tsscds2018: a code for automated discovery of chemical reaction mechanisms and solving the kinetics

Emilio Martínez-Núñez
Department of Physical Chemistry
Santiago de Compostela, SPAIN

FACULTADE DE QUÍMICA

Outline

-Importance of automated computational methods

-Discovery of reaction mechanisms

Other automated methods

tsscds (Transition State Search using Chemical Dynamics Simulation)

-Applications

Gas phase reactions:

Photodissociation experiments

MS measurements

Combustion chemistry

Interstellar chemistry

Organometallic catalysis

Importance of automated methods

Tedious

Dissociation of Difluoroethylenes I J. Phys. Chem. A, Vol. 107, No. 9, 2003 1391

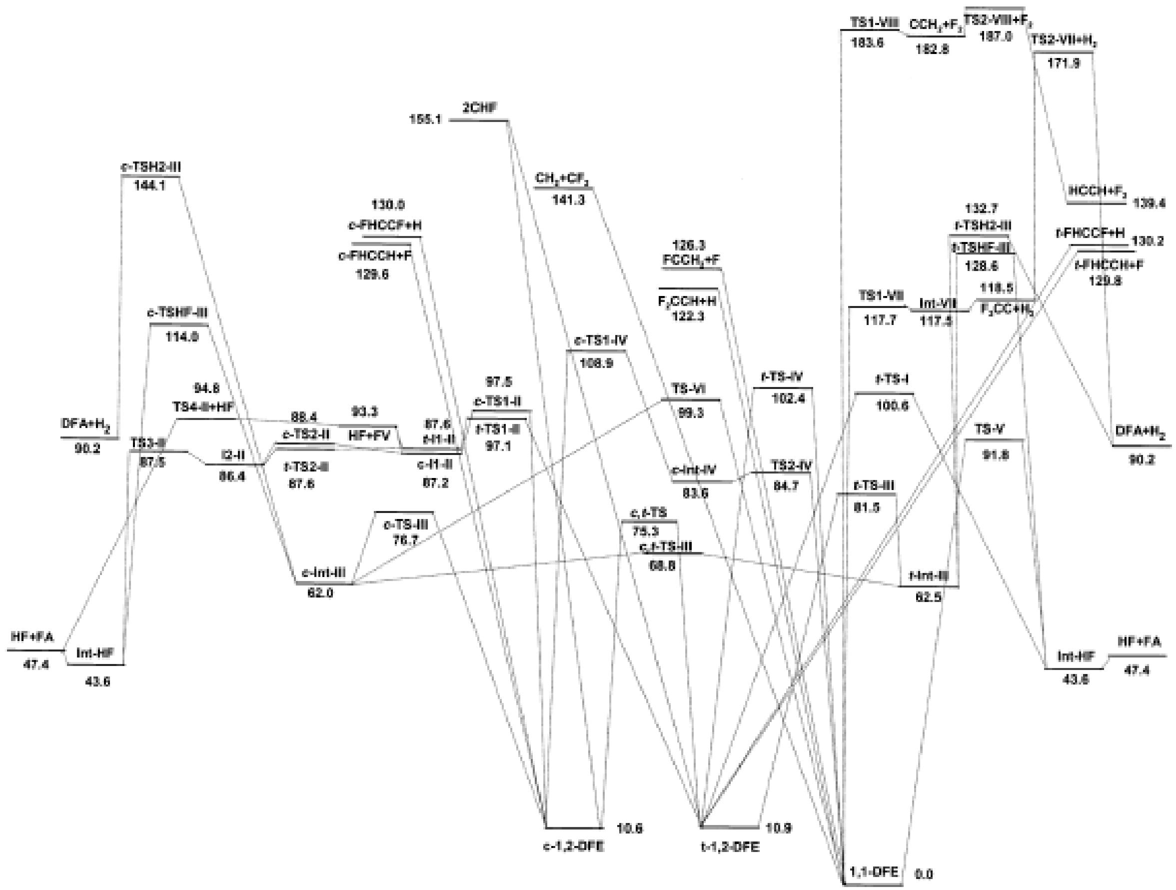
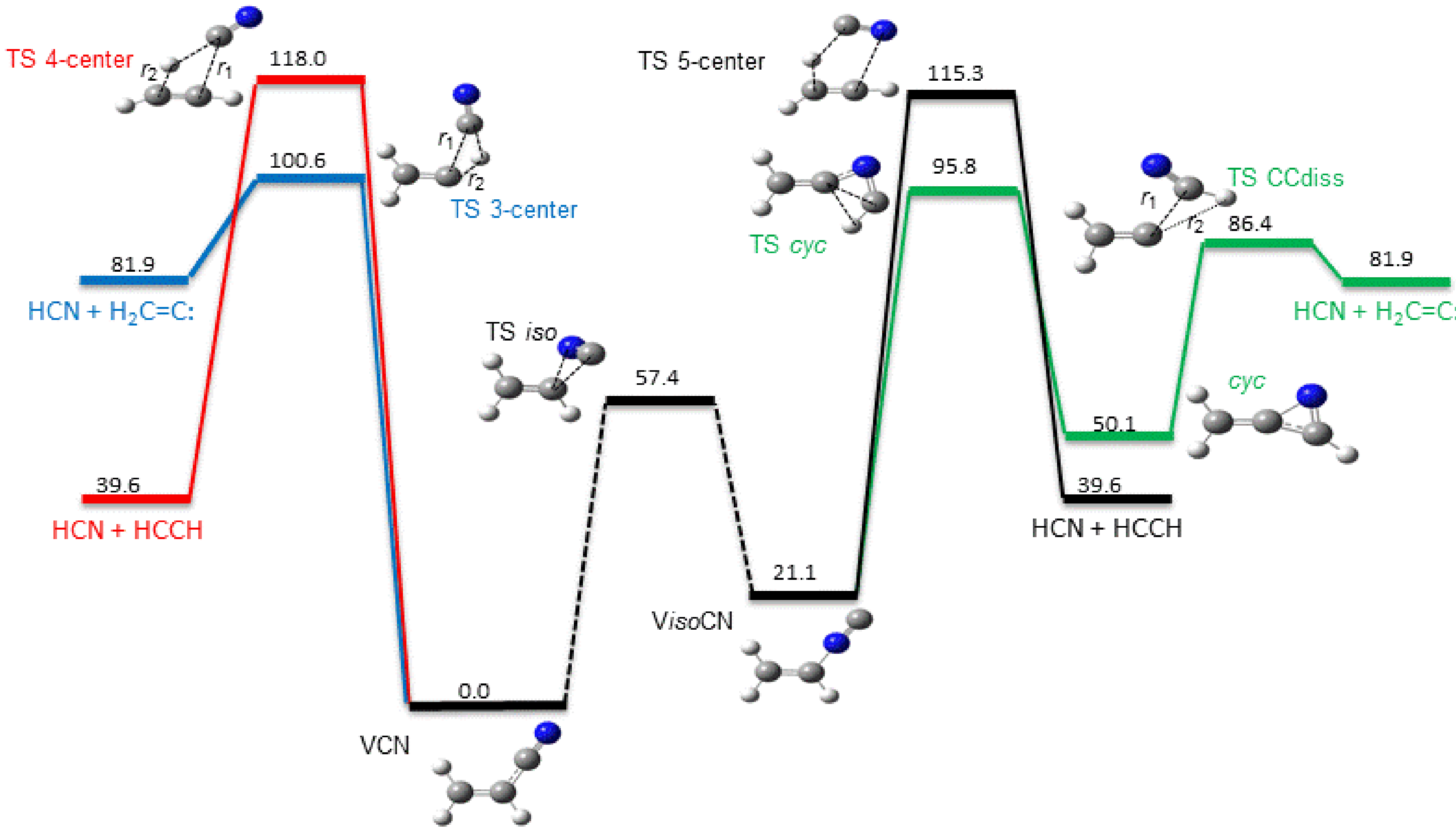


Figure 1. Schematic diagram of the ground electronic state global potential energy surface of difluoroethylenes. The relative energies were calculated at the QCISD(T)/6-311G(2d,2p)//QCISD/6-311G(2d,2p) level.

Importance of automated methods

Unintuitive mechanisms



Discovery of reaction mechanisms

Single-ended methods

(need a guess TS structure)

Berny optimization[1]

Eigenvector Following (EF)[2]

Dimer method[3]

Chain-of-state or Double-ended methods

(knowledge of reactant and products)

Nudged elastic band (NEB)[4]

String method[5]

[1]HN Schlegel *J. Comput. Chem.* **1982**, *3*, 214
[2]CJ Cerjan and WH Miller *JCP* **1981**, *75*, 2800
[3]G. Henkelman and H Jonsson *JCP* **1999**, *111*, 7010
[4]G Henkelman *et al. JCP* **2000**, *113*, 9901
[5]Winan *et al Phys. Rev. B* **2002**, *66*, 052301

Discovery of reaction mechanisms

Automated Methods[1]

Global Reaction Route Mapping (GRRM):

Anharmonic Downward distortion following (ADDF)

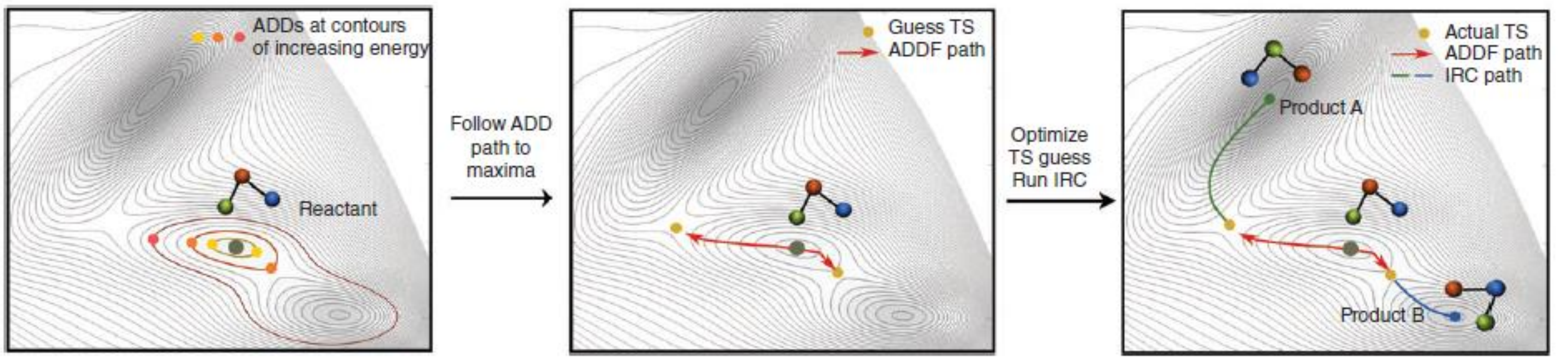


FIGURE 5 | The anharmonic downward distortion following (ADDF) method.

[1]A. L. Dewyer et al WIREs Comput. Mol. Sci, **2018,** 8: e1354, DOI: 10.1002/wcms.1354.

Discovery of reaction mechanisms

Automated Methods[1]

Artificial force induced reaction (AFIR)

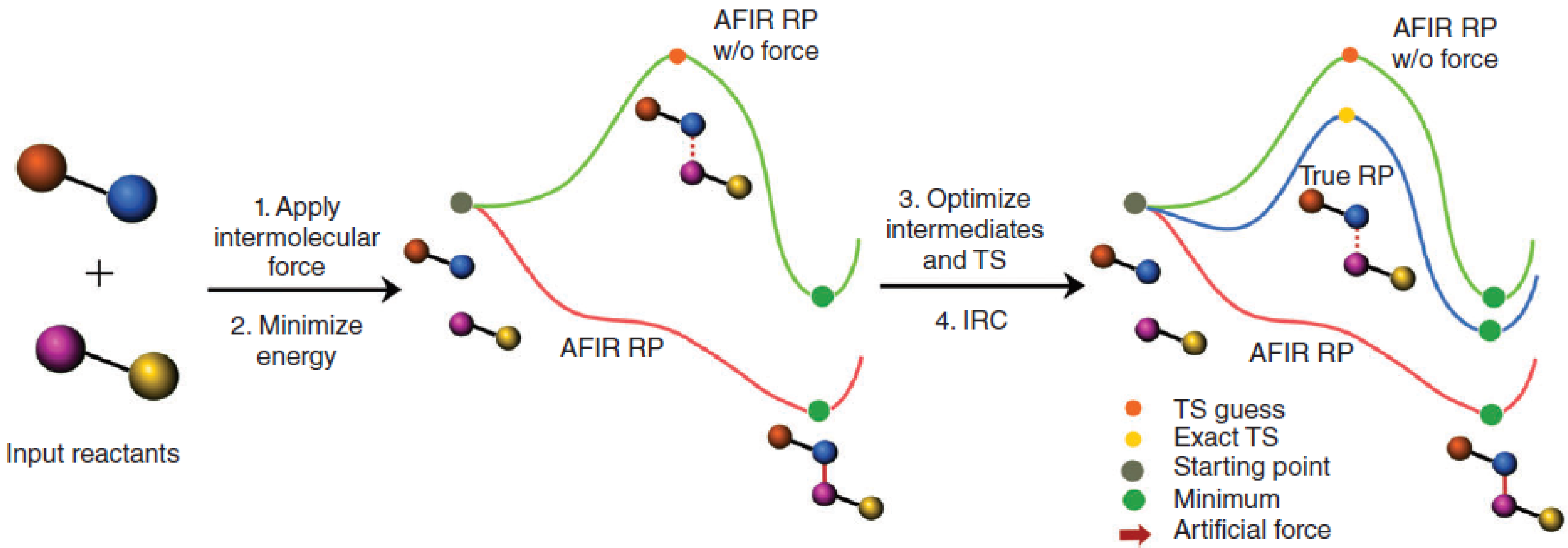


FIGURE 6 | The artificial force-induced reaction (AFIR) method.

