Theor Chem Acc (2012) 131:1170
DOI 10.1007/s00214-012-1170-6

REGULAR ARTICLE

Coordinate reduction for exploring chemical reaction paths

Adam B. Birkholz · H. Bernhard Schlegel

Received: 14 October 2011 / Accepted: 3 January 2012 / Published online: 28 February 2012

Abstract The potential energy surface for the reaction of a typical molecular system composed of N atoms is defined uniquely by $3N$-6 coordinates. These coordinates can be defined by the Cartesian coordinates of the atomic centers (minus overall translation and rotation), or a set of internally defined coordinates such as bond stretches, angle bends, and torsions. By applying principal component analysis to the geometries along a reaction path, a reduced set of coordinates, $d \ll 3N$-6, can be obtained. This reduced set of coordinates can reproduce the changes in geometry along the reaction path with chemical accuracy and may help improve the efficiency of reaction path optimization algorithms.

Keywords Reaction path · PCA · Potential energy surface · Optimization

1 Introduction

A potential energy surface describes the energy of a molecule as a function of its geometrical parameters [1]. The potential energy surface representing the energetics of an N atom molecule is defined in terms of $3N$ Cartesian coordinates or at least $3N$-6 internal coordinates. Minima on the potential energy surface correspond to equilibrium structures such as reactants and products, and first-order saddle points represent transition states for reactions. There are many areas of active study involving the exploration of these surfaces, including geometry optimization [2], reaction path following [3], reaction path optimization [4–6], potential energy surface interpolation [7, 8], and molecular dynamics [9]. The computational difficulty of these algorithms depends heavily upon how many degrees of freedom are used to define the potential energy surface. Any approach to generate a reduced set of coordinates that adequately describes the region of the surface to be explored could improve the efficiency of these algorithms.

A set of (possibly redundant) internal coordinates comprised of the stretching, bending, and torsion of bonds may provide a better chemical description of the structure and flexibility of a molecule than Cartesian coordinates. Delocalized internal coordinates [10] are more compact than primitive redundant internal coordinates and have been employed for geometry optimization and vibrational analysis. They are more general than non-redundant (Z-matrix) internal coordinates, but still use $3N$-6 coordinates to represent the potential energy surface. Reaction paths have been analyzed in terms of Cartesian coordinates, internal coordinates [11], and adiabatic local modes [12]. Because reactions often involve the making and/or breaking of a small number of bonds, it should be possible to represent the paths with far fewer than $3N$-6 coordinates. In this note, we use principal component analysis (PCA) [13] to obtain a reduced number of coordinates to define a reaction path. Principal component analysis has been used successfully to obtain an improved set of internal coordinates for vibrational analysis and to examine conformational changes in molecular dynamics simulations of biomolecules [15, 16].

For a given reaction, one can define a steepest descent reaction path that connects the reactant minimum through the transition state to the products [17]. Reaction path following can be carried out in Cartesian or internal

A. B. Birkholz · H. B. Schlegel (✉)
Department of Chemistry, Wayne State University,
Detroit, MI 48202, USA
e-mail: hbs@chem.wayne.edu

coordinates, with or without mass-weighting [18, 19]. Since a reaction usually involves significant changes in only a few bonds or angles, less than $3N$-6 coordinates should be needed to represent the changes in a molecule along a reaction path. The rest of the coordinates remains approximately constant as the molecule moves along the path from reactants to products. A well-chosen subset of (redundant) internal coordinates may provide a suitable reduction in the number of coordinates needed to represent the path. However, it is often difficult to choose such coordinates manually. Principal component analysis can be applied to the geometries that define a reaction path, described by either internal or Cartesian coordinates. This is a systematic method of obtaining a small set of coordinates that can be used to reproduce the reaction path with chemical accuracy.

In the following sections, we outline a PCA-based coordinate reduction procedure and apply it to a number of reactions for which the reaction path has already been computed in order to provide a benchmark for the method's effectiveness in generating a compact description of a reaction path. The ene reaction is used to illustrate the properties and convergence behavior of the coordinate reduction method. The approach is then tested on a set of model systems that are representative of a variety of different reaction schemes. Finally, the method is applied to a few reactions from the recent literature to demonstrate the coordinate reduction scheme on some larger systems.

2 Methods

Consider a reaction path represented by a set of p structures, with index $1 \le k \le p$. Each structure can be described by $3N$ Cartesian coordinates, where N is the number of atoms. Let c_μ^k be the Cartesian coordinates, where the superscript indicates the sequence number of the structure along the path. Alternatively, internal coordinates, q_μ^k, can be used to describe the structures. Greek subscripts,$1 \le \mu \le m$, are used for the coordinate number. Overall translation and rotation are eliminated from the Cartesian coordinates by moving the structures to a common origin and aligning them to a common axis. Any discontinuities in the torsional coordinates are removed by adding or subtracting 2π as necessary. Linear bends can be avoided by including suitable dummy atoms in the internal coordinate definitions.

