HOKKAIDO UNIVERSITY

Title	Visualization of the Dynamics Effect : Projection of on-the-Fly Trajectories to the Subspace Spanned by the Static Reaction Path Network
Author(s)	Tsutsumi, Takuro; Ono, Yuriko; Arai, Zin et al.
Citation	Journal of chemical theory and computation, 16(7), 4029-4037 https://doi.org/10.1021/acs.jctc.0c00018
Issue Date	2020-07-14
Doc URL	https://hdl.handle.net/2115/81801
Rights	This document is the Accepted Manuscript version of a Published Work that appeared in final form in Journal of chemical theory and computation, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see https://doi.org/10.1021/acs.jctc.0c00018.
Type	journal article
File Information	J. Chem. Theory Comput. 16-7_4029-4037.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Visualization of dynamics effect: projection of on-the-fly trajectories to the subspace spanned by the static reaction path network

Takuro Tsutsumi,[a] Yuriko Ono,[b] Zin Arai,[c] and Tetsuya Taketsugu[a,b,d] *

[a] *Graduate School of Chemical Sciences and Engineering, Hokkaido University, Sapporo 060-0810, Japan*

[b] *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo 001-0021, Japan*

[c] *Academy of Emerging Sciences, Chubu University, Kasugai, Aichi 487-8501, Japan*

[d] *Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan*

ABSTRACT:

Following our recent work to reduce a dimension of a set of reference structures along the intrinsic reaction coordinate (IRC) by a classical multidimensional scaling (CMDS) approach (*J. Chem. Theory Comput.* **2018**, 14, 4263−4270), we propose the method to project on-the-fly trajectories into a reduced-dimension subspace determined by the IRC network, using the out-of-sample extension of CMDS. The method was applied to the S_N2 reaction, OH^- + CH_3F, in which trajectories show a bifurcating nature around the highly-curved region of the IRC path, and to the structural transformation of Au_5 cluster in which the global reaction path network consists of five equilibrium structures and 14 IRCs. It was demonstrated that the present analysis can visualize the dynamics effect by showing a dynamic reaction route on the basis of the static reaction paths.

1. INTRODUCTION

An intrinsic reaction coordinate (IRC)[1] is defined as a steepest descent path in mass-weighted coordinates, which connects two minima (MINs) *via* one transition state (TS) on a potential energy surface (PES). A variation of geometrical structures along the IRC provides an intuitive picture of chemical reaction process. In the last decade, an automated reaction path search approach has been proposed and developed, which includes anharmonic downward distortion following (ADDF) and artificial force induced reaction (AFIR) methods.[2] By employing such automated reaction path search methods, a global reaction route map consisting of all IRC paths could be constructed for a given molecular system. From the viewpoint of dynamics, it is known that a molecule is pushed off the IRC path around a highly-curved region due to a centrifugal force, leading to the excitation of the transverse vibrational modes.[3–5] It is also known that an unstable ridge region sometimes appears on the IRC path through a valley-ridge inflection (or transition) point, which leads to the reaction path bifurcation in some cases.[6–8] Such dynamics effects can be examined by trajectory on-the-fly molecular dynamics (TOF-MD) approach[9–11] in which atomic positions and velocities are evolved by the Newton's equation of motion based on the atomic forces determined from quantum chemical calculations. On-the-fly trajectories explore the PES with a given energy, and the route is not restricted to the IRC paths; they can reach several different isomer structures or dissociative products. To clarify the dynamic reaction routes, we very recently proposed a trajectory mapping method on the basis of a global reaction path network, which was applied to structural transformations of a small gold cluster; we discussed dynamics features of the reactive trajectories, and elucidated mechanism such as an IRC-jump and bifurcations of trajectory ensembles.[12] Very recently, Olasz and Czakó proposed a new approach to analyze on-the-fly trajectories in the coordinate space by combining Eckart-transformation-based approach and stationary-point probability.[13] Vázquez *et al.* also developed a unique approach of transition state search using chemical dynamics simulation (tsscds), which employs a concept of accelerated MD simulation and graph theory to get candidates of TS geometries.[14]

To get a physical interpretation on dynamic reaction routes based on the global reaction path network, the structures in the network must be located in appropriate positions. In other words, similar structures should be located close to each other while largely different structures should be located far away. A classical multidimensional scaling (CMDS) technique[15,16] enables us to locate each structure in an appropriate position in the global reaction path network. Recently, Li *et al.* employed the CMDS method to analyze trajectories obtained by the surface-hopping nonadiabatic on-the-fly dynamics simulation.[17] In our previous study,[18] the CMDS method was employed to visualize the IRC paths for an intramolecular proton transfer in malonaldehyde and of the collision reaction, $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$, as well as the global reaction path network for a small gold cluster, in a reduced dimension.

Once the global reaction path network is prepared in a reduced dimension by the CMDS method, dynamic reaction routes could be discussed by projecting on-the-fly trajectories into the reduced subspace. Mapping a new data set into the subspace spanned by a known data set is possible by employing "out-of-sample extensions" of CMDS formulated by Trosset and Priebe.[19] The out-of-sample extension of CMDS (oCMDS) was applied to a classification of protein conformers where on-the-fly trajectories were embedded into the CMDS configurational space constructed by five crystal structures corresponding to typical conformers.[20] In this paper, we propose the method to study the dynamic reaction routes in the subspace determined by the IRC path and the global reaction path network, using the TOF-MD simulations and oCMDS analysis. We demonstrate the usefulness of the oCMDS analysis through applications to the collision reaction, $OH^- + CH_3F$, and the structural transformation of a small gold cluster.

2. METHODOLOGY

In this section, the CMDS procedure for the reference structures on the IRC is described briefly, and then, the oCMDS approach is introduced to project on-the-fly trajectories into the reduced-dimensional subspace. A geometrical structure of N-atomic molecule can be represented by $3N$ mass-weighted Cartesian coordinates, $\boldsymbol{\xi}$, where the origin of the coordinate axes is taken at the center of mass. Then, the difference of two structures, $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$, can be measured by the distance function defined as,

$$d_{ij} = \left|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right| = \sqrt{\sum_{k=1}^{3N}\left(\xi_i^{(k)} - \xi_j^{(k)}\right)^2} \tag{1}$$

where $\xi_i^{(k)}$ denotes the k-th coordinate of the i-th structure. For a pair of molecular geometries, $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$, the direction of x-y-z coordinate axes are determined to minimize d_{ij} by the Kabsch algorithm.[21] Recently, we proposed the trajectory mapping method[12] which describes the dynamical reaction paths based on the distance functions between the points on the trajectory and the reference structures taken over the static reaction path network.

Hence, we pick up n reference structures along the IRC, denoted by $\{\boldsymbol{\xi}_i^{\mathrm{IRC}}, i = 1, \ldots, n\}$, and employ the CMDS method to determine the subspace based on the reference structural data.[18] The CMDS procedure consists of four steps[15,16]:

Step 1. Generate a distance matrix $\mathbf{D}$ with the ij-th element d_{ij} in eq. (1).

Step 2. Apply a Young-Householder transformation[22] (called a double centering) to the squared distance matrix $\mathbf{D}_{ij}^{(2)} = d_{ij}^2$ in order to get an inner matrix $\mathbf{Q}$.

$$\mathbf{Q} = -\frac{1}{2}\left(\mathbf{E} - \frac{1}{n}\mathbf{1}\right)\mathbf{D}^{(2)}\left(\mathbf{E} - \frac{1}{n}\mathbf{1}\right)^{\mathrm{T}} \tag{2}$$

Here, $\mathbf{E}$ is a unit matrix and $\mathbf{1}$ is a square matrix with all elements as 1.

Step 3. Diagonalize $\mathbf{Q}$. Eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ and the corresponding eigenvectors $\{\mathbf{L}_1, \ldots, \mathbf{L}_n\}$ are ordered in a descending way of the eigenvalues ($\lambda_1 \geq \ldots \geq \lambda_n$).