[1]A. L. Dewyer et al WIREs Comput. Mol. Sci, **2018,** 8: e1354, DOI: 10.1002/wcms.1354.

Discovery of reaction mechanisms

Automated Methods[1]

Heuristic methods

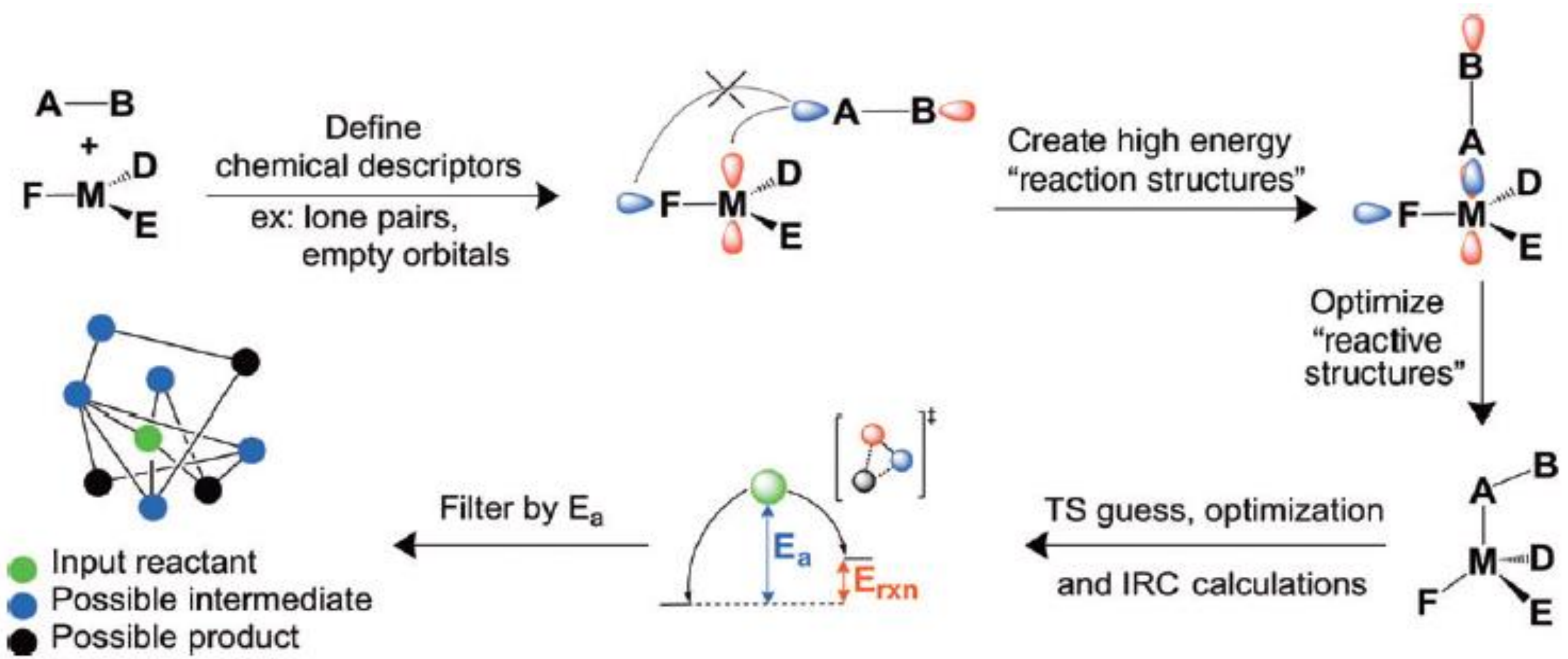


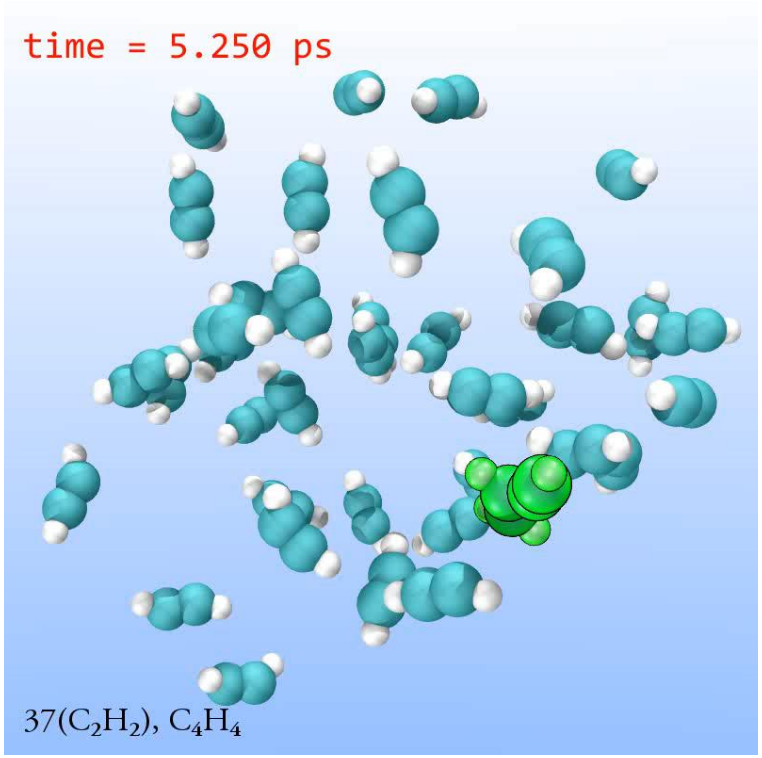
FIGURE 11 | Flow chart for the reaction path discovery method developed by Reiher and coworkers.

[1]A. L. Dewyer et al WIREs Comput. Mol. Sci, **2018,** 8: e1354, DOI: 10.1002/wcms.1354.

Discovery of reaction mechanisms

Automated Methods[1]

Ab initio nanoreactor



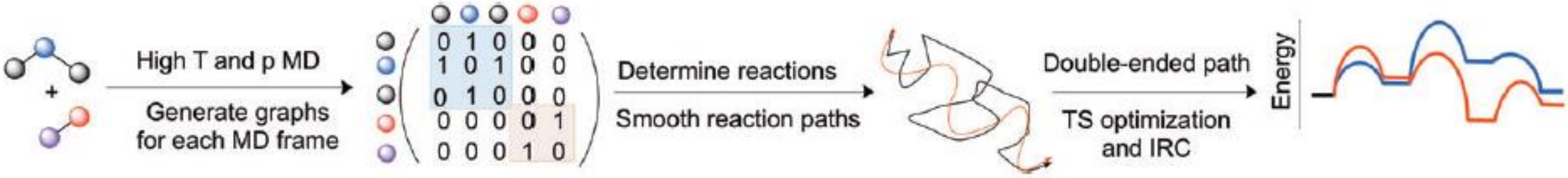


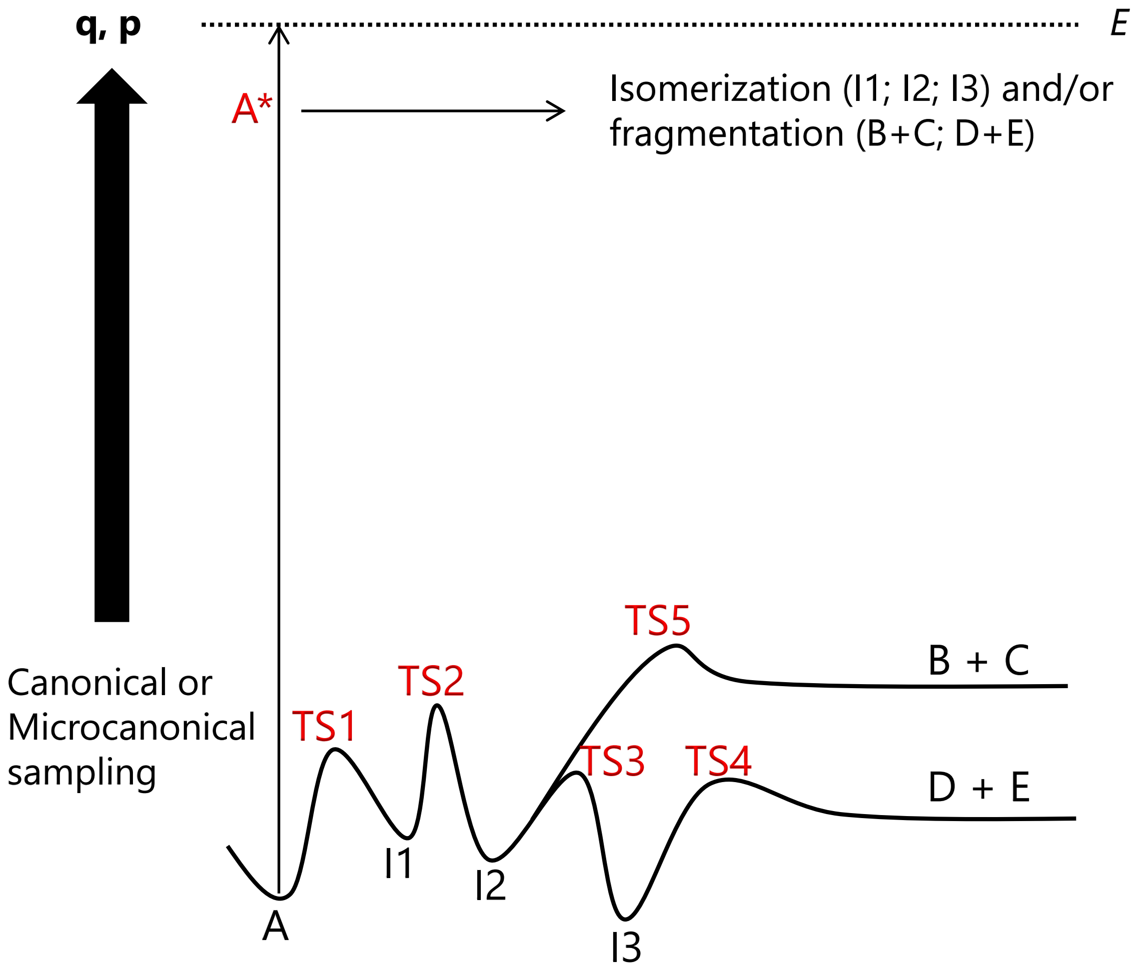
FIGURE 12 | Reaction pathway discovery using the nanoreactor.

[1]A. L. Dewyer et al WIREs Comput. Mol. Sci, **2018,** 8: e1354, DOI: 10.1002/wcms.1354.

Discovery of reaction mechanisms

Transition State Search using Chemical Dynamics Simulations (TSSCDS)[1]

Accelerated Dynamics

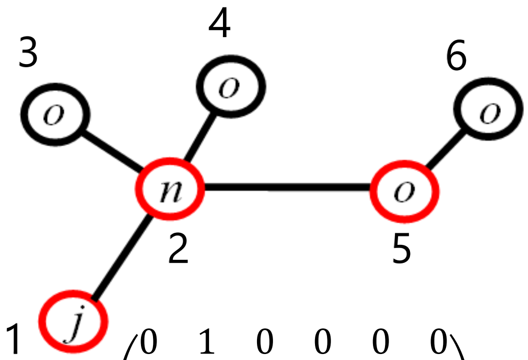


[1]EMN *J Comput Chem* **2015**, *36*, 222

Discovery of reaction mechanisms

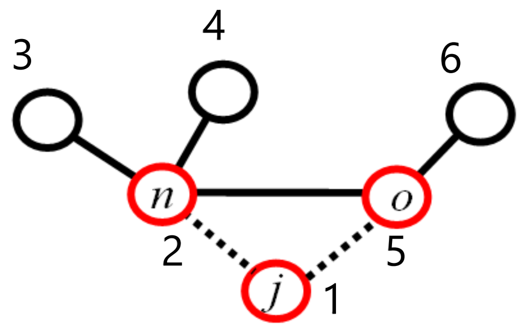
Transition State Search using Chemical Dynamics Simulations (TSSCDS)[1]

Bond Breakage/Formation Search (BBFS)



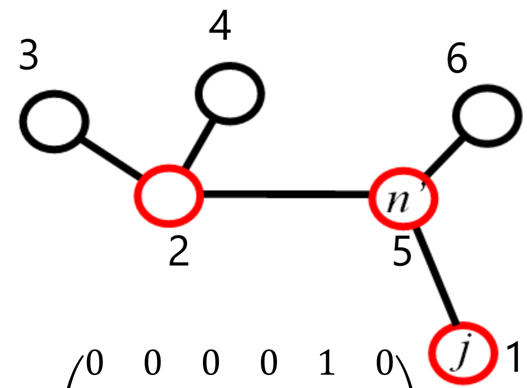
$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

R



$$\begin{pmatrix} 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0.5 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

0.5*(R+P)



$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

P

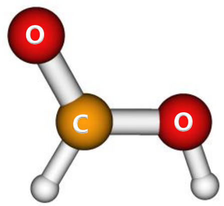
Two levels of theory:

LL (semiempirical calculations)

HL (DFT or *ab initio*)

