

Automated chemical kinetics

Judit Zádor

Combustion Research Facility, Sandia National Laboratories, Livermore, CA, USA

University of the Pacific

October 9, 2018

Stockton, CA

This research was supported by the Exascale Computing Project (ECP), Project Number: 17-SC-20-SC, a collaborative effort of two DOE organizations, the Office of Science and the National Nuclear Security Administration, responsible for the planning and preparation of a capable exascale ecosystem including software, applications, hardware, advanced system engineering, and early test bed platforms to support the nation's exascale computing imperative. The views expressed in the presentation do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525



Outline

1. **Why do we care about chemical kinetics?**
2. Chemical kinetics of complex gas-phase reactions
3. KinBot
4. Sella

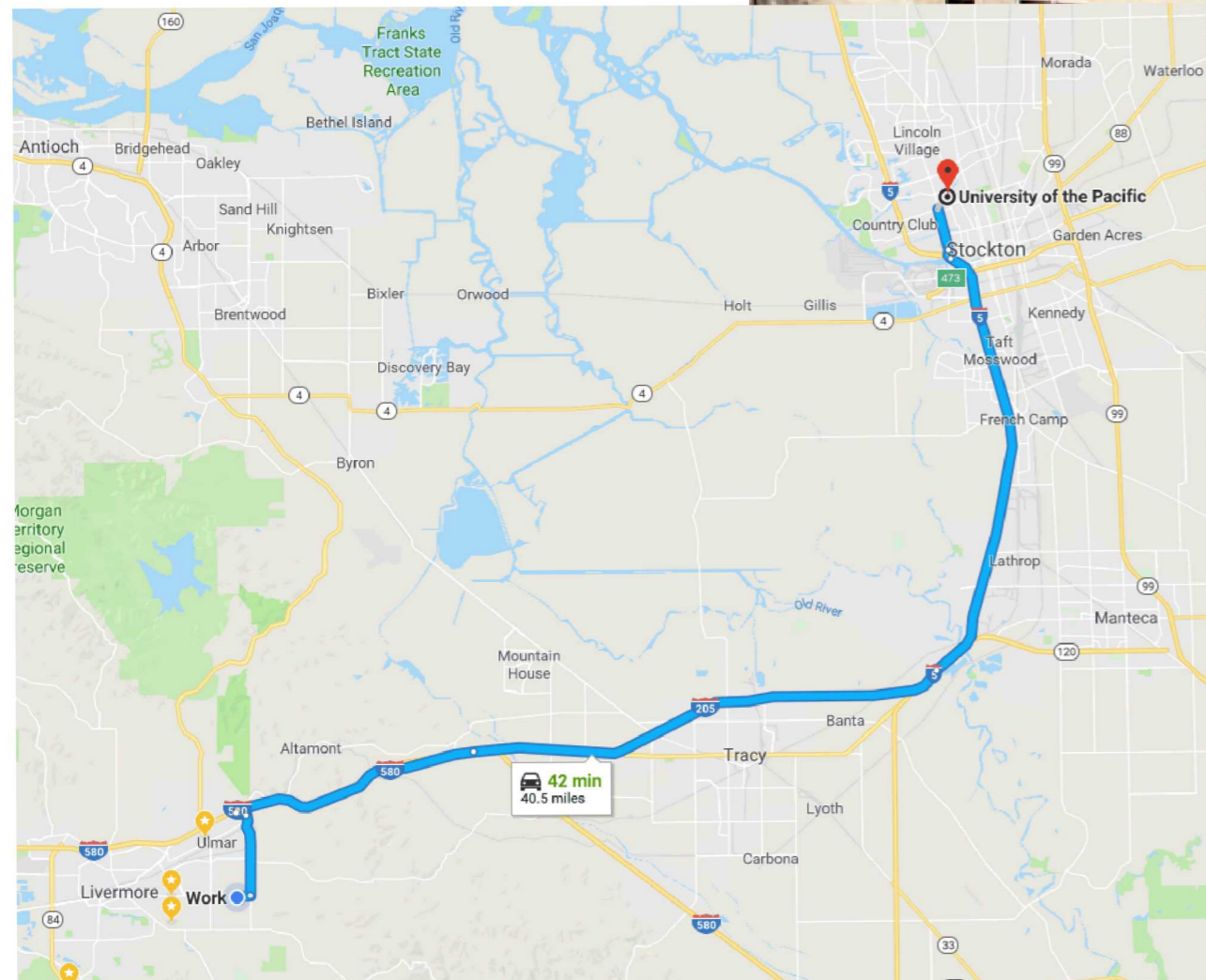


Sandia National Laboratories - Combustion Research Facility



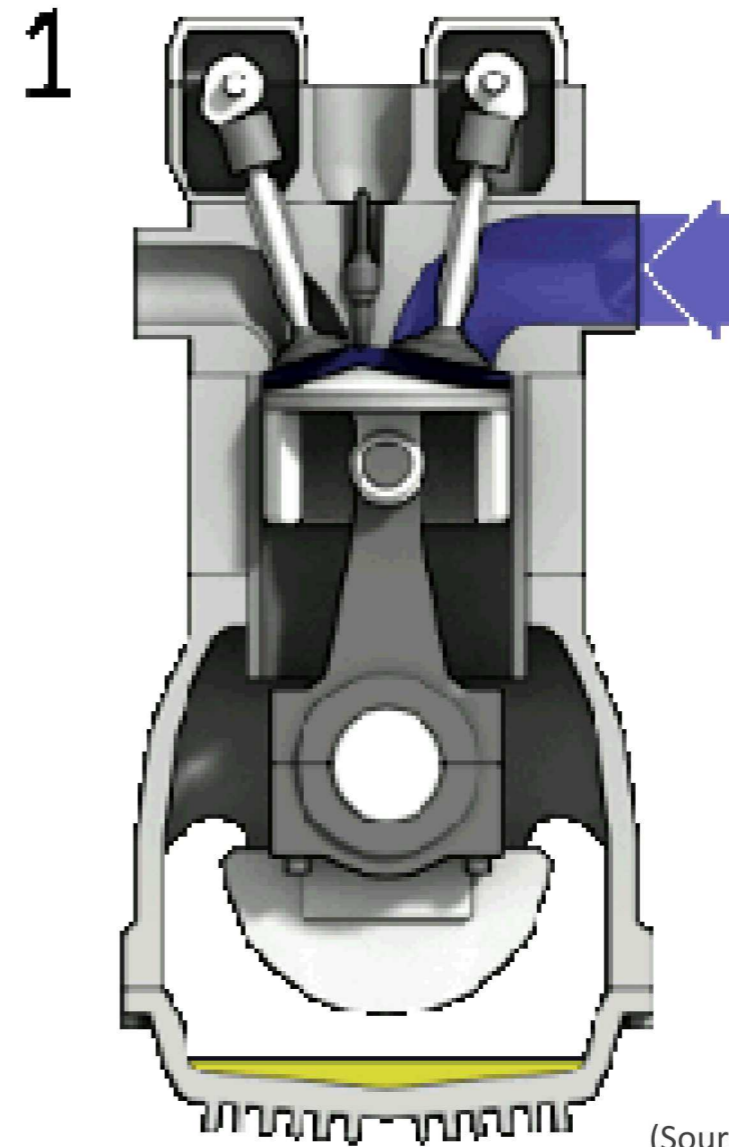
America's Combustion
Research Facility

For over 30 years, the CRF has served
as a national and international leader in
combustion science and technology.



What does the theoretical chemistry have to do with your car's engine?

$$H(t) |\psi(t)\rangle = i\hbar \frac{d}{dt} |\psi(t)\rangle$$

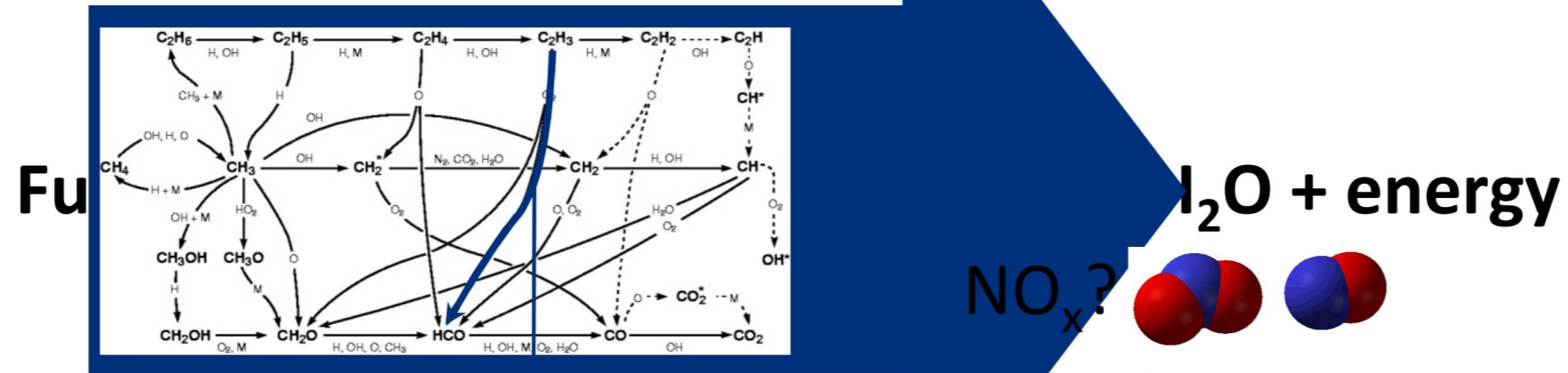


(Source: Wikipedia)

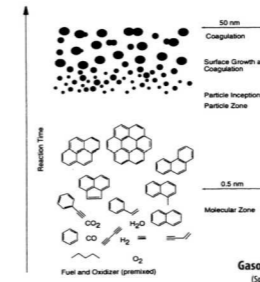
The engine is a very complex chemical reactor.



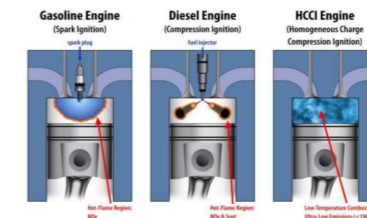
What happens inside an engine?



soot?

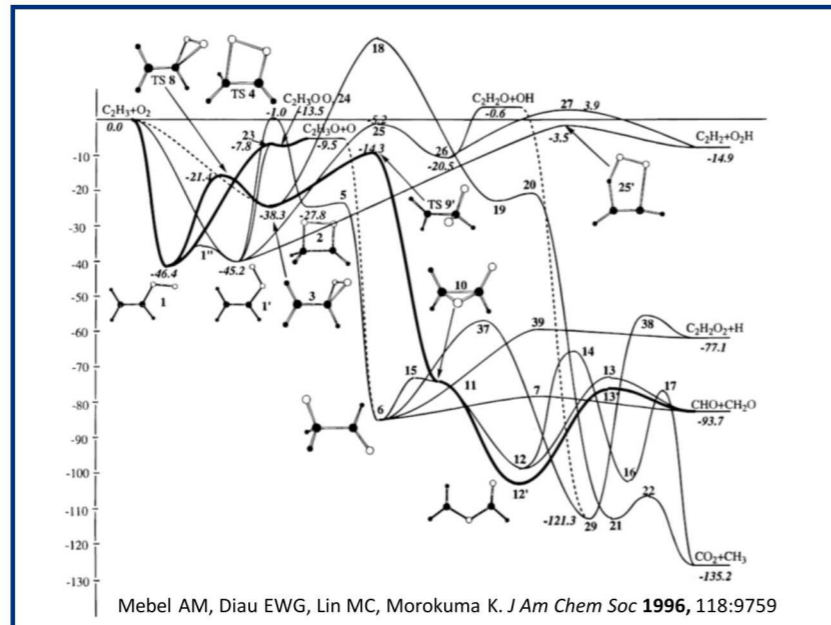


timing?



P and T dependence?

300 – 3000 K, $10 - 10^6$ Torr

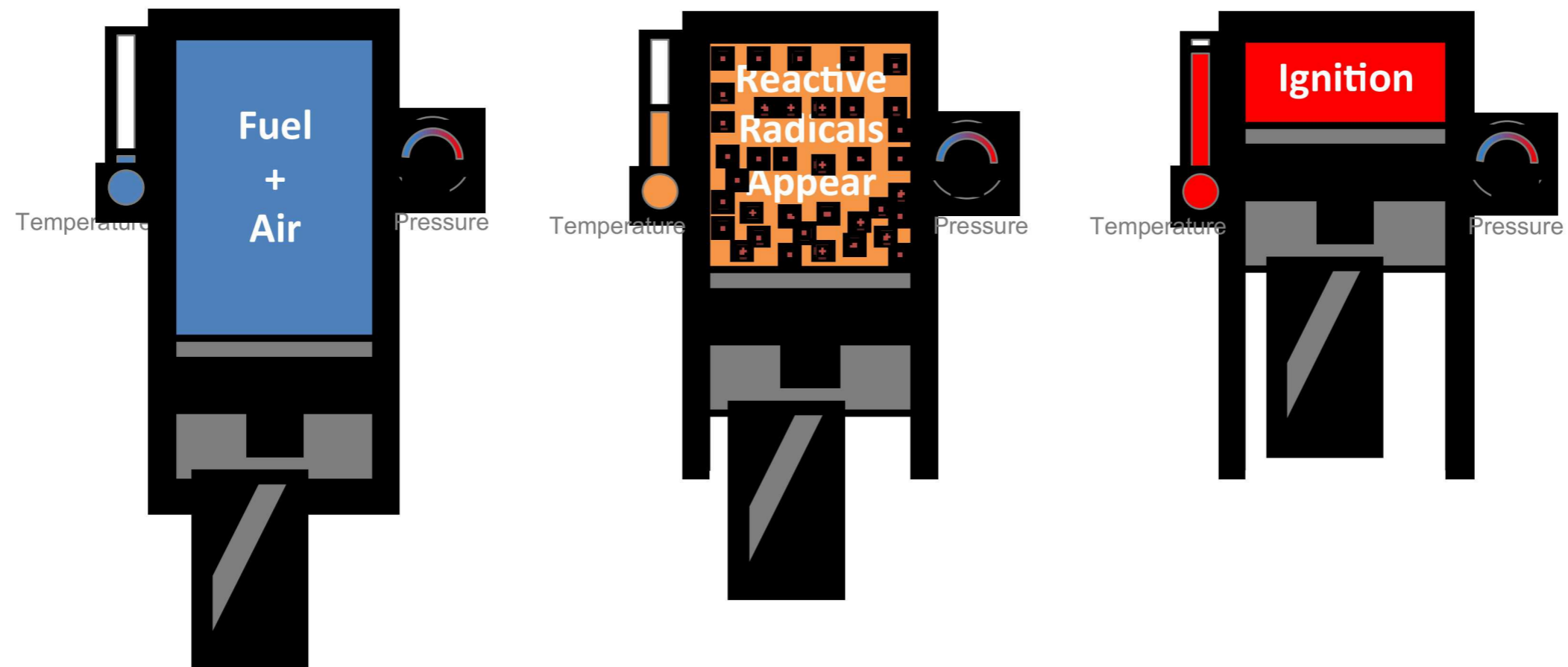


At least 8 papers on the reaction kinetics of vinyl + O_2 !

Details of combustion chemistry are critical in several areas.



Autoignition and engine development



- In spark plug engines compression can cause knock, which damages the engine
- In Diesel engines ignition is controlled by injection
- In Homogeneous Charge Compression Ignition (HCCI) engines combustion is initiated by thermal autoignition

Advanced engine concepts and the increasing use of alternative and non-traditional fuels present new challenges for combustion modeling.

