

Rule-Based Trial Wave Functions for Generalized Valence Bond Theory

Jean-Marc Langlois,[†] Terumasa Yamasaki,[‡] Richard P. Muller,[§] and William A. Goddard III*

Materials and Molecular Simulation Center, Beckman Institute (139–74), Division of Chemistry and Chemical Engineering (CN 8808), California Institute of Technology, Pasadena, California 91125

Received: June 20, 1994; In Final Form: October 4, 1994[⊗]

We present a general method suitable for automatic generation of trial wave functions for generalized valence bond (GVB) descriptions of large molecules. This method uses pseudo-Hartree–Fock (P-HF) molecular orbitals formed from HF atomic orbitals but *without* Fock matrix diagonalization. The occupied P-HF orbitals are projected onto atomic basis functions to obtain GVB first natural orbitals, and the unoccupied HF orbitals are projected to obtain GVB second natural orbitals. This method (denoted GVB-INIT) is fast because no HF wave functions need be calculated and because the localization is piecewise atomic. In conjunction with the recently developed GVB-DIIS method for converging GVB wave functions and the new pseudospectral programs (PS-GVB) for the Fock matrix elements, GVB-INIT makes calculation of highly correlated GVB wave functions quite practical. The efficacy of GVB-INIT is illustrated by application to several cases including GVB wave functions with up to 26 correlated pairs.

1. Introduction

Essentially all methods for calculating the electronic wave functions (i.e., solving the Schrödinger equation) of atoms, molecules, and solids use a first step involving optimization of orbitals (atomic orbitals, molecular orbitals, or band orbitals). The most common starting point for accurate wave functions is the Hartree–Fock (HF) method in which a product of orbitals and spins is antisymmetrized (using a Slater determinant) and the orbitals optimized self-consistently (SCF for self-consistent field). This might then be used as the starting point for obtaining more accurate wave functions [configuration interaction (CI) or perturbation theory] by including excitations where various numbers of electrons are excited from the occupied orbitals of the HF wave function to unoccupied or virtual orbitals.

An alternative approach is the generalized valence bond method^{1–5} (GVB) in which each *pair* of electrons is described with *two* orbitals (rather than one, as in HF), the wave function is generalized from a Slater determinant to a form that ensures that both the Pauli principle and spin symmetry are satisfied, and the orbitals are optimized self-consistently. In general, this GVB approach leads to twice the number of orbitals to be optimized as in HF. The GVB orbitals generally localize uniquely to regions near a bond or atom, and each pair can be viewed as involving one orbital very similar to a localized occupied HF orbital and the other involving a localized unoccupied (virtual) HF orbital.^{1,2} As with HF, many GVB calculations are followed by some kind of CI. However, since the dominant correlating orbitals have been solved self-consistently, the GVB wave function often allows⁶ an accurate description of the wave function to be obtained with very compact CI's.

The two-electron interactions lead to nonlinearity in the optimization of wave functions. Hence, critical to such SCF methods as HF or GVB is prediction of good trial wave

functions followed by an iterative approach that converges quickly and reliably to the desired state.

For HF wave functions the trial wave function is generally taken from an approximate or semiempirical HF based method such as extended Hückel, CNDO/INDO, MINDO, etc. This calculation is very fast and generally leads to adequate trial orbitals. However, for GVB there is not yet an analogous semiempirical method. Generally, GVB calculations have involved constructing localized orbitals from localized HF orbitals (both occupied and unoccupied). However, these methods are not automatic and tend to be unreliable. In addition, the HF and localization calculations can be almost as expensive as the GVB calculations, making the total process unacceptably expensive.

We present here a general rule-based approach, GVB-INIT, that solves the problem of making reliable initial guesses, providing excellent trial functions for general GVB wave functions. Combined with the GVB-DIIS method for converging wave functions,⁷ this leads finally to an automatic procedure for calculating GVB wave functions.

Section 2 reviews the relevant aspects of the GVB wave function, and section 3 outlines the new procedure. The results are discussed in section 4.

2. The GVB-PP Wave Function

The GVB wave function generally involves a product of orbitals with a general spin function χ that is an eigenfunction of the total spin operator [$S^2\chi = S(S+1)\chi$]

$$\Phi^{\text{GVB}} = \mathcal{A}[\phi_1 \dots \phi_N \chi] \quad (2.1)$$

where \mathcal{A} is the antisymmetrizer and both the orbitals and spin functions are optimized. More commonly, the perfect pairing (PP) restriction² is made in which each pair of electrons is written as a valence bond (VB) pair

$$(\phi_a \phi_b + \phi_b \phi_a)(\alpha\beta - \beta\alpha) \quad (2.2)$$

The GVB-PP formalism allows any number of electrons ($2n_{\text{core}}$) to be described with doubly-occupied orbitals (Φ_{core}) and any number of electrons, (n_{open}) to be high-spin coupled ($S = n_{\text{open}}/2$), leading to the composite wave function

[†] Current address: Schrödinger Inc., 80 South Lake Avenue, Suite 735, Pasadena, CA 91101.

[‡] Current address: Asahi Chemical Industry Co., Ltd., Computational Science Department, 2-1 Samejima, Fuji, Shizuoka 416, Japan.

[§] Current address: Department of Chemistry, University of Southern California, Los Angeles, CA.

[⊗] Abstract published in *Advance ACS Abstracts*, December 1, 1994.

$$\Psi = \mathcal{A}[\Psi_{\text{core}} \Psi_{\text{open}} \Psi_{\text{pair}}] \quad (2.3)$$

where

$$\Psi_{\text{core}} = \prod_{i=1}^{n_{\text{core}}} (\phi_i^c \alpha) (\phi_i^c \beta) \quad (2.4)$$

$$\Psi_{\text{open}} = \prod_{i=1}^{n_{\text{open}}} (\phi_i^o \alpha) \quad (2.5)$$

$$\Psi_{\text{pair}} = \prod_{i=1}^{n_{\text{pair}}} (\psi_{ai} \psi_{bi} + \psi_{bi} \psi_{ai}) (\alpha\beta - \beta\alpha) \quad (2.6)$$

Here $\{\phi_i^c\}$ and $\{\phi_i^o\}$ are orthogonal spacial orbitals and α and β are spin functions.