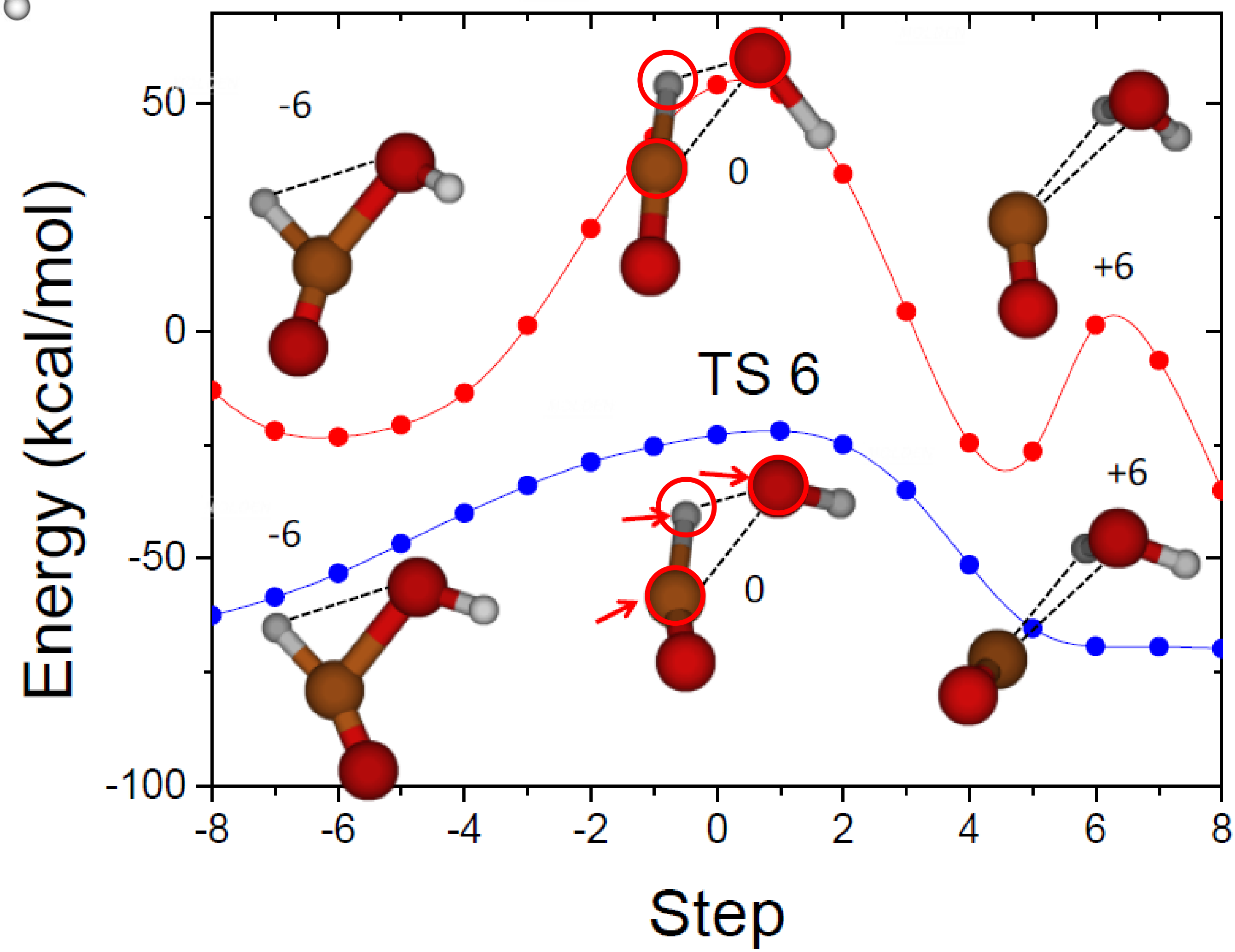
[1]EMN *J Comput Chem* **2015**, *36*, 222

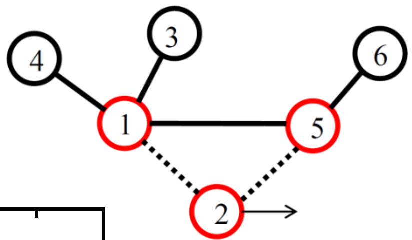
Discovery of reaction mechanisms



Formic acid

$HCOOH \rightarrow CO + H_2O$

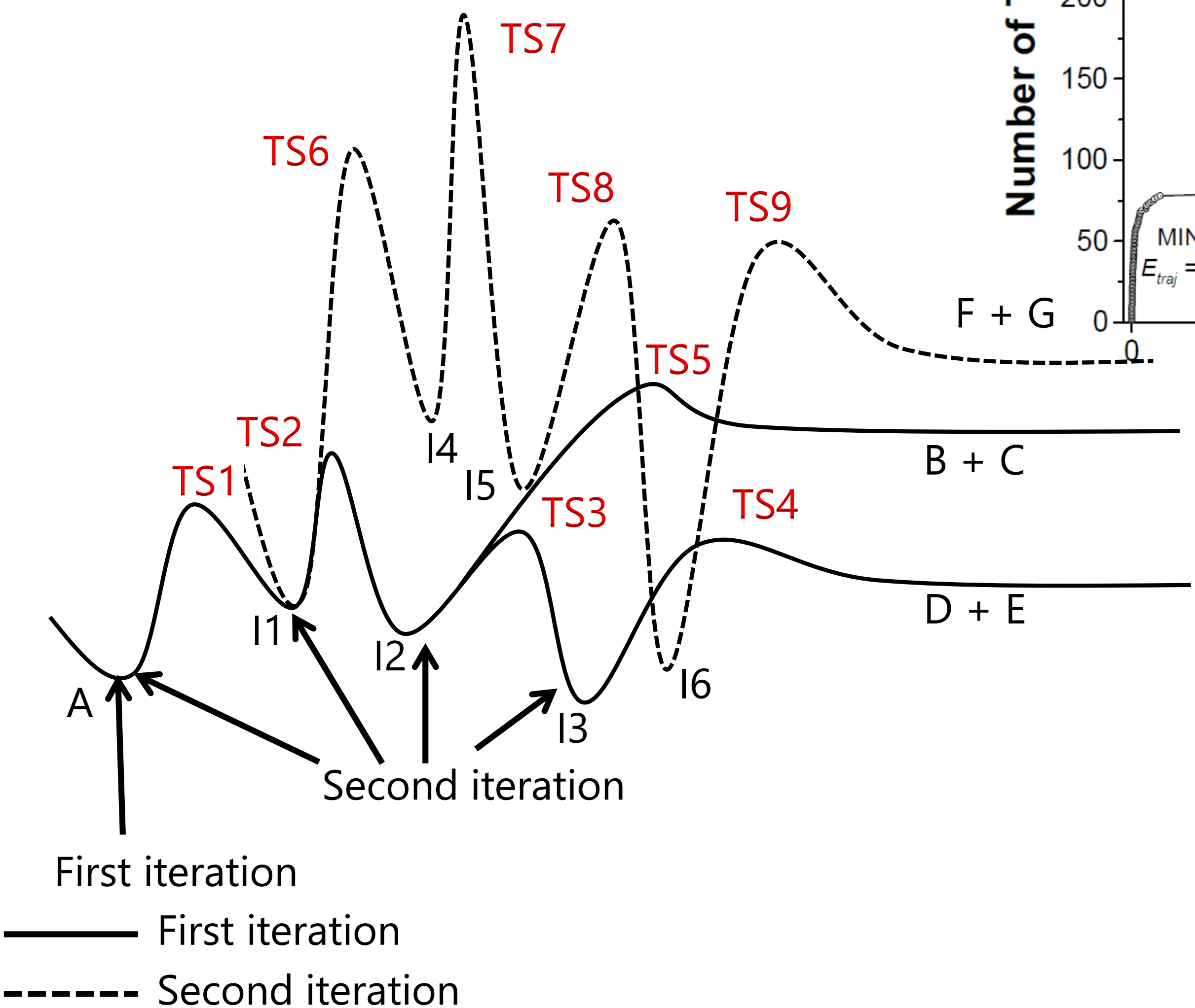


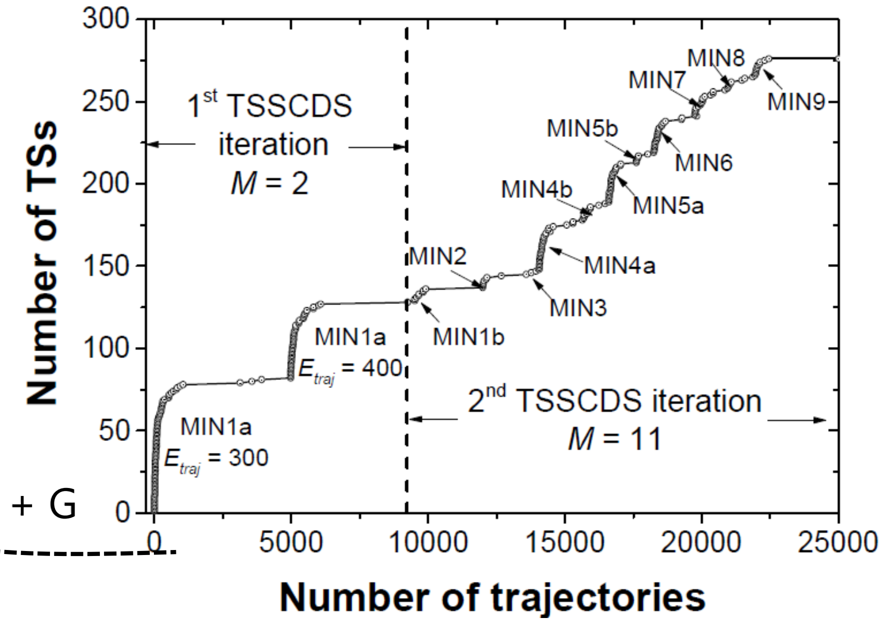


[1]EMN *J Comput Chem* **2015**, *36*, 222

Discovery of reaction mechanisms

Iterative TSSCDS[1]

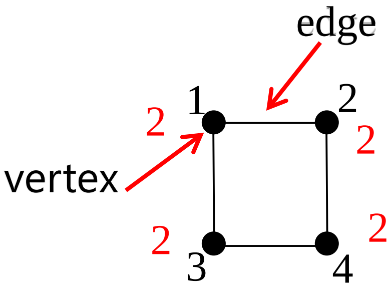




[1]EMN *PCCP* **2015**, *17*, 14912

Discovery of reaction mechanisms

Post-processing analysis tools



Spectral Graph Theory

Degree of each vertex :
Number of bonds

Adjacency Matrix

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$a_{ij} = \frac{1-(\delta_{ij})^n}{1-(\delta_{ij})^m}$$

Laplacian Matrix *L* = *D* – *A*

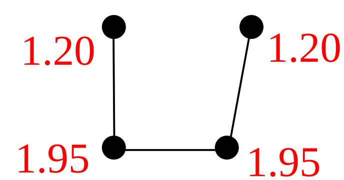
$$\begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

$$\lambda = 0,2,2,4$$

One fragment

SPRINT Coordinates[1]

$$S_i = \sqrt{N}\lambda^{max} v^{max}$$



[1]F Pietrucci and W Andreoni *PRL* **2011**, *107*, 085504

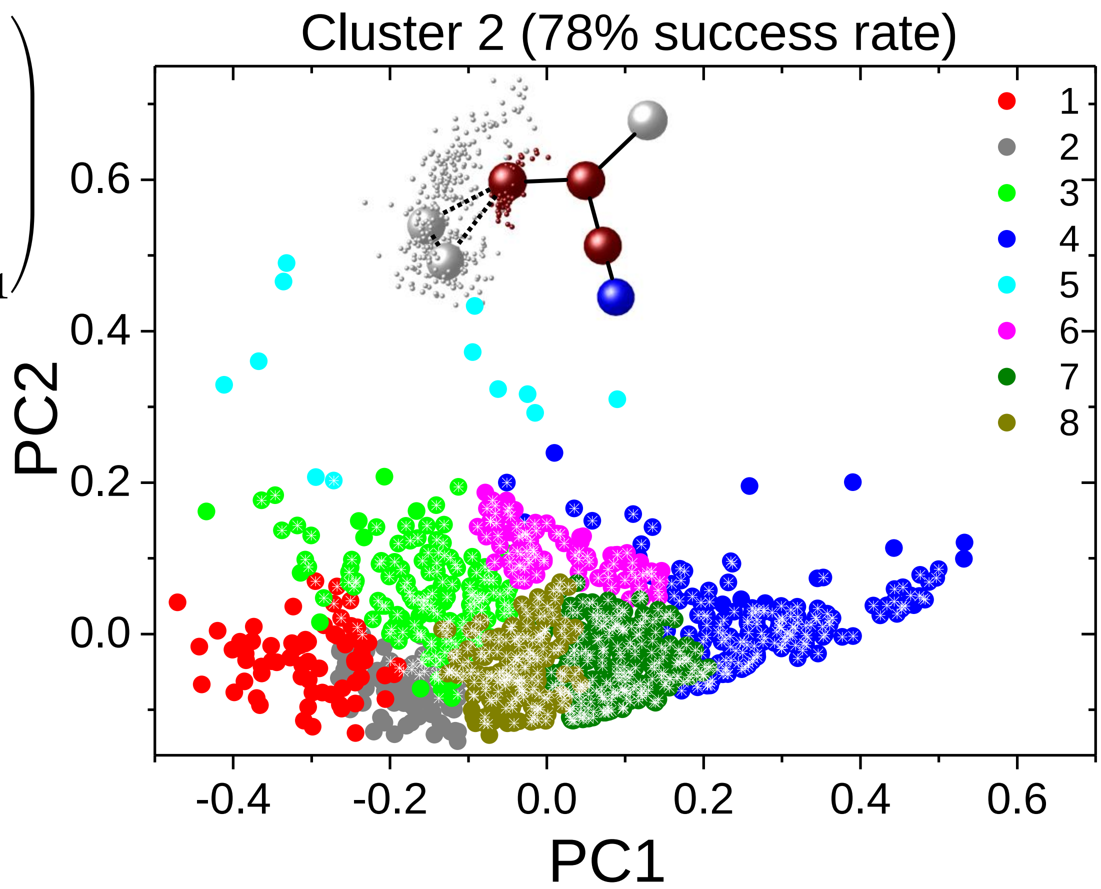
Discovery of reaction mechanisms

Future improvements[1]