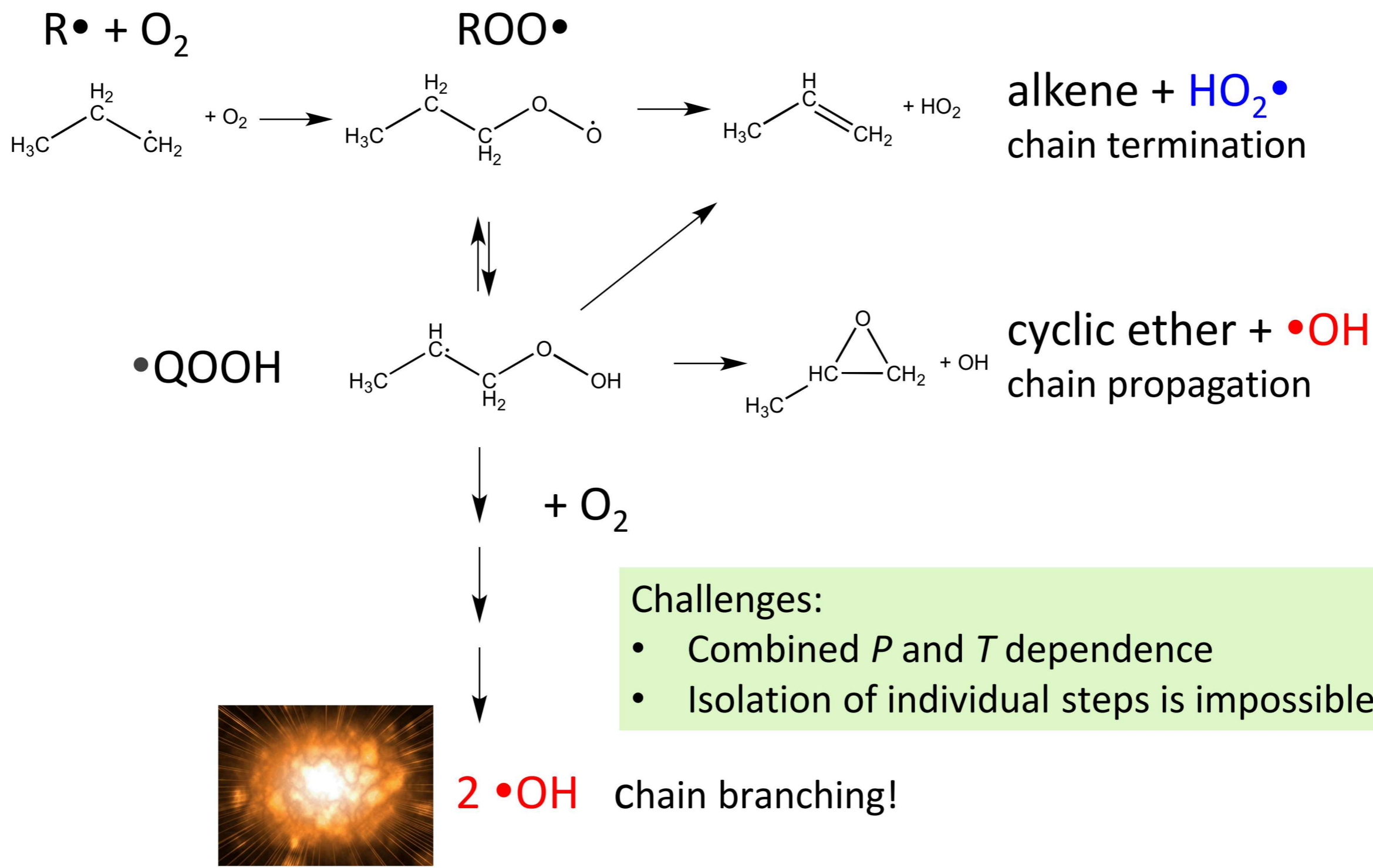


Outline

1. Why do we care about chemical kinetics?
- 2. Chemical kinetics of complex gas-phase reactions**
3. KinBot
4. Sella



Chain reactions of alkyl radicals with O₂ control hydrocarbon autoignition below ~900 K

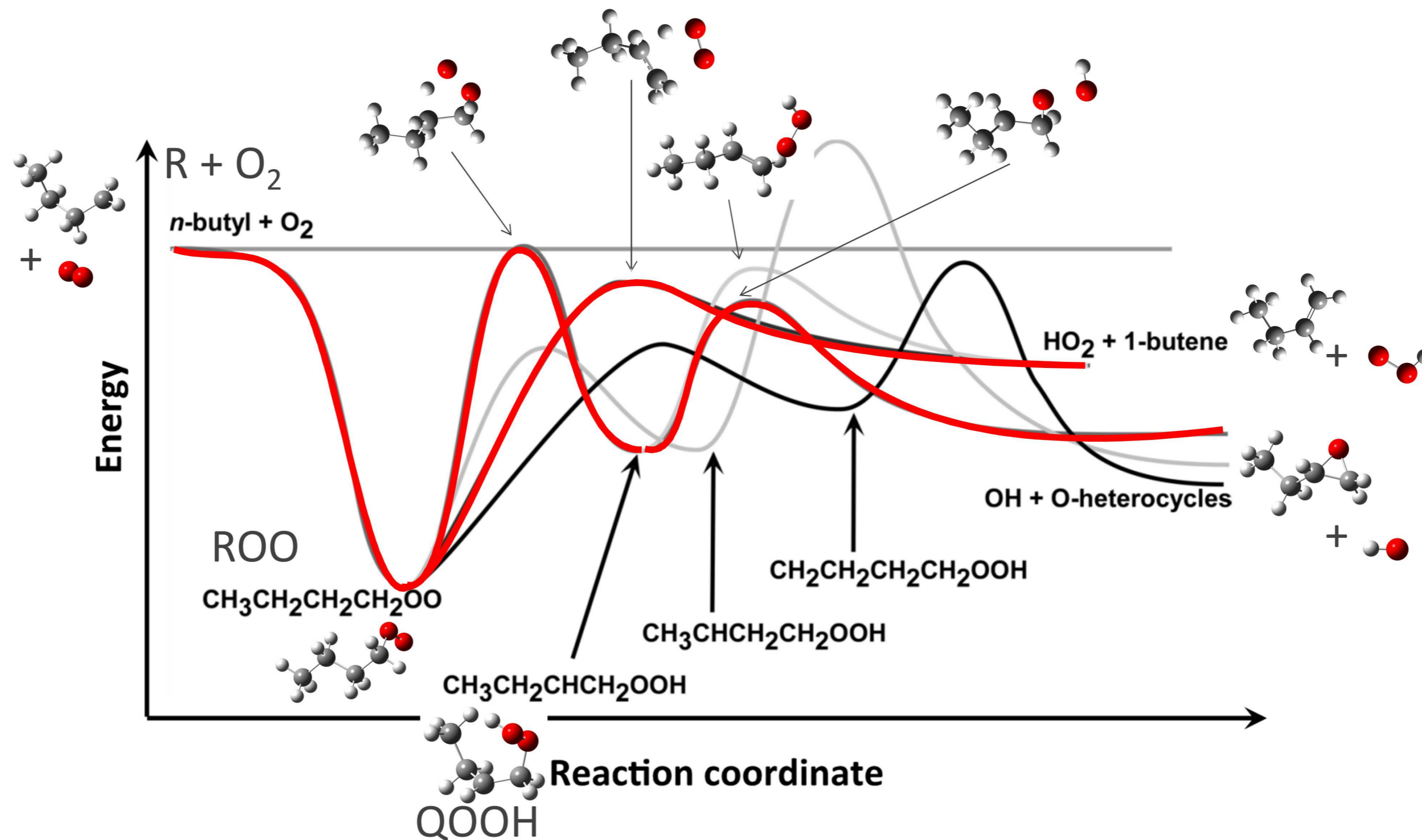


Challenges:

- Combined *P* and *T* dependence
- Isolation of individual steps is impossible



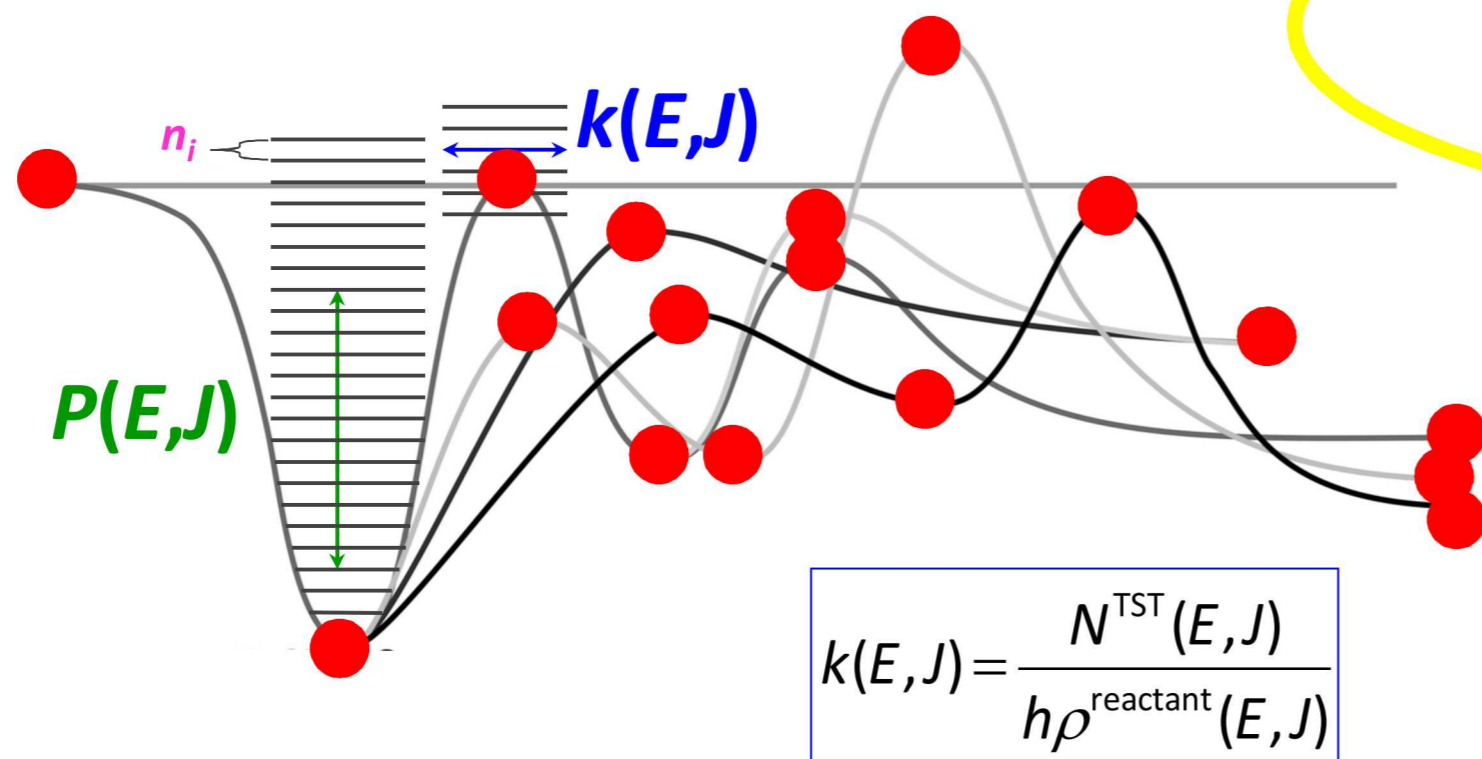
Under the hood: A typical alkyl + O₂ potential energy surface (PES)



We need to know the minima and saddle points to characterize a pathway.



The machinery of calculating rate coefficients from a PES: Master Equation



Energy and other properties of stationary points (minima and saddle points) are calculated from quantum chemistry.

Transition-state theory is used to compute energy and angular-momentum specific rate constants, $k(E,J)$.

Collisional energy transfer redistributes population among energy “levels”.

Multiple-well master equation solution gives time-dependent populations for all species.

$$\frac{dn_i(E,J)}{dt} = Z_i \sum_J \int_{E_{0i}}^{\infty} P_i(E,J;E',J') n_i(E',J') dE' - Z_i n_i(E,J) - \sum_{j \neq i}^M k_{ji}(E,J) n_i(E,J) + \sum_{j \neq i}^M k_{ij}(E,J) n_j(E,J) - k_d(E,J) n_i(E,J) + k_{a_i}(E,J) n_R n_m \rho_{Rm}(E,J) e^{-\beta E} / Q_{Rm} - \sum_{p=1}^{N_p} k_{p_i}(E,J) n_i(E,J)$$

$i = I, \dots, M$

$$\frac{d|w(t)\rangle}{dt} = \mathbf{G}|w(t)\rangle$$



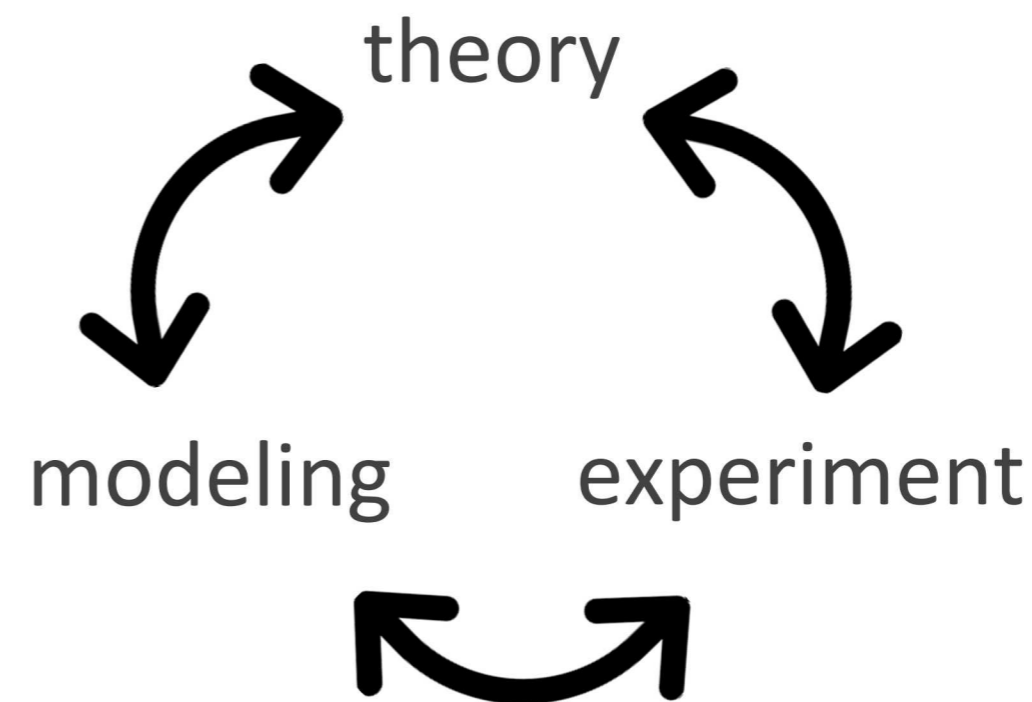
Relevance of PES in other fields

Autoignition and combustion is just one prominent field of research that needs to explore countless reaction pathways.

Other examples:

- atmospheric chemistry
- interstellar chemistry
- catalysis

... just to name a few.



How can we create predictive complex models?

Can we search the PES using the computer only?



