Molecular Physics, Vol. 105, Nos. 19–22, 10 October–20 November 2007, 2719–2729

Taylor & Francis
Taylor & Francis Group

Empirical valence bond models for reactive potential energy surfaces. II. Intramolecular proton transfer in pyridone and the Claisen reaction of allyl vinyl ether

JASON L. SONNENBERG and H. BERNHARD SCHLEGEL*
Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

(*Received 25 May 2007; in final form 6 August 2007*)

Empirical valence bond (EVB) surfaces have been constructed for 2-pyridone-2-hydroxypyridine proton transfer and for the Claisen rearrangement of allyl vinyl ether at the MP2/6-311 + G(d,p) level of theory. A distributed Gaussian approach is used to approximate the interaction matrix elements. Parameters for the distributed Gaussians are determined by fitting to energy, gradient and Hessian data obtained from *ab initio* electronic structure calculations at one to nine points along the reaction path. An efficient DIIS (direct inversion of iterative subspace) method is used to solve the fitting equations. Criteria for choosing internal coordinates for the representation of the potential energy surfaces and for the interaction matrix element are discussed. Practical techniques for determining the placement and exponents of the Gaussians are described. With one set of s-, p- and d-type Gaussians at the transition state, the error in the energy along the reaction path is less than $10\,\text{kJ mol}^{-1}$ for pyridone tautomerization. Five sets of Gaussians reduces the error to less than $5\,\text{kJ mol}^{-1}$ and seven Gaussians drops the error below $1\,\text{kJ mol}^{-1}$. The Claisen rearrangement is more challenging and requires seven Gaussians to achieve an error of less than $4\,\text{kJ mol}^{-1}$ for energies along the reaction path.

Keywords: Potential energy surfaces; Empirical valence bond; EVB; DIIS; GMRES

1. Introduction

A reactive potential energy surface, V, can be modeled by an empirical valence bond (EVB) approach, in which the reactant and product valence bond configurations, ψ_1 and ψ_2, interact via an empirical Hamiltonian, $\mathbf{H}$,

$$\Psi = c_1\psi_1 + c_2\psi_2, \tag{1}$$

$$\mathbf{H} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \tag{2}$$

$$V_{11} = \langle\psi_1|\hat{H}|\psi_1\rangle, \quad V_{12} = V_{21} = \langle\psi_1|\hat{H}|\psi_2\rangle, \\ V_{22} = \langle\psi_2|\hat{H}|\psi_2\rangle, \tag{3}$$

$$V = \frac{1}{2}(V_{11} + V_{22}) - \sqrt{\left[\left(\frac{1}{2}\right)(V_{11} - V_{22})\right]^2 + V_{12}^2}. \tag{4}$$

Good approximations to V_{11} and V_{22} are available from valence force fields such as those used in molecular mechanics methods. The interaction matrix element or resonance integral, V_{12}, is obtainable by fitting to suitable experimental data or electronic structure calculations.

Warshel and Weiss originally formulated the EVB approach to treat reactions in solvated systems [1], and used simple V_{12}^2 approximations to fit barrier heights. Chang and Miller [2] modeled V_{12}^2 using a generalized Gaussian with parameters chosen to fit the energy, geometry and vibrational frequencies of the transition state obtained from electronic structure calculations. Truhlar and co-workers [3–5] used distance weighted interpolants to obtain a V_{12}^2 approximation from electronic structure calculations at several points along the reaction path. Recently, we introduced a method in which V_{12}^2 is represented by a Gaussian times a polynomial at one or more points on the potential energy surface [6].

---

*Corresponding author. Email: hbs@chem.wayne.edu

Molecular Physics
ISSN 0026–8976 print/ISSN 1362–3028 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00268970701622277

In the Chang–Miller approach [2], the generalized Gaussian used to approximate the resonance integral has the form

$$V_{12}(\mathbf{q})^2 = A\exp\left[\mathbf{B}^{\mathrm{T}}\Delta\mathbf{q} - \left(\frac{1}{2}\right)\Delta\mathbf{q}^{\mathrm{T}}\mathbf{C}\Delta\mathbf{q}\right], \quad \Delta\mathbf{q} = \mathbf{q} - \mathbf{q}_{\mathrm{TS}}, \tag{5}$$

where $\mathbf{q}_{\mathrm{TS}}$ is the transition state (TS) geometry, and A, $\mathbf{B}$ (a vector), and $\mathbf{C}$ (a matrix) are parameters [2]

$$A = \left[V_{11}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right]\left[V_{22}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right], \tag{6a}$$

$$\mathbf{B} = \frac{\mathbf{D}_1}{\left[V_{11}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right]} + \frac{\mathbf{D}_2}{\left[V_{22}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right]},$$
$$\mathbf{D}_n = \left.\frac{\partial V_{nn}(\mathbf{q})}{\partial\mathbf{q}}\right|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}} - \left.\frac{\partial V(\mathbf{q})}{\partial\mathbf{q}}\right|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}}, \tag{6b}$$

$$\mathbf{C} = \frac{\mathbf{D}_1\mathbf{D}_1^{\mathrm{T}}}{\left[V_{11}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right]^2} + \frac{\mathbf{D}_2\mathbf{D}_2^{\mathrm{T}}}{\left[V_{22}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})\right]^2} - \frac{\mathbf{K}_1}{V_{11}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})} - \frac{\mathbf{K}_2}{V_{22}(\mathbf{q}_{\mathrm{TS}}) - V(\mathbf{q}_{\mathrm{TS}})},$$
$$\mathbf{K}_n = \left.\frac{\partial^2 V_{nn}(\mathbf{q})}{\partial\mathbf{q}^2}\right|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}} - \left.\frac{\partial^2 V(\mathbf{q})}{\partial\mathbf{q}^2}\right|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}}. \tag{6c}$$

The exponents of the generalized Gaussian are chosen to reproduce the energy, gradient and Hessian of the transition state ($V(\mathbf{q}_{\mathrm{TS}})$, $\partial V(q)/\partial q|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}}$ and $\partial^2 V(q)/\partial q^2|_{\mathbf{q}=\mathbf{q}_{\mathrm{TS}}}$, respectively). Unfortunately, the Chang–Miller approach runs into some difficulties when $\mathbf{C}$ has one or more negative eigenvalues [4, 7, 8].