Once the structures are prepared as above, the principal component method is used to determine the coordinates important to the reaction. To begin, the average value is removed from each image along the path.

$$\bar{x}_\mu = \frac{1}{p}\sum_{l=1}^{p} c_\mu^l \quad \text{or} \quad \bar{x}_\mu = \frac{1}{p}\sum_{l=1}^{p} q_\mu^l \tag{1}$$

$$x_\mu^k = c_\mu^k - \bar{x}_\mu \quad \text{or} \quad x_\mu^k = q_\mu^k - \bar{x}_\mu \tag{2}$$

The covariance matrix, $\mathbf{M}$, its eigenvectors, $\mathbf{V}$, and the diagonal matrix of eigenvalues, $\mathbf{\Sigma}$, are given by

$$M_{\mu\nu} = \frac{1}{p}\sum_{k=1}^{p} x_\mu^k x_\nu^k = \frac{1}{p}\mathbf{X}^T\mathbf{X}, \quad \mathbf{V}^T\mathbf{M}\mathbf{V} = \mathbf{\Sigma} \tag{3}$$

where $X_{k\mu} = x_\mu^k$. The eigenvectors of the covariance matrix are the principal components, and their associated eigenvalues are the variances of the path along the corresponding coordinates. Alternatively, the eigenvectors of $\mathbf{M}$ can be obtained without computing $\frac{1}{p}\mathbf{X}^T\mathbf{X}$ by singular value decomposition of the matrix $\frac{1}{\sqrt{p}}\mathbf{X}$

$$\begin{aligned} &\frac{1}{\sqrt{p}}\mathbf{U}^T\mathbf{X}\mathbf{V} = \mathbf{S}, \\ &\mathbf{S}^T\mathbf{S} = \frac{1}{p}\left(\mathbf{U}^T\mathbf{X}\mathbf{V}\right)^T\left(\mathbf{U}^T\mathbf{X}\mathbf{V}\right) = \mathbf{V}^T\left(\frac{1}{p}\mathbf{X}^T\mathbf{X}\right)\mathbf{V} \\ &= \mathbf{V}^T\mathbf{M}\mathbf{V} = \mathbf{\Sigma} \end{aligned} \tag{4}$$

where $\mathbf{S}$ is an appropriately dimensioned matrix containing the singular values on the diagonal. The singular values correspond to the standard deviation along the associated coordinates in $\mathbf{V}$.

The above procedure yields a new set of coordinate axes that can be used to represent the path

$$\mathbf{R} = \mathbf{X}\mathbf{V}, \quad r_\lambda^k = \sum_{\mu}^{ncrd} x_\mu^k V_{\mu\lambda} \tag{5}$$

where the $R_{k\lambda} = r_\lambda^k$ are the coordinates rotated onto the principal component axes. If the eigenvalues or singular values are arranged in descending order, the associated coordinates have the property

$$\sum_{k=1}^{p} |r_\lambda^k| > \sum_{k=1}^{p} |r_{\lambda+1}^k| \tag{6}$$

for all values of λ. Only the first $d \ll 3N$ of these coordinates have chemically significant values for any of the images along the path. The path can be reconstructed using only these d coordinates without a loss of chemical accuracy.

$$\tilde{x}_\mu^k = \bar{x}_\mu + \sum_{\lambda=1}^{d} r_\lambda^k V_{\lambda\mu} \tag{7}$$

To assess the accuracy of the reduction, the geometries can be reconstructed with different values of d and compared with the original path in the full coordinate

space. The difference, in Cartesian coordinates, between the reconstructed path versus the original path is used as an error metric. To carry out this comparison with paths reconstructed using redundant internal coordinates, the Cartesians are generated iteratively by fitting displacements in Cartesians to displacements in redundant internals [20]. This process begins by using the Cartesians corresponding to the reactants as a starting point. Each image along the path is fit, iterating until the RMS change in the Cartesian coordinates is less than 10^{-4} bohr. The Cartesian coordinates of the reconstructed paths can be used to generate an energy profile for the reaction. Comparison to the energy profile from the original reaction path can then be used to assess the convergence of the energy as additional coordinates are included in the principal component expansion.