Outline

1. Why do we care about chemical kinetics?
2. Chemical kinetics of complex gas-phase reactions
- 3. KinBot**
4. Sella



Current best practice to create combustion models

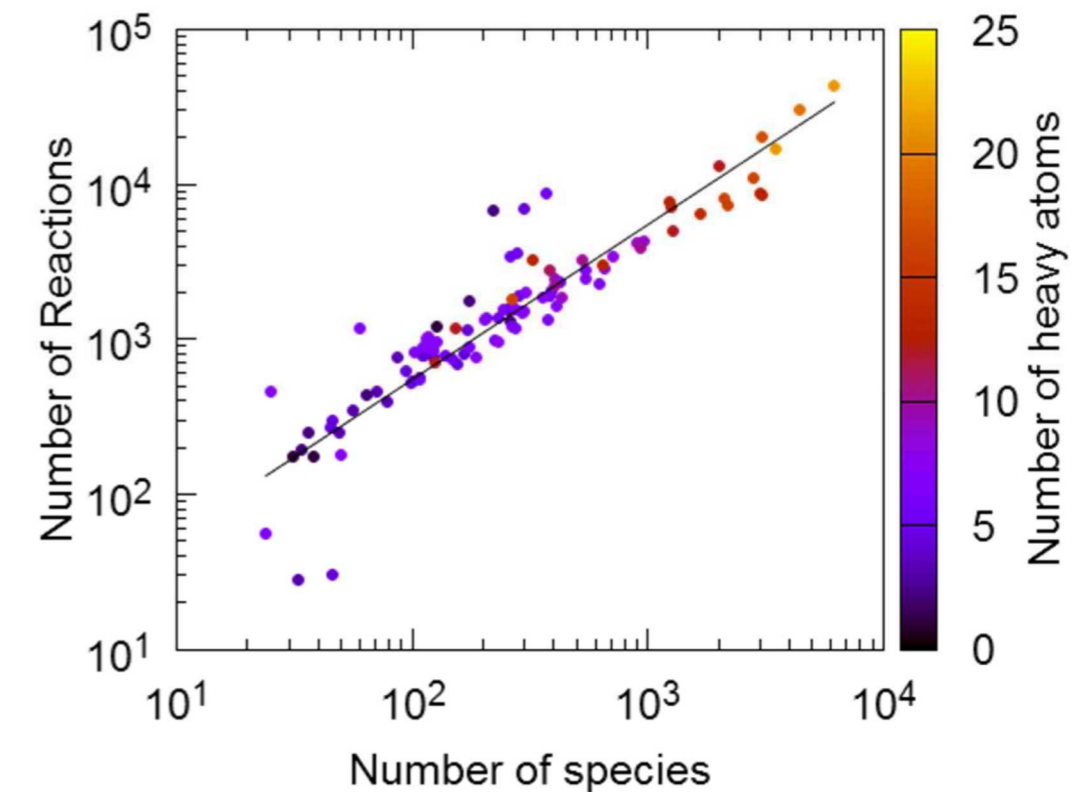
Kinetic models: hundreds, up to thousands, or even ten-thousands of reactions; many rate coefficients are needed

- Use of estimation methods **not always accurate**
- Experimental data: **scarce**
- High-level theoretical data: **scarce**

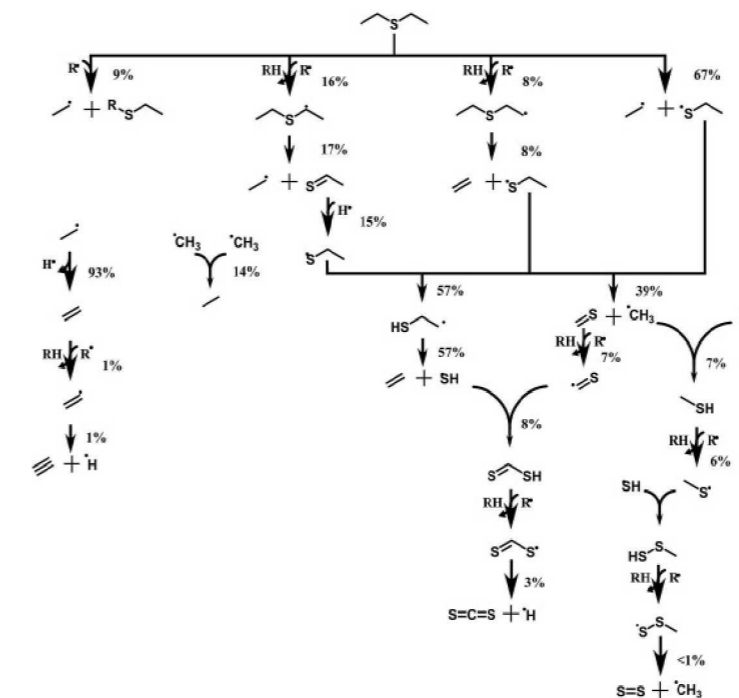
Challenges with reaction mechanisms and kinetic models:

1. Find all kinetically relevant **reaction pathways**
2. For each pathway, calculate **accurate rate coefficients**

- Manually: time-consuming, tedious and not always complete
- Automated searches: faster, systematic, less error-prone
- Development of the **KinBot** software



Van de Vijver, R. et al.; Int. J. Chem. Kinet. 2015, 47 (4), 199-231.



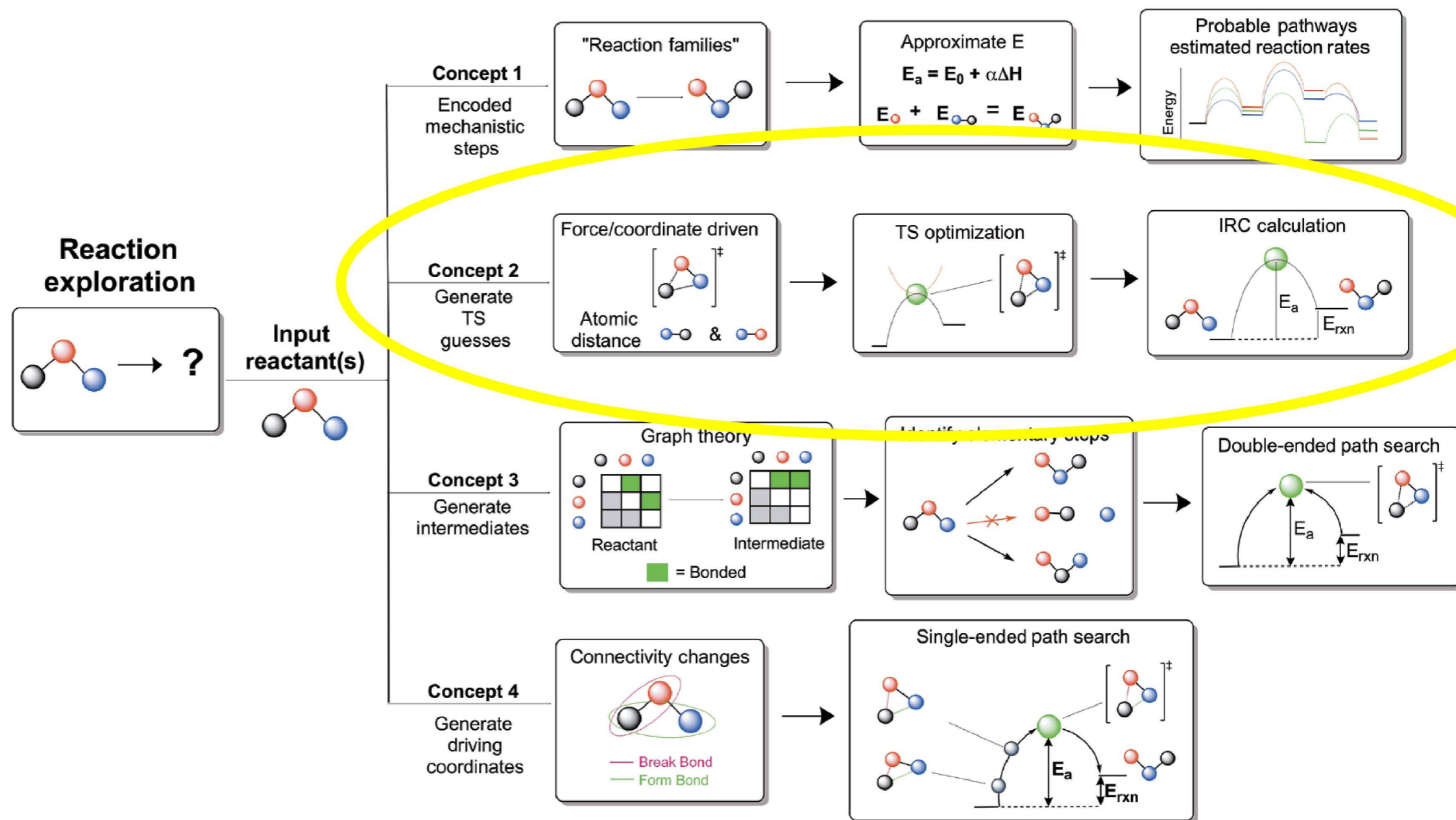
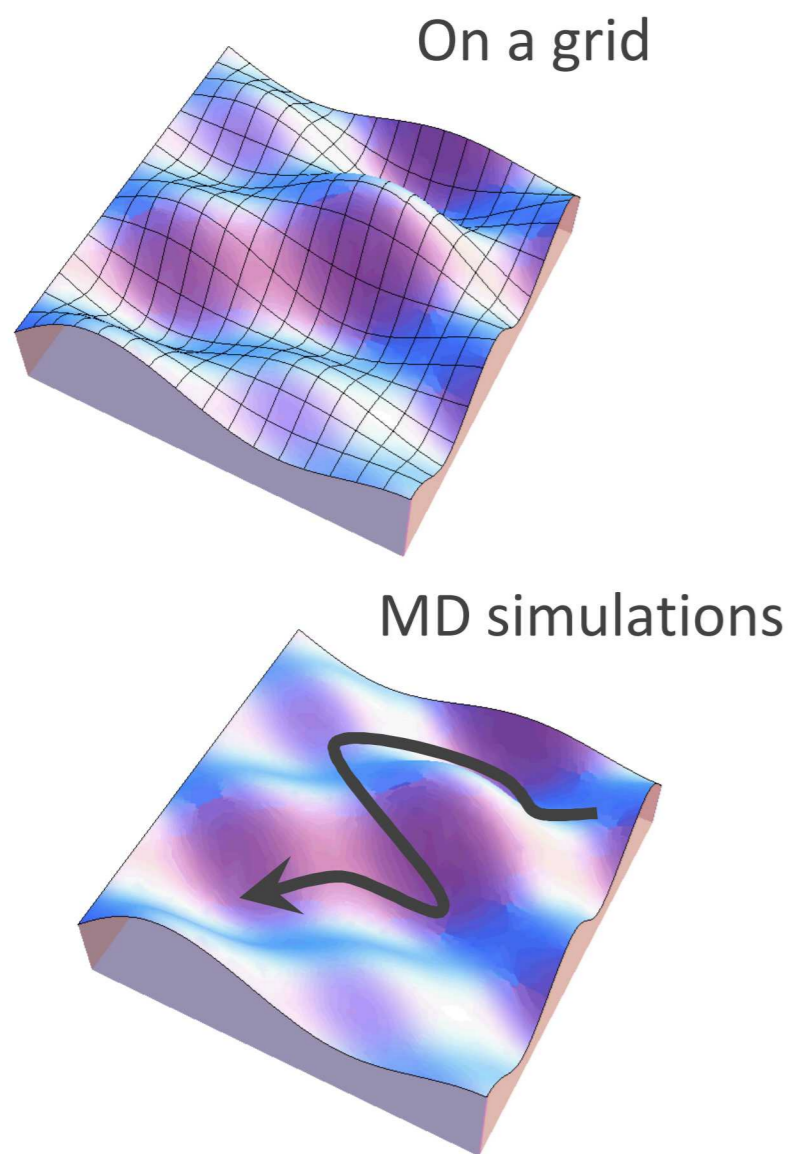
66 species
444 reaction
23 rate coefficients
from ab initio
421 rate coefficients
estimated

Van de Vijver, R. et al.; Chem. Eng. J. 2015, 278, 385-393.



How can we search the PES?

There are many approaches... just a few examples:



KinBot

Goal

Find and characterize reaction pathways (minima and saddle points)

Translate the chemistry we know into computer codes

- Bookkeeping: many reactions, many products
- Identify identical species, pathways
- Account for conformers, anharmonicities, symmetry, etc.

Current focus

- Monomolecular reactions: isomerizations and decompositions (e.g. no intermolecular abstraction reactions)
- C, H, S and O

Python code coupled to:

- Third party quantum chemistry software: Gaussian, NWChem, QChem and MolPro
- Atomic Simulation Environment (ASE)
- Use of open-source chemo-informatics tools: OpenBabel and RDKit
- Master equation solvers: MESS and MESMER
- Workload managers: PBS and SLURM

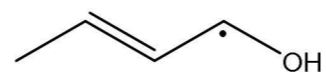


Ruben Van de Vijver

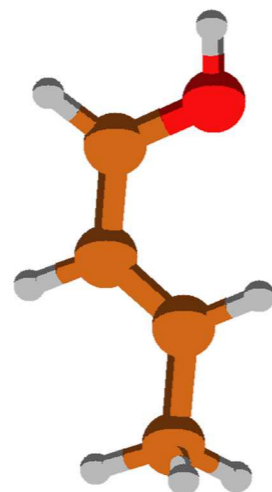


KinBot machinery

- Input
 - Species: SMILES
 - Options
- 3D structure
 - OpenBabel
- Bond matrix
 - Resonance
 - Radical sites
 - Cycle information
 - Atom equivalencies



CC=C[CH]O



C	1.03	0.03	-0.06
C	0.52	-0.09	-1.46
C	-0.29	-1.06	-1.89
C	-0.86	-1.09	-3.20
H	-1.83	-1.50	-3.43
O	-0.27	-0.33	-4.18
H	0.69	-0.78	0.58
H	0.68	0.98	0.37
H	2.13	0.04	-0.06
H	0.82	0.71	-2.14
H	-0.61	-1.85	-1.21
H	-0.77	-0.35	-5.01

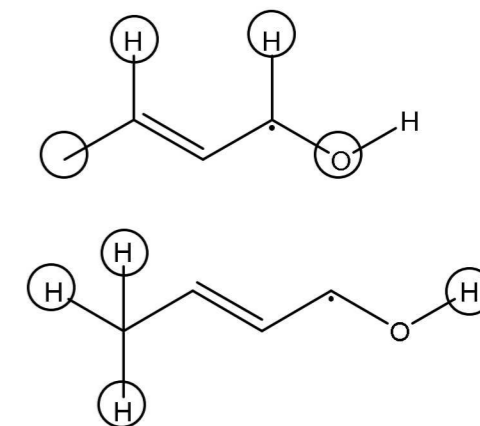
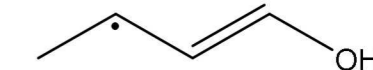
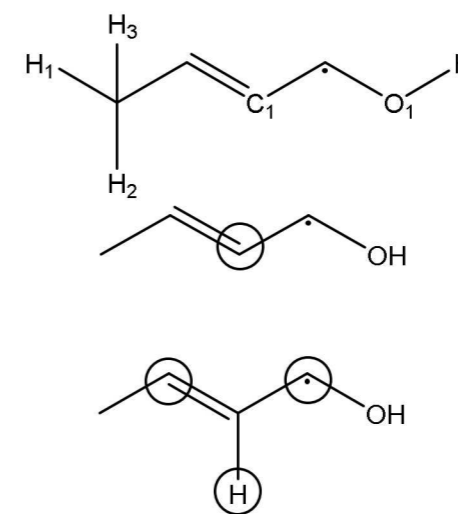
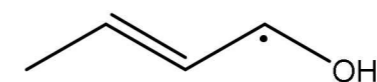
C	0	1	0	0	0	3
C	1	0	1	0	0	1
C	0	1	0	1	0	1
C	0	0	1	0	1	1
O	0	0	0	1	0	1
H	3	1	1	1	1	

C	0	1	0	0	0
C	1	0	2	0	0
C	0	2	0	1	0
C	0	0	1	0	1
O	0	0	0	1	0
Rad	0	0	0	1	0

C	0	1	0	0	0
C	1	0	1	0	0
C	0	1	0	2	0
C	0	0	2	0	1
O	0	0	0	1	0
Rad	0	1	0	0	0

C	0	1	0	0	0
C	1	0	2	0	0
C	0	2	0	1	0
C	0	0	1	0	1
O	0	0	0	1	0
Rad	0	0	0	1	0

	0	1	2	3	Id
C1	12	025	030	004	12025030004
O1	16	013	013	013	16013013013
H1	1	012	014	013	1012014013
H2	1	012	014	013	1012014013
H3	1	012	014	013	1012014013
...					
Mol ID					71195216172