Our first method [6] recasts $V_{12}(\mathbf{q})^2$ as a polynomial in $\triangle\mathbf{q} = \mathbf{q} - \mathbf{q}_{\mathrm{TS}}$ times a spherical Gaussian at the transition state

$$V_{12}(\mathbf{q})^2 = A\left[1 + \mathbf{B}^{\mathrm{T}}\Delta\mathbf{q} + \left(\frac{1}{2}\right)\Delta\mathbf{q}^{\mathrm{T}}(\mathbf{C} + \alpha\mathbf{I})\Delta\mathbf{q}\right] \times \exp\left[-\left(\frac{1}{2}\right)\alpha|\Delta\mathbf{q}|^2\right]. \tag{7}$$

The Gaussian times a polynomial (GP) approximation to V_{12}^2 can be generalized by employing a linear combination of s-, p-, and d-type Gaussians at a number of points on the potential energy surface (PES).

$$g(\mathbf{q}, \mathbf{q}_K, 0, 0, \alpha_K) = \exp\left[-\left(\frac{1}{2}\right)\alpha_K|\mathbf{q} - \mathbf{q}_K|^2\right],$$
$$g(\mathbf{q}, \mathbf{q}_K, i, 0, \alpha_K) = (\mathbf{q} - \mathbf{q}_K)_i\exp\left[-\left(\frac{1}{2}\right)\alpha_K|\mathbf{q} - \mathbf{q}_K|^2\right],$$
$$g(\mathbf{q}, \mathbf{q}_K, i, j, \alpha_K) = (\mathbf{q} - \mathbf{q}_K)_i(\mathbf{q} - \mathbf{q}_K)_j \times \exp\left[-\left(\frac{1}{2}\right)\alpha_K|\mathbf{q} - \mathbf{q}_K|^2\right]. \tag{8}$$

Summing up the contributions from Gaussians at each distributed data point, $\mathbf{q}_K$, yields our distributed Gaussian (DG) approximation to the resonance integral

$$V_{12}(\mathbf{q})^2 = \sum_K \sum_{i\geq j\geq 0}^{NDim} B_{ijK}\, g(\mathbf{q}, \mathbf{q}_K, i, j, \alpha_K). \tag{9}$$

Since the coefficients in equation (9) are linear, $V_{12}(\mathbf{q})^2$ can be fit more readily than using a linear combination of generalized Gaussians with nonlinear coefficients. The number of Gaussians forming a basis at each $\mathbf{q}_K$ is chosen to be equal to the number of energy, first derivative, and second derivative values at $\mathbf{q}_K$. Fitting to the energy and derivative data results in a set of linear equations

$$\mathbf{DB} = \mathbf{F}, \tag{10}$$

where $\mathbf{D}$ is an unsymmetric matrix containing the values of $g(\mathbf{q}_L, \mathbf{q}_K, i, j, \alpha_K)$, $\partial g(\mathbf{q}_L, \mathbf{q}_K, i, j, \alpha_K)/\partial\mathbf{q}|_{\mathbf{q}=\mathbf{q}_L}$, and $\partial^2 g(\mathbf{q}_L, \mathbf{q}_K, i, j, \alpha_K)/\partial\mathbf{q}^2|_{\mathbf{q}=\mathbf{q}_L}$, and $\mathbf{F}$ a column vector containing the corresponding values of $V_{12}^2(\mathbf{q}_L)$, $\partial V_{12}^2(\mathbf{q})/\partial\mathbf{q}_{\mathbf{q}=\mathbf{q}_L}$, and $\partial^2 V_{12}^2(\mathbf{q})/\partial\mathbf{q}^2|_{\mathbf{q}=\mathbf{q}_L}$. In our initial tests, we employed singular value decomposition (SVD) to solve these linear equations. Both the GP and the DG methods were shown to accurately reproduce one- and two-dimensional surfaces and small molecule PESs such as HCN [6]. However, even for simple molecular systems such as 2-pyridone with only five data points the dimensionality of the system quickly exceeds 8000 and thus matrix inversion or SVD are no longer viable linear equation solvers.

In this work, SVD is replaced by the direct inversion of iterative subspace (DIIS) algorithm, which is a type of Krylov-subspace method, to solve equation (10). Pulay and others have shown that the DIIS approach is very effective in solving the equations arising in geometry optimization, SCF convergence, couple-perturbed Hartree–Fock and related problems [9–15]. To facilitate the solution of the fitting parameter linear equations, the s-type Gaussians in equation (8) are replaced by

$$g(\mathbf{q}, \mathbf{q}_K, 0, 0, \alpha_K) = \left(1 + \left(\frac{1}{2}\right)\alpha_K|\mathbf{q} - \mathbf{q}_K|^2\right) \times \exp\left[-\left(\frac{1}{2}\right)\alpha_K|\mathbf{q} - \mathbf{q}_K|^2\right]. \tag{11}$$

This yields identity matrices for the diagonal blocks of the fitting matrix and improves the convergence of the DIIS method.

In the present paper we examine the potential energy surfaces for the intramolecular proton transfer in 2-pyridone-2-hydroxypyridine and for the Claisen rearrangement in allyl vinyl ether. First, some technical aspects of the DIIS solution of the EVB fitting equations are considered. Then, criteria for selecting internal coordinates the potential energy surface and for the approximations to V_{12}^2 are outlined. Finally, practical techniques for choosing α values and placement of the Gaussians are discussed.

2. Methods

Originally developed in the 1950s, Krylov subspace algorithms have become increasingly popular in solving very large systems of linear equations of the form $\mathbf{Ax} = \mathbf{b}$ [16]. The distinguishing feature between different Krylov subspace algorithms is how each method iteratively forms the approximation, $\mathbf{x}^{(k)}$,

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + q_{k-1}(\mathbf{A})\mathbf{r}^{(0)}, \quad \mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}, \tag{12}$$

to $\mathbf{x}$ in the space W_k

$$\begin{aligned} W_k &= \{\mathbf{v} = \mathbf{x}^{(0)} + \mathbf{y}, \mathbf{y} \in K_k(\mathbf{A}; \mathbf{r}^{(0)})\}, \\ K_k(\mathbf{A}; \mathbf{r}^{(0)}) &= \{\mathbf{r}^{(0)}, \mathbf{A}\mathbf{r}^{(0)}, \ldots, \mathbf{A}^{k-1}\mathbf{r}^{(0)}\}, \end{aligned} \tag{13}$$

where $\mathbf{x}^{(0)}$ is the initial guess for $\mathbf{x}$, $q_{k-1}(\mathbf{A})$ a polynomial in $\mathbf{A}$ of degree $k-1$, $\mathbf{r}^{(0)}$ the initial residual vector, and K_k a Krylov subspace of kth order [17]. Two strategies exist for generating $\mathbf{x}^{(k)}$: choose $\mathbf{x}^{(k)}$ such that the residual $\mathbf{r}^{(k)}$ is orthogonal to any vector in $K_k(\mathbf{A}; \mathbf{r}^{(0)})$ or compute $\mathbf{x}^{(k)}$ so as to minimize the Euclidean norm of the residual, $||\mathbf{r}^{(k)}||$. Orthogonalizing the residual against the Krylov subspace results in the full orthogonalization method (FOM) and its variants, while minimizing the norm of the residual generates minimal residual methods.