Step 4. Take p largest eigenvalues $\{\lambda_1, \ldots, \lambda_p\}$ with the corresponding eigenvectors $\{\mathbf{L}_1, \ldots, \mathbf{L}_p\}$, and generate principal coordinates $\mathbf{Y}_a = \mathbf{L}_a\sqrt{\lambda_a} \quad (1 \le a \le p)$.

The contribution of the respective principal coordinate to represent the relative positions of the reference structures is measured by the proportion Λ_a[16] defined as

$$\Lambda_a = \frac{\lambda_a}{\sum_c^b \lambda_c} \tag{3}$$

where b is the number of nonnegative eigenvalues of λ_c. The representability in the reduced dimensional subspace spanned by $\{\mathbf{Y}_1, \ldots, \mathbf{Y}_p\}$ can be evaluated by the cumulated proportion Λ[16],

$$\Lambda = \frac{\sum_c^p \lambda_c}{\sum_c^b \lambda_c} \tag{4}$$

When Λ is close to 1, information of the relative positions of the reference structures is almost kept in a p-dimensional principal coordinate space.

In the oCMDS method, out-of-sample data are projected on the p-dimensional CMDS coordinate space.[19] To illustrate the procedure of the oCMDS method, we introduce distance matrices, $\mathbf{D}_{n\times n}$, $\mathbf{D}_{n\times m}$, and $\mathbf{D}_{m\times m}$, for the reference structures and out-of-sample structures where n and m denote the number of the reference structures and out-of-sample structures, respectively. Then, the squared distance matrix $\mathbf{A}$ involving $\mathbf{D}^{(2)}_{n\times n}$, $\mathbf{D}^{(2)}_{n\times m}$, and $\mathbf{D}^{(2)}_{m\times m}$ is defined as

$$\mathbf{A} \equiv \begin{pmatrix} \mathbf{D}^{(2)}_{n\times n} & \mathbf{D}^{(2)}_{n\times m} \\ \mathbf{D}^{(2)\ \mathrm{T}}_{n\times m} & \mathbf{D}^{(2)}_{m\times m} \end{pmatrix} \tag{5}$$

where $\mathbf{D}^{(2)\ \mathrm{T}}_{n\times m}$ denotes a transposed matrix of $\mathbf{D}^{(2)}_{n\times m}$. Next, a double centering is carried out for $\mathbf{A}$ following the Trosset formula,[19]

$$\mathbf{B} = \tau_{\mathbf{w}}(\mathbf{A}) = -\frac{1}{2}\left(\mathbf{E} - \frac{\mathbf{e}\cdot\mathbf{w}^{\mathrm{T}}}{n}\right)\mathbf{A}\left(\mathbf{E} - \frac{\mathbf{e}\cdot\mathbf{w}^{\mathrm{T}}}{n}\right)^{\mathrm{T}} \tag{6}$$

where $\mathbf{E}$ is a unit matrix, $\mathbf{e}$ is a column vector of $n + m$ dimension with all elements as 1, and $\mathbf{w}$ is a column vector of $n + m$ dimension with 1 (first n components) and 0 (remaining m components). The

out-of-sample problem to obtain p-dimensional coordinates $\mathbf{X}_{(n+m)\times p}$ for out-of-sample structures is now solved by minimization of the following function,[19]

$$\min \left\| \mathbf{B} - \mathbf{X}\mathbf{X}^{\mathrm{T}} \right\|^2 = \min \left\| \begin{pmatrix} \mathbf{B}_{n\times n} - \mathbf{Y}\mathbf{Y}^{\mathrm{T}} & \mathbf{B}_{n\times m} - \mathbf{Y}\mathbf{Z}^{\mathrm{T}} \\ \mathbf{B}_{m\times n} - \mathbf{Z}\mathbf{Y}^{\mathrm{T}} & \mathbf{B}_{m\times m} - \mathbf{Z}\mathbf{Z}^{\mathrm{T}} \end{pmatrix} \right\|^2 \quad (7)$$

where $\mathbf{B}_{n\times n}$, $\mathbf{B}_{n\times m}$, $\mathbf{B}_{m\times n}$, and $\mathbf{B}_{m\times m}$ are a part of the inner matrix of $\mathbf{B}$ with $n + m$ dimension. $\mathbf{X}$ denotes p-dimensional coordinates as

$$\mathbf{X} = \mathbf{X}_{(n+m)\times p} \equiv \begin{pmatrix} \mathbf{Y}_{n\times p} \\ \mathbf{Z}_{m\times p} \end{pmatrix} \quad (8)$$

where $\mathbf{Y}_{n\times p}$ and $\mathbf{Z}_{m\times p}$ are coordinates for n reference structures and m out-of-sample structures, respectively. Hence, variables for optimization in Eq. (7) are only elements of $\mathbf{Z}_{m\times p}$, and through this optimization, the principal coordinates can be determined for out-of-sample structures in the p-dimensional CMDS subspace.

3. RESULTS AND DISCUSSION

3.1 S_N2 Reaction of $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$

As the first target reaction, we chose the collision reaction, $OH^- + CH_3F$. First, the IRC was calculated for this reaction at the MP2/6-31+G* level, and the CMDS method was employed for 101 reference structures taken along the IRC path. Then, TOF-MD simulation was carried out at the same computational level, starting from the TS region. As the initial conditions, normal coordinates and conjugate momenta were determined randomly near the TS based on the Boltzmann distribution at 300 K, and then 50 trajectories which initially moved toward the product region were selected. The time step was taken as 0.1 fs, and the respective trajectories were run over 1000 fs, until the C-F interatomic distance reaches 8 Å, or until the distance function between the point on the trajectory and the product structure P reaches 5 Å amu$^{1/2}$. Lastly, the oCMDS method was employed to project on-the-fly trajectories into the two-dimensional subspace determined for the IRC path by the CMDS method. In the present application, structural data of on-the-fly trajectories was added one by one as out-of-samples to project it onto the coordinate space determined by the IRC. The electronic structure calculations were performed by the GAMESS program,[23] while TOF-MD simulations were performed by the SPPR program[24].

For *N*-atomic molecule, the IRC is regarded as a one-dimensional curved pathway in $3N - 6$ dimensional coordinate space. The reaction path tangent vector **v** for the IRC is defined as

$$\mathbf{v}(s) = \frac{d\boldsymbol{\xi}^{\mathrm{IRC}}(s)}{ds} \tag{9}$$

where s denotes a reaction coordinate and $\boldsymbol{\xi}^{\mathrm{IRC}}(s)$ denotes the mass-weighted Cartesian coordinates of the structure at s on the IRC. A curvature vector $\mathbf{v}'(s)$ and the corresponding scalar curvature $\kappa(s)$ are defined as

$$\mathbf{v}'(s) = \frac{d\mathbf{v}(s)}{ds} = \frac{d^2\boldsymbol{\xi}^{\mathrm{IRC}}(s)}{ds^2} \tag{10}$$

$$\kappa(s) = \left|\mathbf{v}'(s)\right| \tag{11}$$

The reaction path curvature vector can be calculated by numerical differentiation of the reaction path tangent vector, and the scalar curvature is defined as the norm of the curvature vector.[5] **Figure 1** shows variations of the energy (black dots) and scalar curvature (red dots) along the IRC for the reaction, $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$, where A ($s$ = 1.74 Å amu$^{1/2}$) and B (s = 5.36 Å amu$^{1/2}$) correspond to the structures at the maximum of the scalar curvature in the product side. Geometrical changes along the IRC indicate that (1) F^- first attempts to dissociate from CH_3 until reaching the structure A, (2) F^- starts to roam around CH_3OH moiety (from A to B), and (3) a hydrogen-bonded structure is formed with an internal rotation of CH_3 fragment, which corresponds to the IRC terminal (from B to P). The large path curvature around A suggests a possible energy transfer between the C-F bond dissociation mode and F^- roaming mode. In order to discuss in more detail which modes are excited through the curvature of IRC, it is useful to project the reaction path curvature vector to the vibrational modes orthogonal to IRC,[3] or converting it to an internal coordinate representation in a scheme of unified reaction valley approach (URVA).[25]