KinBot machinery

- Reaction families
 - As defined in RMG¹
 - Generalized
- Reaction searches
 - Motif searches

H migration in closed-shell molecules

Reverse Diels Alder

Intramolecular Diels Alder

Endocyclic closed-shell cyclization

Exocyclic closed-shell cyclization

Korcek reaction

Retro-ene reaction

Reverse 1,2 cycloaddition

Reverse 2,2 cycloaddition

Reverse 1,2 R insertion

Reverse 1,3 CO₂ insertion

Reverse 1,3 ROR insertion

H migration in radicals

OH migration

HO₂ elimination

Radical β-scission

CO α-scission

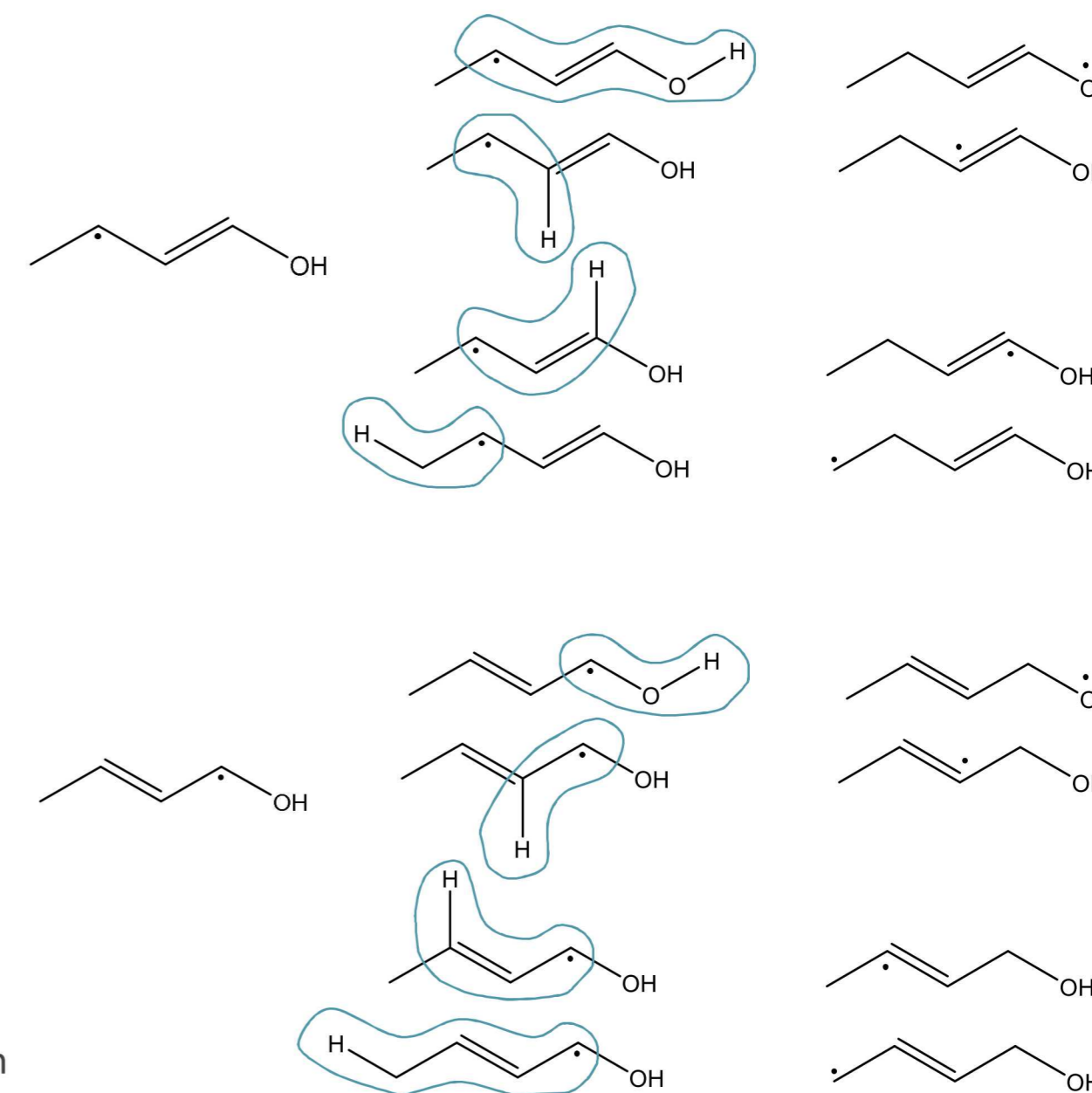
Exocyclic intramolecular radical addition

Endocyclic intramolecular radical addition

Cyclization-elimination

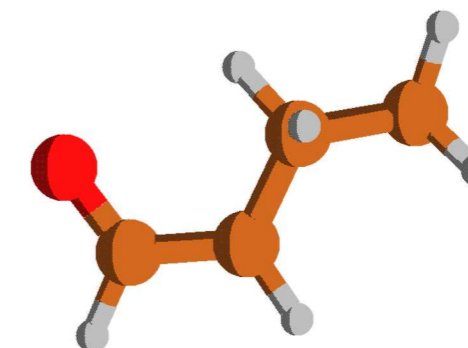
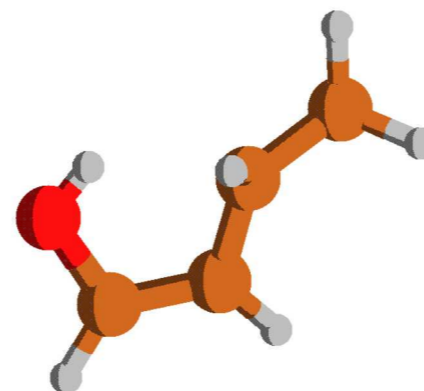
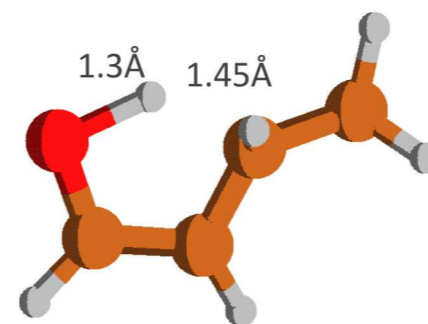
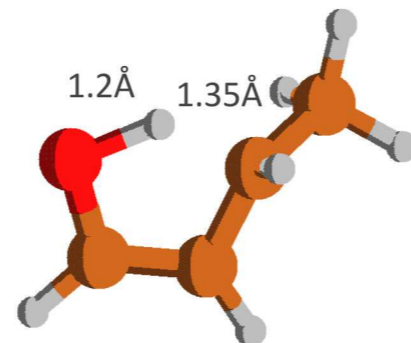
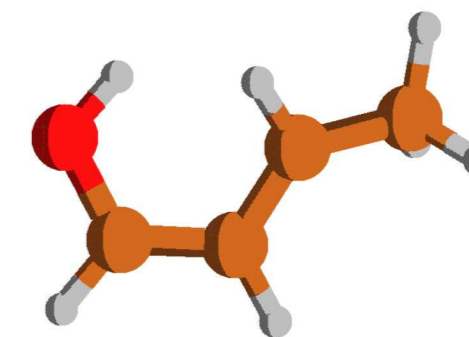
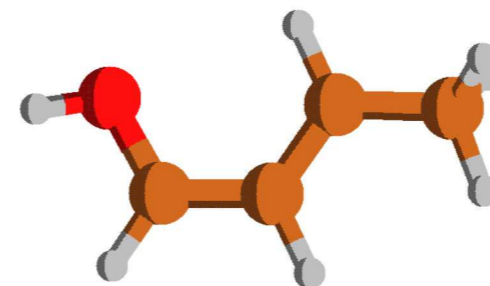
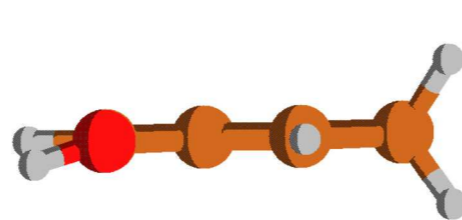
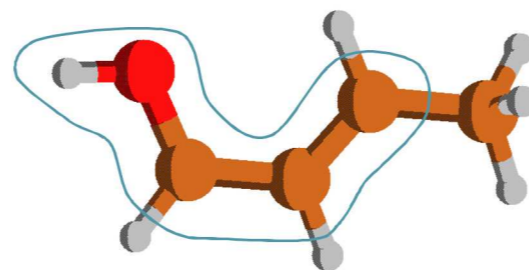
H-migration in radicals:

H-X-...-Ẋ

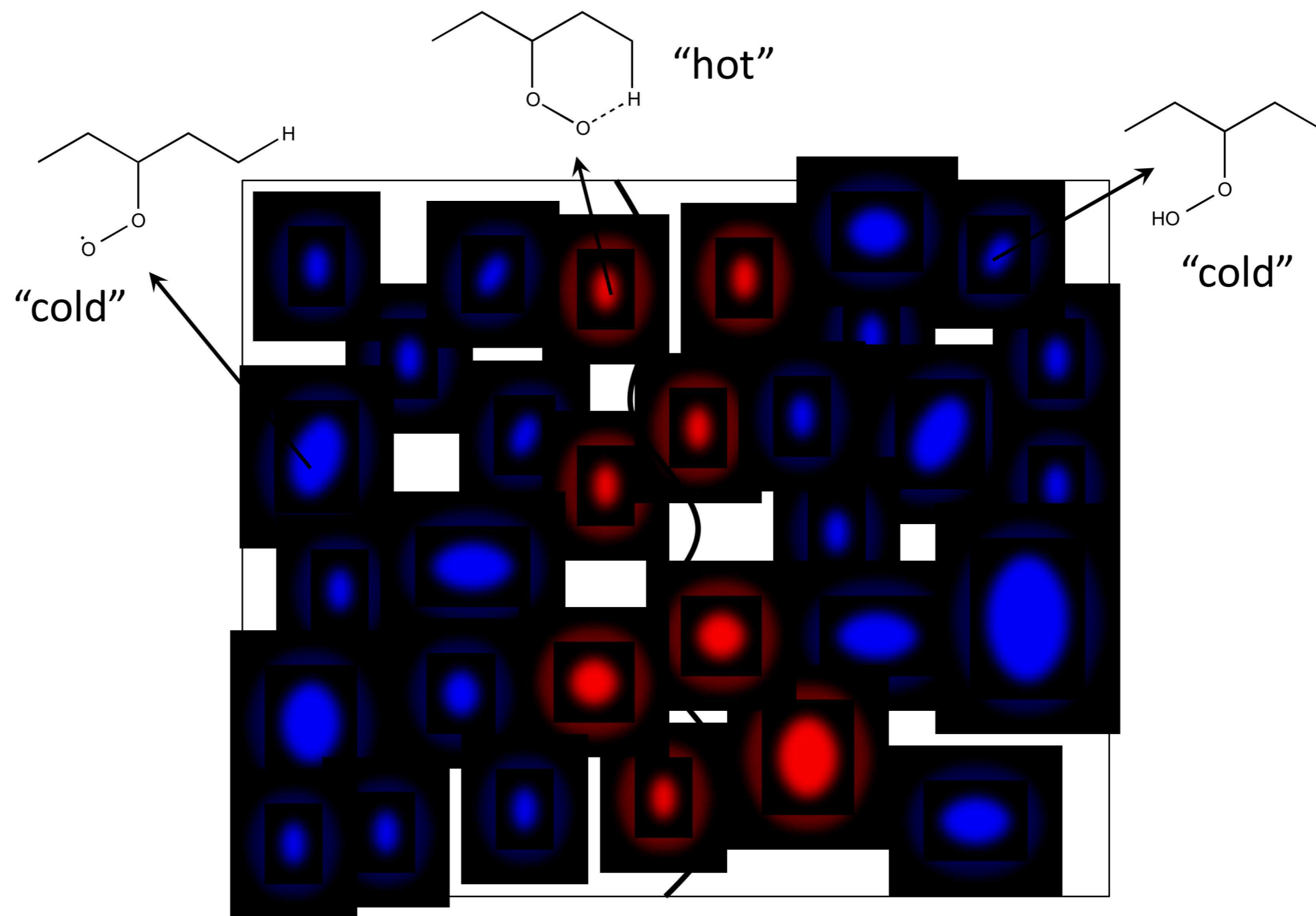


KinBot machinery

- Initial guess for TS geometry
 - Start from reactant geometry
 - Rotate around bonds
 - Fix bond lengths
- Saddle point optimization
- IRC calculations
 - Verify connectivity to well
 - Identify product species



Only certain conformers of reactants and products are connected



Example of internal H-abstraction pathways

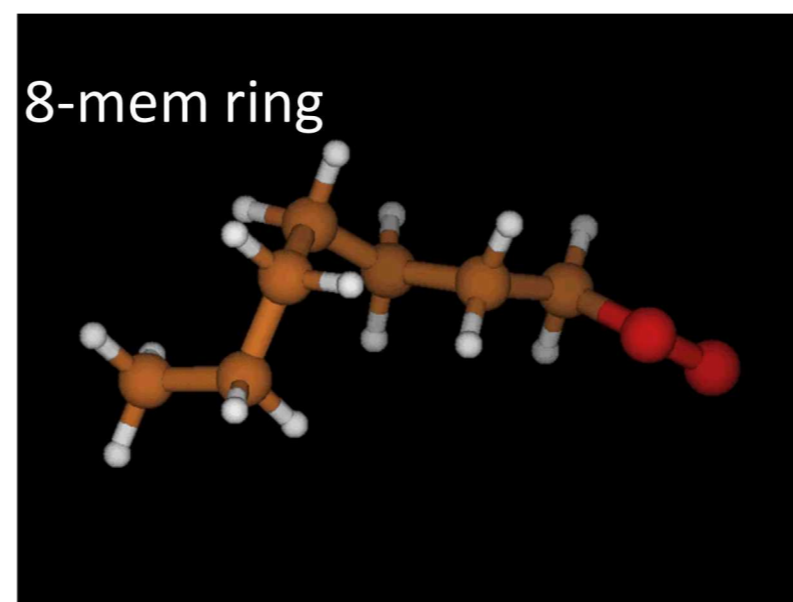
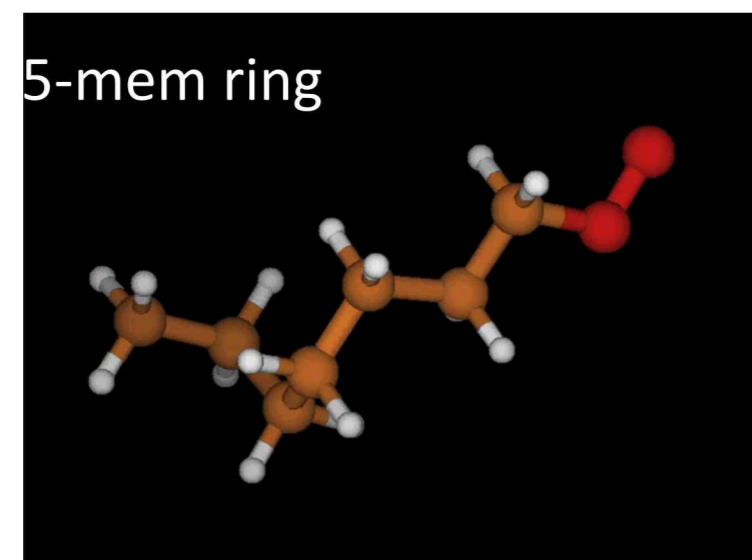
critterion