0.5*(R+P)+Z

$$\begin{pmatrix} 1.7 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 1.6 & 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 1.6 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1.6 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1.1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1.1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1.1 \end{pmatrix}$$

$$a_{ij} = \frac{1-(\delta_{ij})^n}{1-(\delta_{ij})^m}$$

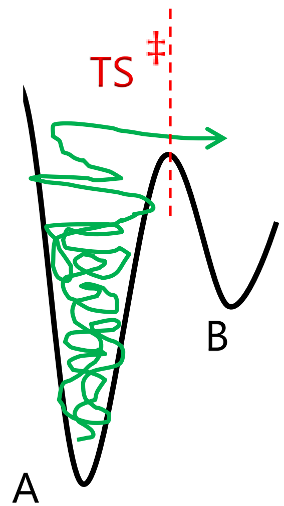


[1]SAV & EMN (work in progress)

Rates of Chemical Reactions

Transition State Theory (TST)

$$k(T) = \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}}$$



Rice-Ramsperger-Kassel-Marcus (RRKM)

$$k(E) = \frac{W^{\ddagger}(E)}{h\rho(E)}$$

TST	RRKM
Boltzmann distribution	Fast IVR
The TS is not re-crossed	
The reaction follows the IRC	
The nuclei behave according to classical mechanics	

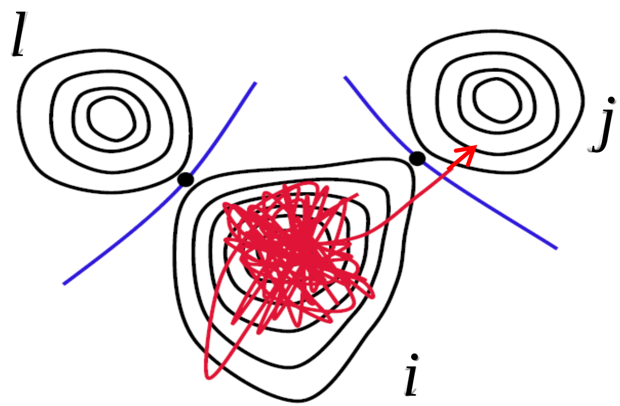
Kinetics: solving the ME

$$\frac{d\mathbf{n}(t)}{dt} = \mathbf{M}\mathbf{n}(t)$$

Kinetic Monte Carlo (KMC)

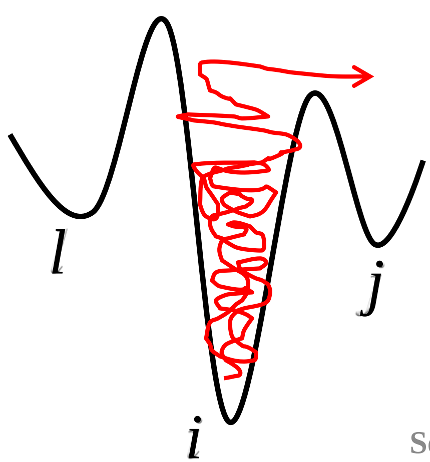
Working out the kinetics knowing rate coefficients (TST or RRKM)

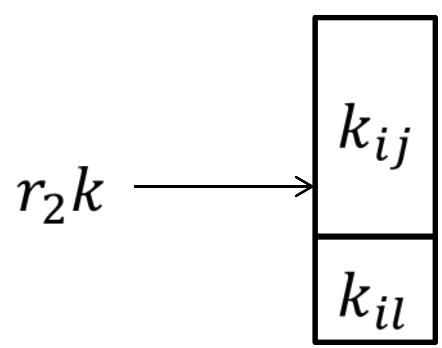
and using random numbers (r_1 *and* r_2)



$$t = -\left(\frac{1}{k}\right) ln(r_1)$$

$$k = k_{ij} + k_{il}$$



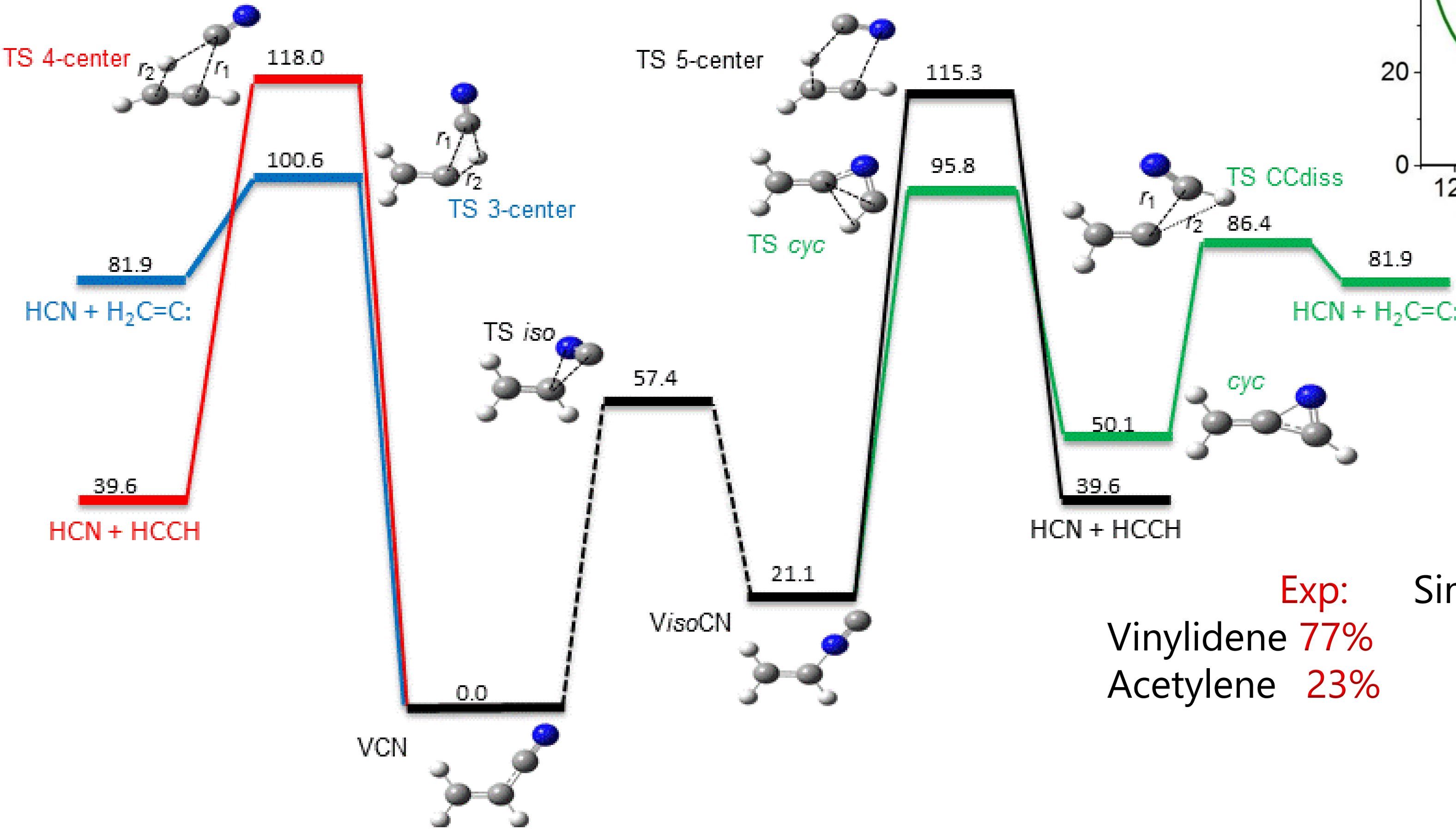


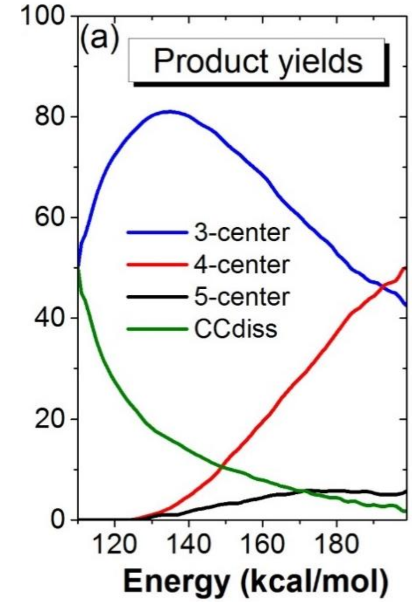
Program wiki

http://forge.cesga.es/wiki/g/tsscds/HomePage

Applications

Vinyl cyanide: HCN elimination channels[1]

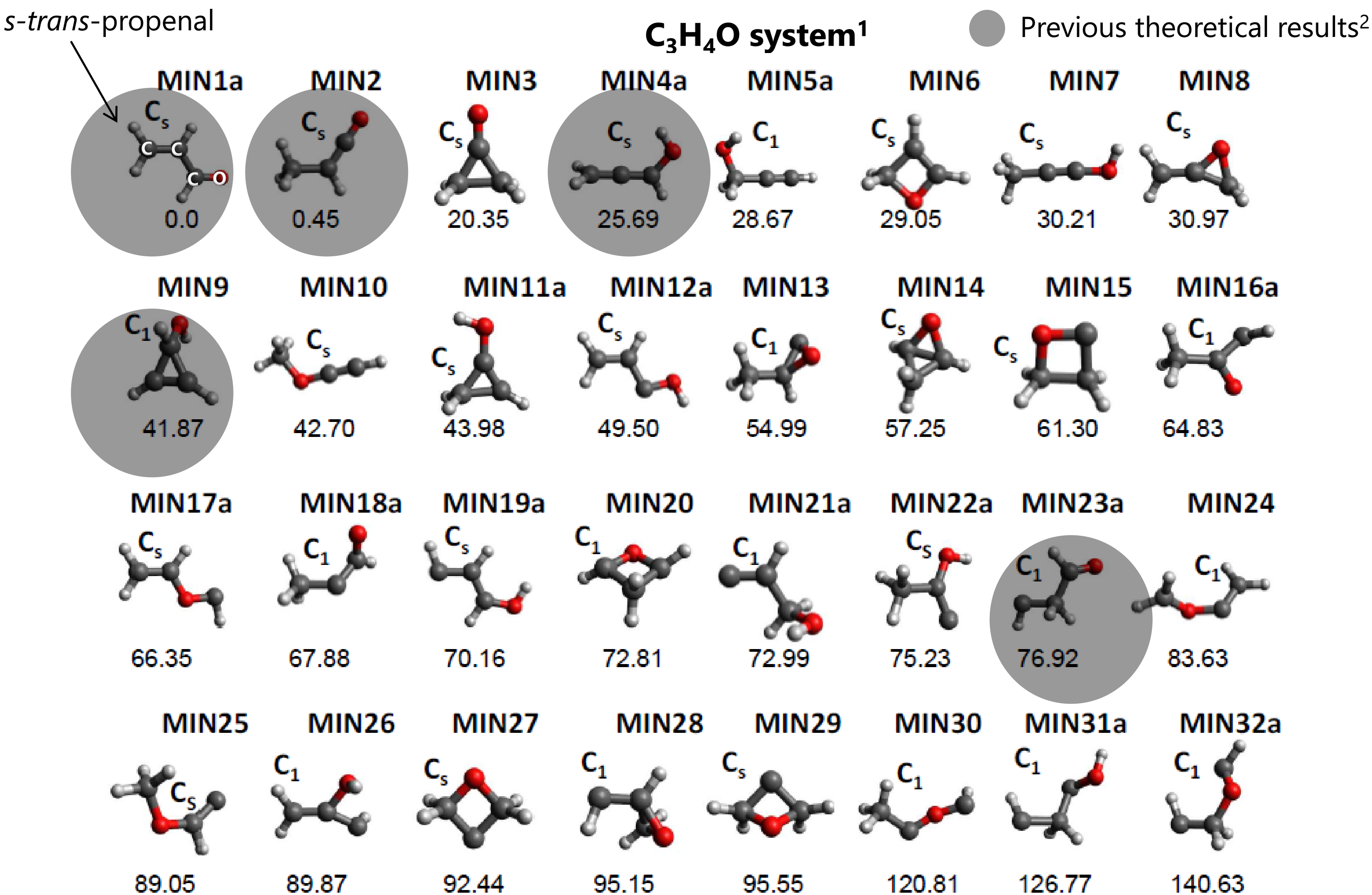




	Exp:	Simulations
Vinylidene	77%	79%
Acetylene	23%	21%

[1]SA Vazquez and EMN *PCCP* **2015**, *17*, 6948
Z Homayoon *JPCA* **2011**,*115*, 979

Applications



[1]EMN *PCCP* **2015**, *17*, 14912
[2]CH Chin and SH Lee *JCP* **2011**, *134*, 044309

