Using Bonding to Guide Transition State Optimization

Adam B. Birkholz and H. Bernhard Schlegel*

Optimization of a transition state typically requires both a good initial guess of the molecular structure and one or more computationally demanding Hessian calculations to converge reliably. Often, the transition state being optimized corresponds to the barrier in a chemical reaction where bonds are being broken and formed. Utilizing the geometries and bonding information for reactants and products, an algorithm is outlined to reliably interpolate an initial guess for the transition state geometry. Additionally, the change in bonding is also used to increase the reliability of transition state optimizations that utilize approximate and updated Hessian information. These methods are described and compared against standard transition state optimization methods. © 2015 Wiley Periodicals, Inc.

Introduction

The optimization of equilibrium geometries and transition states is an important step in the computational study of chemical reactions (see Refs. [1] and [2] for overviews of these methods). In general, geometry optimization begins by computing the energy and forces acting on a guess structure and then explores a local approximation to the potential energy surface (PES) with the goal of locating a structure where the forces are zero (the forces are the negative of the potential energy gradient). This is accomplished by an iterative process, where each successive optimization step produces a geometry closer to the solution, and new energies and forces are computed to update the local approximation. The commonly used Newton–Raphson type optimization methods utilize information about the second derivatives of the PES, the Hessian matrix, to explore the surface with greater confidence than using only the forces alone and to reduce the number of optimization steps required. The full Hessian matrix can be computed analytically at a significantly higher computational cost than required to produce an energy and gradient. Alternatively, a few of the eigenvectors and eigenvalues may be computed analytically or numerically at a reduced cost.[3–5] For quasi-Newton optimizations seeking a minimum energy structure, it is usually sufficient to begin with a diagonal, positive definite estimate based on empirically determined values.[6–8] Estimating the Hessian for transition state optimization by the quasi-Newton method is more difficult, as the Hessian matrix must have exactly one negative eigenvalue.

Quasi-Newton methods begin with an analytical or estimated Hessian matrix, and then update it using gradients calculated during the course of the optimization with one of a variety of different Hessian update schemes (e.g., see Refs. [9–15]). When using a (quasi-)Newton Raphson method to locate transition states, the Hessian matrix must have exactly one negative eigenvalue. The eigenvector corresponding to this eigenvalue is also called the transition vector, and the energy along this vector is a maximum at the transition state. If the initial structure to be optimized is sufficiently close to the actual transition state geometry, the exact Hessian matrix computed by analytical or numerical means will have only one negative eigenvalue and any Hessian update formula that allows for negative eigenvalues[9,14,15] may be used to converge to the transition state geometry. However, it is often difficult to produce such a good guess geometry, and other approaches must be used, such as selecting which eigenvector to follow uphill and then correcting the Hessian before computing the step,[16] or utilizing a reduced potential surface to restrict the transition state search to the dominant coordinates in the reaction.[17–19]

If appropriate estimates of the reactant and product structure are used, the transition vector can be estimated by the tangent of a simple interpolated pathway,[20] and an empirical estimate of the Hessian may be used after some initial steps are taken along the approximate pathway to correct the curvature.[21] However, these estimates tend to be good approximations to the reaction path only when the structures corresponding to the reactant and product in the interpolation lie close to the barrier on the PES. In practice, when fully optimized reactant and product geometries are used, considerable reorientation of the geometry may be required before the reaction can occur. Methods such as the dimer method[22] or one of many reaction path optimization methods[23–27] may be used to explore the surface to locate a region that is likely to contain the transition state. These produce good initial guess structures for quasi-Newton transition state optimizations and in some cases may also be used to generate suitable approximate or updated Hessians[28,29] that rapidly accelerates convergence to the transition state geometry, but at the cost of many gradient calculations prior to the beginning of the transition state search.

A. B. Birkholz, H. B. Schlegel
Department of Chemistry, Wayne State University, Detroit, Michigan 48202
E-mail: hbs@chem.wayne.edu

Previous work[30] demonstrated that a reaction pathway may be accurately described by a small number of composite coordinates constructed as linear combinations of primitive Cartesian or internal coordinates. When the transition state represents the energy barrier for a simple conformational change in geometry, several bending, stretching, and torsional coordinates may contribute significantly to the motion along the path near the transition state, and it can be difficult to determine which of these coordinates are necessary to represent the transition vector based on the reactant and product geometries alone. However, for a chemical reaction, the desired transition state necessarily involves the forming and/or breaking of bonds. In the following work, we attempt to use knowledge about the bonding in reactants and products to produce a guess structure for the geometry at the transition state. Additionally, this bonding information is used to decrease the need for the computation of an analytical Hessian and to improve the selection of the transition vector during the optimization to reduce the number of gradient calculations needed to locate a transition state. To evaluate the performance of our proposed methods, 20 reactions were selected from a variety of published[21,31–33] and unpublished libraries for use as a test set (see Fig. 1 for details). This test set contains characteristic examples of many types of organic reactions including insertions, additions, eliminations, hydrolysis, ring openings, substitutions, cycloadditions, and rearrangements. Reactions containing transition metals will be examined in future work.

Methods

Choice of coordinate system

The first step in any geometry optimization algorithm is the selection of a coordinate system to describe the changes in the geometry. While the Cartesian coordinates of the nuclear centers are an intuitive and straightforward set of coordinates to use, a redundant set of internal coordinates[34] (bond stretches, angle bends, and dihedral torsions) are a more natural basis for describing the potential energy landscape for a chemical system. Such coordinate systems reduce the coupling between the various coordinates, resulting in a much more diagonally dominant Hessian matrix than typically observed in a Cartesian representation. This is especially important for optimization methods that rely on Hessian updating, and hence, redundant internal coordinate systems have enjoyed a widespread use in geometry optimization.

Any number of different coordinate sets may be constructed from the possible combinations of bond stretches, angles, and dihedrals, so long as enough coordinates are included to completely describe the internal degrees of freedom of the system. The standard approach to define a set of redundant internal coordinates for optimization is to begin by determining the bonding skeleton of the structure. This can be input to the optimizer by the user, or it can be generated automatically by computing all of the distances between atoms in the molecule, and considering any two atoms bonded if they are roughly as close as, or closer than, a reference single bond

length for those two atoms. Once the connectivity of the structure is known, the coordinate system may be constructed by including the bond stretches between any bonded atoms, the bends between any pairs of stretches that share an atom, and dihedral angles between bends that share a bond. Once this set is defined, extra care must be taken to be sure to include any coordinates necessary to describe the bending of a near-linear angle, the motion into/out of a plane for atoms in a nearly planar configuration, as well as any coordinates necessary to describe the relative motion of two unconnected fragments in the geometry.