- structure should contain $X^{\bullet} - nX - H$ motif, $n = 3, \dots, 8$

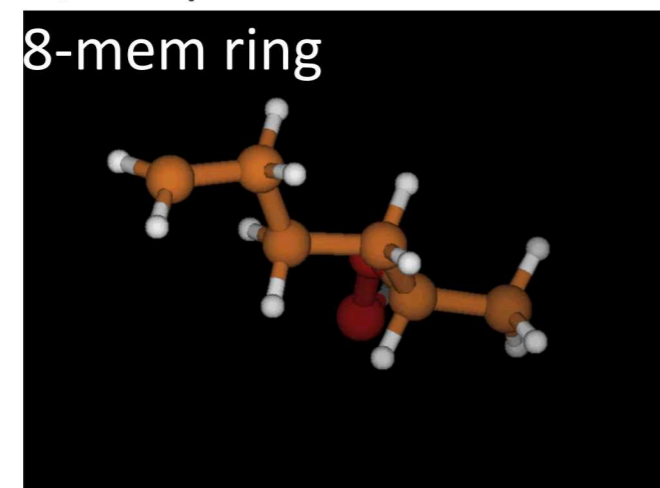
strategy

- fold molecule into a cyclic conformer in a series of steps to generate good initial guess

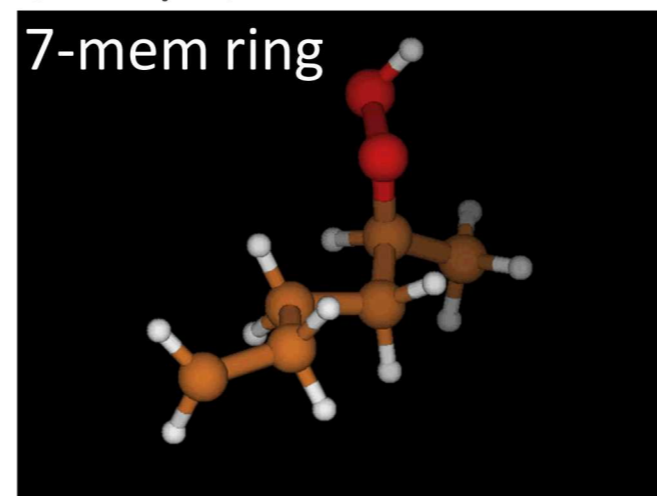
ROO/QOOH isomerization



QOOH/ROO isomerization



QOOH/QOOH isomerization



KinBot machinery

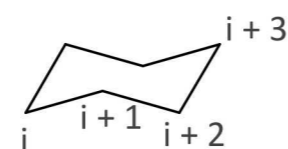
- For each stationary point
 - Conformational search
 - High-level optimization
 - Frequency calculations
 - Hindered rotor calculation
 - Symmetry calculation

- Visualize results

- Write output files

- Summary
- MESS and MESMER input files

Cyclic compounds

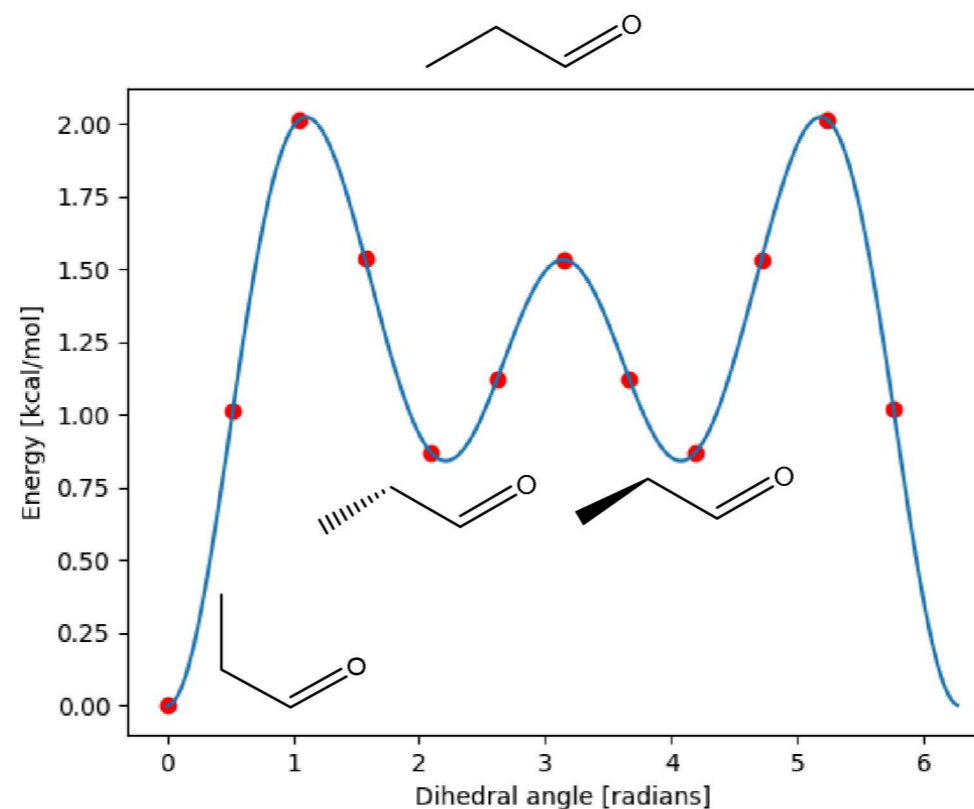
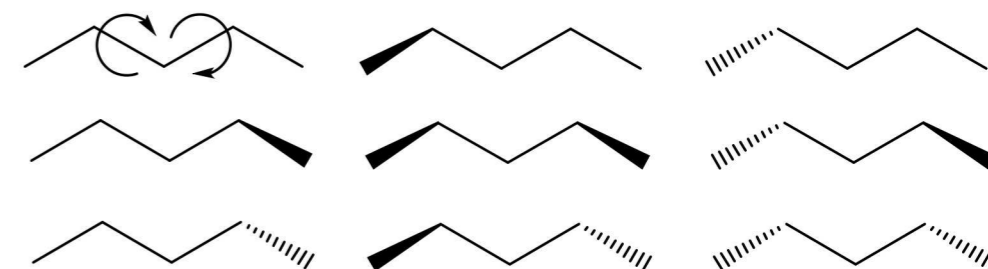


$$f = \frac{1}{N} \sum_{i=1}^N |dih_{i,i+1,i+2,i+3}|$$

Open chain groups

Combine all possible combinations of +0, +120 and +240 degrees for all the asymmetrical rotors

Select N-3 dihedrals randomly
Combinations of these dihedrals to values -f, 0 and +f

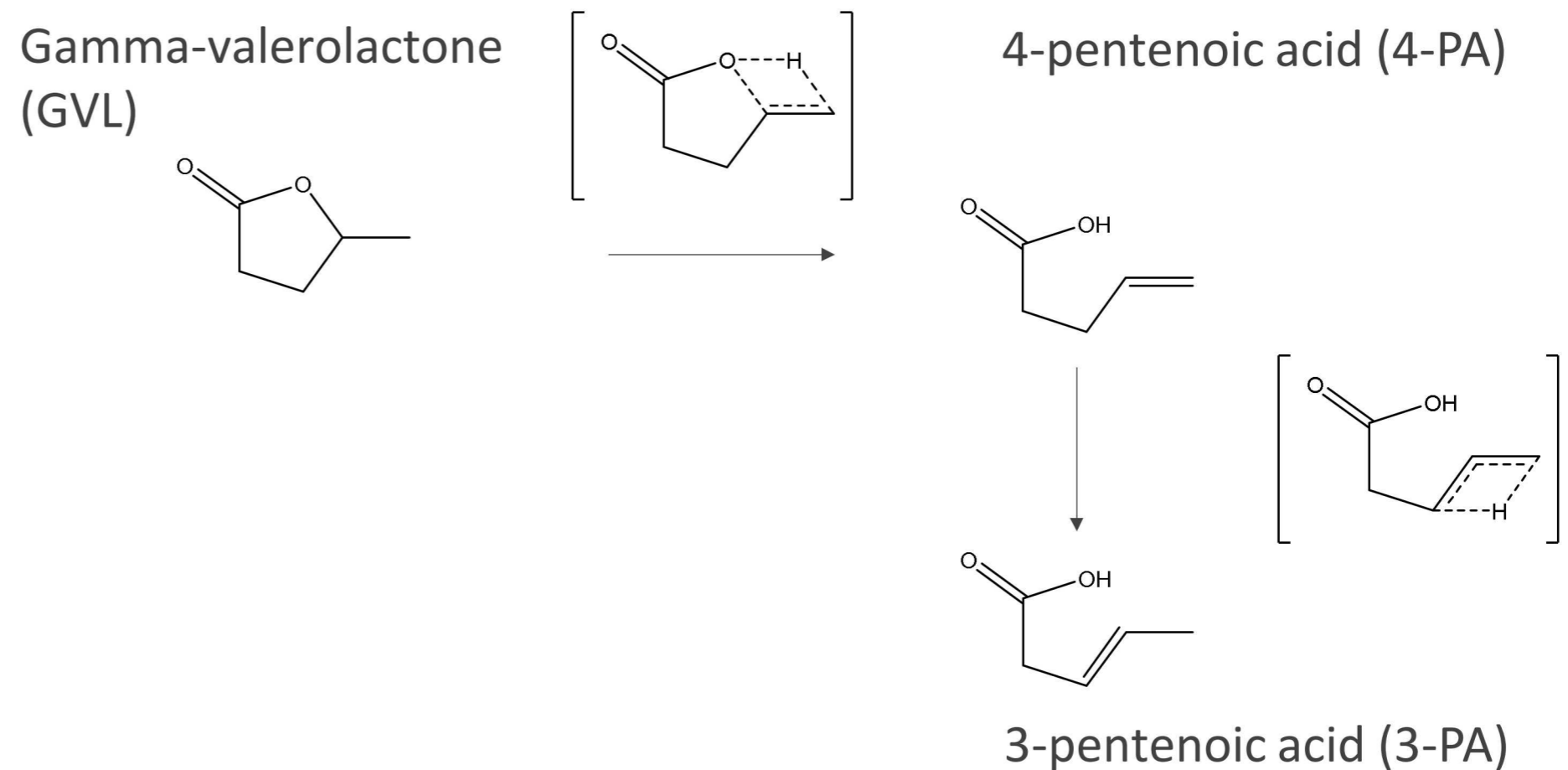


Number of neighbors	Equivalent neighbors	σ_{ext}	n_{opt}
1	0	1	1
2	2	2	1
2	0	1	1
3	3	6	1
3	2	1	1
3	0	1	1
4	4	12	1
4	3	1	1
4	2 × 2	2	1
4	2	1	1
4	0	1	2



Gamma-valerolactone thermal decomposition

- Potential biofuel
- Studied in the past^{1,2}
- Reactions of GVL, 4-PA and 3-PA
- Reaction searches: B3LYP/6-31G
- High-level calculations: B3LYP/6-311++G(d,p)