The GVB orbitals in pair i are denoted as ψ_{ai} and ψ_{bi} . Within a pair, the GVB orbitals overlap, $S_i = \langle \psi_{ai} | \psi_{bi} \rangle$. However, the PP restriction requires that orbitals of pair i be orthogonal to all other pairs (and to $\{\phi_i^c\}$ and $\{\phi_i^o\}$). For computational purposes it is convenient to rewrite the GVB wave function (2.6) as²⁻⁵

$$\Psi_{\text{pair}} = \prod_{i=1}^{n_{\text{pair}}} (C_{gi} \phi_{gi} \phi_{gi} + C_{ui} \phi_{ui} \phi_{ui}) (\alpha\beta - \beta\alpha) \quad (2.7a)$$

where

$$\phi_{gi} = \frac{\psi_{ai} + \psi_{bi}}{\sqrt{2(1+S_i)}} \quad \text{and} \quad \phi_{ui} = \frac{\psi_{ai} - \psi_{bi}}{\sqrt{2(1-S_i)}} \quad (2.7b)$$

are orthogonal GVB natural orbitals. Here the GVB configuration interaction (CI) coefficients satisfy

$$\frac{C_{ui}}{C_{gi}} = -\frac{1-S_i}{1+S_i} \quad (2.8)$$

and

$$C_{gi}^2 + C_{ui}^2 = 1 \quad (2.9)$$

3. Initial Guesses for GVB Wave Functions, GVB-INIT

The generation of the trial wave function is composed of two steps: (i) In the first step a pseudo-HF (P-HF) trial wave function (HF-INIT) is generated based on *atomic* HF orbitals. (ii) In the second step, the GVB trial orbitals are obtained by piecewise localization of the HF trial wave function to form GVB-INIT. This is our standard GVB initial guess. Alternatively, one could use (i) as the starting point for a molecular HF-SCF calculation, which could in turn be used in place of the HF trial function for step two. However, use of the HF optimum orbitals in (ii) generally saves only about one or two iterations in the GVB-SCF calculation. In addition, for stretched bonds or transition metal systems use of HF-SCF orbitals sometimes gives worse results than using atomic based orbitals. Hence, the extra expense of the HF SCF calculation is not justified and is *not* normally a part of GVB-INIT.

3.1. HF-INIT. We find that the HF-SCF orbitals of the individual atoms can be used to form an HF trial wave function for a molecule. This is plausible since the molecular density is not dramatically different from the sum of the atomic densities. The problem is how to go from a set of atomic SCF orbitals (SCF-AO) to a set of molecular orbitals (MO) without actually solving the HF equations for the molecular orbitals.

The procedure we use is the following:

(α) *Atomic HF Orbitals* $\{\theta_n^A\}$. For each atomic basis set $\{\chi_\mu^A\}$, where A is the atom label, we define orthonormal HF atomic orbitals $\{\theta_n^A\}$ obtained from occupation-averaged atomic Hamiltonians (so that orbitals of the same symmetry are equivalent). This is done once and stored with the basis set. The atomic occupation number $\{f_n^A\}$ is associated with each HF-AO. The orbitals are then sorted into core orbitals, $\{\theta_n^{\text{core},A}\}$, valence orbitals, $\{\theta_n^{\text{val},A}\}$, and unoccupied orbitals, $\{\theta_n^{\text{inocc},A}\}$. The core orbitals are (1s) for the C row elements and (1s,2s,2p) for the Si row elements. The valence orbitals are (1s) for hydrogen atoms, (2s,2p) for C row elements, and (3s,3p) for Si row elements.

(β) *VB Assignment.* We next partition the electrons of the system in terms of bond pairs, lone pairs, and core orbitals just as in a simple valence bond structure. This can be done (i) automatically using the autotyping features of commercial molecular dynamics software (e.g., BIOGRAF/POLYGRAF⁸), (ii) by drawing a molecule on a screen and attaching bonds, or (iii) by inputting a list of atoms to be involved in various bonds, lone pairs, etc. GVB-PP is oriented toward molecules with one dominant valence bond structure, and we consider only this case in this paper. GVB does *not* require that every bond pair be correlated. In this case one designates which bonds are to be correlated and second natural orbital (step section 4) are done only for these.

(γ) *Build Occupied HF-MO's* $\{\phi_n^{\text{HF}}\}$.

($\gamma 1$) *Core HF MO's.* Each core HF-AO $\{\theta_n^{\text{core}}\}$ is considered as a doubly-occupied HF MO $\{\phi_n^{\text{core}}\}$.

($\gamma 2$) *Valence HF MO's.* The remaining valence HF-AO's $\{\theta_n^{\text{val}}\}$ are used to construct a subset of HF-MO's $\{\phi_n^{\text{val}}\}$ as follows: We construct the overlap matrix over all valence HF-AO's $\{\theta_n^{\text{val}}\}$,

$$S_{ij} = \langle \theta_i^{\text{val}} | \theta_j^{\text{val}} \rangle \quad (3.1)$$

where ij go over all valence atomic orbitals on all atoms. This matrix is diagonalized, and the eigenvectors with the largest eigenvalues are selected. The number to be selected, n_{occ} , is obtained by adding the atomic occupation numbers for the various atoms and correcting by the total charge of the molecule. The resulting $\{\phi_n^{\text{val}}\}$, which are not orthogonal, include the combinations of HF-AO having the largest overlap, a good first approximation to the best bonding orbitals.

($\gamma 3$) *Form HF-INIT by Orthonormalizing the HF-MOs* $\{\phi_n\}$. Since the core orbitals were not included in the diagonalization of the overlap matrix in $\gamma 2$, we now order the $\{\phi_n\}$ with the core orbitals first (ordered by atomic orbital energy), then the valence orbitals (ordered by eigenvalue from diagonalizing the valence AO overlap matrix), and then Schmidt orthogonalize. We refer to this set of occupied HF orbitals as the HF-INIT space. It is an approximation to the HF wave function of the molecule (which we do *not* calculate).

($\gamma 4$) *Create a New Set of Orthogonal Orbitals* $\{\eta_n^A\}$ *on Each Atom.* Here we want the atomic orbitals ordered according to their projection onto HF-INIT. To do this we construct the atomic density matrix over AO's

$$P_{\mu\nu}^{AA} = \langle \chi_\mu^A | \rho_{\text{HF}} | \chi_\nu^A \rangle \quad (3.2)$$

where ρ_{HF} is the electron density of the occupied HF-MO's

$$Q_{\text{HF}} = \sum_{n=1}^{n_{\text{occ}}} |\phi_n^{\text{HF}}\rangle \langle \phi_n^{\text{HF}}| \quad (3.3)$$

The associated eigenvalue problem to be solved is

$$P^{AA} \mathbf{c} = S^{AA} \mathbf{c} \gamma^{AA} \quad (3.4a)$$

where

$$S_{\mu\nu}^{AA} = \langle \chi_\mu^A | \chi_\nu^A \rangle \quad (3.4b)$$

is the overlap matrix over atomic basis functions. The HF-projected orthogonal orbitals $\{\eta_n^A\}$ are associated with the eigenvectors $\{c_n\}$. Near-unity eigenvalues indicate large projections onto the HF occupied orbitals space. At this point the orbitals in $\{\eta_n^A\}$ are sorted by projections into core orbitals, $\{\eta_n^{\text{core},A}\}$, valence orbitals, $\{\eta_n^{\text{val},A}\}$, and unoccupied orbitals, $\{\eta_n^{\text{unocc},A}\}$. The valence atomic orbitals will define the space used to create the GVB orbitals for both bond pairs (section III.B) and lone pairs (section II.C). For the 2nd GVB natural orbitals, the unoccupied atomic orbitals will be used (section III.C).