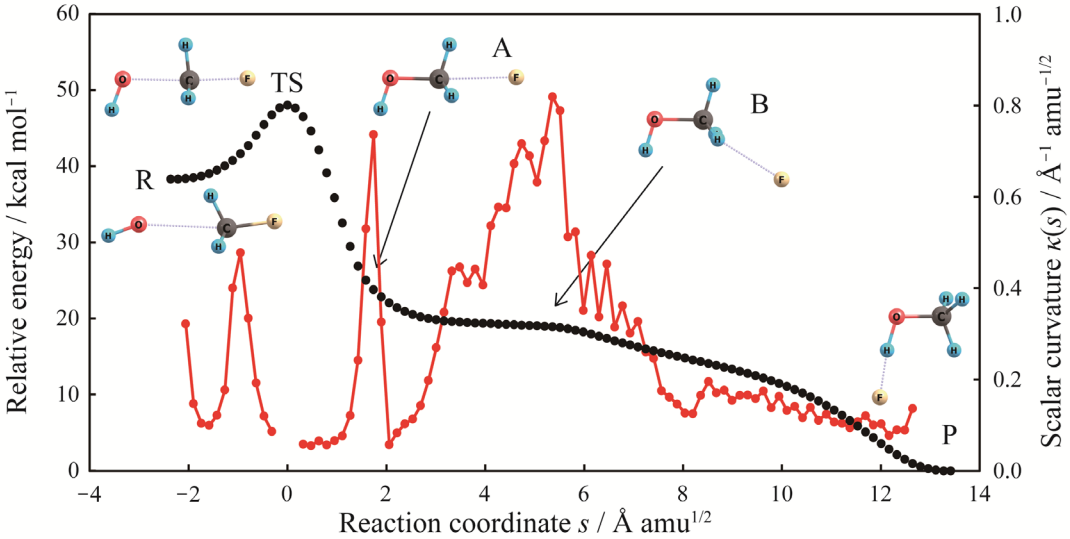


Figure 1. The energy profile (black dots) and the scalar curvature (red dots) along the IRC for $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$. Geometrical structures of reactant (R), TS, two highly-curved regions (A and B), and product (P) are also shown.

In the same way as our previous study,[18] the CMDS method was employed to visualize the IRC path for $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$ in a reduced dimension. In the present computations, 101 geometries along the IRC were picked up, and the distance matrix **D** was generated for these reference geometries with the ij-th element d_{ij} defined in Eq. (1). Following the steps 2 ~ 4 in Section 2, the principal coordinates (denoted by PCo) were determined to represent the IRC path effectively in a reduced-dimensional subspace. **Figure 2** shows the IRC path in the two-dimensional principal coordinates, PCo1 and PCo2. The proportions for PCo1 and PCo2 are $\Lambda_1 = 0.819$ and $\Lambda_2 = 0.134$, respectively, and the cumulated proportion Λ for these two-dimensions is 0.953, indicating that most geometrical information of the IRC path is represented in this two-dimensional subspace. In Fig. 2, positions of reactant (R), transition state (TS), highly-curvature points (A, B), and product (P) are also shown where the region around B shows a high curvature but the region around A shows a calm curvature because of a reduction of the dimension. As discussed in our previous study,[18] PCo1 correlates with an O-C-F bond angle (related to F^- roaming) while PCo2 correlates with a C-F interatomic distance (related to C-F dissociation).

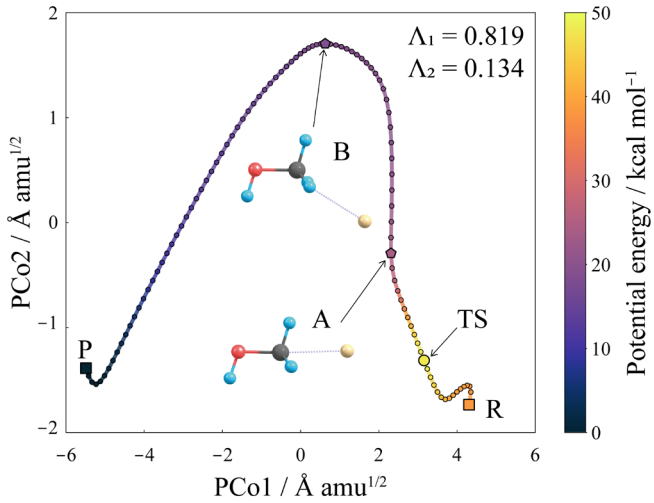


Figure 2. The IRC for $OH^- + CH_3F \rightarrow [CH_3OH\cdots F]^-$ projected to the two-dimensional principal coordinate subspace. R, TS, and P denote the reactant, transition state, and product, respectively, while A and B denote the points with maximum of the reaction path curvature. The color shows the potential energy relative to the energy of P.

Next, we turn to a dynamical aspect of this reaction on the basis of on-the-fly trajectories. Sun *et al.*[10] reported a pioneering work for this reaction to demonstrate the significance of non-IRC path by TOF-MD simulations at the MP2 level, where only 10% of trajectories reach the terminal of the IRC path, $[CH_3OH\cdots F]^-$, while 90% of trajectories lead to the dissociated products, $CH_3OH + F^-$. The origin of non-IRC path was discussed based on the inefficient intramolecular vibrational energy distribution (IVR); the trajectories proceeding along the IRC need to turn their direction to the deep minimum of $[CH_3OH\cdots F]^-$ through the IVR between the O-C---F^- bending mode and the relative translation motion of CH_3OH and F^-, but this IVR is very inefficient, resulting in the energy release to the relative translational motion of CH_3OH and F^- rather than the O-C---F^- bending mode. In the present TOF-MD simulations at the same MP2 level, 50 trajectories were run from the TS toward the product side, where 11 trajectories proceed near the IRC while 39 trajectories lead to the dissociation of F^- from CH_3OH. Among the 39 trajectories leading to $CH_3OH + F^-$, two trajectories show a return of F^- to the moiety of CH_3OH, making a complex of $[CH_3OH\cdots F]^-$ near the structures A and B. Following our previously proposed method,[12] we calculated distance functions for all the points along the trajectories from the reference structures, R, TS, A, B, and P, and plotted those as a function of time. **Figure 3** shows changes of the distance functions for (a) a typical trajectory proceeding near the IRC, (b) a typical trajectory leading to $CH_3OH + F^-$, and (c) a typical trajectory where F^- leaves once and comes back to the CH_3OH moiety. Fig. 3a indicates that the trajectory proceeds along the IRC as TS → A → B → P (the closest structure changes in this order); the trajectory calculation was stopped at 346.9 fs when the distance function from P reached 5 Å $amu^{1/2}$. In the case of Figs. 3b and 3c, the trajectories depart from the IRC in the region between A and B, suggesting that they follow the non-IRC route of C-F dissociation due to an inertial force. This feature is related to the inefficient IVR[10] around the highly-curved region. In a case of Fig. 3b, the C-F interatomic distance reaches 8 Å at 279.2 fs, while in a case of Fig. 3c, F^- could not escape from CH_3OH and returns to the structure A at 740.0 fs, due to an electrostatic interaction (F^- continues leaving and returning motions).

