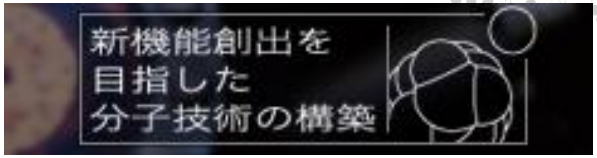
CMD workshop

Finding reaction pathways and their networks by first principle calculations

Satoshi Meada

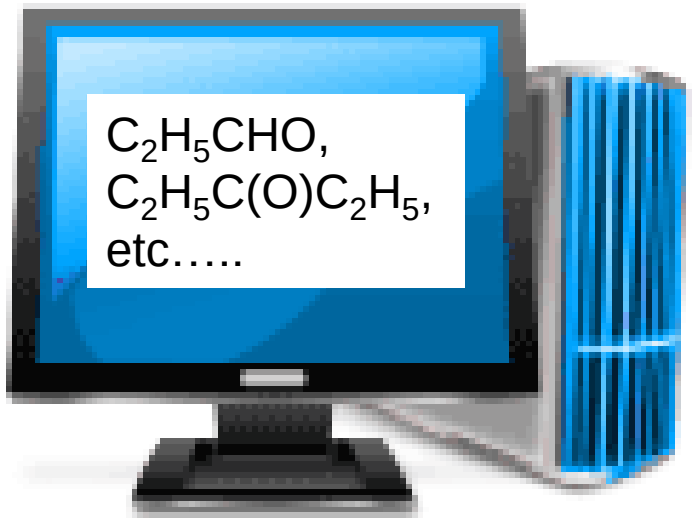
Hokkaido Univ., NIMS

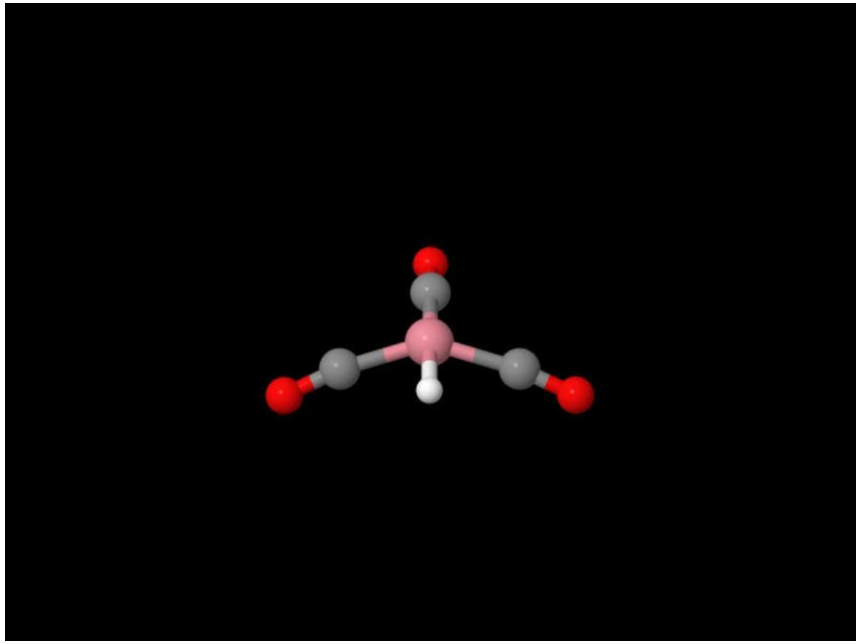


Introduction

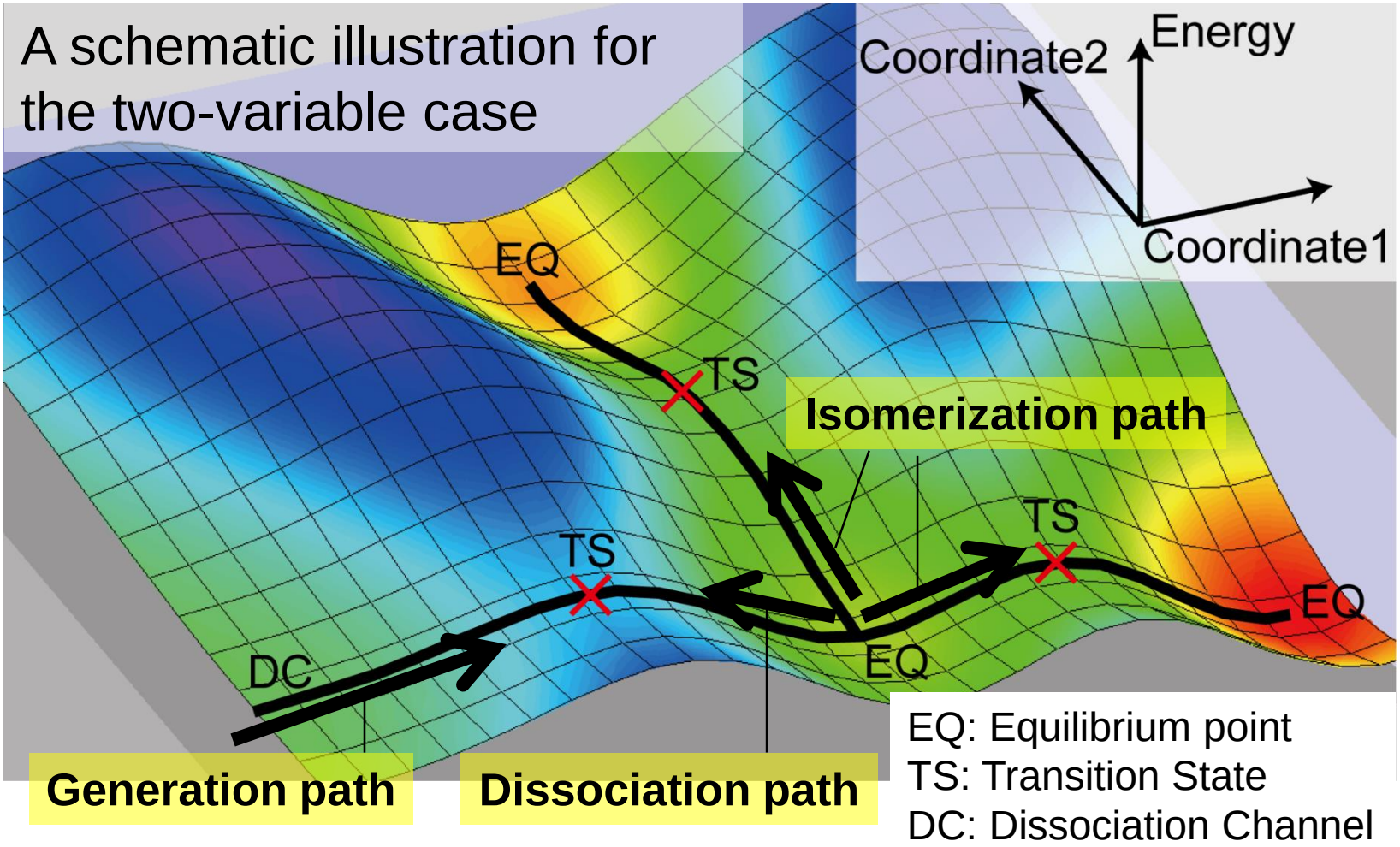
Our objective

Establishing a computational method by which products and all possible byproducts are predicted together with their generation mechanisms based on first principle calculations, just by giving reactants, catalysts, additives, etc. together with reaction conditions.





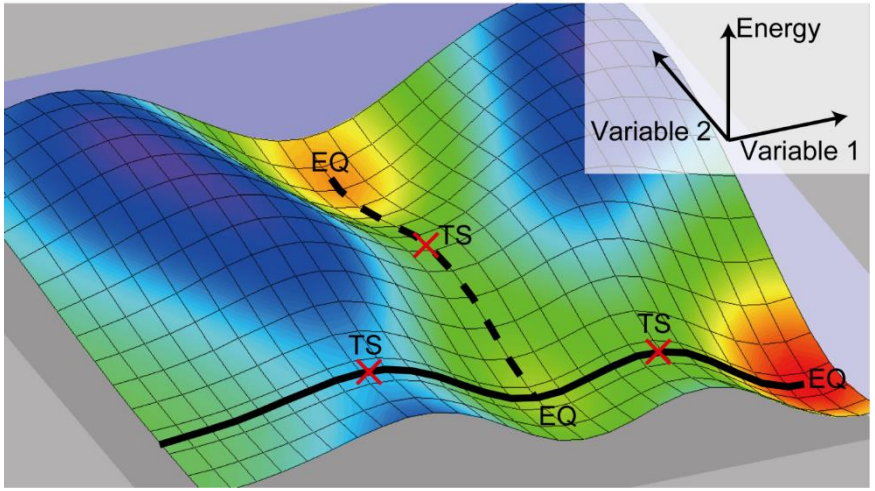
Potential energy surface (PES)



Reaction paths can be predicted by systematically exploring the PES

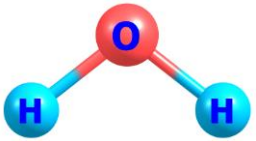
Search by the grid method

It takes a huge cost to acquire the entire network with grid search

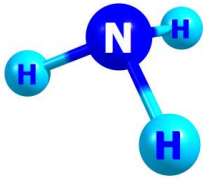


Grid points ($10 \times 10 \times 10 \times \cdots$ grid) on the potential energy surface

3 atom	(3 variables)	→	10^{3} points
4 atom	(6 variables)	→	10^{6} points
5 atom	(9 variables)	→	10^{9} points
6 atom	(12 variables)	→	10^{12} points
10 atom	(24 variables)	→	10^{24} points
60 atom	(174 variables)	→	10^{174} points



OH distance × 2
HOH angle
3 variables



NH distance × 3
HNH angle × 3
6 variables

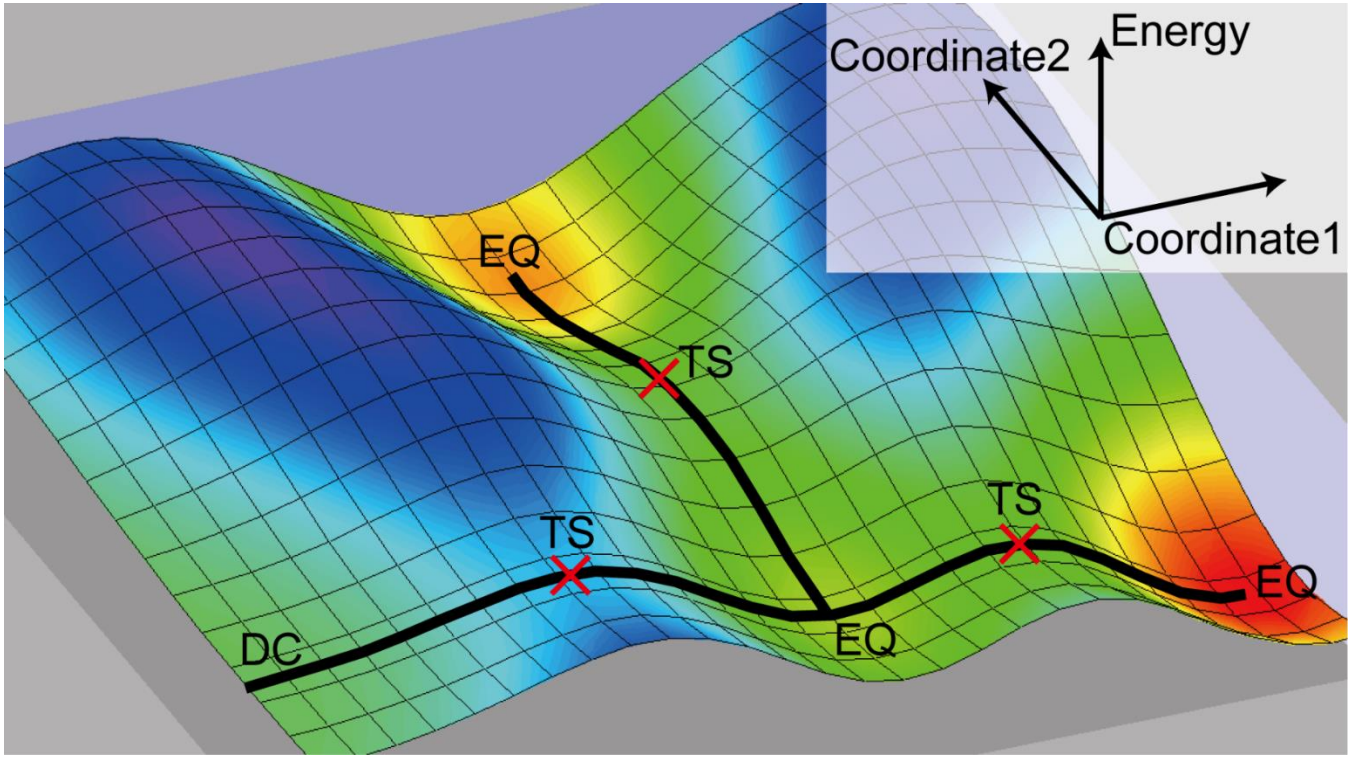
The number of variables increases by 3 as the number of atoms increases

If it takes one minutes to calculate 1 point, calculation at 10^{24} points
takes 1.9×10^{18} years!!
(cf., from the birth of the universe: ~1.37×10^{10} years)

A speedup of ~10^{6} times by a supercomputer cannot be a sufficient help

Exploring PESs by tracing paths

This seems to be the best strategy, though…



It was shown to be impossible to follow the reaction path defined as intrinsic reaction coordinates from minima

Jensen, F. Introduction to Computational Chemistry, Wiley, Chichester, **1998**.
Schlegel, H. B. *J. Comput. Chem.* **2003**, *24*, 1514.

Our strategy

1. Trace a path similar to the reaction path (Find an entrance of *approximate* reaction path and trace it)
2. Avoid searching parts away from the reaction path as much as possible

What method is effective for this purpose?