This approach is complicated when selecting a coordinate set for optimization of the transition state for a chemical reaction. In this case, the bond lengths corresponding to bonds breaking would be much longer in the product structure, and bond lengths corresponding to bonds forming are going to be significantly longer in the reactants. If the reactant or product structures are multimolecular, additional virtual bonds must be added, usually joining the pair of atoms in two molecules with the shortest distance between them. This will result in a coordinate set that includes coordinates to describe the relative motion between fragments, but the added virtual bond may not necessarily be a bond involved in the reaction if significant reorientation of the molecules is required to move from the minimum structure to the region of the PES where the transition state lies. Including such coordinates can cause problems in an optimization and should be avoided. Likewise, special coordinates such as those needed to describe the bending motions of angles that are nearly linear might be appropriate for describing the reactant and/or the product even when the corresponding angle is nonlinear at the transition structure, and the inclusion of such coordinates can also frustrate an optimizer. For these reasons, taking the union of coordinates defined at the reactants and products may not be the best approach for a transition state optimization.

The approach described in this work, instead, merges the bonding skeletons from the reactant and product structures, and then uses the merged skeleton to define the redundant internal coordinates in the standard fashion. For the purpose of interpolating an initial geometry, an initial coordinate set may be defined by adding or removing the necessary coordinates to describe the linear or planar structures in the reactants and the products. Once the initial guess for the transition state is generated, the coordinate set is redefined using the geometry of the guess, so that only the linear and planar coordinates necessary to describe the interpolated structure are included. Defining the coordinates in this way, the number of reactions that fail to produce an initial geometry by linear interpolation of the reactants and products is reduced from 6 to 2 for the 20 reactions in our test suite. To further improve the reliability of interpolating reactant and product structures to produce an approximate transition state geometry, a new approach must be introduced.

Bond order interpolation with relaxation

To find the transition state corresponding to a reaction where the minimized structures of the reactants and products are
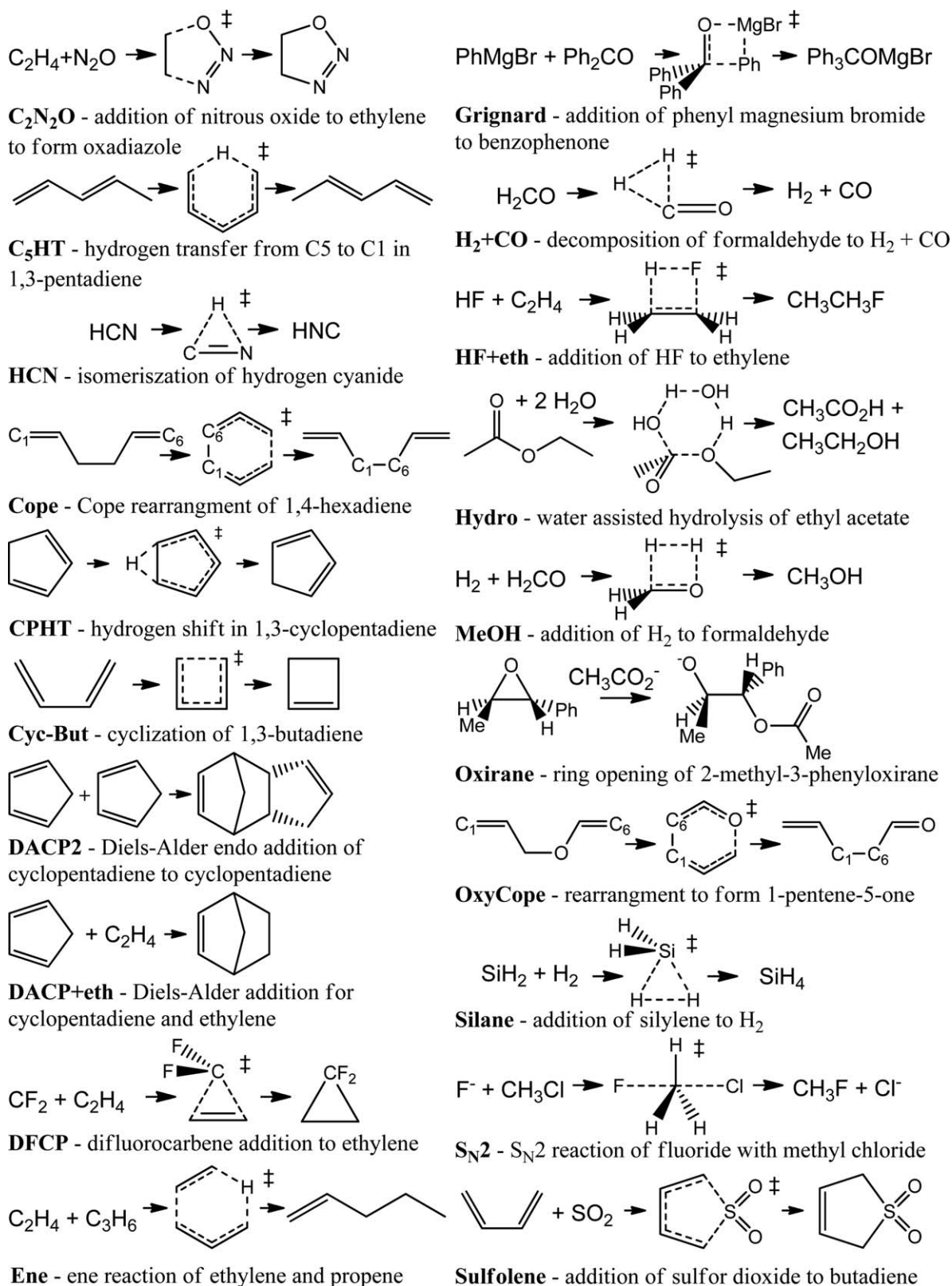
Figure 1. Reactions for transition state optimization.

already known, the typical approach is to begin by interpolating between the two structures to generate a better guess for the transition state optimization. This is usually done by a sim-

ple linear interpolation, often in distance matrix or internal coordinates, to the midpoint of the line connecting the reactants and products. This approach includes an explicit

dependence on the reactant and product geometries, which can result in wildly different guess structures when different, equally valid, geometries are input. A simple example would be in the case of a bimolecular reaction, where the interpolated midpoint is highly dependent on the distance between the two reactants. Typically, only a few of the total number of coordinates that describe the system are involved with the reaction, and it is better to limit the interpolation to that reduced set. Intuitively, the bonds that are breaking or forming over the course of the reaction must belong to that set, and it should be sufficient to interpolate along only those coordinates.