Applications

C_3H_4O system

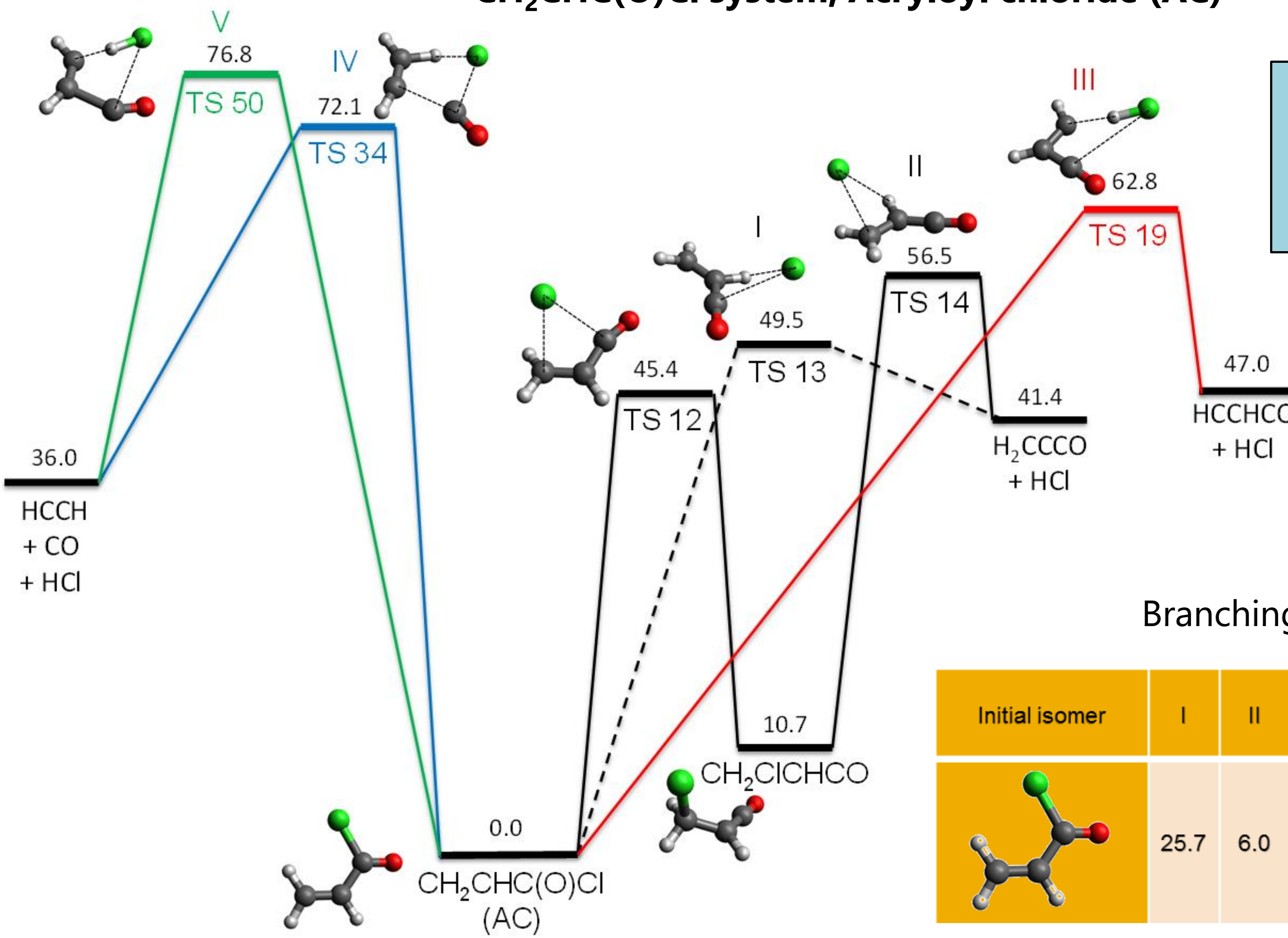
Branching ratios

Channel	Previous Th. results[1]	*i*TSSCDS[2]	Exp[3]	
H_2O	0.01	0.03	0.07	
CH_2O	0.65	0.20	0.07	
H_2	0.09	0.19	0.00	
CO	1.00	1.00	1.00	
Triple	6.82	1.49	1.10	$CO + H_2 + HCCH$

[1]CH Chin and SH Lee *JCP* **2011**, *134*, 044309
[2]EMN *PCCP* **2015**, *17*, 14912
[3]C Chaudhuri and SH Lee, *PCCP* **2011**, *13*, 7312

Applications

$CH_2CHC(O)Cl$ system, Acryloyl chloride (AC)[1,2]



174 Minima

527 TSs

20 Diss. channels

Branching ratios

Initial isomer	I	II	III	IV	V	CO
	25.7	6.0	27.5	13.0	20.3	7.5

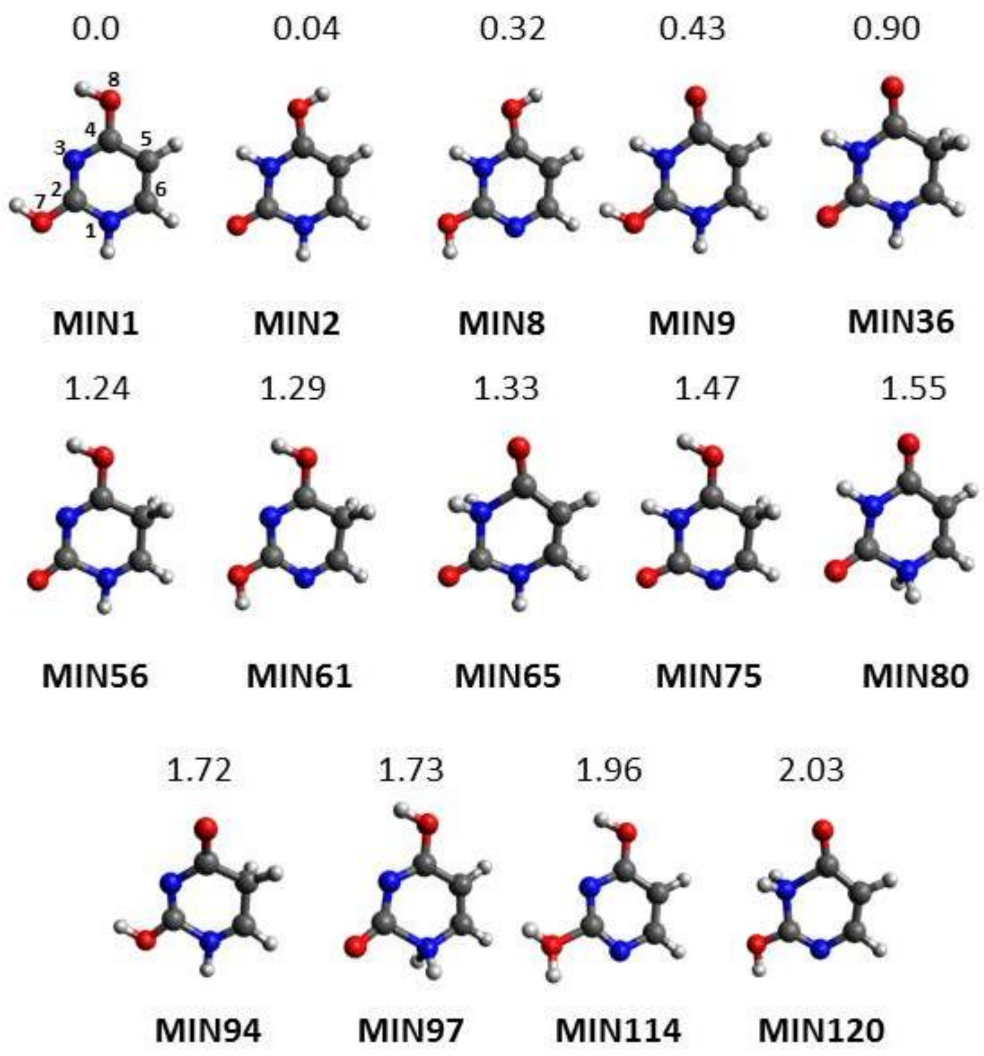
[1]R Pérez-Soto et al. *PCCP* **2016**, *18*, 5019
[2]P.-W. Lee et al. *JPC A* **2015**, *119*, 7293.

Applications

Protonated uracil[1]

1398 stationary points

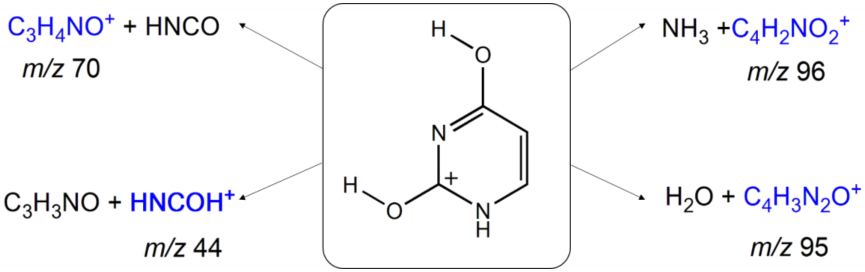
751 TSs

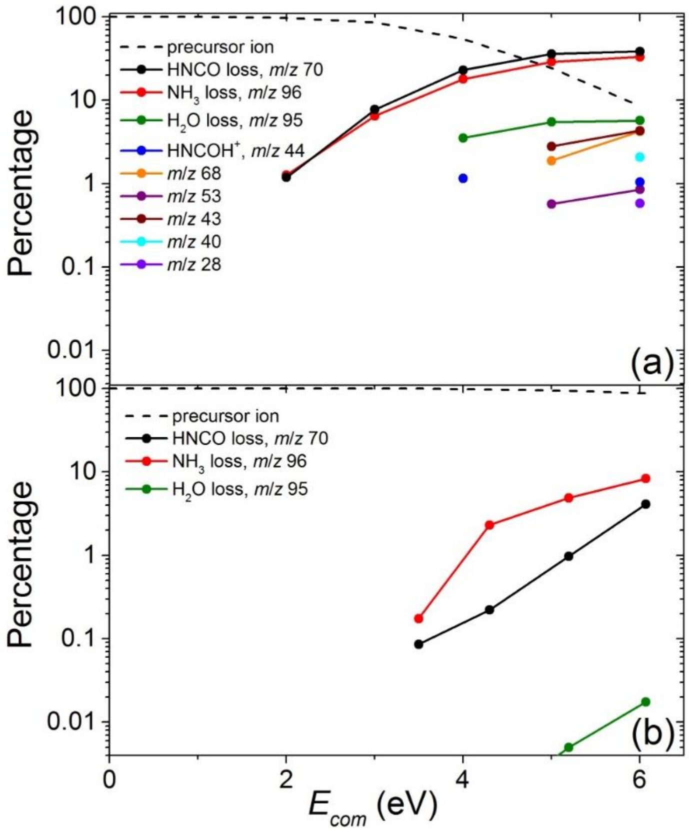


[1]E Rossich-Molina *et al*. *PCCP* **2016**, *18*, 14980

Applications

Major dissociation channels (MS experiments)

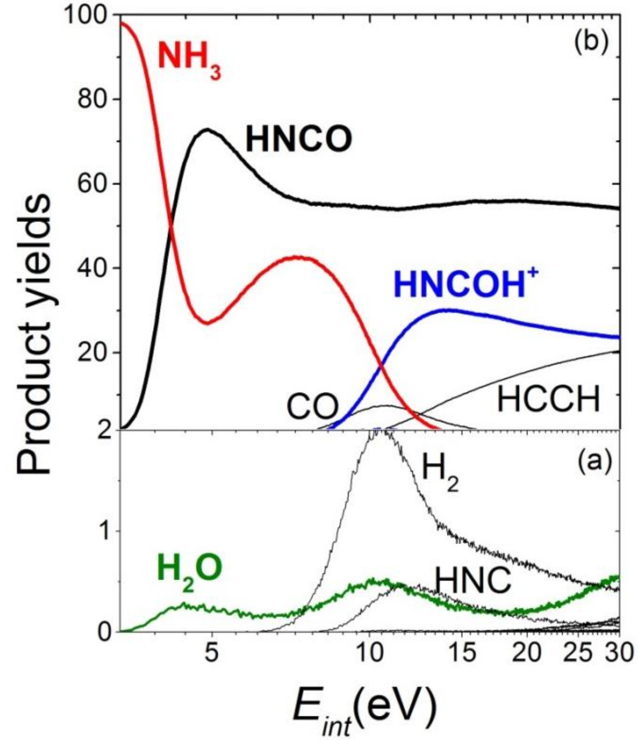




MS Experiments

Statistical simulations

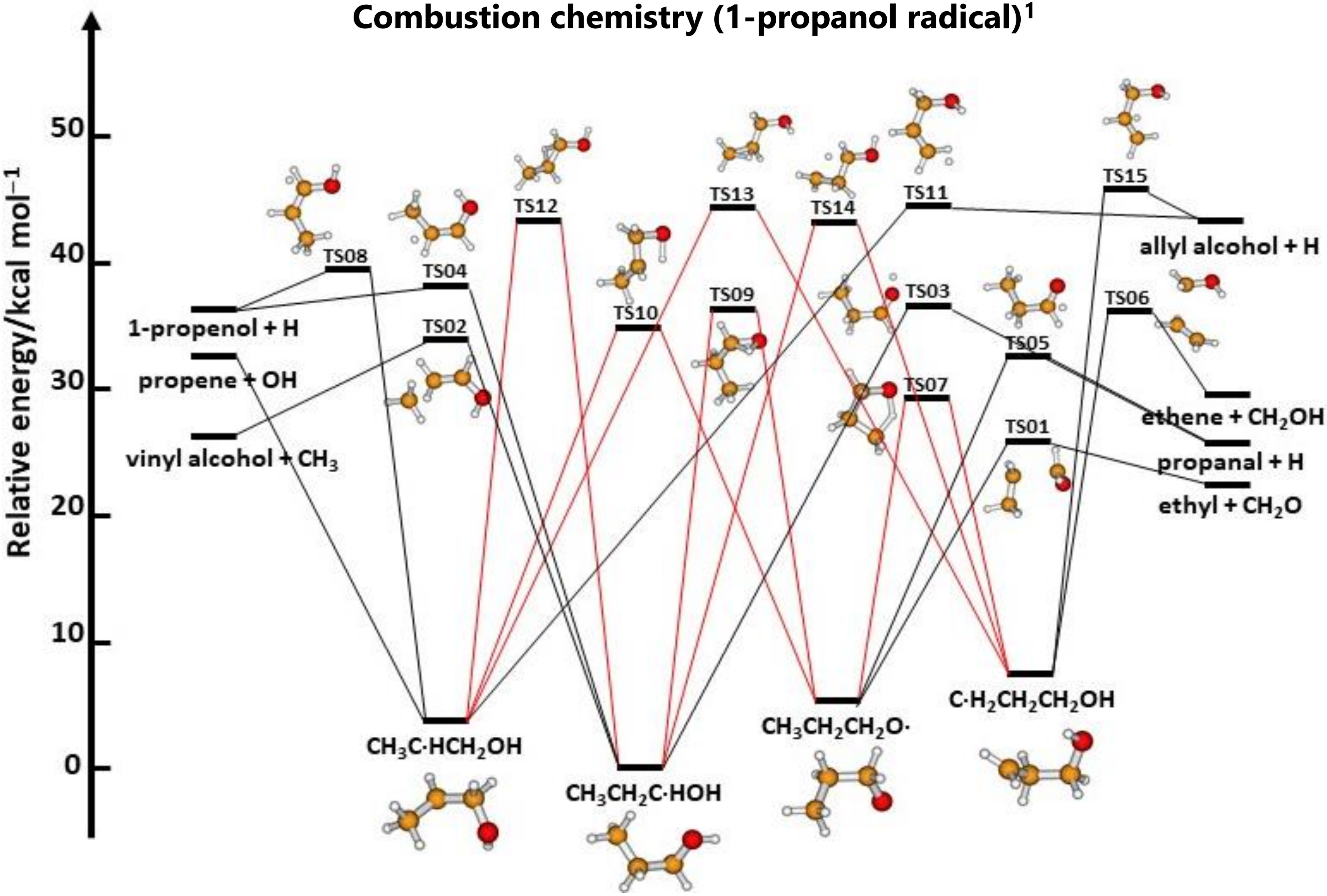
Branching ratios (statistical results)