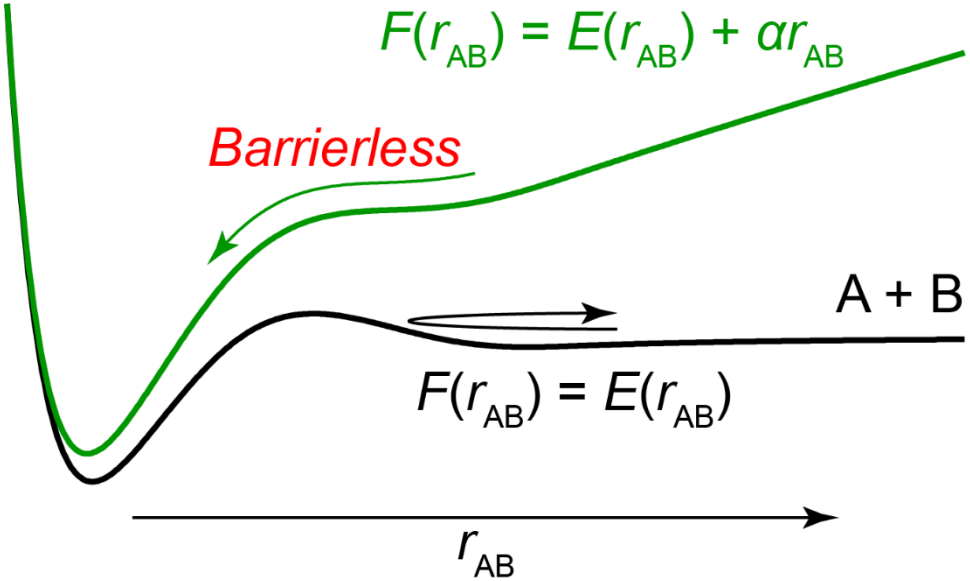
1. Anharmonic downward distortion following (ADDF) method, since 2004
2. Artificial force induced reaction (AFIR) method, since 2010

The artificial force induced reaction (AFIR) method: its basics and ideas to apply it to complex chemical reactions

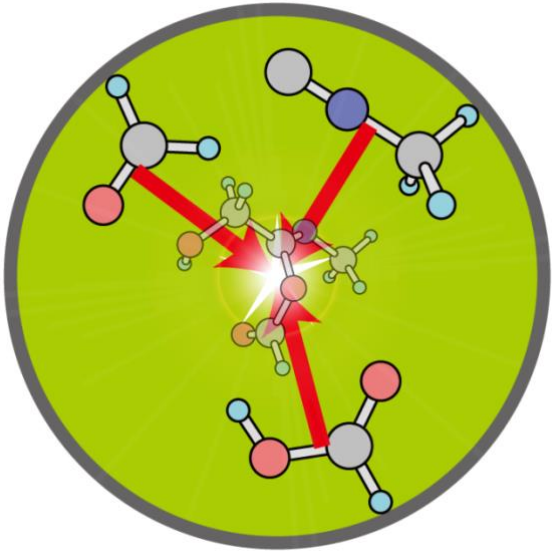
The AFIR method

Idea: *Push reactants together (or pull them apart)*

A case for atom-atom reaction



Intermolecular reaction



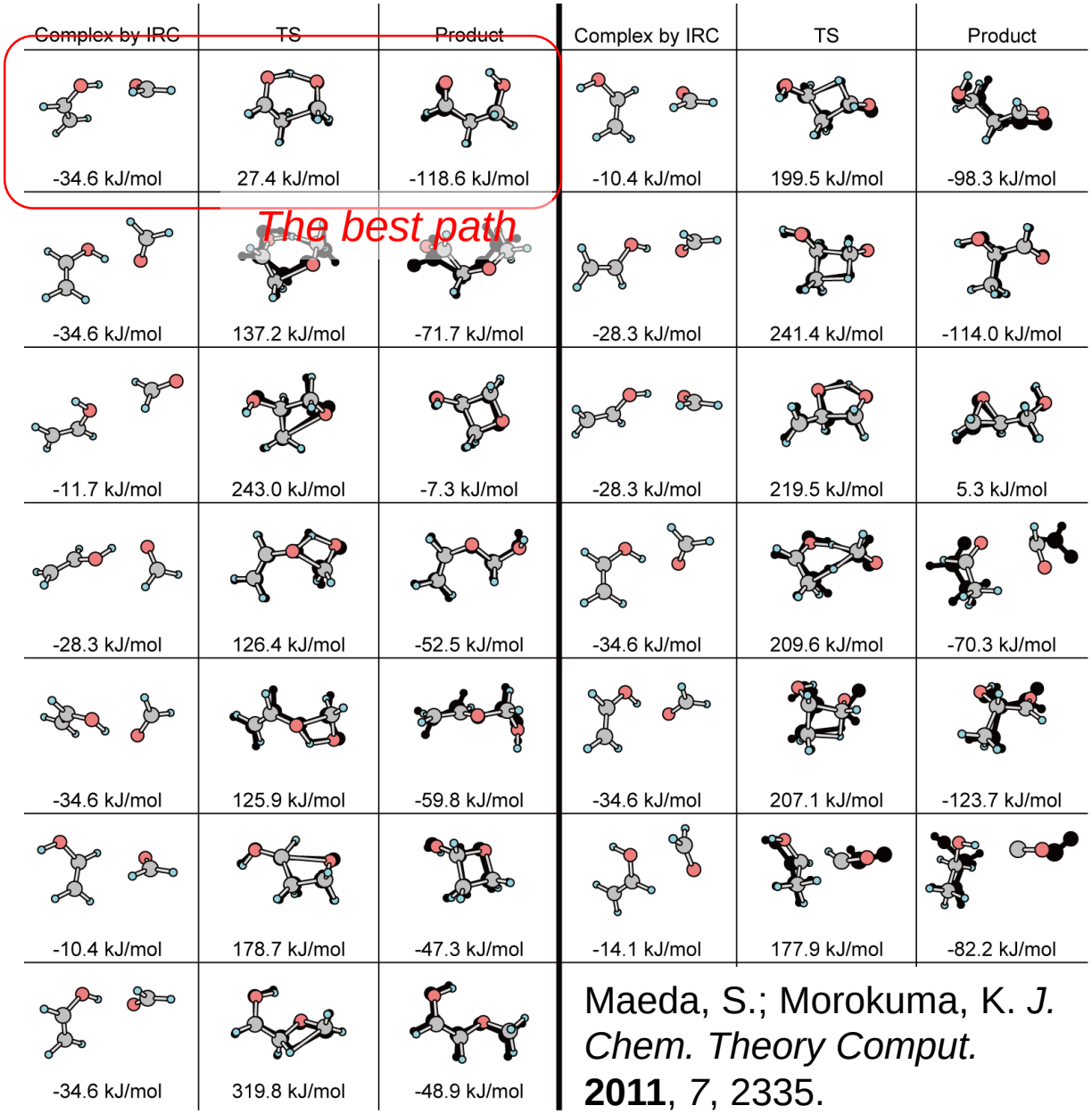
The AFIR function: Designed to trace *approximate* reaction paths

$$F(\mathbf{Q}) = E(\mathbf{Q}) + \alpha \frac{\sum_{i\in A}\sum_{j\in B}\left[\left(R_i + R_j\right)/r_{ij}\right]^p r_{ij}}{\sum_{i\in A}\sum_{j\in B}\left[\left(R_i + R_j\right)/r_{ij}\right]^p}$$

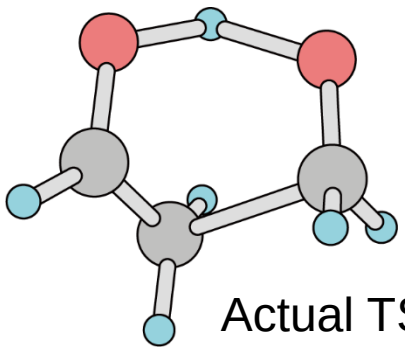
$E(\mathbf{Q})$**:** PES, α: strength of force, R_i: covalent radii, r_{ij}: *i-j* distance, p: an integer (set to 6)

Maeda, S.; Morokuma, K. *J. Chem. Phys.* **2010**, *132*, 241102.

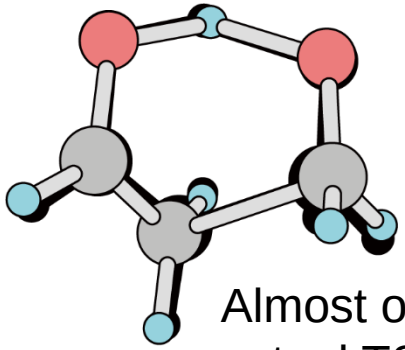
Aldol reaction (HCHO + CH_2=CHOH)



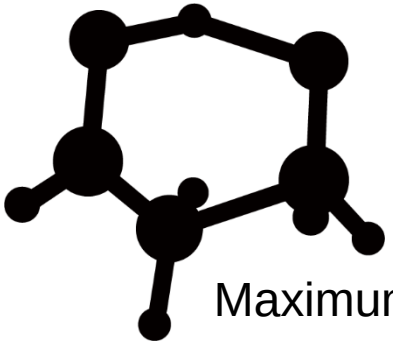
Maeda, S.; Morokuma, K. *J. Chem. Theory Comput.* **2011**, *7*, 2335.



Actual TS



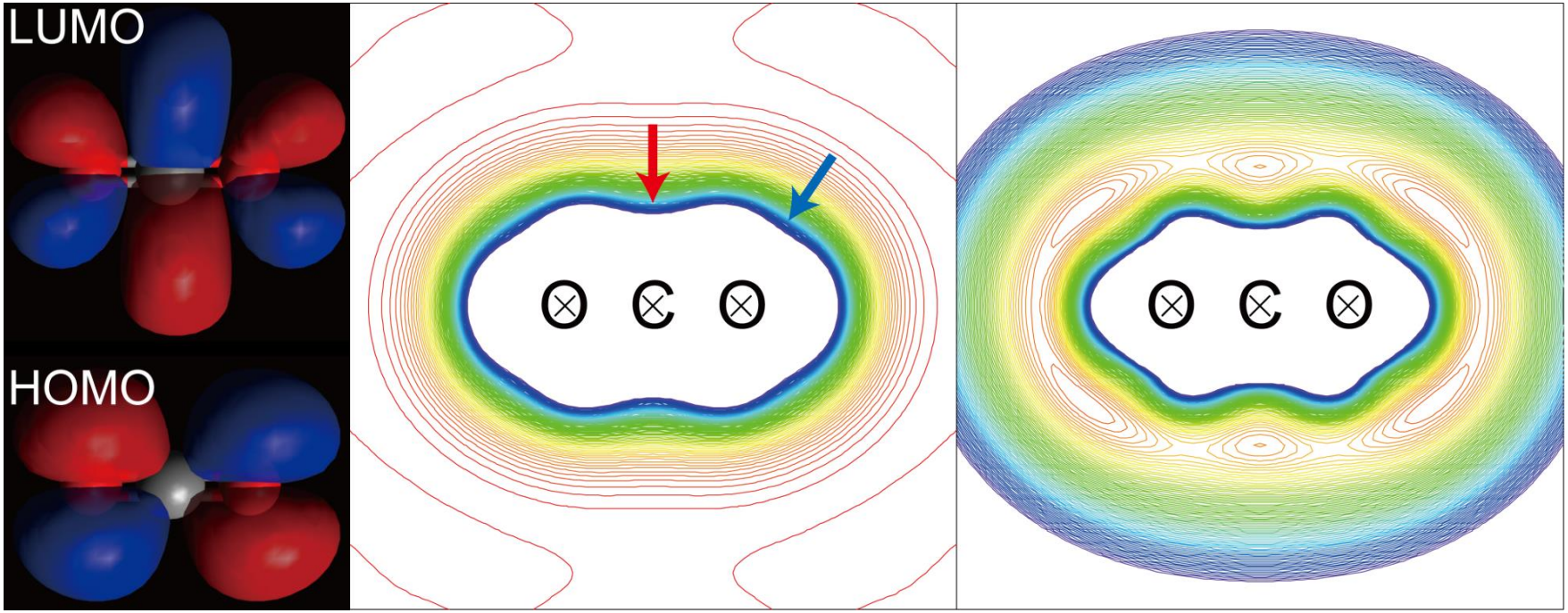
Almost on the actual TS



Maximum energy point

The AFIR function

$$F(\mathbf{Q}) = E(\mathbf{Q}) + \alpha \frac{\sum_{i\in A}\sum_{j\in B}\left[\left(R_i + R_j\right)/r_{ij}\right]^p r_{ij}}{\sum_{i\in A}\sum_{j\in B}\left[\left(R_i + R_j\right)/r_{ij}\right]^p}$$



The AFIR function gives local minima around each reactive site, and its minimization induces a structural change along a path similar to the actual reaction path

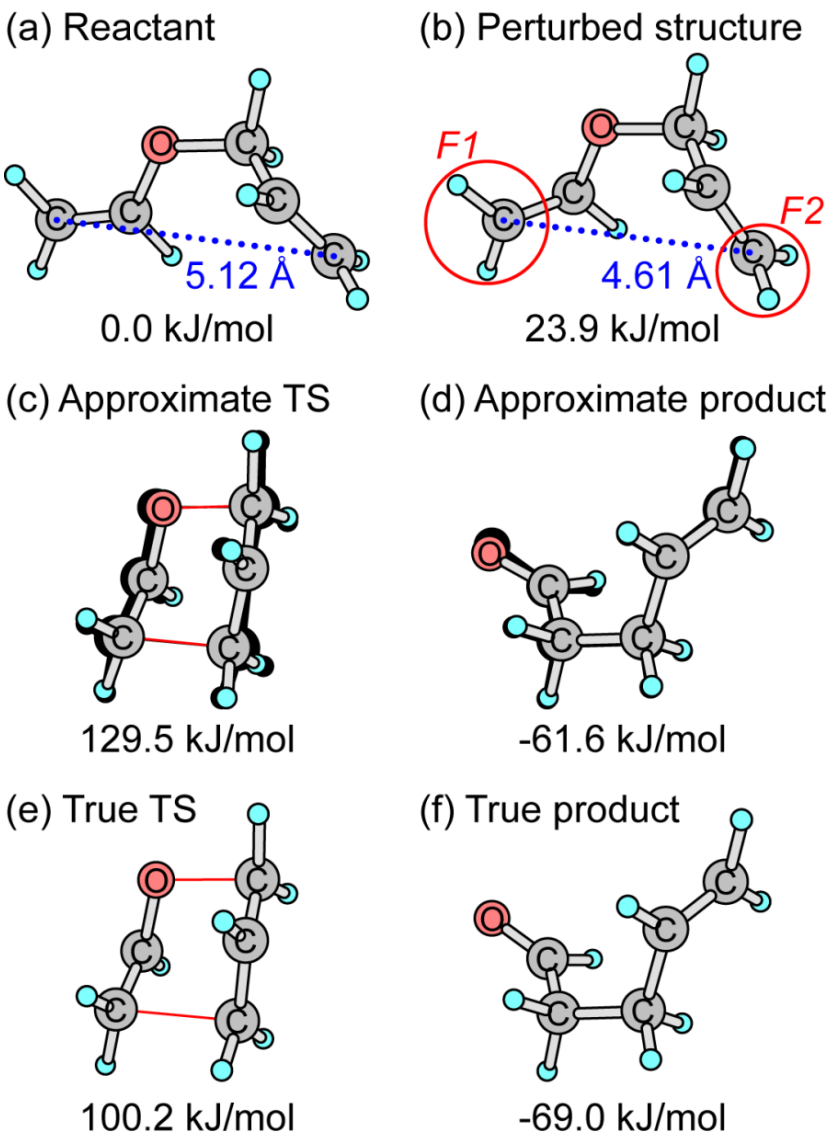
Single-component AFIR (SC-AFIR)

- **Algorithm for a single step**

1. An atom-atom pair is chosen
2. Generate a perturbed structure by reducing the atom-atom distance by 20% and a least square fitting for other distances
3. Define fragments around the chosen atoms
4. Push the fragments by AFIR
5. Optimize a product
6. Optimize a transition state (TS)
7. IRC from an obtained TS