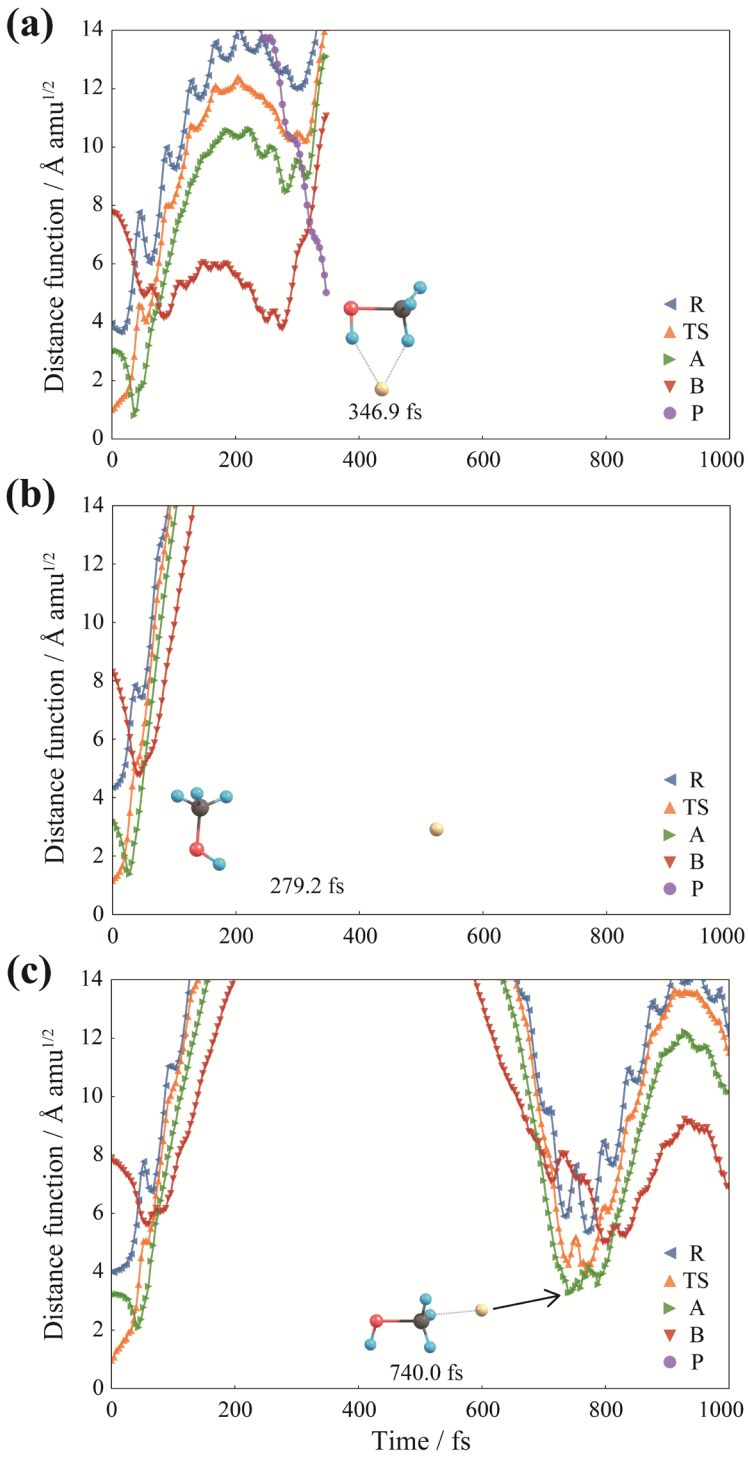


Figure 3. Variations of the distance functions from the reference structures, R, TS, A, B, and P, for (a) a typical trajectory proceeding near the IRC, (b) a typical trajectory leading to $CH_3OH + F^-$, and (c) a typical trajectory where F^- leaves once and comes back to the CH_3OH moiety.

Hence, we employed the oCMDS method to map on-the-fly trajectories into the reduced two-dimensional subspace determined for the IRC by the CMDS method. According to the procedures in a previous section, geometrical structures along the trajectories are treated as out-of-sample data one by one, and the principal coordinates are determined for each out-of-sample structure by the optimization procedure in Eq. (7), using a modified Powell algorithm.[26] **Figure 4** shows examples of three different type of trajectories mapped on the two-dimensional subspace by the oCMDS method, corresponding to trajectories in Figs. 3a, 3b, and 3c: (a) a trajectory proceeding near the IRC, (b) a trajectory leading to $CH_3OH + F^-$, and (c) a trajectory where F^- leaves once and comes back to the CH_3OH moiety. In the case of Fig. 4a, the molecule starting from TS attempts to depart from the IRC around the highly-curved region of B, but it comes back to the IRC path due to an attraction force from the IRC, and enters the well region around P. In the case of Fig. 4b, the molecular system largely departs from the IRC around the highly-curved region of B, and F^- leaves from a moiety of CH_3OH. In the case of Fig. 4c, the molecular system attempts to depart from the IRC around the highly-curved region of B, but it comes back to the region near A. The difference in trajectories between Fig. 4b and Fig. 4c may be caused by the energy distribution of F^- and CH_3O at the highly-curved region around B, which works as a bifurcation point in a dynamical sense. The centrifugal force due to a large curvature of the IRC pushes the molecule in the direction orthogonal to the IRC path in which major trajectories depart from the IRC (leading to the $CH_3OH + F^-$ dissociation) while minor trajectories return to the IRC (leading to the $CH_3OH\cdots F^-$ complex with a highly-excited vibrational energy).

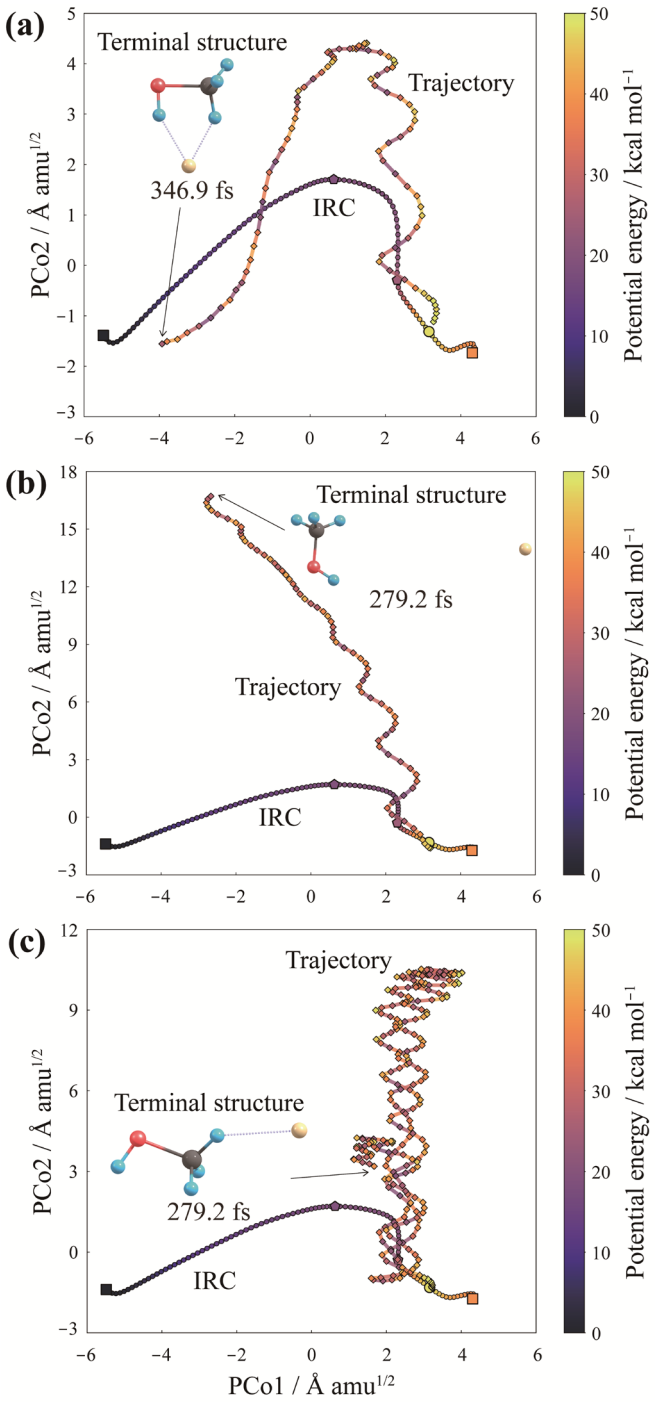


Figure 4. The trajectories projected to the two-dimensional subspace determined for the IRC path: (a) a trajectory which proceeds near the IRC and reaches the region around P; (b) a trajectory which departs from the IRC at a highly-curved region; (c) a trajectory which could not escape from the IRC due to the attraction force from the IRC. The color on the IRC and trajectories denotes the potential energy values.