Inspired by Pople *et al.*'s use of the Arnoldi process [18] (aka FOM) to solve for wave function derivative coefficients [19], Pulay formulated his direct inversion of iterative subspace (DIIS) to accelerate SCF convergence [20, 21]. DIIS is a Krylov subspace method aiming to minimize the norm of the residual subject to the constraint that the expansion coefficients, c_i, sum to one

$$\mathbf{r}^{(k)} = \sum_{i=1}^{k-1} c_i \mathbf{e}^{(i)}, \quad \sum_{i=1}^{k-1} c_i = 1. \tag{14}$$

Ionova and Carter showed that the definition of the kth error tensor, $\mathbf{e}^{(k)}$, varies depending upon the application, but always approximates one of the forms in equation (15) as a result of Banach's principle [22]

$$\begin{aligned} &\mathbf{e}_1^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \quad \mathbf{e}_2^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}, \\ &\mathbf{e}_3^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}. \end{aligned} \tag{15}$$

After orthogonalizing the error vectors, the expansion coefficients are determined by solving

$$\begin{pmatrix} h_{1,1} & \cdots & h_{1,k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ h_{k,1} & \cdots & h_{k,k} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_k \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad h_{i,j} = \langle \mathbf{e}^{(i)} | \mathbf{e}^{(j)} \rangle, \tag{16}$$

where $h_{i,j}$ is an appropriate inner product in the subspace and λ is a Lagrangian multiplier. With the expansion coefficients in hand, a new $\mathbf{x}^{(k)}$ can be formed and the cycle repeated

$$\mathbf{x}^{(k)} = \sum_{i=1}^{k-1} c_i \mathbf{x}^{(i)}. \tag{17}$$

Conceptually, DIIS is the forefather of modern minimal residual methods [17], but the true beauty of DIIS lies in the generality of the algorithm: $\mathbf{e}^{(k)}$ is not restricted to a particular exact form and $\mathbf{A}$ can be symmetric or unsymmetric, definite or indefinite. Such flexibility facilitates easy coupling with other numeric techniques to form the set of preeminent iterative equation solvers for electronic structure methods [9–15]. With the current work, Pulay's inspiration comes full circle as we employ DIIS in constructing a chemically accurate, EVB PES for a variety of reactions.

Electronic-structure calculations were computed at MP2/6-311 + G(d,p) [23–25] with the development version of the Gaussian suite [26]. All stationary points were optimized employing the Berny algorithm [27, 28] while the intrinsic reaction path was explored via the damped velocity-Verlet [29] and Hessian-based predictor-corrector integrators of Hratchian and Schlegel [30, 31]. Transition states and data points along the IRC were tightly optimized until the maximum force was 2×10^{-6} a.u., whereas all other stationary points were optimized to 4.5×10^{-4} a.u. V_{11} and V_{22} are represented by quadratic expansions about the reactant and product minima using the geometry and Hessians obtained from the *ab initio* calculations. EVB fitting and analysis was carried out using Mathematica 5.2 [32].

To solve equation (10) with DIIS for a given set of α values, an initial guess, $\mathbf{x}_0$, for $\mathbf{B}$ is used to generate a normalized error vector via $\mathbf{e}^{(1)} = (\mathbf{F} - \mathbf{D}\mathbf{x}_0)/||\mathbf{F} - \mathbf{D}\mathbf{x}_0||$. The generalized minimum residual (GMRES) algorithm [33] utilizing the stabilized Gram–Schmidt procedure is then employed to iteratively form an orthonormal set of error vectors in K_m

$$\begin{aligned} h_{i,k} &= \langle \mathbf{e}^{(i)} | \mathbf{A}\mathbf{e}^{(k)} \rangle, \quad \text{for } i = 1 \ldots k, \quad k = 1 \ldots m, \\ \mathbf{e}^{(k+1)} &= \left(\mathbf{A}\mathbf{e}^{(k)} - \sum_{i=1}^{k} h_{i,k}\mathbf{e}^{(i)} \right) (h_{k+1,k})^{-1}, \\ h_{k+1,k} &= \left\| \mathbf{A}\mathbf{e}^{(k)} - \sum_{i=1}^{k} h_{i,k}\mathbf{e}^{(i)} \right\|. \end{aligned} \tag{18}$$

At each step $\mathbf{H}_k$ forms a $k+1 \times k$ Hessenberg matrix that is directly decomposed using standard QR decomposition to produce $\mathbf{Q}_k$ and $\mathbf{R}_k$ such that $\mathbf{Q}_k{}^{\mathrm{T}}\mathbf{R}_k = \mathbf{H}_k$. The residual at step k can then be approximated according to $\mathbf{Q}_k||\mathbf{B} - \mathbf{D}\mathbf{x}_0||\mathbf{v}$, where $\mathbf{v}$ is the first unit vector in $\mathbf{R}^{k+1}$. Once the residual is smaller than 1×10^{-9} the solution is computed in the following manner:

$$\begin{aligned} \mathbf{c} &= (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}(\mathbf{H}^{\mathrm{T}}||\mathbf{B} - \mathbf{D}\mathbf{x}_0||\mathbf{v}), \\ \mathbf{x} &= \mathbf{x}_0 + \mathbf{E}^{\mathrm{T}}\mathbf{c}, \end{aligned} \tag{19}$$

where $\mathbf{E}$ is the matrix formed from the $\mathbf{e}^{(k)}$ column vectors.

3. Results and discussion

3.1. *Choice of redundant internal coordinates*

Even though both PG and DG methods (equations (7) and (9)) can utilize any coordinate system, we employ redundant internal coordinates rather than Cartesian coordinates to eliminate standard orientation requirements for the resulting energy hypersurface. Bond lengths are in bohr and angles are in radians. Determining a suitable set of n redundant internal coordinates for the hypersurface is entirely analogous to choosing redundant internal coordinates for geometry optimizations with a few additional caveats.

(1) The intersection of coordinates for all $\mathbf{q}_K$ must be included in the universal set.
(2) Coordinates corresponding to the largest components of the transition vector must be included. For reactions involving bond breaking/forming, this will automatically include the partially formed bond coordinate(s) if the TS coordinate system was well chosen.
(3) Linear bend coordinates and valence angles passing through linear along the path from reactants to products are excluded. The range of dihedral angles is also controlled so that no abrupt jumps (e.g. $-179.9^\circ \rightarrow +180.0^\circ$) occur along the reaction path.

The union minus the intersection of coordinate systems for all $\mathbf{q}_K$ produces a set of coordinates that must be carefully examined for adherence to point 3 before addition to the universal set.

$V_{12}(\mathbf{q})^2$ usually does not depend upon the full set of redundant internal coordinates necessary for an accurate representation of the energy hypersurface. As such, fitting $V_{12}(\mathbf{q})^2$ in a subspace of the $\mathbf{R}^n$ internal coordinate vector space can offer large computational savings. An appropriate subspace for most reactions is easily determined from the difference vector between reactants and products in $\mathbf{R}^n$. In reactions where the reactants and products are nearly identical, the difference vectors between the transition state and either the reactants or products may be more appropriate. In most cases, any coordinate corresponding to a zero element of the difference vector can be removed. Experience has shown that other coordinates can be eliminated if the corresponding difference vector element has a value less than a given tolerance value. Tolerance values of 1×10^{-1}–10^{-3} provide reasonable results with less than $0.5\,\mathrm{kJ\,mol^{-1}}$ changes in the maximum energy error along the reaction path.

3.2. *Placement of the Gaussian centers, $\mathbf{q}_K$*

The Gaussian times a polynomial scheme only requires *ab initio* data at the TS, while the distributed Gaussian method can handle any number of data points. The simplest set of data points for the DG method is the reactant, product and TS since it is readily available from standard *ab initio* investigations. While a three point DG hypersurface for a well-behaved case is usually accurate to better than $10\,\mathrm{kJ\,mol^{-1}}$ for points along the reaction path, a tighter accuracy may be required for a particular application. It is important to note that accuracy along the reaction path does not imply the same accuracy in directions perpendicular to said path. If the surface will be used for molecular dynamics, where a uniform description of the hypersurface is essential, additional Gaussian centers should be added in pairs: one data point at roughly the same place on both sides of the TS. As seen in Truhlar's work [3, 4] points at 1/2 and 1/4 down both sides of the reaction path are also good locations for placement of DG data points. Stationary points occurring in the redundant internal gradient norm curve along the reaction coordinate provide points for placement of additional Gaussian centers.