To use the bonding information to interpolate a guess of the structure at the transition state, it is necessary to have a convenient description of bonding. Using the concept of bond order, a relationship between a bonds length and approximate strength can be described. Pauling[35] suggested using the following equation

$$o(r) = e^{(r_0 - r)/a} \qquad (1)$$

$$r(o) = r_0 - a \ln(o) \qquad (2)$$

where o is the order, r_0 is a reference single bond length for the atoms involved, and a is a positive constant determined by fitting to known experimental or computational data. With $a = 0.3$ Å, this equation agrees well with typical values for double and triple bonds. Unfortunately, Pauling bond order initially decays quite rapidly when r is longer than a single bond length, and there is a smooth decrease in the bond order out to infinity. Using eq. (2), a bond of order 0.5 would be roughly 114% the length of the reference single bond length, while the associated bond length will increase slowly to infinity as the order approaches zero. These are characteristics that are undesirable for interpolation of bonds in transition state geometries.

Instead, this work will use the following approximation for interpolation of the bond order

$$\tilde{o}(r) = \max\left[0, \frac{b(r_0/r) - 1}{b - 1}\right]$$

$$r(\tilde{o}) = \frac{b}{(b-1)\tilde{o} + 1} r_0, \quad \tilde{o} \geq 0 \qquad (3)$$

While this form is not as accurate as the Pauling bond order for bond orders greater than one, it decays much more quickly as the bond length increases, and will assign a bond order of zero to all bond lengths greater than a particular value (determined by selection of b). A typical bond length at a transition state for a bond being broken or formed is between roughly 120–150% the length of the corresponding standard bond length, and it is usually safe to assume that when the distance between two atoms is greater than twice the standard bond length they are not bonded. Using a value of 2 for b in the equations above results in a bond length 133% of the reference at bond order 0.5, and a bond order of zero for bond lengths greater than or equal to 200% the reference length. Table 1 lists the mean, maximum and minimum bond order of

Table 1. Barrier distances versus the bond order at the PM6 transition state geometry.

Reaction	Barrier			CTS bond order[a]			
	Forward (mH)	Reverse (mH)	Distance[b] (%)	n_{CTS}	Mean	Max	Min
C_2N_2O	4.43	8.68	34	2	0.32	0.33	0.31
C_5HT	6.44	6.44	50	2	0.52	0.52	0.52
HCN	13.70	11.58	54	2	0.49	0.66	0.33
Cope	5.99	5.99	50	2	0.75	0.75	0.75
CPHT	6.13	6.13	50	2	0.55	0.55	0.55
Cyc-but	7.11	6.36	53	1	0.45	0.45	0.45
DACP2	5.33	8.88	37	2	0.43	0.55	0.32
DACP + eth	4.89	8.84	36	2	0.44	0.44	0.44
DFCP	1.49	7.37	17	2	0.58	0.77	0.39
Ene	6.16	10.03	38	3	0.57	0.69	0.42
Grignard	3.74	6.81	35	2	0.73	1.07	0.39
$H_2 + CO$	12.34	9.32	57	3	0.59	0.94	0.24
HF + eth	8.23	10.24	45	3	0.49	0.51	0.45
Hydro	3.23	3.36	49	6	0.54	0.71	0.40
MeOH	13.39	13.54	50	3	0.41	0.62	0.23
Oxirane	3.42	2.66	56	2	0.45	0.46	0.44
Oxycope	5.10	6.80	43	2	0.62	0.65	0.59
Silane	4.06	8.46	32	3	0.65	0.92	0.20
S_N2	0.35	6.02	6	2	0.49	0.73	0.24
Sulfolene	2.65	8.59	24	2	0.54	0.54	0.54

[a] n_{CTS} is the number of bonds breaking or forming in the reaction. [b] The energetic distance of the barrier is defined as the ratio of the forward barrier height to the sum of the forward and reverse heights. A distance much smaller than 50% can be viewed as an indicator of an early transition state, while a distance much larger than 50% can be viewed as an indicator of a late transition state.

the bonds being broken/formed for the geometries in the test set computed at the PM6 level of the theory and demonstrates that a bond order of roughly 0.5 produces reasonable bond lengths for a wide variety of reactions.

Using this definition of bond order, the linear interpolation can be restricted to the bond order of only the bonds being broken or formed during the reaction. A direct, least-squares fit to the change in the bonding coordinates that is necessary to achieve the interpolated structure may cause other parts of the geometry to change in unfavorable ways, and bring atoms too close to one another. To maintain reasonable values for the remaining coordinates, a relaxed scan that minimizes an inexpensive approximate potential energy orthogonal to the set of bonds is used.

Molecular mechanics potentials like the Universal Force Field (UFF)[36] are among the least expensive approximations to the energy of a chemical system, and should be sufficient for the purpose of relaxing the coordinates orthogonal to the bonds being broken or formed. The UFF potential is simple and parameterized for a large number of atom types, and when the combined connectivity skeleton from the reactants and products is used to select the atom types, it is generally sufficient to produce bonds and angles that avoid very unfavorable regions of the actual potential. Alternatively, semiempirical electronic structure methods such as PM6[37] may produce structures that are better suited for optimization of transition states using Hartree–Fock or density functional theory (DFT).

To avoid significantly impacting the total cost of the optimization, the low level method used during interpolation and relaxation should be at least an order of magnitude faster than the electronic structure method used for the transition state optimization. For different classes of reactions other than the representative organic reactions discussed in this work, such as those involving transition metals, the selection of the appropriate low level method for this interpolation process may require further investigation.

