JCTC Journal of Chemical Theory and Computation

Tight-Binding Configuration Interaction (TBCI): A Noniterative Approach to Incorporating Electrostatics into Tight Binding

Mark A. Iron,[†,§] Andreas Heyden,[†,||] Grażyna Staszewska,[†,‡] and Donald G. Truhlar*[,†]

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431, and Institute of Physics, Nicolaus Copernicus University, ul. Grudziądzka 5, 87-100 Toruñ, Poland

Received December 18, 2007

Abstract: We present a new electronic structure approximation called Tight Binding Configuration Interaction. It uses a tight-binding Hamiltonian to obtain orbitals that are used in a configuration interaction calculation that includes explicit charge interactions. This new method is better capable of predicting energies, ionization potentials, and fragmentation charges than the Wolfsberg−Helmholz Tight-Binding and Many-Body Tight-Binding models reported earlier (Staszewska, G.; Staszewski, P.; Schultz, N. E.; Truhlar, D. *Phys. Rev. B* **2005**, *71*, 045423). The method is illustrated for clusters and nanoparticles containing aluminum.

1. Introduction

Tight-binding (TB) theory[1–3] (also called extended Hückel theory) is well-suited to modeling the electronic structure and potential energy surfaces of materials and nanoparticles in that it incorporates quantum mechanical effects with a minimum of computational expense. It is, however, a very approximate theory and can only provide useful accuracy if the matrix elements are empirically parametrized. It is well-known in a variety of contexts that empirically parametrized theories work best when the functional or operational form of the theory incorporates the dominant physical interactions. Tight-binding theory has the functional form of a one-electron eigenvalue problem; this class of functional forms can be derived[4,5] from Kohn−Sham density functional theory[6] by assuming that the densities of the atoms of a given atomic number are all close to a reference density for that atomic number. This assumption breaks down for many problems, in particular those involving bond breaking and formation, variable coordination numbers, multiple oxidation states, and polar or ionic bonds. Thus, the question arises of whether there might be other functional or operational forms that are only slightly more computationally expensive but better incorporate the atomic interactions.

Kohn−Sham theory itself has the form of a pseudoeigenvalue problem in that the Kohn−Sham matrix to be diagonalized depends on the occupied orbitals, which in turn depend on the field produced by the orbital eigenvectors. This kind of problem can be solved by an iterative process, leading to the familiar self-consistent field problem.[7–9] Various versions of tight binding with occupancy-dependent terms in the Hamiltonian (e.g., iterative extended Hückel theory,[10] the Anderson−Newns model,[11] the Hubbard model,[12] the Grimley−Pisani model,[13] self-consistent charge density functional tight binding (SCC-DFTB),[14,15] and the generalized Hubbard model[16–18]) have been proposed, and some of these models have enjoyed considerable success. Nonetheless, occupancy-dependent Hamiltonians require an iterative method for their solution, and an iterative process raises the cost and complexity of the method, especially if one is computing the gradients and/or Hessians[19] for dynamics. Furthermore, there is the added inconvenience that one must develop computational strategies for dealing with unstable cases where the iterative process diverges, stalls, or oscillates. The present article is devoted to trying to solve some of the same issues that motivated iterative tight-binding

* Corresponding author e-mail: truhlar@umn.edu.
† University of Minnesota.
‡ Nicolaus Copernicus University.
§ Current address: Computational Chemistry Unit, Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel 76100.
|| Current address: Department of Chemical Engineering, University of South Carolina, Columbia, SC 29208.

10.1021/ct700343t CCC: $40.75
Published on Web 04/25/2008

schemes, especially finding the optimum charge distribution, but with noniterative methods.

In general, tight binding has almost always been considered a way to approximate a single-configuration wave function, such as the ones in Hartree−Fock,[20–24] Hartree−Fock−Slater,[25] Kohn−Sham,[6] or dynamical Hartree−Fock[26] theories. One can make Kohn−Sham theory more accurate by improving the one-electron Hamiltonian (by using better density functionals), and such an approach is analogous in some respects to better parametrization of the tight-binding Hamiltonian, just as the iterative refinement of the Kohn−Sham Hamiltonian is the direct analogue of the iterative tight-binding methods discussed above. However, the more traditional approach in quantum chemistry has been configuration interaction (CI), either by variational theory,[27] by perturbation (or many-body) theory,[28,29] or by coupled cluster theory;[29] in these theories, the wave function is a linear combination of configuration state functions (CSFs), each built from Hartree−Fock orbitals (as in the so-called post-Hartree−Fock correlation methods[30]), from non-self-consistent orbitals (as in valence bond theory[31,32]), or from Kohn−Sham orbitals.[33] This suggested to us that one might improve tight-binding theory by using the tight-binding orbitals to perform an approximate configuration interaction calculation. We propose and test such an approach in the present article.

The main reason why one usually prefers using self-consistent-field orbitals for configuration interaction is that it makes the matrix elements connecting the self-consistent configurations to single excitations all equal to zero.[34–38] Since we are using tight-binding orbitals, which are not self-consistent, we will include single excitations. In fact, the single excitations are the ones most responsible for charge redistribution. Since our main goal is to include charge redistribution effects without resorting to iterative methods, the single excitations are the key element of the new method.

A second motivation for the development of the new method, called tight-binding configuration interaction (TBCI), is that single-configuration methods like tight-binding often lead to incorrect electron distributions upon bond breaking.[39,40] The new method is designed to improve this so that the new tight-binding method is more reasonable in describing bond-breaking dynamical processes.

Aluminum clusters and nanoparticles were chosen as the initial application for evaluating the new method. Recently, we have devoted considerable effort to studying these systems.[39,41–48] Analytical potential energy functions (PEFs) were developed that accurately predict cluster energies.[42,45] A number of TB and many-body TB (MBTB) models were also parametrized.[39,44,46] In addition, when potentials based on these methods were used, the properties of aluminum were explored by molecular dynamics and Monte Carlo calculations.[49,50] There are a number of reasons for the interest in this metal. It is of technological and industrial value as a lightweight, rust-resistant metal, and it is an ingredient for high-energy fuels (e.g., as a component of solid rocket propellant) and potentially as a hydrogen-storage medium.[51] The use of aluminum nanoparticles, instead of bulk metal, is expected to enhance the latter two properties.[51–54] Metal nanoparticles provide a challenging test for approximate electronic structure theories because it has been demonstrated that the properties of nanometer-sized particles can depend dramatically on size, occasionally in an unpredictable manner, as compared to those of bulk metal.[55–65]

2. Theory

First, we present the theory behind the new electronic structure method for a system involving only a single element. Afterward, the extension of the method to heteroatomic systems is described.

The CI wave function is written as

$$\Psi = \sum_j C_j \Phi_j \tag{1}$$

where C_j is a CI coefficient and Φ_j is a CSF. We follow the common practice in TB of using the frozen-core approximation and thus consider only the valence electrons. Each CSF is a Hartree product:

$$\Phi_j = \varphi_1^{o_1^{(j)}} \varphi_2^{o_2^{(j)}} \cdots \varphi_M^{o_M^{(j)}} \tag{2}$$

where φ_k is an orbital, $o_k^{(j)}$ is the occupancy number (0, 1, or 2) of orbital k in CSF j, and M is the number of orbitals. The sum of the occupancy numbers in any CSF equals the number of valence electrons:

$$N_{\mathrm{val}} = \sum_k o_k^{(j)} \tag{3}$$

The one-electron density matrix of the CI wave function is approximated as

$$P_{mn} = \sum_j |C_j|^2 P_{mn}^{(j)} \tag{4}$$

where $P_{mn}^{(j)}$ is the density matrix of CSF j. The total CI energy E is the lowest eigenvalue of the CI matrix. Rather than construct and diagonalize the full CI matrix, which is far too large even if we knew or had approximations for the off-diagonal elements, the lowest eigenvalue is approximated using a weighted version of the degeneracy-corrected perturbation theory[66] energy:

$$E = \sum_j |C_j|^2 \Gamma_j \tag{5}$$

$$\Gamma_j = E_j + \frac{1}{2}\sum_i \left[(E_i - E_j) - \sqrt{(E_i - E_j)^2 + 4\chi_{ij}^2}\right] \tag{6}$$

where E_j is the energy of CSF j, and χ_{ij} is the *coupling element* between CSFs i and j, when $i \neq j$, and is a constant called the *diagonal coupling constant* when $i = j$ (the $i = j$ term is included in the sum to ensure size extensivity even in the presence of degeneracies). This form (eqs 5 and 6) was chosen for several reasons: it leads to a size-consistent method, it gives the correct answer for a 2×2 CI matrix with known off-diagonal elements, the ground-state energy is below the lowest CSF energy, and the ground-state energy approaches the energy of the lowest CSF in the limit of well-separated CSF energies.

In lieu of eq 5 for determining the TBCI energy, one could consider using

$$E = \sum_j |C_j|^2 E_j \tag{7}$$

However, this equation will lead to total energies that are higher than that of the lowest-energy CSF, whereas eq 5 will always give total energies that, as one finds in a variational calculation, are lower. Another possible formula for the total energy is based on the free energy formula:

$$E = -\Delta \ln\left[\sum_j \exp\left(-\frac{E_j}{\Delta}\right)\right] \tag{8}$$

Equation 8 gives overall energies that are lower than that of the lowest-energy CSF, but it is not size-extensive. We found that the combination of eqs 5 and 24 (*vide infra*) is the most satisfactory formula for the configuration interaction energy, and we did not use eqs 7 or 8.

The remaining tasks are to develop approximations for $|C_j|^2$, χ_{ij}, and E_j.

The molecular orbitals (MOs) are expanded in a minimal basis set $\{\eta_a\}$ of atomic orbitals:

$$\varphi_k = \sum_a c_{ak}\eta_a \tag{9}$$

The atomic orbitals are assumed to be orthonormal. The orbital coefficients c_{ak} are obtained by solving the tight-binding orbital eigenvalue equations, given in matrix form by

$$\mathbf{Hc} = \boldsymbol{\varepsilon}\mathbf{c} \tag{10}$$

where $\mathbf{H}$ is the $M \times M$ tight-binding Hamiltonian matrix (i.e., the one-electron Hamiltonian) with elements H_{ab}, $\mathbf{c}$ is an $M \times M$ matrix in which each column is an eigenvector, and $\boldsymbol{\varepsilon}$ is a diagonal matrix of orbital eigenvalues. The elements of $\mathbf{c}$ are the orbital coefficients in eq 9. On the basis of previous work,[39,44,46] we will assume zero differential overlap (ZDO)[67,68] so that the density matrix of CSF j in the atomic orbital representation becomes

$$P_{ab}^{(j)} = \delta_{ab}\sum_k^{MO} o_k^{(j)} |c_{ak}|^2 \tag{11}$$

The ZDO approximation is the reason that the atomic-orbital overlap matrix is neglected in eq 10. In a previous work on TB, it was found that neglecting the overlap gave a more balanced set of approximations.[46]

The energy of a given CSF is given by

$$E_j = E_{\mathrm{val}}^{(j)} + E_{\mathrm{CB}}^{(j)} + V_{\mathrm{rep}} + E^{(0)} \tag{12}$$

The valence energy, $E_{\mathrm{val}}^{(j)}$, is given by

$$E_{\mathrm{val}}^{(j)} = \sum_k o_k^{(j)}\varepsilon_k \tag{13}$$

In some versions[39,46,69] of TB, the valence energy contains an additional term, $\delta_{o_k,2}\mu_{\mathrm{penalty}}$, which introduces a penalty (u_{penalty}) for two electrons occupying the same orbital. This allows for ground states that are not singlets or doublets and for homolytic bond dissociations of singlet molecules. One motivation behind developing the TBCI model is to find a more satisfactory means of allowing for these possibilities.