1. De Bruycker, R. et al.; Proc. Combust. Inst. 2015, 35 (1), 515-523.

2. Ye, L. et al. ; RSC Adv. 2018, 8 (23), 12975-12983.

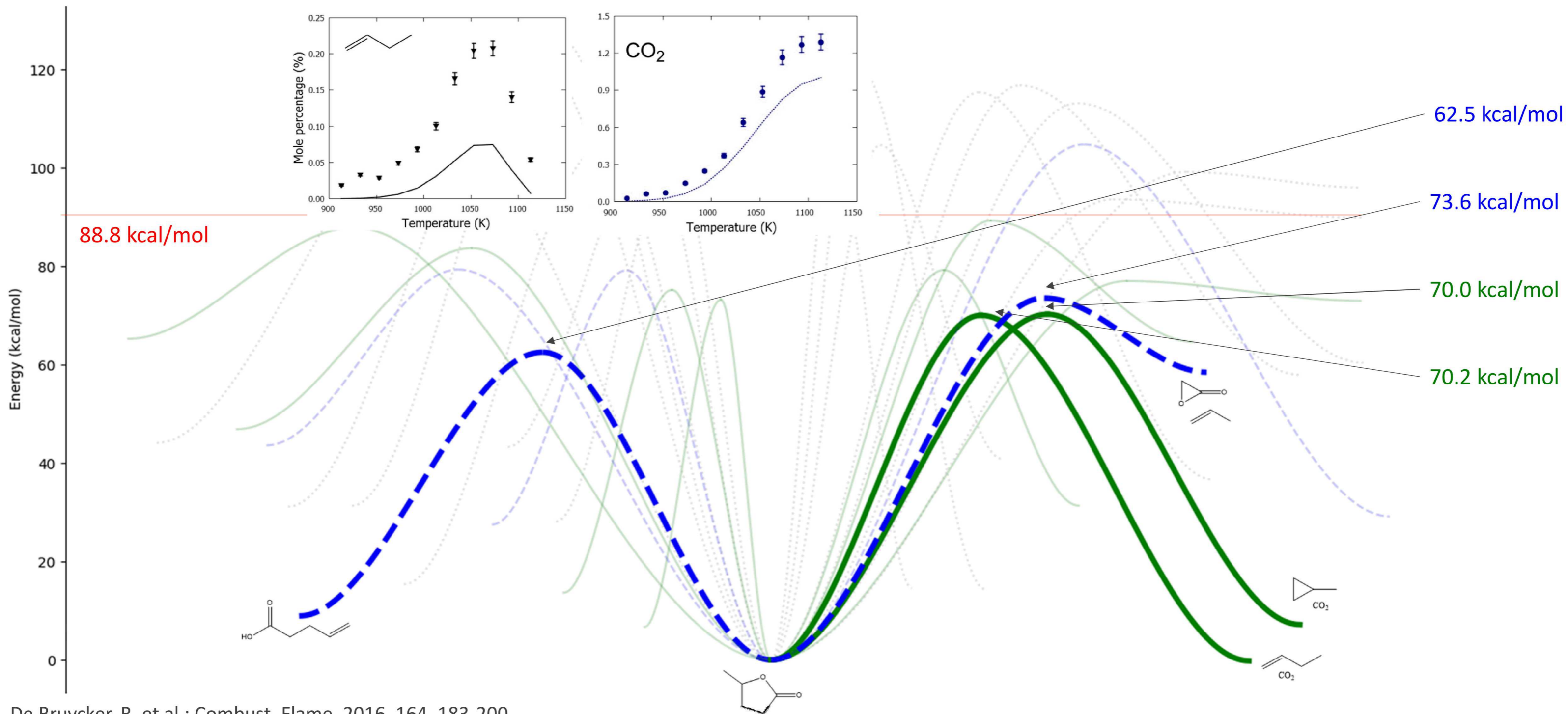


Gamma-valerolactone

Below 1050K: Molecular chemistry

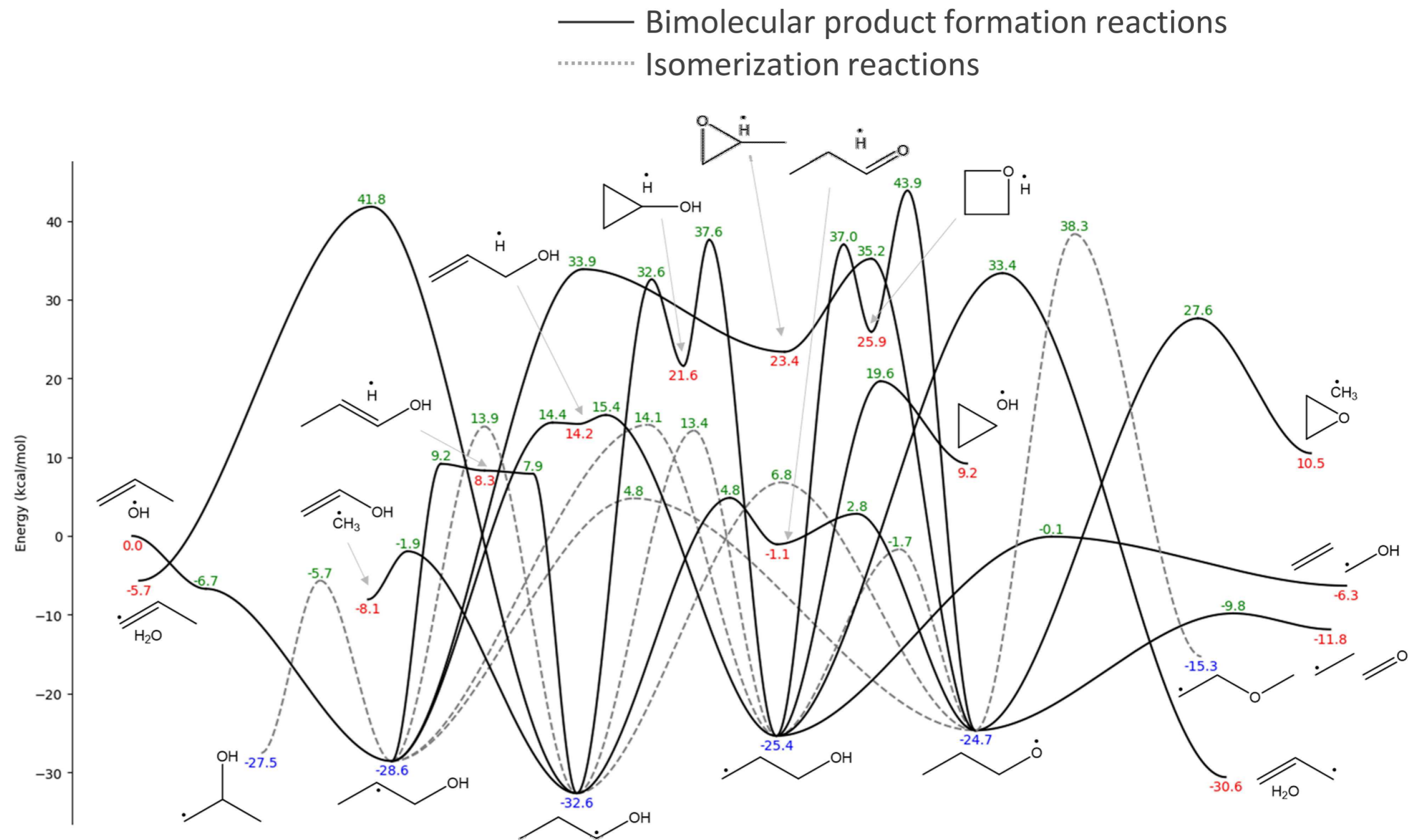
Above 1050K: Radical chemistry

- “New reactions” by KinBot
- - - Previously reported reaction (also found by KinBot)
- ⋯ “New reactions” with high barrier



C₃H₆+OH

- Important reaction in combustion
- Studied in the past^{1,2}
- Reaction searches:
B3LYP/6-31G
- High-level calculations:
B3LYP/6-311++G(d,p)

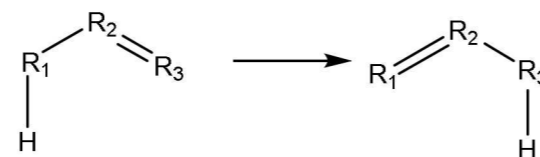


1. Zador, J. et al.; PCCP 2009, 11 (46), 11040-11053.
2. Kappler, C. et al.; Z. Phys. Chem., 2011, 225, 1271.

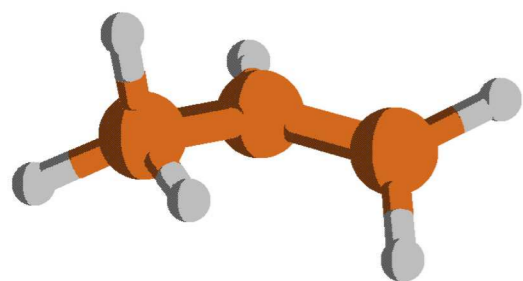


Sigmatropic [1,3] hydrogen shift

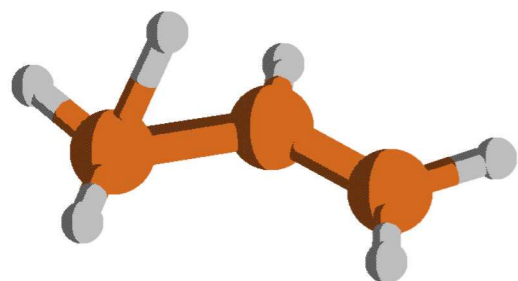
- Woodward-Hoffmann
- Möbius-Hückel
- Antrafacial is thermally allowed
- Suprafacial is thermally forbidden



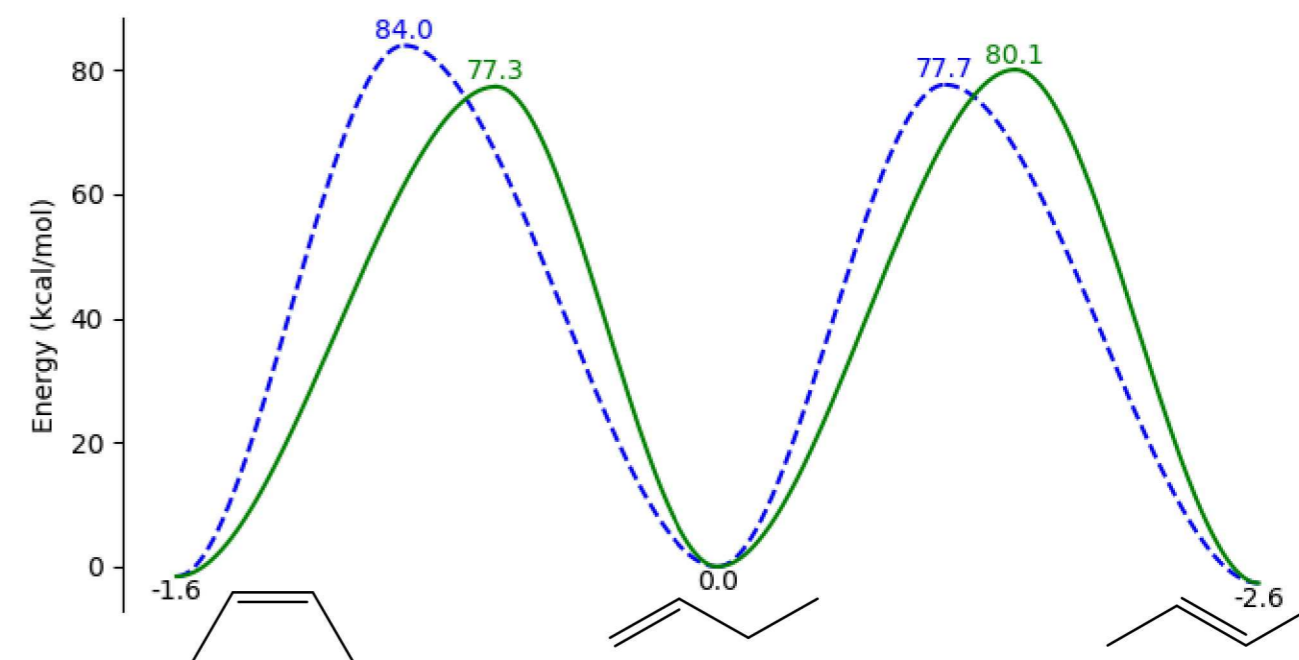
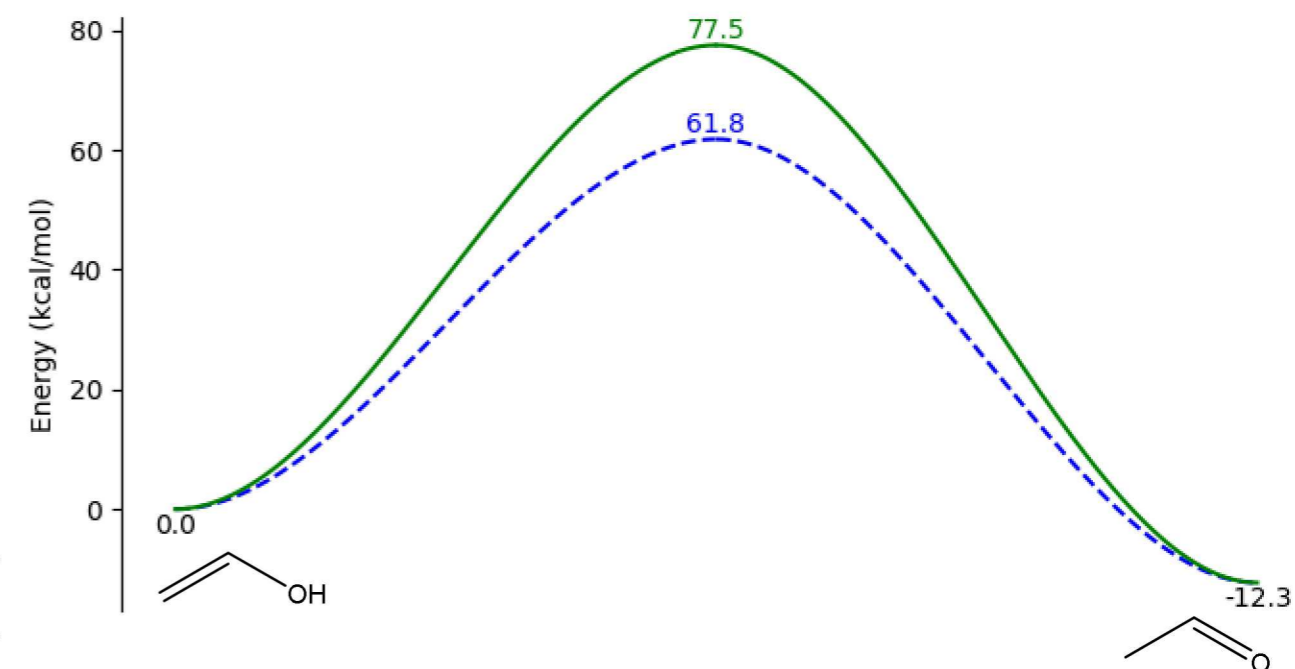
antrafacial



suprafacial

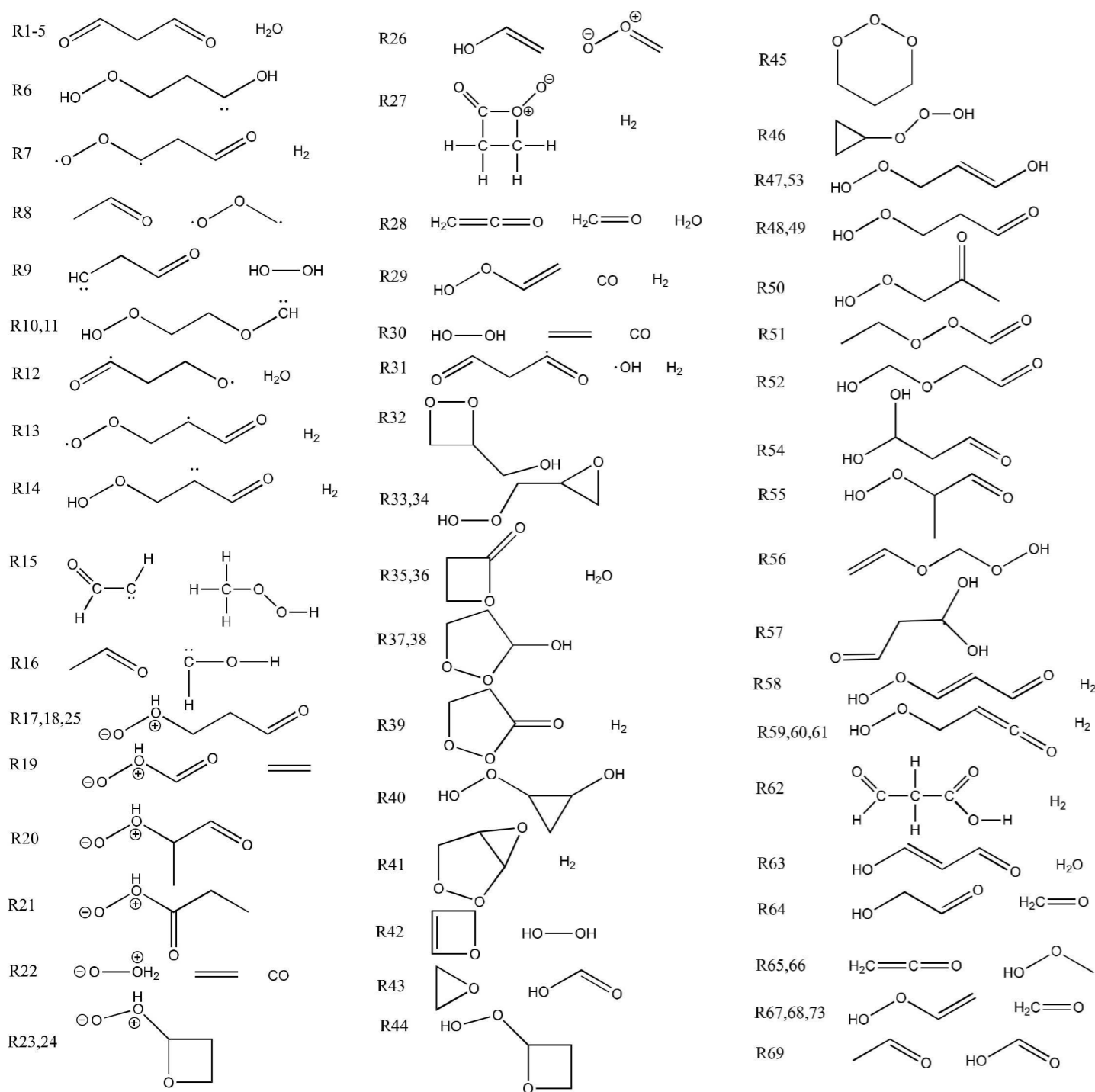


Reactant	Antra	Supra	Product
Vinyl alcohol	61.8	77.5	Acetaldehyde
Propene	81.2	83.5	Propene
Formic acid	38.5		Formic acid
2-methyl-1-butene	75.7	78.9	2-methyl-2-butene
Propanal	72.7	83.9	Trans-1-propenol
Propanal	77.1	82.3	Cis-1-propenol
1-butene	84.0	77.3	Cis-2-butene
1-butene	77.7	80.1	Trans-2-butene
3-methyl-1-butene	81.9	74.9	3-methyl-2-butene
4-pentenoic acid	87.6	78.7	Cis-3-pentenoic acid
4-pentenoic acid	81.8	82.3	Trans-3-pentenoic acid



Comparative ketohydroperoxide study

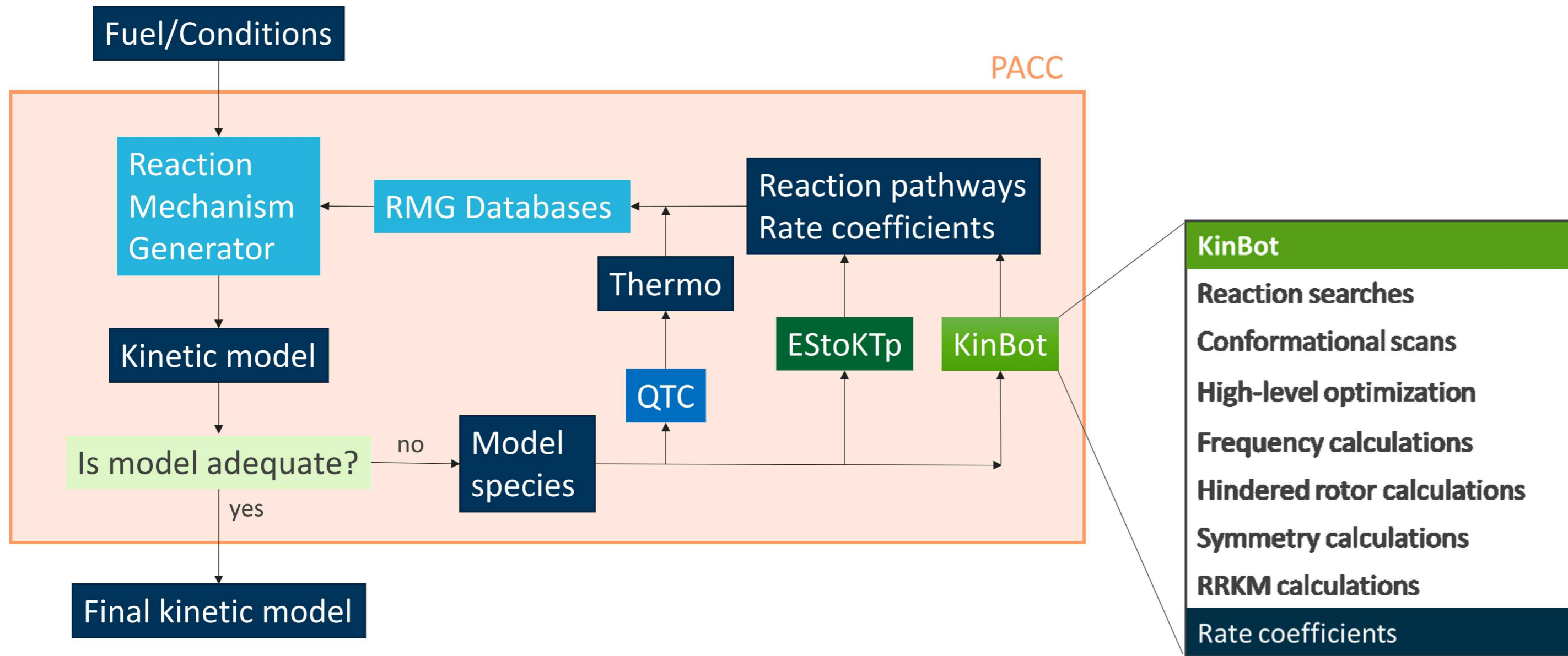
We discovered 75 reaction pathways for a simple molecule in a combined study. KinBot found 48 and was the cheapest method.