The proposed interpolation algorithm is as follows:

1. Determine which atom pairs are bound for both the reactant and product. Compare the lists of bonds to determine which bonds are breaking/forming, and combine the lists of bonds to construct a redundant coordinate set.
2. Compute the bond orders for the breaking/forming set in both reactants and products, and compute a set of goal bond lengths for these bonds from the average value of the bond order.
3. Clean up the reactant and product structures by fitting the displacements to bring any bond lengths longer than two times the reference length down to two times the reference length, and then minimize the low level energy orthogonal to the breaking/forming set.
4. Beginning with the reactant structure, take a series of steps no greater than 0.5 bohr along the displacement to the goal lengths. Following each step, the low level energy is minimized by updating the coordinates orthogonal to the breaking/forming set.
5. Repeat 4, beginning with the product structure.
6. Select the guess with the lower energy, evaluated by the low level method, as the initial geometry to optimize to a transition state on the actual potential.

When force field methods are used for interpolation and relaxation, some care is needed for defining the atom types for atoms that undergo a change in bonding. The atom types for these atoms are selected based on the number of connected groups in the combined connectivity skeleton if such an atom type is defined, otherwise the atom type is set to the atom type with the largest number of connected groups for that atom.

When this algorithm is applied to interpolate an initial guess structure for each of the 20 reactions in the test set, a structure with reasonable bond lengths and angles is produced even when beginning with fully minimized reactant and product structure. In some cases, however, the optimization of the structures either fails to converge or converges to a transition state for an alternate reaction, particularly when Hessian updating is used. To improve the reliability of the transition state search, new methods for selecting the transition vector, computing the optimization step, and estimating the Hessian for a transition state optimization were developed.

Using connectivity change to approximate the transition vector

Once an approximate structure has been obtained, a transition state optimization needs to be carried out to find the structure

that maximizes energy along the transition vector while minimizing the energy along all other coordinates. The Newton–Raphson method can accomplish this using a shifted Hessian to compute the displacement toward the stationary point with the desired curvature

$$\Delta \mathbf{q} = -(\mathbf{H} + \mathbf{S})^{-1} \mathbf{g} \tag{4}$$

where $\mathbf{g}$ and $\mathbf{H}$ are the gradient and Hessian of the PES, and $\mathbf{S}$ is a matrix that is typically constructed to control the direction and the size of the displacement step. For a transition state optimization, the shifted Hessian must have exactly one negative eigenvalue. The rational function optimization (RFO) method defines $\mathbf{S}$ as a scaled identity matrix (i.e., $\mathbf{S} = \lambda\ \mathbf{I}$), where the scaling factor is the predicted energy change of the step on a local rational function approximation to the PES.[16] This scaling factor is computed straightforwardly as an eigenvalue of the augmented Hessian

$$\mathbf{H}_{\mathrm{aug}} = \begin{pmatrix} \mathbf{H} & \mathbf{g} \\ \mathbf{g}^T & 0 \end{pmatrix} \tag{5}$$

For a geometry defined by n variables, the augmented Hessian has $n + 1$ eigenvalues, and using the nth most negative eigenvalue results in a shifted Hessian with $n - 1$ negative eigenvalues.

While the RFO method can compute a single correction such that the shifted Hessian has the correct number of negative eigenvalues, this will always result in following the eigenvector of the Hessian with the most negative eigenvalue uphill. The partitioned RFO (pRFO) method selects a specific eigenvector to follow uphill, which may not necessarily be the one corresponding to the most negative eigenvalue, and defines $\mathbf{S}$ as the following sum

$$\mathbf{S}_{\mathrm{pRFO}} = -\lambda_1\ \mathbf{v}_{\mathrm{TS}} \mathbf{v}_{\mathrm{TS}}^T - \lambda_2 \sum_{i}^{n-1} \mathbf{v}_i \mathbf{v}_i^T \tag{6}$$

where $\mathbf{v}_{\mathrm{TS}}$ is the eigenvector whose corresponding eigenvalue should be negative and λ_1 is the most positive eigenvalue of the augmented Hessian constructed in the eigenspace defined by $\mathbf{v}_{\mathrm{TS}}$, while the $\mathbf{v}_i$ are the remaining $n - 1$ eigenvectors of the Hessian and λ_2 is the most negative eigenvalue of the augmented Hessian defined in that space.

The most difficult part of a transition state optimization that uses the pRFO method is the selection of the eigenvector along which the energy should be maximized. One existing strategy that has been used effectively in the past is to compare the eigenvectors of the Hessian to the tangent of a simple approximation to the reaction path. These approximate pathways are typically functions of the geometries corresponding to the reactant and product, with some also including a dependence on the geometry at the current step of the optimization. Three such functions[20,21,38] that have been used in the past are a linear path connecting reactants ($\mathbf{q}_{\mathrm{R}}$) and products ($\mathbf{q}_{\mathrm{P}}$), a quadratic path connecting $\mathbf{q}_{\mathrm{R}}$ and $\mathbf{q}_{\mathrm{P}}$ through this estimate of the transition state ($\mathbf{q}_{\mathrm{TS}}$), and an arc of a circle

connecting $\mathbf{q}_R$ and $\mathbf{q}_P$ through $\mathbf{q}_{TS}$. The tangent for the linear (τ_{lin}), quadratic (τ_{quad}), and arc of a circle (τ_{arc}) paths are given by

$$\tau_{lin} = \mathbf{q}_P - \mathbf{q}_R \tag{7}$$

$$t_{lin} = \frac{\tau_{lin}^T (\mathbf{q}_{TS} - \mathbf{q}_R)}{\tau_{lin}^T \tau_{lin}} \tag{8}$$

$$\tau_{quad} = \tau_{lin} + \frac{(1 - 2t_{quad})((\mathbf{q}_P - \mathbf{q}_R) - t_{quad}\tau_{lin})}{t_{quad} - t_{quad}^2} \tag{9}$$

$$\tau_{arc} = \frac{(\mathbf{q}_P - \mathbf{q}_{TS})}{|\mathbf{q}_P - \mathbf{q}_{TS}|^2} - \frac{(\mathbf{q}_R - \mathbf{q}_{TS})}{|\mathbf{q}_R - \mathbf{q}_{TS}|^2} \tag{10}$$

see Bell and Crighton[38] for the derivation of the quadratic vector, and Peng et al.[31] for the derivation of the arc vector. All coordinate values are expressed in redundant internal coordinates.