The principal components produced by the above approach will naturally depend on the number of images used in the analysis. A subset of the original path can be constructed by taking the transition state plus every *n*th image downhill in each direction. The PCA method can be applied to these paths to obtain a new set of $\bar{\mathbf{x}}$, $\mathbf{X}$, and $\mathbf{V}$. These new average coordinate values and principal components can then be used to reconstruct the path, using all of the images, and compare it with the original path. This provides a relationship between the number of images used in the PCA, the number of components used in the reconstruction of the path, and the reconstruction error.

A test suite of representative reactions was chosen to study using this coordinate reduction procedure. The transition state for the reaction was found, and a steepest descent reaction path was calculated using the Hessian-based predictor corrector integration method [21] in the Gaussian 09 suite of programs [22]. The step size was set to 0.05 Å, and up to 200 points were computed in each direction. Mathematica [23] was used to analyze the data and calculate the reduced coordinates. For redundant internal coordinates, a single optimization step was taken from the reactant and product geometries found by the reaction path calculation in order to generate a set of internal coordinate definitions. The union of these two sets was used to analyze the reaction in redundant internal coordinates.

3 Discussion and results

The convergence behavior of the coordinate reduction procedure was tested with the ene reaction. This reaction is a condensation of a molecule with a double bond and an allylic hydrogen and a molecule containing a multiple bond (Scheme 1). This system has only a few atoms, but the reaction requires a concerted rearrangement of two double bonds and a hydrogen transfer. In the simplest ene reaction, propene and ethylene react to produce 1-pentene.

The effect of the PCA on the coordinates representing the path is illustrated in Fig. 1. The Cartesian coordinates with the first, second, third, and tenth largest variance for

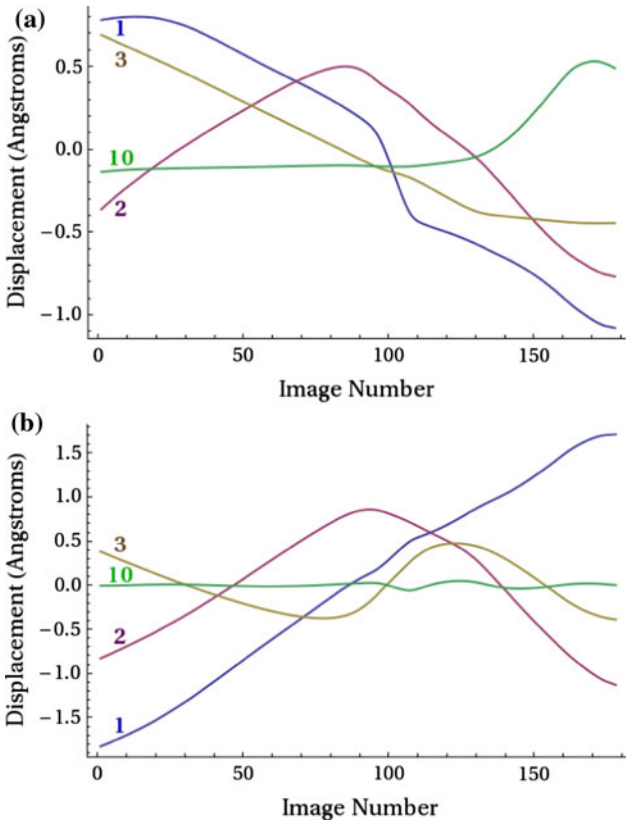
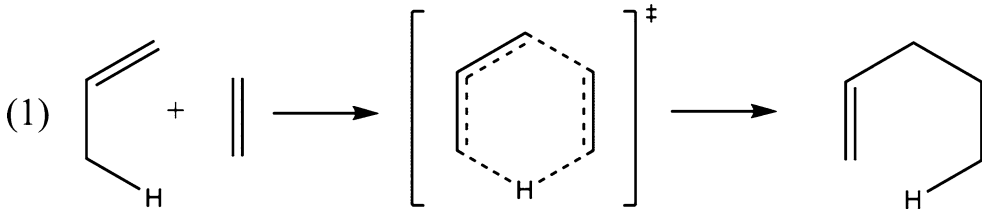


Fig. 1 Reaction path for the ene reaction. **a** The three Cartesian coordinates with the largest variance and the Cartesian coordinate with the 10th largest variance. **b** The three Cartesian principal components with the largest variance and the Cartesian principal component with the 10th largest variance

Scheme 1 Ene reaction



the ene reaction are plotted in Fig. 1a. Note that each coordinate varies around its average value by ca. ±1 Å and that the tenth coordinate varies roughly as much as the first three. Figure 1b shows the first, second, third, and tenth Cartesian principal components. The first principal component is almost linear and has a range of nearly 3 Å. The tenth principal component is nearly constant when compared to the first three components.