3.2. GVB Bond Pairs. GVB-INIT uses a piecewise localization to obtain trial GVB orbitals. Each pair of GVB orbitals is generated independently. They are later combined and symmetrically orthogonalized.

The most common case is a GVB bond pair localized between two nearest-neighbor atoms. This is described by two GVB orbitals $[\psi_{ai}$ and ψ_{bi} of (2.6)], one localized on each atom, which overlap. We construct instead the GVB natural orbitals $[\phi_{gi}$ and ϕ_{ui} of (2.7)], which are orthogonal.

The task here is to choose the localized GVB natural orbitals as suitable linear combinations of HF-MO's. The most important characteristic of the GVB orbitals forming a bond pair, $(\psi_{ai}$ and $\psi_{bi})$, is the large overlap between them. The GVB orbitals can thus be found by searching for the linear combination of the valence atomic projected HF orbitals $\{\eta_n^{\text{val},A}\}$ (which will form ψ_{ai}) and $\{\eta_n^{\text{val},B}\}$ (which will form ψ_{bi}) so that ψ_{ai} and ψ_{bi} have maximum probability of sharing electrons.⁹ The procedure is as follows:

(δ) Consider one by one each bond and its pair of atoms (A and B). We will obtain one pair of orbitals $\{\psi_{ai}\}$ and $\{\psi_{bi}\}$ for each bond between A and B.

($\delta 1$) Evaluate the off-diagonal density matrix elements of the occupied HF-MO's in the basis of the orthogonal valence atomic orbitals $\{\eta_i^{\text{val},A}\}$ and $\{\eta_j^{\text{val},B}\}$. This defines the bond space from which we create the overlapping GVB orbitals:⁹

$$P_{ij}^{AB} = \langle \eta_i^{\text{val},A} | Q_{\text{HF}} | \eta_j^{\text{val},B} \rangle \quad (3.5)$$

($\delta 2$) Find the pair of unitary transformations, U^A and U^B , that makes each orbital on atom A overlap with only one orbital on atom B. This biorthogonalization process¹⁰ transforms the density matrix P^{AB} into a diagonal matrix Λ^{AB}

$$(U^A)^\dagger P^{AB} U^B = \Lambda^{AB} \quad (3.6)$$

where the largest diagonal element of Λ indicates the largest overlap for the pair of localized orbitals (one each on atom A and B). Since the matrix P^{AB} is always nonsymmetric, (3.6) is solved by singular value decomposition.

($\delta 3$) Starting with the largest singular value of Λ^{AB} , select GVB-like orbitals $(\bar{\psi}_{ai}$ and $\bar{\psi}_{bi})$ as biorthogonal pairs. [Here the bar indicates that the $\bar{\psi}_{ai}$ contains only basis functions on center A.] For multiple-bonded cases, select the appropriate

number of pairs. Generally the highest Λ^{AB} corresponds to a σ bond, the next two largest eigenvalues correspond to π bonds, the next two to δ bonds, etc.

($\delta 4$) The two overlapping GVB orbitals of each pair ($\bar{\psi}_{ai}$ and $\bar{\psi}_{bi}$) from $\delta 3$ are combined to form the GVB natural orbitals $[\phi_{gi}$ and ϕ_{ui} of (2.7)]. The actual GVB orbitals $(\psi_{ai}$ and $\psi_{bi})$ will be linear combinations of the atomic $\bar{\psi}_{ai}$ and $\bar{\psi}_{bi}$ ($\psi_{ai} = \bar{\psi}_{ai} + \lambda_a \bar{\psi}_{bi}$ and $\psi_{bi} = \bar{\psi}_{bi} + \lambda_b \bar{\psi}_{ai}$ where $\lambda_i \approx 0.1$) and thus the GVB-NO's, $(\phi_{gi}$ and $\phi_{ui})$, will be combinations of $\bar{\psi}_{ai}$ and $\bar{\psi}_{bi}$ as in (2.7) but with coefficients that are not ± 1 . The 1st GVB-NO is written as $\phi_{gi} = c_a \bar{\psi}_{ai} + c_b \bar{\psi}_{bi}$ where the coefficients c_a and c_b are chosen to produce the largest projection onto the occupied HF space. The 2nd GVB-NO, ϕ_{ui} , is just the orthogonal linear combination. This is done by solving the corresponding 2×2 eigenvalue problem:

$$P^{\text{GVB}} \mathbf{c} = S^{\text{GVB}} \mathbf{c} \lambda^{\text{GVB}} \quad (3.7a)$$

where

$$P_{\alpha\beta}^{\text{GVB}} = \langle \bar{\psi}_{ai} | Q_{\text{HF}} | \bar{\psi}_{bi} \rangle \quad (3.7b)$$

and

$$S_{\alpha\beta}^{\text{GVB}} = \langle \bar{\psi}_{ai} | \bar{\psi}_{bj} \rangle \quad (3.7c)$$

Generally the projection of the 1st GVB-NO is very large ($\lambda_{1\text{st}}^{\text{GVB}} > 0.95$), and we replace the GVB-NO by its projection to the occupied HF-MO space. On the other hand, the projection of the 2nd GVB-NO is very small. Thus, the 1st GVB-NO can be viewed as a linear combination of the initial occupied HF-MO's whereas the 2nd GVB-NO corresponds to some linear combination of the initially unoccupied MO's.

3.3. GVB Lone Pairs. For a molecule well described in terms of a single valence bond structure, the above procedure leads to an excellent starting guess for a full GVB wave function. In this case lone pair orbitals (and core orbitals) of the wave function remain doubly occupied (not correlated). However, for many purposes it is also desirable to correlate the lone pairs. In this case, the 1st GVB-NO consists of a localized lobelike orbital (say, an sp^3 -like hybrid corresponding to an orbital centered off the atom). The 2nd GVB-NO has the same overall direction and location, but with a nodal plane passing through the middle of the first NO, leading to an off-center p-like orbital. With GVB-INIT we choose the lobe 1st GVB-NO from a linear combination of the atomic valence orbitals $\{\eta_n^{\text{val}}\}$ not previously used in making GVB bond-pair orbitals. The lobe 2nd GVB-NO is chosen as the linear combination of unoccupied atomic orbitals $\{\eta_n^{\text{unocc}}\}$ having the largest gradient matrix element with the lobe 1st GVB-NO. The overall method is summarized as follows:

($\epsilon 1$) Evaluate the overlap matrix between the valence orthogonal atomic orbitals $\{\eta_i^{\text{val},A}\}$ in the subspace of the GVB bond-pair orbitals involved with this atom:

$$P_{ij}^{AA} = \delta_{ij} - \langle \eta_i^A | \sum_{n=1}^{n_{\text{pair}}} |\bar{\psi}_{an}\rangle \langle \bar{\psi}_{an}| | \eta_j^A \rangle \quad (3.8)$$

where n_{pair} is the number of GVB pairs involving atom A and $\{\bar{\psi}_{an}\}$ are the GVB orbitals. The matrix P^{AA} is diagonalized, and the appropriate number of lobe 1st GVB-NO's, $\{\phi_{gi}^{\text{lobe}}\}$, are chosen as the eigenvectors with the largest eigenvalue. If there is more than one (e.g., O, F), they are Boys localized to obtain sp^3 -like orbitals.