These approaches can produce an accurate estimate of the transition vector when the reaction is relatively simple, or when the reactant-like and product-like structures supplied are located quite close to the transition state, but they tend to be less reliable when fully optimized structures are used, and when the reaction path has multiple, distinct regions with different curvature. When the transition state involves bond breaking and forming, bond stretch coordinates tend to be the dominant coordinates in the transition vector. As an alternative to the above approaches, a simple delta bonding (Δb) vector is proposed, whose elements are defined in the following way:

1. Stretches corresponding to bonds forming in the reaction have the value 1.
2. Stretches corresponding to bonds being broken in the reaction have the value −1.
3. All other coordinates have a value of 0.

This vector is trivial to compute and does not depend explicitly on the geometry at the reactant, product, or transition state. Additionally, for the reactions in the test set, Δb has a large overlap with the transition vector while typically having a small overlap with all remaining eigenvectors of the Hessian computed at the transition state geometry. Table 2 compares the linear, quadratic, and arc of circle tangents to the Δb vector for the set of reactions shown in Figure 1. The vectors, in redundant internal coordinates, are computed according to the eqs. (7–10) using the transition state geometry, and then projected into the locally nonredundant space

$$\tau_{NR} = \mathbf{B}\mathbf{B}^+ \tau \tag{11}$$

where $\mathbf{B}$ is the Wilson B-matrix, which contains the partial derivatives of the internal coordinates with respect to a change in the Cartesian coordinates $B_{ij} = (\partial q_i / \partial x_j)$. $\mathbf{B}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{B}$. The Δb vector not only has a large overlap with the transition vector in the majority

Table 2. Comparison of overlap between different tangent approximations and eigenvectors of the Hessian computed at the transition state of the PM6 PES.

Reaction	Transition vector[a]				Other vector[b]			
	τ_{lin}	τ_{quad}	τ_{arc}	Δb	τ_{lin}	τ_{quad}	τ_{arc}	Δb
C_2N_2O	0.69	0.28	0.68	**0.89**	0.38	0.32	0.39	0.38
C_5HT	0.13	0.13	0.13	**0.82**	0.67	0.67	0.67	0.36
HCN	0.84	0.77	0.83	**0.91**	0.37	0.37	0.38	0.29
Cope	0.35	0.35	0.35	**0.78**	0.19	0.19	0.19	0.43
CPHT	0.28	0.21	0.28	**0.83**	0.24	0.37	0.24	0.15
Cyc-but	0.13	0.10	0.07	**0.70**	0.59	0.29	0.58	0.52
DACP2	0.20	0.04	0.14	**0.84**	0.56	0.22	0.47	0.36
DACP + eth	0.17	0.03	0.12	**0.83**	0.37	0.16	0.30	0.20
DFCP	0.48	0.52	0.53	**0.89**	0.20	0.22	0.15	0.21
Ene	0.30	0.28	0.35	**0.84**	0.35	0.24	0.33	0.22
Grignard	0.11	0.09	0.11	**0.69**	0.32	0.39	0.32	0.43
$H_2 + CO$	0.60	0.75	0.79	**0.82**	0.25	0.50	0.33	0.40
HF + eth	0.44	0.44	0.44	**0.96**	0.41	0.41	0.41	0.14
Hydro	0.30	0.27	0.30	**0.92**	0.29	0.28	0.29	0.17
MeOH	0.55	0.47	0.54	**0.87**	0.34	0.36	0.30	0.29
Oxirane	0.15	0.18	0.18	**0.84**	0.38	0.26	0.36	0.33
Oxycope	0.26	0.19	0.25	**0.79**	0.26	0.23	0.26	0.34
Silane	0.63	0.56	0.66	**0.72**	0.68	0.48	0.68	0.59
S_N2	**0.96**	0.01	0.83	0.92	0.18	0.74	0.45	0.35
Sulfolene	0.30	0.02	0.23	**0.83**	0.52	0.31	0.37	0.26

[a] Transition vector refers to eigenvector with a corresponding negative eigenvalue. The bold numbers indicate which approximation has the largest overlap with the transition vector. [b] The other vector columns report the largest overlap with the tangent approximation among the remaining eigenvectors.

of the reactions but it also tends to have a small overlap with the remaining coordinates.

Initial approximation of the Hessian

For minimizations, a simple valence force field Hessian tends to work well as it is a rough, but reasonable, approximation to the exact Hessian at a minimum. This reduces the amount of work that the optimizer and update formulas must do to improve it. Additionally, when the initial geometry is poor and forces are large, the shift matrix $\mathbf{S}$ in eq. (4) will tend to dominate, resulting in similar optimization steps regardless of whether the Hessian used is exact or approximate. For transition state optimizations, however, the direction of the step is considerably more sensitive to the accuracy of the Hessian, particularly in how the transition vector couples to the minimization modes when far from the solution. As the initial few steps of transition state optimizations will be largely downhill, it would be beneficial to make use of the increased stability of the downhill RFO steps and Broyden-Fletcher-Goldfarb-Shanno (BFGS) updates to get as close as possible to the transition state when using approximate or updated Hessian information.

The method described earlier to interpolate an estimated transition state geometry along the bonds being broken or formed during the reaction utilizes a less expensive method to approximate the PES for the purpose of relaxing the geometry along the remaining coordinates. While these methods generally do a good job at predicting bond lengths and angles, they may neglect or simplify many of the nonbonding

interactions that are more correctly accounted for in Hartree–Fock or DFT calculations. The resulting forces may still be quite large in the displacements corresponding primarily to the non-bonded interactions. Relaxing these modes on the same PES that the transition state optimization will be carried out on prior to beginning the transition state search should help to reduce some of the risks associated with using approximated and updated Hessian information.

As it would be difficult to attempt to define a set of coordinates corresponding only to the nonbonded interactions, the minimization phase is carried out by choosing a small number of coordinates to be frozen while there remaining are allowed to relax. The frozen set contains the bonds being broken or formed, and angles between any pairs of these bonds that share an atom. If the convergence criterion on the minimization phase is too loose, the geometry may not change enough to make a meaningful impact on the transition state optimization. Likewise, convergence criteria that are too tight may result in an unnecessarily large number of optimization steps in the minimization phase with no improvement to the transition state optimization, particularly in any case where the actual transition vector is not entirely contained in the frozen set. Using the root mean square (RMS) of the gradient in the active set of coordinates as the sole convergence metric, a threshold of 3.2×10^{-4} hartree/bohr was determined to provide a good balance between the time spent in the minimization phase and the efficiency and reliability of the transition state optimization. This is comparable to the convergence threshold used for regular geometry optimization.