Figure 2 shows the energy profiles corresponding to the paths defined by the first d principal components for $d = 1$ through 4 for Cartesian and redundant internal coordinates. With both coordinate systems, the region near the transition state is first to converge. Because the transition state occurs in roughly the middle of the reaction, the mean value of the coordinates is a better approximation to the transition state than it is to either of the minima. The redundant internal coordinates converge more quickly to the true energy profile because they are a more natural representation of chemical motion, particularly when the principal motion involves the rotation of a dihedral angle. In either coordinate system, though, only a few coordinates are needed to converge to the energy. Figure 3 shows the energy convergence for the ene reaction as a function of the number of principal components used to reconstruct the path. The dashed horizontal line marks the chemical accuracy threshold of 1 kcal/mol.

In general, fewer principal components are needed to achieve chemical accuracy in the energy than in the coordinates. Figure 4 demonstrates the convergence behavior of Cartesian and internal coordinates relative to the number of principal components needed to achieve chemical accuracy. The dashed horizontal line marks the chemical accuracy threshold of 0.005 Å. With this level of accuracy, the scalar curvature along the reaction path in the reduced coordinates is essentially the same as for the full set of coordinates.

The effect of using fewer images to represent the path is shown in Fig. 5, where the error in the coordinates is plotted as a function of the number of images and the number of Cartesian principal components used to reconstruct the path. There is very little difference in the

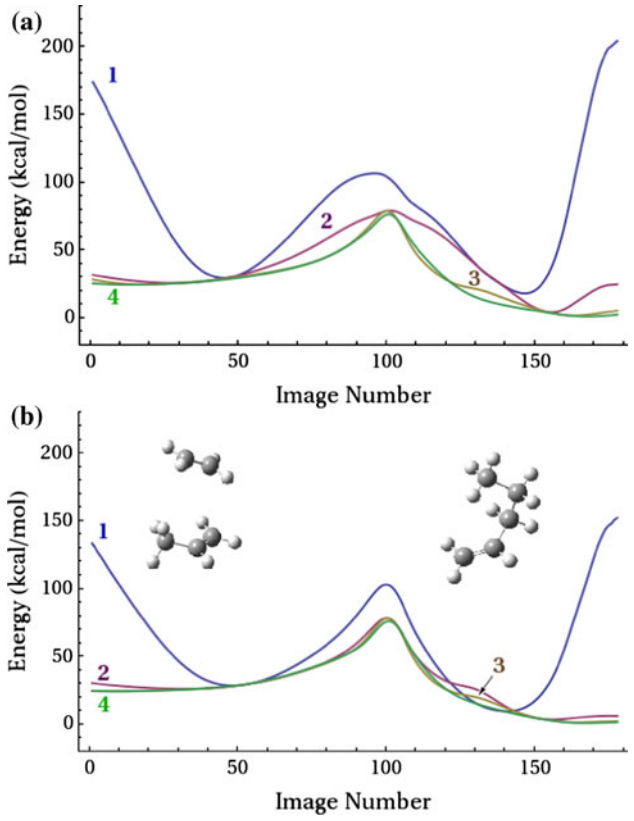


Fig. 2 Energy along the reaction path for the ene reaction using the first d principal components to represent the reaction path, $d = 1$–4. **a** Cartesian coordinates. **b** Redundant internal coordinates

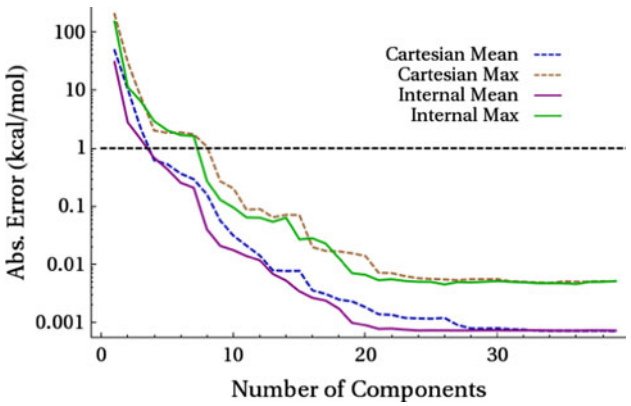


Fig. 3 Mean absolute error and maximum absolute error in the energy along the path for the ene reaction as a function of the number of principal components, the *dashed horizontal line* indicates chemical accuracy of 1 kcal/mol