These are applied to all atom pairs

Looks N^2 algorithm, **but much less than N^2 actually**

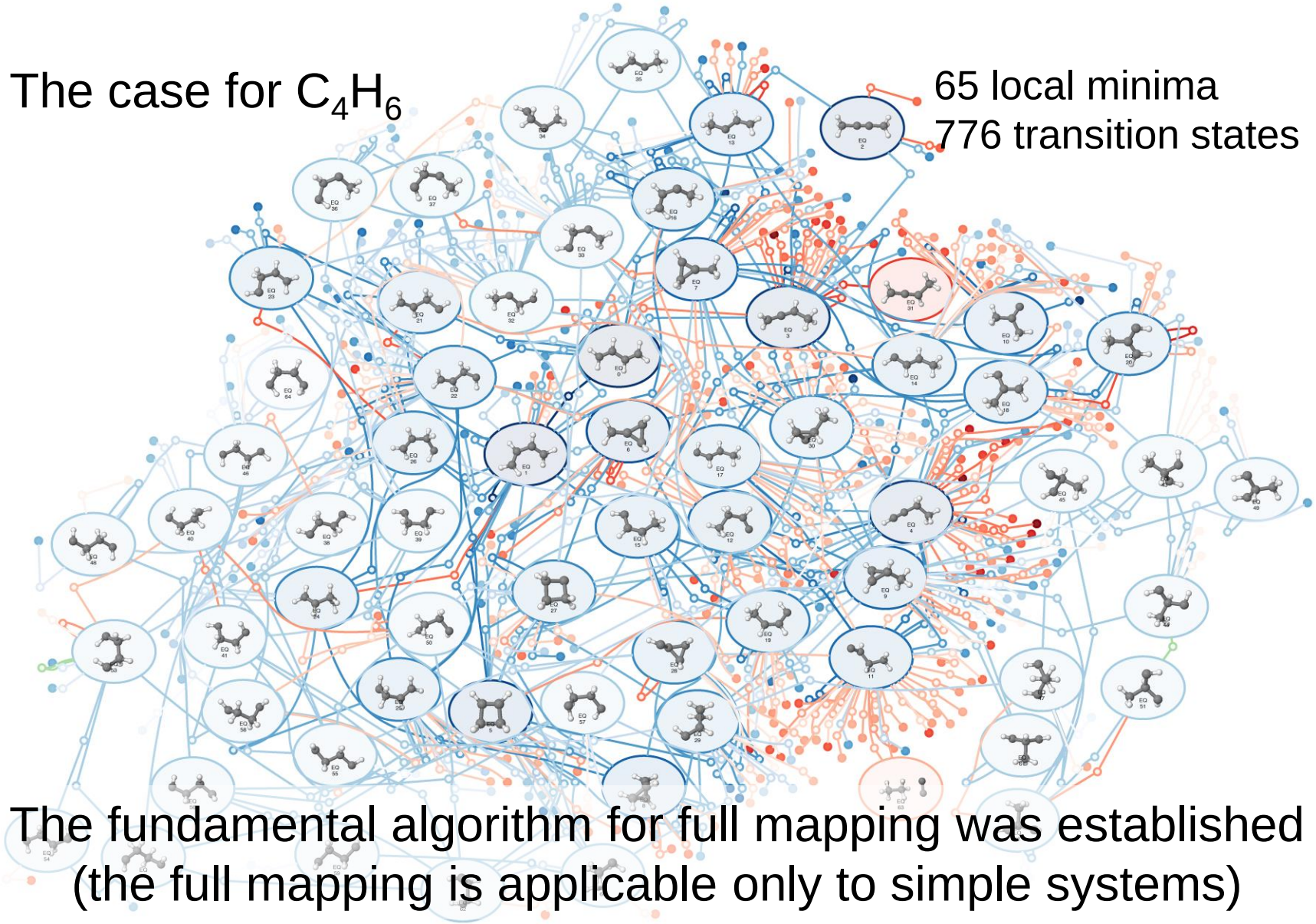


Maeda, S.; Taketsugu, T.; Morokuma, K. *J. Comput. Chem.* **2014**, *35*, 166.

Global reaction route map (GRRM)

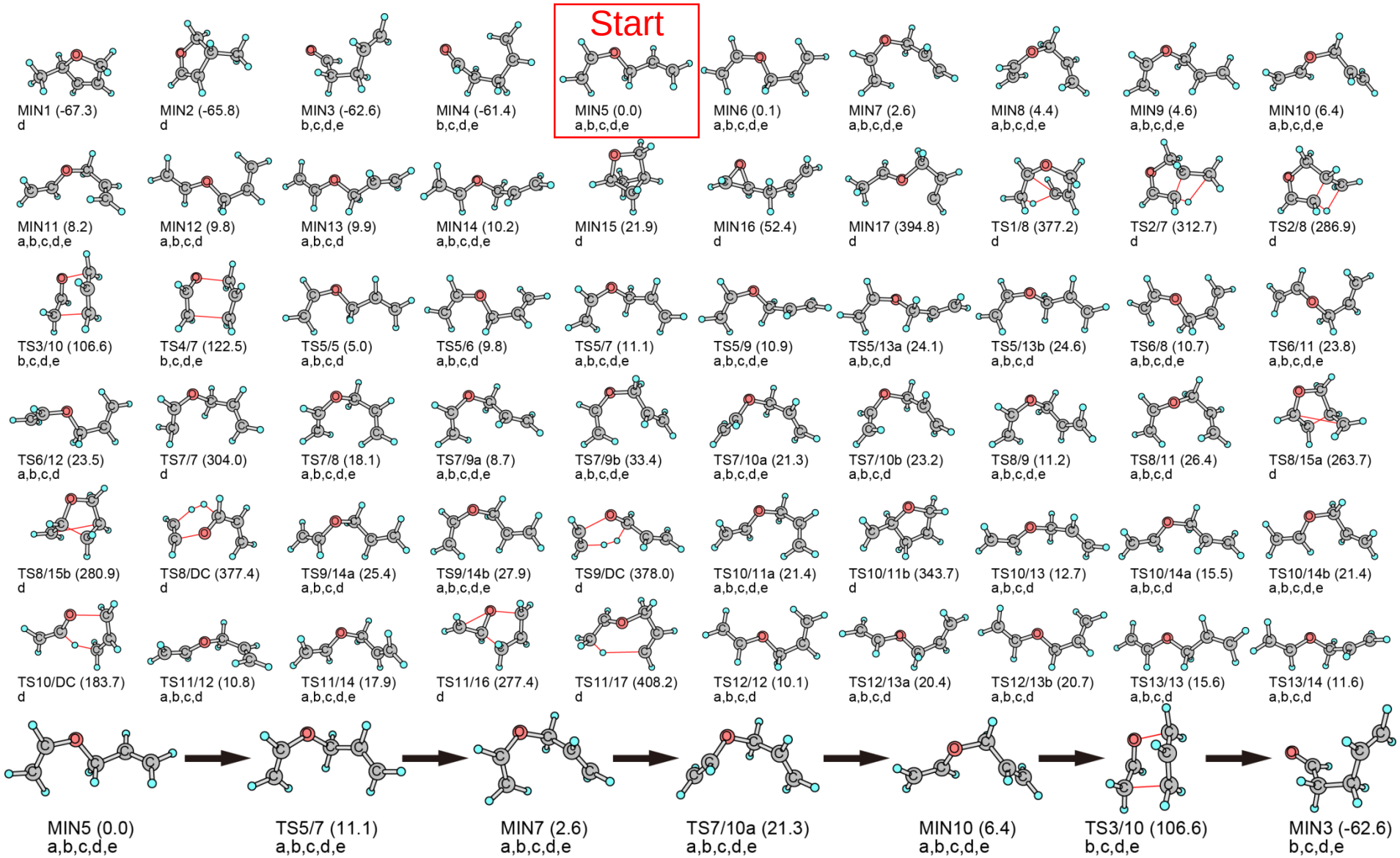
The case for C_4H_6

65 local minima
776 transition states



The fundamental algorithm for full mapping was established
(the full mapping is applicable only to simple systems)

Limit the search by bond patterns



The best path starting from a single conformer of the reactant molecule

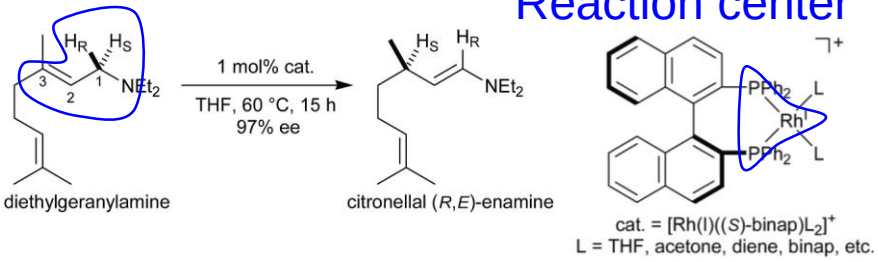
By specifying the reaction center

E.g.) Rh-BINAP catalyzed asymmetric isomerization

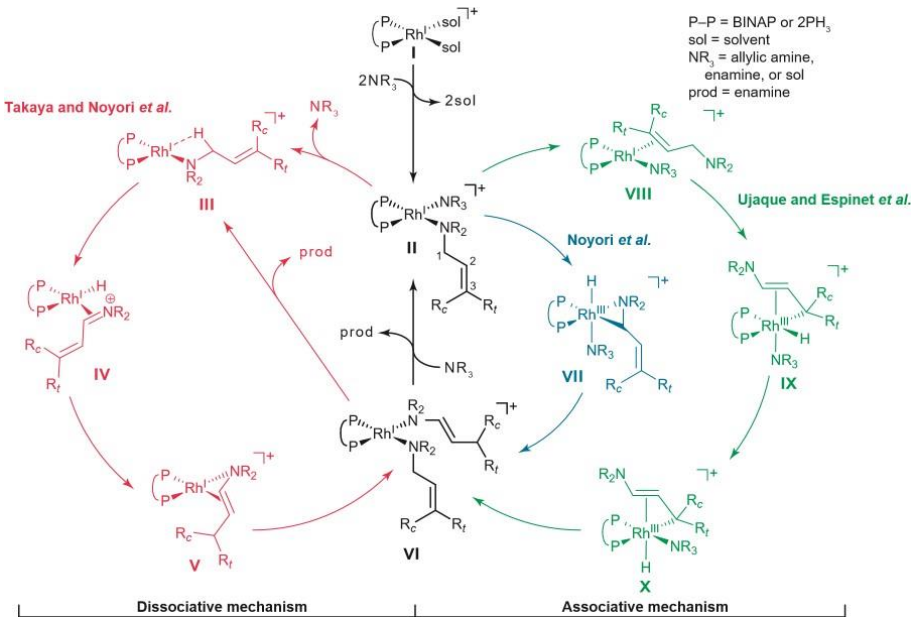
Noyori, R.; Takaya, H. *Acc. Chem. Res.* **1990**, *23*, 345.