In eq 12, V_{rep}, is the core−core repulsion, which is assumed to be a function of geometry but not of the CSF (i.e., index j). It is modeled in the present work by a pairwise-additive potential:

$$V_{\mathrm{rep}} = \kappa\sum_A \sum_{B>A} \exp(-\tau R_{AB}) \tag{14}$$

where A and B are atom labels, and κ and τ are empirical parameters. This form is similar to that used in our previous TB work[39] where we also included a prefactor $(R_{AB})^{-u_{AB}}$ with u_{AB} being a positive constant parameter; this prefactor was omitted here since the optimized u_{AB} tended to be small, and the prefactor was found in preliminary work to have an insignificant effect. $E^{(0)}$ is a constant, independent of both geometry and CSF, that determines the zero of energy. The charge balance (CB) energy, $E_{\mathrm{CB}}^{(j)}$, is primarily responsible for describing the electrostatics in the system. In order to determine E_j, what remains is to define $\mathbf{H}$, $E_{\mathrm{CB}}^{(j)}$, and $E^{(0)}$.

The Hamiltonian $\mathbf{H}$ is modeled by the Wolfsberg−Helmholz approximation.[39,70] The diagonal elements are

$$H_{aa} = -U_Z^l \tag{15}$$

where U_Z^l is a parameter, different for each type of orbital. Note that these matrix elements depend on the atomic number Z and the orbital angular momentum quantum number l. In our previously published TB models, U_Z^l were taken as valence-state ionization potentials (VSIP or I_Z^l).[39] In TBCI, because of the CI treatment and CB term, one cannot make the same assumption, and the equations used for assigning U_Z^l are given later in this section.

The off-diagonal elements, called either transfer or hopping integrals, are[39,70]

$$H_{ab} = K_{l_a l_b m}\frac{H_{aa} + H_{bb}}{2}S_{ab} \tag{16}$$

where S_{ab} is the atomic-orbital overlap integral and $K_{l_i l_j m}$ is an empirical constant that depends on the orbital angular momenta l_i and angular momentum projection quantum number m on the axis connecting atoms A and B. This approximation was initially proposed by Mulliken[71] and has been used by Hoffmann[2,72−75] and Wolfsberg and Helmholz[70] and in the TB and MBTB models for aluminum.[39] We take each basis function η_a to be a single Slater-type orbital.[76] Then, the overlap matrix elements can be readily calculated. For example, when the orbital exponents on the two centers are equal, one can use the expressions determined by Jones[77]

$$S_{ab} = P_6(\zeta R_{AB})\exp(-\zeta R_{AB}) \tag{17}$$

where $P_6(x)$ is a sixth-order polynomial, A and B label the atoms on which orbitals a and b are centered, R_{AB} is the interatomic distance, and ζ is the Slater-orbital exponential parameter. This expression is sufficient for Al because fortuitously the s- and p-exponential parameters, as determined in the minimal basis set Hartree−Fock calculations by Clementi and Raimondi,[78] happen to be nearly the same; for the s-p overlap, the average of the two exponential parameters is used, in a manner similar to that of our previous work.[39] Jones has also provided equations for the overlap

for the more general case where the two exponential parameters (ζ_a and ζ_b) are different.[79]

As in previous work on TB,[39] various models can be formulated by using different sets of $K_{l_i l_j m}$ and $\zeta_{l_i l_j m}$. For aluminum, there are four relevant combinations of l_i, l_j, and m: $ss\sigma$, $sp\sigma$ (equivalent to $ps\sigma$), $pp\sigma$, and $pp\pi$. One could optimize either a single $K_{l_i l_j m}$ parameter for all four combinations (the Wolfsberg−Helmholz or WH model) or separate $K_{l_i l_j m}$ parameters for each permutation (the extended Wolfsberg−Helmholz or EWH model). In the EWH model, we use Clementi and Raimondi's ζ_a exponential parameters (except that for Al, we average the *s*- and *p*-exponential parameters in the $sp\sigma$ case in order to use the simpler expression for the overlap in eq 17, *vide supra*). In the optimized Wolfsberg−Helmholz or OWH model, the four Slater exponential parameters and four separate $K_{l_i l_j m}$ parameters are optimized.[39]

In order to calculate the charge balance energy, one must first calculate the partial atomic charges for each CSF. These are determined using a Mulliken−Coulson population analysis,[80–82] which, using eq 11, yields

$$q_A^{(j)} = Z_A^* - \sum_{a \in A} P_{aa}^{(j)} \tag{18}$$

where Z_A^* is the shielded nuclear charge of atom A (i.e., the number of valence electrons in the neutral atom, which is three for aluminum). The notation "$a \in A$" means that the sum in eq 18 is over all orbitals a on center A.

The charge balance energy is written as the sum of on-site and intersite interactions:

$$E_{\mathrm{CB}}^{(j)} = \sum_A \sum_{B \geq A} \gamma(R_{AB})\, q_A^{(j)} q_B^{(j)} \tag{19}$$

where $\gamma(R_{AB})$ is a Coulomb integral (*vide infra*) and $R_{AA} \equiv 0$. The diagonal terms in eq 19 are called on-site interactions. These interactions are the key element in Hubbard[12,16–18] and DFT+U[83,84] theories. The present theory also includes the intersite interactions. If the latter are neglected, the method is called TB+U, while the full theory is termed TBCI.

A key difference from SCC-DFTB[14,15] is that the electrostatic terms are added to the energy of each configuration after determining the orbitals, rather than iteratively while determining the orbitals. Their effects are included by the charge balance term of eq 19 and the configuration interaction approximation of eq 5. *Thus, TBCI is a noniterative method.* The slowest step in either SCC-DFTB or TBCI is the diagonalization of the one-electron Hamiltonian matrix in eq 10, which approximately scales as N^3, where N is the number of atoms in the system. Since SCC-DFTB repeats this step in each iterative cycle, whereas TBCI only does this step once, TBCI is significantly faster. Furthermore, the absence of an iterative step makes the analytical gradients particularly stable and precise. Finally, the neglect of the overlap integrals is a further feature that makes TBCI faster.

The on-site interactions $\gamma(R_{AA})$ could be approximated by the Pariser formula:[85]

$$\gamma(R_{AA}) \equiv \gamma_{AA} = \mathrm{IP}_A - \mathrm{EA}_A \tag{20}$$

where IP_A and EA_A are, respectively, the ionization potential and electron affinity of atom A; alternatively, they can be taken as empirical parameters. In the current implementation, the second option was chosen. It is convenient to use atomic units and write

$$\gamma_{AA} = \frac{e^2}{\alpha_A} \tag{21}$$

where e is the charge of an electron (unity in atomic units), and α_A is a parameter with units of length that will be called the *atomic Coulomb radius*.

The intersite charge interaction terms can be approximated in various ways. For example, one possible approximation for the Coulomb integral is based on a dielectric screening model[86,87] yielding, in atomic units

$$\gamma(R_{AB}) = e^2 \left[\alpha_A \alpha_B \exp\left(\frac{-R_{AB}^2}{d \alpha_A \alpha_B} \right) + R_{AB}^2 \right]^{-1/2} \tag{22}$$

where d is a parameter (originally 4[86]). Note that for $A = B$, this reduces to eq 21. If $d = \infty$ and the geometric mean $\alpha_A \alpha_B$ is replaced by an arithmetic mean, this becomes the Ohno−Klopman formula that is widely used in semiempirical molecular orbital theory.[88–90]

Omitting the configuration interaction step, that is, using the lowest E_j rather than using eq 5, corresponds to assuming that the eigenvectors of the CI matrix are the configurations Φ_j and that the ground-state wave function is the CSF with the lowest energy. This procedure leads to a nondifferentiable potential energy surface whenever the identity of the lowest-energy CSF switches. The use of eq 5 eliminates this problem.

The total number of configurations of the form of eq 2 is factorially large and thus unmanageable. Therefore, we limit the sums in eq 5 to a subset of configurations. In the present work, we consider two options. In both cases, the reference CSF is the "aufbau" state where all electrons are paired up in the MOs with the lowest $E_{\mathrm{val}}^{(j)}$; systems with an odd number of electrons will have one MO with a single electron. The first set of excitations, whose use yields a method called TBCI-S, involves all single excitations from the occupied orbitals ($o_k^{(i)} = 1$ or 2) to the unoccupied or singly occupied orbitals ($o_k^{(j)} = 0$ or 1). In the second option, denoted TBCI-SPD, in addition to the single excitations, all paired double (PD) excitations are also included, where a paired double excitation is defined as an excitation of two paired electrons (i.e., electrons in the same orbital) from one doubly occupied orbital ($o_k^{(i)} = 2$) to the same unoccupied orbital ($o_k^{(j)} = 0$).

Finally, we consider $E^{(0)}$. The zero of energy is taken as the energy of the system with infinitely separated neutral atoms. Thus,

$$E^{(0)} = -N E_{\mathrm{atom}} \tag{23}$$

where N is the number of atoms and E_{atom} is the energy of a neutral atom. This same zero of energy is used regardless of the overall charge of the system.

The next issue is to approximate the CI coefficients in eqs 1, 4, and 5. We do this by using the following approximation:

$$|C_j|^2 = \frac{\exp\left(\frac{-E_j}{\Delta}\right)}{\sum_k \exp\left(\frac{-E_k}{\Delta}\right)} \tag{24}$$

where Δ is a parameter called the *resonance integral*. (Note that if one defines $\Delta = k_B T$, where k_B is the Boltzmann constant, then one would have the classic equation for a Boltzmann distribution.) These same weights are used in eq 4 to determine the partial atomic charges and in eq 5 to determine the overall energy. The role of Δ is to determine how much each CSF contributes to the total energy E. Those CSFs where $E_j - \min(E_j) \gg \Delta$ will have practically zero contributions, while those with $E_j - \min(E_j) \ll \Delta$ will have nearly equal contributions. Thus, one would like to have Δ on the same order as the spread of CSFs that one wants to contribute to the energy.