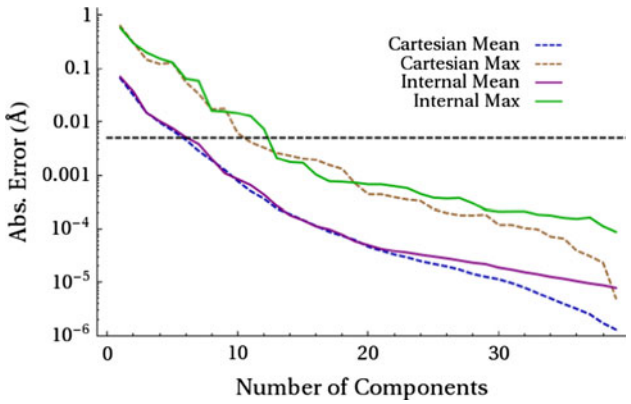


Fig. 4 Mean absolute error and maximum absolute error in the Cartesian coordinates along the path for the ene reaction as a function of the number of principal components. The *dashed horizontal line* indicates chemical accuracy of 0.005 Å

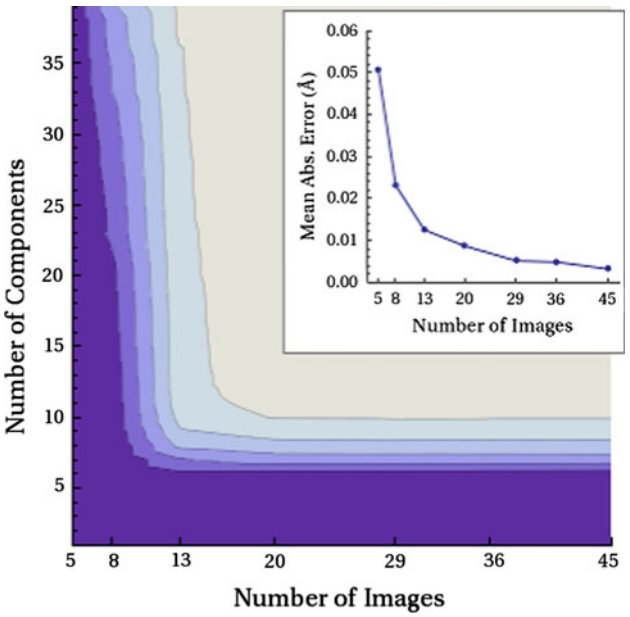


Fig. 5 Contour plot of the mean absolute error in the coordinates along the path for the ene reaction as a function of the number of principal components and the number of images used to generate the principal components, with the darkest region corresponding to an error of >0.005 Å. The *inset* shows the mean absolute error in the computation of the average coordinates versus the number of images in the subset path

accuracy of the path versus the number of components used when more than 20 images are used, out of a total of 178 in the original path. When fewer images are used, the error increases regardless of the number of principal components used in the reconstruction of the path. The inset in the figure shows the mean absolute error in the average coordinates, $\bar{x}_\mu$, as a function of the number of images in the path, and this error does indeed reflect the error in the contour plot. This systematic error is a consequence of the mean-centered data used in the PCA method. For the most reliable results, the images should be spaced to provide an adequate representation of the average value of the coordinates. Furthermore, the number of images should be at least two times the number of principal components needed to achieve chemical accuracy.

Six representative test reactions were chosen to study the properties of the coordinate reduction method (see Schemes 1, 2):

1. The ene reaction of ethylene with propene.
2. The Diels–Alder reaction of ethylene with cyclobutadiene to form norbornene.
3. Hydrolysis of acetamide by water with an additional water molecule.
4. The ring closing step of a Robinson annulation [24].
5. The trans-gauche rotational barrier of 1,2-dichlorethane.
6. The isomerization for alanine dipeptide from the 7_{ax} minima to the α_L minima [25].

These reactions were all carried out at the HF/3-21G level of theory except reactions 2 and 4, which were computed using the PM6 semiempirical method [26]. These examples were analyzed for the convergence to chemical accuracy of both the energy and the coordinates. The first four reactions have cyclic transition states which involve the concerted motion of many atomic centers. Reactions 5 and 6 are driven primarily by the rotation about 1 and 2 dihedral angles, respectively. Dihedral rotations are difficult to model using Cartesian coordinates, but are one of the primary coordinate types used to represent molecules when using internal coordinates.

Table 1 summarizes the energy convergence behavior for the reactions in the above test set, showing the number of principal components needed to achieve chemical accuracy in mean absolute error and max absolute error for every image along the path. In every case, the number of principal components is much smaller than the total number of coordinates. Redundant internal coordinates do a bit better than the Cartesians, especially for reactions dominated by torsions. The convergence behavior for the coordinates is examined in Table 2. The redundant internals perform a little bit worse in this case; but again, there is a large difference between the total number of coordinates and the number in the reduced set for each reaction studied.