[1]E Rossich-Molina *et al.* *PCCP* **2016**, *18*, 14980

Applications

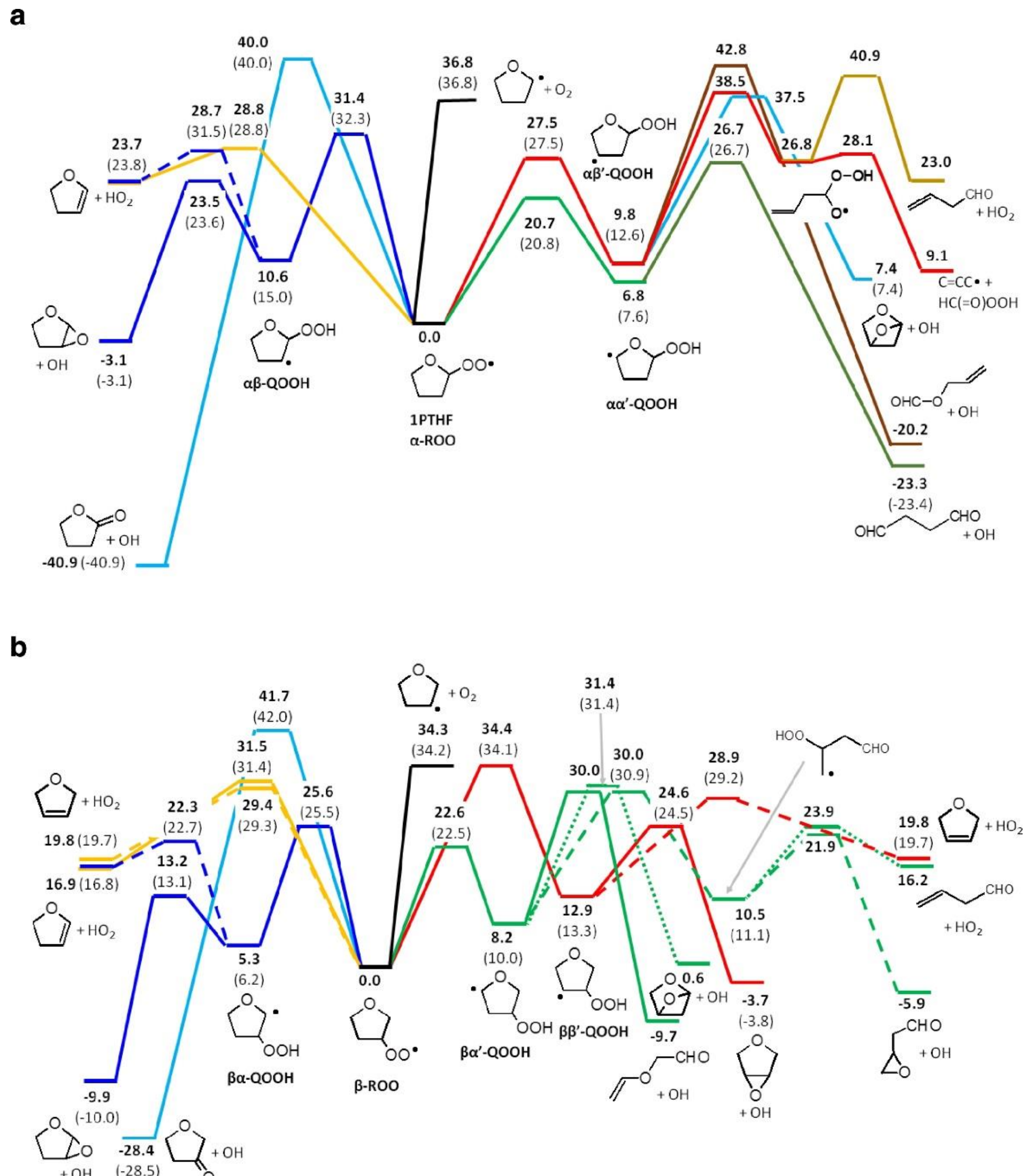
Combustion chemistry (1-propanol radical)[1]



[1]D. Ferro-Costas et al. JPCA (submitted)

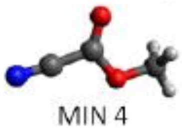
Applications

Combustion chemistry (Tetrahydrofuran oxidation)[1]

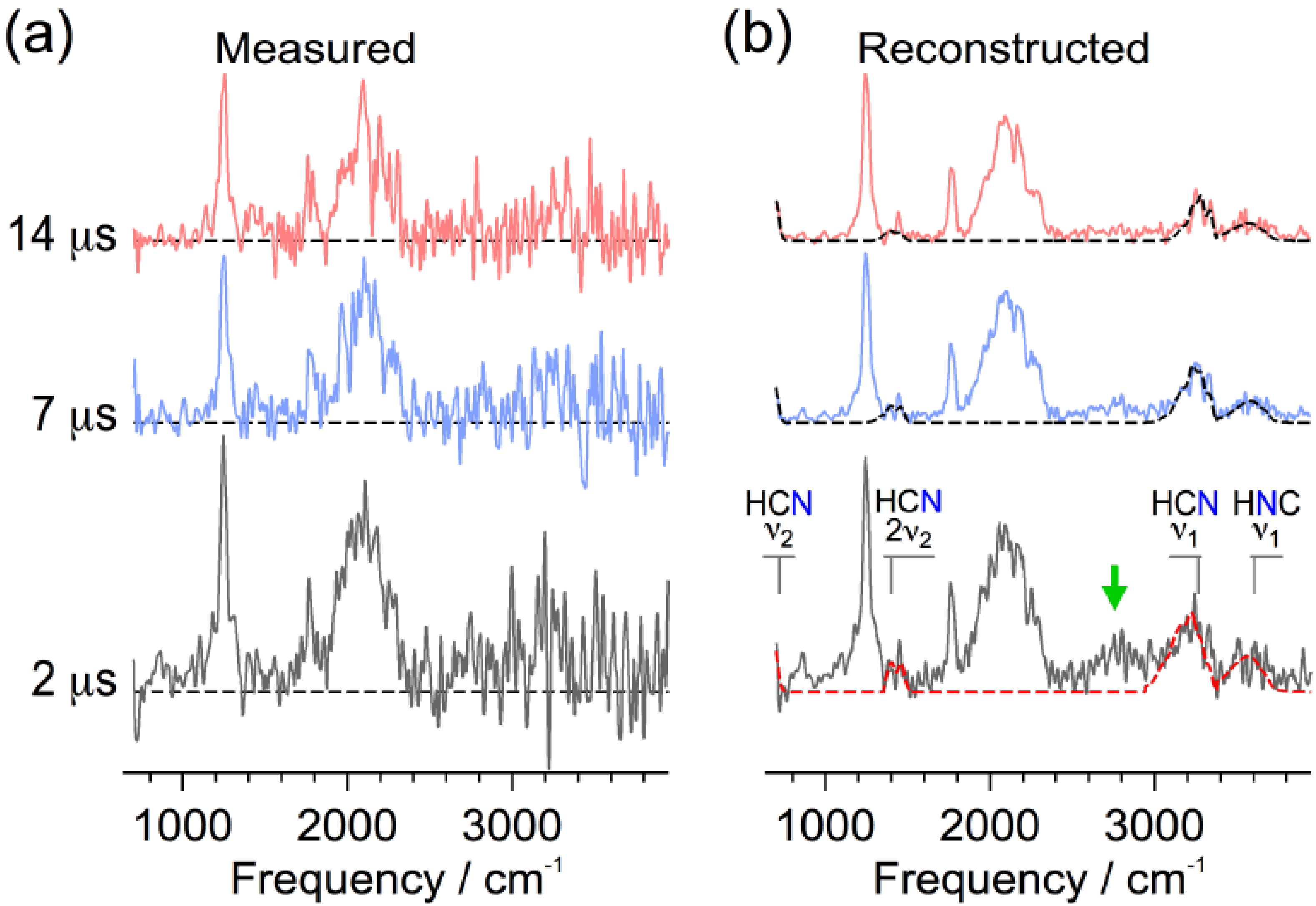


[1]Y. Fenard et al. Combust. Flame **2018**, *191*, 252

Applications



Interstellar chemistry[1]

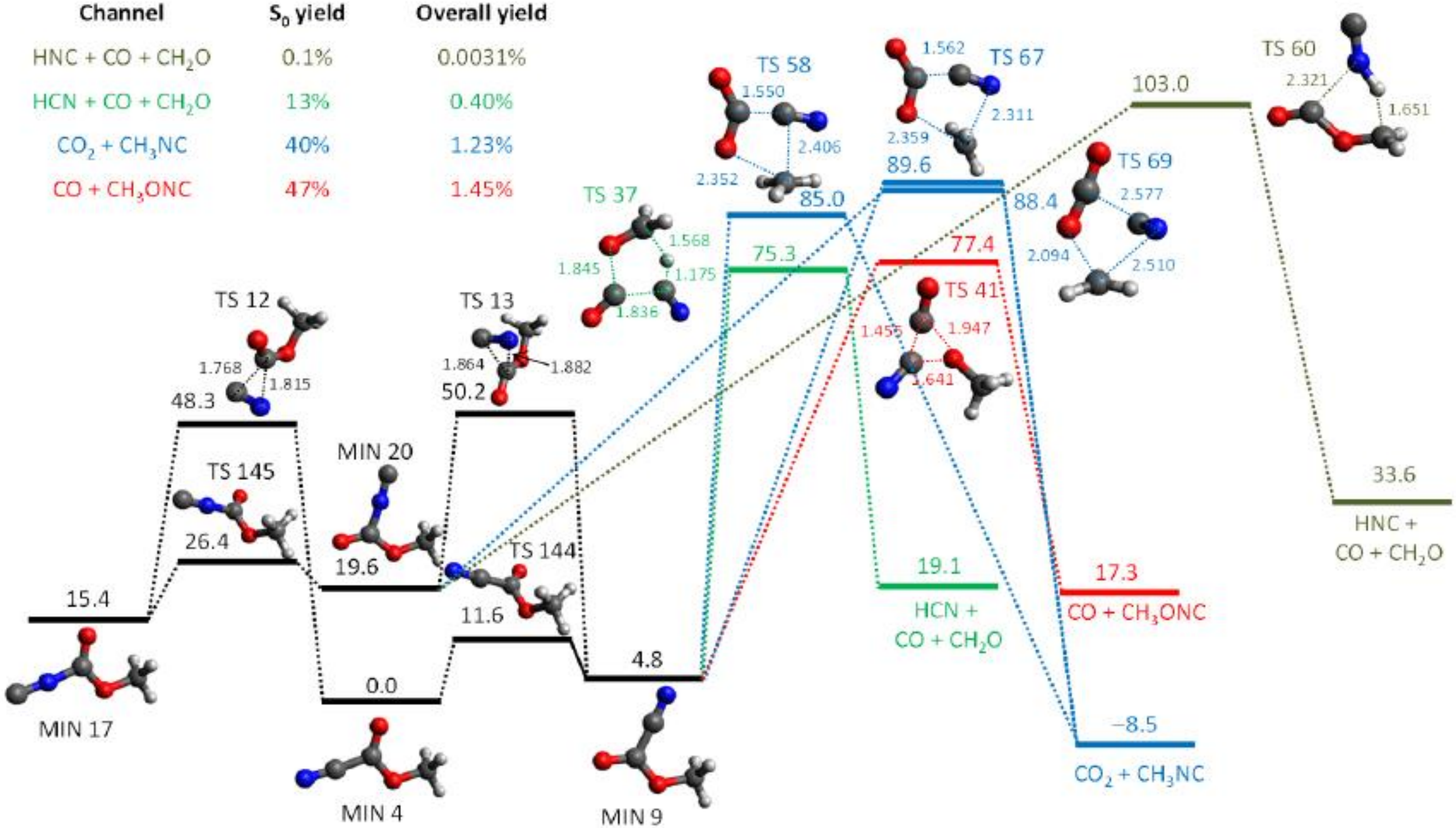


[1]M.J. Smith et al . *ApJ* **2017**, *849*, 15

Applications

Interstellar chemistry[1]