Method	FSM			GSM			SSM		SC-AFIR			KinBot First run (Generic)		KinBot Second run (Expanded)				
Total number of channels tested (N_{channel})	562			562			562		74 ^l			200		443				
Total number of gradient calls (excluding IRC)	65420 ^a			756227 ^a			649589 ^a		N/A			8046 ^a (10093 ^b)		23458 ^a (25333 ^c)				
Total number of detected valid (i.e., reactive) saddle points (N_{detected})	153	27% ^d	-	215	38% ^d	-	371	66% ^d	-	17	23% ^d	-	83	42% ^d	-	162	37% ^d	-
<ul style="list-style-type: none"> Unique saddle points Exact duplicates^g Equivalent hydrogens^h 	39	7% ^d	26% ^e	46	8% ^d	21% ^e	50	9% ^d	14% ^e	7	10% ^d	41% ^e	32	16% ^d	39% ^e	48	11% ^d	30% ^e
<ul style="list-style-type: none"> Intended reactions^l Unintended reactions 	19	3% ^d	12% ^e	24	4% ^d	11% ^e	27	5% ^d	7% ^e	14	19% ^d	82% ^e	-	-	-	-	-	-
Wrong reactant	85	15% ^d	-	39	7% ^d	-	23	4% ^d	-	4	5% ^d	-	13	6% ^d	-	22	5% ^d	-
Conformational saddle points	3	0.5% ^d	-	13	2% ^d	-	0	0% ^d	-	53	72% ^d	-	48	24% ^d	-	69	16% ^d	-
Crashed searches (N_{crashes})	321	57% ^d	-	295	52%	-	168	30% ^d	-	-	-	-	56 ^m	28% ^d	-	190 ^m	43% ^d	-
<ul style="list-style-type: none"> Berny optimization failure Gradient failure in string method^j Max number of iterations reached Dissociative channel^k 	230	41% ^d	72% ^f	36	6% ^d	12% ^f	20	4% ^d	12% ^f	-	-	-	-	-	-	-	-	-
	91	16% ^d	28% ^f	88	16% ^d	30% ^f	124	22% ^d	74% ^f	-	-	-	-	-	-	-	-	-
	0	0% ^d	0% ^f	171	30% ^d	58% ^f	18	3% ^d	11% ^f	-	-	-	-	-	-	-	-	-
	0	0% ^d	0% ^f	6	0% ^d	0% ^f	6	1% ^d	4% ^f	-	-	-	-	-	-	-	-	-



KinBot in Predictive Automated Combustion Chemistry



Argonne National Laboratories (Stephen J. Klippenstein)

Politecnico di Milano (prof. Carlo Cavallotti)

Massachusetts Institute of Technology (prof. William H. Green)

Sandia National Laboratories (Judit Zádor)

Goal: Using exascale computers create a whole reaction mechanism automatically



KinBot for propane combustion

Propane combustion model from RMG: 11589 reactions
between 320 species

b3lyp/6-31G reaction searches in Gaussian

Note: no bimolecular reactions nor barrierless reactions
considered

20 Species without any reaction

36 Species that do not exist

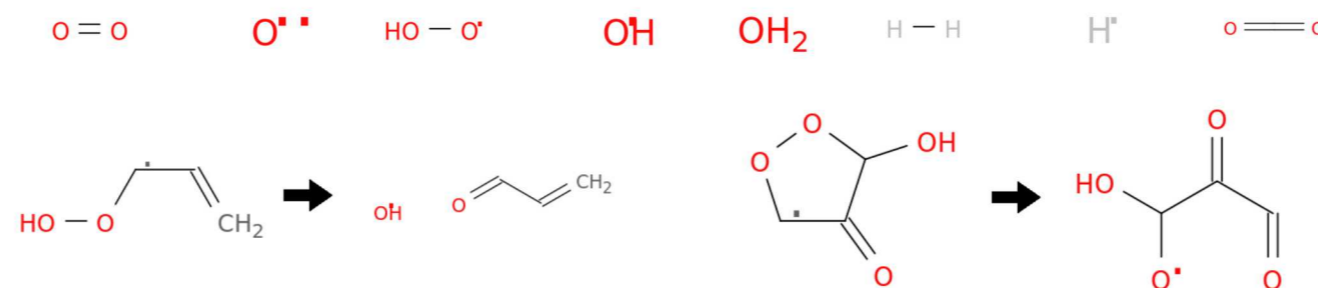
567 Reactions reported by RMG

1332 Reactions found by KinBot

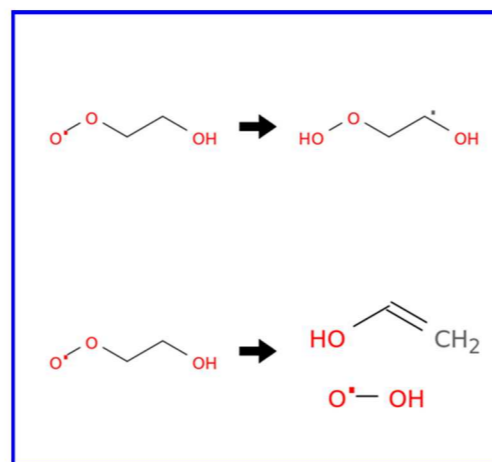
330 Reactions in KinBot and RMG

1002 “New” reactions

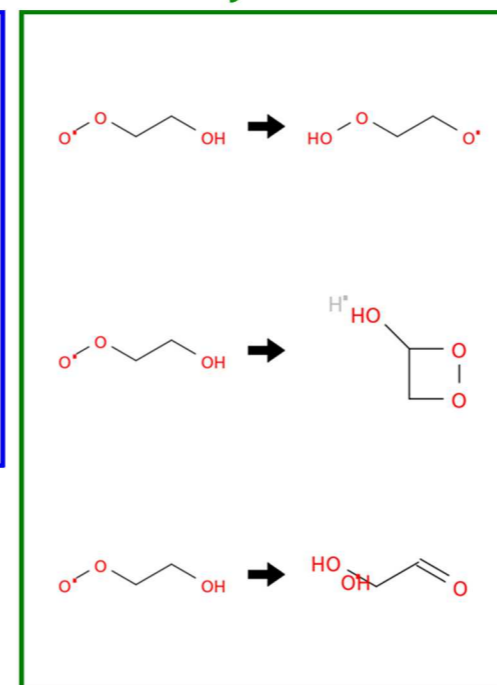
237 Reactions missed by KinBot



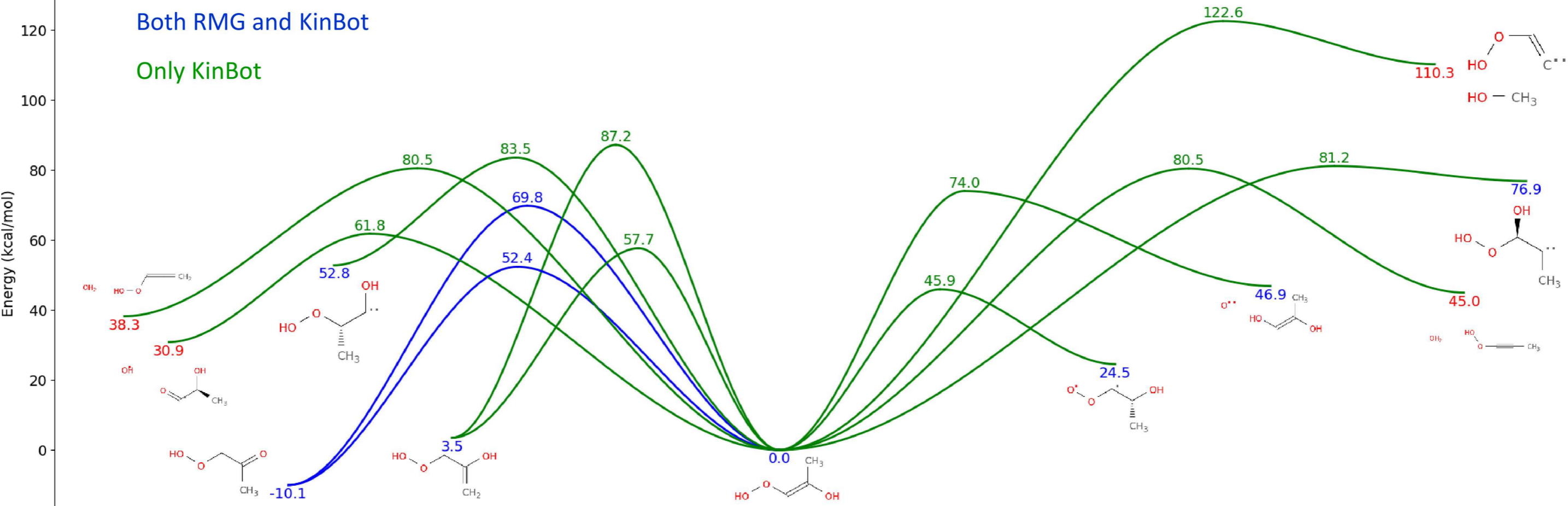
Both RMG and KinBot



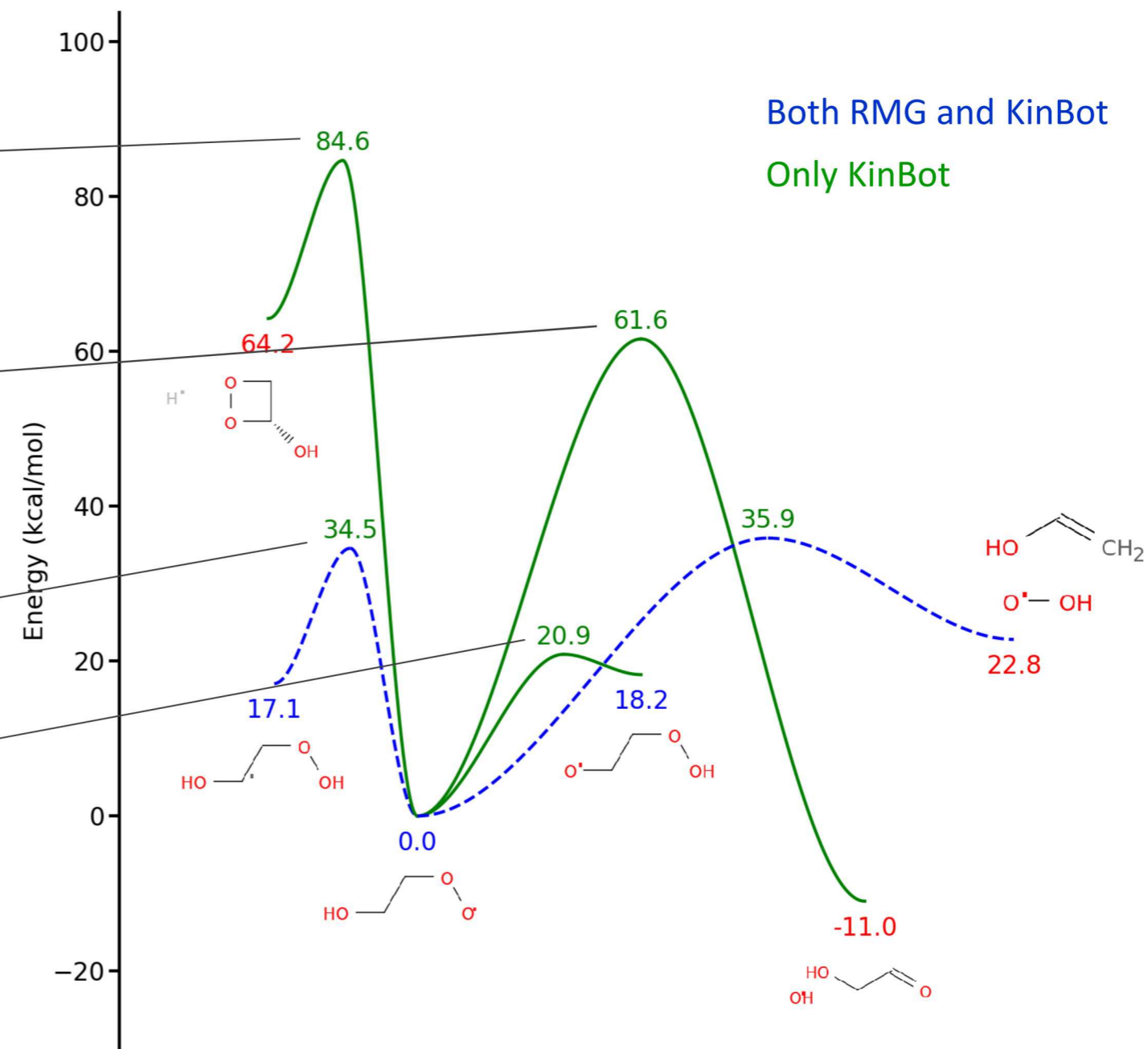
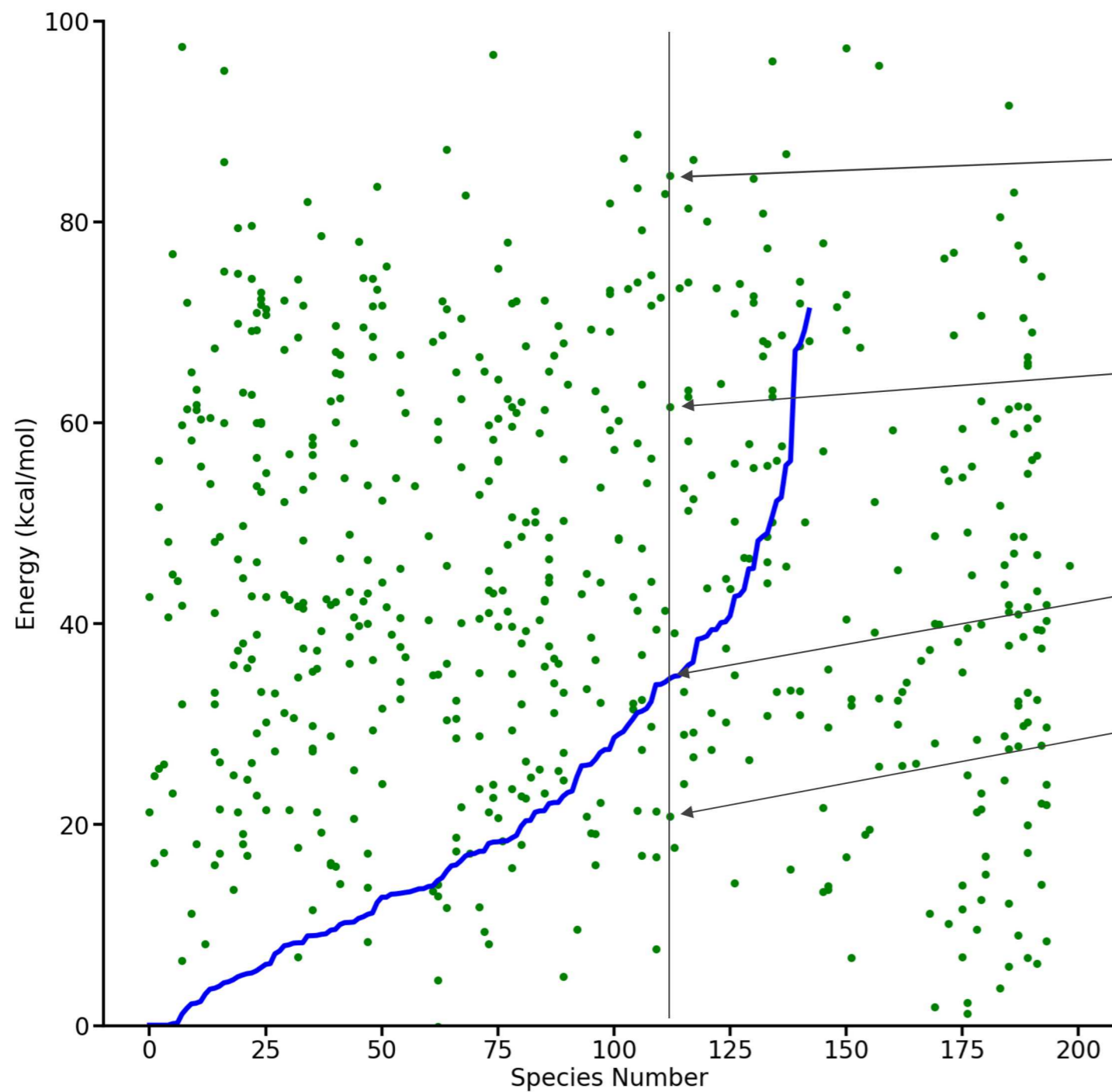
Only KinBot