We would introduce the significant work related to the present study. Very recently, Hare *et al.*[27] proposed a method to reduce the dimensionality of the reaction paths for polyatomic reactions based on a principal component analysis (PCA), and demonstrated the methodology by applying it to several reactions which involve the collision reaction, $OH^- + CH_3F$, following our previous report.[18] In their approach, the Cartesian coordinates for the reference structures on the IRC are directly used to determine the principal components, and then, the coordinates for the corresponding principal components are assigned to the points along on-the-fly trajectories. Because of the employment of Cartesian coordinates, their approach is more straightforward than our oCMDS approach in which the distance functions evaluated for all the pairs of the reference structures are used to determine the principal coordinates. However, the PCA approach requires the common x-y-z Cartesian coordinate axes, which makes it challenging to apply for largely-different structures, and also, the treatment of nuclear-permutation inversion isomers becomes more complicated. We believe that it is valuable to develop both of oCMDS and PCA approaches as the new methodology to discuss dynamics effects from on-the-fly trajectories mapped on the reduced dimensional subspace determined for the static reaction paths.

3.2 Global reaction path network of Au_5 cluster

As the second application, oCMDS is applied to the structural changes of Au_5 cluster which was investigated previously by ADDF,[8] TOF-MD,[12] and CMDS approaches.[18] As shown in **Fig. 5**, Au_5 cluster has five MINs (MIN*i*; *i* = 1-5) and 14 TSs and the global reaction path network is composed of 14 IRCs at the PBE/LanL2DZ level,[8] where TS*i*-*j* denotes the TS connecting MIN*i* and MIN*j*, and TS*i*-*i* denotes the TS connecting two nuclear permutation-inversion (NPI) isomers of MIN*i*; the TSs which connect the same pair of MIN*i* and MIN*j* are distinguished by a letter (a, b, c, and d). Several valley-ridge transition (VRT) points related to a reaction path bifurcation were located on the IRCs. In previous study,[12] dynamics related to reaction path bifurcation was investigated by TOF-MD simulations at the PBE/LanL2DZ level in which 200 trajectories calculated from TS1-1d were analyzed by the distance functions. Very recently, we applied[18] CMDS to all MINs and TSs involved in the global reaction path network of Au_5 cluster, and visualized the network onto the two principal coordinate subspace. In the present study, we reuse the global reaction path network and the TOF-MD trajectories from TS1-1d at the PBE/LanL2DZ level.[8,12,18] We first employ CMDS to visualize the global reaction path network in a reduced dimension by involving not only MINs and TSs but also the molecular structures along the IRCs, and then, the trajectories from TS1-1d are projected onto the principal coordinate subspace spanned by the IRC network to verify the dynamic reaction route.

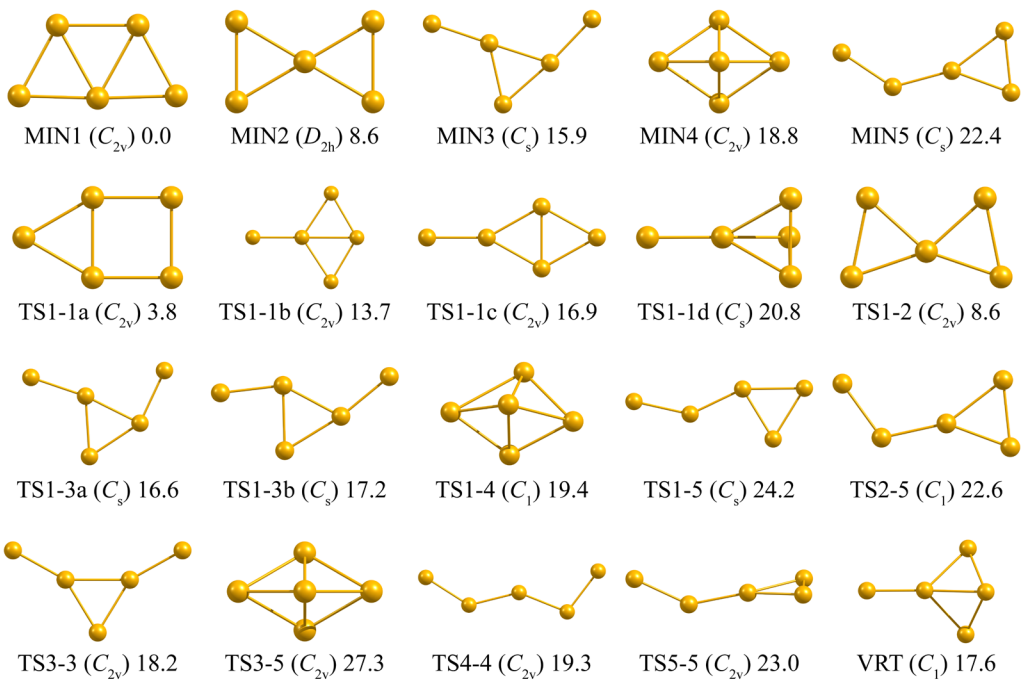


Figure 5. All minima, TS structures, and one VRT structure near TS1-1b for Au_5 calculated by PBE/LanL2DZ.[8] The point group and relative energy (in kcal/mol) are also given for each structure.

First, we examine a case of the reaction path bifurcation related to VRT along the IRC from TS1-1d to MIN1. By distinguishing NPI isomers,[8] three NPI isomers of TS1-1d are located very closely (denoted by TS1-1d, TS1-1d', and TS1-1d"), and three IRC paths from the respective TS1-1d structures connect two NPI isomers of MIN1 (denoted by MIN1 and MIN1', MIN1" and MIN1"", and MIN1"' and MIN1""', respectively). As discussed in the previous study,[8] VRT has the corresponding transition state, TS1-1b, which also connects two NPI isomers of MIN1 (MIN1 → TS1-1b → MIN1", MIN1"" → TS1-1b' → MIN1""', and MIN1' → TS1-1b" → MIN1"'). Then, CMDS is used to determine principal coordinates for the structures along three IRCs connected by TS1-1d, TS1-1d', and TS1-1d" and those along three IRCs connected by TS1-1b, TS1-1b', and TS1-1b" (in total, 387 structures). **Figure 6a** shows six IRCs which are projected on the two-dimensional principal coordinate subspace by PCo1 and PCo2 where circles, squares, diamonds, and smaller circles denote MINs, TSs, VRTs, and IRC points, respectively. The proportions of PCo1 and PCo2 show the same

value, 0.365, indicating that these two principal coordinates belong to a doubly-degenerate symmetry representation. The cumulative proportion of the two-dimensional subspace is 0.730, and the IRCs for NPI isomers of TS1-1b are not clearly shown in Fig. 6a. Then, we applied CMDS to the reference structures in the IRC from TS1-1d to MIN1, the IRC from TS1-1d' to MIN1", and the IRC from MIN1 to MIN1" *via* TS1-1b (in total, 130 structures). Figure 6b shows these partial IRCs which are projected on the two-dimensional subspace. The proportions for PCo1 and PCo2 are $\Lambda_1 = 0.528$ and $\Lambda_2 = 0.432$, respectively, and the cumulated proportion is 0.960, indicating the almost all configurational information is represented in this two-dimensional subspace.