($\epsilon 2$) The lobe 2nd GVB-NO's, $\{\phi_{li}^{\text{lobe}}\}$, are formed by linear combinations of unoccupied orthogonal atomic orbitals

$\{\eta_i^{\text{unocc},A}\}$ having the maximum overlap with the gradient of the 1st GVB-NO's:

$$\left\langle \phi_{ui}^{\text{lobe}} \left| \frac{\bar{R}_{\text{lobe}} \cdot \bar{\nabla}}{|r - R_{\text{lobe}}|^3} \right| \phi_{gi}^{\text{lobe}} \right\rangle \quad (3.9)$$

where the \bar{R}_{lobe} are the centroids of the lobe first GVB-NO's:

$$\bar{R}_{\text{lobe}} = \langle \phi_{gi}^{\text{lobe}} | \bar{r} | \phi_{gi}^{\text{lobe}} \rangle \quad (3.10)$$

The $1/|r - R_{\text{lobe}}|^3$ term was found useful to restrict the effect of the gradient to a small region of space near the centroid location. For computational simplicity the $1/r^3$ term was approximated by an expansion over eight Gaussian s functions (see Table 7).

3.4. Symmetric Orthogonalization. So far we have described how to obtain each pair of GVB natural orbitals independently. We must now merge these GVB orbitals with the existing HF orbitals. The procedure is as follows.

(λ1) Symmetrically orthogonalize all 1st GVB-NO's $\{\phi_{gi}\}$; then symmetrically orthogonalize all 2nd GVB-NO's $\{\phi_{ui}\}$.

(λ2) Form the remaining occupied orbitals by projecting out the bond and lone pair 1st GVB-NO's $\{\phi_{gn}^{\text{GVB}}\}$ from the initial set of HF-MO's, $\{\phi_i^{\text{HF}}\}$. This is done by selecting the near-unity eigenvalues of the following matrix:

$$P_{ij}^{\text{HF}} = \delta_{ij} - \langle \phi_i^{\text{HF}} | \sum_{n=1}^{n_{\text{pair}}} |\phi_{gn}^{\text{GVB}}\rangle \langle \phi_{gn}^{\text{GVB}}| | \phi_j^{\text{HF}} \rangle \quad (3.11)$$

(λ3) Combine all occupied GVB orbitals in the following sequence: (i) all doubly occupied orbitals (core, uncorrelated lobe, and bonding orbitals), (ii) all the 1st GVB-NO's $\{\phi_{gi}^{\text{GVB}}\}$, (iii) any singly-occupied (high-spin) molecular orbitals, and (iv) all 2nd GVB-NO's $\{\phi_{ui}^{\text{GVB}}\}$.

(λ4) Schmidt orthogonalize the 2nd GVB-NO's to all doubly-occupied orbitals and redo the symmetric orthogonalization of the 2nd GVB-NO's to remove any overlap between the 2nd GVB-NO's introduced by the Schmidt orthogonalization.

In steps λ1 and λ4 it is important to use symmetrical orthogonalization in order to preserve symmetry between the GVB orbitals for equivalent bonds. The 1st GVB-NO's are symmetrically orthogonalized independently of the 2nd GVB-NO since the former have much larger occupation.

Summarizing, the sequence of steps and notation is as follows: (i) The atomic basis functions are $\{\chi_{\mu}\}$. (ii) The atomic HF-SCF orbitals are $\{\theta_n^A\}$, where f_n^A are the occupation numbers. Step α. (iii) The molecular orbitals of the HF-INIT wave function are $\{\phi_n\}$. This contains only the occupied MO's for the HF wave functions. Steps λ1, γ2, and γ3. (iv) Atomic projections of the HF-MO's are $\{\eta_n^A\}$. These are sorted into core, valence, and unoccupied orbitals according to occupations. Step γ4. (v) Bond pairs (if correlated): 1st and 2nd GVB-NO $\{\phi_{gi}^{\text{GVB}}\}$ and $\{\phi_{ui}^{\text{GVB}}\}$ are formed by combining the $\{\bar{\psi}_{ai}, \bar{\psi}_{bi}\}$ obtained by maximizing the overlap between the $\{\eta_n^A\}$ and $\{\eta_n^B\}$ valence orbitals. Steps δ1, δ2, δ3, and δ4. (vi) Lone pairs (if correlated): 1st GVB-NO $\{\phi_{gi}^{\text{lobe},A}\}$ selected from valence orbitals $\{\eta_n^{\text{val},A}\}$ after eliminating $\{\bar{\psi}_{gi}^{\text{GVB}}\}$ on center A; 2nd GVB-NO $\{\phi_{ui}^{\text{lobe},A}\}$ obtained from unoccupied orbitals $\{\eta_n^{\text{unocc},A}\}$ by maximizing overlap with gradient operator applied to $\{\phi_{gi}^{\text{lobe},A}\}$. Steps ε1 and ε2. (vii) Symmetric orthogonalization with highest occupations first (steps λ1, λ2, λ3, and λ4) to form GVB-INIT.