During the minimization phase, the empirical approximation to the Hessian will work as well as it does in a standard minimization, and use of the BFGS update formula should be sufficient to achieve convergence. Following the minimization phase, however, it is likely that the approximate Hessian will be positive definite, making it a poor choice to begin the transition state search. Given that Δb is a good approximation to the transition vector, the initial Hessian for the transition state optimization may be constructed from the updated, positive definite Hessian from the minimization phase ($\tilde{\mathbf{H}}$) by reversing the sign and adjusting the magnitude of the Rayleigh coefficient along the locally nonredundant projection of Δb. Using τ_{NR} from (11), the initial Hessian $\mathbf{H}_0$ for the transition state optimization is given by

$$\mathbf{H}_0 = \tilde{\mathbf{H}} - 1.5 \frac{\tau_{NR}^T \mathbf{H} \tau_{NR}}{\tau_{NR}^T \tau_{NR}} \frac{\tau_{NR} \tau_{NR}^T}{\tau_{NR}^T \tau_{NR}} \qquad (12)$$

This simple transformation produces an approximate Hessian with the correct curvature for a transition state search, and that guarantees that the eigenvector with a negative eigenvalue will have a large overlap with Δb.

Divided RFO

The shifted Hessian computed with the pRFO shift matrix (6) guarantees that the computed step will be downhill in the minimization space and uphill along the chosen transition vec-

tor. This is a consequence of the fact that the scale factor in the minimization space is more negative than the most negative eigenvalue in that space, and the scale factor along the transition vector is more positive than its corresponding eigenvalue. When the Hessian information is accurate, the step size computed by this approach tends to be reasonable even when the minimization space is not positive definite. However, during an optimization that utilizes Hessian updating, and particularly when the initial Hessian is approximate, the sign of some of the eigenvalues of the Hessian may be incorrect. In this case, the pRFO method produces a very ill-conditioned shifted Hessian that results in very large steps that are primarily along the eigenvectors with incorrect curvature. Even when the step size is scaled back to a reasonable limit prior to taking the step, it may take the optimizer a number of steps to correct the Hessian sufficiently and continue to make good progress toward the solution. Additionally, it is possible that updates to the Hessian will reduce the overlap of the transition vector with Δb, or increase the overlap of other vectors with Δb, complicating the selection of the transition vector in subsequent steps.

To increase the stability of the transition state optimization, the following divided RFO (dRFO) approach is proposed. First, the eigenvectors of the Hessian are compared against Δb, and any with an overlap of greater than 0.5 are considered to be part of the transition vector space. The augmented Hessian is constructed in each space, and λ_1 is defined as the second most negative eigenvalue of the augmented Hessian in the transition vector space, while λ_2 is defined as the most negative eigenvalue of the minimization space. In the case where the transition vector space has only one vector, this is identical to the pRFO approach. In addition, the augmented Hessian is constructed and diagonalized in the full space, and the second most negative eigenvalue is used as the shift factor λ_3 that is applied to the entire space

$$\mathbf{S}_{cpRFO} = -(\lambda_1 + \lambda_3) \sum_i^{n_{TS}} \mathbf{v}_i \mathbf{v}_i^T - (\lambda_2 + \lambda_3) \sum_i^{n_{min}} \mathbf{v}_i \mathbf{v}_i^T \qquad (13)$$

Just as in the case of the pRFO method, all three of the scale factors in the dRFO method tend toward zero as the magnitude of the gradient is reduced, but the additional correction makes it less likely that the shifted Hessian will be ill-conditioned. This results in an optimization that favors steepest descent in the minimization space over a step along a vector with the wrong curvature when the gradient in the minimization space is still large. A second consequence of the additional correction is that motion in the transition vector space tends to be favored, helping the optimizer to locate the ridge corresponding to the transition vector before taking large steps downhill, which has been shown[39] to increase the stability of transition state optimizations in the past.

Implementation and Discussion of Results

The interpolation and optimization methods described in the previous sections will be referred to collectively as the

Journal of COMPUTATIONAL CHEMISTRY

Table 3. Number of surface evaluations required to converge at PM6 level of theory.

| | QST2 Hessian? | | | CTS (UFF interpolation) Hessian? | | | | |
| | | | | Minimization | | Total[a] | | |
Reaction	Always	Once	Never	Aways	Never	Always	Once	Never
C$_2$N$_2$O	9	13	17	5	5	13	10	20
C$_5$HT	10	16	15	3	9	11	34	48
HCN	Fail[b]	Fail[b]	Fail[b]	4	5	7	8	9
Cope	Fail[c]	Fail[c]	Fail[c]	14	20	19	30	36
CPHT	8	10	16	5	7	7	9	10
Cyc-but	6	10	9	4	5	9	17	21
DACP2	Fail[d]	Fail[d]	38	5	6	11	12	27
DACP + eth	11	32	24	5	6	7	13	18
DFCP	Fail[c]	Fail[c]	Fail[c]	4	5	7	10	11
Ene	17	35	35	7	9	11	21	30
Grignard	Fail[c]	Fail[c]	Fail[c]	6	10	14	35	78
H$_2$ + CO	10	14	15	7	5	18	23	15
HF + eth	12	19	17	4	5	10	14	17
Hydro	Fail[c]	Fail[c]	Fail[c]	4	5	15	51	42
MeOH	15[e]	20[e]	27[e]	3	5	8	14	14
Oxirane	Fail[c]	Fail[c]	Fail[c]	8	15	19	71	73
Oxycope	Fail[c]	Fail[c]	Fail[c]	8	11	12	17	25
Silane	Fail[f]	Fail[f]	Fail[f]	5	6	11	20	18
S$_N$2	6	7	9	2	3	7	9	9
Sulfolene	14	30	25	11	7	14	23	24

[a] The Min columns contains only the number of PES evaluations required to converge the minimization phase, while the Total columns contain the PES evaluations required for both the minimization phase and the transition state optimization. [b] Convergence of the internal coordinate gradient was achieved on a structure with a nonzero Cartesian gradient, indicating an internal coordinate definition error. [c] Linear interpolation failed to produce an initial geometry. [d] Exceeded the default maximum number of optimizaiton steps, which was set as the larger of the number of coordinates being optimized or 100. [e] Optimization converged to the transition state for the decomposition of methanidyloxidanium (CH$_2$OH$_2$) to formaldahyde and H$_2$. [f] Optimizer exit with an error due to very small Hessian eigenvalue.