For simplicity, in the present work, we approximate χ_{ij} in eq 6 as a constant Δ, the same as the resonance integral in eq 24. Whether one might obtain better results by approximating χ_{ij} in terms of the overlap of configurations i and j is a subject for a future study. The roles of Δ and χ_{ij}, in their respective equations, are responsible (*vide infra*) for determining which CSFs contribute significantly. While the approximation that χ_{ij} is constant and is equal to Δ was initially made to reduce the parameter space during the parametrization, it also helps ensure size consistency.

Since both of the sums in eq 6 are over all excitations, the number of which scales as N^2, the computational effort of this method (eqs 5 and 6) scales as N^4. To reduce the scaling of the problem, only those terms j are included for which $|C_j|^2 \geq 10^{-8}$, as the rest contribute negligibly to the overall energy and its derivatives. For small systems, most or all of the excitations contribute; in this case, the computational effort for the method does scale as N^4, but this is not a concern due to the small sizes of the systems. This truncation, however, has a very dramatic effect on the scaling and computational times for larger systems, as most of the CSFs do not have a significant contribution to the overall energy. For example, for TBCI-S on Al_{177}, only 54 of the 117 838 CSFs (i.e., 0.05%) are considered. For TBCI-SPD, only 81 of the 235 410 (i.e., 0.03%) CSFs pass the cutoff criterion. In general, over 98% of the CSFs were found to be neglected. One thus has a very small prefactor on the N^4 step, with the rest of the work in eq 6 scaling as N^2, so that the computational cost for many systems will be dominated by the diagonalization of the tight-binding Hamiltonian matrix, an N^3-scaling process.

It was noted above that in eq 15 U_Z^l is taken as a parameter rather than as the experimental VSIP. This is motivated by considering the ionization potential of the Al atom, which should be just the VSIP of the Al 3p orbital. In TBCI, there is, however, the E_{CB} term (*vide supra*) that contains $\gamma_{AA} q_A q_A$. Since $q_A = 1$ in the atomic cation and $q_A = 0$ in the neutral atom, this has a positive contribution to the cation and zero contribution to the neutral atom. In fact, one can show that for the Al atom:

$$U_{Al}^s = -I_{Al}^s + \frac{1}{\alpha_{Al}} \tag{25}$$

$$U_{Al}^p = -I_{Al}^p + \frac{1}{\alpha_{Al}} + 2\chi_{ij} \tag{26}$$

Equations 25 and 26 are used to determined the U_Z^l parameters in this work.

The method as described so far is suitable for only homonuclear systems, such as particles or clusters containing only aluminum atoms. When the theory is extended to heteronuclear systems, a number of entities become dependent on the atomic numbers of the atoms. Some of the variables—specifically U, α_A, and E_{atom}—depend only on a single atom's atomic number and become U_{Z_A}, α_{Z_A}, and $E_{atom}(Z_A)$, respectively, where Z_A is the atomic number of atom A. Others depend on two atomic numbers, which may or may not be different, and become $\kappa_{Z_A Z_B}$, $\tau_{Z_A Z_B}$, $K_{Z_i l_i Z_j l_j m}$, $\gamma_{Z_A Z_B}(R_{AB})$, and $d_{Z_A Z_B}$. Furthermore, the definition of the $E^{(0)}$ (eq 23) becomes

$$E^{(0)} = -\sum_Z N_Z E_{atom}(Z) \tag{27}$$

where N_Z is the number of atoms of atomic number Z, and $E_{atom}(Z)$ is the energy of a single atom. We will, however, continue in this article using the simpler notation that suffices for a homonuclear system.

One of the motivations behind developing TBCI was the problem, in TB, of improper charges when molecules dissociate. The fragment charges are obtained as the sum of the atomic charges of all of the atoms of a fragment. For the TBCI wave function, the atomic charges are obtained by combining eqs 4 and 18:

$$q_A = Z_A^* - \sum_{a \in A} \sum_j |C_j|^2 P_{aa}^{(j)} \tag{28}$$

where the weighting factors given by eq 24.

3. Analytical Gradients of TBCI

The derivative of the TBCI energy E (eq 5) with respect to a nuclear Cartesian coordinate X_C of atom C is

$$\frac{\partial E}{\partial X_C} = \sum_j \left(\frac{\partial |C_j|^2}{\partial X_C} \Gamma_j + |C_j|^2 \frac{\partial \Gamma_j}{\partial X_C} \right) \tag{29}$$

From the equations for $|C_j|^2$ (eq 24) and Γ_j (eq 6), one finds that their derivatives are

$$\frac{\partial |C_j|^2}{\partial X_C} = \frac{|C_j|^2}{\Delta} \left(\frac{\sum_k \left[\exp\left(-\frac{E_k}{\Delta}\right) \frac{\partial E_k}{\partial X_C} \right]}{\sum_k \exp\left(-\frac{E_k}{\Delta}\right)} - \frac{\partial E_j}{\partial X_C} \right) \tag{30}$$

$$\frac{\partial \Gamma_j}{\partial X_C} = \frac{\partial E_j}{\partial X_C} + \frac{1}{2} \sum_i \left(\frac{\partial E_i}{\partial X_C} - \frac{\partial E_j}{\partial X_C} \right) \left(1 - \frac{E_i - E_j}{\sqrt{(E_i - E_j)^2 + 4\chi_{ij}^2}} \right) \tag{31}$$

From eq 12 for the energy of each CSF, one obtains an expression for $\partial E_j / \partial X_C$ that has three terms; since $E^{(0)}$ is independent of geometry, its gradients are zero. The gradients of the core−core repulsion energy (eq 14) are

$$\frac{\partial V_{rep}}{\partial X_C} = -\sum_A \sum_{B>A} \tau V_{rep}^{AB} \frac{\partial R_{AB}}{\partial X_C} \tag{32}$$

The gradients of the valence energy can be found using the Hellmann−Feynman theorem.[38] This theorem states that

$$\frac{\partial}{\partial X_C}\langle\Psi|\mathbf{H}|\Psi\rangle = \langle\Psi|\frac{\partial \mathbf{H}}{\partial X_C}|\Psi\rangle \tag{33}$$

with $\langle\Psi|\Psi\rangle = 1$.[91,92] Therefore, if

$$\mathbf{c}_k^{\dagger}\mathbf{H}\mathbf{c}_k = \varepsilon_k \tag{34}$$

where $\mathbf{c}_k$ is a column vector of the matrix with elements c_{ak}, then

$$\frac{\partial \varepsilon_k}{\partial X_C} = \mathbf{c}_k^{\dagger}\frac{\partial \mathbf{H}}{\partial X_C}\mathbf{c}_k \tag{35}$$

The one-electron Hamiltonian gradient matrix can be found by taking the derivatives of eqs 15−17, as in previous TB work.[93,94]

Finally, consider the gradients of the charge balance term. Applying the product rule to eq 19 yields

$$\frac{\partial E_{\mathrm{CI}}^{(j)}}{\partial X_C} = \sum_A \sum_{B\geq A}\left[\frac{\partial \gamma(R_{AB})}{\partial X_C}q_A^{(j)}q_B^{(j)} + \gamma(R_{AB})\frac{\partial q_A^{(j)}}{\partial X_C}q_B^{(j)} + \gamma(R_{AB})\, q_A^{(j)}\frac{\partial q_B^{(j)}}{\partial X_C}\right] \tag{36}$$

where the gradients of the Coulomb integrals $\gamma(R_{AB})$ can be found by taking the derivatives of eq 22:

$$\frac{\partial \gamma(R_{AB})}{\partial X_C} = e^2 R_{AB}\left[\alpha_A\alpha_B \exp\left(\frac{-R_{AB}^2}{d\alpha_A\alpha_B}\right) + R_{AB}^2\right]^{-3/2} \times \left(\frac{\exp\left(\frac{-R_{AB}^2}{d\alpha_A\alpha_B}\right)}{d} - 1\right)\frac{\partial R_{AB}}{\partial X_C} \tag{37}$$

Clearly, $\partial\gamma(R_{AB})/\partial X_C$ is nonzero only if X_C is a Cartesian coordinate of either atom A or B. The CSF atomic charges $q_A^{(j)}$ were found using a Mulliken−Coulson population analysis[80−82] (eq 18), and thus,

$$\frac{\partial q_A^{(j)}}{\partial X_C} = -\sum_{a\in A}\frac{\partial q_A^{(j)}}{\partial P_{aa}^{(j)}}\frac{\partial P_{aa}^{(j)}}{\partial X_C} \tag{38}$$

In the Supporting Information to a paper by Giesen et al.,[95] it is shown that

$$\frac{\partial q_A^{(j)}}{\partial P_{aa}^{(j)}} = -\sum_{b\in A}\delta_{ab} \tag{39}$$

From eq 11,

$$\frac{\partial P_{aa}^{(j)}}{\partial X_C} = \sum_k^{MO} 2\cdot o_k^{(j)} c_{ak}\frac{\partial c_{ak}}{\partial X_C} \tag{40}$$

The derivatives of the eigenvector matrix can be found using a unitary transformation as shown by Dykstra and Jasien:[96]

$$\mathbf{c}^{X_C} = \mathbf{c}\mathbf{U}^{X_C} \tag{41}$$

where $\mathbf{c}^{X_C}$ is the matrix of the derivatives of the c_{ak} coefficients with respect to the nuclear coordinate X_C, and $\mathbf{U}^{X_C}$ is the unitary transformation. Dykstra and Jasien showed that the off-diagonal elements of $\mathbf{U}^{X_C}$ are[96]

$$U_{ab}^{X_C}(\varepsilon_a - \varepsilon_b) = R_{ab}^{X_C}\varepsilon_b - G_{ab}^{X_C} \tag{42}$$

where

$$\mathbf{R}^{X_C} = \mathbf{c}^{\dagger}\mathbf{S}^{X_C}\mathbf{c} \tag{43}$$

$$\mathbf{G}^{X_C} = \mathbf{c}^{\dagger}\mathbf{H}^{X_C}\mathbf{c} \tag{44}$$

Since the overlap matrix is neglected in TBCI (i.e., $S_{ab} = \delta_{ab}$, *vide supra*), $\mathbf{R}^{X_C} = 0$. The diagonal elements of $\mathbf{U}^{X_C}$ are zero since[96]

$$U_{aa}^{X_C} = -\frac{1}{2}R_{aa}^{X_C} = 0 \tag{45}$$

4. The Aluminum Databases

Four databases were used in the parametrization and evaluation of the TBCI models. Each database has three components: an energy database, an IP database, and a cluster dissociation database. The complete databases are provided in the Supporting Information.