Computationally, the method described above is very efficient. The most time-consuming steps are the construction of the projection operator over the HF orbitals, the construction

TABLE 1

(a) Population Difference $n_{\text{occ}} - n_{\text{guess}}$ [See Eq 4.1] between Initial and Converged HF Wave Functions

molecule	no. of occupied orbitals	HF-INIT	GAUSSIAN90	geometry
water	5	0.149	0.066	STO-3G optimized
formaldehyde	8	0.167	0.084	STO-3G optimized
ethylene	8	0.141	0.077	ref 14
Si ₂ H ₆	17	0.297	0.249	STO-3G optimized
glycine ^a	20	0.455	0.224	STO-3G optimized
glutamine ^b	39	0.808	0.454	ref 15
TTF ^c	52	0.819	0.707	6-31G** optimized
porphine ^d	81	1.213	0.832	ref 16

(b) Initial Energies (hartrees) Produced by HF-INIT and G90-INIT

molecule	HF-INIT	GAUSSIAN90
water	-75.608	-75.761
formaldehyde	-113.420	-113.451
ethylene	-77.629	-77.621
Si ₂ H ₆	-580.892	-579.613
glycine	-281.625	-281.713
glutamine	-526.308	-526.330
TTF	-1818.091	-1815.625
porphine	-979.832	-978.393

^a See Figure 1a. ^b See Figure 1b. ^c See Figure 1c. ^d See Figure 1d.

and diagonalization of the matrix P^{HF} of (3.11), and the final Schmidt orthogonalization. All three steps scale as M^3 . Note that this method does not require a global localization of all the orbitals of the molecule^{11,12} (an iterative process) but only piecewise noniterative localization based on simple chemical bonding ideas.

4. Results and Discussion

The GVB-INIT method described above has been used as the starting point for a number of studies of various wave functions. In particular, a recent paper⁷ on GVB-DIIS uses it for a number of systems, including glycine correlated with up to 10 GVB pairs. In this paper we have concentrated on making comparisons between the GVB-INIT wave functions and the converged wave functions in order to assess the soundness of the approach.

4.1. Hartree-Fock SCF Calculations. In Table 1 we compare the efficacy of the HF initial guess method defined in section III.A (called HF-INIT hereafter) with that of the standard initial guess from¹³ GAUSSIAN90 (G90-INIT). All calculations used the 6-31G** basis sets with the geometries¹⁴⁻¹⁶ specified in Table 1a and Figure 1. As a measure of the efficacy, we evaluate the overlap population between the initial guess wave function and the final SCF wave function. The overlap population, n_{guess} , is defined as follows

$$n_{\text{guess}} = \sum_{ij}^{n_{\text{occ}}} \langle \phi_i^{\text{guess}} | \phi_j^{\text{HF-SCF}} \rangle^2 \quad (4.1)$$

where the $\{\phi_i^{\text{guess}}\}$ are occupied orbitals of the initial guess wave function and the $\{\phi_i^{\text{HF-SCF}}\}$ are occupied orbitals of the final HF-SCF wave function. Thus $0 \leq n_{\text{guess}} \leq n_{\text{occ}}$. The initial energies from HF-INIT and G90-INIT are given in Table 1b. For smaller molecules G90-INIT is better by up to 0.16 hartree (water) while for larger molecule HF-INIT is better by up to 2.4 hartrees (TTF).

The test cases shown in Table 1 include small and intermediate sized systems up to porphine (with 81 occupied orbitals and 430 basis functions). Both methods do quite well with overlaps between the initial guess orbitals and the final SCF orbitals always greater than 98%. The average error per orbital is 1.2%

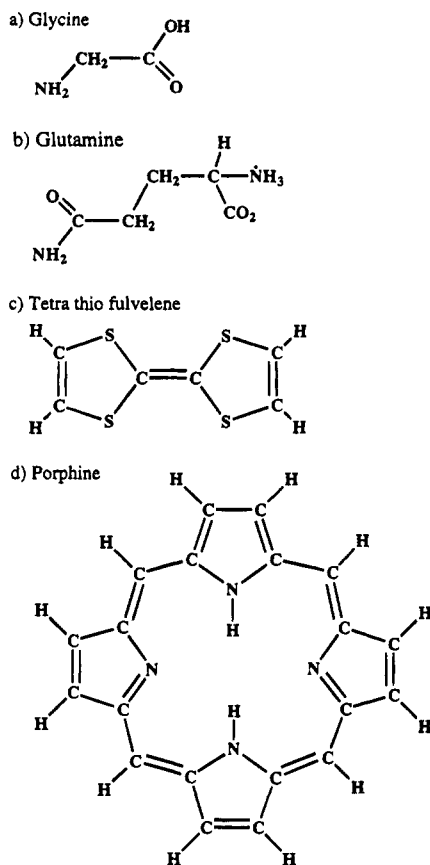


Figure 1.

for GAUSSIAN90, which is slightly better than the value of 1.7% for HF-INIT.

Given the simplicity of the HF-INIT method it is encouraging that the accuracy is almost as good as the semiempirical method used in GAUSSIAN90, particularly since HF-INIT involves no adjustable parameters while the semiempirical approach in G90-INIT includes some parametrization.

The success of HF-INIT validates the basic hypotheses that (i) the molecular density is not significantly different from the sum of the atomic densities and (ii) the HF-INIT orbitals can be selected out of the space spanned by the valence atomic orbitals by choosing the linear combination of valence atomic orbitals with maximal overlap.

It should be noted that the HF-INIT orbitals are different than the orbitals (HF-DENS) obtained by diagonalizing the sum of atomic densities. For HF-DENS the sum of atomic densities is given by

$$Q_{\text{average}} = \sum_{A=1}^{n_{\text{atoms}}} \sum_{n=1}^{n_{\text{occ}}(A)} |\theta_n^A\rangle f_n^A \langle \theta_n^A| \quad (4.2)$$

where $n_{\text{occ}}(A)$ is the number of occupied HF-AO on atom A. In this case associated initial guess orbitals would be the eigenvectors of the corresponding eigenvalue problem:

$$Pc = Sc\gamma \quad (4.3a)$$

with

$$P_{\mu\nu} = \langle \chi_\mu | Q_{\text{average}} | \chi_\nu \rangle \quad (4.3b)$$

and

$$S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle \quad (4.3c)$$

TABLE 2: Description of GVB-PP Wave Functions Used for Tests, Including the Number of Iterations to Convergence^a from GVB-INIT to an Energy Criterion of 5×10^{-5} hartrees

molecule	no. of GVB pairs	no. of iterations	pair distribution			pair correlated
			σ	π	lone	
water	2	7	2	0	0	all bond pairs
water	4	9	2	0	2	all bond and lone pairs
formaldehyde	2	8	1	1	0	CO double bond pairs
formaldehyde	4	7	3	1	0	all bond pairs
formaldehyde	6	9	3	1	2	all bond and lone pairs
ethylene	2	6	1	1	0	CC double bond pairs
ethylene	6	7	5	1	0	all bond pairs
glycine	5	9	4	1	0	all NC, CC, CC, CO bond pairs
glycine	10	9	9	1	0	all bond pairs
glycine	15	11	9	1	5	all bond and lone pairs
TTF	18	9	15	3	0	all bond pairs
TTF	26	11	15	3	8	all bond and lone pairs
porphine	22	14	11	11	0	all double bond pairs

^a Using the PS-GVB electronic structure program (see ref 17) with GVB-DIIS.

where the $\{\chi_\mu\}$ are the molecular basis functions. The occupied orbitals are chosen according to their eigenvalue which correspond to projection onto the space formed by Q_{average} . Those density selected orbitals, HF-DENS, are different from the HF-SCF orbitals by the same amount as the HF-INIT orbitals, i.e., an average orbital error of 1.7% when compared with the converged HF-SCF orbitals, but the average difference per orbital between the HF-INIT wave function and the wave function coming from the density selected orbitals is 1.3%.