Reaction scheme

Reaction center

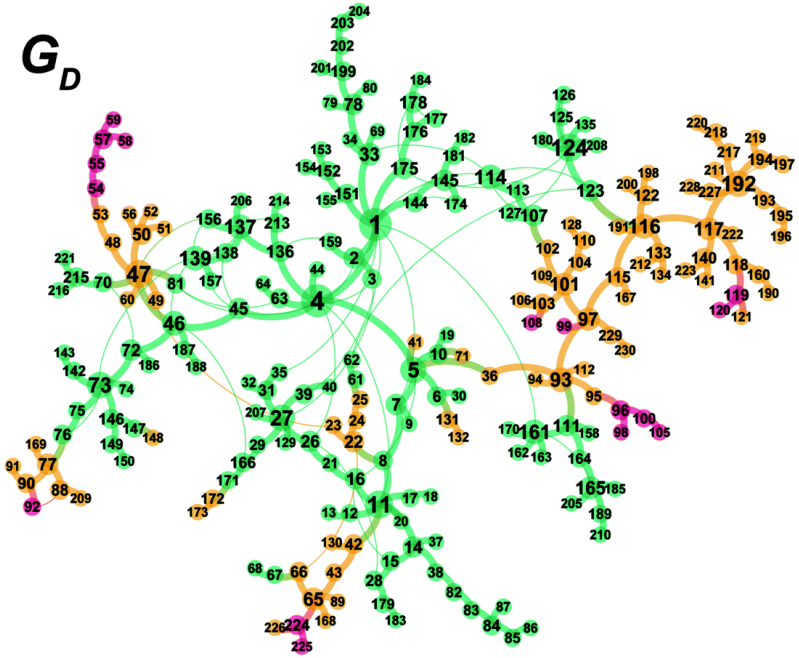


Previously proposed mechanisms



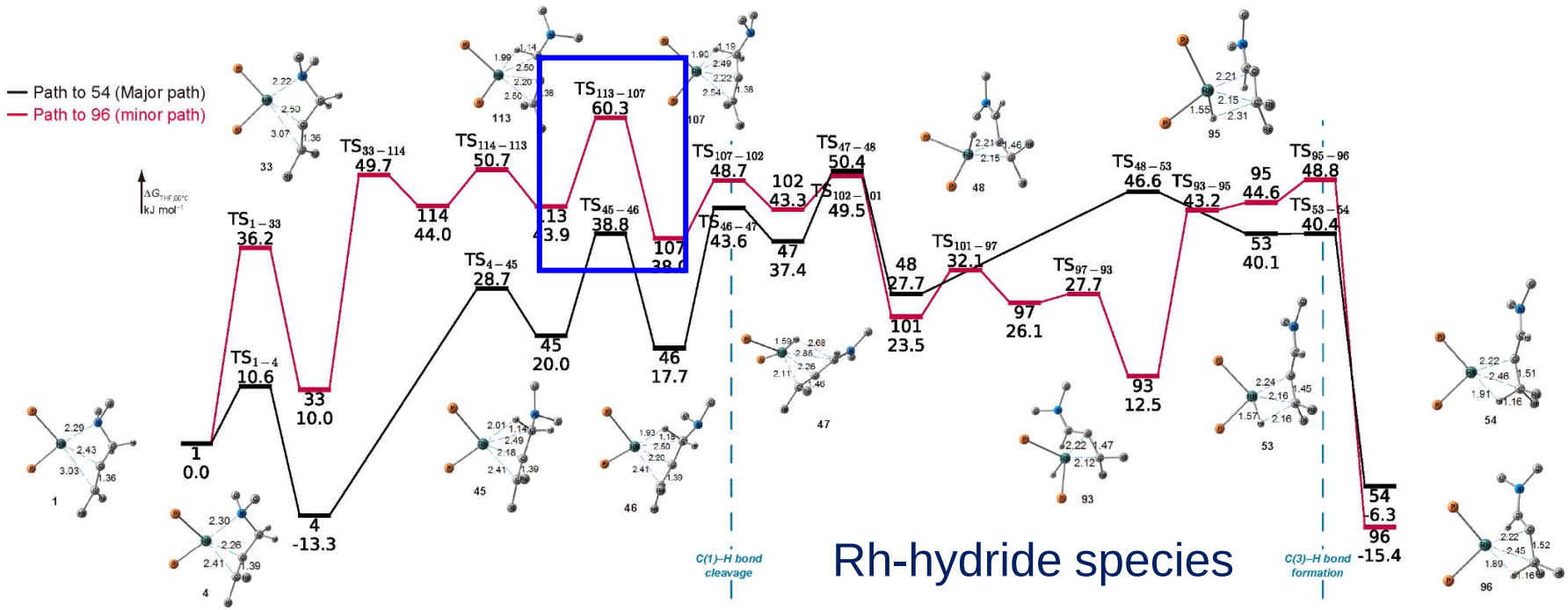
Reaction route map

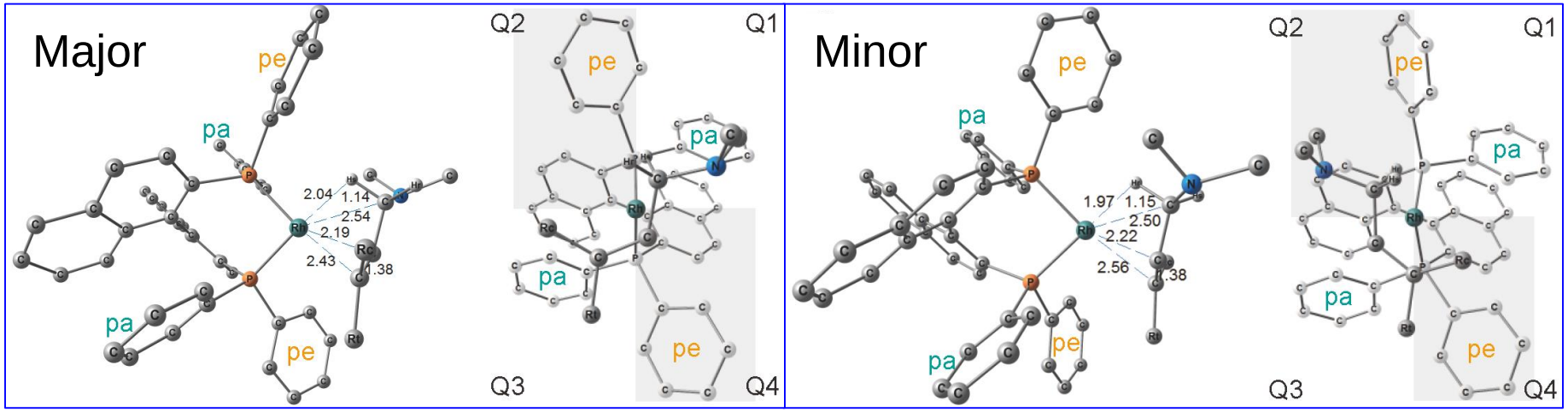
Reaction path network (the minimum spanning tree) for the Rh-BINAP catalyzed asymmetric isomerization



Yoshimura, T.; Maeda, S.; Taketsugu, T.; Sawamura, M.; Morokuma, K.; Mori, S. *Chem. Sci.* **2017**, *8*, 4475.

The best path found from the map

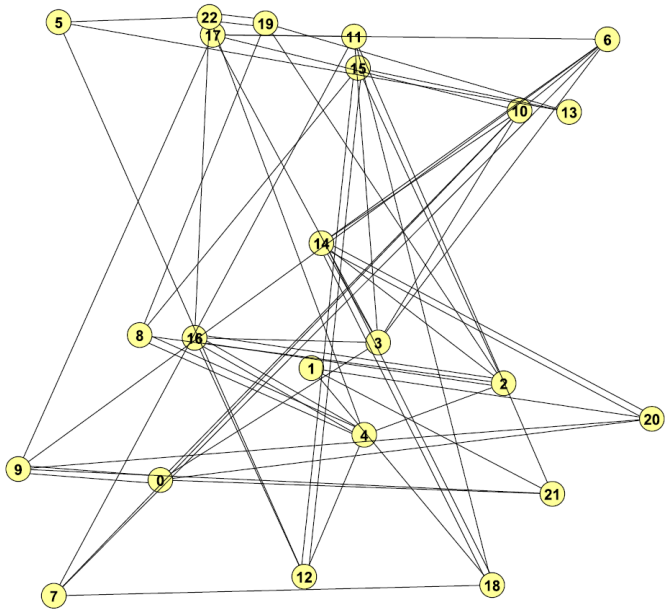


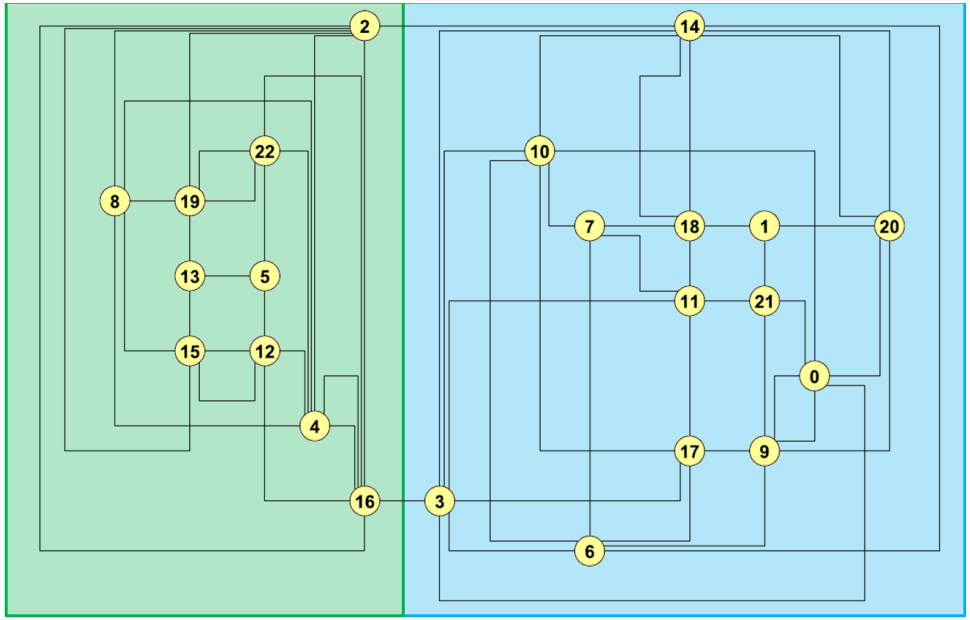


Kinetic simulation on the path map

Coarse-graining the path map

E.g.) The path network for Claisen rearrangement

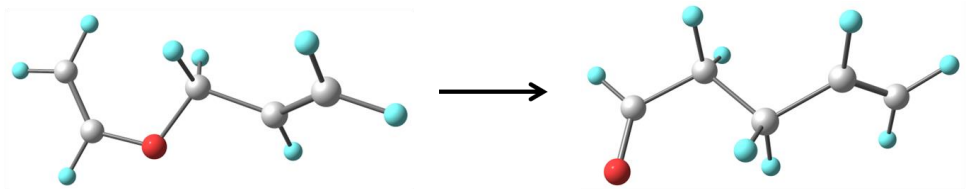




Cannot solve the rate equations!

Transform the path network into a simpler one which consists of a few superstates

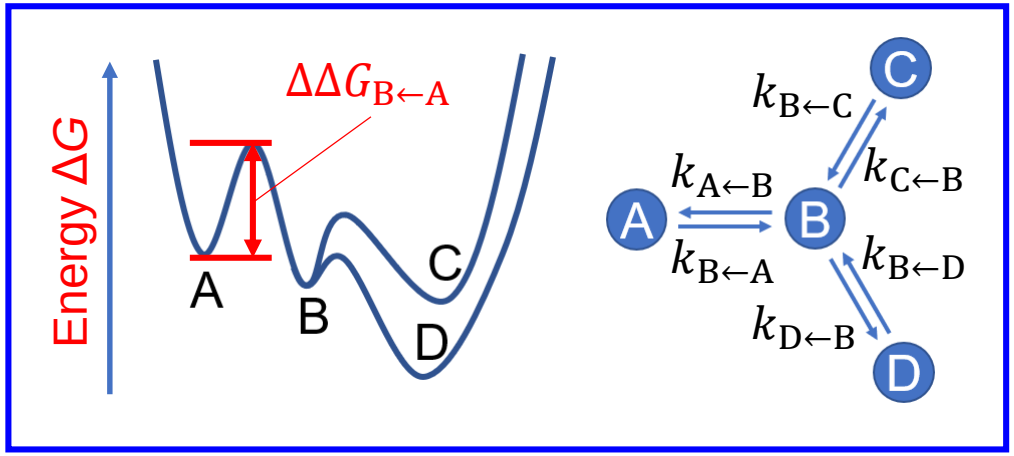
Superstate-1 Superstate-2



Describe reactions as transitions among the superstates

Sumiya, Y.; et al. *J. Phys. Chem. A* **2015**, *119*, 11641.; *J. Comput. Chem.* **2017**, *38*, 101.

Kinetic simulation



- Rate constant for A→B

$$k_{B\leftarrow A} = \varepsilon exp\left[-\frac{\Delta\Delta G_{B\leftarrow A}}{\delta}\right]$$

- Rate for A→B

$$k_{B\leftarrow A}[A]$$

(ε, δ: const., $[A]$: A's concentration)

The rate equations for time evolution of concentrations of A, B, C, D

$$\begin{bmatrix}[\dot{A}]\\[\dot{B}]\\[\dot{C}]\\[\dot{D}]\end{bmatrix} = \begin{bmatrix} -\sum_{X\neq A} k_{X\leftarrow A} & k_{A\leftarrow B} & 0 & 0 \\ k_{B\leftarrow A} & -\sum_{X\neq B} k_{X\leftarrow B} & k_{B\leftarrow C} & k_{B\leftarrow D} \\ 0 & k_{C\leftarrow B} & -\sum_{X\neq C} k_{X\leftarrow C} & 0 \\ 0 & k_{D\leftarrow B} & 0 & -\sum_{X\neq D} k_{X\leftarrow D} \end{bmatrix} \begin{bmatrix}[A]\\[B]\\[C]\\[D]\end{bmatrix}$$

($[\dot{A}]$: the first derivative on *t* of [A])

In general, the rate equations are solved numerically (Runge–Kutta method)

Problem and our solution

Rate equation: $$\begin{bmatrix} [\dot{\mathrm{A}}] \\ [\dot{\mathrm{B}}] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} -\sum_{\mathrm{X}\neq\mathrm{A}} k_{\mathrm{X}\leftarrow\mathrm{A}} & k_{\mathrm{A}\leftarrow\mathrm{B}} & \cdots \\ k_{\mathrm{B}\leftarrow\mathrm{A}} & -\sum_{\mathrm{X}\neq\mathrm{B}} k_{\mathrm{X}\leftarrow\mathrm{B}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}}_{\text{Rate constant matrix}} \begin{bmatrix} [\mathrm{A}] \\ [\mathrm{B}] \\ \vdots \end{bmatrix}$$

Rate constants with highly different magnitudes (10^{-3}~10^{13} s^{-1}) coexist

Rate constant matrix contraction (RCMC) method

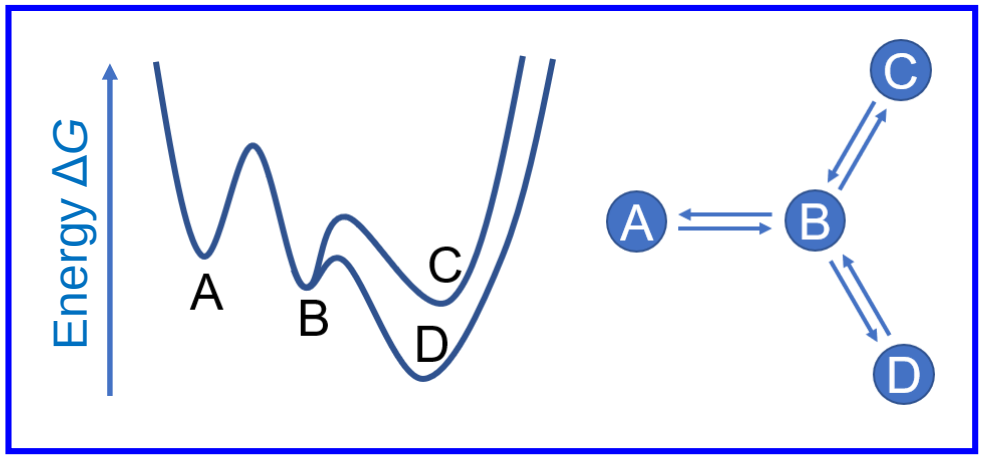
$$\begin{bmatrix} -k_{11}^0 & k_{12}^0 & \cdots & k_{1(n-1)}^0 & k_{1n}^0 \\ k_{21}^0 & -k_{22}^0 & \cdots & k_{2(n-1)}^0 & k_{2n}^0 \\ \vdots & & \ddots & & \vdots \\ k_{(n-1)1}^0 & k_{(n-1)2}^0 & \cdots & -k_{(n-1)(n-1)}^0 & k_{(n-1)n}^0 \\ k_{n1}^0 & k_{n2}^0 & \cdots & k_{n(n-1)}^0 & -k_{nn}^0 \end{bmatrix} \rightarrow \begin{bmatrix} -k_{11}^1 & k_{12}^1 & \cdots & k_{1(n-1)}^1 \\ k_{21}^1 & -k_{22}^1 & \cdots & k_{2(n-1)}^1 \\ \vdots & & \ddots & \vdots \\ k_{(n-1)1}^1 & k_{(n-1)2}^1 & \cdots & k_{(n-1)(n-1)}^1 \end{bmatrix}$$

Recursive contraction

$$\rightarrow [\quad] \rightarrow [\quad] \rightarrow \cdots \rightarrow \begin{bmatrix} -k_{11}^{(n-3)} & k_{12}^{(n-3)} & k_{13}^{(n-3)} \\ k_{21}^{(n-3)} & -k_{22}^{(n-3)} & k_{23}^{(n-3)} \\ k_{31}^{(n-3)} & k_{32}^{(n-3)} & -k_{33}^{(n-3)} \end{bmatrix} \rightarrow \begin{bmatrix} -k_{11}^{(n-2)} & k_{12}^{(n-2)} \\ k_{21}^{(n-2)} & -k_{22}^{(n-2)} \end{bmatrix}$$

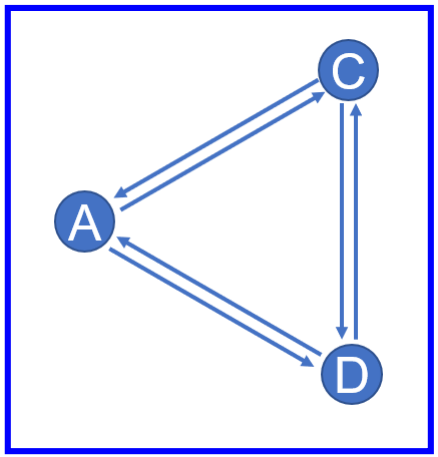
Sumiya, Y.; et al. *J. Phys. Chem. A* **2015**, *119*, 11641.; *J. Comput. Chem.* **2017**, *38*, 101.

The contraction



Contract B

Merge short life states to the others



$$\begin{bmatrix} -\sum_{X\neq A} k^{(0)}_{X\leftarrow A} & k^{(0)}_{A\leftarrow B} & 0 & 0 \\ k^{(0)}_{B\leftarrow A} & -\sum_{X\neq B} k^{(0)}_{X\leftarrow B} & k^{(0)}_{B\leftarrow C} & k^{(0)}_{B\leftarrow D} \\ 0 & k^{(0)}_{C\leftarrow B} & -\sum_{X\neq C} k^{(0)}_{X\leftarrow C} & 0 \\ 0 & k^{(0)}_{D\leftarrow B} & 0 & -\sum_{X\neq D} k^{(0)}_{X\leftarrow D} \end{bmatrix} \rightarrow \begin{bmatrix} -\sum_{X\neq A} k^{(1)}_{X\leftarrow A} & 0 & k^{(1)}_{A\leftarrow C} & k^{(1)}_{A\leftarrow D} \\ 0 & 0 & 0 & 0 \\ k^{(1)}_{C\leftarrow A} & 0 & -\sum_{X\neq C} k^{(1)}_{X\leftarrow C} & k^{(1)}_{C\leftarrow D} \\ k^{(1)}_{D\leftarrow A} & 0 & k^{(1)}_{D\leftarrow C} & -\sum_{X\neq D} k^{(1)}_{X\leftarrow D} \end{bmatrix}$$

$$k^{(m+1)}_{C\leftarrow A} = \frac{k^{(m)}_{B\leftarrow A} k^{(m)}_{C\leftarrow B} \sigma^{(m)}_{B}}{1 + k^{(m)}_{B\leftarrow A} \sigma^{(m)}_{B}}, \sigma^{(m)}_{B} = \frac{1}{k^{(m)}_{A\leftarrow B} + k^{(m)}_{C\leftarrow B} + k^{(m)}_{D\leftarrow B}}$$

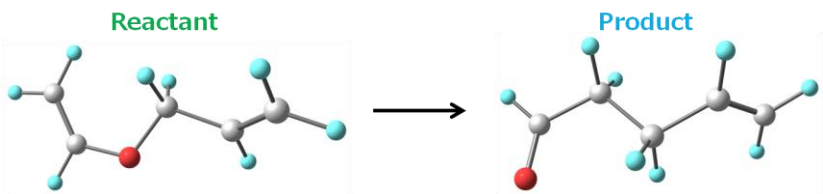
Derived from the steady state approximation and the conservation condition of the Boltzmann distribution

Sumiya, Y.; et al. *J. Phys. Chem. A* **2015**, *119*, 11641.; *J. Comput. Chem.* **2017**, *38*, 101.

Claisen rearrangement

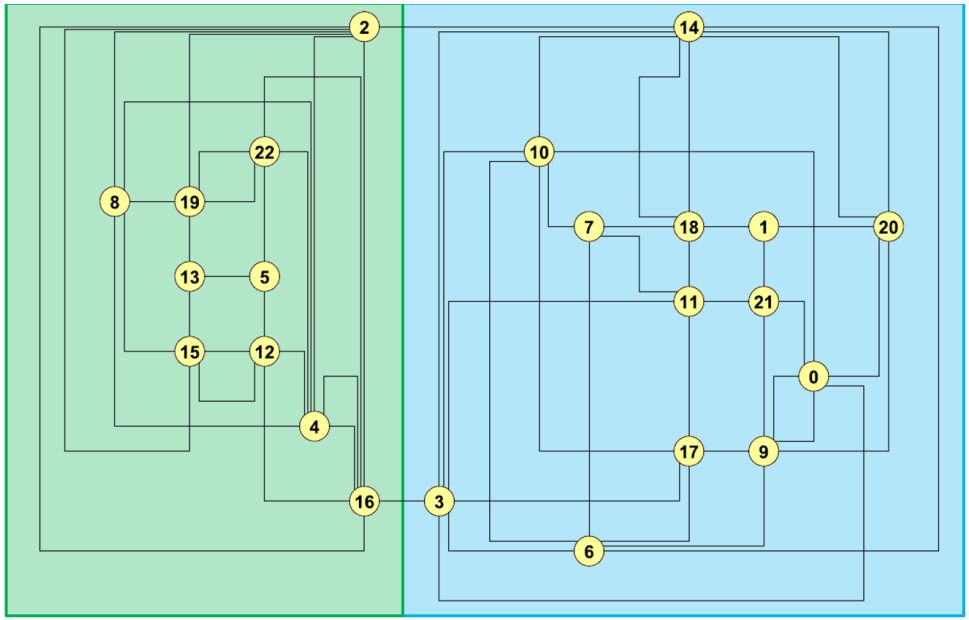
Rate Constant Matrix Contraction (RCMC)

Reaction path network:
23 MINs and 63 TSs
23×23 matrix → 2×2 matrix



Superstate-1 Superstate-2

CCSD(T)-F12a/jul-cc-pVTZ//M062X/6-311+G(2d,p)



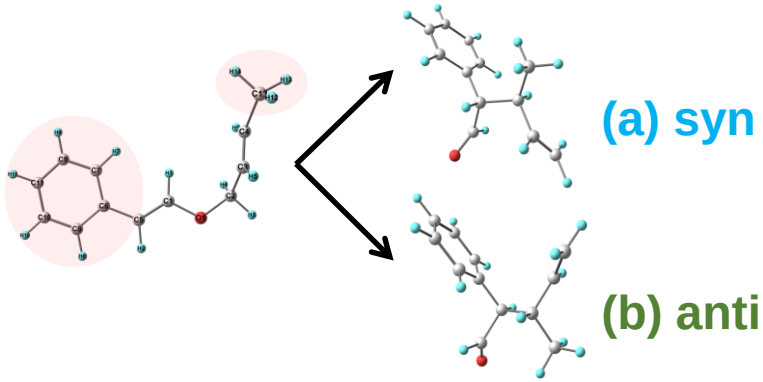
Superstate-1 Superstate-2

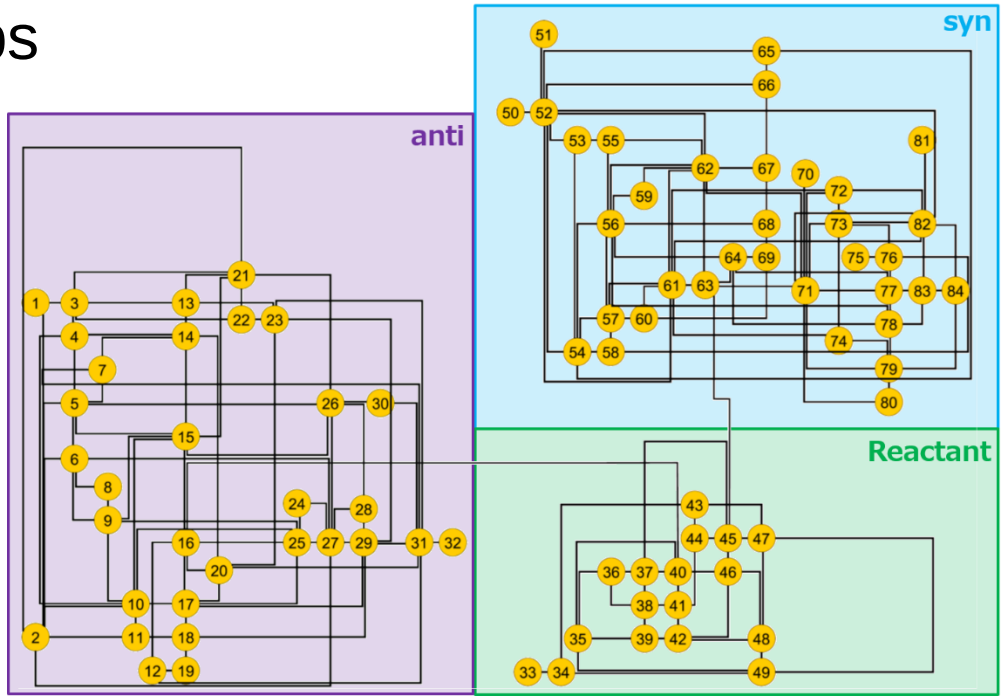
	k / s^{-1}	$\Delta\Delta G$ / kJ mol^{-1}	Error / kJ mol^{-1}
Conventional	1.486×10^{-2}	133.1	6.4
This work (RCMC)	1.794×10^{-3}	141.3	−1.8
Experiment	2.875×10^{-3}	139.5	

The difference (8.2 kJ mol^{-1}) is originated from the conformational entropy

With a substituted reactant

Introduce Me and Ph groups





Reaction path network:
84 MINs and 278 TSs
84×84 matrix → 3×3 matrix

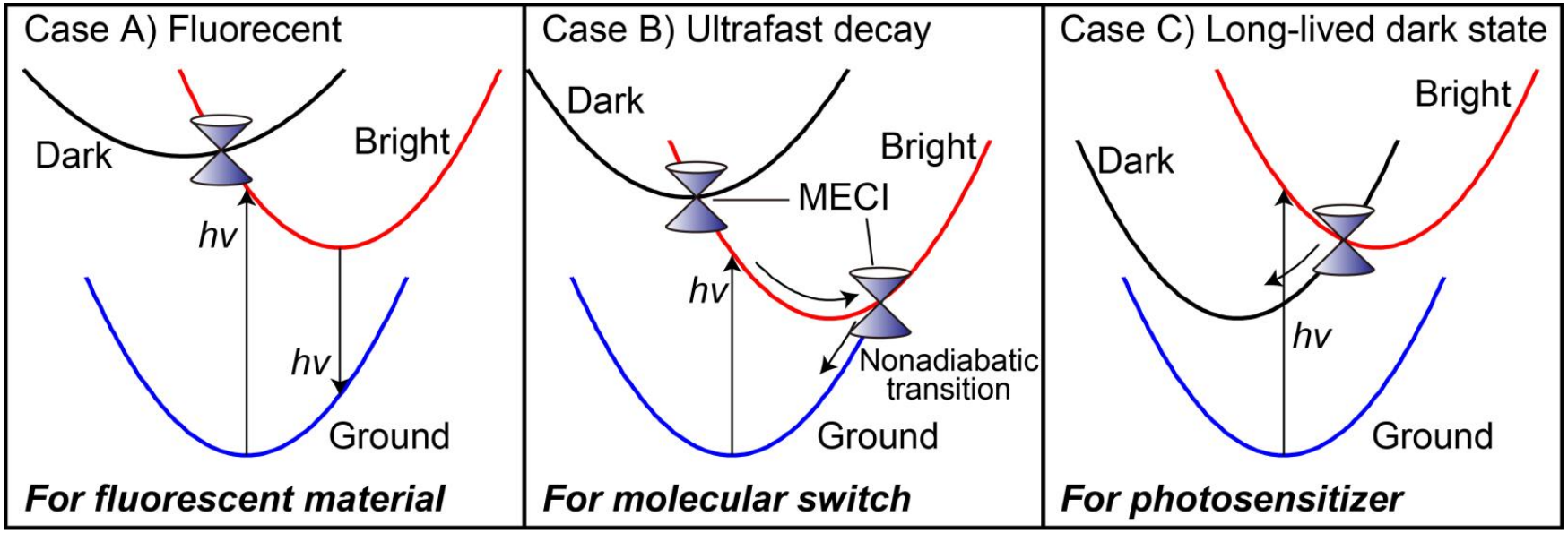
M062X/6-311+G(2d,p)

	Syn-product		*Anti*-product	
	k / s^{-1}	ΔG / kJ mol^{-1}	k / s^{-1}	ΔG / kJ mol^{-1}
RDES model	2.231×10^{-2}	131.5	4.848×10^{-4}	146.4
This work	1.434×10^{-3}	142.2	2.601×10^{-5}	157.8

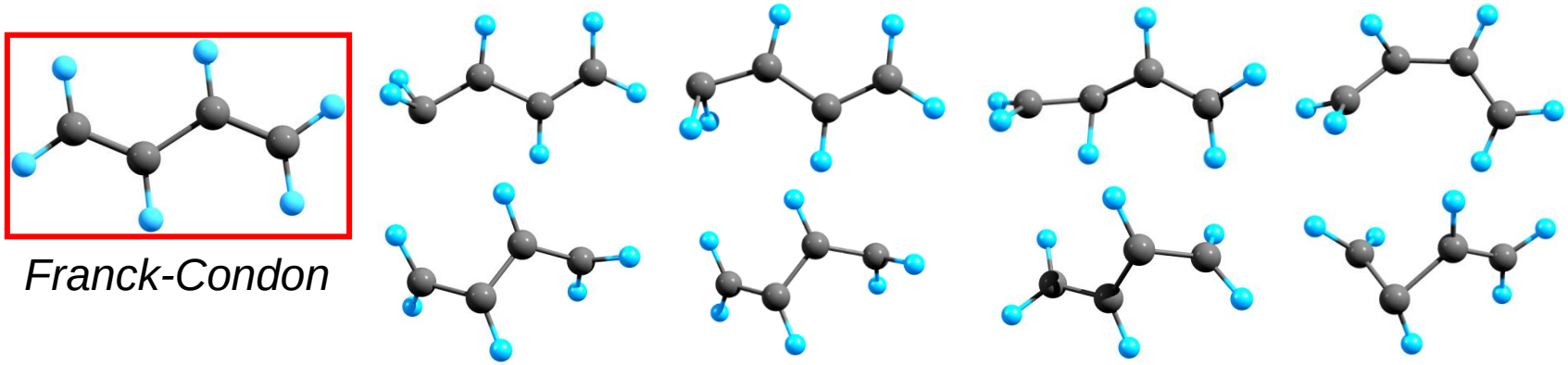
The difference (10.7 kJ mol^{-1}) is originated from the conformational entropy
Impact of the conformational entropy on the selectivity is small in this system

Applying to photochemistry

Photochemical reaction by AFIR



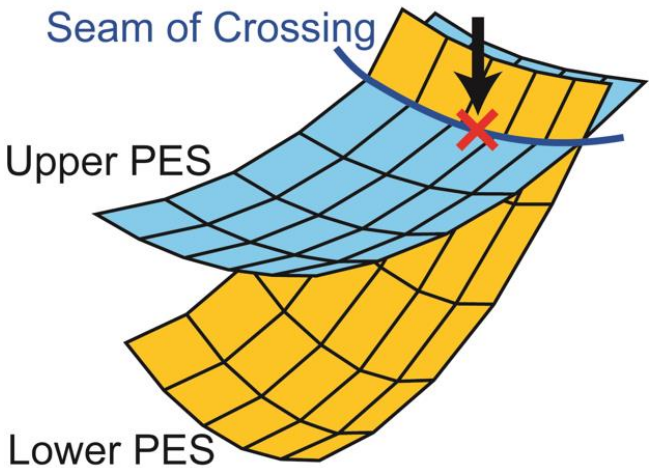
Finding all potential crossing points near the FC point



Potential crossing points near the FC point of *trans*-1,3-butadiene

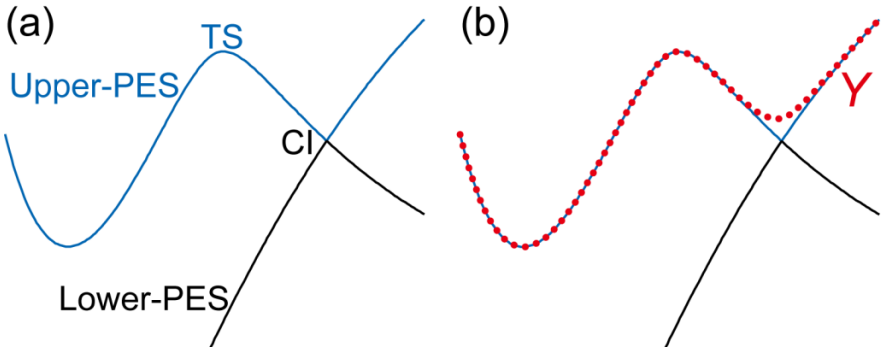
Techniques for photoreactions

SMF for automated search of surface crossing points



$$X(\mathbf{Q}) = \frac{1}{2}\left(E^{\text{State1}}(\mathbf{Q}) + E^{\text{State2}}(\mathbf{Q})\right) + \frac{\left(E^{\text{State1}}(\mathbf{Q}) - E^{\text{State2}}(\mathbf{Q})\right)^2}{\beta}$$