KinBot for propane combustion



KinBot for propane combustion



Outline

1. Why do we care about chemical kinetics?
2. Chemical kinetics of complex gas-phase reactions
3. KinBot
4. **Sella**



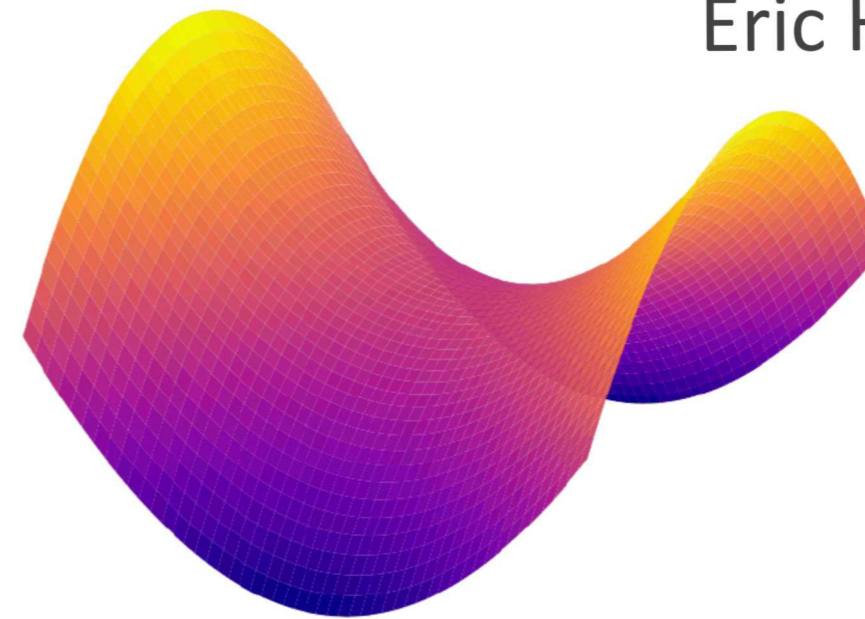
First order saddle points

If \mathbf{q} is a first order saddle point, then $|\mathbf{g}(\mathbf{q})| = 0$, and the Hessian (second derivative matrix) \mathbf{H} has exactly one negative eigenvalue.

While it is possible to optimize minima by following the negative gradient, finding first order saddle points is much more difficult, as it requires information about \mathbf{H} .



Eric Hermes



Local saddle point optimization

There are many existing methods for finding saddle points:

- Double-ended string methods (e.g. NEB)¹
- (Quasi-)Newton (QN) methods:
 - Rational function optimization² (RFO) and its variants
 - Berny's method³
- Gentlest ascent (GA) methods
 - Dimer⁴ and its variants
 - Activation relaxation technique⁵ (ART) and its variants

¹Jónsson, H.; Mills, G.; Jacobsen, K. W. *Classical and Quantum Dynamics in Condensed Phase Simulations*; WORLD SCIENTIFIC, 1998; pp 385–404

²Culot, P.; Dive, G.; Nguyen, V. H.; Ghuysen, J. M. *Theoretica chimica acta* **1992**, *82*, 189–205

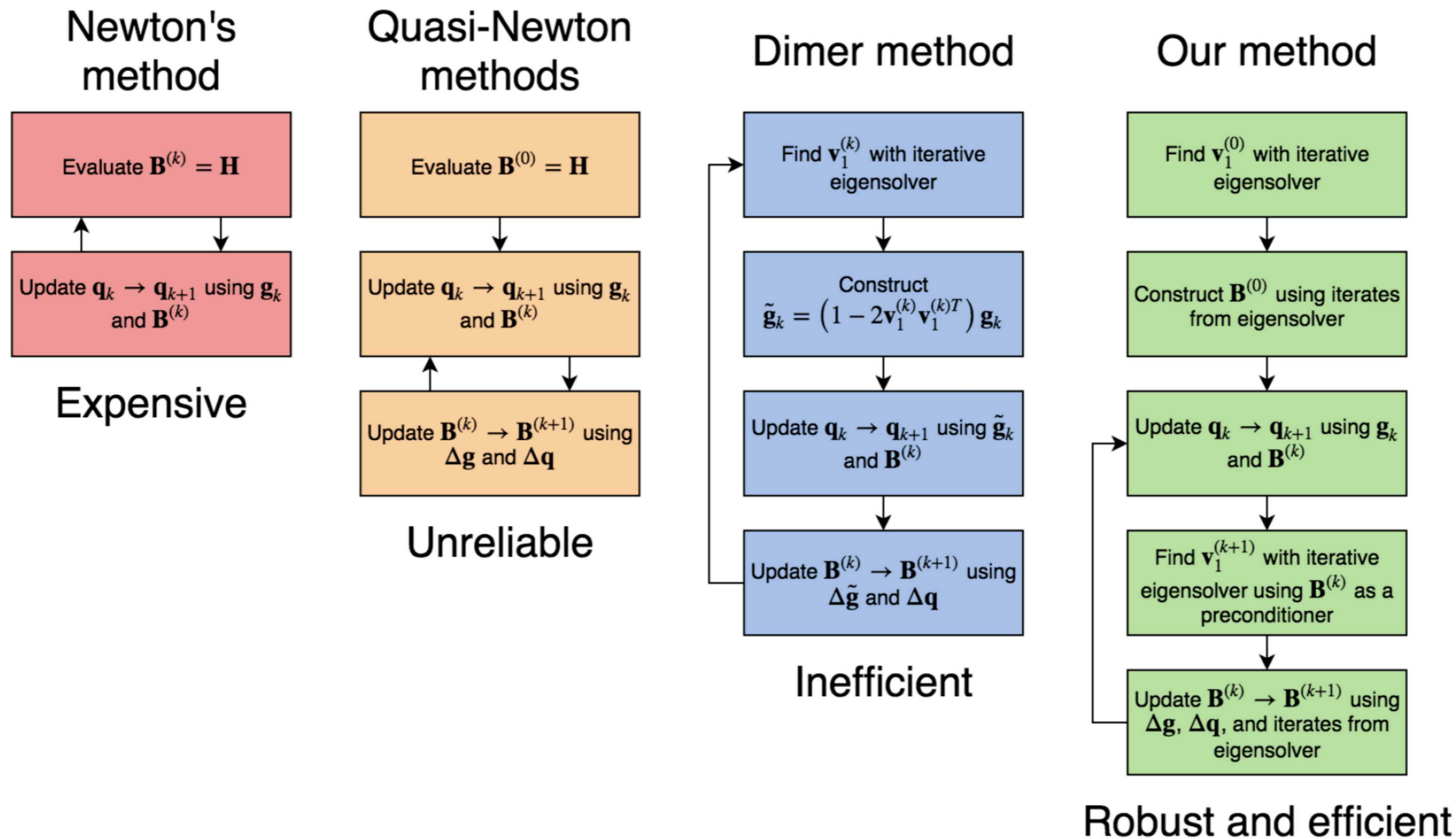
³Schlegel, H. B. *Journal of Computational Chemistry* **1982**, *3*, 214–218

⁴Henkelman, G.; Jónsson, H. *The Journal of Chemical Physics* **1999**, *111*, 7010–7022

⁵Mousseau, N.; Barkema, G. T. *Physical Review E* **1998**, *57*, 2419–2424

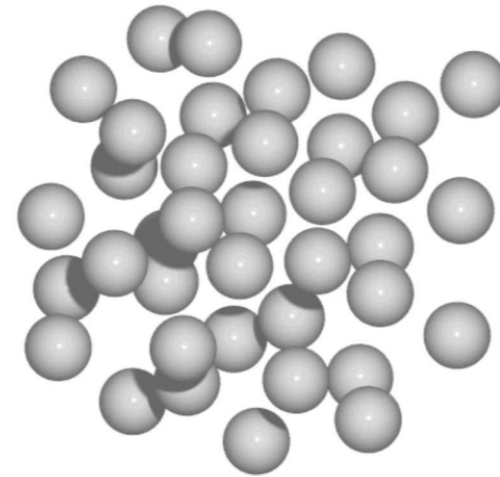


Method comparison



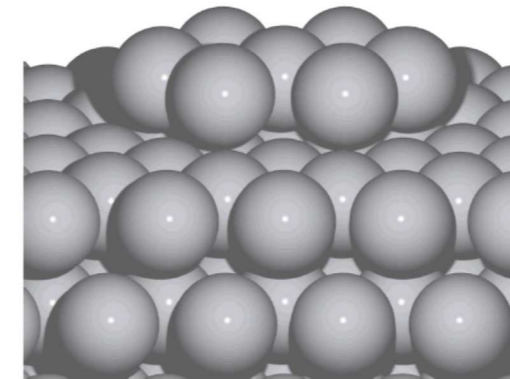
The approximate Hessian is used both as a preconditioner for the eigensolver and for QN steps.





Algorithm	<FCs>	min(FCs)	median(FCs)	max(FCs)	nfailed
optim_ffd	145	57	119	565	0
pele_ffd	192	59	150	1488	0
eon-lanczos-improved ¹	237	65	181	1898	0
eon-lanczos ¹	321	83	220	8857	0
eon-dimer ¹	528	92	388	3581	0

Show/Hide Additional Entries...



Algorithm	<FCs>	min(FCs)	median(FCs)	max(FCs)	nfailed
optim_ffd	71	43	70	143	0
pele	88	52	78	198	0
optim_cd	90	52	86	199	0
eon-lanczos-improved ¹	106	71	100	163	0
eon-dimer ¹	116	83	114	160	0
eon-lanczos ¹	143	106	139	292	0

Show/Hide Additional Entries...

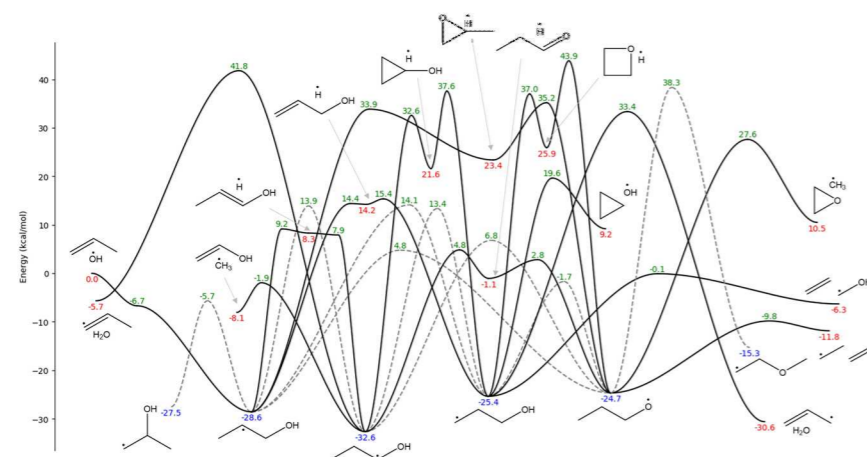
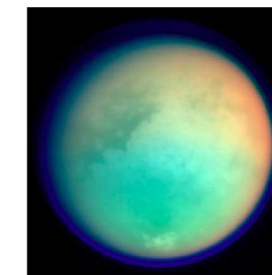
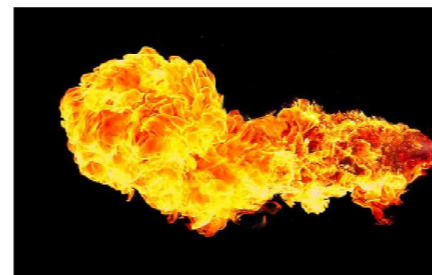
Test set	<FCs>	min(FCs)	median(FCs)	max(FCs)	nfailed
LJ38	92	37	86	205	0
Pt(111)	61	39	57	107	0

¹⁰Chill, S. T.; Stevenson, J.; Ruehle, V.; Shang, C.; Xiao, P.; Farrell, J. D.; Wales, D. J.; Henkelman, G. *Journal of Chemical Theory and Computation* **2014**, *10*, 5476–5482



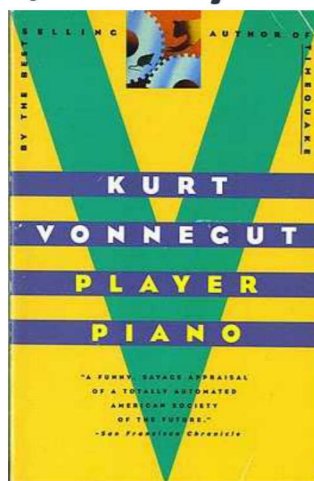
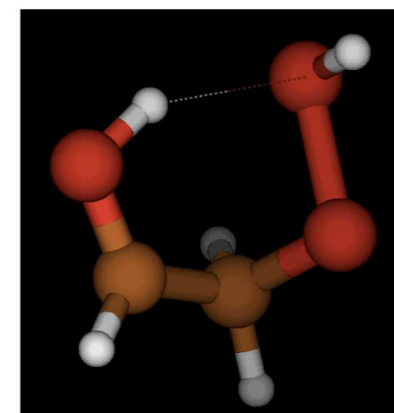
Perspective

To meet the challenges of fuel development, air pollution control, automation is inevitable on all levels.



Automation is possible in elementary kinetics calculations, mainly to speed up research, and generalize knowledge.

Fundamentally new pathways often involve difficult electronic structure problems, very hard to automate blindly.



Discoveries will continue to be made by (human) experimentalists, theorists and modelers.