4.2. GVB-SCF Calculations. The GVB-INIT approach has been applied to a variety of molecules and wave functions, as summarized in Table 2. Most systems converge in less than 12 iterations. We use two criteria to assess the soundness of the approach.

1. *Orbital Space:* (a) how well GVB-INIT separates the space of core orbitals from the space of 1st GVB-NO's, and (b) how well GVB-INIT separates the space of 2nd GVB-NO's from the space of unoccupied orbitals. We test these by comparing the initial guess orbitals to the final SCF orbitals. In this orbital space test, we consider only the space of 1st GVB-NO's (or 2nd GVB-NO's) and do *not* distinguish if the 1st GVB-NO's (or the 2nd GVB-NO's) are mixed together by some unitary transformation.

2. *Orbital Pairs:* how well each 1st GVB-NO (or 2nd GVB-NO) from GVB-INIT compares with the self-consistent GVB orbitals.

The *orbital space* efficacy for GVB-INIT (separating core orbitals from the 1st GVB-NO's and the 2nd GVB-NO's from the unoccupied orbitals) is given in Table 3 for all systems studied.

4.2.1. Use of GVB-INIT. The efficacy for each set of orbitals (core orbitals, 1st GVB-NO's, 2nd GVB-NO's) was defined as $n_{\text{guess}}/n_{\text{orbs}}$ where n_{orbs} is the number of orbitals in that set and n_{guess} is defined in eq 4.1. For orbital space separation of the core orbitals from the first GVB-NO's, the efficacies range from 77% to 98% for GVB-INIT. For the second GVB-NO orbital space, the GVB-INIT efficacies range from 82% to 96%. This indicates the advantage of obtaining the second GVB-NO's directly from the first GVB-NO's (as in the GVB-INIT).

More relevant for convergence is the orbital pair criterion, that is, how well each GVB-NO is predicted. To assess this, we evaluate the sum of populations over the 1st GVB-NO's or the 2nd GVB-NO's.

TABLE 3: Efficacies for GVB-INIT^a

molecule	no. of electrons	no. of GVB pairs	orbital space efficacies (fraction)			orbital pair efficacies	
			core orb	1st GVB-NO	2nd GVB-NO	1st GVB-NO	2nd GVB-NO
water	10	2	0.8632	0.7862	0.8432	1.547 (77.3%)	1.681 (84.0%)
water	10	4	0.9998	0.9629	0.8935	3.435 (85.9%)	3.349 (83.7%)
formaldehyde	16	2	0.9272	0.8085	0.8978	1.617 (80.8%)	1.795 (89.7%)
formaldehyde	16	4	0.9008	0.8831	0.9084	3.530 (88.2%)	3.633 (90.8%)
formaldehyde	16	6	0.9998	0.9737	0.9180	5.523 (92.0%)	5.269 (87.8%)
ethylene	16	2	0.9522	0.8980	0.9654	1.796 (89.8%)	1.930 (96.5%)
ethylene	16	6	0.9999	0.9766	0.9375	5.858 (97.6%)	5.624 (93.7%)
glycine	40	5	0.9118	0.7715	0.8906	3.853 (77.1%)	4.447 (88.9%)
glycine	40	10	0.8827	0.8696	0.9005	8.647 (86.5%)	8.975 (89.7%)
glycine	40	15	0.9998	0.9693	0.9124	13.542 (90.3%)	13.105 (87.4%)
TTF	104	18	0.9485	0.8954	0.9249	16.061 (89.5%)	16.623 (92.3%)
TTF	104	26	0.9996	0.9670	0.8208	23.891 (91.9%)	20.921 (80.5%)
porphine	162	22	0.9435	0.8667	0.9445	19.021 (86.5%)	20.737 (94.3%)

^a The orbital space efficacy is given as the fraction of the total. Thus, for porphine there are 162 electrons, of which 118 are doubly occupied and 44 are in GVB pairs. The 94.35% for the core orbitals implies that $n_{\text{occ}} = 55.67$ out of 59, while the efficacy for GVB-NO's is $n_{\text{occ}} = 0.8667 \times 22 = 18.07$ for the 1st NO and $n_{\text{occ}} = 0.9445 \times 22 = 20.78$ for the 2nd NO. The orbital pair efficacy is the sum of populations between initial and converged wave functions (in n_{occ} units). Thus, for porphine the average projection is 86.46% for first NO and 94.26% for second NO.

TABLE 4: Same as Table 3 Except Using Boys Localization Instead of GVB-INIT

molecule	no. of GVB pairs	orbital space efficacies			orbital pair efficacies		GVB-Boys 1st GVB-NO
		core orb	1st GVB-NO	2nd GVB-NO	1st GVB-NO	2nd GVB-NO	
water	2	0.9201	0.8813	0.4607	1.756	0.919	1.756
water	4	0.9998	0.9629	0.4904	3.724	1.419	3.724
formaldehyde	2	0.9570	0.8992	0.7039	0.899	0.704	1.798
formaldehyde	4	0.9951	0.9379	0.6562	2.836	1.918	3.744
formaldehyde	6	0.9998	0.9737	0.5983	4.807	2.501	5.761
ethylene	2	0.9732	0.9630	0.7508	0.963	0.782	1.926
ethylene	6	0.9999	0.9766	0.6004	4.867	2.883	5.851
glycine	5	0.9244	0.8154	0.6622	3.168	3.252	4.470
glycine	10	0.8866	0.8750	0.6014	7.816	5.944	9.112
glycine	15	0.9998	0.9693	0.5549	11.662	7.139	14.053
TTF	18	0.9594	0.9162	0.7077	13.501	10.447	16.434
TTF	26	0.9996	0.9670	0.6399	17.749	13.897	24.095
porphine	22	0.9647	0.9242	0.6946	10.839	09.040	19.626

$$n_{\text{sum pop.}} = \sum_i^{n_{\text{pair}}} \langle \phi_i^{\text{guess}} | \phi_i^{\text{GVB-SCF}} \rangle^2 \quad (4.4)$$

where the $\{\phi_i^{\text{guess}}\}$ are GVB-NO's of the initial guess wave function and the $\{\phi_i^{\text{GVB-SCF}}\}$ are GVB-NO's of the final HF-SCF wave function.