3.3. *Optimization of α values for each $\mathbf{q}_K$*

Determining the optimal values of α at a particular $\mathbf{q}_K$ is a crucial part of both the GP and DG algorithms. Although both methods permit separate values of α for each Gaussian function in the basis, this work utilizes the same exponent for all basis functions at a particular data point. Initial values of α are set to the average redundant internal Hessian eigenvalue. If a solution to equation (10) is not found within 75 iterations or the maximum coefficient in **B** becomes too large, each α is set equal to the previous value times a multiplier. Values of the multiplier range from 1.1 to 2.0 depending upon the final accuracy required in the resulting EVB hypersurface.

3.4. *2-Pyridone-2-hydroxypyridine intramolecular proton transfer*

This section focuses on the methodological issues of fitting EVB surfaces using the GP and DG methods for the gas-phase interconversion of 2-pyridone to 2-hydroxypyridine via an intramolecular proton transfer (figure 1). Future work will explore the rates, isotope effects and molecular dynamics of the pyridone system along with the effect of additional water molecules and bulk solvent on the potential energy surface [34]. *Ab initio* energies, gradients and Hessians were computed for nine points along the 2-pyridone to 2-hydroxypyridine reaction path. In addition to the TS, reactant, and product, data was also computed at approximately 1/2, 1/4, and 1/8 of the way toward each minimum along the reaction path, calculated in mass-weighted Cartesian coordinates. The corresponding values along the reaction path, starting from the reactants with the TS at zero, were −3.1712, −1.6549, −0.8321, −0.4162, 0.0, 0.4164, 0.8332, 3.6134, and 5.1053 atomic units. To evaluate the quality of the fitted EVB hypersurfaces, IRC data was calculated at 117 steps along the reaction path. A set of 56 redundant internal coordinates was employed.

Six EVB fits were performed with the number of data points along the reaction path ranging from $K=1$ to $K=9$. The α values and error analysis are reported in tables 1 and 2, respectively. Based upon the redundant internal coordinate difference vector for the reactant and product, four different subspaces, $\mathbf{R}^n$, were examined for each value of K. The subspace with 32 coordinates, $\mathbf{R}^{32}$, reflects the elimination of all dihedral angles from $\mathbf{R}^{56}$, while $\mathbf{R}^n$ values of 31, 25, and 10

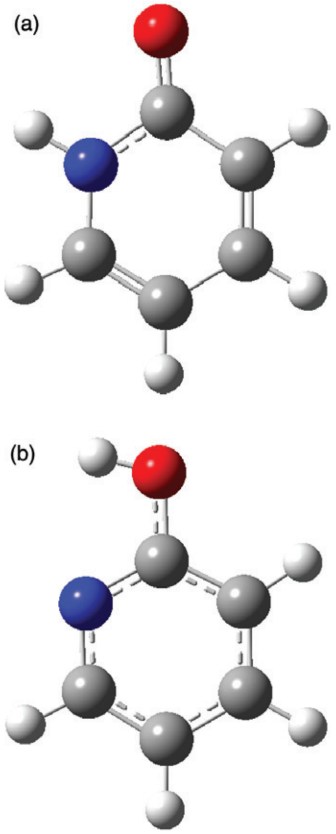

Figure 1. Optimized structures of (a) 2-pyridone and (b) 2-hydroxypyridine.

Table 1. Optimized α values for gas-phase pyridone tautomerism EVB fits.

K	TS	Reactant	Product	−1/2	1/2	−1/4	1/4	−1/8	1/8
1	7.1893								
3	1.9100	0.1290	0.1290						
5	2.4000	1.5000	1.2000	2.0000	2.0000				
7	9.3000	10.0000	10.000	10.100	9.9000	9.8000	9.8000		
9	17.5692	10.6879	10.6879	7.3205	7.3205	11.7128	11.7128	12.8841	12.8841

Table 2. Gas-phase pyridone tautomerism EVB analysis. Errors are determined from *ab initio* IRC data. Energies are in kJ mol^{-1}, V_{12}^2 is in a.u.^{2}, and angles are in degrees.

K	$\mathbf{R}^n$	Norm [**B**]	Max. **B** element	Max. energy error	Max. V_{12}^2 error	Max. grad. norm error	Mean angle between gradients
1	32	0.09738	0.06067	7.125	1.477×10^{-3}	2.946×10^{-2}	17.5
1	31	0.09737	0.06067	7.125	1.477×10^{-3}	2.947×10^{-2}	17.5
1	25	0.09728	0.06067	7.123	1.476×10^{-3}	2.950×10^{-2}	18.0
1	10	0.09425	0.06067	7.239	1.458×10^{-3}	2.965×10^{-2}	20.1
3	32	4.169	2.394	4.157	8.911×10^{-4}	4.864×10^{-3}	11.1
3	31	4.147	2.394	4.157	8.911×10^{-4}	4.865×10^{-3}	11.1
3	25	4.018	2.396	4.114	8.819×10^{-4}	4.894×10^{-3}	12.0
3	10	3.863	2.527	4.432	8.230×10^{-4}	2.965×10^{-2}	16.2
5	32	11.280	1.661	3.316	4.899×10^{-4}	6.549×10^{-3}	7.03
5	31	11.094	1.659	3.328	4.902×10^{-4}	6.556×10^{-3}	7.06
5	25	10.374	1.724	3.792	5.001×10^{-4}	6.749×10^{-3}	7.93
5	10	41.767	7.054	86.49	1.417×10^{-2}	9.309×10^{-2}	23.7
7	32	1.900	0.2773	0.4365	7.468×10^{-5}	1.137×10^{-2}	7.18
7	25	1.570	0.2613	0.4315	7.271×10^{-5}	1.161×10^{-2}	7.84
7	10	11.94	2.725	1.812	5.261×10^{-4}	2.965×10^{-2}	12.9
9	32	0.8095	0.1196	0.2645	7.751×10^{-5}	1.810×10^{-3}	1.97
9	25	0.6836	0.1129	0.2568	7.527×10^{-5}	4.436×10^{-3}	3.53
9	10	3.013	0.6279	1.262	3.719×10^{-4}	2.965×10^{-2}	11.3

correspond to difference vector tolerance values of 10^{-3}, 10^{-2}, and 10^{-1}, respectively. The PG method ($K = 1$) requires a rather tight α value in order to correct the large values of V_{11} and V_{22} near the TS. Six measures were employed to evaluate the goodness of the EVB fits: the norm of **B**, the magnitude of the largest coefficient in **B**, the maximum error in total energy, the maximum error in $V_{12}(\mathbf{q})^2$, the maximum difference in the gradient norm along the IRC, and the mean angle between the EVB internal coordinate gradients and the *ab inito* internal gradients along the IRC. The **B** norm and the maximum **B** coefficient become very large when the α values require further refinement or when the subspace employed in the $V_{12}(\mathbf{q})^2$ fit is too small. The total energy and $V_{12}(\mathbf{q})^2$ have large errors when the α values are poorly refined or when K should be increased. When the gradient norm error and the mean angle between predicted and actual gradients are sufficiently small, then the fit is satisfactory and no additional $\mathbf{q}_K$ are required.