In addition to the above test reactions, three examples from the recent literature were also analyzed for their convergence to chemical accuracy of the coordinates (see Scheme 3):

1. The tautomerization of 2-pyridone assisted by two water molecules [27].
2. NO insertion into the Co–CH_3 bond of CpCo(CH_3)(NO) [28].
3. Matrix metalloproteinase 2 (MMP2) inhibition by (*S*)–SB-3CT [29].

The first reaction involves two water molecules in a concerted proton shuttle to convert 2-pyridone to 2-hydroxypyridine. It was calculated at a high level of theory (MP2/6-311 + G(d,p)) and illustrates that the coordinate reduction methodology works equally well when more accurate levels of theory are employed. The second reaction is an example inorganic reaction involving NO migratory insertion into a Co–CH_3 sigma bond and was computed with the B3PW91 functional [30–33] and the 6-311G(d) basis set. The third reaction is a large-scale QM/MM study of the mechanism of SB-3CT reacting in the active site of MMP2 and inhibiting the enzyme. The reaction path involves a concerted proton abstraction, opening of a three-membered ring, and tight binding of a thiolate with the zinc ion in the active site. This example

Scheme 2 Test reactions

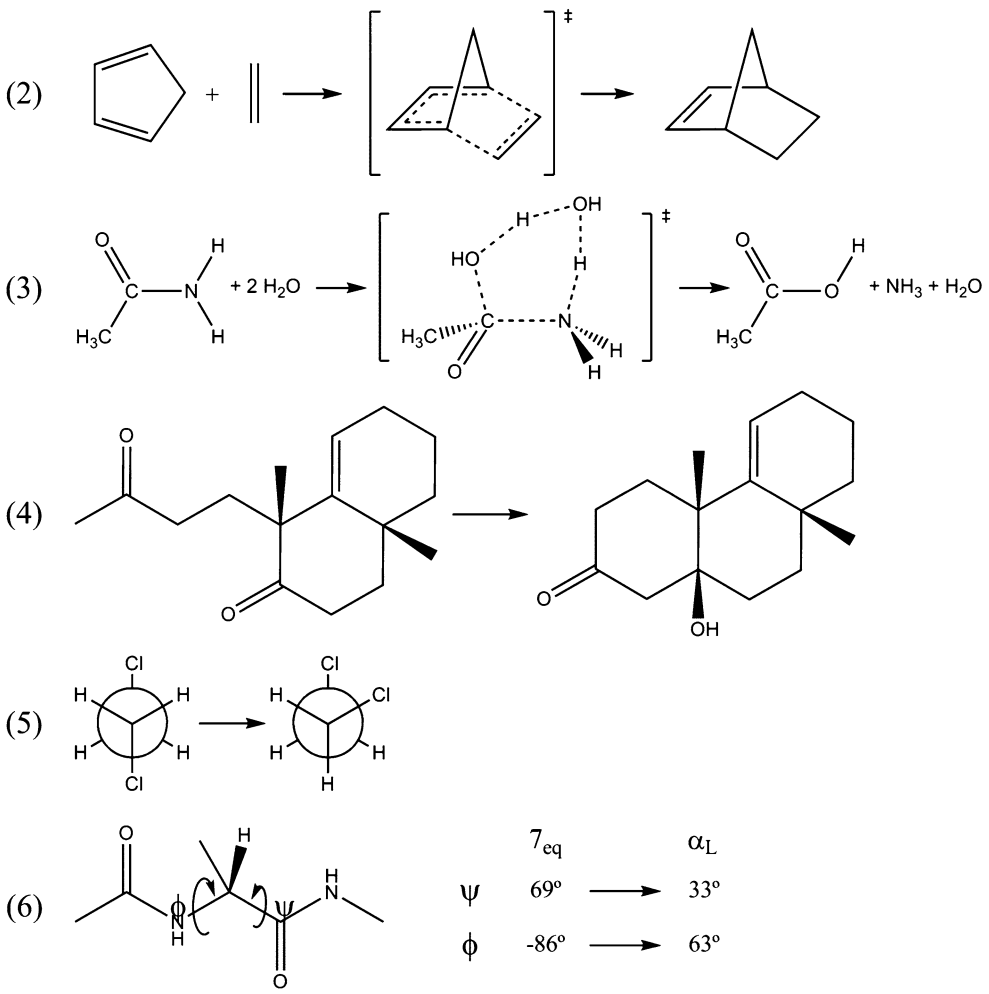


Table 1 Comparison of the number of principal components needed to achieve chemical accuracy (<1 kcal/mol) in the mean and maximum absolute error in the energy of the reduced path

Reaction	Cartesian			Redundant internal coordinates		
	# Total	# PCA mean	# PCA max	# Total	# PCA mean	# PCA max
Ene	45	4	9	79	4	8
Diels–Alder	57	3	5	244	3	5
Amide hydrolysis	45	5	8	112	5	7
Robinson annulation	126	5	8	283	4	7
Dichloroethane rotation	24	2	2	28	1	1
Alanine dipeptide	66	3	5	98	1	2

illustrates that the reaction path for a very large system can depend on only a few degrees of freedom.

Since the energy converges before the coordinates in every example from the test set, only the coordinates were analyzed in the three examples from the literature. These reactions show behavior similar to those of the test set, where fewer than 15 principal components are needed even though many more coordinates are required to define the full system. Even in the extreme case of the QM/MM reaction path where there are thousands of atoms involved, only six Cartesian principal components are necessary to achieve chemical accuracy. For large systems, the maximum error is a better measure of the number of coordinates needed, since the mean absolute error is an average over a large number of nearly stationary atoms.

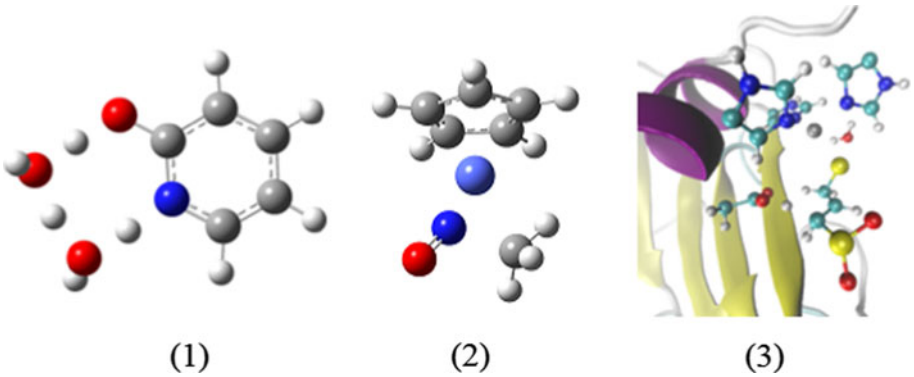
4 Conclusions

Principal component analysis is an effective technique to significantly reduce the number of coordinates necessary to accurately reproduce the geometric changes along a reaction

Table 2 Comparison of the number of principal components needed to achieve chemical accuracy (<0.005 Å) in the mean and maximum absolute error in the Cartesian coordinates of the reduced path

Reaction	Cartesian			Redundant internal coordinates		
	# Total	# PCA Mean	# PCA Max	# Total	# PCA Mean	# PCA Max
Ene	45	6	11	79	7	13
Diels–Alder	57	3	8	244	3	7
Amide hydrolysis	45	7	14	112	7	15
Robinson annulation	126	5	12	283	6	14
Dichloroethane rotation	24	2	4	28	1	3
Alanine dipeptide	66	4	8	98	5	8
Pyridone + $2H_2O$	54	4	9	–	–	–
$CpCo(NO)(CH_3)$	51	3	5	–	–	–
MMP2	7,932	1	6	–	–	–

Scheme 3 Transition states for additional examples



path for a chemical reaction. For the representative reactions studied here, the reduction often exceeded 90%. Such a dramatic reduction may be useful in reaction path optimization, provided the reduced space is allowed to evolve during the optimization. Additionally, coordinate reduction may also be advantageous in reducing the cost of constructing and evaluating high accuracy interpolated potential energy surfaces for studying reaction dynamics.

Acknowledgments This work was supported by a grant from the National Science Foundation (CHE0910858). Wayne State University's computing grid provided computational support.

