Implementation and Performance of the Artificial Force Induced Reaction Method in the GRRM17 Program

Satoshi Maeda,[a] Yu Harabuchi,[a,b] Makito Takagi,[c] Kenichiro Saita,[a] Kimichi Suzuki,[a,d] Tomoya Ichino,[a] Yosuke Sumiya,[c] Kanami Sugiyama,[c] and Yuriko Ono[a]

This article reports implementation and performance of the artificial force induced reaction (AFIR) method in the upcoming 2017 version of GRRM program (GRRM17). The AFIR method, which is one of automated reaction path search methods, induces geometrical deformations in a system by pushing or pulling fragments defined in the system by an artificial force. In GRRM17, three different algorithms, that is, multicomponent algorithm (MC-AFIR), single-component algorithm (SC-AFIR), and double-sphere algorithm (DS-AFIR), are available, where the MC-AFIR was the only algorithm which has been available in the previous 2014 version. The MC-AFIR does automated sampling of reaction pathways between two or more reactant molecules. The SC-AFIR performs automated generation of global or semiglobal reaction path network. The DS-AFIR finds a single path between given two structures. Exploration of minimum energy structures within the hypersurface in which two different electronic states degenerate, and an interface with the quantum mechanics/molecular mechanics method, are also described. A code termed SAFIRE will also be available, as a visualization software for complicated reaction path networks. © 2017 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.25106

Introduction

Equilibrium (EQ) structures of molecules, clusters, organometallic complexes, and so forth, correspond to local minima on the adiabatic potential energy surface (PES). First-order saddle points on the PES have been calculated as transition state (TS) of chemical reactions.[1,2] Intrinsic reaction coordinate (IRC) of Fukui, that is, steepest descent path in the mass-weighted coordinates starting from TS, has been calculated as paths of chemical reactions.[3,4] The nonadiabatic transition takes place efficiently on the geometrical hyperspace in which two different electronic states degenerate. Minimum energy points within the crossing seam between their PESs, that is, minimum energy seam of crossing (MESX) or minimum energy conical intersection (MECI), have been calculated as critical points of a nonadiabatic transition,[5,6] where the seam for two electronic states having the same spin and space symmetry is called conical intersection.[7,8]

Geometry optimization has been implemented in most quantum chemistry program packages and widely used for obtaining these geometries.[9] In general, geometry optimization requires some guesses regarding the reaction mechanism. Only with geometry optimization, it has been hard to study mechanisms involving so many reaction steps or to predict mechanisms from scratch. To overcome this problem, methods that can find these structures systematically have been strongly desired.

Considerable efforts have been devoted to the development of such methods.[9–34] We have developed automated reaction path search methods.[35–47] Among various automated reaction path search methods, the artificial force induced reaction (AFIR) method would be the one which has been applied most extensively to organic reactions.[47] In addition, the anharmonic downward distortion following (ADDF) and AFIR methods are only methods that have been applied to automated exploration of MESX and MECI geometries,[45,47] and only the AFIR method has been applied to practical systems consisting of more than 30 atoms. Our methods have been implemented in the global reaction route mapping (GRRM) program.[48] Originally, the ADDF method was the only method available in the GRRM program.[35–37] In an update made in 2014, the AFIR method was partly implemented.

[a] S. Maeda, Y. Harabuchi, K. Saita, K. Suzuki, T. Ichino, Y. Ono
Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan
E-mail: smaeda@eis.hokudai.ac.jp
[b] Y. Harabuchi
Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Saitama 332-0012, Japan
[c] M. Takagi, Y. Sumiya, K. Sugiyama
Graduate School of Chemical Sciences and Engineering, Hokkaido University, Sapporo 060-8628, Japan
[d] K. Suzuki
Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103 Japan

Contract grant sponsor: JST-CREST; Contract grant number: JPMJCR14L5; Contract grant sponsor: JST-ACT-C; Contract grant number: JPMJCR12YN; Contract grant sponsor: JST-PRESTO (to Y.H.); Contract grant number: JPMJPR16N8; Contract grant sponsor: JSPS Research Fellowships for Young Scientists (DC1; to Y.S.); Contract grant sponsor: MEXT-FRAGSHIP2020 (the priority study 5; to Y.O.)

However, only the multicomponent algorithm (MC-AFIR) has been available in the 2014 version (GRRM14).[40,41]

In this article, we report full implementation of the AFIR method for its application to molecular systems. In the upcoming 2017 version (GRRM17) released soon, in addition to MC-AFIR, the single-component algorithm (SC-AFIR) and the double-end algorithm termed double-sphere AFIR (DS-AFIR) will be available. MC-AFIR is an algorithm for systematic sampling of pathways through which two or more molecules react together, SC-AFIR is for systematic exploration of pathways that occur in a given system, and DS-AFIR is for calculation of a single path connecting a given pair of reactant and product structures. These new features are described with simple application examples. In addition, a software for visualizing complicated reaction path networks generated by SC-AFIR is introduced.

AFIR Method

AFIR function

The AFIR method induces various structural deformations by pushing fragments A and B together or by pulling them apart.[40] The artificial force is applied to the target system by the following AFIR function.

$$F^{\mathrm{AFIR}}(\mathbf{Q}) = E(\mathbf{Q}) + \rho\alpha \frac{\sum_{i\in \mathrm{A}} \sum_{j\in \mathrm{B}} \omega_{ij} r_{ij}}{\sum_{i\in \mathrm{A}} \sum_{j\in \mathrm{B}} \omega_{ij}} \quad (1)$$

In eq. (1), $E(\mathbf{Q})$ is the PES of geometrical parameters $\mathbf{Q}$ and the second term applies the artificial force to the system. The constant parameter α defines the strength of the force, and ρ is set to either 1.0 or -1.0. The second term is given as a weighted sum of the interatomic distance r_{ij} between atoms i and j, and the weight ω_{ij} is defined as,

$$\omega_{ij} = \left[\frac{R_i + R_j}{r_{ij}}\right]^6 \quad (2)$$

where R_i and R_j are the covalent radii of atoms i and j, respectively. The force parameter α is obtained using eq. (3).

$$\alpha = \frac{\gamma}{\left[2^{-\frac{1}{6}} - \left(1 + \sqrt{1 + \frac{\gamma}{\varepsilon}}\right)^{-\frac{1}{6}}\right] R_0} \quad (3)$$

In this equation, R_0 and ε are parameters of the Lennard–Jones potential for Ar–Ar interaction and set to the standard value, that is, 3.8164 Å and 1.0061 kJ/mol, respectively. The α of eq. (3) corresponds to the mean force that an Ar–Ar pair feels while it moves from the minimum point to the turning point in their direct collision with collision energy γ. The model collision energy parameter γ, which provides an approximate upper limit of the barrier height that the system can overcome by the AFIR method, is defined by the user.

In application to thermal reactions, the value of γ is often decided using the following equation,

$$\gamma = -RT\ln\left(\frac{h}{tk_{\mathrm{B}}T}\right) \quad (4)$$

where R, T, h, t, and k_{B} are the gas constant, temperature, the Plank constant, reaction time, and the Boltzmann constant, respectively. Values of t and T depend on reaction conditions. To allow for paths having somewhat high barriers, t should be set to a larger value than the actual reaction time t_{actual}, for example, $t = 10t_{\mathrm{actual}}$.

AFIR path

The path through which the system passes during the minimization of the AFIR function is called AFIR path. The AFIR path changes depending on the following two conditions: the initial structure and the fragment pair A and B. In the MC-AFIR algorithm, different AFIR paths are obtained by calculating them starting from different initial structures. Conversely, the SC-AFIR algorithm finds different AFIR paths by calculating them for different fragment pairs. More details of these two algorithms are described in the corresponding sections.

Approximate EQs and TSs for $E(\mathbf{Q})$ can be obtained as local minima and maxima, respectively, along the AFIR path. To improve the accuracy, relaxation of the AFIR path may be made before optimization of TSs. The locally updated planes (LUP) method is implemented in GRRM17. In the LUP method,[49] path points are evenly distributed along a given path, and moved them to lower energy points in the hyperplane perpendicular to the path tangent, where the two end points and all local maxima are optimized directly to minima and maxima, respectively, without the constraint to the hyperplane.[50] In default, the path point relaxation is done ten times with a relatively large path point interval (1.0 Å) and then five times with a small path point interval (0.5 Å). Along the resulting LUP path, better approximate EQs and TSs are obtained as local minima and maxima, respectively. Then, approximate EQs and TSs either on the AFIR path or on the LUP path are optimized to actual EQs and TSs by any standard quasi-Newton method. Finally, IRC calculations are made starting from all obtained TSs. As long as an option which disables calculation of exact Hessian is not applied, normal mode analysis is done at all obtained EQs and TSs.

In GRRM17, minimization of either $E(\mathbf{Q})$ or $F(\mathbf{Q})$ and TS optimization are done by hybrid rational function optimization and trust radius methods.[51,52] Steepest descent path calculations either in mass-weighted or nonmass-weighted coordinates are done by the local quadratic approximation (LQA) method.[53] Exact Hessian matrix, if available in reasonable costs, will accelerate the search. In default, exact Hessian is calculated once in fifty steps in minimization, once in five steps in TS optimization, once in ten steps in steepest descent path calculation, where these settings can be adjusted by users. In steps in which exact Hessian is not calculated, a model Hessian obtained by a Hessian updating method is used, where Farkas and Schlegel's method is used in minimization and Bofill's method in TS optimization and steepest descent path calculation.[54,55]

Table 1. SC-AFIR options that specify paths to be computed.[a]

	Default	EQOnly	EQOnly&KeepSCPath	EQOnly&KeepLUPPath
EQ	YES	YES	YES	YES
AFIR path	NO[b]	NO[b]	YES	NO[b]
LUP path	YES	NO	NO	YES
IRC path	YES	NO	NO	NO

[a] The AFIR path is a crude approximation of the IRC paths, and the LUP path is the better approximation obtained by refining the AFIR path by the LUP method (see text). [b] AFIR paths are calculated but do not appear in main output files.

In MC-AFIR calculations, a set of AFIR paths are searched first with the algorithm described in the next section. Then, all obtained AFIR paths are read and a "RePATH" calculation is executed. The "RePATH" calculation automatically applies all the other procedures, that is, LUP path optimization, EQ optimization, TS optimization, and IRC calculation, to all paths (either AFIR path, LUP path, or IRC path) that are read from previous results. Obtained cartesian coordinates and energies of EQs and TSs are shown in list files, that is, EQ-list and TS-list, respectively. After IRC calculation, labels (EQ numbers) of two EQs that are linked by the IRC path appear below the corresponding TS in the TS-list.

In SC-AFIR calculations, all the above procedures, that is, AFIR path calculation, LUP path optimization, EQ optimization, TS optimization, and IRC calculation, are applied automatically in default. After completion of the calculation, one can draw the IRC path network using data in EQ-list and TS-list. Conversely, in many practical applications, some of these procedures may be skipped. When a keyword "EQOnly" is provided, LUP path optimization, TS optimization, and IRC calculation are skipped. This calculation gives many EQs including the global minimum quickly. When a keyword "KeepSCPath" is provided together with "EQOnly," LUP path optimization, TS optimization, and IRC calculation are skipped but all AFIR paths are stored on disk. This calculation gives the AFIR path network, in the same computational costs as those required in the "EQOnly" calculation. The corresponding IRC path network can be obtained afterward by applying the "RePATH" calculation to the AFIR path network. When a keyword "KeepLUPPath" is provided together with "EQOnly," only TS optimization and IRC calculation are skipped, and all LUP paths are stored on disk to generate the LUP path network. The corresponding IRC path network can also be obtained afterward by applying the "RePATH" calculation to the LUP path network. Paths that are obtained by the SC-AFIR search with these options are summarized in Table 1. For example, one may perform the SC-AFIR calculation with "EQOnly" and "KeepSCPath" keywords at the B3LYP/6-31G level, and then do the "RePATH" calculation at the B3LYP/6-311G** level, to reduce the total computational cost obtaining the IRC path network at the final computational level (B3LYP/6-311G** in this case).

It would be beneficial to users to note here that the "RePATH" calculation or the SC-AFIR calculation without the "EQOnly" option provides both the LUP path network and the IRC path network. The difference in the connectivity among EQs between the LUP path network and the IRC path network suggests cases in which TS optimization failed. Further application of the "RePATH" calculation with "PTOnly" keyword to the results of the previous SC-AFIR or "RePATH" calculation applies LUP path optimization, EQ optimization, TS optimization, and IRC calculation, only to LUP paths for which IRC paths of the corresponding connection are missing and those having the LUP top lower in energy than actual TS of the corresponding connection. It is also noted that the "RePATH" calculation can be applied to the IRC path network. One can apply the "RePATH" calculation to the IRC path network of the lower computational level to obtain the IRC path network of the higher computational level.

MC-AFIR Algorithm

In the MC-AFIR algorithm,[41] AFIR paths are sampled starting from various initial structures, where the fragments A and B are fixed to given ones throughout. The set of initial structures are generated automatically by giving mutual orientations and positions of two or more inputted molecules randomly. Optionally, the structure of each molecule can be read from a EQ-list prepared separately, where one of structures in the EQ-list is chosen randomly from the EQ-list and used as the corresponding part of the initial structure. In addition, there is an option with which initial structures are read from a structure list prepared by the user.