The results in Table 3 show that GVB-INIT gives excellent orbital pair results for both the 1st GVB-NO's and 2nd GVB-NO's. The average error per GVB orbitals for the 1st NO ranges from a high of 23% for water to a low of 2% for ethylene. As the number of GVB pairs is increased for a given molecule, the average error always decreases. This indicates that GVB-INIT performs best when the entire valence space is described with GVB pairs (including lone pairs).

For the 2nd NO, the range of the average error (4–16%) is similar to the 1st NO, but the dependence of the number of GVB pair is different. The presence of the lone pair GVB orbitals always increases the average error. This means that the second GVB-NO's for lone pairs are less accurately predicted. This is indicated by evaluating the average error per orbital for each type of GVB pair.

For example, glycine with 15 GVB pairs leads to an average error per orbital of 8.1% for bond pair 1st GVB-NO's, 13.0% for lone pair 1st GVB-NO's, 11.7% for bond pair 2nd GVB-NO's, and 48.2% for lone pair 2nd GVB-NO's. Unlike the bond pair 2nd GVB-NO's, the lone pair 2nd GVB-NO's are not constructed directly from the first GVB-NO's. In the bond pair case, the first GVB-NO and second GVB-NO are obtained from the set of two localized atomic orbitals centered on the nearest-neighbor atoms. The first GVB-NO corresponds to the

bonding orbital and the second GVB-NO to the antibonding orbital. For lone pair, we have only one atomic-centered orbital involved from which we can only construct the first GVB-NO. The second GVB-NO must be obtained by some extra criterion. The guiding principle is that the first GVB-NO is dominated by s character and the second GVB-NO is p-like and pointing away from the nucleus in the direction of the centroid for the first GVB-NO. Two simple operators, the dipole moment and the gradient, can be used to transform an s-like orbital to a p-like orbital. If the s-like orbital is composed of a single s Gaussian function, then the two operators will generate the same p-like orbital. If the s-like function contains many different Gaussian functions, the p-like orbital generated by the gradient operator emphasizes the tight Gaussian functions more than the dipole moment operator. The effect of these operators is limited to the small region of space around the 1st GVB-NO centroid by multiplying the operator by $1/|r - R_{\text{lobe}}|^n$. We find that the best results are obtained with the gradient operator and $n = 3$; see Table 5.

Above we compared directly the initial guess wave function to the final SCF wavefunction. A second important criterion is the rate of convergence obtained with the initial guess wave functions. In Table 6 the GVB-INIT guess is used for three different convergence methods in the case of formaldehyde with four GVB pairs. The convergence criterion requires the sum of square of change in the orbital coefficients to be less than 10^{-9} :

$$\sum_{\mu=1}^{n_{\text{basis}}} \sum_{i=1}^{n_{\text{occ}}} (\Delta C_{\mu i})^2 \leq 10^{-9} \quad (4.5)$$

TABLE 5: Efficacies for Lone Pair Second Natural Orbital^a

molecule	no. of GVB pairs	operator	power of $1/r^n$			
			$n = 0$	$n = 1$	$n = 2$	$n = 3$
water	4	dipole	1.857	1.762	1.627	1.760
		gradient	1.809	2.007	2.826	3.334
formaldehyde	6	dipole	3.856	3.840	3.730	3.746
		gradient	3.766	3.882	3.949	4.147
glycine	15	dipole	9.383	9.292	9.010	9.401
		gradient	9.221	9.630	10.722	11.477
TTF	26	dipole	17.132	16.609	17.492	16.549
		gradient	16.972	18.244	20.452	20.724

^a This is defined as the sum of populations between initial and converged second GVB natural orbitals. n is the exponent in eq 3.9.

TABLE 6: Number of Iterations for Converging the GVB Wave Functions of Formaldehyde

trial function	no. of GVB pairs	convergence method ^a		
		GVB-DIIS	GVB2P5	GAUSSIAN90
GVB-INIT	2	15	28	34
Boys	2	21 ^b	83	95
GVB-INIT	4	15	58	69
Boys	4	21 ^b	>300 ^c	>300 ^c

^a Convergence criterion of 10^{-9} . ^b Converged to nearest root. ^c Unconverged after 300 iterations.

In all cases the GVB-INIT initial guess leads to excellent convergence, with the fastest convergence obtained using the GVB-DIIS method.

4.2.2. Boys Localization and GVB-Boys. As an alternative to GVB-INIT, we have considered Boys localization¹¹ of the occupied orbitals from HF-INIT to form core orbitals and 1st NO's and Boys localization of the unoccupied orbitals for HF-INIT to form 2nd NO's. [Boys localization maximizes the distances between the centroids of the various orbitals.]

In order to have as close as possible a comparison between GVB-INIT and the Boys method, the Boys orbitals were obtained as follows: (1) A Boys localization was performed on the HF-INIT guess orbitals with the unoccupied orbitals localized independently from the occupied orbitals and the inner core orbitals (i.e., 1s orbital for C, N, O atoms and 1s, 2s, 2p for S atoms) kept frozen from HF-INIT. (2) The separation of the core orbitals from the 1st GVB-NO's and the 2nd GVB-NO's from the unoccupied orbitals was done by selecting the Boys orbitals having the maximum overlap with the SCF orbitals.

For Boys localization one needs some form of analysis of the location and composition of each Boys orbitals in order to properly assign which are core orbitals and which are lone pairs or 1st GVB-NO's. GVB-INIT automatically does this assignment. For the tests we did this analysis for Boys by hand. For all systems studied here the assignment for the Boys localized

orbitals is easy since the electron distributions are consistent with simple valence bond structures. Assigning the 2nd GVB-NO's in terms of Boys localized orbital becomes very difficult since the 2nd GVB-NO's must be localized in the same region as the corresponding 1st GVB-NO.

Table 4 shows that the Boys method is slightly better than GVB-INIT for orbital space separation of the core orbitals from the first GVB-NO's. The Boys efficacies vary from 81% to 98% for the first GVB-NO's, whereas the efficacies vary from 77% to 98% for GVB-INIT. The advantage of the Boys method is largest when only some valence orbitals are correlated. When all possible bond and lone pair orbitals are correlated, then the core space is reduced to the inner core orbitals, and the Boys and GVB-INIT methods become equivalent for the core orbitals and the first GVB-NO's. The reason for the slight disadvantage with GVB-INIT is that the first GVB-NO's are formed from projections without any mechanism to deliberately make the orbital localized. Of course, GVB-INIT does automatic assignment of the orbitals whereas this was done manually for Boys.

For the second GVB-NO orbital space, the GVB-INIT method is markedly superior to the Boys method. Here the Boys efficacies range from 46% to 75%, whereas the GVB-INIT efficacies range from 82% to 96%. This indicates that the second GVB-NO's should be obtained directly from the first GVB-NO's as in the GVB-INIT method. In the Boys method, the localization principle does not necessarily lead to 2nd GVB-NO's that correspond well with the first GVB-NO's.