The polynomial times a Gaussian yields a maximum error of ca. 7 kJ mol^{-1} along the IRC and is essentially constant for all subspaces (table 2, $K = 1$). Although the energy error is less than 10 kJ mol^{-1}, the errors in $V_{12}(\mathbf{q})^2$ and the angle between the predicted EVB internal gradient and the actual gradient indicate that a better fit may be desirable.

For $K = 3$ and 5, the maximum error in energy is reduced to less than 4.2 kJ mol^{-1}, thereby achieving chemical accuracy along the IRC. In both cases the error in $V_{12}(\mathbf{q})^2$ has decreased by an order of magnitude to 10^{-4} a.u. but the error in the gradient norm has increased by an order of magnitude to 10^{-3}. The DG fits were insensitive to the subspace employed for $V_{12}(\mathbf{q})^2$ except for $\mathbf{R}^{10}$ which was just too small. As such, $\mathbf{R}^{31}$ was not examined for $K > 5$, and figures 2–4 were generated from data in $\mathbf{R}^{25}$. Increasing K to seven drops the maximum error in the energy along the path to less than 0.5 kJ mol^{-1} and reduces the error in $V_{12}(\mathbf{q})^2$ by another order of magnitude to 10^{-5}. The potential energy surface as a function of the N–H and O–H distances is shown in figure 5. Although the energy and $V_{12}(\mathbf{q})^2$ are quite accurate for $K = 7$, the gradient norm error is only slightly better than the PG method generated. To improve the gradient norm fit, two more data points lying very near to the TS were necessary. It should be noted that, for $K > 3$, all additional points along the reaction path correspond to stationary points or inflection points in the gradient norm curve, suggesting that the gradient norm curve can provide information on where to place additional Gaussian centers to improve the fit.

As the number of data points used in the DG EVB fit increases for the pyridone system, the α value for a particular $\mathbf{q}_K$ always increases. This increase in α values indicates that the Gaussian basis set at $\mathbf{q}_K$ is becoming more localized and therefore less able to correct errors in the hypersurface fit far away from $\mathbf{q}_K$.

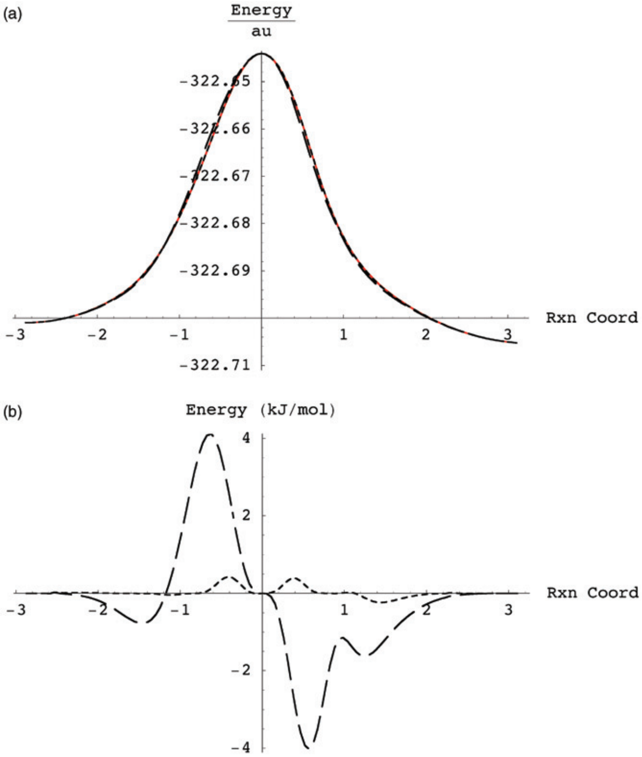

Figure 2. EVB energy along the reaction path for 2-pyridone↔2-hydroxypyridine with $K=3$ (long dash) and $K=7$ (short dash). The two surfaces are plotted along with IRC energies (red) in (a), while differences between the IRC and predicted energies are in part (b).

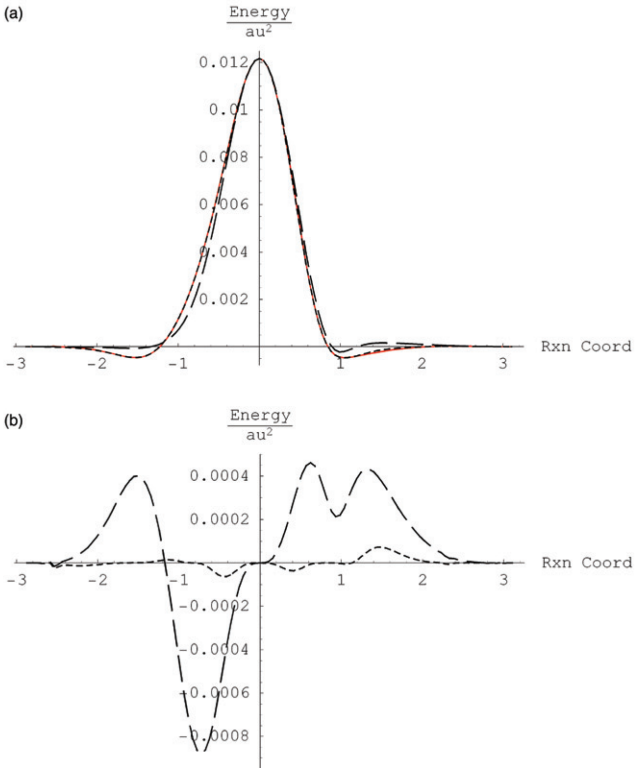

Figure 3. EVB $V_{12}(\mathbf{q})^2$ energy along the reaction path for 2-pyridone↔2-hydroxypyridine with $K=3$ (long dash) and $K=7$ (short dash). The two surfaces are plotted along with IRC $V_{12}(\mathbf{q})^2$ energies (red) in (a), while differences between the IRC and predicted $V_{12}(\mathbf{q})^2$ energies are in part (b).

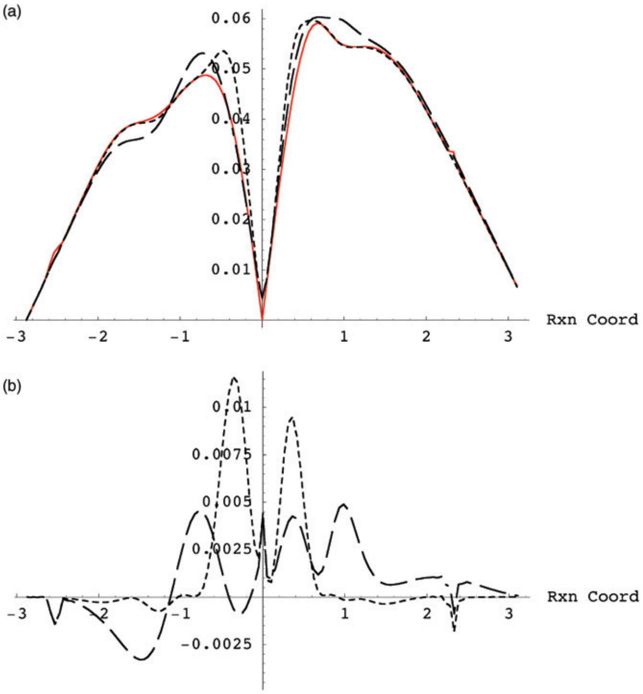

Figure 4. Norm of the EVB redundant internal gradient along the reaction path for 2-pyridone↔2-hydroxypyridine with $K=3$ (long dash) and $K=7$ (short dash). The two surfaces are plotted along with the norm of the IRC redundant internal gradient (red) in (a), while differences between the IRC and predicted norms are in part (b).