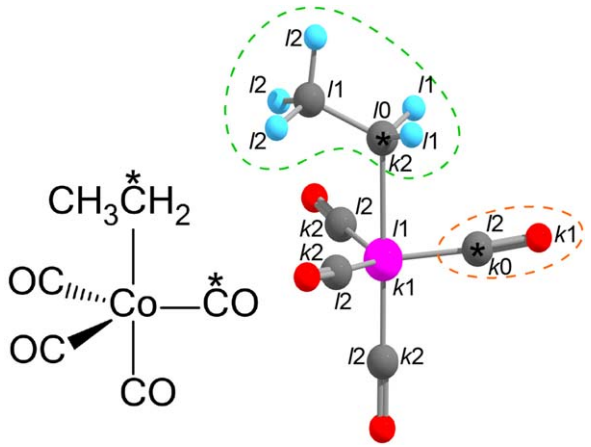
In cases where initial structures are randomly generated, two stopping criteria are available. In one criterion, the calculation is terminated when the last N_{fault} AFIR paths do not give any new product, where N_{fault} is a parameter defined by the user. In the other criterion, the calculation is terminated when N_{sample} AFIR paths are calculated. When initial structures are prepared by the user, the calculation is terminated once AFIR paths from all the prepared initial structures are finished.

The γ value in eq. (3) is set to a given value γ_{start} or to a random value between γ_{max} and γ_{min}, and gradually increased to γ_{max}. These AFIR path calculations can be done in parallel when many CPU-cores are available. MC-AFIR has been available in the previous version GRRM14, and its benchmarking is not made in this article which mainly focuses on new features in the upcoming GRRM17.

SC-AFIR Algorithm

Fragment generation

In the SC-AFIR algorithm,[44] AFIR paths are calculated starting from EQs. The EQ is not necessarily a single molecule, but also H-bond clusters, metal clusters, van der Waals complexes,



Scheme 1. Fragment generation around carbon atoms with asterisk marks. [Color figure can be viewed at wileyonlinelibrary.com]

organometallic complexes, and so forth. From a single EQ, many different AFIR paths for different fragment pairs are searched. At each EQ structure, fragment pairs are defined systematically by the following algorithm.

Focusing on two atoms k and l, a perturbed structure is generated by the following two steps: the distance of the two atoms is reduced by ζ% or increased by $\zeta + 30$% (in default, $\zeta = 20$), and then with fixing the k-l distance the positions of all atoms are optimized by the least squares fitting in which the root-mean-square error between elements in the distance matrix of the original structure and those of the perturbed structure is minimized. In the perturbed structure, fragments A and B are defined around the atoms k and l. First, k and l are included in A and B, respectively. Then, all atoms connected to A or B are included in A or B, respectively, where atoms i and j are regarded to be connected when $r_{ij}/(R_i + R_j) < 1.5$. This selection is done twice, and atoms directly connected to k and l (first layer) and those connected to them (second layer) are included in A and B, respectively. Finally, atoms i in A and j in B are excluded from A and B, respectively, when $r_{ij}/(R_i + R_j) < r_{kl}/(R_k + R_l)$.

In Scheme 1, a schematic which illustrates the fragment generation is shown for an alkyl-Co-carbonyl complex, where the carbon atoms indicated by asterisk marks are the two atoms k and l. The atom k itself, atoms connected to the atom k (first layer), and those connected to the first layer atoms (second layer) are indicated by $k0$, $k1$, and $k2$, respectively. Likewise, those for the atom l are indicated by $l0$, $l1$, and $l2$, respectively. Excluding atoms ki ($i > 0$) and lj ($j > 0$) that exist in between $k0$ and $l0$, the fragments indicated by dashed circles are obtained. These fragments correspond to the carbonyl and alkyl ligands in this organometallic complex, and minimization of eq. (1) with these fragments as A and B induces a reaction that connects these two ligands.

A series of fragment pairs are generated by applying this procedure to all atom pairs except for those with very long distances $r_{ij}/(R_i + R_j) > 8.0$. Then, AFIR paths for all the fragment pairs are calculated. This indicates that the computational costs required to complete the search around a single EQ increases in proportion to N^2, where N is the number of atoms. However, the actual cost increase is much smaller. At each minimization step of the AFIR function, the value of the second term in eq. (1) is compared to the value at the perturbed structure. The AFIR path is regarded to be "not reactive" at a given γ and just discarded, if its value at 10 minimization steps is larger than the value at the perturbed structure. This treatment allows one to avoid calculations of many AFIR paths that are not reactive at a given γ, and reduces the total computational cost. The other criteria to terminate the AFIR path integration in the SC-AFIR are: when a local minimum on the AFIR function is reached, when two energy maxima are passed along the AFIR path, or when the system decomposed into dissociated fragments (see the subsection "dissociation channel"). Different AFIR paths can also be calculated in parallel.

Difficult case and its solution

We point out here paths that are not found just by applying the artificial force to a single fragment pair, and present our way to overcome this problem. The AFIR function which imposes the artificial force only to a single fragment pair works well even in cases where two or more bonds rearrange synchronously between the fragments. For example, we showed in our early study that the AFIR method described well a path of Diels-Alder reaction in which two C—C bonds generated synchronously.[40] The AFIR function also works in cases where several bonds rearrange spontaneously following a single trigger event. For example, the AFIR method successfully gave the path of Claisen rearrangement; a C—O bond dissociation occurred spontaneously after a C—C bond formation induced by an artificial force applied to a fragment pair that did not include the dissociating C—O atoms.[44] Conversely, there exist cases in which bond formations and/or breakings occur in two or more sites that do not have strong causality, and in such cases, additional force terms are needed as the third or latter terms of eq. (1).

To deal with such cases systematically, a simple scheme is introduced. In this scheme, the terminal point of the AFIR path is stored as the rank-2 dummy EQ (r2d-EQ) when the path meets the following three conditions simultaneously: (I) there is no energy maximum point along the resulting LUP path, (II) the terminal point of the AFIR path corresponds to a local minimum on the AFIR function, and (III) there is at least one bond generated or broken through the AFIR path, where atom pairs with $r_{ij} < 1.25 \times (R_i + R_j)$ are regarded to be bonded in this judgement. Information of the fragment pair A and B that gave the r2d-EQ through the AFIR path is stored together with the geometry of r2d-EQ. Then, the SC-AFIR procedure is applied not only to EQs but also to r2d-EQs, where the force term which yielded the corresponding r2d-EQ was added to eq. (1) as the second force term, in addition to the first force term for the fragment pair defined at the r2d-EQ. The α values of the first and second force terms are scaled by 1.0 and 0.5, respectively. With this treatment, two independent artificial forces are applied to the system. Dummy EQs

generated in the same way from a rank-($n-1$) dummy EQ are called rank-n dummy EQ, and the maximum rank n_{MR} (at most 10 in GRRM17) can be specified by the user, where $n_{MR}=1$ (generation of dummy EQs is disabled) is the default.

A search with the rank-n_{MR} algorithm (with $n_{MR}>1$) is denoted as SC-AFIRn_{MR}. In general, the number of combinations of choosing $2n_{MR}$ atoms from N atoms increases in proportion to $N^{2n_{MR}}$. However, the actual cost increase is much less. Although, in a 10 atom system, the cost of SC-AFIR2 was 100 times larger than SC-AFIR1 if the cost increased in proportion to $N^{2n_{MR}}$, the actual cost increase was about six times in the example of C_4H_6 shown below. This is because AFIR paths that fulfill the above three conditions (I)–(III) simultaneously are limited, and most AFIR paths do not generate the dummy EQ. A SC-AFIRn_{MR} search with $n_{MR}>1$ is required only when one needs paths of high barriers, as discussed further in the subsection "Global search including high energy regions" with a practical example.

How to expand the path network

Each path can be distinguished by three numbers i, j, and ρ; a path from EQi obtained by applying the artificial force to the j-th fragment pair with positive or negative ($\rho=1$ or -1) sign of eq. (1). During the search, the next path to be calculated is chosen by a simple stochastic algorithm, where the same path is not calculated twice or more. In this algorithm, the following parameter μ_i is computed for all EQs in the EQ-list.

$$\mu_i=\xi_i\frac{\exp\left[-\Delta G_i/RT_R\right]}{\sum_j^{\text{all EQs}}\exp\left[-\Delta G_j/RT_R\right]} \quad (5)$$

Then, an EQ which has the maximum value is chosen. In eq. (5), ξ_i is a random number between 0 and 1 and ΔG_i is Gibbs free energy of EQi. The T_R is a model temperature parameter, which determines how frequently high energy EQs are chosen, is usually set to a much higher value than the corresponding experimental temperature (in default, $T_R=2981.5$ K). When an option to disable exact Hessian calculation is used, electronic energy ΔE_i is adopted instead of ΔG_i because ΔG_i is available only when exact Hessian is computed. Then, the combination of j and ρ is selected randomly. This path selection scheme controls the order of finding of EQs and TSs. It preferentially executes the search in low energy region, and EQs and TSs in such regions, therefore, tend to be found earlier than those in high energy regions. Conversely, it does not affect the final results when all paths are computed. In other words, the search by SC-AFIR is deterministic as long as one does not terminate the calculation before completion of the search.

Several options are available to limit the search areas. The "FirstOnly" option disables searches starting from EQi with $i\geq 1$, and the search is terminated after calculating all paths starting from EQ0 (the initial structure). The "NoBondRearrange" option disables searches starting from EQs having bond-connectivity different from EQ0. With the "Target" option, atoms that are considered as k or l in the fragment generation procedure can be specified (all atoms in the entire system in default), and this option is used when atoms in the reaction-center can be identified. With the "PriorityPath" option, paths for specified k-l combinations are calculated first, and it, therefore, is useful to find desired paths quickly when atom-pairs that form/dissociate bonds are known beforehand. With the "rTarget" option, one can apply different (usually large) γ only to the paths of atom pairs between specified reactive target atoms, where "Target" and "rTarget" can be used simultaneously.

Local search

If the SC-AFIR algorithm is applied to all obtained EQs with sufficiently large γ, an entire reaction path network of a given atomic composition can be obtained. However, only local areas that are relevant to a target reaction would be of interest for many users. In GRRM17, various options are available to restrict the search areas as aforementioned. These options are introduced here using an elementary reaction step in Co-catalyzed hydroformylation shown in Figure 1a,[56] where electronic structure calculations were made by B3LYP/6-31G method with Gaussian 09 program.[57] In default (without the "Target" option), all possible atom pairs are considered in the fragment generation procedure. With the "Target" option, the user can choose atoms that are considered in this procedure. The simplest case is choosing only two atoms, and an example is shown here for the case in which the C atom in the CH_2CH_3 group directly connected to the Co center and the C atom in one of equatorial CO ligands are chosen. In this case, only the path of CO insertion shown in Figure 1a was obtained and then the calculation finished. In this calculation, another keyword "FirstOnly" was also used, where this option disables application of the SC-AFIR procedure to EQs newly obtained in the search.

This test calculation was done with two different γ for benchmarking. The case with $\gamma=153$ kJ/mol assumes a reaction condition $T=130$°C and $t=10$ days ($t_{actual}=1$ day) in eq. (4). Another case with $\gamma=100$ kJ/mol is also shown to see the effect of the γ value to the resulting AFIR and LUP paths. The highest energy structures along AFIR and LUP paths (denoted top of AFIR path and top of LUP path, respectively) are listed in Figure 1b. These structures were all close enough to the actual TS, and TS optimization starting from these structures converged to the actual TS. Figure 1c compares the energy profile along the corresponding IRC path with those along AFIR and LUP paths obtained through the searches with different γ. The AFIR and LUP paths were calculated in the non-mass-weighted coordinate, and the horizontal axis was converted to the mass-weighed coordinate afterward. Comparing the two AFIR paths to each other, the path of the larger γ value is higher in energy at any point than the one of the lower γ. This is a general trend seen in the AFIR method, since the larger artificial force allows the system to pass the higher energy region. Path optimization by the LUP method improved the entire path, and energy values at the top of LUP paths are very close to the value at the actual TS (the top of IRC path). The deviation in positions of the top points between the LUP path and the IRC path along the horizontal axis is not