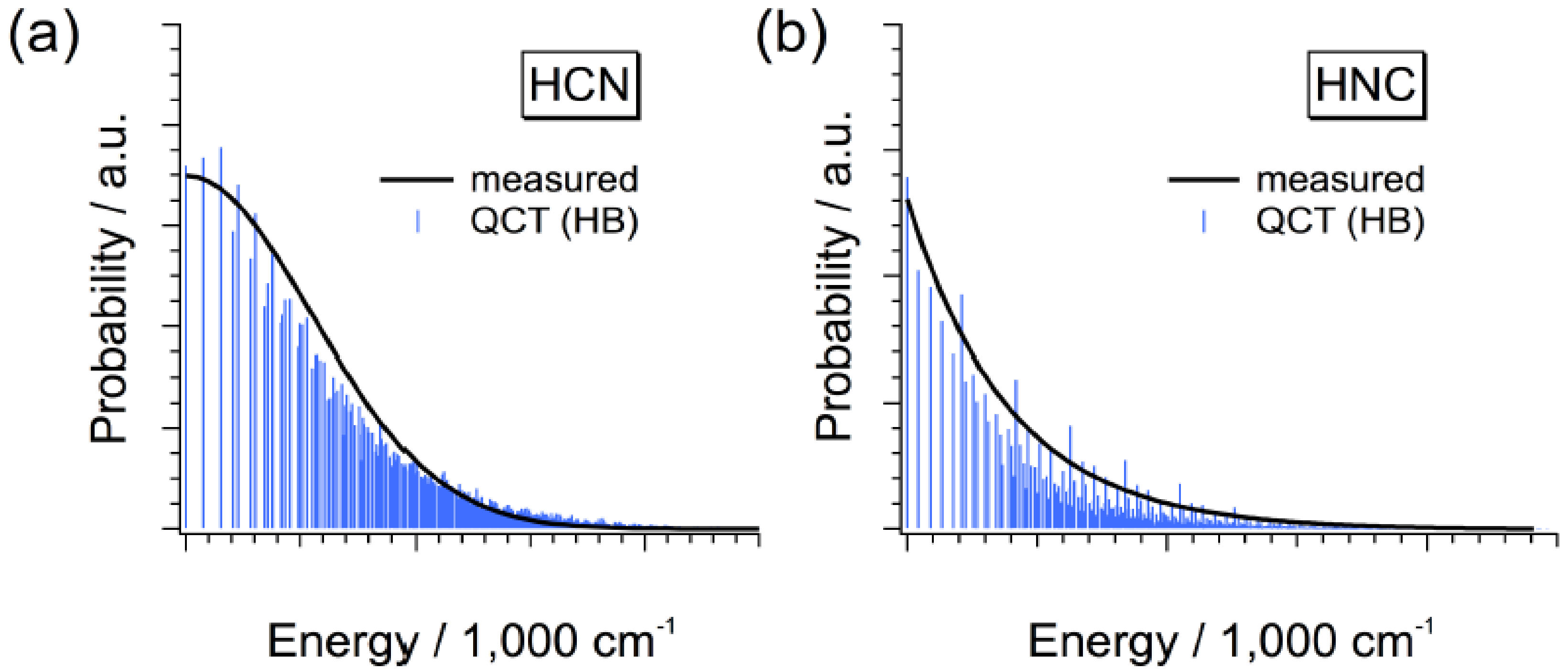
Channel	S_0 yield	Overall yield
HNC + CO + CH_2O	0.1%	0.0031%
HCN + CO + CH_2O	13%	0.40%
CO_2 + CH_3NC	40%	1.23%
CO + CH_3ONC	47%	1.45%



[1]M.J. Smith et al . *ApJ* **2017**, *849*, 15

Applications

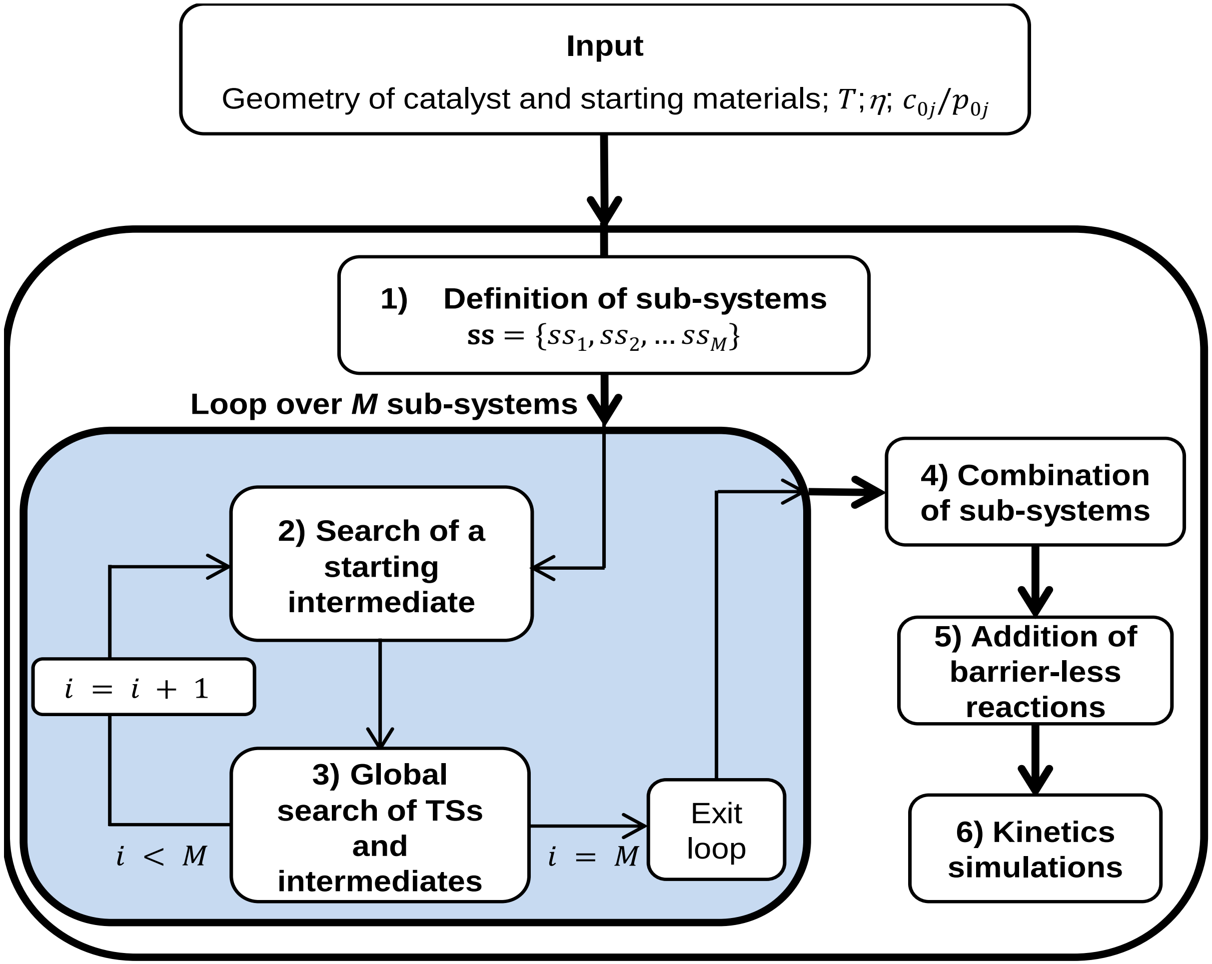
Interstellar chemistry[1]



[1]M.J. Smith et al . *ApJ* **2017**, *849*, 15

Applications

Organometallic catalysis[1]



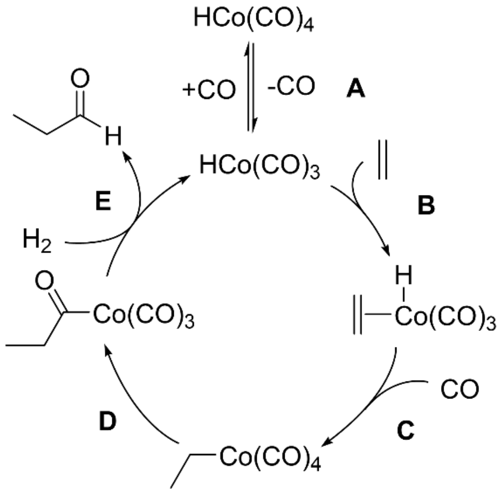
T = temperature of the system; η = viscosity of the solvent
c_{0j}/p_{0j} = initial concentration/pressure of species j
ss_i = Catalyst or any combination of the catalyst with the starting materials

[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Test case[1]

Co-catalyzed hydroformylation of ethylene



Heck and Breslow mechanism

Input
Catalyst and starting materials: $HCo(CO)_3$ (**1**), ethylene (2), CO (**3**) and H_2 (**4**)
Temperature: 423 K
Solvent: Toluene ($\eta = 2.09 \times 10^{-4}$ Pa s)
$c_{0,cat}$=0.0004-0.02 M
$c_{0,eth}$=0.04-2 M
p_0=1-60 bar CO and H_2

[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Method[1]

1) *Definition of sub-systems.*

$\mathrm{ss} = \{ss_1, ss_2, \ldots, ss_M\}$; $M = 1 + \sum_{n=1}^{N} \frac{N!}{(N-n)!n!}$; N starting materials

Definition of the sub-systems studied in this work, and species A and B utilized in step 2 of the method.

sub-system		A[a]	B[a]
ss_1	$HCo(CO)_3$	-	-
ss_2	$HCo(CO)_3/C_2H_4$	$HCo(CO)_3$	C_2H_4
ss_3	$HCo(CO)_3/CO$	$HCo(CO)_3$	CO
ss_4	$HCo(CO)_3/H_2$	$HCo(CO)_3$	H_2
ss_5	$HCo(CO)_3/C_2H_4/CO$	$HCo(CO)_3/C_2H_4$	CO
ss_6	$HCo(CO)_3/C_2H_4/H_2$	$HCo(CO)_3/C_2H_4$	H_2
ss_7	$HCo(CO)_3/CO/H_2$	$HCo(CO)_3/CO$	H_2
ss_8	$HCo(CO)_3/C_2H_4/CO/H_2$	$HCo(CO)_3/C_2H_4/CO$	H_2

2) *Starting intermediate.*

-Generate 10^2 guess A····B structures ramdonly rotating each unit about pivot points.

-Optimize with MOPAC (PM7).

-Choose the best starting intermediate (energy and valence of the metal center)

[1]JA Varela *et al.* *Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Method[1]

3) *Global search for TSs and intermediates.*
-*i*TSSCDS algorithm
-Locally modified versión MOPAC (larger step size, termination criteria). T = 10^4 K, Δt = 0.5 fs, PM7.
-High level: B3LYP/6-31G(d,p). TSs with ΔG>40 kcal/mol are not considered.

4) *Combination of sub-systems*.
Obtain single lists: **TS**, **I**, **P**
-Setting a common G scale.
-Identifying common molecules in different sub-systems. Tools of Graph Theory.

5) *Addition of barrierless reactions*
-Analysis of **I** to check if an element of **P** (P_j) is within **I**
-The distance r between metal and P_j is doubled. Partial optimization and follow the gradient.

6) *Kinetics simulations*.

Over a barrier

$$k_l(T) = \sigma \frac{k_B T}{h} \left(\frac{RT}{p_0}\right)^{\Delta n} e^{-\frac{\Delta G_l^{\ddagger}}{RT}}$$

Barrierless association

$$k_{diff}(T) = \frac{8 k_B T}{3\eta}$$

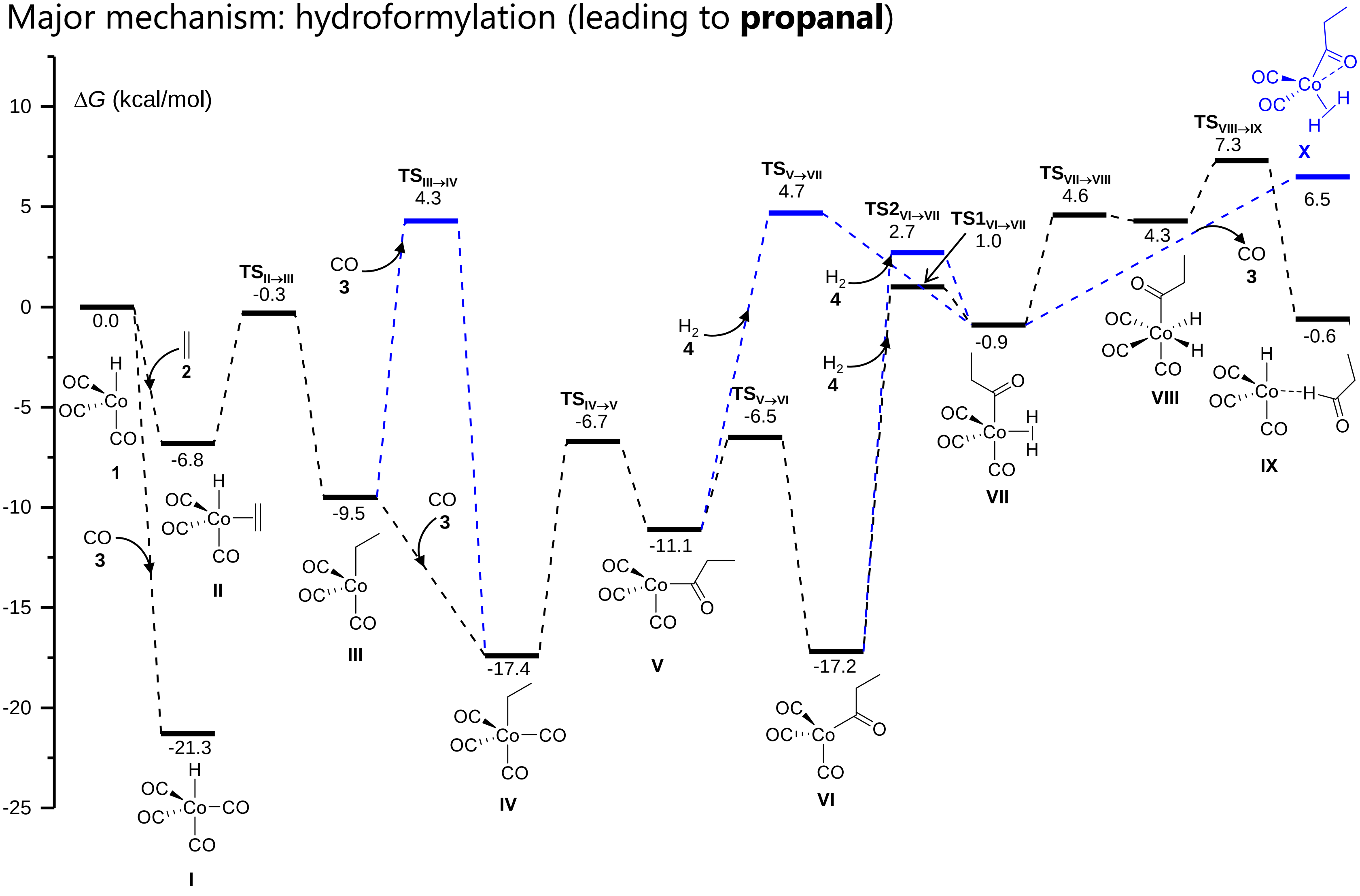
KMC simulations: StochKit2.0[2]