More relevant for convergence is the orbital pair criteria, eq 4.4. For both the first GVB-NO's and second GVB-NO's, the GVB-INIT method gives much better orbital pair results than the Boys method. [The single exception is for the first GVB NO's of water.] One problem with Boys localization is that double bonds and triple bonds become "banana"-like bonds with two or three equivalent bonds, whereas the optimum GVB-NO's generally have the character of σ and π orbitals. A similar situation occurs for the lone pair orbitals on oxygen and sulfur atoms where Boys localization sometimes selects the opposite set of lone pair orbitals.

That Boys localization for water is better for the orbital space of 1st GVB-NO's, which indicates that choosing the most localized orbitals within the 1st GVB-NO space might lead to better trial GVB orbitals. This is referred to as GVB-BOYS. To test this hypothesis, we performed a series of calculations where the Boys localization is followed by simple linear combination of the 1st GVB-NO's in order to reproduce the GVB-INIT orbitals as closely as possible. This consistently gives the best results (see Table 4 under the GVB-Boys heading).

In all cases the GVB-INIT initial guess leads to much better convergence, eq 4.5, than Boys localization (see Table 6).

TABLE 7: For Computational Simplicity the $1/r^n$ Terms Used in Obtaining the Lobe Second GVB-NO's in Step ϵ_2 Are Represented by an Expansion over Eight Unnormalized Gaussian Functions,

$$\frac{1}{r^n} = \sum_{i=1}^8 e^{-\alpha_i r^2} C_i$$

α	$1/r$	$1/r^2$	$1/r^3$	$1/r^4$
0.1	1.306 843 88	0.265 230 22	-2.069 809 60	-102.104 369
0.3	-1.784 365 54	0.288 542 06	15.751 962 06	744.469 961
0.9	3.556 331 02	-0.038 622 20	-74.135 563 64	-3589.785 994
2.7	-2.773 267 67	8.826 3 852 6	327.117 677 48	15351.157 764
8.1	6.266 879 83	-16.84 2 755 57	-1284.246 761 30	-62488.952 048
24.3	-1.276 731 07	129.72 6 918 39	5275.259 827 38	244239.329 549
72.9	5.887 516 94	-287.533 946 28	-17199.924 487 36	-839793.781 169
218.7	20.283 583 83	1201.461 779 54	48908.533 565 32	2085688.141 678

5. Conclusions

The GVB-INIT method of obtaining initial guesses for GVB wave functions is inexpensive (using only piecewise localization) and yet provides trial guesses very close to the final orbitals with the correct localization properties for GVB wave functions.

GVB-INIT together with GVB-DIIS provides a robust automatic procedure to obtain wave functions describing bond distortion and dissociation processes for large molecules. These highly correlated wave functions can also be used to calculate molecular properties such as charges, dipole moments, and solvation energies.

The GVB-INIT automatic scheme for GVB trial wave functions and the GVB-DIIS convergence scheme are implemented in the PS-GVB electronic structure program,¹⁷ allowing automatic GVB calculations on very large systems.

Acknowledgment. This research was partially funded by the National Science Foundation (Che 91-100284 and ASC 92-17368). The facilities at the MSC/BI are also supported by grants from DOE-AICD, Allied-Signal Corp., Asahi Chemical, Asahi Glass, BP America, Chevron Petroleum Technology Co., B.F. Goodrich, Teijin Ltd., Hughes Research Labs., Xerox Corp., and Beckman Institute. Some of these calculations were carried out on the NSF Pittsburgh Supercomputing Center and some on the JPL Cray YMP.

References and Notes

(1) Ladner, R. C.; Goddard, W. A., III *J. Chem. Phys.* **1969**, *51*, 1073. Goddard, W. A., III; Dunning, T. H., Jr.; Hunt, W. J.; Hay, P. J. *Acc. Chem. Res.* **1973**, *6*, 368. Goddard, W. A., III; Harding, L. B. *Annu. Rev. Phys.*

Chem. **1978**, *29*, 363. Goddard, W. A., III *Science* **1985**, *227*, 917.

(2) Hunt, W. J.; Hay, P. J.; Goddard, W. A., III *J. Chem. Phys.* **1972**, *57*, 738. Hay, P. J.; Hunt, W. J.; Goddard, W. A., III *J. Am. Chem. Soc.* **1972**, *94*, 8293.

(3) Bobrowicz, F. W.; Goddard, W. A., III In *Modern Theoretical Chemistry: Methods of Electronic Structure Theory*; Schaefer, H. F., III Ed.; Plenum: New York, 1977; Vol. 3, p 79.

(4) Yaffe, L. G.; Goddard, W. A., III *J. Chem. Phys.* **1977**, *67*, 1777.

(5) Hunt, W. J.; Goddard, W. A., III; Dunning, T. H., Jr. *Chem. Phys. Lett.* **1970**, *6*, 147. See also: *Chem. Phys. Lett.* **1969**, *4*, 231 and *Chem. Phys. Lett.* **1969**, *3*, 606.

(6) Carter, E. A.; Goddard, W. A., III *J. Phys. Chem.* **1988**, *92*, 2109. Shin, S. K.; Goddard, W. A., III *J. Chem. Phys.* **1990**, *93*, 4986. Bair, R. A. PhD Thesis, Caltech, Chemistry, 1981.

(7) Muller, R. P.; Langlois, J.-M.; Ringnalda, M. N.; Friesner, R. A.; Goddard, W. A., III *J. Chem. Phys.* **1994**, *100*, 1226.

(8) BIOGRAF/POLYGRAF from Molecular Simulation Inc., Burlington, MA.

(9) Polack, R. *Int. J. Quantum Chem.* **1970**, *4*, 271.

(10) Amos, A. T.; Hall, G. G. *Proc. R. Soc. London* **1961**, *A263*, 483. See also: Goddard, W. A., III *J. Chem. Phys.* **1968**, *48*, 450.

(11) Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.

(12) Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457.

(13) GAUSSIAN 90; Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. P.; Schlegel, H. B.; Raghavachari, K.; Robb, M. A.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1990.

(14) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(15) Koetzle, T. F.; Frey, M. N.; Lehmann, M. S.; Hamilton, W. C. *Acta Crystallogr.* **1973**, *B29*, 2571.

(16) Chen, B. M. L.; Tulinsky, A. *J. Am. Chem. Soc.* **1972**, *94*, 4144.

(17) Ringnalda, M. N.; Langlois, J.-M.; Greeley, B. H.; Russo, T. V.; Muller, R. P.; Marten, B.; Won, Y.; Donnelly, R. E., Jr.; Pollard, W. T.; Miller, G. H.; Goddard, W. A., III; Friesner, R. A. PS-GVB v1.0, Schrödinger, Inc., Pasadena, CA, 1994.