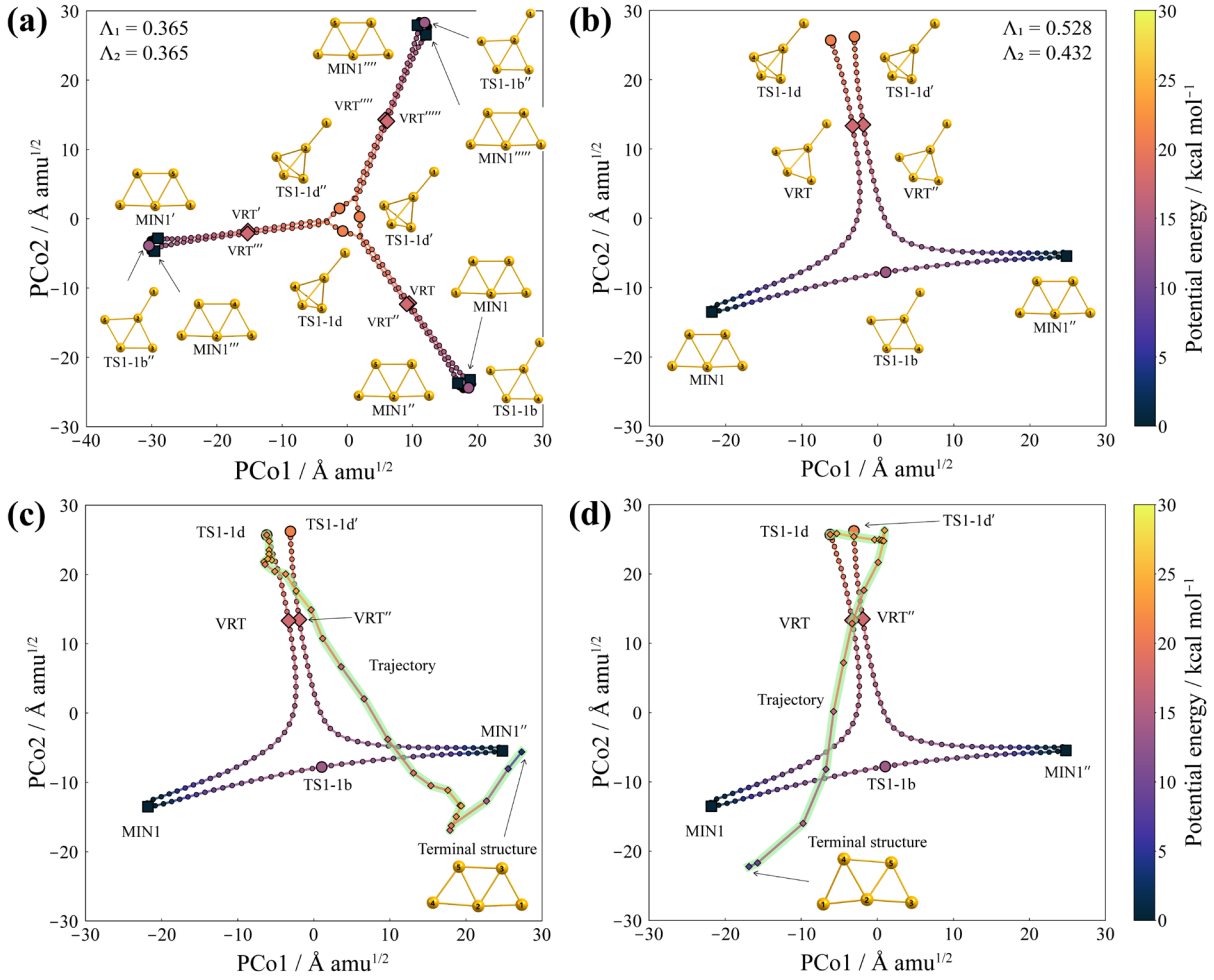


Figure 6. The IRCs and TOF-MD trajectories for Au_5 structural transformation projected on the two-dimensional principal coordinate subspace determined by CMDS and oCMDS analyses where the NPI isomers are distinguished, and TSs, MINs, and VRT are denoted by circles, squares, and diamonds, respectively, while the structures along the IRCs and trajectories (highlighted by light green) are denoted by smaller circles and smaller diamonds, respectively; the color shows the potential energy relative to the energy of MIN1: (a) Three IRCs for MIN1 → TS1-1d → MIN1 and three IRCs for MIN1 → TS1-1b → MIN1; (b) Two IRCs for TS1-1d → MIN1 and one IRC for MIN1 → TS1-1b → MIN1; (c), (d) TOF-MD trajectory projected on the two-dimensional principal coordinate subspace determined by the IRC network.

On the basis of the static reaction path network in Fig. 6b, we employed oCMDS to project dynamical trajectories on this two-dimensional principal coordinate subspace. Figures 6c and 6d show two examples of the trajectories departing from TS1-1d toward MIN1; in case c, the trajectory shows an IRC-jump before reaching VRT, reaching MIN1" (conventional bifurcating mechanism), while in case d, the trajectory initially jumps to the side of TS1-1d', but returns to the original IRC around the VRT", reaching MIN1. In both cases, the dynamic reaction routes include an IRC-jump process, which was introduced in previous study.[12]

Next, we discuss the dynamic reaction route based on the reduced principal coordinate subspace determined by the static reaction path network in which NPI isomers are not distinguished. In our previous study,[18] we generated the two-dimensional principal coordinate subspace for the structural transformation of Au_5, in which five MINs and 14 TSs are included as reference structures in CMDS analysis (in total, 19 structures); the result is shown in **Fig. 7a** where MINs and TSs are denoted by circles and squares, respectively, and the IRC connection is represented by a straight line. In this study, we added the structures along the 14 IRCs in the reference structures (in total, 336 structures), and carried out CMDS analysis to determine the two-dimensional principal coordinates; the result is shown in Fig. 7b. In comparison with Fig. 7a, the relative positional relation of stationary points and the proportions are not so different in spite of much larger number reference structures. In Fig. 7b, the cumulated proportion for two principal coordinates is 0.642, and the IRC profiles toward MIN2, MIN3, MIN4, and MIN5 are well described but the IRCs around MIN1 are not clearly shown. To zoom in the low-energy region around MIN1, we picked up seven IRCs which are connected with MIN1 (passing TS1-1a, TS1-1b, TS1-1c, TS1-1d, TS1-2, TS1-3a, or TS1-3b), and carried out CMDS analysis for the reference structures chosen from these IRC points (in total, 260 structures). The results are shown in Fig. 7c (two-dimension) and Fig. 7d (three-dimension). The proportions for three principal coordinates are $\Lambda_1 = 0.385$, $\Lambda_2 = 0.269$, and $\Lambda_3 = 0.153$, and thus, the cumulated proportions for two- and three-dimensional map are 0.654 and 0.807, respectively. In comparison with Fig. 7b, Fig. 7c clearly represents the network structure of the IRCs in the low-energy region. Moreover,

considering the third principal coordinate PCo3 in Fig. 7d, the details of the IRC network structure are more clarified. This analysis indicates that one can define more appropriate reduced principal coordinate subspace for the target region by designing the reference structures.