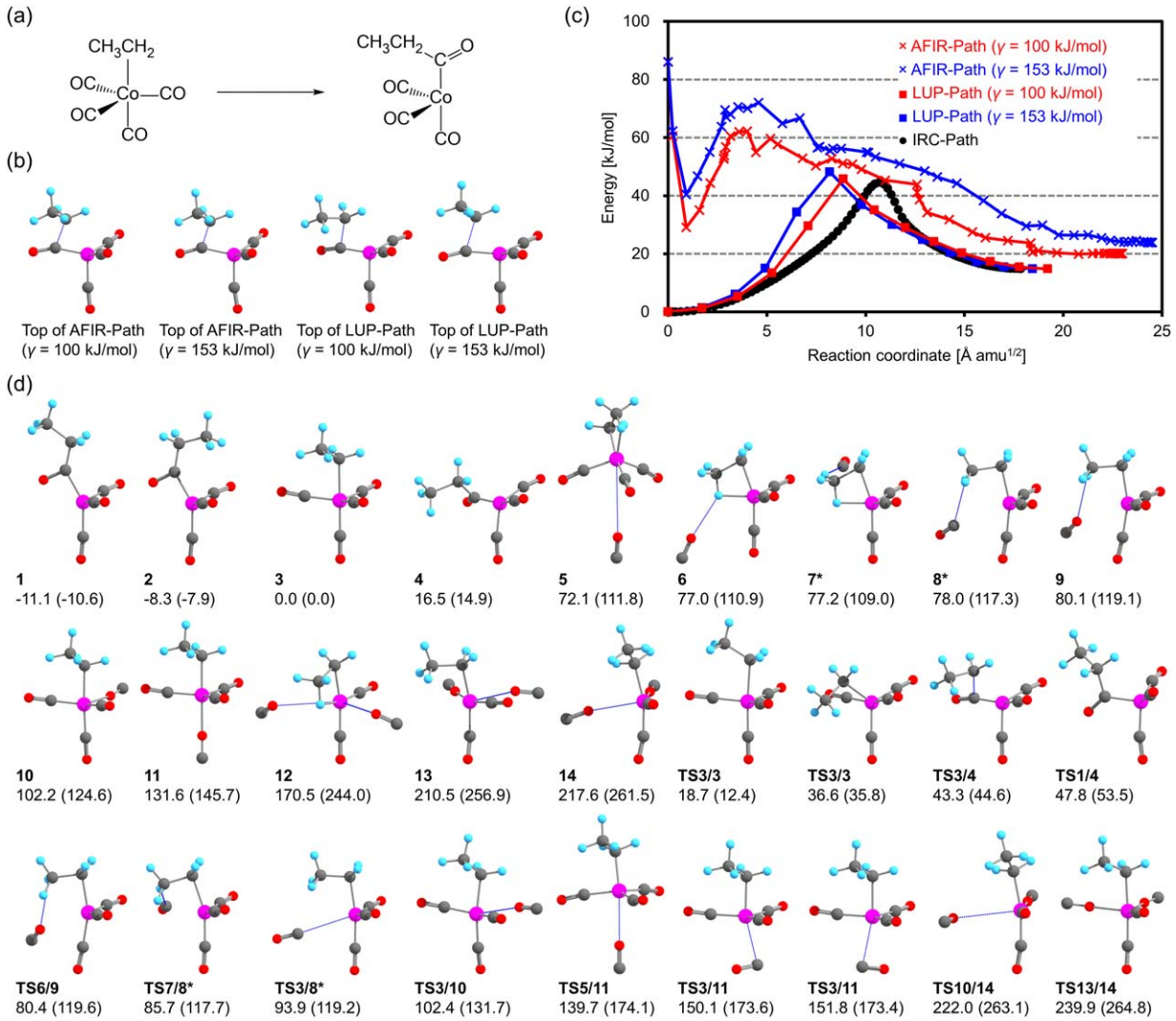


Figure 1. a) an elementary reaction step in Co-catalyzed hydroformylation, b) the highest energy structures along AFIR and LUP paths, c) an energy profile along the IRC path and those along AFIR and LUP paths obtained through the searches with different γ (see text), d) all obtained EQs and TSs by the (restricted) SC-AFIR search, where those obtained by the subsequent "RePATH" calculation are indicated by asterisk marks (see text). Relative Gibbs free energy (130°C, 1 atm) values and relative electronic energy values in parentheses are shown in kJ/mol below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

important, when considering the difference in the coordinate system adopted in the path integration.

A more systematic search can be done by executing the calculation with taking account of atom pairs among the Co center and all atoms connected to the Co center. In this case, not only chemical bond rearrangements but also pseudorotation pathways around the Co center and ligand dissociation paths are found. In this calculation, another keyword "NoBondRearrange" was also used, where this option disables application of the SC-AFIR procedure to EQs having different bond connectivity from the initial EQ. Figure 1d lists all obtained EQs and TSs through this search. Among these EQs, EQ3 is the initial EQ, and there is no other EQ that has the same bond connectivity to EQ3. Among these EQs and TSs, those shown without asterisk mark were found in the search. Some TSs that are expected to be found from existence of the corresponding LUP paths were missing. To find these TSs, the output files of the search calculation were read and a "RePATH" calculation was performed. In the "RePATH" calculation, the LUP method was applied to LUP paths connecting an EQ-EQ pair for which connection via TS is missing and also to those having the top lower in energy than actual TS of the corresponding connection. Such missing connections in the IRC path network are due to a failure in TS optimization from the top of LUP paths, where such a failure tends to occur for shallow TSs having the imaginary frequency of a small magnitude. After the "RePATH" calculation, EQs and TSs indicated with asterisk marks were obtained. Although the second "RePATH" calculation can further be done reading the outputs of the first "RePATH" calculation, it was not performed. This is because all connections related to the target EQ (EQ3 in this case as restricted by the "NoBondRearrange" option) and seen in the LUP path network have already been obtained in the IRC path network. Some EQs, which were not directly

Table 2. The number of atoms that are considered in the fragment generation N_{target}, the number of AFIR paths computed N_{path}, the value of model collision energy parameter adopted γ (in kJ/mol), the number of gradient calculations required $N_{gradient}$, and the number of Hessian calculations required $N_{Hessian}$, in the three SC-AFIR calculations with different setups (see text).

N_{target}	N_{path}	γ	$N_{gradient}$	$N_{Hessian}$
2	2	153	359	50
2	2	100	310	43
6	18	153	6687 (1911)[a]	792 (343)[a]

[a] The numbers in parentheses are those required in the subsequent "RePATH" calculation.

connected with EQ3 via IRC path, were obtained because of too strong artificial force.

The numbers of gradient and Hessian calculations required to complete each calculation discussed above are listed in Table 2. Comparing the former two cases, the smaller the total cost was the smaller γ was. This is a general trend of the AFIR method, although one should keep in mind that too small γ is not enough to induce the target reaction. The calculation in which atom pairs were selected from the six atoms in the fragment generation, required much more gradient and Hessian calculations. The subsequent "RePATH" that was performed to find shallow TSs, was only a fraction of cost as shown in

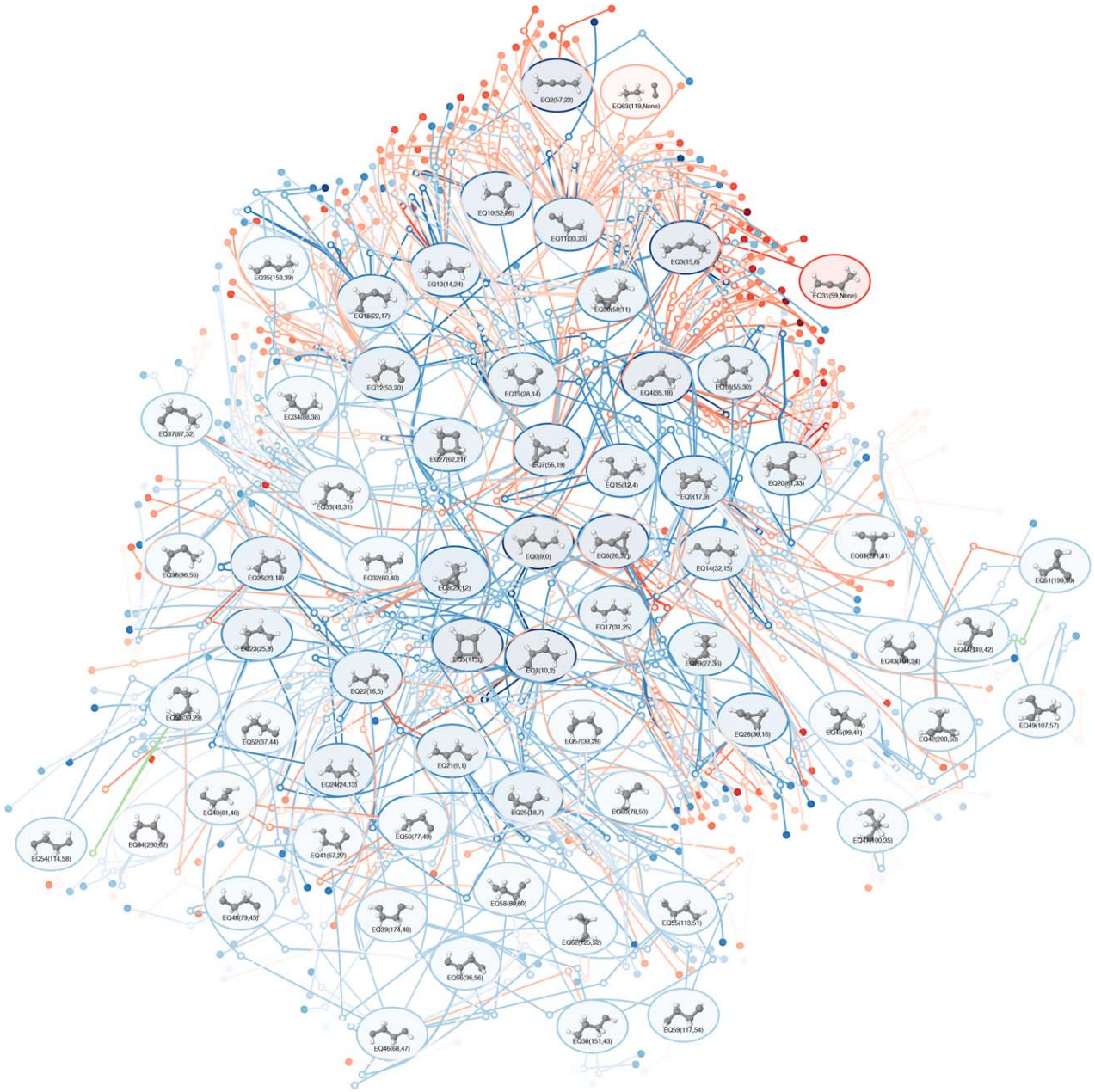


Figure 2. The global reaction path network of the atomic composition C_4H_6, where the part obtained by both the SC-AFIR1 and SC-AFIR2 searches are shown in blue, that obtained only in the SC-AFIR1 search in green, and that obtained only in the SC-AFIR2 search in red. Low and high energy regions are indicated by dark and light colors, respectively, where relative energy values in kJ/mol are shown below each EQ. EQs are indicated by ellipses in which the corresponding molecular structures are depicted. TSs and dissociated products are indicated by small open and filled circles, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

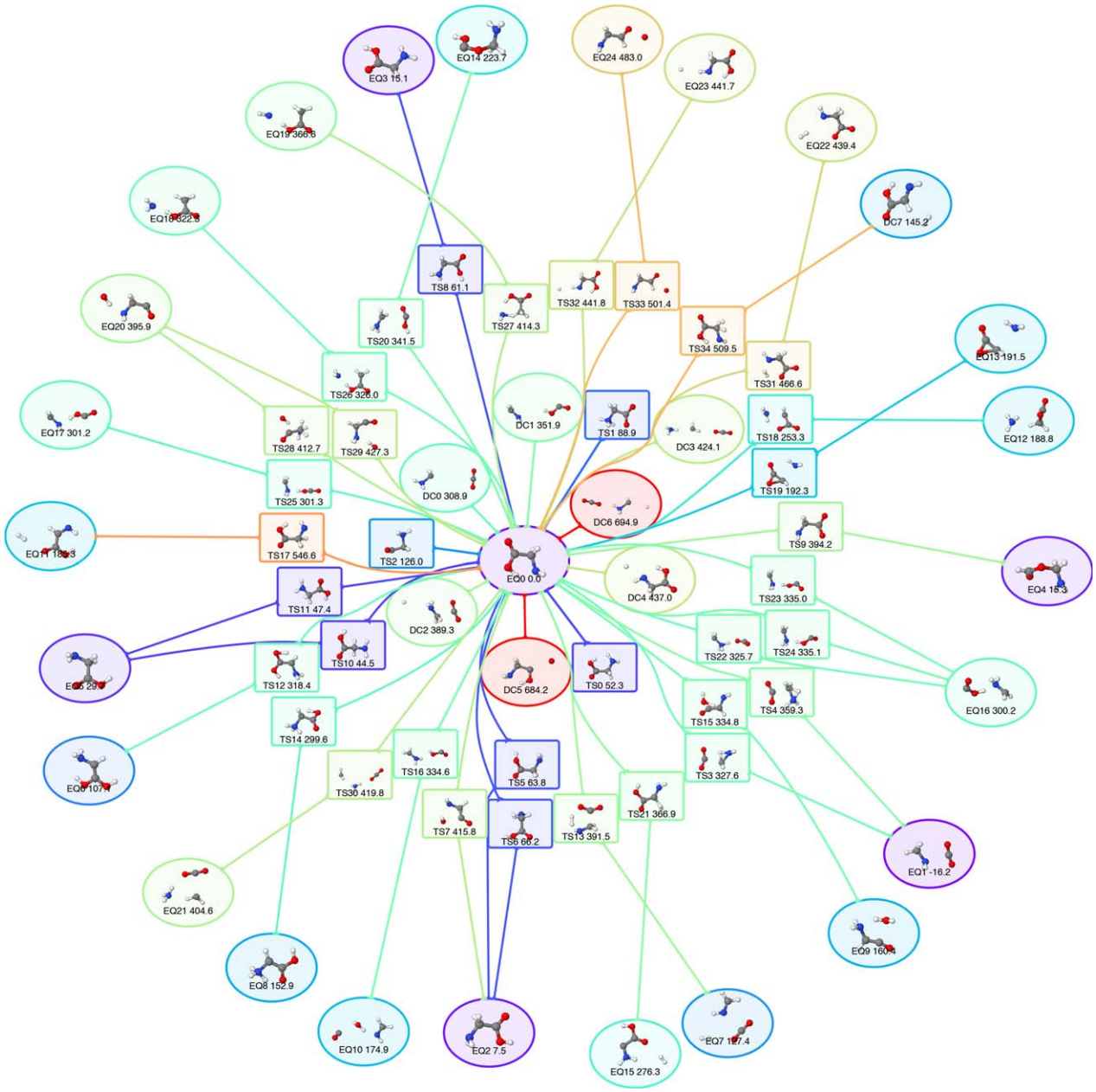


Figure 3. The reaction path network around EQ0, a conformer of glycine molecule. Although the search found many other TSs and DCs that are not connected to EQ0, only those directly connected to EQ0 are depicted for simplicity. Relative energy is depicted according to the rainbow spectrum, where the lowest energy EQ is indicated by purple. Relative energy values in kJ/mol are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

parentheses in Table 2 because there were only two paths having such TSs. As shown in this subsection, the user can change the search area and the total cost flexibly by choosing the atoms that are considered in the fragment generation, by changing the γ value, and/or by using the options such as "NoBondRearrange" and "FirstOnly" restricting EQs to which the SC-AFIR procedure is applied.