The reduced ability of each basis set to correct long-range errors in the hypersurface is not problematic so long as K is large enough. For the pyridone system this occurs at K values of 7 and 9 where the individual α values are converging to an average value. Convergence to average α values suggests that a single α may be appropriate at all $\mathbf{q}_K$ as K becomes large. Starting with α values set equal to the average of the α values listed for $\mathbf{R}^{25}$ in table 2 and incrementing until convergence, new EVB fits were generated for $K = 3$–9. As seen in table 3 the average α values increase the error by a factor of 2 for $K = 3$ and a factor of ~1.5 for $K = 5$–9. Depending upon the accuracy required for a particular application, an average α value may be appropriate for production level work.

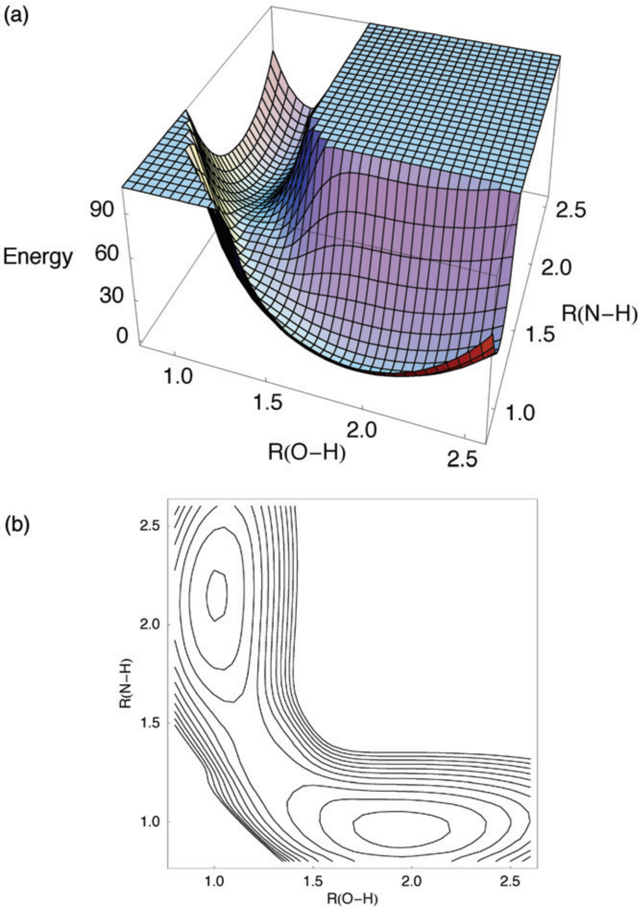

Figure 5. EVB hypersurface for 2-pyridone↔2-hydroxypyridine with $K = 7$.

3.5. *Claisen rearrangement of allyl vinyl ether*

As seen above, proton transfer reactions may be particularly well suited for DG EVB fitting due to the natural Gaussian-like shape of the energy and $V_{12}(\mathbf{q})^2$ along the reaction path. As a more challenging case, we investigated the Claisen rearrangement of allyl vinyl ether to form pent-4-enal employing a set of 61 redundant internal coordinates (figure 6). The optimized alpha values are given in table 4. Although the potential energy profile appears simple enough, $V_{12}(\mathbf{q})^2$ is considerably more complex (see figures 7–9). For the pyridone tautomerism, $V_{12}(\mathbf{q})^2$ is essentially a 'one-humped camel', but in the Claisen reaction $V_{12}(\mathbf{q})^2$ is a 'multi-humped camel' and therefore more difficult to reproduce. With just one Gaussian center at the TS, the PG method can only reproduce $V_{12}(\mathbf{q})^2$ near the TS and does very poorly elsewhere, as seen in table 5.

Four additional Gaussian centers, −6, −1.5, 1.5, and 6 a.u., were placed at points indicated by the shape of the gradient norm and $V_{12}(\mathbf{q})^2$. As K is increased from 3 to 7, the DG EVB fit improves until chemical accuracy is achieved at $K = 7$. As with the pyridone system, reducing the subspace used for $V_{12}(\mathbf{q})^2$ by employing a tolerance value of 10^{-2} had minimal effects on the final hypersurface. The gradient norm error and the mean angle between predicted and actual gradients at $K = 7$ indicate that more data points will be required along the path if the resulting EVB hypersurface is to be used in applications requiring gradients.

Closer inspection of the potential energy surface suggests that the shape and magnitude of $V_{12}(\mathbf{q})^2$ are due to shortcomings in V_{11} and V_{22}. In particular, motion from the allyl vinyl ether reactant to the transition state involves substantial rotation about the C–O bonds. Representing these torsions as three-fold rotation potentials rather than simple quadratic

Table 3. Analysis of gas-phase pyridone tautomerism EVB fits employing average α values. Errors are determined from *ab initio* IRC data. Energies are in kJ mol^{-1}, V_{12}^2 is in a.u.^{2}, and angles are in degrees.

K	$\mathbf{R}^n$	α	Max. energy error	Max. V_{12}^2 error	Max. grad. norm error	Mean angle between gradients
3	25	0.72267	8.302	2.144×10^{-3}	1.128×10^{-2}	15.0
5	25	2.0020	4.868	6.728×10^{-4}	7.067×10^{-3}	9.15
7	25	9.8429	0.7572	8.380×10^{-5}	1.270×10^{-2}	8.13
9	25	12.5620	0.4356	1.274×10^{-4}	4.436×10^{-3}	3.80

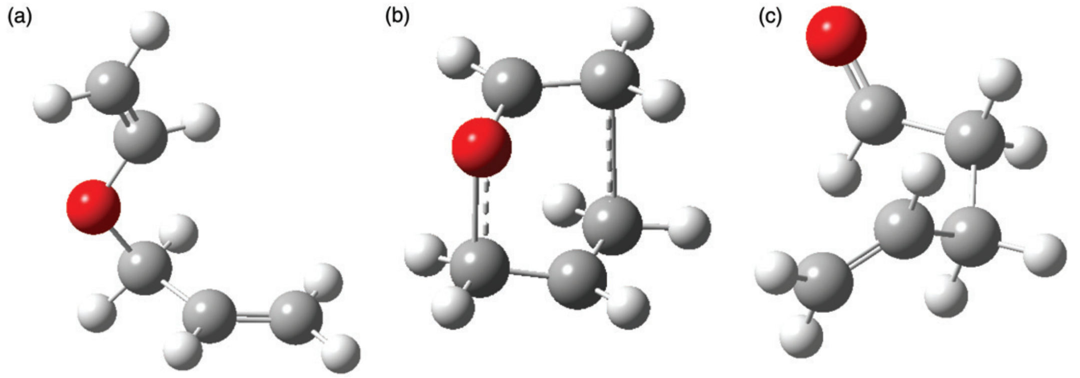

Figure 6. Optimized structures of (a) allyl vinyl ether, (b) the transition state, and (c) pent-4-enal in the Claisen reaction.