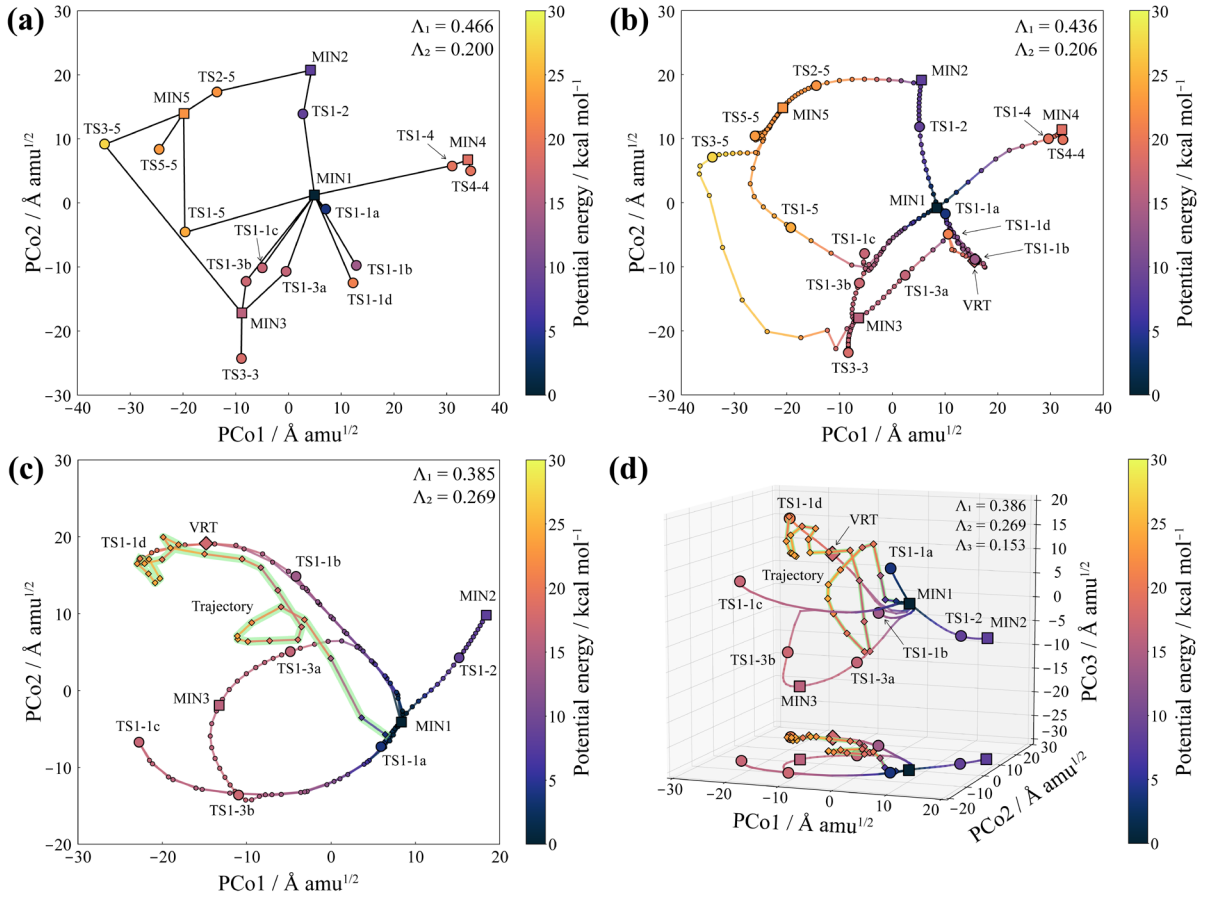


Figure 7. The stationary points, IRCs, and TOF-MD trajectory for Au_5 structural transformation projected on the two- or three-dimensional principal coordinate subspace determined by CMDS and oCMDS analyses where the NPI isomers are not distinguished, and TSs, MINs, and VRT are denoted by circles, squares, and diamonds, respectively, while the structures along the IRCs and trajectories (highlighted by light green) are denoted by smaller circles and smaller diamonds, respectively: (a) five MINs and 14 TSs in the two-dimensional subspace; (b) 336 reference structures along the 14 IRCs in the two-dimensional subspace; (c) 260 reference structures along the seven IRCs in the two-dimensional subspace, with one TOF-MD trajectory; (d) 260 reference structures along the seven IRCs in the three-dimensional subspace, with one TOF-MD trajectory. In (d), the two-dimensional projection on PCo1-PCo2 is also given at the bottom. The color shows the potential energy relative to the energy of MIN1.

Finally, we discuss a dynamic reaction route of the TOF-MD trajectory on the basis of the static reaction path network in a reduced dimension. The previous study based on the distance function[12] showed that, among 200 trajectories starting from TS1-1d toward MIN1, 62 trajectories initially approached to TS1-3a before reaching MIN1. To examine this behavior, we chose one trajectory which approaches to TS1-3a, as an example, and projected it in the two-dimensional subspace (Fig. 7c) and in the three-dimensional subspace (Fig. 7d) by the oCMDS approach. Figure 7c indicates that the trajectory proceeds along the IRC from TS1-1d to TS1-1b region, and deviates from the IRC and approaches to TS1-3a region; then, it wonders around TS1-3a, and approaches to MIN1. Figure 7d shows the dynamic reaction route in the three-dimensional subspace. In this figure, MIN3, TS1-3a, TS1-3b, MIN2, TS1-2, and TS1-1b are located in the negative PCo3 region, and the trajectory seems to leave the IRC after passing VRT, approaching to TS1-3a, and returns to the original IRC, reaching MIN1 via TS1-1a region. The present analysis reveals the dynamic reaction route on the static reaction path network in a reduced dimension.

5. CONCLUSION

The IRC is a uniquely-determined one-dimensional reaction path to describe an elementary reaction process from reactant to product minima *via* a first-order saddle point (= TS geometry) on the PES. The reaction path concept has played a significant role to understand the reaction mechanism, but the definition of IRC depends on only the geometrical feature of PES and does not consider the dynamics effect. However, dynamics often play an important role in the elementary reaction process, pushing the moving molecule out of the IRC and changing the reaction path to a non-IRC pathway. Such a dynamical picture can be analyzed by examining the dynamic reaction route of on-the-fly trajectories relative to the static IRC pathway, but the multidimensional nature of the coordinate space makes its analysis complex and difficult.

In this study, we propose a new approach to visualize a dynamic reaction route in the reduced-dimensional subspace determined for the IRCs, by employing the out-of-sample CMDS (oCMDS) method. As demonstration, the method was applied to the S_N2 reaction of $OH^- + CH_3F$ and the structural transformation of Au_5 cluster, which were investigated by CMDS in our previous study. For the former reaction, the CMDS analysis shows that the IRC is well described by two principal coordinates, and on-the-fly trajectories are mapped into this two-dimensional principal coordinate subspace. It is shown that the highly-curved region along the IRC works as a bifurcation point leading to two different paths ($CH_3OH + F^-$ dissociation and $CH_3OH\cdots F^-$ complex) in a dynamical sense, which coincides with a chemical intuition. For the Au_5 structural transformation, we carried out CMDS and oCMDS analyses in two different ways. First, we discussed the reaction path bifurcation case by distinguishing the NPI isomers, and confirmed several patterns of IRC-jump from one IRC to another IRC. Second, we generated the two- and three-dimensional principal coordinate subspace by CMDS for five MINs, 14 TSs, and succeeding structures along the 14 IRCs where the NPI isomers are not distinguished, and verified that the profile of two-dimensional network is very similar to the previous one determined for only five MINs and 14 TSs. The trajectory was projected in the two- and

three-dimensional principal coordinate subspaces by oCMDS, and was analyzed on the basis of the static reaction path network. The present oCMDS approach combined with the IRCs and TOF-MD simulation can be a robust tool to elucidate the significance of dynamics effect in chemical reaction process.

AUTHOR INFORMATION

Corresponding Author

*E-mail: take@sci.hokudai.ac.jp.

ORCID

Tetsuya Taketsugu: 0000-0002-1337-6694

Funding

TT (Tsutsumi) was supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT, Japan) through Program for Leading Graduate Schools (Hokkaido University "Ambitious Leader's Program") and Grant-in-Aid for Japan Society for the Promotion of Science (JSPS, Japan) Fellows Grant Number JP18J20856. This work was partly supported by MEXT as "Priority Issue on Post-K computer" (Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use), JSPS KAKENHI Grant Number JP18K18723, the Photo-excitonix Project in Hokkaido University, and JST CREST Grant Number JPMJCR1902, Japan. A part of calculations was performed using the Research Center for Computational Science, Okazaki, Japan.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Fukui, K. Formulation of the Reaction Coordinate. *J. Phys. Chem.* **1970**, *74*, 4161–4163.

(2) Maeda, S.; Ohno, K.; Morokuma, K. Systematic Exploration of the Mechanism of Chemical Reactions: The Global Reaction Route Mapping (GRRM) Strategy Using the ADDF and AFIR Methods. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683.

(3) Kato, S.; Morokuma, K. Potential Energy Characteristics and Energy Partitioning in Chemical Reactions: Ab Initio MO Study of Four-centered Elimination Reaction $CH_3CH_2F \rightarrow CH_2{=}CH_2{+}HF$. *J. Chem. Phys.* **1980**, *73*, 3900–3914.

(4) Taketsugu, T.; Gordon, M. S. Dynamic Reaction Path Analysis Based on an Intrinsic Reaction Coordinate. *J. Chem. Phys.* **1995**, *103*, 10042–10049.

(5) Taketsugu, T.; Gordon, M. S. Reaction Path Hamiltonian Based on a Reaction Coordinate and a Curvature Coordinate. *J. Chem. Phys.* **1996**, *104*, 2834–2840.

(6) Taketsugu, T.; Tajima, N.; Hirao, K. Approaches to Bifurcating Reaction Path. *J. Chem. Phys.* **1996**, *105*, 1933–1939.

(7) Yanai, T.; Taketsugu, T.; Hirao, K. Theoretical Study of Bifurcating Reaction Paths. *J. Chem. Phys.* **1997**, *107*, 1137–1146.

(8) Harabuchi, Y.; Ono, Y.; Maeda, S.; Taketsugu, T. Analyses of Bifurcation of Reaction Pathways on a Global Reaction Route Map: A Case Study of Gold Cluster Au_5. *J. Chem. Phys.* **2015**, *143*, 014301.

(9) Taketsugu, T.; Gordon, M. S. Dynamic Reaction Coordinate Analysis: An Application to $SiH_4 + H^- \rightarrow SiH_5^-$. *J. Phys. Chem.* **1995**, *99*, 8462–8471.

(10) Sun, K. S.; Hase, W. L. A S_N2 Reaction That Avoids Its Deep Potential Energy Minimum. *Science* **2002**, *296*, 875–878.

(11) Pratihar, S.; Ma, X.; Homayoon, Z.; Barnes, G. L.; Hase, W. L. Direct Chemical Dynamics Simulations. *J. Am. Chem. Soc.* **2017**, *139*, 3570–3590.

(12) Tsutsumi, T.; Harabuchi, Y.; Ono, Y.; Maeda, S.; Taketsugu, T. Analyses of Trajectory On-the-Fly Based on the Global Reaction Route Map. *Phys. Chem. Chem. Phys.* **2018**, *20*, 1364–1372.

(13) Olasz, B.; Czakó, G. Uncovering the Role of the Stationary Points in the Dynamics of the F^- + CH_3I Reaction. *Phys. Chem. Chem. Phys.* **2019**, *21*, 1578–1586.

(14) Vázquez, S. A.; Otero, X. L.; Martinez-Nunez, E. A Trajectory-Based Method to Explore Reaction Mechanisms. *Molecules* **2018**, *23*, 1–21.

(15) Cox, T.; Cox, M. In *Multidimensional Scaling, Second Edition*, 2nd Edit.; C&H/CRC Monographs on Statistics & Applied Probability; Chapman and Hall/CRC: New York, 2000; Vol. 88.

(16) Härdle, W. K.; Simar, L. In *Applied Multivariate Statistical Analysis*, 3rd Edit.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2015.

(17) Li, X.; Xie, Y.; Hu, D.; Lan, Z. Analysis of the Geometrical Evolution in On-the-Fly Surface-Hopping Nonadiabatic Dynamics with Machine Learning Dimensionality Reduction Approaches: Classical Multidimensional Scaling and Isometric Feature Mapping. *J. Chem. Theory Comput.* **2017**, *13*, 4611–4623.

(18) Tsutsumi, T.; Ono, Y.; Arai, Z.; Taketsugu, T. Visualization of the Intrinsic Reaction Coordinate and Global Reaction Route Map by Classical Multidimensional Scaling. *J. Chem. Theory Comput.* **2018**, *14*, 4263–4270.

(19) Trosset, M. W.; Priebe, C. E. The Out-of-Sample Problem for Classical Multidimensional Scaling. *Comput. Stat. Data Anal.* **2008**, *52*, 4635–4642.

(20) Pisani, P.; Caporuscio, F.; Carlino, L.; Rastelli, G. Molecular Dynamics Simulations and Classical Multidimensional Scaling Unveil New Metastable States in the Conformational Landscape of CDK2. *PLoS One* **2016**, *11*, 1–23.

(21) Kabsch, W. A Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallogr. Sect. A* **1976**, *32*, 922–923.

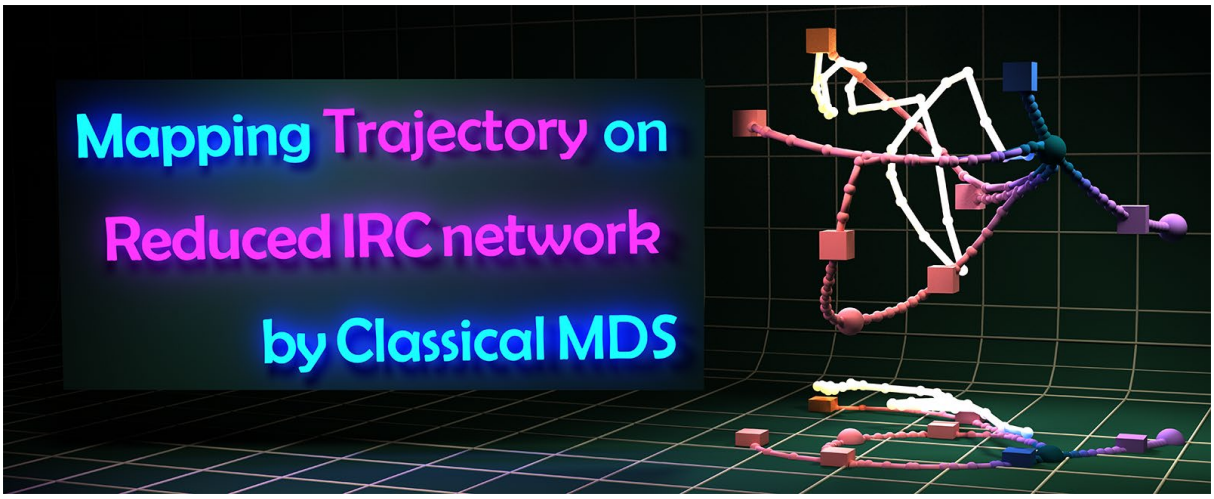
(22) Young, G.; Householder, A. S. Discussion of a Set of Points in Terms of Their Mutual Distances. *Psychometrika* **1938**, *3*, 19–22.

(23) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(24) Harabuchi, Y.; Okai, M.; Yamamoto, R.; Tsutsumi, T.; Ono, Y.; Taketsugu, T. SPPR. Hokkaido University: Sapporo, Japan 2020.

(25) Zou, W.; Sexton, T.; Kraka, E.; Freindorf, M.; Cremer, D. A New Method for Describing the Mechanism of a Chemical Reaction Based on the Unified Reaction Valley Approach. *J. Chem. Theory Comput.* **2016**, *12*, 650–663.

(26) Powell, M. J. D. An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives. *Comput. J.* **1964**.

(27) Hare, S. R.; Bratholm, L. A.; Glowacki, D. R.; Carpenter, B. K. Low Dimensional Representations along Intrinsic Reaction Coordinates and Molecular Dynamics Trajectories Using Interatomic Distance Matrices. *Chem. Sci.* **2019**.

For Table of Contents Only