Global search

In small systems, an exhaustive search over the entire PES can be done by the SC-AFIR method. GRRM17 performs such a calculation, when a sufficiently large γ (e.g., $\gamma = 1000$ kJ/mol) is applied with neither specifying atoms that are considered in the fragment generation by "Target" nor using any EQ selection option such as "NoBondRearrange" and "FirstOnly." We show here a ten atom case for the atomic composition C_4H_6, at the RHF/6-31G level with Gaussian 09.[57] As the initial structure, trans-1,3-butadiene was chosen. Two different searches were performed with setting the maximum rank n_{MR} discussed in the "difficult cases and its solution" section to 1 or 2 (in default, $n_{MR} = 1$), where these two calculations are denoted as SC-AFIR1 and SC-AFIR2, respectively. Figure 2 shows the global reaction path network of this system, where the part obtained by both of the two different searches are shown in blue, that obtained only in the SC-AFIR1 search in green, and that

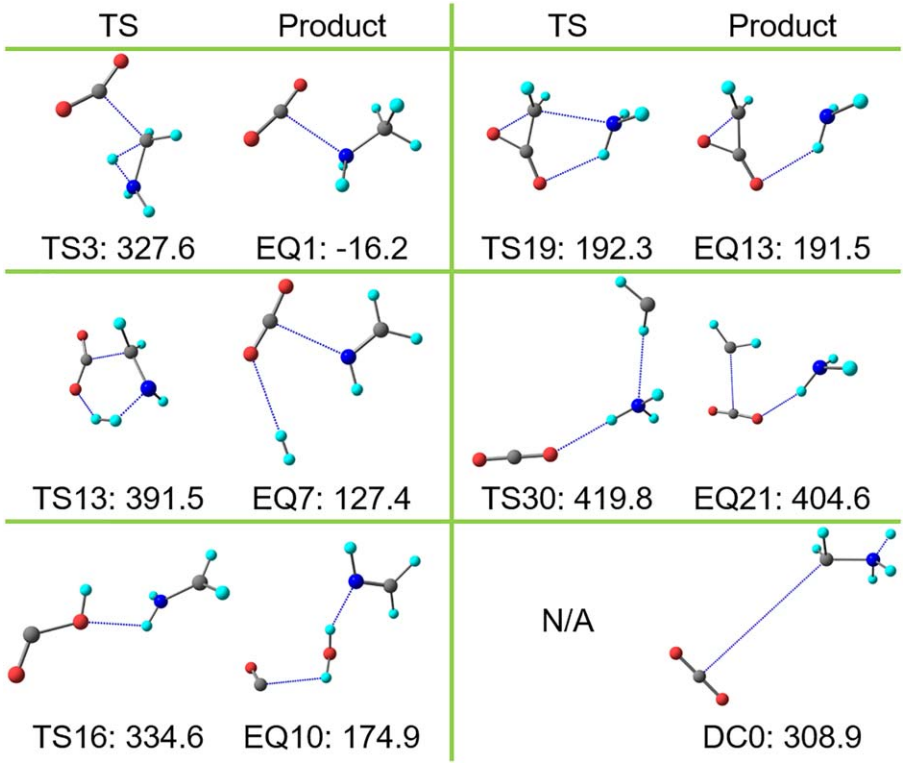


Figure 4. Six dissociation channels starting from EQ0, a conformer of glycine molecule. Relative energy values in kJ/mol are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

obtained only in the SC-AFIR2 search in red. Low and high energy regions are indicated by dark and light colors, respectively. In the path network, EQs are indicated by ellipses in which the corresponding molecular structures are depicted. TSs and dissociated products are indicated by small open and filled circles, respectively, where $H_2 + H_2C = CH\text{-}C{\equiv}CH$, $H_2C = CH_2 + HCCH$, etc., correspond to dissociated products (see the next subsection for the definition of dissociated products).

After the two calculations, 65/776 independent EQ/TS structures were found in total, where 575301/79408 and 3070732/402070 gradient/Hessian calculations were required to complete the SC-AFIR1 and SC-AFIR2 calculations, respectively. Comparing to the results of SC-AFIR2, two shallow EQs were missed by SC-AFIR1. One such EQ is EQ31, which is 370.7 kJ/mol less stable than trans-1,3-butadiene and easily isomerize into EQ3 (1,2-butadiene) through a tiny barrier of 1.5 kJ/mol. Another such EQ is EQ63, which is 716.5 kJ/mol less stable than trans-1,3-butadiene and also easily isomerize into EQ11 (CH_3CH_2CHC) through a tiny barrier of 0.6 kJ/mol. Such very shallow minima are difficult to find because the PES area from which geometry optimization converge to them is very narrow. The lowest TS which was missed in the SC-AFIR1 search is the one between EQ3 and EQ31, where this TS is not kinetically important considering the extreme shallowness of EQ31. The second lowest missed TS connecting different EQs is the one between EQ18 and EQ20 and higher in energy by 419.1 kJ/mol than trans-1,3-butadiene, and this TS again is not kinetically important because there is the other, lower energy TS connecting EQ18 and EQ20. Most TSs missed in the SC-AFIR1 search are high lying, kinetically unimportant ones. To our experiences, SC-AFIR1 is enough in many purposes such as analysis of organic reaction in mild condition. In case where high lying TSs play roles, such as combustion, SC-AFIRn_{MR} ($n_{MR} > 1$) is useful.

The SC-AFIR2 search found all EQs obtained by the SC-AFIR1 search. Conversely, there are two missing TSs in the network by SC-AFIR2 indicated by green in Figure 2. Through detailed analysis of all the output files, we found that these two TSs were found also in the SC-AFIR2 search. However, these two were regarded to be identical to the other TSs that were missed in the SC-AFIR1 search. This is because both of these two TSs are connected to a very shallow EQ and are very close to these EQs. These TSs were regarded to be identical to the other TS that is also connected to the shallow EQs and located very close to the EQs. We also confirmed that these two TSs can be found even in the SC-AFIR2 search by adopting the "tight" structure comparison criteria in GRRM17. We warn that the "tight" structure comparison criteria may generate many essentially identical but only slightly different structures, especially when the calculation is done with DFT and the density of DFT grid adopted is not high enough.

Dissociation channel

In many cases, dissociation channel (DC) is involved as one of important processes. In GRRM17, whether a structure is dissociated or not is judged by looking a bond connectivity matrix, where atom *i* and *j* are regarded to be connected if their

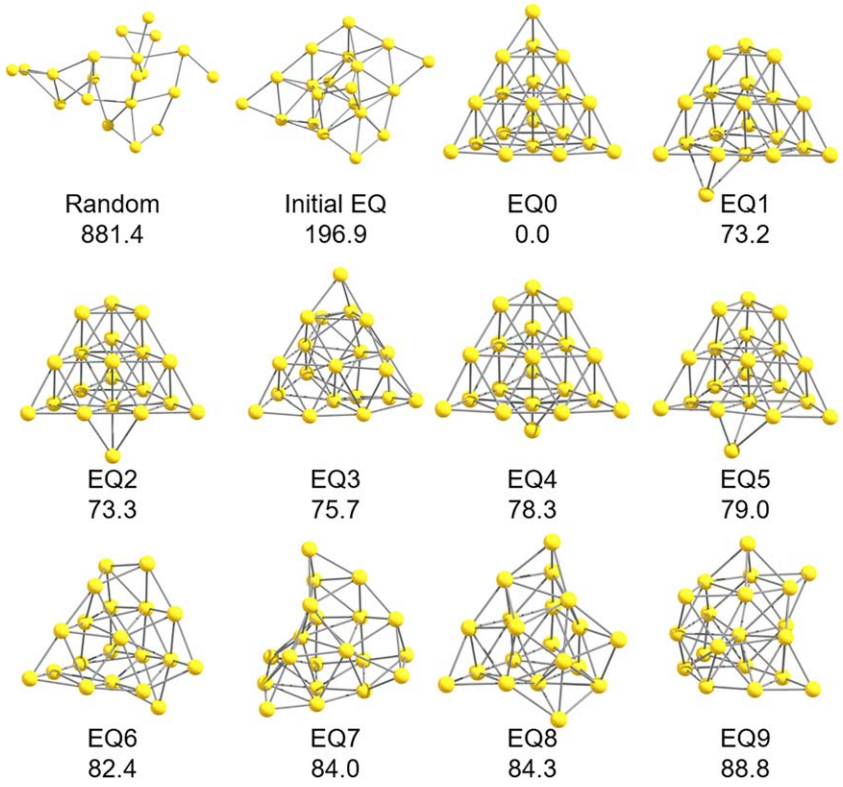


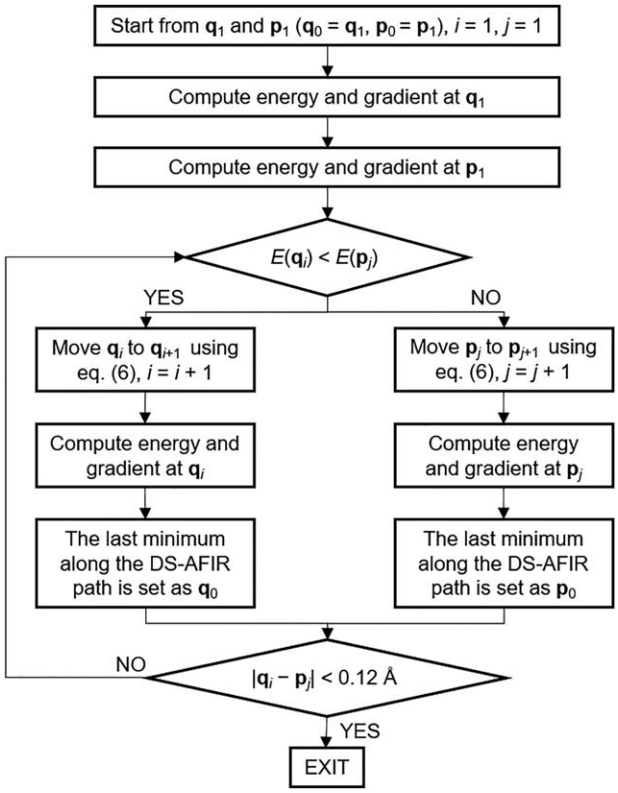
Figure 5. A random structure, the initial EQ obtained by optimizing the random structure, and the ten lowest energy EQs among 186 obtained by the EQ only search. Relative energy values in kJ/mol are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

distance is shorter than $2 \times n_D \times (R_i + R_j)$. In default, n_D is set to 8, and the search in the last section was done with this default setting. In this case, most of weak complexes are regarded to be dissociated. In Figure 2, EQ63, a complex between C_2 and C_2H_6, is the only complex obtained as EQ and the other weak complexes were regarded to be dissociated. If one intends to search PES areas of H-bond complexes, $n_D = 12$ is a recommended value. If PES areas for van der Waals complexes and/or those for the roaming pathways[6] are the research focus, $n_D = 15$ is a recommended value.

There are two types in DC. One is a type passing through a TS, and another is a type which does not have any peak along the energy profile. DCs of the former type are included in the reaction path network in Figure 2. Although SC-AFIR also searches for paths of DC of the latter type, in default AFIR paths for such DCs are not examined further. Conversely, AFIR paths for the latter type DCs are examined further if an option "SC = FindDC" is used. In this case, the terminal point of the AFIR path is optimized fixing the center positions of dissociated fragments, and the meta-IRC (steepest descent path starting from non-TS point)[4] is computed starting from the structure obtained by the constrained optimization. Then, the structure obtained by the constrained optimization is stored in the DC-list file with the path connection information, that is, the EQ to which the meta-IRC reached. Such DCs usually generate high energy species such as radical pairs and may not be very important in many purposes.

As a simple example, we performed an SC-AFIR2 search starting from one conformer of glycine molecule, with $n_D = 15$, $\gamma = 1000$ kJ/mol, and the "SC = FindDC" and "FirstOnly" options. The calculation was done at the UB3LYP/6–31G level with Gaussian 09,[57] where the stability check of molecular orbitals was performed during the search to detect biradical character in some structures. The calculation required 274972/30932 gradient/Hessian calculations. Figure 3 shows the reaction path network connected to the initial EQ (EQ0). Although the search found many other TSs and DCs that are not connected to EQ0 due to the too strong artificial force, only those directly connected to EQ0 are depicted in Figure 3 for simplicity. In Figure 3, various DCs of both of the aforementioned two types were found. Because of the large n_D (= 15) value, some weakly bound complexes among dissociated products were obtained as EQs in the output file.