The largest database—Al974—is the union of several subsets. The first subset is called the Al808 database, and it consists of the 808 aluminum clusters and their energies given in a previously published database;[45] the composition of the 808 cluster database is given in Table 1. In addition, there are 22 ionization potentials of small Al_{1-13} clusters and of three larger clusters (Al_{19}, Al_{43}, and Al_{55}). There are also 34 dissociations of clusters ranging from Al_2 to Al_{16} and 15 additional dissociations involving larger clusters. The small neutral and cationic clusters that form the Al82 database (*vide infra*) are also included. This database (Al974) is used in the evaluation of the various theoretical methods.

The next database—Al824—is a subset of the larger database. It contains 686 of the 808 clusters, specifically Al_n

Table 1. Number of Clusters (n_k) of Each Size (N_k) in the Al808 Cluster Database

N_k	n_k	N_k	n_k	N_k	n_k	N_k	n_k	N_k	n_k
2	44	16	1	39	1	64	1	92	1
3	402	17	1	42	2	65	1	93	1
4	79	18	4	43	14	68	1	104	1
7	42	19	27	50	2	69	1	105	1
9	1	20	1	51	7	78	2	113	1
10	1	21	5	54	2	79	7	128	1
11	1	26	1	55	12	80	1	129	1
12	4	27	11	56	1	81	1	134	3
13	65	28	5	57	6	86	3	135	3
14	2	35	5	58	1	87	10	141	1
15	7	38	1	59	6	89	1	177	1

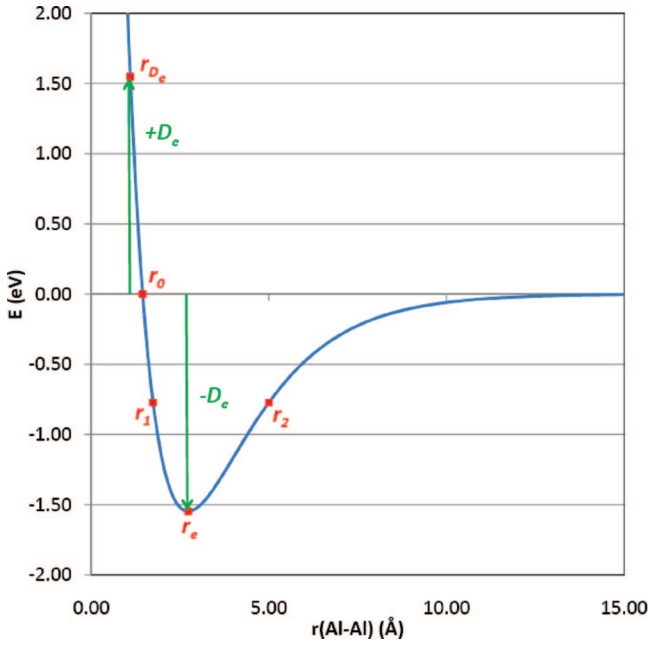

Figure 1. Diagram of a Morse curve showing r_e, r_{D_e}, r_0, r_1, and r_2.

where $n = 2-13$, 19, 35, 55, and 86. It also contains the Al82 database (*vide infra*). To this are added the 22 IPs and 34 dissociations involving small clusters only (*vide supra*). This database was used during the initial optimizations of the TBCI and TB+U methods.

The Al711 database, a subset of Al824, contains only the 711 clusters Al_n for which $n = 2-13$ (including those of appropriate size from the Al82 database). After significant preliminary work, it was found that this database is sufficient for optimizing the TBCI and TB+U methods (*vide infra*). This database does not include any IP or dissociation data.

Al82 is a database that was generated to fit the repulsive interaction of eq 14, as it is more balanced than the larger database for this purpose. For a number of reaction coordinates (stretch of Al_2, linear and perpendicular approach of Al to Al_2, symmetric stretch of D_{3h} Al_3, parallel and perpendicular approach of two Al_2 units, stretch of I_h Al_{13}, and an approach of two Al_9 units derived from a face-centered-cubic Al_{13} unit), five points (see Figure 1) were determined for both the neutral and cationic systems. First the minimum (r_e) of the potential energy surface was determined, along with the corresponding equilibrium dissociation energy (D_e). Then, using $V(r_e - 0.3$ Å) and $V(r = \infty)$, a Morse curve[97] was fit and the points $r_{D_e}(V = +D_e)$, r_0 ($V = 0$), and r_1 and r_2 ($V = -D_e/2$) were determined. The energies of all five points were determined at the PBEh/6-311+G(3d2f) or PBEh/MEC level of theory as appropriate (*vide supra*). For the smaller clusters ($n \leq 13$), this procedure was repeated for both the neutral and cationic systems; for the larger clusters, this procedure was done once for the neutral species, and the cationic species employ the same geometries as in the neutral case. Two extra points were added when, in two cases, the energy of the points r_{D_e} were too far from the predicted Morse curve.

The cluster energies and IPs were calculated at the PBEh/6-311+G(3d2f)[98–100] level of theory for clusters up to Al_{13} in size and at the PBEh/MEC level for larger clusters. The PBEh[101] hybrid exchange-correlation functional (Adamo and Barone's hybrid version[101] of the Perdew−Burke−Ernzerhof functional,[102,103] also called PBE0 or PBE1PBE) was chosen on the basis of comparisons[41] to accurate MCG3/3[104,105] energies for small clusters. The MEC basis set-relativistic effective core potential was designed to yield accurate energies for large aluminum clusters.[43]

The accuracy of the PBEh IP predictions was checked by comparing the IPs obtained for small clusters to IPs calculated using the coupled cluster *ab initio* method with all single and double substitutions[106] with a quasiperturbative estimate of the effect of the connected triple substitutions[107]—CCSD(T)—extrapolated to the complete-basis-set limit with the sequence of aug-cc-pV(n+d)Z basis sets. Dunning's aug-cc-pV(n+d)Z basis sets (n = D, T, Q)[108] were used as recommended by Martin et al. for the elements Al−Ar.[109,110] In particular, the Hartree−Fock (HF) energy, the CCSD correlation energy, and the connected triple excitations—(T)—contributions were extrapolated using the Weizmann-1 extrapolation scheme recently proposed by Martin and Parthiban:[111,112]

$$E_\infty = E_n + \frac{E_n - E_{n-1}}{(n/n-1)^\beta - 1} \tag{46}$$

where the HF energy and CCSD contributions are determined with the two larger basis sets (i.e., $n = 4$) with β values of 5 and 3.22, respectively, while the (T) contribution is extrapolated with the two smaller basis sets (i.e., $n = 3$) and with $\beta = 3$. The comparison presented in Table 2 between the PBEh/6-311+G(3d2f) and extrapolated CCSD(T) IPs shows that the former are sufficiently accurate. The geometries of these specific Al clusters are given in the Supporting Information.

5. The Fitting Procedure

The error function for a given model and database has a number of components and is similar to that used in previous work.[39,41–43,45] The mean unsigned error per atom in the energies of a set of n_k aluminum clusters of size N_k is

$$\varepsilon_{N_k} = \frac{1}{2N_k}\left(\frac{\sum_{i=1}^{n_k} w_i \Delta E_i^k}{\sum_{i=1}^{n_k} w_i} + \frac{\sum_{i=1}^{n_k-1}\sum_{j=i+1}^{n_k} w_i w_j \Delta\Delta E_{ij}^k}{\sum_{i=1}^{n_k-1}\sum_{j=i+1}^{n_k} w_i w_j}\right) \tag{47}$$

where w_i is the weight of cluster i (*vide infra*) and

$$\Delta E_i^k = |E_i^{k,\mathrm{PBEh}} - E_i^{k,\mathrm{TBCI}}| \tag{48}$$

$$\Delta\Delta E_{ij}^k = |\Delta E_i^k - \Delta E_j^k| \tag{49}$$

and $E_i^{k,\mathrm{PBEh}}$ and $E_i^{k,\mathrm{TBCI}}$ are the PBEh and TBCI energies, respectively, of cluster i of size N_k. The second term in eq 47 was found in previous work[39,41,42,45] to be important in order to obtain a better fit with respect to the relative energies within a set of clusters of the same size. The total mean unsigned error in the energies is

$$\varepsilon_{\mathrm{EN}} = \frac{1}{N_{N_k}}\sum_{k=1}^{N_{N_k}} N_k \varepsilon_{N_k} \tag{50}$$

Table 2. Comparison of the PBEh/6-311+G(3d2f) and Extrapolated CCSD(T) Ionization Potentials (eV)

	extrap. CCSD(T)[a]	PBEh	ΔIP_i[b]	%Δ[c]
Al_2 (R_{Al-Al} = 1.9 Å)	7.701	7.761	0.060	0.8
Al_2 (R_{Al-Al} = 2.528202 Å)	6.570	6.600	0.030	0.5
Al_2 (R_{Al-Al} = 2.7 Å)	6.362	6.352	−0.010	−0.2
Al_3 $C_{\infty v}$ (R_{Al-Al} = 2.863, 1.699 Å)	6.068	6.053	−0.014	−0.2
Al_3 D_{3h} (R_{Al-Al} = 2.59 Å)	6.514	6.545	0.031	0.5
Al_3 D_{3h} (R_{Al-Al} = 2.5066 Å)	6.454	6.533	0.080	1.2
Al_4 (edge-on approach of Al (R_{Al-Al} = 2.629 Å) to Al_3 D_{3h} R_{Al-Al} = 2.5066 Å)	6.450	6.410	−0.040	−0.6
Al_4 (vertex approach of Al (R_{Al-Al} = 1.800 Å) to Al_3 D_{3h} R_{Al-Al} = 2.863 Å)	6.465	6.376	−0.089	−1.5
Al_4 (top-on approach of Al (R_{Al-Al} = 3.08 Å) to a slightly distorted Al_3 D_{3h} R_{Al-Al} = 2.5 Å)	6.038	6.098	0.060	1.0
Al_4 (rhomboid (D_{2d}) R_{Al-Al} = 2.551 Å)	6.690	6.560	−0.110	−1.6
Al_5 (C_{2v} planar)	6.616	6.576	−0.040	−0.6
MUE[d]		0.0513		

[a] Extrapolated CCSD(T); see text. [b] Difference in IPs; see eq 54. [c] Percent difference in IP. [d] Mean unsigned error in eV.

where N_{N_k} is the number of different cluster sizes. At certain times during the parametrization (*vide infra*), the contribution of $\varepsilon_{N_k=2}$ to ε_{EN} was increased by a factor of 5; when this is done, $\varepsilon_{N_k=2}$ is multiplied by 5, but N_{N_k} and n_k were not changed.

The databases contain structures that are high in energy due to small interatomic distances. To prevent these structures from dominating the fit, weights (w_i) are included in eq 47. A scheme previously used[45] in the parametrization of analytical potential energy functions for aluminum was chosen. This scheme is defined as follows. If R_i is the smallest interatomic distance in a given system, then w_i is defined as

$$w_i = \begin{cases} 1 & R_i \geq R_{nc} \\ \dfrac{V_2(R_{nc})}{V_2(R_i)} & R_i < R_{nc} \end{cases} \tag{51}$$

where R_{nc} is the nonclose radius, the cutoff for the definition of "too close," and $V_2(R)$ is the energy of the aluminum dimer at separation R. Rather than calculate, using DFT, the diatomic energy for every R_i, all of the available Al_2 DFT data, with $R \leq 3.0$ Å, were fit to a polynomial. The obtained polynomial, with a fitness of $R^2 = 0.998$, is

$$V_2(R_i) = -15.66R_i^3 + 114.0R_i^2 - 275.4R_i + 219.2 \tag{52}$$

where the units of V_2 and R_i are electronvolts and Ångstroms, respectively. R_{nc} is chosen such that $V_2(R_{nc}) = -V_2(r_e)$, which results in $R_{nc} = 1.798$ Å.

The MUE for the IPs is

$$\varepsilon_{IP} = \frac{1}{2}\left(\frac{1}{n_{IP}}\sum_{i=1}^{n_{IP}} \Delta IP_i + \frac{2}{n_{IP}(n_{IP}-1)}\sum_{i=1}^{n_{IP}-1}\sum_{j=i+1}^{n_{IP}} \Delta\Delta IP_{ij}\right) \tag{53}$$

where n_{IP} is the number of clusters in the IP database and

$$\Delta IP_i = |IP_i^{PBEh} - IP_i^{TBCI}| \tag{54}$$

$$\Delta\Delta IP_{ij} = |\Delta IP_i - \Delta IP_j| \tag{55}$$

The MUE of the fragment charges upon dissociation, ε_{dis}, is:

$$\varepsilon_{dis} = \frac{\Omega \dfrac{\text{hartree}}{\text{unit charge}}}{n_{dis}} \sum_{i=1}^{n_{dis}} |q_i| \tag{56}$$

where n_{dis} is the number of dissociating clusters in the database, q_i is the charge on each fragment of dissociated cluster i as determined by TBCI, and Ω is a constant taken as 0.1 on the basis of initial estimates of the relative magnitudes of ε_{EN}, ε_{IP}, and ε_{dis}. The errors in the dissociation charges are $|q_i|$ since experimentally only neutral fragments are observed. The total error function ε is defined as

$$\varepsilon = \frac{N_{N_k}\varepsilon_{EN} + \varepsilon_{IP} + \varepsilon_{dis}}{N_{N_k} + 1 + \Omega} \tag{57}$$

The error function in eq 47 is in units of electronvolts per atom, while all other error functions are in units of electronvolts.

In order to fit the parameters in the TBCI models, a microgenetic algorithm[113] was used, specifically version 1.7a of Carroll's FORTRAN code.[114] locally modified with our own fitness function and designed to run in parallel using the message-passing interface (MPI).[115,116] Because genetic algorithms, by definition, maximize a given function, the fitness function, f, used was minus the total error function ε.

There are different components to the TBCI model. Rather than optimize all of the parameters at once, it was decided to optimize the model in stages. Thus, for a given component being optimized, initial values for the other parameters are chosen on the basis of reasonable values—either from physically reasonable values or from a previous optimization—that were kept fixed during the optimization. The parameters in V_{rep} (i.e., κ and τ) were the first to be optimized. During the initial stages, values for the remaining parameters were chosen as follows:

• In E_{val}, the Wolfsberg−Helmholz parameters were taken from a previously published TB model,[46] specifically the third entry in Table S1 of this reference, which corresponds to a TB-WH model, where all of the Wolfsberg−Helmholz constants ($K_{l_il_jm}$, eq 16) are given by a single K_0. The Slater-

Table 3. Final Parameters of the TBCI and TB+U Models

	model				model		
parameter (units)	S	SPD	TB+U	parameter (units)	S	SPD	TB+U
$K_{ss\sigma}$ (unitless)	0.033858	0.052569	0.045132	κ (eV)	796.05	573.02	564.15
$K_{sp\sigma}$ (unitless)	0.019876	0.37467	0.54078	τ (Å^{-1})	2.8216	2.6602	2.5734
$K_{pp\sigma}$ (unitless)	1.6238	1.8041	1.4896	α (Å)	3.7804	3.7804	3.7804
$K_{pp\pi}$ (unitless)	1.9347	1.2179	1.5931	d (unitless)	1.0	1.0	1.0
$\zeta_{ss\sigma}$ (Å^{-1})	2.5935	2.5935	2.5935	Δ (eV)	0.05	0.05	0.05
$\zeta_{sp\sigma}$ (Å^{-1})	2.5772	2.5772	2.5772	$E^{(0)}$ (eV)	−15.851	−15.951	−15.851
$\zeta_{pp\sigma}$ (Å^{-1})	2.5610	2.5610	2.5610	U_s (eV)	6.8110	6.8110	6.8110
$\zeta_{pp\pi}$ (Å^{-1})	2.5610	2.5610	2.5610	U_p (eV)	1.9770	1.7770	2.0770

type orbital exponents (ζ_a) were taken from the Hartree−Fock calculations by Clementi and Raimondi.[78] Thus, $K_0 = 0.409\,61$, $\zeta_s = 2.5935$ Å^{-1}, and $\zeta_p = 2.5610$ Å^{-1}.

• The diagonal Hamiltonian elements U_a (eq 15) were determined using eqs 25 and 26, where α_{Al} was determined as below, and the VSIPs were taken from experiments and are $I^s_{Al} = 10.620$ eV and $I^p_{Al} = 5.986$ eV;[117,118] thus, $U^s_{Al} = 6.811$ eV and $U^p_{Al} = 1.977$ eV.

• For γ_{AB} (eq 22), the three parameters were chosen on the basis of physical intuition; for simplicity, $d = 1.0$. Since α_{Al} is the distance where the repulsion switches from an r^{-1} to an e^{-r} behavior, and since this occurs as the two electron clouds start to overlap, α_{Al} was chosen as twice the van der Waals radius for Al, which has been experimentally determined to be 1.89 Å.[119]

• The parameter Δ controls how many CSFs contribute significantly to the configuration interaction wave function. It was chosen to be approximately 2 orders of magnitude smaller than the ionization potential of the aluminum atom; in particular, $\Delta = 0.10$ eV.

• The zero of energy, $E^{(0)}$, was determined from a calculation on the aluminum atom with the given set of parameters.

With the above parameters frozen, the two parameters in V_{rep} were optimized. For this optimization, the Al82 database was used and the quantity that was minimized was ε_{EN} (eq 50). Once these parameters were optimized, they were frozen and K_0 was then optimized. This was done using Al824 and fitting to ε (eq 57) rather than ε_{EN}. This cycle of optimizations was repeated. Finally, all three parameters in each model were allowed to vary. From these values, models were optimized with different values of Δ.

While the TBCI models based on TB-WH—especially TBCI-SPD—showed reasonable results, notably in the fragmentation charges, they are not sufficiently reliable. Therefore, the next level of TB approximation—EWH—was used for the TBCI and TB+U models. In this model, four different $K_{l,l,m}$ (eq 16) values are used. Because preliminary evaluations showed that the previous models were most deficient in the performance for small clusters, initially the TBCI models were optimized against ε_{EN} for the Al711 database. Finally, for both TBCI and TB+U, each with several fixed values of Δ, we simultaneously optimized six parameters (four Wolfsberg−Helmholz and two repulsion parameters) against ε for our Al824 database.

It was noted during the evaluation of the obtained models that, while the overall performance was satisfactory, the models predicted Al_2 to be too strongly bound by over half an electronvolt. By increasing the relative weight of ε_2 in ε_{EN} (see eq 50) 5-fold, models were obtained that showed improved performance for Al_2 without significantly compromising the fits to the rest of the data.

6. Computational Methods and Software

The CCSD(T) calculations were performed using the MOLPRO 2006.2 *ab initio* program package.[120] All DFT calculations were carried out using *Gaussian 03*,[121] except for the IP calculations for the clusters larger than Al_{55}, which were calculated using NWChem, version 4.5.[122] The TBCI calculations were done using an in-house code. The comparisons to the previously published NP-A and NP-B potentials[45] were done using published routines.[123] The comparisons to the TB and MBTB models[39,44] were done using the TB 2.0 code.[124]

The newly developed TBCI and TB+U models are implemented in TBPAC 2007,[125] which is available from the authors at http://comp.chem.umn.edu/tbpac/.

7. Results and Discussion

Three different TBCI models were examined: TBCI-S, TBCI-SPD, and TB+U; for TB+U, we used only single excitations. On the basis of extensive tests, we found that $\Delta = 0.05$ eV led to better results than $\Delta = 0.10$ eV and dramatically better results than $\Delta = 0.20$ eV; therefore, we chose $\Delta = 0.05$ eV for the final optimizations. We also set $\chi_{ij} = \Delta$. For all parameter sets in the present article, we also constrained $d = 1$ without optimization. All other parameters were optimized or frozen as discussed above. The final parameters are in Table 3.

In addition to comparisons between the TBCI models, the TBCI results are compared in Table 4 to three other kinds of results, with the comparison in all three cases based on comparing the errors measured against the PBEh Al974 database. The first and second kinds of methods to which we compare are the set of results obtained using TB. Previously, six different TB models were parametrized—three based on the Wolfsberg−Helmholz model in eqs 15 and 16 (TB-WH, TB-EWH, and TB-OWH) and three that contain many-body terms (MBTB, specifically TB-S, TB-CN, and TB-BA that include, respectively, screening, coordination number and bond angle many-body effects). Five parametrizations of the WH model were used, one from the original TB paper (herein denoted as TB-WH(SSST))[39] and four (denoted TB-WH(JSTi), $i = 1-4$) that are the first four entries in Table S-1 of ref 46. The deficiencies of these TB

Table 4. Mean Errors[a]

model	ε_{EN} (0)	ε (0)	ε_{EN}	ε_{IP}	ε_{dis}	ε	max(q)[b]
TBCI-S	1.01	65.52	1.08	704.8	1.07	65.59	±6(×5)
TBCI-SPD	0.94	54.40	1.01	571.1	0.89	54.48	±5(×7)
TB+U[c]	0.99	86.71	1.12	1287.2	1.29	86.83	±6(×6)
TB-WH(SSST)	2.33	93.00	2.44	570.3	1.66	93.11	±6
TB-WH(JST1)	2.11	102.53	2.23	582.3	1.86[d]	102.65	±7(×4)[d]
TB-WH(JST2)	2.00	93.19	2.12	598.6	1.66[d]	93.30	±7(×2)[d]
TB-WH(JST3)	1.93	90.39	2.05	591.3	1.61[d]	90.50	±7[d]
TB-WH(JST4)	1.15	150.34	1.30	450.3	2.91	150.49	±10
TB-EWH	1.05	110.95	1.15	529.9	2.07	111.05	±8
TB-OWH	0.71	380.59	0.80	514.8	7.64	380.69	±36
MBTB-S	1.74	114.24	1.87	482.3	2.14	114.37	±9(×2)
MBTB-CN	2.01	357.36	2.09	403.8	7.18	357.45	±34
MBTB-BA	1.57	249.32	1.68	488.9	4.93	249.44	±18
NP-A[e]	0.71						
NP-B[e]	0.80						

[a] ε_{EN}, ε_{IP}, and ε are in units of meV. ε_{dis} is in units of charge. The first two columns with (0) exclude the data in Al82. [b] Maximum fragment charge of the clusters in the dissociation part of Al974. The number in parentheses denotes the number of instances of this maximum charge. [c] TB+U is with single excitations only. [d] All clusters for the methods in this row (in the case of TB-WH(JST3)-all but one) have at least ±1 charge. [e] Energies of neutral clusters only; see text.

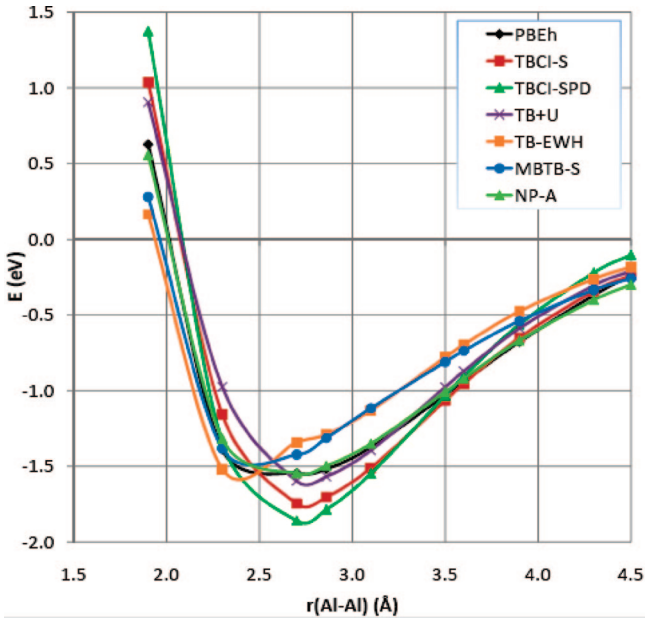

Figure 2. Plot of the Al_2 stretch potential energy surface for various theoretical methods.

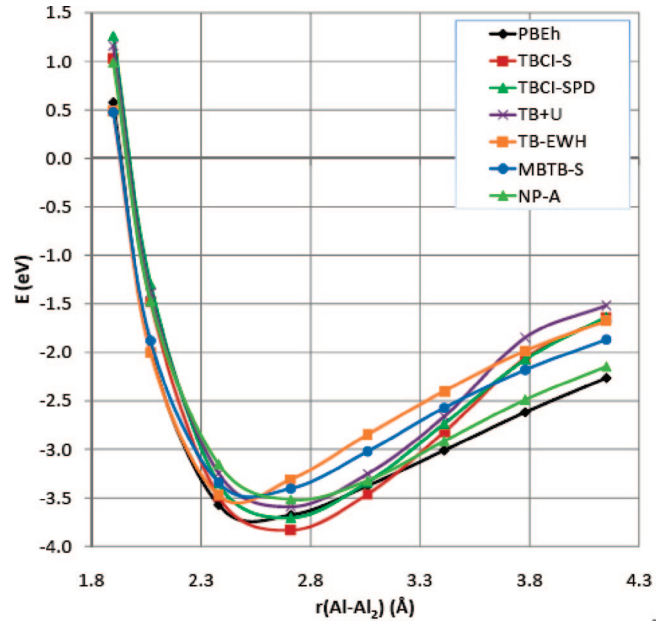

Figure 3. Plot of the perpendicular Al_2 + Al (C_{2v}, $r(Al_2)$ = 2.863 Å) potential energy surface for various theoretical methods.

methods, as well as the anticipated difficulty of extending them to heteronuclear systems, were the impetus behind this work; for a full description of these methods, see refs 39 and 46. The third kind of method to which we compare is the analytic PEFs, in particular NP-A and NP-B, which we previously developed.[45] These PEFs were found to accurately estimate the energy of aluminum clusters; since they do not contain any information on electrons or charges, they are incapable of predicting IPs or charges. NP-A is the more accurate of the two PEFs, and it includes many-body terms, while NP-B is an order of magnitude less computationally expensive but nearly as accurate.

Table 4 presents a comparison of the new and previous methods. This evaluation is over Al974. Also given in this table is an evaluation over Al974 with Al82 excluded. This was done because NP-A and NP-B cannot handle charged systems.

7.1. Preliminary Observations. Before discussing the final versions of the three new methods, we first mention some observations made during the parametrization.

First, we note that requiring all four Wolfsberg−Helmholz $K_{l,l,jm}$ constants to be the same (i.e., using the WH approximation, *vide supra*) leads to significantly larger (about a factor of 2 to 3) values of ε_{EN}, the average error in the energies of the clusters and nanoparticles. The optimized value of a single K is typically in the range of 0.55−0.8, but Table 3 shows that the final optimized values range from 0.03 to 1.93. In light of this wide range of optimized values, it is not surprising that a single compromise value is much worse. Removing the restriction that all four $K_{l,l,jm}$'s be equal (i.e., moving to the EWH approximation, *vide supra*) also lowered ε_{IP} by about a factor of 2.

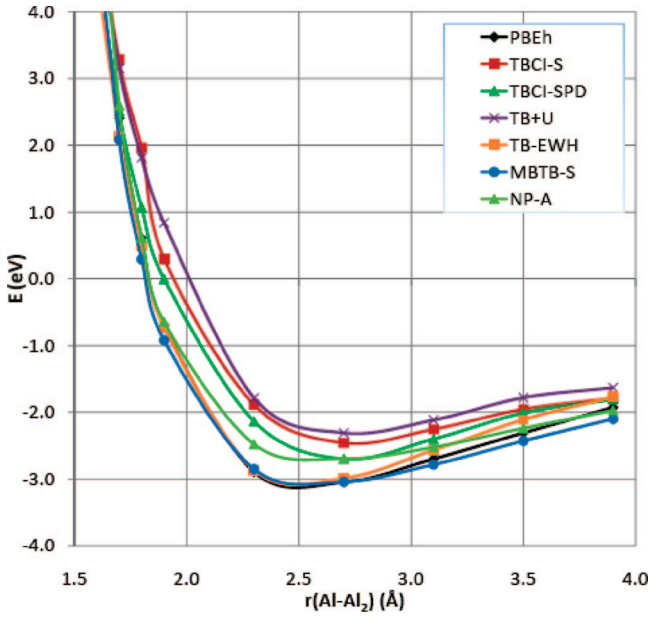

Figure 4. Plot of the linear Al + Al_2 ($C_{\infty v}$, $r(Al_2)$ = 2.863 Å) potential energy surface for various theoretical methods.

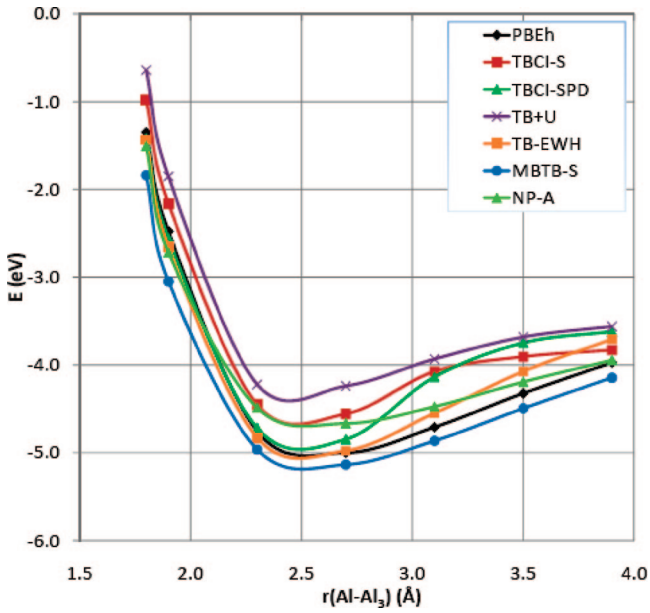

Figure 5. Plot of the on-top approach of Al to Al_3 (D_{3h}, r(Al−Al) = 2.863 Å) potential energy surface for various theoretical methods.

Another observation is that optimizing against only ε_{EN} for Al_n with $n = 2-13$ (i.e., Al711) is capable of yielding average errors (ε) for Al974 that are quite close (∼5−20% larger) to those obtained by optimizing over the larger Al824 database. This is an indication of the robustness of the methods, and thus the final optimizations were over this smaller subset.

It was also noted that, while the methods performed very well for the energies of Al clusters and nanoparticles, one glaring exception was Al_2. The initial parametrizations predicted Al_2 to be too strongly bound by more than 0.5 eV. A 5-fold increase of the relative weight of the Al_2 data

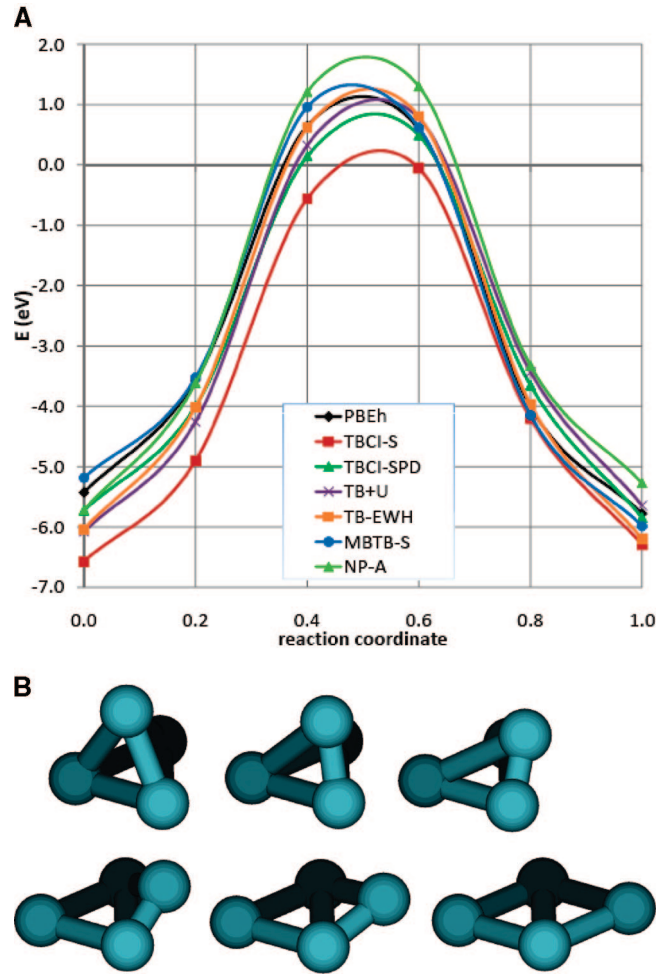

Figure 6. (A) Plot of the potential energy surface for the transition between tetrahedral (T_d) and rhomboid (D_{2h}) Al_4 for different theoretical methods. (B) Images of the clusters in the transition between tetrahedral and rhomboidal Al_4.

in the evaluation of ε_{EN} (eq 50) provided TBCI and TB+U models that gave better predictions of the Al_2 potential energy surface with minimal (insignificant) deterioration of the remaining data predictions. A 10-fold increase provided excellent prediction of the Al_2 curve but resulted in poor potential energy curves for the other small clusters. Therefore, the final optimizations were against ε_{EN} for Al_{2-13} with a 5-fold increase in the weights for Al_2.