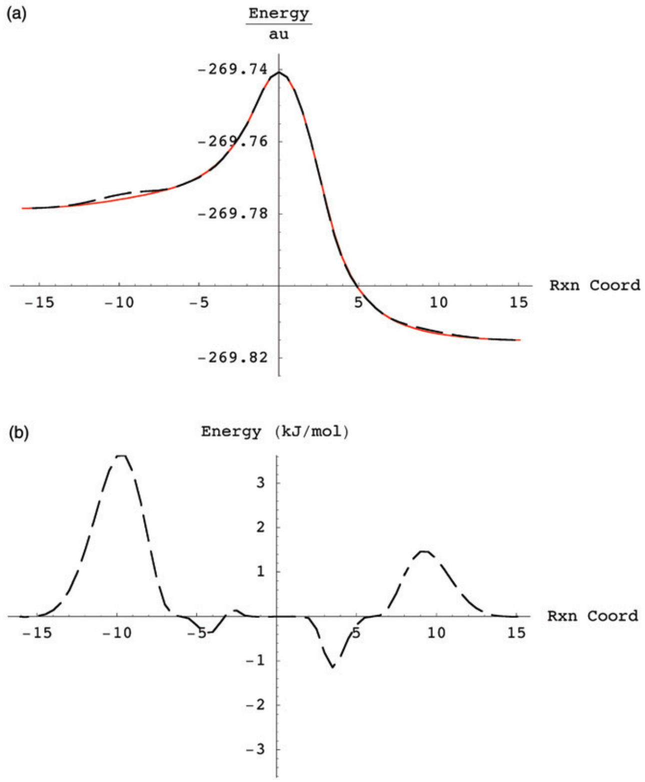

Figure 7. EVB energy along the reaction path for the Claisen reaction with $K=7$ (long dash) and $\mathbf{R}^n=61$. The two surfaces are plotted along with IRC energies (red) in (a), while differences between the IRC and predicted energies are in part (b).

functions of the dihedral angle will improve V_{11} and V_{22} significantly. In turn, this should simplify the shape and diminish the magnitude of $V_{12}(\mathbf{q})^2$ needed in the EVB model to reproduce the reaction path energy profile. The choice of functions for V_{11} and V_{22} for *ab initio* and molecular mechanics based EVB models will be examined in a subsequent paper [34].

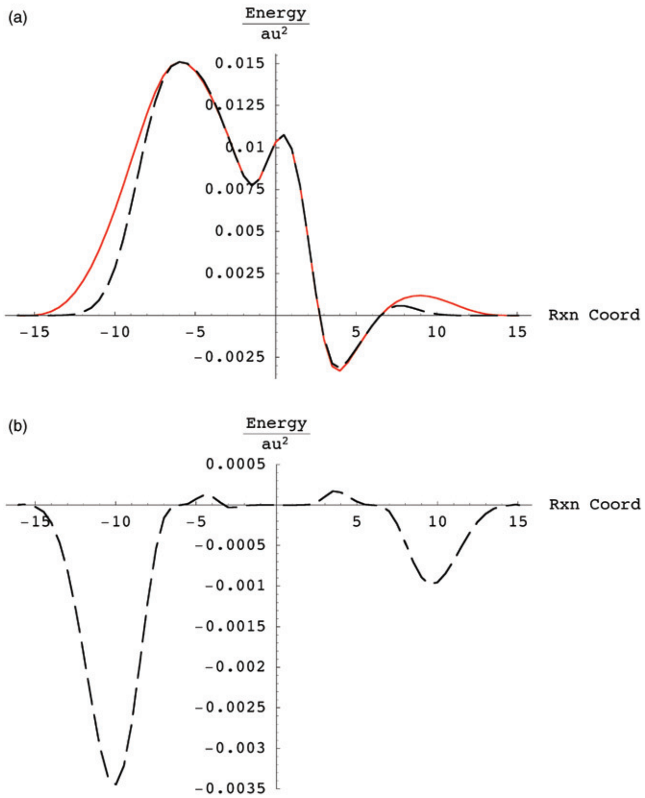

Figure 8. EVB $V_{12}(\mathbf{q})^2$ along the reaction path for the Claisen reaction with $K=7$ (long dash) and $\mathbf{R}^n=61$. The two surfaces are plotted along with IRC energies (red) in (a), while differences between the IRC and predicted energies are in part (b).

4. Conclusions

Distributed Gaussians are convenient for fitting EVB surfaces to energy, gradient and Hessian data from *ab initio* calculations. The large sets of linear equations arising in the fitting procedure can be solved efficiently with a DIIS approach. Redundant internal coordinates

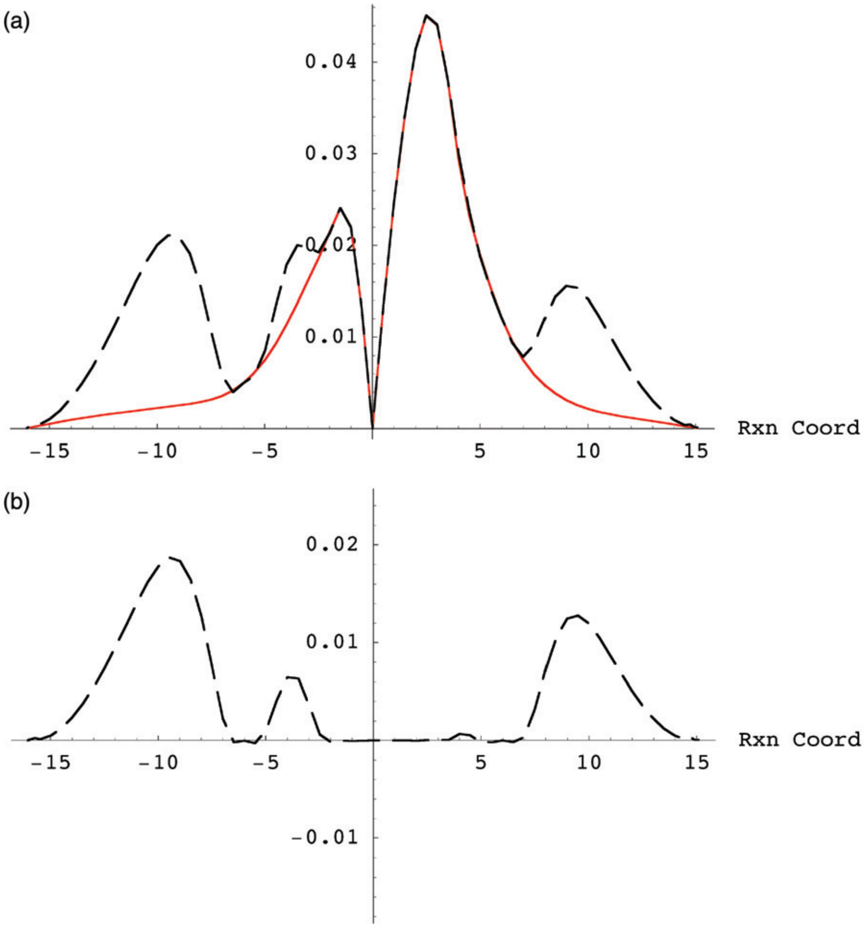

Figure 9. EVB gradient norm along the reaction path for the Claisen reaction with $K=7$ (long dash) and $\mathbf{R}^n=61$. The two surfaces are plotted along with IRC energies (red) in (a), while differences between the IRC and predicted energies are in part (b).