We discuss here six DCs listed in Figure 4. Four different paths giving CO_2 are shown in Figure 4. Those through TS3, TS13, TS19, and TS30 gave weak complexes among simple molecules such as CO_2 and NH_3. Conversely, the path to DC0 directly gave $CO_2 + CH_2NH_3$ (ammonium ylide) without passing any TS. The meta-IRC starting from DC0 was confirmed (automatically during the search) to give EQ0 through a C—C bond generation and a subsequent proton transfer. It is also noted that DC0 is a structure obtained by constrained optimization keeping the center positions CO_2 and CH_2NH_3 starting from the final point of the AFIR path at which the structure was judged to be dissociated. The path through TS16 is so called roaming path,[6] which once gives a radical pair COOH—CH_2NH_2 and subsequently undergoes their recombination affording closed-shell molecules. We confirmed that TS16



Scheme 2. Calculation flow of DS-AFIR.

had diradical character through analysis of its electronic structure. The decomposition through TS19 giving ammonia and acetolactone is energetically the most favorable one among many bond reorganization paths in Figure 3.

EQ only search

When the focus is to find only EQs, an option "EQOnly" is used. With this option, LUP path optimization, TS optimization, and IRC calculation are disabled, and, therefore, only EQs are found. The stochastic EQ selection scheme using eq. (5) allows for preferential searching of low energy regions. An example of a systematic search of low-lying isomers of Au_{20} clusters is shown here. The calculation was done at RI-PBE/def2-SV(P) level using the TURBOMOLE 7.0 *ab initio* program package,[58] with $\gamma = 300$ kJ/mol and $T_R = 300$ K. The search was terminated when the lowest n_L EQs are not updated in the last p_L paths, where $n_L = 60$ and $p_L = 200$ were adopted. We usually set n_L and p_L to $3N_T$ and $10N_T$, respectively, where N_T is the number of target atoms specified by the "Target" option (all the 20 atoms in this case). The search was initiated from a random structure shown in Figure 5. Geometry optimization gave the initial EQ after 164 gradient calculations. Then, 157,984 gradient calculations were required before meeting the termination criteria. The search generated 186 EQs, and the lowest 10 are listed in Figure 5. The lowest one (EQ0) corresponds to the known global minimum with regular tetrahedron shape.[59] Most of low-lying ones are those with small deformation from the tetrahedron shape. Use of larger n_L, p_L, γ, and T_R values will provide more EQs performing the more comprehensive search.

DS-AFIR Algorithm

Another useful algorithm termed DS-AFIR is available in GRRM17. DS-AFIR can find a single path linking two given structures. As shown below, it can also be applied to multistep paths that have more than one TS. Like the other double-end methods,[49,50,60–64] the obtained path does not necessarily correspond to the kinetically most preferable one but to the shortest one, when multiple paths exist between given end points. A DS-AFIR calculation also yields EQ-list, TS-list, and AFIR or LUP path data in the same format to those the SC-AFIR calculation gives. The "RePATH" can also be applied, to find TSs along these paths with a different computational level or to automatically recover TSs that are missed in the initial application (if any).

In the DS-AFIR, the following function is minimized.

$$F^{\mathrm{DS-AFIR}}(\mathbf{q}_i) = E(\mathbf{q}_i) + XY|\mathbf{q}_i - \mathbf{p}_j| - X(1-Y)|\mathbf{q}_i - \mathbf{q}_0| \tag{6}$$

In eq. (6), $\mathbf{q}_i$ and $\mathbf{p}_j$ correspond to the current positions of the two end points, $\mathbf{q}_0$ is a reference point explained below. The parameter X is given as,

$$X = \frac{\delta}{|\mathbf{u}|} - \frac{\mathbf{g}_i \cdot \mathbf{u}}{|\mathbf{u}|^2} \tag{7}$$

where $\mathbf{g}_i$ is the PES gradient at $\mathbf{q}_i$ and δ is a parameter. The vector $\mathbf{u}$ is,

$$\mathbf{u} = Y\frac{(\mathbf{q}_i - \mathbf{p}_j)}{|\mathbf{q}_i - \mathbf{p}_j|} - (1-Y)\frac{(\mathbf{q}_i - \mathbf{q}_0)}{|\mathbf{q}_i - \mathbf{q}_0|}, \tag{8}$$

where the parameter Y is given as,

$$\begin{aligned} Y &= \frac{Z}{1+Z} \qquad \text{if } Z > 0 \\ Y &= 0 \qquad \text{if } Z \le 0, \end{aligned} \tag{9}$$

using the parameter Z defined below.

$$Z = \frac{|\mathbf{q}_i - \mathbf{q}_0|}{|\mathbf{q}_i - \mathbf{p}_j|} + \frac{(\mathbf{q}_i - \mathbf{p}_j) \cdot (\mathbf{q}_i - \mathbf{q}_0)}{|\mathbf{q}_i - \mathbf{p}_j||\mathbf{q}_i - \mathbf{q}_0|} \tag{10}$$

In eq. (6), there are two force terms. The first force term applies an artificial force that shortens the distance between the two end points. The other pushes the current point $\mathbf{q}_i$ away from the reference point $\mathbf{q}_0$, where $\mathbf{q}_0$ corresponds to the latest local minimum $\mathbf{q}_k$ along the AFIR path and initially set to the given end-point $\mathbf{q}_1$. The parameter X in eq. (7) puts a condition that the component of derivative of eq. (6) along the vector of the sum of two force terms becomes δ. Eq. (9) and (10) control the weight of the second force term to become larger than that of the first force term around local minima. Conversely, the first force term gets the larger weight around ridges between two minima. Such a way of the weight controlling is justified below from nature of the two force terms.

The AFIR path of eq. (6) is integrated by the steepest descent method, where the steepest descent path is integrated by the LQA method in GRRM17. The actual procedure

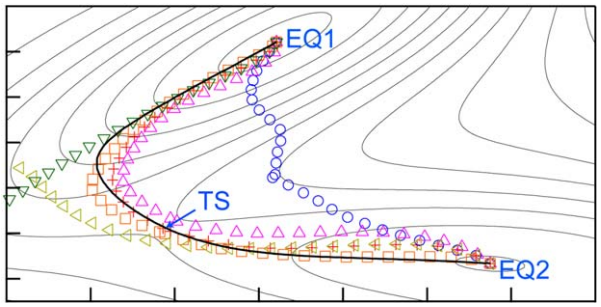


Figure 6. The IRC path (thick black line) and six paths of different settings of eq. (6) on the Müller–Brown potential. The path indicated by circles was obtained by integrating the AFIR path of eq. (6) with $\delta = 300.0$ and $Y = 1.0$ throughout. Paths indicated by inverse triangles and triangles facing sideways were obtained by integrating the AFIR path of eq. (6) with $\delta = 300.0$ and $Y = 0.0$ throughout, starting from EQ1 and EQ2, respectively. Paths indicated by triangles, crosses, and squares are those obtained by integrating the AFIR path of eq. (6) with $\delta = 100.0$, 300.0, and 500.0, respectively, and Y determined by eq. (9). [Color figure can be viewed at wileyonlinelibrary.com]

for constructing the DS-AFIR path is illustrated in Scheme 2. At first, energy and gradient are computed at the two given end points $\mathbf{q}_1$ and $\mathbf{p}_1$. At an arbitrary step, the one with lower energy is chosen out of $\mathbf{q}_i$ and $\mathbf{p}_j$ and moved to $\mathbf{q}_{i+1}$ or $\mathbf{p}_{j+1}$ by the LQA method on the function of eq. (6), where the second force term in eq. (6) is omitted at $\mathbf{q}_1$ and $\mathbf{p}_1$ because these points are identical to the reference points $\mathbf{q}_0$ and $\mathbf{p}_0$, respectively. Then, energy and gradient at the newly obtained point are computed. Before entering to the next cycle, the last local minimum along the potential curve is set as $\mathbf{q}_0$ or $\mathbf{p}_0$. This procedure is repeated until the two end points approach to each other sufficiently, where the threshold is 0.12 Å $> |\mathbf{q}_i - \mathbf{p}_j|$.

The two force terms are explained using the Müller–Brown (MB) potential with the standard parameters.[60] Figure 6 shows a contour map of the MB surface. There are two local minima, EQ1 and EQ2, which are connected by an IRC path indicated by a thick line via the TS. The IRC and the other paths shown in Figure 6 were integrated by the LQA method with the step size 0.05. Six different paths of different settings of eq. (6) are shown. A path indicated by circles was obtained by integrating the AFIR path of eq. (6) with $\delta = 300.0$ and $Y = 1.0$ throughout. In the limit of $\delta = 0.0$, this path should resemble the path of the sphere optimization algorithm by Müller and Brown.[60] The path with $\delta = 300.0$ lead to a relatively high energy area of the ridge. Paths indicated by inverse triangles and triangles facing sideways were obtained by integrating the AFIR path of eq. (6) with $\delta = 300.0$ and $Y = 0.0$ throughout, starting from EQ1 and EQ2, respectively. When $Y = 0.0$, the path does not have any influence from the other end-point, and, therefore, these two paths were calculated independently. In the limit of $\delta = 0.0$, these paths should resemble the path of the sphere optimization algorithm by Abashkin and Russo.[13] The path starting from EQ1 diverges, which is a typical behavior of the softest mode following from EQ1. The other path from EQ2 passed near the TS. Paths indicated by triangles, crosses, and squares are those obtained by integrating the AFIR path of eq. (6) with $\delta = 100.0$, 300.0, and 500.0, respectively, and Y determined by eq. (9). These paths followed the softest mode around EQs, and converged smoothly to the ridge in the TS area, and consequently gave the paths resembling the IRC path. This behavior was intended by combining the two force terms that resembling the two different sphere optimization methods, and this was the reason why this algorithm is called DS-AFIR.

A benchmark test was made for a test set proposed by Baker.[65] The Baker set consists of 25 elementary steps in different reactions. In this test, an option "DS = Single" was used, where this option disables LUP and IRC calculations and optimize single TS from the highest energy point along the AFIR path. At first, all TSs in the Baker set reported in the original paper (optimized at HF/STO-3G)[65] were reoptimized at the present B3LYP/6-31G** level with Gaussian 09.[57] Then, IRC calculations were performed at the same level starting from the optimized TSs. The two terminal points (optimized local minima) of the IRC paths were used as the initial two end points of the DS-AFIR calculation. In all the 25 cases, the DS-AFIR found the corresponding TS. Table 3 lists computational costs required to find the TS for each elementary step starting from the initial two end points. Two cases are compared; a case where the exact Hessian was computed only once at the starting point of TS optimization and the other in which the exact Hessian was never computed, where a model Hessian constructed by Bofill's Hessian updating algorithm was used at points where the exact Hessian was not

Table 3. DS-AFIR; Baker's set performances. N_g and N_h mean number of gradient calculations and that of Hessian, respectively.

		RUN1[a]		RUN2[b]	
Entry	Reaction investigated	N_g	N_h	N_g	N_h
1	HNC ↔ HCN	57	1	61	0
2	HCCH ↔ CCH_2	44	1	50	0
3	$H_2 + CO$ ↔ H_2CO	97	1	100	0
4	CH_2OH ↔ CH_3O	49	1	54	0
5	ring closing cyclopropyl	48	1	57	0
6	ring opening bicyclo[1.1.0] butane TS 1	97	1	113	0
7	ring closing bicyclo[1.1.0] butane TS 2	133	1	131	0
8	1,2-migration-(formyloxy)ethyl	45	1	69	0
9	cyclohexene ↔ butadiene + ethylene	122	1	148	0
10	s-tetrazine ↔ $2HCN + N_2$	181	1	201	0
11	trans-butadiene ↔ cis-butadiene	74	1	84	0
12	ethylene + H_2 ↔ ethane	181	1	182	0
13	fuluoroethane ↔ ethylene + HF	85	1	97	0
14	acetaldehyde ↔ vinyl alcohol	84	1	93	0
15	HCOCl ↔ HCl + CO	54	1	60	0
16	$H_2O + PO_3^-$ ↔ $H_2PO_4^-$	104	1	124	0
17	$CH_2CHCH_2CH_2CHO$ ↔ CH_2CHCH_2—O—$CHCH_2$	108	1	135	0
18	$SiH_2 + CH_3CH_3$ ↔ $SiH_3CH_2CH_3$	85	1	126	0
19	HNCCS ↔ HNC + CS	108	1	133	0
20	$CO + NH_4^+$ ↔ $HCONH_3^+$	140	1	119	0
21	rotational TS in acrolein	81	1	97	0
22	HOCHNOH ↔ HCONHOH	69	1	73	0
23	H_2CNH ↔ $HNC + H_2$	92	1	101	0
24	$HCNH_2$ ↔ H_2CNH	119	1	118	0
25	H_2NCH ↔ $HCN + H_2$	113	1	101	0

[a] Exact Hessian was computed once at the initial structure of TS optimization. [b] Exact Hessian was never computed.

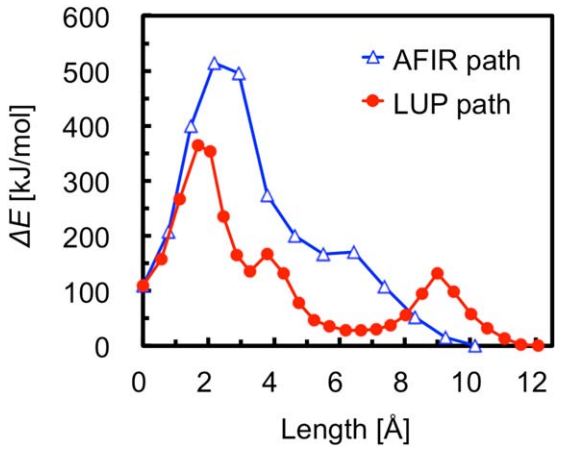


Figure 7. Energy profiles along the AFIR path obtained by applying DS-AFIR to pentaprismane and [10]annulene and the LUP path obtained by optimizing the AFIR path. [Color figure can be viewed at wileyonlinelibrary.com]

computed.[55] As seen in this table, the exact Hessian improved the overall performance only a little, probably because the model Hessian was well trained by Bofill's Hessian updating method during integrating the AFIR path before reaching the ridge region. The worst case was found in the entry 10. The AFIR path was relatively long in this reaction, because in its reactant complex the three molecules are only very weakly bound to each other and linearly arranged as N≡C − H—N≡N—H − C≡N. The other worse cases such as entries 9 and 12 are also those having a weak van der Waals complex as one of initial end points, and in such a case the AFIR path tends to be long and the total cost increases.

Another example is an isomerization between pentaprismane and [10]annulene, at the B3LYP/6-31G** level with Gaussian 09.[57] There are several elementary steps between these two isomers. In this application, at first the DS-AFIR algorithm was applied to the two initial end points, pentaprismane and [10]annulene. Then, the obtained AFIR path was optimized by the LUP method of our implementation explained above. All maximum energy points along the final LUP path were optimized to TSs, and finally IRC calculations were made starting from all obtained TSs. Figure 7 presents energy profiles along the initial AFIR path and the final LUP path. On the AFIR path, there are two peaks corresponding to the first and the third peaks, respectively, on the final LUP path. The second peak appeared through the path optimization by the LUP method. Then, starting from the three peaks on the LUP path, three TSs were obtained. Finally, IRC calculations were done starting from all the three TSs, and the four EQs and three TSs appeared in the EQ-list and TS-list, respectively. IRC profiles and these EQs and TSs are depicted in Figure 8. In default, these multistep calculations are done automatically from a single input file. The final LUP path is also available in the final output, and further "RePATH" calculation can also be done if there is any discontinuous part in the IRC profile due to missing TSs. To complete the entire calculation, that is, AFIR path calculation, LUP path optimization, TS optimizations, and IRC calculations, 1142 gradient calculations were required.

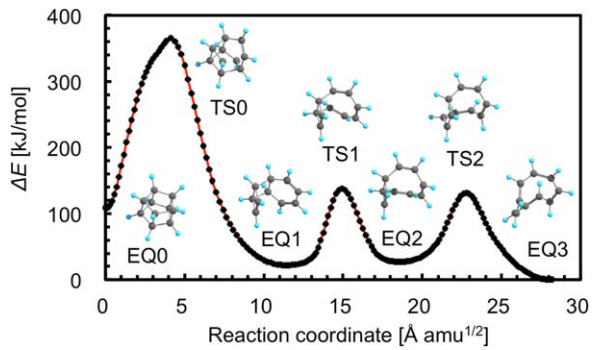


Figure 8. The energy profile along IRC paths between pentaprismane and [10]annulene. [Color figure can be viewed at wileyonlinelibrary.com]

Search for MESX/MECI Geometries

In GRRM17, structures of MESX and MECI can be explored by the SC-AFIR method.[66,67] MESX is a local minimum on the crossing seam between two adiabatic PESs of different spin or space symmetry. MECI is a local minimum on the conical intersection between two adiabatic PESs of the same spin and space symmetry. To search for MESX or MECI structures by the SC-AFIR method, a conventional MESX/MECI optimization method called gradient projection (GP) method was combined.[68] The GP method uses a composed gradient $\mathbf{g}^{GP}$ given in eq. (11),

$$\mathbf{g}^{GP}=2\left[E^{X}(\mathbf{Q})-E^{Y}(\mathbf{Q})\right]\frac{\mathbf{v}^{DGV}}{|\mathbf{v}^{DGV}|}+\mathbf{P}\frac{1}{2}\left[\frac{dE^{X}(\mathbf{Q})}{d\mathbf{Q}}+\frac{dE^{Y}(\mathbf{Q})}{d\mathbf{Q}}\right] \quad (11)$$

where E^{X} and E^{Y} are adiabatic PESs of the two target states, and $\mathbf{v}^{DGV}$ corresponds to the difference gradient vector between the two PESs. The matrix $\mathbf{P}$ is so called projection matrix, and $\mathbf{P}$ for MESX optimization is defined as follows.

$$\mathbf{P}=\mathbf{1}-\frac{\mathbf{v}^{DGV}\left(\mathbf{v}^{DGV}\right)^{T}}{|\mathbf{v}^{DGV}|^{2}} \quad (12)$$

Conversely, $\mathbf{P}$ for MECI optimization is defined as,

$$\mathbf{P}=\mathbf{1}-\frac{\mathbf{v}^{DGV}\left(\mathbf{v}^{DGV}\right)^{T}}{|\mathbf{v}^{DGV}|^{2}}-\frac{\mathbf{v}'(\mathbf{v}')^{T}}{|\mathbf{v}'|^{2}} \quad (13)$$

where $\mathbf{v}'$ is an arbitrary vector not parallel to $\mathbf{v}^{DGV}$ on the branching plane (BP), and the derivative coupling vector or a vector obtained by the BP updating (BPU) method is used.[69] In GRRM17, $\mathbf{v}'$ is obtained by BPU as a unit vector perpendicular to $\mathbf{v}^{DGV}$.

In the combined GP/AFIR method, the composed gradient $\mathbf{g}^{GP/AFIR}$ is,

$$\mathbf{g}^{GP/AFIR}=2\left[E^{X}(\mathbf{Q})-E^{Y}(\mathbf{Q})\right]\frac{\mathbf{v}^{DGV}}{|\mathbf{v}^{DGV}|}+\mathbf{P}\frac{1}{2}\left[\frac{dE^{X}(\mathbf{Q})}{d\mathbf{Q}}+\frac{dE^{Y}(\mathbf{Q})}{d\mathbf{Q}}+2\frac{dJ(\mathbf{Q})}{d\mathbf{Q}}\right] \quad (14)$$

where $J(\mathbf{Q})$ is the artificial force term in eq. (1). This equation can be obtained just by substituting F^{AFIR} of eq. (1) for E^{X} and

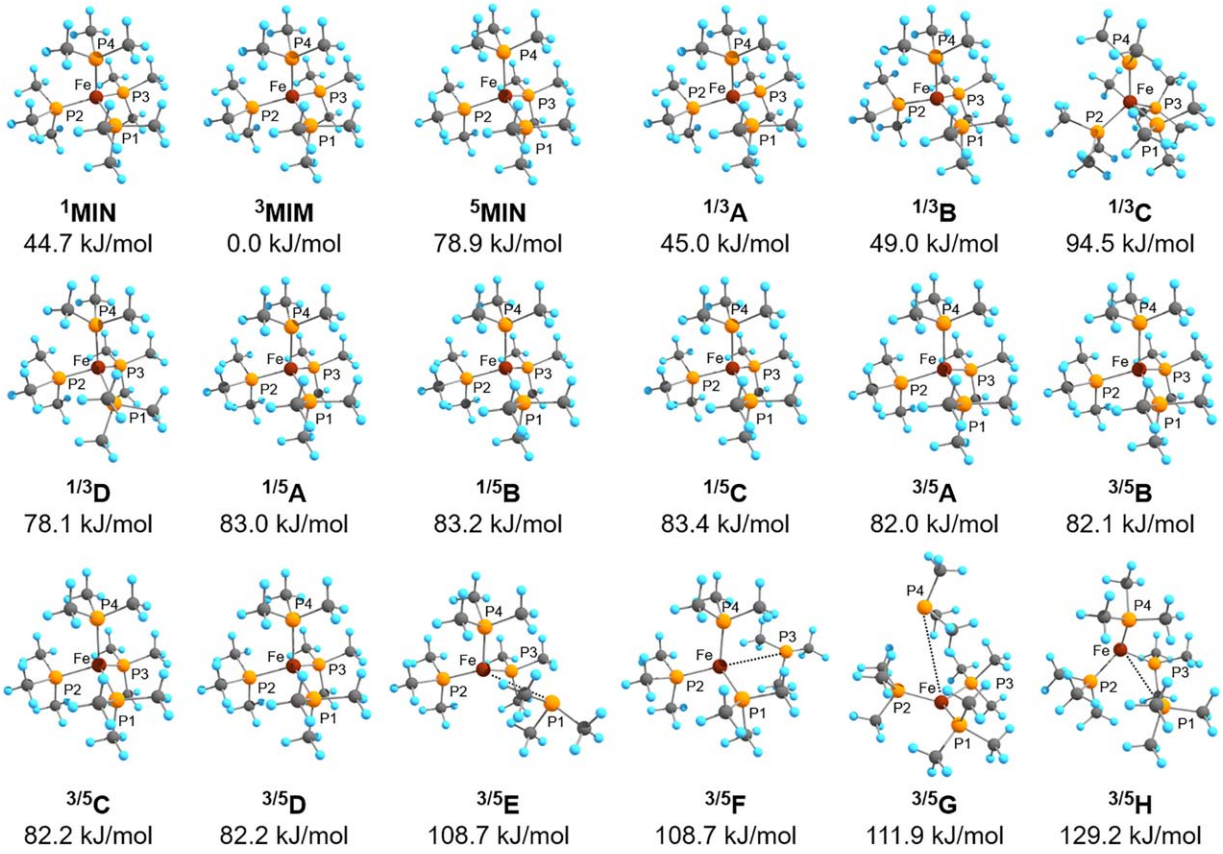


Figure 9. Structures of singlet ^{1}**MIN**, triplet ^{3}**MIN**, and quintet ^{5}**MIN**, and all obtained MESX structures by the GP/SC-AFIR searches, for $Fe(PMe_3)_4$. MESXs for singlet-triplet, singlet-quintet, and triplet-quintet pairs are denoted as $^{1/3}$**A–D**, $^{1/5}$**A–C**, $^{3/5}$**A–H**, respectively. Relative energy values in kJ/mol are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

E^Y into E^X and E^Y in eq. (11), respectively. In the SC-AFIR search, the GP/AFIR path is calculated by inducing structural deformations using $\mathbf{g}^{GP/AFIR}$. A GP/AFIR path leads to an approximate MESX/MECI structure within the seam of crossing/conical intersection hyperspace, where different structures can be obtained depending on a fragment pair in the force term, as in the same way in the SC-AFIR search on a single PES. Therefore, the SC-AFIR algorithm using $\mathbf{g}^{GP/AFIR}$ generates various approximate MESX/MECI structures automatically. These approximate structures are further optimized to actual MESX/MECI automatically during the search. Thus, many actual MESX/MECI structures are obtained and appear in the EQ-list.

An example in which MESX structures of $Fe(PMe_3)_4$ are explored is shown. In this calculation, only Fe and four P atoms were selected by the "Target" option and the "FirstOnly" option was adopted to restrict the search area. Three different searches were performed for three different spin state pairs, that is, singlet-triplet, singlet-quintet, and triplet-quintet, where these searches were initiated from local minima on the singlet, singlet, and triplet PESs, respectively. The γ value was set to 100 kJ/mol. The search was done at the UB3LYP-D3/SDD(Fe)+6–31G*(H, C, P) level with Gaussian09.[57] We note that the system tends to have an open-shell singlet character at MESXs, and the stability check of molecular orbitals implemented in Gaussian 09 was done to detect such a character during the search. Figure 9 shows all obtained MESX structures and structures of singlet ^{1}**MIN**, triplet ^{3}**MIN**, and quintet ^{5}**MIN**. MESXs for singlet-triplet, singlet-quintet, and triplet-quintet pairs are denoted as $^{1/3}$**A-D**, $^{1/5}$**A-C**, $^{3/5}$**A-H**, respectively. Since the two target spin states have the same energy (at least up to the first digit below the decimal point in kJ/mol), only a single energy value is shown below each structure. In these searches for MESXs of singlet-triplet, singlet-quintet, and triplet-quintet pairs, $\mathbf{g}^{GP/AFIR}$ or $\mathbf{g}^{GP}$ were computed at 3505, 1795, and 2856 structures, respectively, where at each structure two gradient calculations were done. Among the three spin states, the triplet is the ground state. By performing meta-IRC calculations starting from $^{1/3}$**A**, $^{1/5}$**A**, and $^{3/5}$**A** on both of the two spin states that degenerate, an energy profile shown in Figure 10 was obtained.

Another example is the search for MECI structures between the singlet ground S_0 and first excited S_1 electronic state of naphthalene. The search was done starting from the S_0 local minimum (Franck–Condon structure) with the "FirstOnly" option, where the γ value was set to 100 kJ/mol. Electronic structure calculations were done by the Spin-Flip TDDFT with BHHLYP functional and 6-31G* basis set using GAMESS program.[70] The search generated MECI structures listed in Figure 11. In this search, $\mathbf{g}^{GP/AFIR}$ or $\mathbf{g}^{GP}$ were computed at 5539 structures, where at each structure two gradient calculations were done. To identify the most feasible nonadiabatic decay path, TSs that separate the S_1 local minimum and the MECI

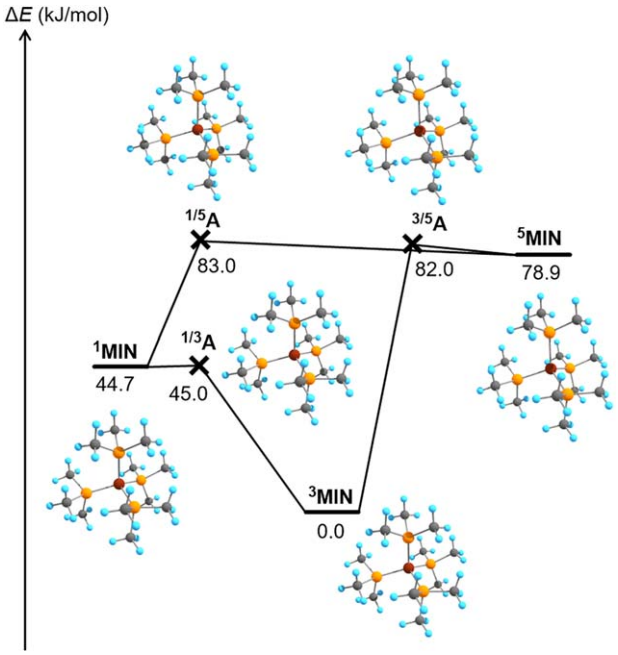


Figure 10. Connections among lowest energy MESX structures, that is, $^{1/3}$**A**, $^{1/5}$**A**, and $^{3/5}$**A** in Figure 9, through meta-IRC paths. [Color figure can be viewed at wileyonlinelibrary.com]

structures on S_1 were searched by applying the DS-AFIR between them. The DS-AFIR was applied in order from MECI with low energy. The application to the lowest MECI1 gave TS1 which is 5.76 eV relative to S_0. The next calculation for the second lowest MECI2 gave TS2 with relative energy 5.59 eV. At this point, MECIs having higher energy than 5.59 eV can be omitted. Further application of the DS-AFIR to only remaining one MECI3 gave TS3 with 5.75 eV. Finally, the most feasible nonradiative decay path was identified to be the one going through MECI2 and TS2.

Interface with ONIOM

The ONIOM (our own *N*-layered integrated molecular orbital and molecular mechanics) is a flexible quantum mechanics (QM)/molecular mechanics (MM) method.[71,72] The ONIOM allows to combine different QM methods and/or three or more QM and MM methods. The microiteration technique has been used in geometry optimization by ONIOM.[73,74] In this technique, positions of atoms in the surrounding parts that are usually treated by low computational level (MM or semi empirical QM) in ONIOM are optimized with fixing positions of atoms in the reaction center part, before changing atom positions in the reaction center part. This treatment greatly reduces the number of gradient calculations for the reaction center part. It is assumed in microiteration that optimization of the surrounding parts is computationally less expensive than single gradient calculation in the reaction center part, and with this assumption it can reduce the total computational cost.

The microiteration has already been implemented in GRRM14.[75,76] In GRRM17, it is also interfaced with the SC-AFIR method. In the microiteration, the whole system is divided into two parts: the reaction center part and the surrounding part. Similarly to the microiteration in ordinary geometry optimization, the surrounding part is once optimized with fixing atom positions in the reaction center part, and then positions of atoms in the reaction center part are changed using the gradient vector of eq. (1). We note that with the microiteration the artificial force can be applied only to atoms in the reaction center part. In the microiteration, the following effective

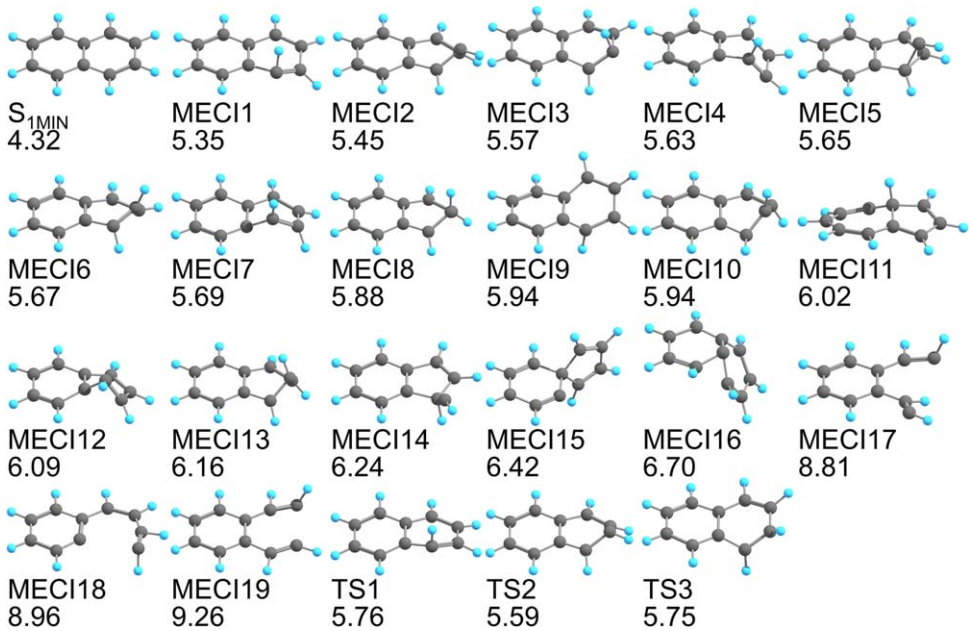


Figure 11. MECI structures for naphthalene obtained by the GP/SC-AFIR search and TSs on S_1 linking the S_{1MIN} and the lowest lying three MECIs. Relative energy values in eV are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

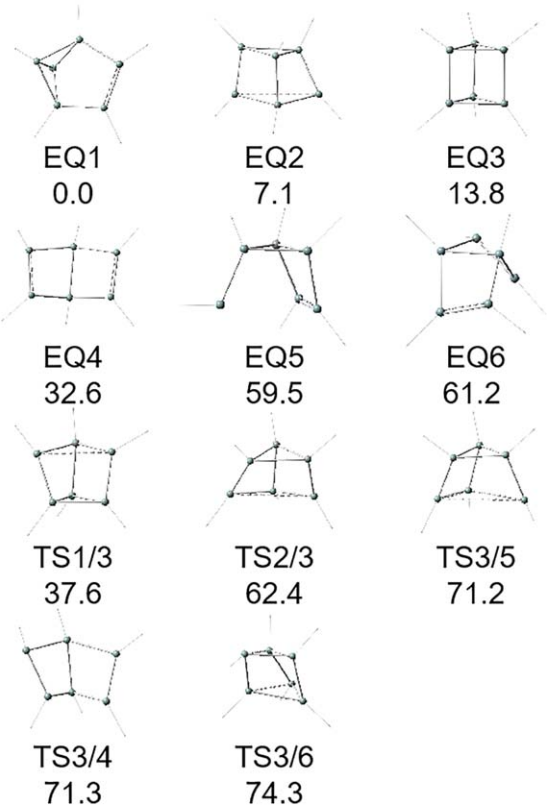


Figure 12. EQs and TSs obtained by the SC-AFIR method combined with ONIOM and microiteration technique, starting from hexasilaprismane (EQ3). The six bulky 2,6-diisopropylphenyl groups are not depicted for clarity. Relative energy values in kJ/mol are shown below each structure. [Color figure can be viewed at wileyonlinelibrary.com]

Hessian is used to take the coupling between the surrounding part and the reaction center part,

$$\frac{\partial E(\mathbf{Q})}{\partial Q_i \partial Q_j} = h_{ij}^{\mathrm{R}} - \sum_{m=3p+1}^{3q} \sum_{n=3p+1}^{3q} h_{im}^{\mathrm{C}} W_{nm}^{\mathrm{S}} h_{nj}^{\mathrm{C}} \tag{15}$$

where $\mathbf{Q}$ is the atomic coordinates $\{Q_i\}$ for atoms in the reaction center part, h_{ij}^{R} and h_{im}^{C} are Hessian matrix elements in the block for the reaction center part and those in the coupling block between the surrounding part and the reaction center part, respectively, p and q are the number of atoms in the reaction center part and the total number of atoms, respectively, and W_{nm}^{S} is an inverse matrix element of the block for the surrounding part.

Isomerization pathways of hexasilaprismane ($Si_6C_{72}H_{102}$), in which 2,6-diisopropylphenyl group is attached to each six Si atom constructing the triangular prism skeleton,[77] were explored by the SC-AFIR with the "FirstOnly" option and $\gamma = 200$ kJ/mol. In this application, the six Si atoms and C atoms directly bonded to the Si atoms were the reaction center and the remaining atoms were the surrounding part. Atoms that are considered in the fragment generation were restricted only to the six Si atoms by the "Target" option. In the ONIOM setting, the six Si atoms and the six link H atoms were dealt as model-QM (B3LYP/6–31G*), and the universal force field was adopted in the model-MM and real-MM calculations. The ONIOM calculation was done with the mechanical embedding framework, where the atom charge was determined at the initial structure by the restrained electrostatic potential method.

Figure 12 shows obtained EQs and TSs in the search, where only the reaction center part is depicted for clarity. In this calculation, 7897 gradient and 1077 Hessian calculations were done, where optimization of the surrounding part was done before computing these gradient and Hessian calculations. Actually, 33 EQs and 15 TSs were found in total. If focusing only the reaction center part (arrangement of the six Si atoms), the 33 can be classified into the 6 EQs shown in Figure 12, and the 6 EQs are those having the lowest energy among similar EQs differing only in the conformations in 2,6-diisopropylphenyl group. Except for TSs connecting same EQs, the 5 TSs in Figure 12 were obtained. Although a few TSs and EQs that were not directly connected to EQ3 were also found due to too strong artificial force applied, such structures are not shown for clarity.

Visualization of Reaction Path Networks

As shown above, the SC-AFIR search generates a database of EQs, TSs, and DCs automatically. The database can be big even in small systems like C_4H_6, and data processing such as creation of the reaction path network like Figure 2 can be a hard task for users. Moreover, all data are not necessarily useful in each specific purpose. To analyze the database and also to grasp the entire image easily and accurately, a tool is required. In this section, a code termed SAFIRE (Simple AFIR-calculation-results Extractor) which was used to draw reaction path networks in Figures 2 and 3 is described.

The SAFIRE code, which is written in the python 3 programing language, draws the reaction path network using Graphviz (AT&T, Eclipse Public License-v1.0).[78] SAFIRE reads EQs, TSs, and DCs as well as connections between EQs through TS from output files of GRRM17, and generates files in the "dot" format which can be read by Graphviz. There are some options to control whether to sort EQs depending on their energy, whether to draw 3D structures of EQs, how to change color of the network depending on energy, and so forth. The pictures of 3D structures of EQs are generated by Jmol program.[79] Each structure is automatically rotated to the "best" orientation for viewing by Jmol, such that those atoms' containing box is thinnest along the line of sight (z), and widest horizontally (x). It is also possible to focus a local area if one's interest is a local area between specific EQs or around a specific EQ.

Concluding Remarks

In this article, we have described new features in GRRM17 compared to the previous version. The main update is implementation of the SC-AFIR algorithm which enables extensive reaction path searching and generation of global or semiglobal reaction path network of a given system automatically. There are several options to limit the search area to accelerate obtaining results depending on one's focus, and some representative usages were introduced with simple applications. In addition, the double-end version termed DS-AFIR, an implementation of SC-AFIR for MESX/MECI exploration, and an

interface of SC-AFIR with QM/MM-ONIOM were introduced with simple examples. Options used in these examples are summarized in Appendix. Input files of all these calculations will be available as an attachment to GRRM17 in its distribution. Although some very simple examples that can be a test set of GRRM17 were presented in this article, we note that much more complicated systems such as multicomponent organic reactions and organometallic catalysis involving about 100 atoms can be practical targets as have been demonstrated in our recent studies.[80,81] SAFIRE, a tool to visualize complex reaction path network generated by GRRM17, will also be available. GRRM17 can be combined with any *ab initio*, DFT, semi-empirical, or molecular mechanics program by preparing a simple interface script. GRRM17 with these features will be a useful tool to explore unknown chemical structures and reaction pathways based on quantum chemical calculations.

We note that these calculations are done for a given, specific atomic composition. Which species to include in a calculation will be a choice of users. The computational cost increases rapidly as the number of species involved increases. It is, therefore, advised in each reaction step to delete species that are not directly participate in the reaction step. However, one should also keep in mind that there are cases where species that formally do not participate in the reaction step have significant contribution. Another parameter that need to be chosen carefully by users is computational level. Although one can adopt a computational level which is fast but not very accurate in the initial search for fast screening, it should be noted that too low computational level may mislead a conclusion.

In the SC-AFIR algorithm, fragments for eq. (1) are automatically defined for different systems. Only parameter that users need to specify is the γ value. The γ value is generally decided by eq. (4) in thermal reactions or based on excess energy available in the excited state in photoreactions. One may set γ to a very large value such as 1000 kJ/mol if the target is all possible paths. Users can specify atoms involved in the reaction center, and by doing so the computational cost decreases dramatically. How to choose the reaction center depends on the system. In the example regarding the Co-carbonyl complex, the central Co atom and atoms bonded to it were chosen as the reaction center. In defining a rational reaction center, user's intuition is required. It is also noted that the periodic boundary conditions (PBCs) option is not available in GRRM17. Development of the PBCs options for AFIR is under progress toward its application to crystals and interfaces,[82] and the options will be available in the next update.