References

1. Wales D (2004) Energy landscapes: applications to clusters, biomolecules and glasses (Cambridge Molecular Science). Cambridge University Press, Cambridge
2. Schlegel HB (2011) Geometry optimization. WIREs Comput Mol Sci 1:790–809. doi:10.1002/wcms.34
3. Heidrich D (1995) The reaction path in chemistry: current approaches and perspectives. Understanding chemical reactivity, vol 16. Kluwer Academic Publishers, Dordrecht
4. Elber R, Karplus M (1987) A method for determining reaction paths in large molecules: Application to myoglobin. Chem Phys Lett 139:375–380
5. Sheppard D, Terrell R, Henkelman G (2008) Optimization methods for finding minimum energy paths. J Chem Phys 128:134106. doi:10.1063/1.2841941
6. Burger SK, Yang WT (2006) Quadratic string method for determining the minimum-energy path based on multiobjective optimization. J Chem Phys 124:054109
7. Ischtwan J, Collins MA (1994) Molecular potential energy surfaces by interpolation. J Chem Phys 100:8080–8088. doi:10.1063/1.466801
8. Schatz GC (1990) The analytical representation of potential energy surfaces for chemical reactions. Advances in molecular electronic structure theory. JAI Press, London
9. Bolton K, Hase WL, Peslherbe GH (1998) Direct dynamics of reactive systems. Modern methods for multidimensional dynamics computation in chemistry. World Scientific, Singapore
10. Baker J, Kessi A, Delley B (1996) The generation and use of delocalized internal coordinates in geometry optimization. J Chem Phys 105:192–212. doi:10.1063/1.471864
11. Kato S, Morokuma K (1980) Potential-energy characteristics and energy partitioning in chemical-reactions—ab initio MO study of 4-centered elimination-reaction CH3CH2F–CH2 = CH2 + HF. J Chem Phys 73:3900–3914. doi:10.1063/1.440576
12. Konkoli Z, Kraka E, Cremer D (1997) Unified reaction valley approach mechanism of the reaction CH3 + H-2- > CH4 + H. J Phys Chem A 101:1742–1757. doi:10.1021/jp962877j
13. Jolliffe IT (2002) Principal component analysis, vol XXIX, 2nd edn. Springer Series in Statistics, NY
14. Jensen F, Palmer DS (2011) Harmonic vibrational analysis in delocalized internal coordinates. J Chem Theor Comput 7:223–230. doi:10.1021/ct100463a

15. Balsera MA, Wriggers W, Oono Y, Schulten K (1996) Principal component analysis and long time protein dynamics. J Phys Chem 100:2567–2572. doi:10.1021/jp9536920
16. Palmer DS, Jensen F (2011) Predicting large-scale conformational changes in proteins using energy-weighted normal modes. Proteins 79:2778–2793. doi:10.1002/prot.23105
17. Fukui K (1981) The path of chemical-reactions—the IRC approach. Acc Chem Res 14:363–368. doi:10.1021/ar00072a001
18. Gonzalez C, Schlegel HB (1990) Reaction-path following in mass-weighted internal coordinates. J Phys Chem 94:5523–5527. doi:10.1021/j100377a021
19. Gonzalez C, Schlegel HB (1989) An improved algorithm for reaction-path following. J Chem Phys 90:2154–2161. doi:10.1063/1.456010
20. Pulay P, Fogarasi G (1992) Geometry optimization in redundant internal coordinates. J Chem Phys 96:2856–2860
21. Hratchian HP, Schlegel HB (2005) Using Hessian updating to increase the efficiency of a Hessian based predictor-corrector reaction path following method. J Chem Theory Comput 1:61–69. doi:10.1021/ct0499783
22. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratman RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochierski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Mayhall NJ, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2010) Gaussian 09 Version B.01; Gaussian, Inc., Wallingford, CT
23. Mathematica Version 7.0 (2008) Wolfram Research, Inc. Champaign, IL
24. Heathcock CH, Mahaim C, Schlecht MF, Utawanit T (1984) A synthetic approach to the quassinoids. J Org Chem 49:3264–3274. doi:10.1021/jo00192a004
25. Headgordon T, Headgordon M, Frisch MJ, Brooks CL, Pople JA (1991) Theoretical-study of blocked glycine and alanine peptide analogs. J Am Chem Soc 113:5989–5997
26. Stewart J (2007) Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. J Mol Model 13:1173–1213. doi:10.1007/s00894-007-0233-4
27. Sonnenberg JL, Wong KF, Voth GA, Schlegel HB (2009) Distributed gaussian valence bond surface derived from ab initio calculations. J Chem Theory Comput 5:949–961. doi:10.1021/ct800477y
28. Niu SQ, Hall MB (2000) Theoretical studies on reactions of transition-metal complexes. Chem Rev 100:353–405
29. Zhou J, Tao P, Fisher JF, Shi Q, Mobashery S, Schlegel HB (2010) QM/MM studies of the matrix metalloproteinase 2 (MMP2) inhibition mechanism of (S)-SB-3CT and its oxirane analogue. J Chem Theory Comput 6:3580–3587. doi:10.1021/ct100382k
30. Becke AD (1993) Density-functional thermochemistry. 3. The role of exact exchange. J Chem Phys 98:5648–5652
31. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C (1992) Atoms, molecules, solids, and surfaces—applications of the generalized gradient approximation for exchange and correlation. Phys Rev B 46:6671–6687
32. Perdew JP (1991) Generalized gradient approximations for exchange and correlation—a look backward and forward. Phys B 172:1–6
33. Perdew JP (1991) Electronic structure of solids. Akademie Verlag, Berlin