Table 4. Optimized α values for gas-phase Claisen reaction EVB fits.

K	TS	Reactant	Product	1.5	−1.5	6	−6
1	26.718						
3	2.0000	1.3000	6.0000				
5	5.1000	0.8000	0.8400	0.4000	0.4000		
7	2.299	2.299	2.299	2.299	2.299	2.299	2.299

are attractive for representing the $V_{12}(\mathbf{q})^2$ matrix elements in the EVB method, since they avoid the standard orientation problem encountered with the use of Cartesian coordinates. Some care is needed in selecting internal coordinates to ensure that they vary smoothly from reactant to transition state to products. A subset of the internal coordinates may be sufficient for representing $V_{12}(\mathbf{q})^2$. The gradient norm and the energy difference along the reaction path are useful guides to placing additional Gaussians to improve the EVB surface fit. Care should be taken to test that the accuracy both on and off the reaction path is appropriate for the application at hand. With five or more Gaussians, optimizing a single value for α provides a fit nearly as good as optimizing the different α's separately. Tests on pyridone tautomerism and on the Claisen reaction show that EVB surfaces reproduce the energy along the reaction path to chemical accuracy and can be fit with less than 10 *ab initio* data points.

Acknowledgements

The Office of Naval Research (N00014-05-1-0457) supported this research. JLS thanks Professors Fatih Celiker and Martin McClain for their invaluable instruction in the subjects of numerical analysis and

Table 5. Gas-phase Claisen reaction EVB analysis. Errors are determined from *ab initio* IRC data. Energies are in kJ mol^{-1}, V_{12}^2 is in a.u.^{2}, and angles are in degrees.

K	$\mathbf{R}^n$	Norm[**B**]	Max. **B** element	Max. energy error	Max. V_{12}^2 error	Max. grad. norm error	Mean angle between gradients
1	61	0.09091	0.03820	59.27	1.511×10^{-2}	1.054×10^{-1}	77.1
3	61	0.09091	0.03820	39.16	1.494×10^{-2}	5.462×10^{-2}	74.5
5	61	14.85	3.155	10.53	6.672×10^{-3}	1.874×10^{-2}	56.1
5	59	1.782	0.2801	10.40	2.563×10^{-3}	2.008×10^{-2}	63.5
7	61	5.187	0.4660	3.625	3.446×10^{-3}	1.871×10^{-2}	51.4
7	59	24.42	2.492	3.826	3.544×10^{-3}	1.885×10^{-2}	62.1

Mathematica, respectively. The authors are grateful for discussions with Dr. Hrant Hratchian at Indiana University and for computer time from the Wayne State University Grid.

References

[1] A. Warshel and R. M. Weiss, J. Am. Chem. Soc. **102,** 6218 (1980).

[2] Y.-T. Chang and W. H. Miller, J. Phys. Chem. **94,** 5884 (1990).

[3] T. V. Albu, J. C. Corchado, and D. G. Truhlar, J. Phys. Chem. A **105,** 8465 (2001).

[4] Y. Kim, J. C. Cochado, J. Villa, J. Xing, and D. G. Truhlar, J. Chem. Phys. **112,** 2718 (2000).

[5] H. Lin, J. Pu, T. V. Albu, and D. G. Truhlar, J. Phys. Chem. A **108,** 4112 (2004).

[6] H. B. Schlegel and J. L. Sonnenberg, J. Chem. Theory Comput. **2,** 905 (2006).

[7] J. M. Anglada, E. Besalu, J. M. Bofill, and R. Crehuet, J. Comp. Chem. **20,** 1112 (1999).

[8] Y. T. Chang, C. Minichino, and W. H. Miller, J. Chem. Phys. **96,** 4341 (1992).

[9] H. Sellers, Chem. Phys. Lett. **180,** 461 (1991).

[10] H. Sellers, Int. J. Quant. Chem. **45,** 31 (1992).

[11] M. Kawata, C. M. Cortis, and R. A. Friesner, J. Chem. Phys. **108,** 4426 (1998).

[12] X. Li, J. M. Millam, G. E. Scuseria, M. J. Frisch, and H. B. Schlegel, J. Chem. Phys. **119,** 7651 (2003).

[13] Ö. Farkas and H. B. Schlegel, Phys. Chem. Chem. Phys. **4,** 11 (2002).

[14] X. Li and M. J. Frisch, J. Chem. Theory Comput. **2,** 835 (2006).

[15] K. N. Kudin, G. E. Scuseria, and E. Cancès, J. Chem. Phys. **116,** 8255 (2002).

[16] H. A. van der Vorst, Comput. Sci. Engng **2,** 32 (2000).

[17] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics* (Springer, New York, 2000).

[18] W. E. Arnoldi, Q. Appl. Math. **9,** 17 (1951).

[19] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quant. Chem. **13,** 225 (1979).

[20] P. Pulay, Chem. Phys. Lett. **73,** 393 (1980).

[21] P. Pulay, J. Comp. Chem. **3,** 556 (1982).

[22] I. V. Ionova and E. A. Carter, J. Comp. Chem. **17,** 1836 (1996).

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

[24] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. **72,** 650 (1980).

[25] T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. V. R. Schleyer, J. Comp. Chem. **4,** 294 (1983).

[26] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, G. Scalmani, B. Mennucci, V. Barone, G. A. Petersson, M. Caricato, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, X. Li, H. P. Hratchian, J. E. Peralta, A. F. Izmaylov, K. N. Kudin, J. J. Heyd, E. Brothers, V. Staroverov, G. Zheng, R. Kobayashi, J. Normand, J. L. Sonnenberg, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. C. Burant, 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, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, W. Chen, M. W. Wong, and J. A. Pople, *Gaussian DV, Revision F.02* (Gaussian, Inc., Wallingford, CT, 2007).

[27] H. P. Hratchian and H. B. Schlegel, in *Theory and Applications of Computational Chemistry the First Forty Years*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, New York, 2005).

[28] H. B. Schlegel, J. Comp. Chem. **3,** 214 (1982).

[29] H. P. Hratchian and H. B. Schlegel, J. Phys. Chem. A **106,** 165 (2002).

[30] H. P. Hratchian and H. B. Schlegel, J. Chem. Phys. **120,** 9918 (2004).

[31] H. P. Hratchian and H. B. Schlegel, J. Chem. Theory Comput. **1,** 61 (2005).

[32] Mathematica; 5.2, Wolfram Research, 2005.

[33] Y. Saad and M. H. Schultz, SIAM J. Sci. Stat. Comput. **7,** 856 (1986).

[34] J. L. Sonnenberg, K. F. Wong, F. Paesani, W. Zhang, J. Liu, D. Case, T. E. Cheatham, W. H. Miller, G. A. Voth, and H. B. Schlegel, Unpublished work, 2007.