Appendix: GRRM17 Keywords Used

MC-AFIR: Systematic sampling of reaction pathways through which two or more molecules react to each other is made.

SC-AFIR: Systematic exploration of pathways of reactions that occur in a given system is done. All types of paths including those for intramolecular and intermolecular chemical bond rearrangements, hydrogen bond rearrangement, conformational rearrangement, coordination bond rearrangement, and so forth, are found.

DS-AFIR: A single (in general shortest) path connecting a given pair of reactant and product structures is calculated.

RePATH: LUP path optimization, EQ optimization, TS optimization, and IRC calculation, are applied to all paths (either AFIR path, LUP path, or IRC path) read from previous results.

EQOnly: LUP path optimization, TS optimization, and IRC calculation are skipped. Only EQs are obtained if this option is used without KeepSCPath nor KeepLUPPath.

KeepSCPath: When used in SC-AFIR together with EQOnly, AFIR paths and their peak structures are printed in output files.

KeepLUPPath: When used together with EQOnly, LUP path optimization (disabled by EQOnly) is done and LUP paths and their peak structures are printed in output files.

PTOnly: In RePATH, LUP path optimization, EQ optimization, TS optimization, and IRC calculation, are applied only to LUP paths read from previous results, where LUP paths that connect an EQ-EQ pair for which a lower barrier IRC path connection is already available are omitted.

FirstOnly: In SC-AFIR, SC-AFIR is applied only to a given initial structure.

NoBondRearrange: In SC-AFIR, SC-AFIR is applied only to local minima having the same bond connectivity as a given initial geometry has.

Target: In SC-AFIR, paths (an atom-atom pair used in generating a pair of fragments) are chosen among atoms specified by this option.

rTarget: In SC-AFIR, particularly reactive paths (an atom-atom pair used in generating a pair of fragments) are chosen among atoms specified by this option.

PriorityPath: In SC-AFIR, paths (an atom-atom pair used in generating a pair of fragments) specified with this option will be computed prior to the other paths.

SC = FindDC: In SC-AFIR, dissociation channels that do not have TS are calculated and stored in the DC-list file (in default, such paths are ignored).

DS = Single: In DS-AFIR, LUP path optimization is skipped and only the highest energy point along the AFIR path is optimized to TS.

Acknowledgment

This research partly used computational resources provided by Academic Center for Computing and Media Studies, Kyoto University.

Keywords: potential energy surface · transition state · intrinsic reaction coordinate · nonadiabatic transition · QM/MM

How to cite this article: S. Maeda, Y. Harabuchi, M. Takagi, K. Saita, K. Suzuki, T. Ichino, Y. Sumiya, K. Sugiyama, Y. Ono. *J. Comput. Chem*. **2018**, *39*, 233–250. DOI: 10.1002/jcc.25106

[1] H. B. Schlegel, *J. Comput. Chem.* **2003**, *24*, 1514.
[2] F. Jensen, *Introduction to Computational Chemistry*, 2nd ed.; Wiley: Chichester, UK, **2007**.

[3] K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 363.
[4] S. Maeda, Y. Harabuchi, Y. Ono, T. Taketsugu, K. Morokuma, *Int. J. Quantum Chem.* **2015**, *115*, 258.
[5] J. N. Harvey, *WIREs Comput. Mol. Sci.* **2014**, *4*, 1.
[6] S. Maeda, T. Taketsugu, K. Ohno, K. Morokuma, *J. Am. Chem. Soc.* **2015**, *137*, 3433.
[7] D. R. Yarkony, *Rev. Mod. Phys.* **1996**, *68*, 985.
[8] F. Bernardi, M. Olivucci, M. A. Robb, *Chem. Soc. Rev.* **1996**, *25*, 321.
[9] H. B. Schlegel, *WIREs Comput. Mol. Sci.* **2011**, *1*, 790.
[10] H. L. Davis, D. J. Wales, R. S. Berry, *J. Chem. Phys.* **1990**, *92*, 4308.
[11] J.-Q. Sun, K. Ruedenberg, *J. Chem. Phys.* **1993**, *98*, 9707.
[12] C. J. Tsai, K. D. Jordan, *J. Phys. Chem.* **1993**, *97*, 11227.
[13] Y. Abashkin, N. Russo, *J. Chem. Phys.* **1994**, *100*, 4477.
[14] K. Bondensgård, F. Jensen, *J. Chem. Phys.* **1996**, *104*, 8025.
[15] J. P. K. Doye, D. J. Wales, *Z. Phys. D* **1997**, *40*, 194.
[16] W. Quapp, M. Hirsch, O. Imig, D. Heidrich, *J. Comput. Chem.* **1998**, *19*, 1087.
[17] M. Černohorský, S. Kettou, J. Koča, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 705.
[18] K. M. Westerberg, C. A. Floudas, *J. Chem. Phys.* **1999**, *110*, 9259.
[19] D. J. Wales, J. P. K. Doye, M. A. Miller, P. N. Mortenson, T. R. Walsh, *Adv. Chem. Phys.* **2000**, *115*, 1.
[20] K. K. Irikura, R. D. Johnson, III, *J. Phys. Chem. A* **2000**, *104*, 2191.
[21] E. M. Müller, A. de Meijere, H. Grubmüller, *J. Chem. Phys.* **2002**, *116*, 897.
[22] M. Dallos, H. Lischka, E. V. D. Monte, M. Hirsch, W. Quapp, *J. Comput. Chem.* **2002**, *23*, 576.
[23] J. Baker, K. Wolinski, *J. Comput. Chem.* **2011**, *32*, 43.
[24] P. M. Zimmerman, *J. Comput. Chem.* **2013**, *34*, 1385.
[25] D. Rappoport, C. J. Galvin, D. Yu. Zubarev, A. Aspuru-Guzik, *J. Chem. Theory Comput.* **2014**, *10*, 897.
[26] B. Schaefer, S. Mohr, M. Amsler, S. Goedecker, *J. Chem. Phys.* **2014**, *140*, 214102/1-13.
[27] D. J. Wales, *J. Chem. Phys.* **2015**, *142*, 130901/1-12.
[28] S. Habershon, *J. Chem. Phys.* **2015**, *143*, 094106/1-14.
[29] E. Martínez-Núñez, *J. Comput. Chem.* **2015**, *36*, 222.
[30] P. M. Zimmerman, *J. Comput. Chem.* **2015**, *36*, 601.
[31] M. Bergeler, G. N. Simm, J. Proppe, M. Reiher, *J. Chem. Theory Comput.* **2015**, *11*, 5712.
[32] X.-J. Zhang, Z.-P. Liu, *Phys. Chem. Chem. Phys.* **2015**, *17*, 2757.
[33] L.-P. Wang, R. T. McGibbon, V. S. Pande, T. J. Martinez, *J. Chem. Theory Comput.* **2016**, *12*, 638.
[34] M. Yang, J. Zou, G. Wang, S. Li, *J. Phys. Chem. A* **2017**, *121*, 1351.
[35] K. Ohno, S. Maeda, *Chem. Phys. Lett.* **2004**, *384*, 277.
[36] S. Maeda, K. Ohno, *J. Phys. Chem. A* **2005**, *109*, 5742.
[37] K. Ohno, S. Maeda, *J. Phys. Chem. A* **2006**, *110*, 8933.
[38] K. Ohno, S. Maeda, *Phys. Scr.* **2008**, *78*, 058122/1.
[39] S. Maeda, K. Ohno, K. Morokuma, *J. Phys. Chem. A* **2009**, *113*, 1704.
[40] S. Maeda, K. Morokuma, *J. Chem. Phys.* **2010**, *132*, 241102/1–4.
[41] S. Maeda, K. Morokuma, *J. Chem. Theory Comput.* **2011**, *7*, 2335.
[42] S. Maeda, K. Ohno, K. Morokuma, *Adv. Phys. Chem.* **2012**, *2012*, 268124/1-13.
[43] S. Maeda, K. Ohno, K. Morokuma, *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683.
[44] S. Maeda, T. Taketsugu, K. Morokuma, *J. Comput. Chem.* **2014**, *35*, 166.
[45] S. Maeda, T. Taketsugu, K. Morokuma, K. Ohno, *Bull. Chem. Soc. Jpn.* **2014**, *87*, 1315.
[46] W. M. C. Sameera, S. Maeda, K. Morokuma, *Acc. Chem. Res.* **2016**, *49*, 763.
[47] S. Maeda, Y. Harabuchi, M. Takagi, T. Taketsugu, K. Morokuma, *Chem. Rec.* **2016**, *16*, 2232.
[48] S. Maeda, Y. Harabuchi, Y. Osada, T. Taketsugu, K. Morokuma, K. Ohno, GRRM14. Available at: http://grrm.chem.tohoku.ac.jp/GRRM/index_e.html (accessed on August 19, **2017**).
[49] C. Choi, R. Elber, *J. Chem. Phys.* **1991**, *94*, 751.
[50] P. Y. Ayala, H. B. Schlegel, *J. Chem. Phys.* **1997**, *107*, 375.
[51] A. Banerjee, N. Adams, J. Simons, R. Shepard, *J. Phys. Chem.* **1985**, *89*, 52.
[52] P. Culot, G. Dive, V. H. Nguyen, J. M. Ghuysen, *Theor. Chim. Acta* **1992**, *82*, 189.
[53] M. Page, J. W. McIver, Jr., *J. Chem. Phys.* **1988**, *88*, 922.
[54] Ö. Farkas, H. B. Schlegel, *J. Chem. Phys.* **1999**, *111*, 10806.
[55] J. M. Bofill, *J. Comput. Chem.* **1994**, *15*, 1.
[56] F. Hebrard, P. Kalck, *Chem. Rev.* **2009**, *109*, 4272.
[57] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision A.02; Gaussian, Inc.: Wallingford, CT, **2013**.
[58] TURBOMOLE Version 7.0, 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007; TURBOMOLE GmbH, since 2007. Available at: http://www.turbomole.com (accessed on August 19, **2017**).
[59] J. Li, X. Li, H.-J. Zhai, L.-S. Wang, *Science* **2003**, *299*, 864.
[60] K. Müller, L. D. Brown, *Theor. Chim. Acta* **1979**, *53*, 75.
[61] R. Elber, M. Karplus, *Chem. Phys. Lett.* **1987**, *139*, 375.
[62] G. Henkelman, B. P. Uberuaga, H. Jónsson, *J. Chem. Phys.* **2000**, *113*, 9901.
[63] W. E.,W. Ren, E. Vanden-Eijnden, *Phys. Rev. B* **2002**, *66*, 052301/1–4.
[64] B. Peters, A. Heyden, A. T. Bell, A. Chakraborty, *J. Chem. Phys.* **2004**, *120*, 7877.
[65] J. Baker, F. Chan, *J. Comput. Chem.* **1996**, *17*, 888.
[66] S. Maeda, Y. Harabuchi, T. Taketsugu, K. Morokuma, *J. Phys. Chem. A* **2014**, *118*, 12050.
[67] Y. Harabuchi, T. Taketsugu, S. Maeda, *Chem. Phys. Lett.* **2017**, *674*, 141.
[68] M. J. Bearpark, M. A. Robb, H. B. Schlegel, *Chem. Phys. Lett.* **1994**, *223*, 269.
[69] S. Maeda, K. Ohno, K. Morokuma, *J. Chem. Theory Comput.* **2010**, *6*, 1538.
[70] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347.
[71] L. W. Chung, H. Hirao, X. Li, K. Morokuma, *WIREs Comput. Mol. Sci.* **2012**, *2*, 327.
[72] L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, K. Morokuma, *Chem. Rev.* **2015**, *115*, 5678.
[73] T. Vreven, K. Morokuma, O. Farkas, H. B. Schlegel, M. J. Frisch, *J. Comput. Chem.* **2003**, *24*, 760.
[74] T. Vreven, M. J. Frisch, K. N. Kudin, H. B. Schlegel, K. Morokuma, *Mol. Phys.* **2006**, *104*, 701.
[75] S. Maeda, K. Ohno, K. Morokuma, *J. Chem. Theory Comput.* **2009**, *5*, 2734.
[76] S. Maeda, E. Abe, M. Hatanaka, T. Taketsugu, K. Morokuma, *J. Chem. Theory Comput.* **2012**, *8*, 5058.
[77] A. Sekiguchi, T. Yatabe, C. Kabuto, H. Sakurai, *J. Am. Chem. Soc.* **1993**, *115*, 5853.
[78] E. R. Gansner, S. C. North, *Softw. Pract. Exp.* **2000**, *30*, 1203.
[79] Jmol: an open-source Java viewer for chemical structures in 3D. Available at: http://www.jmol.org/ (accessed on August 22, **2017**).
[80] K. Vong, S. Maeda, K. Tanaka, *Chemistry* **2016**, *22*, 18865.
[81] T. Yoshimura, S. Maeda, T. Taketsugu, M. Sawamura, K. Morokuma, S. Mori, *Chem. Sci.* **2017**, *8*, 4475.
[82] M. Takagi, T. Taketsugu, H. Kino, Y. Tateyama, K. Terakura, S. Maeda, *Phys. Rev. B* **2017**, *95*, 184110/1-11.

Received: 26 August 2017
Revised: 16 October 2017
Accepted: 18 October 2017
Published online on 14 November 2017