Connectivity Transition State (CTS) method for the following discussion. The CTS method was implemented in the development version of Gaussian 09.[40] Reactant and product structures for the 20 reactions in Figure 1 were optimized on the PM6 and B3LYP[41–44]/6–31G(d,p) PES. The results of these optimizations were then used as input for the CTS interpolation prior to the transition state search. Table 3 summarizes the performance of the CTS method using UFF during interpolation and optimizing to the transition state on the PM6 PES, while Table 4 summarizes the performance of the CTS method using UFF or PM6 during interpolation and optimizing to the transition state on the B3LYP/6-31G(d,p) PES. In each case, the CTS results are compared with the two input Quadratic Synchronous Transit (QST2) method and are measured by the total number of optimization steps required for convergence.

The PM6 results in Table 3 are shown with analytical Hessians computed at every step (Always), an analytical Hessian computed once at the beginning of the optimization (Once), and estimated force constants without computing an analytical Hessian (Never). In the latter two cases, the Hessian approximation is updated at each step using the combined SR1/PSB method proposed by Bofill.[15] The number of steps required for convergence of only the initial minimization phase of the CTS method is also reported separately. For the CTS method, it was determined that computing the Hessian at the beginning of the transition state optimization performed better than computing it at the beginning of the initial minimization phase. Hence, the Once and Never results for CTS both use estimated Hessians for the minimization phase.

The most dramatic difference between the two algorithms is in the interpolation phase. As mentioned previously, the linear interpolation in redundant internal coordinates used by the QST2 method fails to produce a structure in 6 of the 20 reactions using the PM6 minima structures, while the CTS interpolation succeeds for all 20 reactions. An additional four QST2 optimizations fail to converge to the correct transition state: DACP2 and Silane follow the incorrect eigenvector and wander away from the region containing the transition state, HCN incorrectly converges to a structure with nonzero Cartesian forces due to an error in defining the internal coordinates, and MeOH converges to the transition state for a different reaction. The QST2 method is usually used with reactant-like and product-like structures which are chosen to more closely

Table 4. Number of surface evaluations required to converge at B3LYP/6–31G(d,p) level of theory.

| | QST2 Hessian? | | CTS | | | |
| | | | PM6 interpolation Hessian? | | UFF interpolation Hessian? | |
Reaction	Always	Never	Aways	Never	Always	Never
C$_2$N$_2$O	Fail[a]	Fail[a]	8	9	11	13
C$_5$HT	11	14	6	7	11	16
HCN	Fail[b]	Fail[b]	6	7	9	10
Cope	Fail[a]	Fail[a]	7[c]	9[c]	30	37
CPHT	Fail[d]	Fail[d]	8	9	11	12
Cyc-but	9	9	16	17	17	18
DACP2	45	30	18	31	19	26
DACP + eth	Fail[a]	Fail[a]	8	10	8	12
DFCP	13	14	31	47	19	19
Ene	Fail[a]	Fail[a]	9	16	15	16
Grignard	Fail[a]	Fail[a]	50	51	30	57
H$_2$ + CO	Fail[d]	28	Fail[d]	18	37	26
HF + eth	Fail[a]	Fail[a]	9	12	12	15
Hydro	Fail[a]	Fail[a]	14	24	15	25
MeOH	Fail[a]	Fail[a]	19	20	12	14
Oxirane	Fail[a]	Fail[a]	20	41	39	64
Oxycope	Fail[a]	Fail[a]	11	18	19	25
Silane	24	23	Fail[d]	Fail[d]	Fail[d]	Fail[d]
S$_N$2[e]	25	25	18	30	21	21
Sulfolene	26	27	8	20	32	40

[a] Linear interpolation failed to produce an initial geometry. [b] Convergence of the internal coordinate gradient was achieved on a structure with a nonzero Cartesian gradient, indicating an internal coordinate definition error. [c] Converged to a transition state with a boat configuration, approximately 1.1 mH higher in energy than the chair configuration. [d] Failed to converge within the maximum number of iterations, which was set as the larger of the number of coordinates being optimized or 100. [e] The transition state for methyl chloride + fluoride does not exist at the B3LYP/6-31G(d,p) level of theory. The tert-butyl chloride + fluoride reaction, wihch is energetically similar to the methyl chloride reaction, was used instead.

resemble the expected geometry at the transition state than the fully optimized reactants and products. This could account for some of the difficulties that the method has with the interpolation and optimization of the reactions in the test set. The CTS method manages to converge to the correct transition state for each reaction at the PM6 level theory when starting from optimized reactant and product structures, while performing similarly to QST2 in most cases where both methods manage to converge.

In a number of cases, part or all of the geometry in the CTS guesses using the UFF energy for interpolation had higher symmetry than what is observed in the transition state. For example, the initial CTS structure for 1,3-pentadiene hydrogen transfer (C$_5$HT), has C$_{2V}$ while the transition state only has C$_S$ symmetry. The initial CTS structure for the Cope rearrangement is a nearly planar ring twisted by ca 35$°$ and has D$_2$ symmetry, while the transition state has a chair-like geometry with C$_{2h}$ symmetry. Other reactions with symmetries in the initial CTS structures that are not present in the final geometry include the ring closing of 1,3-butadiene (Cyc-But), the ene reaction of propene and ethylene, and the addition of sulfur dioxide to 1,3-butadiene (Sulfolene). The optimizations used to generate the initial CTS structures would seem to favor high symmetry. As the gradient in the direction of the atomic motion required to break symmetry is zero, it can be difficult for an optimizer to step away from the symmetric geometry and progress toward the transition state geometry, especially when exact Hessian information is not used. This can also be a problem in cases where the initial CTS structure does not necessarily have a higher degree of symmetry than the transition state, but does possess incorrect symmetry that must be broken. This is observed in the CTS structure for the Cope reaction as well as the decomposition of formaldehyde to H$_2$ to carbon monoxide, which had the initial position of the H$_2$ rotated by 90$°$ relative to the carbon monoxide, compared with the entirely planar transition structure geometry.

For the majority of the reactions in the test set, the PM6 transition state and the B3LYP/6-31G(d,p) transition state are quite similar, with the exception of the S$_N$2 reaction which does not have a barrier on the DFT PES. The S$_N$2 results in Table 4 are for the energetically similar reaction of *tert*-butyl chloride with a fluoride ion. While the transition states were generally similar, the minimum energy structures often differed significantly, particularly in the distance or relative position between bimolecular reactants, with the H$_2$ + CO and Silane reactions showing the largest difference. The relative orientation of the bimolecular reactant structures for both of these reactions at the DFT level provides a much greater challenge for the interpolation than the corresponding minima on the PM6 surface. Additionally, some reactions where linear interpolation of the PM6 minima succeeded, failed using the DFT minima (C$_2$N$_2$O, DACP + eth and Ene), while some reactions where linear interpolation of the PM6 minima failed succeed using the DFT minima (DFCP, HF + eth). In most cases, the CTS interpolation using PM6 as the low level energy resulted in a structure that was as good as or better for optimization of the DFT transition state than when UFF was used, with DFCP as the

only reaction significantly favoring UFF interpolated structure. The PM6 interpolation for the Cope structure does converge more quickly than the UFF interpolation, but it converges to a higher energy boat configuration. The symmetry issues discussed above are not seen during the PM6 interpolation.

Summary

The QST2 method implemented in the Gaussian suite of programs is a powerful tool for locating transition state geometries from known reactant and product configurations. The methods described in this article seek to improve on the stability and efficiency of locating the transition state by using bonding information to improve the construction of a redundant internal coordinate set to be used for the optimization, and to remove the explicit dependence on the reactant and product geometries from the estimation of the initial transition state geometry and from the computation of subsequent steps throughout the optimization. Whether exact Hessians are computed or not, the CTS method locates transition states with an efficiency comparable with the QST2 method on reactions where both are able to locate the correct transition state, but also succeeds in optimizing most transition states where the QST2 method fails.

Acknowledgment

[1] D. Wales, Energy Landscapes: Applications to Clusters, Biomolecules and Glasses; Cambridge University Press, Cambridge, **2003**.
[2] H. B. Schlegel, *WIREs Comput. Mol. Sci.* **2011**, *1*, 790.
[3] R. Olsen, G. Kroes, G. Henkelman, A. Arnaldsson, H. Jónsson, *J. Chem. Phys.* **2004**, *121*, 9776.
[4] S. M. Sharada, A. T. Bell, M. Head-Gordon, *J. Chem. Phys.* **2014**, *140*, 164115.
[5] Y. Zeng, P. Xiao, G. Henkelman, *J. Chem. Phys.* **2014**, *140*, 044115.
[6] H. B. Schlegel, *Theor. Chim. Acta.* **1984**, *66*, 333.
[7] T. H. Fischer, J. Almlof, *J. Phys. Chem.* **1992**, *96*, 9768.
[8] J. M. Wittbrodt, H. B. Schlegel, *J. Mol. Struct.: Theochem* **1997**, *398*, 55.
[9] B. A. Murtagh, *Comput. J.* **1970**, *13*, 185.
[10] C. G. Broyden, *Math. Comput.* **1970**, *24*, 365.
[11] R. Fletcher, *Comput. J.* **1970**, *13*, 317.
[12] D. Goldfarb, *Math. Comput.* **1970**, *24*, 23.
[13] D. F. Shanno, *Math. Comput.* **1970**, *24*, 647.
[14] M. J. D. Powell, *Math. Prog.* **1971**, *1*, 26.
[15] J. M. Bofill, *J. Comput. Chem.* **1994**, *15*, 1.
[16] A. Banerjee, N. Adams, J. Simons, R. Shepard, *J. Phys. Chem.* **1985**, *89*, 52.
[17] J. M. Bofill, J. M. Anglada, *Theor. Chem. Acc.* **2001**, *105*, 463.

[18] S. K. Burger, P. W. Ayers, *J. Chem. Theory Comput.* **2010**, *6*, 1490.

[19] S. K. Burger, P. W. Ayers, *J. Chem. Phys.* **2010**, *132*, 234110.

[20] T. A. Halgren, W. N. Lipscomb, *Chem. Phys. Lett.* **1977**, *49*, 225.

[21] C. Peng, H. B. Schlegel, *Isr. J. Chem.* **1993**, *33*, 449.

[22] G. Henkelman, H. Jónsson, *J. Chem. Phys.* **1999**, *111*, 7010.

[23] R. Elber, M. Karplus, *Chem. Phys. Lett.* **1987**, *139*, 375.

[24] S. K. Burger, W. Yang, *J. Chem. Phys.* **2006**, *124*, 054109.

[25] D. Sheppard, R. Terrell, G. Henkelman, *J. Chem. Phys.* **2008**, *128*, 134106.

[26] A. Behn, P. M. Zimmerman, A. T. Bell, M. Head-Gordon, *J. Chem. Phys.* **2011**, *135*, 224108.

[27] P. Plessow, *J. Chem. Theory Comput.* **2013**, *9*, 1305.

[28] S. M. Sharada, P. M. Zimmerman, A. T. Bell, M. Head-Gordon, *J. Chem. Theory Comput.* **2012**, *8*, 5166.

[29] P. Zimmerman, *J. Chem. Theory Comput.* **2013**, *9*, 3043.

[30] A. B. Birkholz, H. B. Schlegel, *Theor. Chem. Acc.* **2012**, *131*, 1170.

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

[32] J. Baker, F. Chan, *J. Comput. Chem.* **1996**, *17*, 888.

[33] D. H. Ess, *J. Chem. Educ.* **2012**, *89*, 817.

[34] P. Pulay, G. Fogarasi, *J. Chem. Phys.* **1992**, *96*, 2856.

[35] L. Pauling, *J. Am. Chem. Soc.* **1947**, *69*, 542.

[36] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard, III, W. M. Skiff, *J. Am. Chem. Soc.* **1992**, *114*, 10024.

[37] J. J. Stewart, *J. Mol. Model.* **2007**, *13*, 1173.

[38] S. Bell, J. S. Crighton, *J. Chem. Phys.* **1984**, *80*, 2464.

[39] J. W. May, J. D. Lehner, M. J. Frisch, X. Li, *J. Chem. Theory Comput.* **2012**, *8*, 5175.

[40] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, B. G. Janesko, F. Lipparini, G. Zheng, J. L. Sonnenberg, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, P. V. Parandekar, N. J. Mayhall, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian Development Version, Rev. H.35 Gaussian Inc.: Wallingford, CT, **2010**.

[41] S. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200.

[42] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[43] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.

[44] P. Stephens, F. Devlin, C. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623.