[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843
[2]https://github.com/StochSS/StochKit

Applications

Organometallic catalysis: Mechanisms[1]

Major mechanism: hydroformylation (leading to **propanal**)

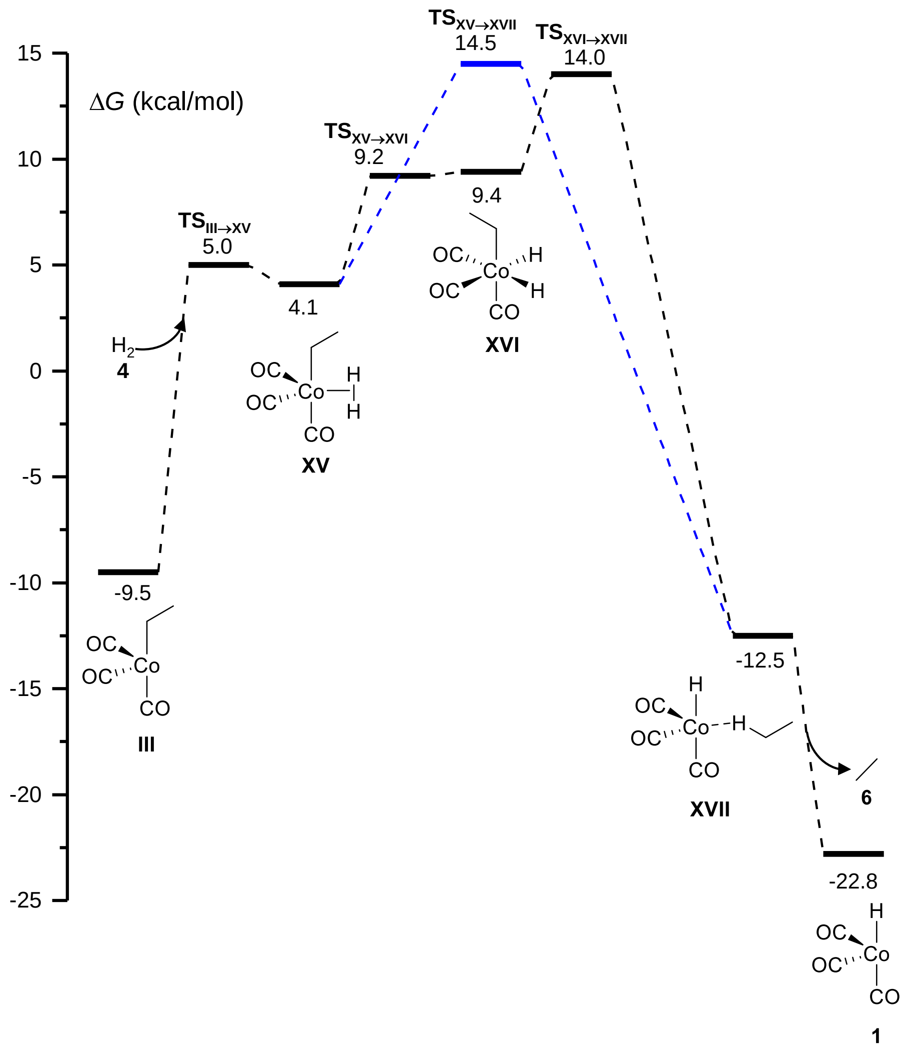


[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Mechanisms[1]

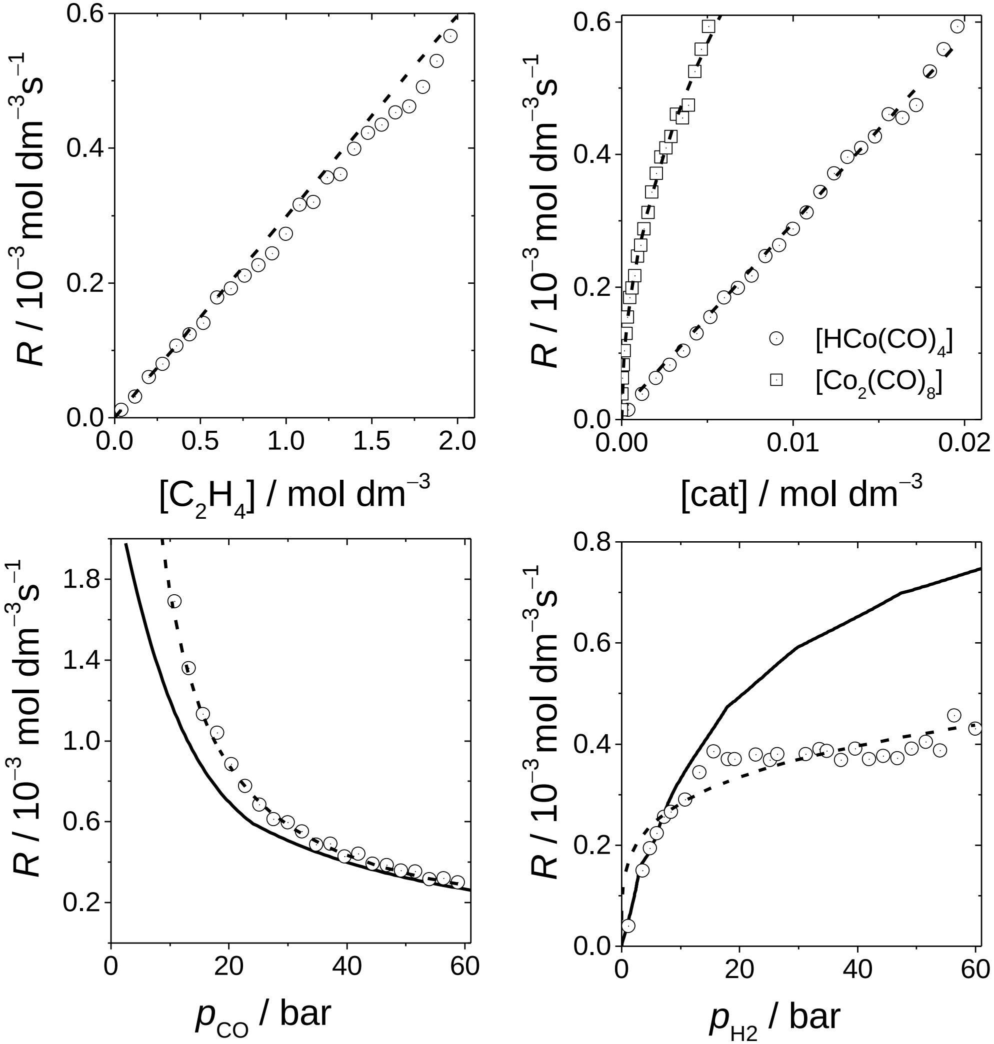
Side products: **ethane**; formaldehyde; propene; and water.



[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Rate law[1]



$$R_{exp} = k\frac{[H_2]^{0.6}[CO][cat]^{0.8}[alkene]}{(1+K[CO])^2}$$

$$R_{Harvey} = k\frac{[H_2]^{0.5}[cat]^{0.5}[alkene]}{[CO]}$$

$$R_{Habershon} = k\frac{[H_2][CO][cat][alkene]^{0.55}}{1+b[CO]+c[CO]^2}$$

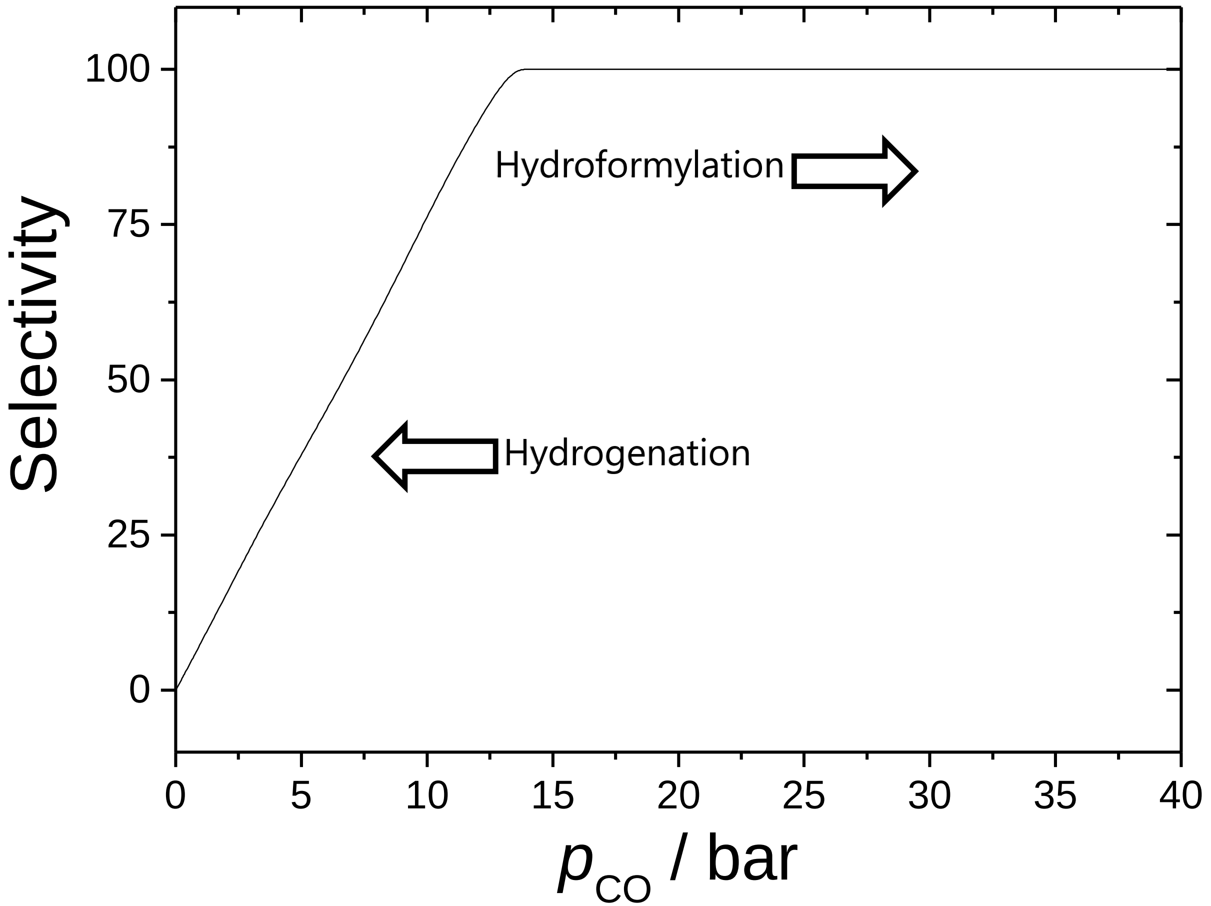
$$R_{TSSCDS} = k\frac{[H_2]^{0.4}[cat]^{0.5}[alkene]}{1+a_1[CO]+a_2[CO]^2}$$

[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Applications

Organometallic catalysis: Side reactions[1]

Hydroformylation vs Hydrogenation



[1]JA Varela *et al. Chem. Sci.* **2017**, *8*, 3843

Conclusions

-Efficient and highly parallelizable method to find TSs and reaction pathways

-Furnished with KMC codes to solve the kinetics.

-Successfully applied to different systems.

Future work

-Improve efficiency

-New electronic structure codes

-Barrierless reactions

Acknowledgements

Saulo Vázquez. Vinyl cyanide

Raúl Pérez-Soto and Saulo Vázquez. Acryloyl chloride

Estefania Rossich-Molina, and Riccardo Spezia. Protonated uracil

David Ferro, Antonio Fernández-Ramos and Saulo Vázquez. Combustion

Michael Wilhelm and Saulo Vázquez. Metcyanoformate

Jesús Varela and Saulo Vázquez. Catalysis

Developers:

Aurelio Rodríguez, Roberto Rodríguez-Fernández, Saulo Vázquez,

George Barnes and Jimmy Stewart

CESGA
Centro de Supercomputación de Galicia

XUNTA DE GALICIA

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD