the complex.
```

```
      The calculation of KIE (hf/6-31+g*) using the BEBO program
                   data  for  reactants CH3Cl
   5    9    0    2    1     1 298.00000
   1    0.10000
   1    0  0.000000   0    0.00000    0    0.00000
   2    1  1.786354   0    0.00000    0    0.00000
   3    1  1.077917   2 108.39700    0    0.00000
   4    1  1.077917   2 108.39700    3 120.00000
   5    1  1.077917   2 108.39700    3-120.00000
 0.248986  0.005698  0.005698  0.005698  0.056753  0.056752
 0.056752  0.000000  0.000000
 0.392879  0.002869  0.002869 -0.012519 -0.006115 -0.006115
 0.015399 -0.015399
 0.392879  0.002869 -0.006115 -0.012519 -0.006115 -0.000001
 0.015399
 0.392879 -0.006115 -0.006115 -0.012519 -0.015399  0.000001
 0.260580  0.037599  0.037599 -0.054086  0.054086
 0.260581  0.037600 -0.000001 -0.054086
 0.260581  0.054086  0.000001
 0.207015 -0.103509
 0.207015
   5    1    0
 12.00000  34.96885   1.00783   1.00783   1.00783
 12.00000  34.96885   2.01410   2.01410   2.01410
                        normal
                        (h3-h4-h5) D
     Isotop           sdddd1

 1       C-1  0.000000  0.000000  0.000000
 2        17  0.000000  0.000000  1.786354
 3         1  1.022829  0.000000 -0.340186
 4         1 -0.511425 -0.885789 -0.340186
 5         1 -0.511425  0.885789 -0.340186
0
```

The calculation of KIE (hf/6-31+g*) using the BEBO program
data for transition state Cl-CH3-Cl

```
    6   12    0    2    1    1 298.00000
    1    0.10000
    1    0  0.000000   0   0.00000    0    0.00000
    2    1  2.393962   0   0.00000    0    0.00000
    3    1  1.061534   2  90.00000    0    0.00000
    4    1  1.061534   2  90.00000    3 120.00100
    5    1  1.061534   2  90.00000    3-120.00100
    6    1  2.393962   3  90.03100    2 180.00000
 0.029621 -0.002717 -0.002720 -0.002720  0.026934  0.038001
 0.038010  0.038010  0.000006  0.000000  0.000000  0.000000
 0.447139 -0.000367 -0.000367 -0.002715 -0.013117  0.000000
 0.000000 -0.013113  0.014543 -0.014543  0.000000
 0.447137 -0.000367 -0.002721  0.006564  0.000002 -0.000001
 0.006564  0.000004  0.014540  0.011370
 0.447137 -0.002721  0.006564 -0.000001  0.000002  0.006564
-0.014540 -0.000004 -0.011370
 0.029621 -0.038042 -0.038032 -0.038032 -0.000092 -0.000005
 0.000005  0.000000
 0.177039  0.062491  0.062491  0.076255 -0.000515  0.000515
 0.000000
 0.177380  0.024186  0.038290  0.000010  0.000010  0.066348
 0.177380  0.038290 -0.000010 -0.000010 -0.066348
 0.152798 -0.000522  0.000522  0.000000
 0.220526 -0.110263  0.000314
 0.220526  0.000314
 0.152895
   6    1    0
 12.00000  34.96885   1.00783   1.00783   1.00783  34.96885
 12.00000  34.96885   2.01410   2.01410   2.01410  34.96885
    Isotop          sdddd1

  1       C-1  0.000000   0.000000  0.000000
  2        17  0.000000   0.000000  2.393962
  3         1  1.061534   0.000000  0.000000
  4         1 -0.530775  -0.919311  0.000000
  5         1 -0.530775   0.919311  0.000000
  6        17 -0.001304   0.000000 -2.393962
0
```