If we consider the TBCI-SPD calculations on all of the clusters in the energy and dissociation subsets of Al974, we find that the reference CSF (i.e, the aufbau CSF) is the dominant CSF in the CI expansion in only 57% of the cases.

7.2. Comparison of the Methods. Figures 2−7 depict potential energy profiles for various one-dimensional cuts through the potential energy surfaces of Al_2 to Al_7. These figures compare the results to PBEh, TB-EWH, MBTB-S, and NP-A. On the basis of these figures and Table 4, we can draw some conclusions.

First of all, we see that TBCI with the SPD configurational selection scheme is slightly better than the S scheme, on average, with the final ε decreasing by 17%. Both TBCI models yield similar errors in the cluster and nanoparticle energies (i.e., ε_{EN}), but TBCI-SPD is better in predicting IPs

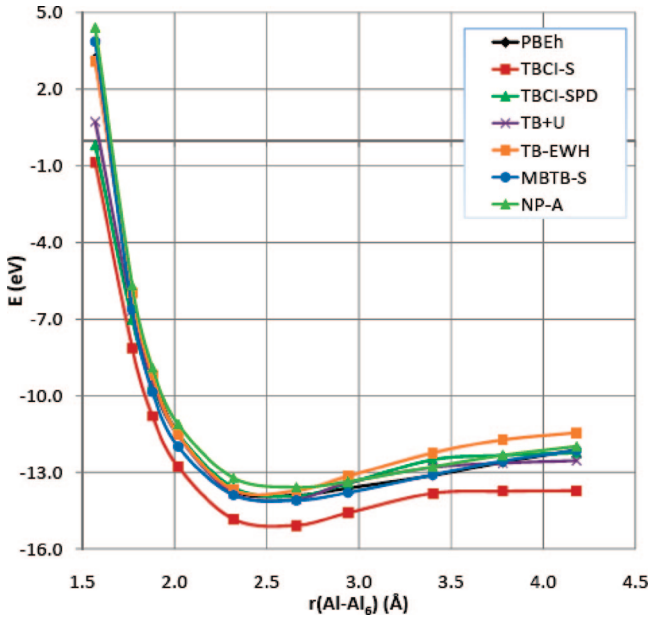

Figure 7. Plot of the edge approach of Al to Al_6 (O_h, r(Al−Al) = 2.8635 Å) potential energy surface for various theoretical methods.

(i.e., ε_{IP}) and fragment charges (i.e., ε_{dis}). TB+U, which is slightly faster than the TBCI models, performs equally well on predicting cluster and nanoparticle energies but is much worse in predicting fragmentation charges and is woeful in predicting IPs. In contrast to the traditional Hubbard model,[12] the TB+U method does not give the correct dissociation limit because the Hubbard (+U) correction is not introduced self-consistently in the present formalism.

In general, compared to the previously published TB and MBTB models,[39] the new methods are superior in predicting cluster and nanoparticle energies; the sole exception is the TB-OWH model. While TBCI-SPD and, to a lesser extent, TBCI-S are capable of predicting IPs with similar accuracy to that of the TB and MBTB methods, the new methods are superior when considering fragmentation charges. The TBCI and TB+U methods predict many of the cases to be neutral, while some of the TB methods are incapable of predicting any neutral fragments. In fact, TB-OWH predicts 15 of the dissociations to have fragment charges of ±10 or more (12 with ±20 or more).

It would be challenging to explain how TBCI is capable of eliminating such large charges. These charges are intrinsic artifacts of the methods, and not of our implementation, and affect also the older work of Slater and Koster,[1] Wolfsberg and Helmholz,[70] and Hoffmann.[2,72–75] In the TB models, the large charges result from the density of states in the valence energy region being more localized on one dissociation fragment. This would result in more electron density, and hence negative charge, being localized on this fragment, which in turn results in a positive charge localized on the other fragment. The excitations in TBCI could alleviate this charge imbalance by transferring electrons from one cluster to the other. However, only single and pair double excitations are considered, and thus, a maximum of two units of charge could be transferred. Nonetheless, much larger charges are being alleviated in the TBCI models. The most likely explanation is that the charge balance term prevents the Wolfsberg−Helmholz ($K_{l_il_jm}$) and repulsion (κ and τ) parameters from entering regions of parameter space, during their optimizations, that would result in large charge imbalances. Even more remarkable is the observation that reasonable fragmentation charges are obtained despite the fact that they were not included in the optimization of the TBCI models. In fact, when the TBCI models are evaluated against Al974 using the TB-OWH $K_{l_il_jm}$, $\zeta_{l_il_jm}$, and repulsion parameters, similarly large fragment charges are obtained as with TB-OWH. Thus, the charge balance term is clearly responsible for dampening, during the parametrization, any large charges that may occur. Note that, in this test of TB-OWH with CB terms, instead of eq 14, the following model for the repulsion was used here as in TB-OWH:

$$V_{rep} = \kappa \sum_A \sum_{B>A} \frac{\exp(-\tau R_{AB})}{R_{AB}^u} \tag{58}$$

where u is a constant.[39]

8. Summary

In summary, a new TB method has been proposed and developed. This model, called TBCI, improves on TB by applying a configuration-interaction-like procedure based on the TB orbitals. In such a manner, partial charges are incorporated into the calculation in a noniterative manner. This new TBCI model was optimized for aluminum nanoclusters and found to give exceptional performance with a low average error. The method is also applicable to other kinds of systems.

Acknowledgment. Research was supported in part by the Defense University Research Initiative on Nanotechnology (DURINT) through a grant managed by the Army Research Office. Computer resources were provided by the Minnesota Supercomputing Institute. M.A.I. also acknowledges support of a research grant from the Herbert J. Seligman Charitable Trust. A.H. gratefully acknowledges the Minnesota Supercomputing Institute for a research scholarship.

Supporting Information Available: The Al82 database, the ionization potentials and fragment dissociation components of Al824, the geometries of the clusters in Table 2, and the structures of the clusters in Figure 6. This material is available free of charge from the authors or via the Internet at http://pubs.acs.org.

References

(1) Slater, J. C.; Koster, G. F. *Phys. Rev.* **1954**, *94*, 1498–1524.

(2) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397–1412.

(3) *Tight-Binding Approach to Computational Materials Science*; Turchi, P. E. A., Gonis, A., Colombo, L., Eds.; MRS Symposium Proceedings 491, Materials Research Society: Warendale, PA, 1998.

(4) Harris, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *31*, 1770–1779.

(5) Foulkes, W. M. C.; Haydock, R. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1989**, *39*, 12520–12546.

(6) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133−A1138.

(7) Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 111–132.

(8) Slater, J. C. *Phys. Rev.* **1930**, *35*, 210–211.

(9) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(10) Rein, R.; Fukuda, N.; Win, H.; Clarke, G. A. *J. Chem. Phys.* **1966**, *45*, 4743–4744.

(11) Newns, D. M. *Phys. Rev.* **1969**, *178*, 1123–1135.

(12) Hubbard, J. *Proc. R. Soc. London, Ser. A* **1963**, *276*, 238–257.

(13) Grimley, T. B.; Pisani, C. *J. Phys. C: Solid State Phys.* **1974**, *7*, 2831–2848.

(14) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenhein, T.; Suhai, S.; Seifert, G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 7260–7268.

(15) Frauenhein, T.; Seifert, G.; Elstner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. *Phys. Status Solidi B* **2000**, *217*, 41–62.

(16) Micnas, R.; Ranninger, J.; Robaszkiewicz, S. *Rev. Mod. Phys.* **1990**, *62*, 113–171.

(17) Strack, R.; Vollhardt, D. *Phys. Rev. Lett.* **1993**, *70*, 2637–2640.

(18) Arrachea, L.; Aligia, A. A. *Phys. Rev. Lett.* **1994**, *73*, 2240–2243.

(19) Pulay, P. *Mol. Phys.* **1969**, *17*, 197–204.

(20) Fock, V. *Z. Phys.* **1930**, *61*, 126–148.

(21) Hartree, D. R.; Hartree, W. *Proc. R. Soc. London, Ser. A* **1935**, *150*, 9–33.

(22) Hartree, D. R.; Hartree, W. *Proc. R. Soc. London, Ser. A* **1936**, *154*, 588–607.

(23) Hartree, D. R. *Rep. Prog. Phys.* **1948**, *11*, 113–143.

(24) Hartree, D. R. *The Calculation of Atomic Structures*; John Wiley & Sons: New York, NY, 1957.

(25) Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

(26) Grimley, T. B. In *Adsorption at Solid Surfaces*; King, D. A., Woodruff, D. P., Eds.; Elsevier: Amsterdam, 1983; p 333.

(27) Shavitt, I. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum Press: New York, NY, 1977; Vol. 4.

(28) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(29) Adams, G. F.; Bent, G. D.; Bartlett, R. J.; Purvis, G. D. In *Potential Energy Surfaces and Dynamics Calculations for Chemical Reactions and Molecular Energy Transfer*; Truhlar, D. G., Ed.; Plenum Press: New York, NY, 1981; pp 133−167.

(30) Head-Gordon, M. *J. Phys. Chem.* **1996**, *100*, 13213–13225.

(31) Shaik, S.; Hiberty, P. C. *Rev. Comput. Chem.* **2004**, *20*, 1–100.

(32) Truhlar, D. G. *J. Comput. Chem.* **2007**, *28*, 73–86.

(33) Watts, J. D.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 8718–8733.

(34) Brillouin, L. *Actual. Sci. Ind.* **1933**, *71*.

(35) Brillouin, L. *Actual. Sci. Ind.* **1934**, *159*.

(36) Nesbet, R. K. *Proc. R. Soc. London, Ser. A* **1955**, *230*, 312–321.

(37) Levy, B.; Berthier, G. *Int. J. Quantum Chem.* **1968**, *2*, 307–319.