AMF for automated search of excited state TSs



$$Y(\mathbf{Q}) = \frac{1}{2}\left(E^{\text{State1}}(\mathbf{Q}) + E^{\text{State2}}(\mathbf{Q})\right) + \frac{1}{2}\sqrt{\left(E^{\text{State1}}(\mathbf{Q}) - E^{\text{State2}}(\mathbf{Q})\right)^2 + 4U^2}$$

$$U(\mathbf{Q}) = \frac{\chi}{2}\exp\left[-\left(\frac{E^{\text{state1}}(\mathbf{Q}) - E^{\text{state2}}(\mathbf{Q})}{\chi}\right)^2\right]$$

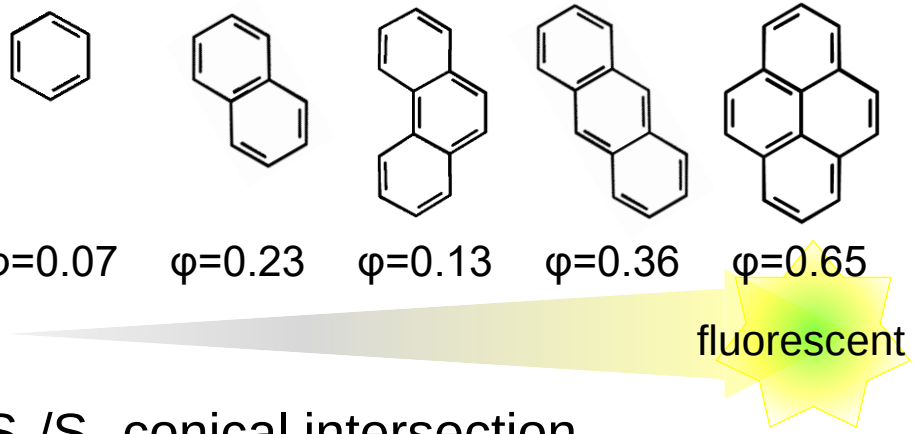
GP/BPU for MECI optimization without derivative coupling

$$\mathbf{y}_k = \frac{(\mathbf{y}_{k-1}\cdot\mathbf{x}_k)\mathbf{x}_{k-1} - (\mathbf{x}_{k-1}\cdot\mathbf{x}_k)\mathbf{y}_{k-1}}{\sqrt{(\mathbf{y}_{k-1}\cdot\mathbf{x}_k)^2 + (\mathbf{x}_{k-1}\cdot\mathbf{x}_k)^2}}$$

Predicting luminescence ability

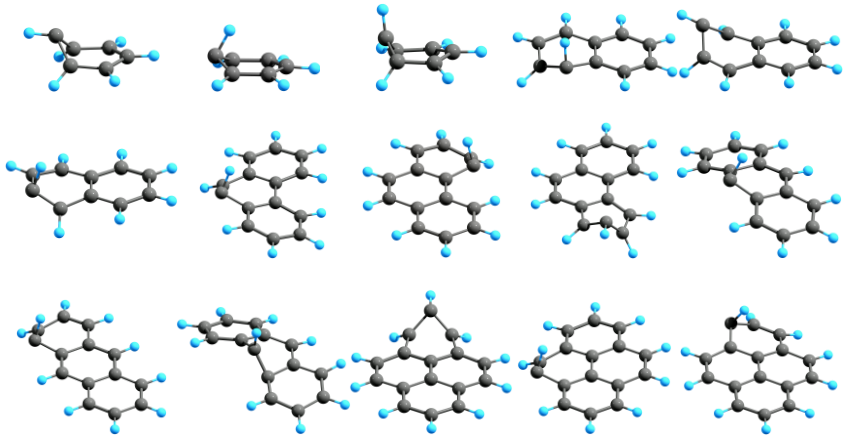
Obsd. fluorescence quantum yield φ[1]

[1] CSJ, Kagaku Binran (Kisohen 1) 5th ed., Maruzen, Tokyo, 2004.



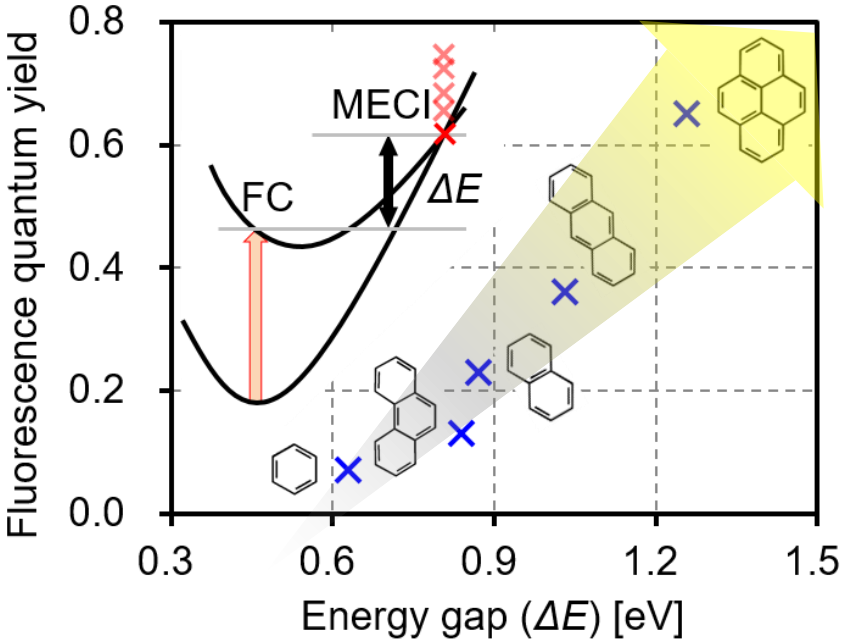
S_0/S_1 conical intersection

*) The three lowest energy structures



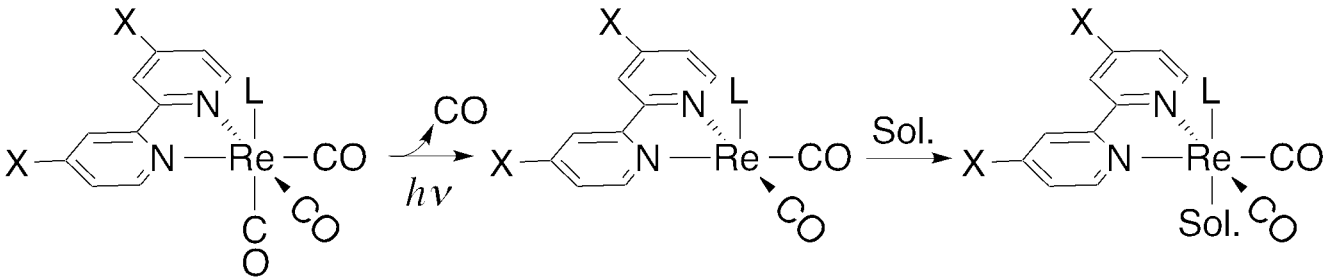
φ vs. energy of crossing points

A clear correlation was established between experimental φ and energy values of the lowest crossing points identified by the automated search



Harabuchi, Y.; Taketsugu, T.; Maeda, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22561.

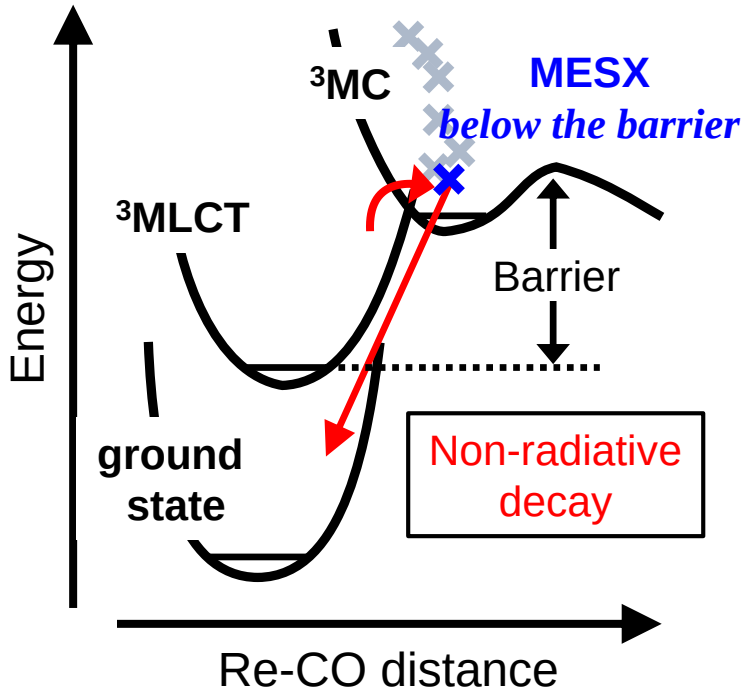
PLS reaction of Re(I) complex



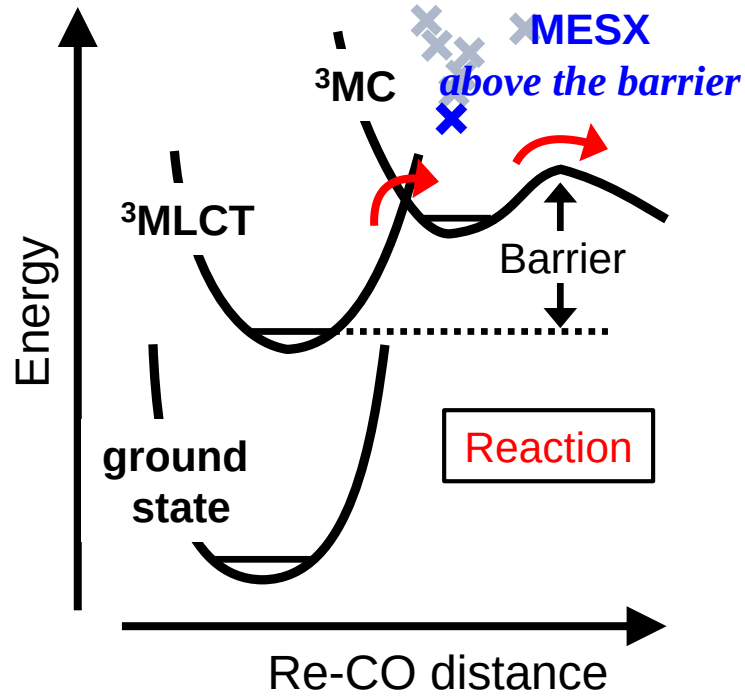
L = P(OMe)$_3$: *YES*
L = Cl$^-$: *NO*

Koike, K.; Okoshi, N.; Hori, H.; Takeuchi, K.; Ishitani, O.; Tsubaki, H.; Clark, I. P.; George, M. W.; Johnson, F. P. A.; Turner, J. J., J. Am. Chem. Soc. 2002, 124, 11448.

L = Cl$^-$ (photostable)



L = P(OMe)$_3$ (photoreactive)



Saita, K.; et al., *Phys. Chem. Chem. Phys.* **2016**, *18*, 17557.

Implementation of AFIR in GRRM17

AFIR for molecular systems will be available in GRRM17

- Global / semiglobal reaction path network
- Local reaction path search
- Local / global potential crossing points search
- Interface with QM/MM-ONIOM
- Runs with any ab initio code (a simple interface script is needed)

Software News and Updates: S. Maeda, Y. Harabuchi, M. Takagi, K. Saita, K. Suzuki, T. Ichino, Y. Sumiya, K. Sugiyama, Y. Ono, *J. Comput. Chem*. **2018**, *39*, 233 (***Open Access***).

AFIR web (under construction): afir.sci.hokudai.ac.jp

For *elucidating* reaction mechanisms,
predicting unknown structures and pathways,
generating databases of structures and pathways, etc..

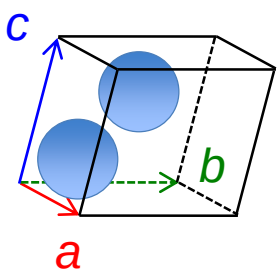
Applying to periodic systems: crystal and surface

Implementation for PBCs

Periodic boundary conditions (PBCs)

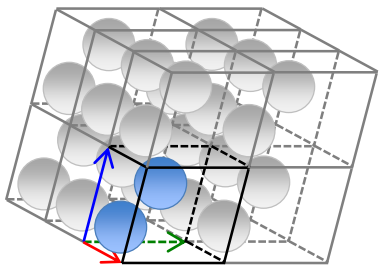
Represent a system as repetition of an unit-cell, in first principle calculations

Unit-cell



repeat

Crystal, surface



Deformation of the unit-cell

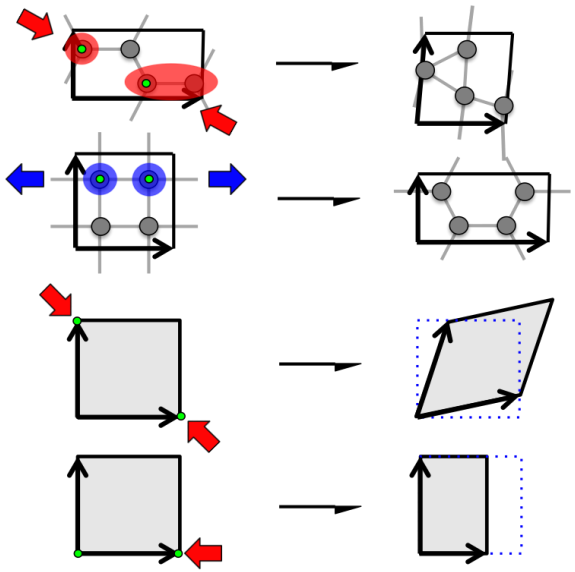
The gradient of coordinates of translational vectors (TVs) is obtained from the stress tensor

$$\left.\frac{\partial E}{\partial \mathbf{h}}\right|_{\mathbf{h}_0} = -V_0 \sigma^t \mathbf{h}_0^{-1}$$

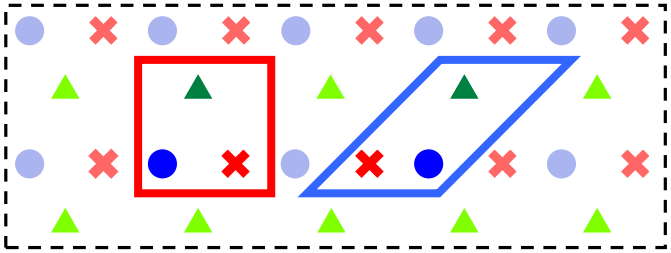
E : Energy
h: TVs
V_0: Cell volume
s: Stress tensor

Issues specific for periodic systems

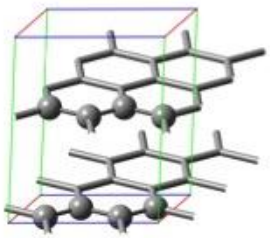
How to induce deformations of TVs (treat $3N+9-6$ variables)



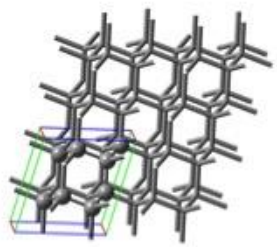
Two or more representations possible for one structure



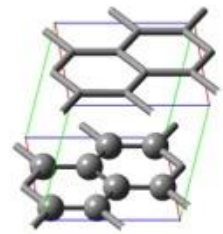
Crystal structure search (C_8/cell)



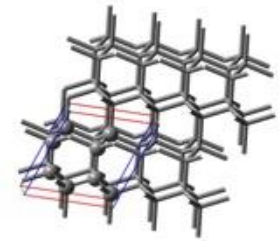
3D-min0 graphite
$P6_3/mmc$
0.0 kJ/mol·atoms



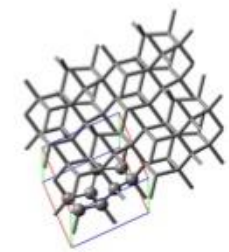
3D-min1 diamond
$Fd\bar{3}m$
0.1 kJ/mol ·atoms



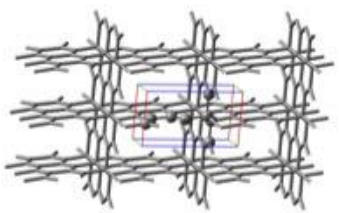
3D-min2 graphite
$R\bar{3}m$
0.9 kJ/mol·atoms



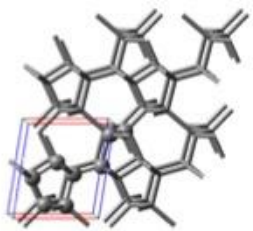
3D-min3 lonsdaleite
$P6_3/mmc$
2.8 kJ/mol ·atoms



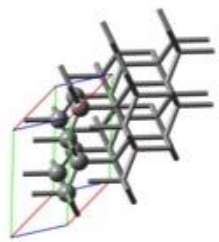
3D-min4 Cco-C_8
$Cmmm$
12.7 kJ/mol·atoms



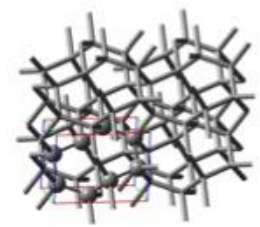
3D-min5 crossed graphene
$P1$
15.2 kJ/mol·atoms



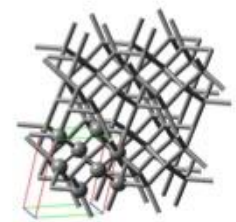
3D-min6 *M*-carbon
$P2/m$
16.2 kJ/mol·atoms



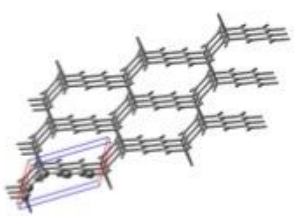
3D-min7 8-tetra(2,2) tubulane
$I4/mmm$
17.9 kJ/mol·atoms



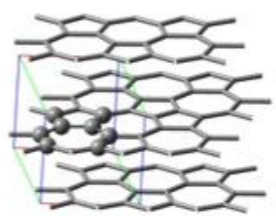
3D-min8 LA5
$P1$
18.9 kJ/mol·atoms



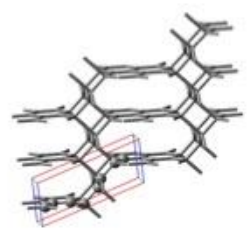
3D-min9 *C*2/*m*-16
$P1$
18.9 kJ/mol·atoms



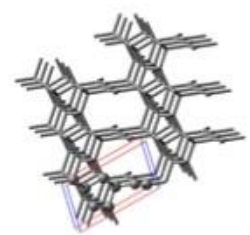
3D-min10
$P2/m$
21.5 kJ/mol·atoms



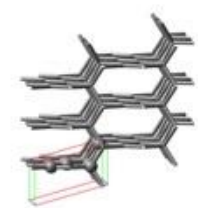
3D-min11 net2*b*
$P\bar{1}$
23.0 kJ/mol·atoms



3D-min12
$P1$
24.5 kJ/mol·atoms



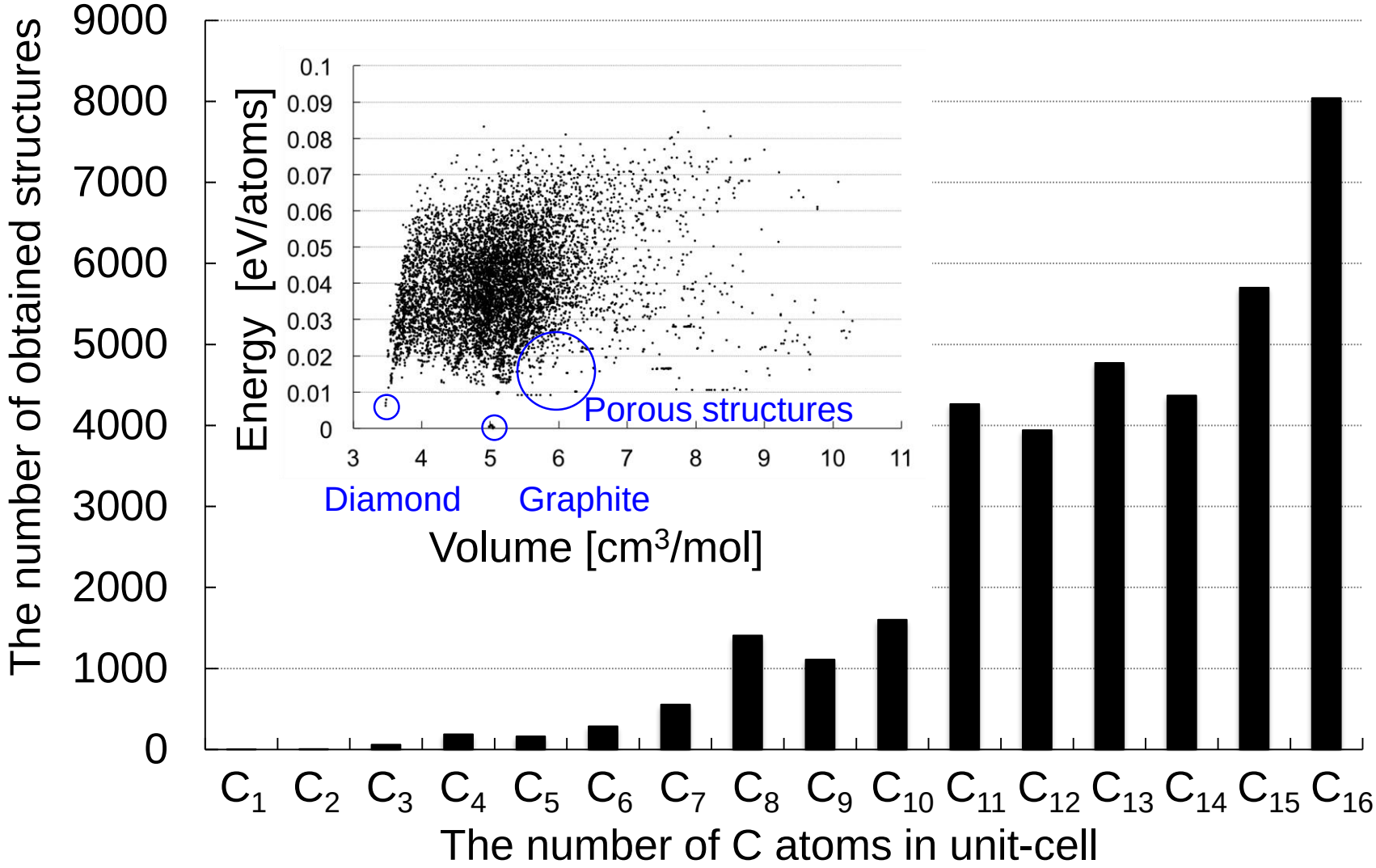
3D-min13
$P1$
25.0 kJ/mol·atoms



3D-min14 3D-(5,0)
$P2_1/m$
26.0 kJ/mol·atoms

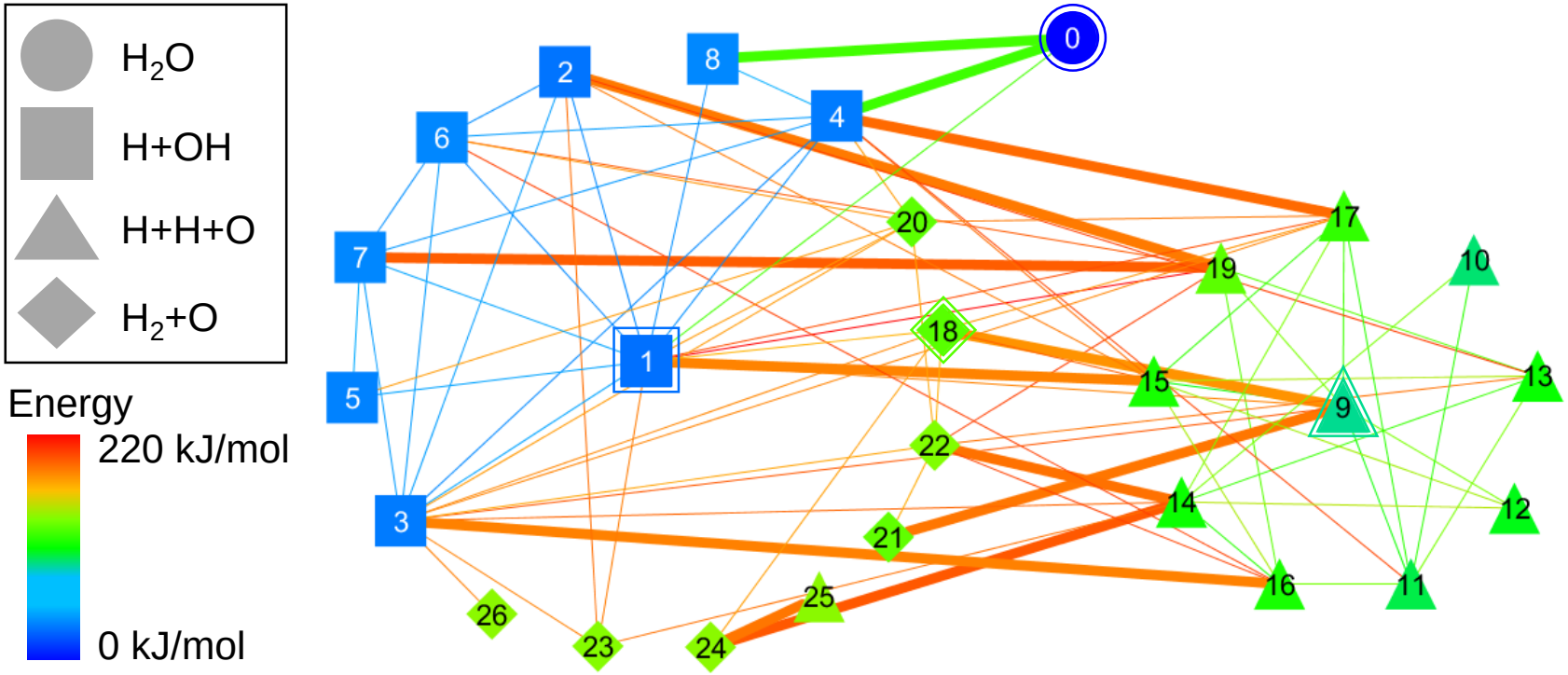
Takagi, M.; Taketsugu, T.; Kino, H.; Tateyama, Y.; Terakura, K.; Maeda, S. *Phys. Rev. B* **2017**, *95*, 184110.

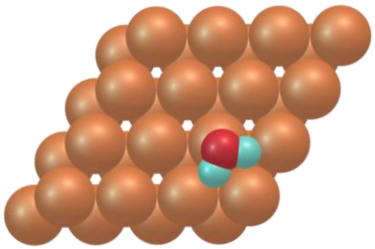
Crystal structure search ($C_{1\text{-}16}$/cell)



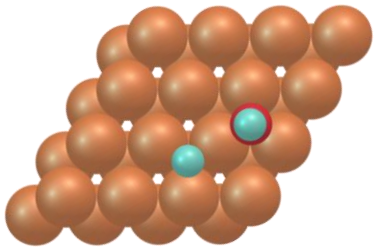
More than 10000 independent structures were predicted by the search

The path map for H_2O@Cu(111)

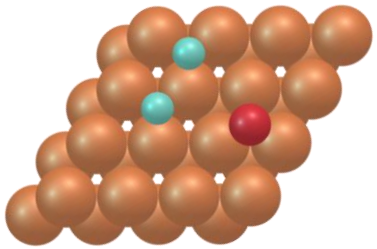




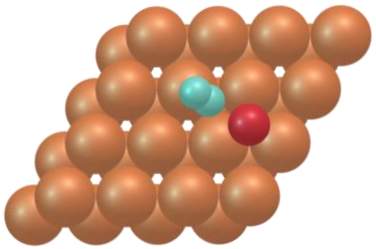
MIN0 H_2O
0.0 kJ/mol



MIN1 H+OH
32.6 kJ/mol



MIN9 H+H+O
95.9 kJ/mol

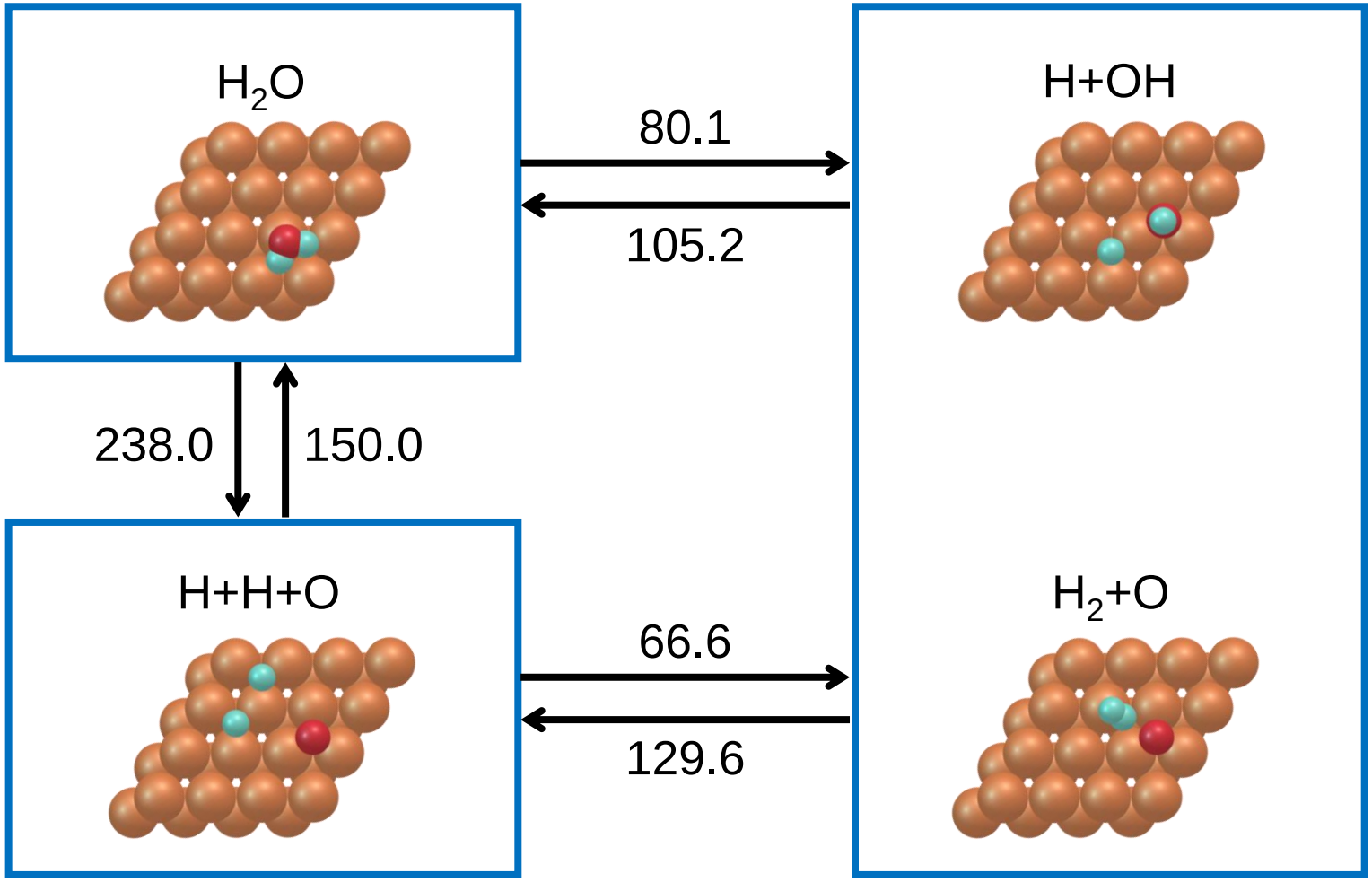


MIN18 H_2+O
131.0 kJ/mol

Maeda, S.; Sugiyama, K.; Sumiya, Y.; Takagi, M.; Saita, K. *Chem. Lett.* **2018**, (in press).

Kinetic analysis of H_2O@Cu(111)

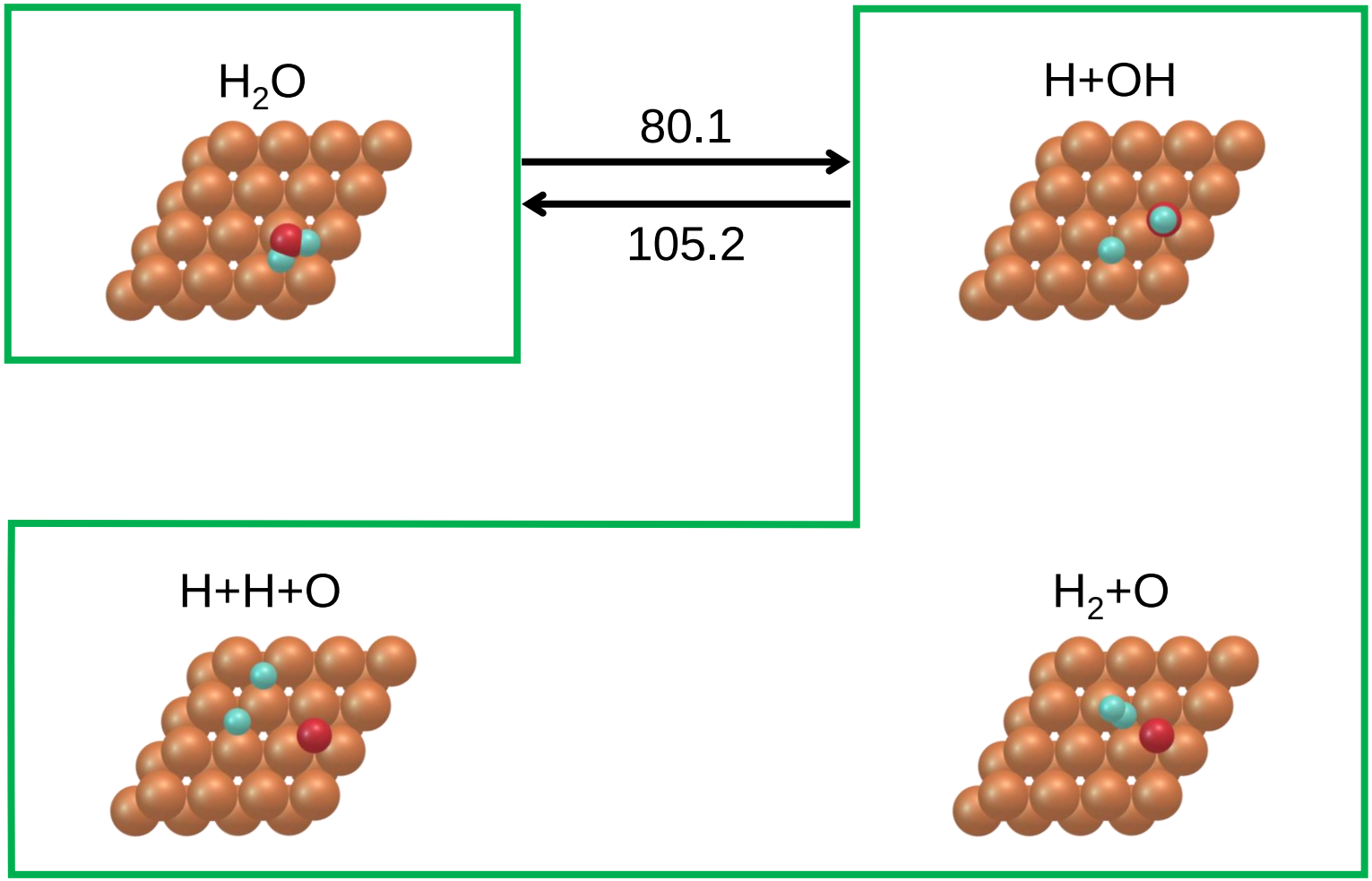
Superstates obtained by RCMC with t = 0.001 sec.



$\Delta\Delta G^{\ddagger}$ in kJ/mol

Kinetic analysis of H_2O@Cu(111)

Superstates obtained by RCMC with *t* = 1.0 sec.

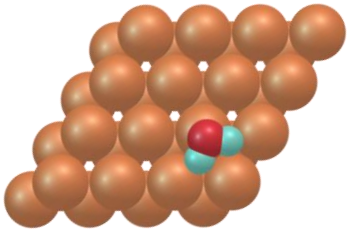


$\Delta\Delta G^{\ddagger}$ in kJ/mol

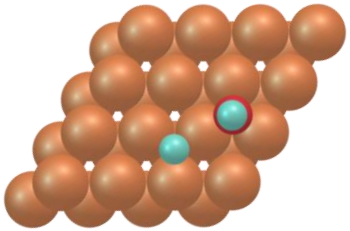
Kinetic analysis of H_2O@Cu(111)

Superstates obtained by RCMC with t = 1000.0 sec.

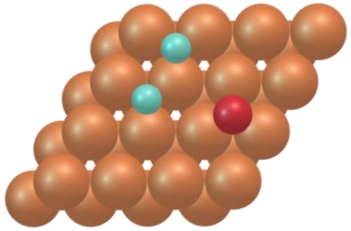
H_2O



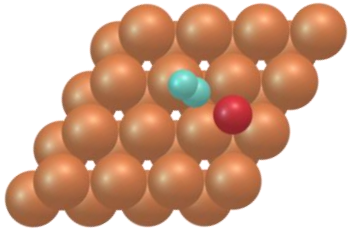
H+OH



H+H+O



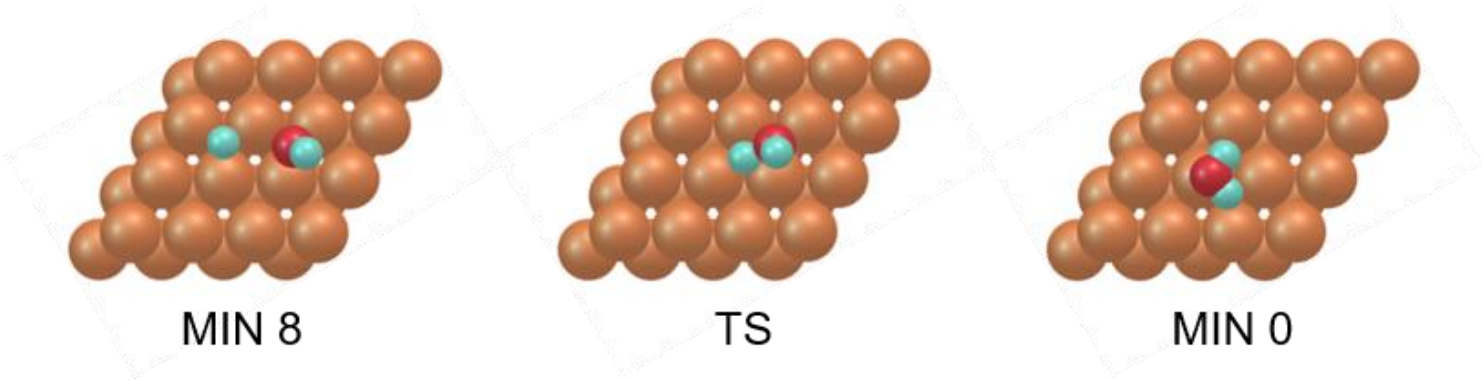
H_2+O



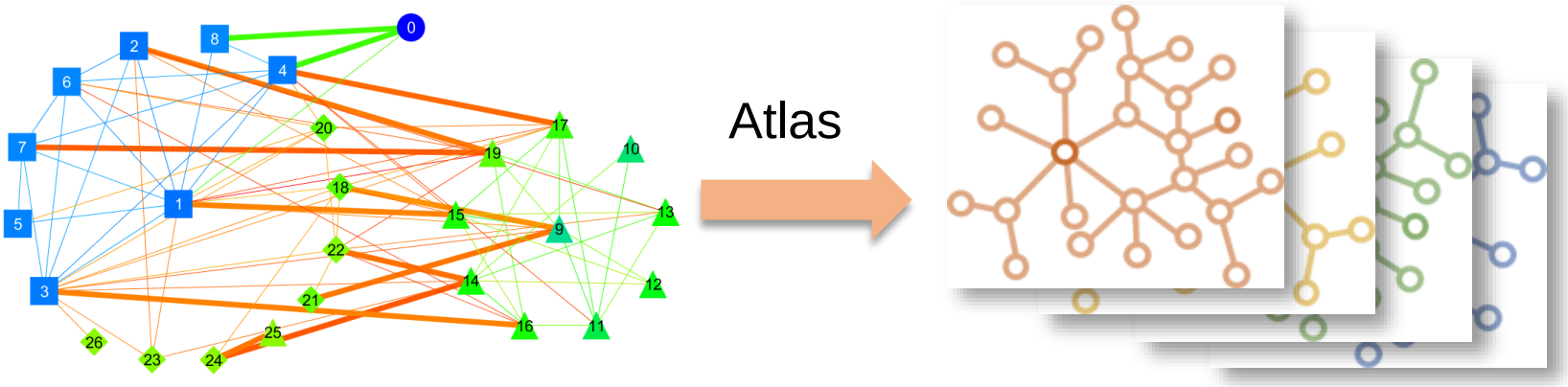
$\Delta\Delta G^{\ddagger}$ in kJ/mol

The impact

Intermediates (~2000) → reaction paths that can be assumed (present)



This study (automated search by the AFIR method) → reaction path map



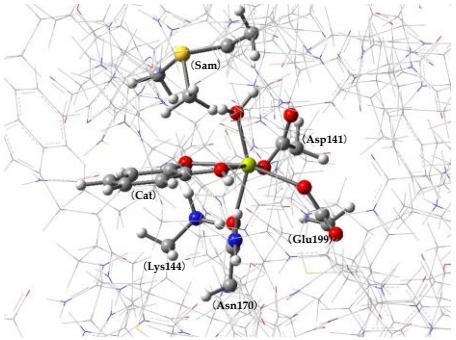
Each map can be obtained in ~1 week using single computer

→ Each single map to atlas toward automated discovery of new reactions

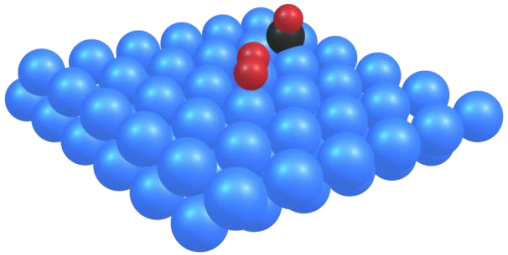
Application scope of AFIR

AFIR is applicable to reactions of various types

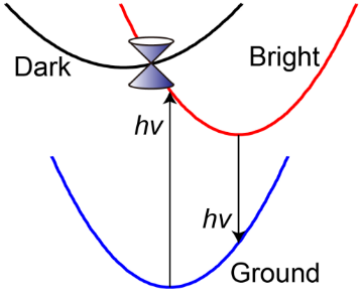
Enzyme (QM/MM)



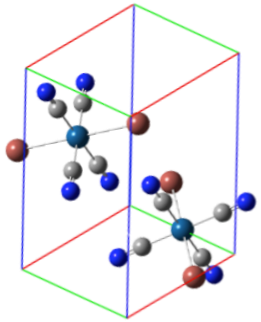
Surface (PBCs)



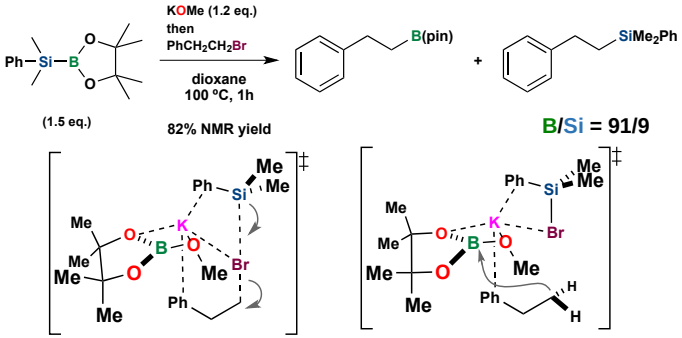
Photoreaction



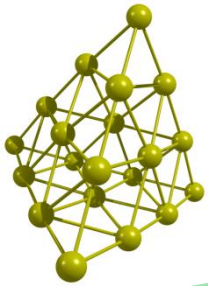
Photoreaction in crystal (PBCs)



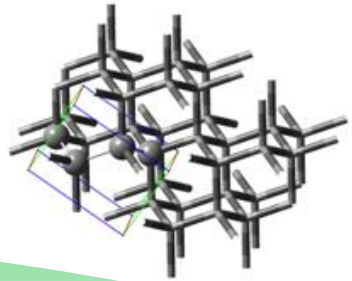
Organic reaction



Metal cluster catalysis



Crystal (PBCs)



Artificial Force Induced Reaction (AFIR)