(38) Simons, J.; Nichols, J. A. In *Quantum Mechanics in Chemistry*; Oxford University Press: New York, NY, 1997; p 442.

(39) Staszewska, G.; Staszewski, P.; Schultz, N. E.; Truhlar, D. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *71*, 045423.

(40) Valone, S. M.; Atlas, S. R. *J. Chem. Phys.* **2004**, *120*, 7262–7273.

(41) Schultz, N. E.; Staszewska, G.; Staszewski, P.; Truhlar, D. G. *J. Phys. Chem. B* **2004**, *108*, 4850–4861.

(42) Jasper, A. W.; Staszewski, P.; Staszewska, G.; Schultz, N. E.; Truhlar, D. G. *J. Phys. Chem. B* **2004**, *108*, 8996–9010.

(43) Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 41–53.

(44) Staszewska, G.; Staszewski, P.; Schultz, N. E.; Truhlar, D. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 039903(E).

(45) Jasper, A. W.; Schultz, N. E.; Truhlar, D. G. *J. Phys. Chem. B* **2005**, *109*, 3915–3920.

(46) Jasper, A. W.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 210–218.

(47) Li, Z. H.; Jasper, A. W.; Truhlar, D. G. *J. Am. Chem. Soc.* **2007**, *129*, 14899–14910.

(48) Li, Z. H.; Bhatt, D.; Schultz, N. E.; Siepmann, J. I.; Truhlar, D. G. *J. Phys. Chem. C* **2007**, *111*, 16227–16242.

(49) Bhatt, D.; Schultz, N. E.; Jasper, A. W.; Siepmann, J. I.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 26135–26142.

(50) Bhatt, D.; Jasper, A. W.; Schultz, N. E.; Siepmann, J. I.; Truhlar, D. G. *J. Am. Chem. Soc.* **2006**, *128*, 4224–4225.

(51) Züttel, A.; Wenger, P.; Sudan, P.; Mauron, P.; Orimo, S.-i. *Mater. Sci. Eng., B* **2004**, *108*, 9–18.

(52) Ramaswamy, A. L.; Kaste, P.; Trevino, S. F. *J. Energ. Mater.* **2004**, *21*, 1–24.

(53) Ramaswamy, A. L.; Kaste, P. *J. Energ. Mater.* **2005**, *23*, 1–25.

(54) Galfetti, L.; De Luca, L. T.; Severini, F.; Meda, L.; Marra, G.; Marchetti, M.; Regi, M.; Bellucci, S. *J. Phys.: Condens. Matter* **2006**, *18*, S1991−S2005.

(55) Barnett, R. N.; Yannouleas, C.; Landman, U. *Z. Phys. D: At., Mol. Clusters* **1993**, *26*, 119–125.

(56) Mulder, F. M.; Thiel, R. C.; de Jongh, L. J.; Gubbens, P. C. M. *Nanostruct. Mater.* **1996**, *7*, 269–292.

(57) Kara, A.; Rahman, T. S. *Phys. Rev. Lett.* **1998**, *81*, 1453–1456.

(58) Link, S.; El-Sayed, M. A. *Int. Rev. Phys. Chem.* **2000**, *19*, 409–453.

(59) Voisin, C.; Del Fatti, N.; Christofilos, D.; Vallée, F. *J. Phys. Chem. B* **2001**, *105*, 2264–2280.

(60) *Metal Nanoparticles: Synthesis, Characterization, and Applications*; Feldheim, D. L., Foss, C. A., Jr., Eds.; Marcel Dekker: New York, NY, 2002.

(61) *Nanopartilces: From Theory to Application*; Schmid, G., Ed.; Wiley-VCH: Weinheim, Germany, 2004.

(62) Mitev, P.; Papageorgiou, D. G.; Lekka, C. E.; Evangelakis, G. A. *Surf. Sci.* **2004**, *566*−*568*, 937–943.

(63) Mazzone, A. M.; Morandi, V. *Comput. Mater. Sci.* **2007**, *38*, 830–837.

(64) Sun, C. Q. *Prog. Solid State Chem.* **2007**, *35*, 1–159.

(65) Jiang, Q.; Ao, Z. M.; Zheng, W. T. *Chem. Phys. Lett.* **2007**, *439*, 102–104.

(66) Assfeld, X.; Almlöf, J. E.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *241*, 438–444.

(67) Pople, J. A.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S136−S151.

(68) Klopman, G.; Evans, R. C. In *Semiempirical Methods of Electronic Structure Theory, Part A*; Segal, G. A.; Plenum Press: New York, NY, 1977; pp 29−67.

(69) Wang, Y.; Mak, C. H. *Chem. Phys. Lett.* **1995**, *235*, 37–46.

(70) Wolfsberg, M.; Helmholz, L. *J. Chem. Phys.* **1952**, *20*, 837–843.

(71) Mulliken, R. S. *J. Chim. Phys. Phys. Chim. Biol.* **1949**, *46*, 497–542.

(72) Hoffmann, R. *J. Chem. Phys.* **1964**, *40*, 2047–2048.

(73) Hoffmann, R. *J. Chem. Phys.* **1964**, *40*, 2474–2480.

(74) Hoffmann, R. *J. Chem. Phys.* **1964**, *40*, 2480–2488.

(75) Hoffmann, R. *J. Chem. Phys.* **1964**, *40*, 2745–2745.

(76) Slater, J. C. *Phys. Rev.* **1930**, *36*, 57–64.

(77) Jones, H. W. *Int. J. Quantum Chem.* **1981**, *19*, 567–574.

(78) Clementi, E.; Raimondi, D. L. *J. Chem. Phys.* **1963**, *38*, 2686–2689.

(79) Jones, H. W. *Int. J. Quantum Chem.* **1980**, *18*, 709–713.

(80) Mulliken, R. S. *J. Chem. Phys.* **1935**, *3*, 564–573.

(81) Coulson, C. A.; Longuet-Higgins, H. C. *Proc. R. Soc. London, Ser. A* **1947**, *191*, 39–60.

(82) Maslen, V. W.; Coulson, C. A. *J. Chem. Soc.* **1957**, 4041–4049.

(83) Liechtenstein, A. I.; Anisimov, V. I.; Zaanen, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *52*, R5467–R5470.

(84) Leung, K.; Rempe, S. B.; Schultz, P. A.; Sproviero, E. M.; Batista, V. S.; Chandross, M. E.; Medorth, C. J. *J. Am. Chem. Soc.* **2006**, *128*, 3659–3668.

(85) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 767–776.

(86) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

(87) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, *6*, 1–72.

(88) Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219–227.

(89) Klopman, G. *J. Am. Chem. Soc.* **1964**, *86*, 4550–4557.

(90) Tucker, S. C.; Truhlar, D. G. *Chem. Phys. Lett.* **1989**, *157*, 164–170.

(91) Hellman, J. *Einführung in die Quantenchemie*; Deuticke & Co.: Leipzig, 1937.

(92) Feynman, R. P. *Phys. Rev.* **1939**, *56*, 340–343.

(93) Liu, T. A Tight Binding Model for the Energetics of Hydrocarbon Fragments on Metal Surfaces. Ph.D. Thesis, University of Minnesota, Minneapolis, MN, May 2000.

(94) Liu, T.; Truhlar, D. G. Unpublished results.

(95) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057–1068.

(96) Dykstra, C. E.; Jasien, P. G. *Chem. Phys. Lett.* **1984**, *109*, 388–393.

(97) Morse, P. M. *Phys. Rev.* **1929**, *34*, 57–64.

(98) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294–301.

(99) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265–3269.

(100) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

(101) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(102) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(103) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

(104) Fast, P. L.; Sánchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407–410.

(105) Tratz, C. M.; Fast, P. L.; Truhlar, D. G. *PhysChemComm* **1999**, *2*, 70–79.

(106) Purvis, G. D., III; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(107) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(108) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244–9253.

(109) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129–4141.

(110) Martin, J. M. L. *THEOCHEM* **2006**, *771*, 19–26.

(111) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843–1856.

(112) Martin, J. M. L.; Parthiban, S. In *Quantum Mechanical Prediction of Thermochemical Data*; Cioslowski, J., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001; pp 31−65.

(113) Carroll, D. L. In *Developments in Theoretical and Applied Mechanics*; Wilson, H., Batara, R., Bert, C., Davis, A., Schapery, R., Stewart, D., Swinson, F., Eds.; School of Engineering, The University of Alabama: Tuscaloosa, AL, 1996; Vol *XVII*, p 411.

(114) Carroll, D. L. GA_version 1.7a: FORTRAN Genetic Algorith, Driver; CU Aerospace: Urbana, IL, 2001. See: http://cuaerosopace.com/carroll/ga.html (accessed Mar 2008).

(115) Message Passing Interface Forum. MPI: A message-passing interface standard. Computer Science Department Technical Report CS-93-214; University of Tennessee, Knoxville, TN, November 1993.

(116) Gropp, W.; Lusk, E.; Skjellum, A. *Using MPI: Portable Parallel Programming with the Message-Passing Interface*, 2nd ed.; MIT Press: Cambridge, MA, 1999.

(117) *CRC Handbook of Chemistry and Physics*, 78th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 1997.

(118) Moore, C. E. *Atomic Energy Levels as Derived from the Analyses of Optical Spectra, Circular of the National Bureau of Standards 467*; U.S. Department of Commerce: Washington, DC, 1949; Vol. 1.

(119) Behm, J. M.; Blume, T.; Morse, M. D. *J. Chem. Phys.* **1994**, *101*, 5454–5463.

(120) Werner, H. J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. MOLPRO, version 2006.2, a package of ab initio programs. See: http://www.molpro.net (accessed Mar 2008).

(121) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.01; Gaussian, Inc.: Pittsburgh, PA, 2003.

(122) Straatsma, T. P.; Apra, E.; Windus, T. L.; Dupuis, M.; Bylaska, E. J.; de Jong, W.; Hirata, S.; Smith, D. M. A.; Hackler, M. T.; Pollack, L.; Harrison, R. J.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Valiev, M.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem, A Computational Chemistry Package for Parallel Computers*, version 4.5; Pacific Northwest National Laboratory: Richland, WA, 2003.

(123) Duchovic, R. J.; Volobuev, Y. L.; Lynch, G. C.; Jasper, A. W.; Truhlar, D. G.; Allison, T. C.; Wagner, A. F.; Garrett, B. C.; Espinosa-García, J.; Corchado, J. C. POTLIB-online. http://comp.chem.umn.edu/potlib (accessed Mar 2008).

(124) Staszewska, G.; Liu, T.; Truhlar, D. G. *TB*, version 2.0; University of Minnesota: Minneapolis, MN, 2003.

(125) Iron, M. A.; Staszewska, G.; Liu, T.; Jasper, A. W.; Truhlar, D. G. *TBPAC 2007*; University of Minnesota: Minneapolis, MN, 2007.

CT700343T