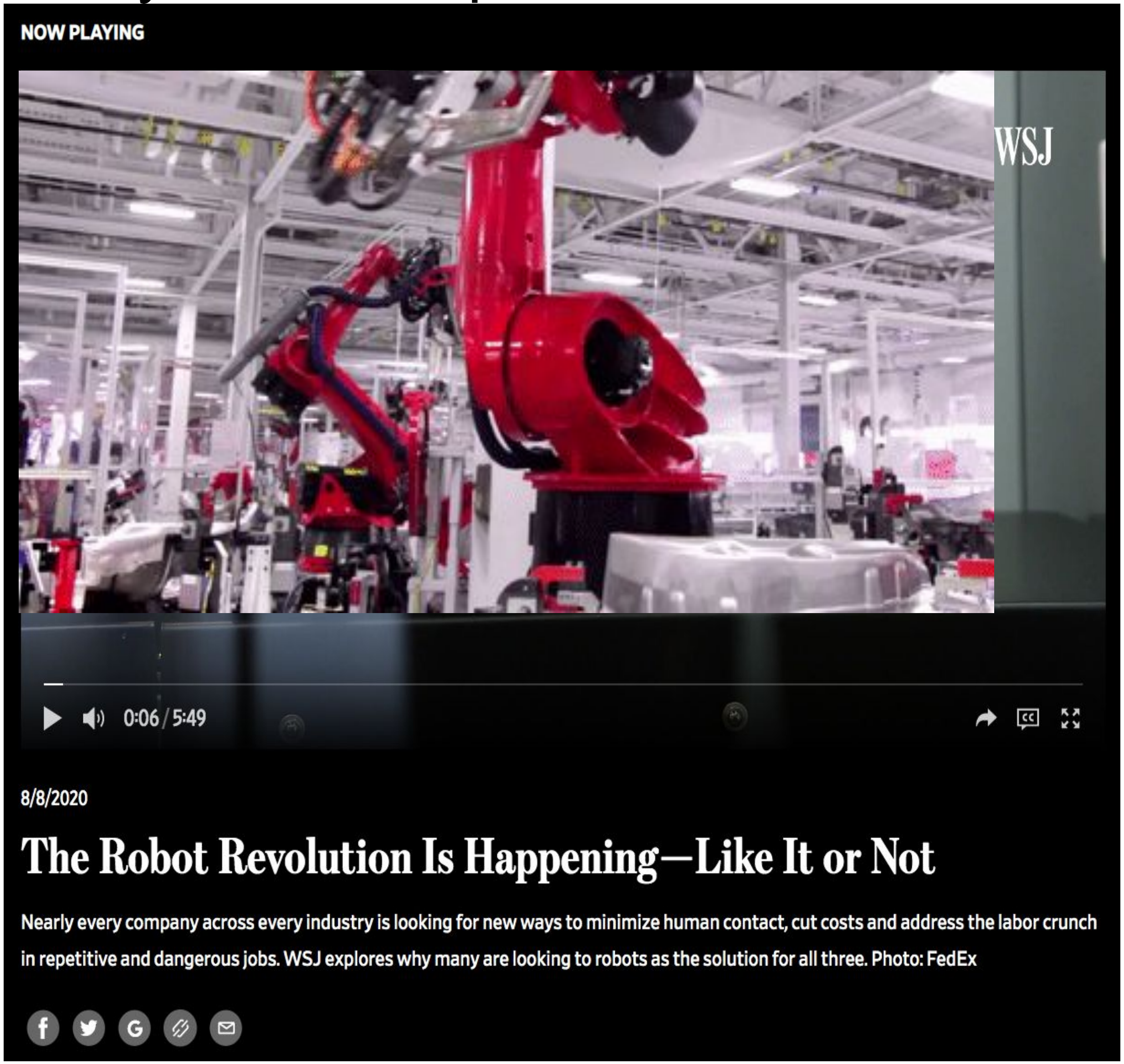
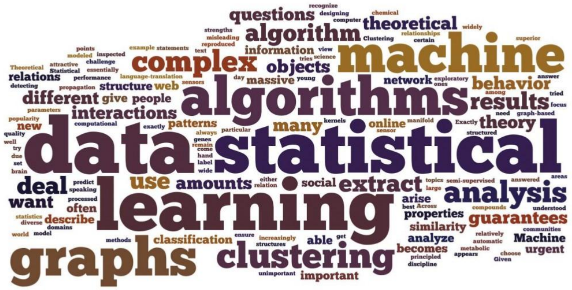
Accelerate? Yes but how?

Physics + computer simulation +



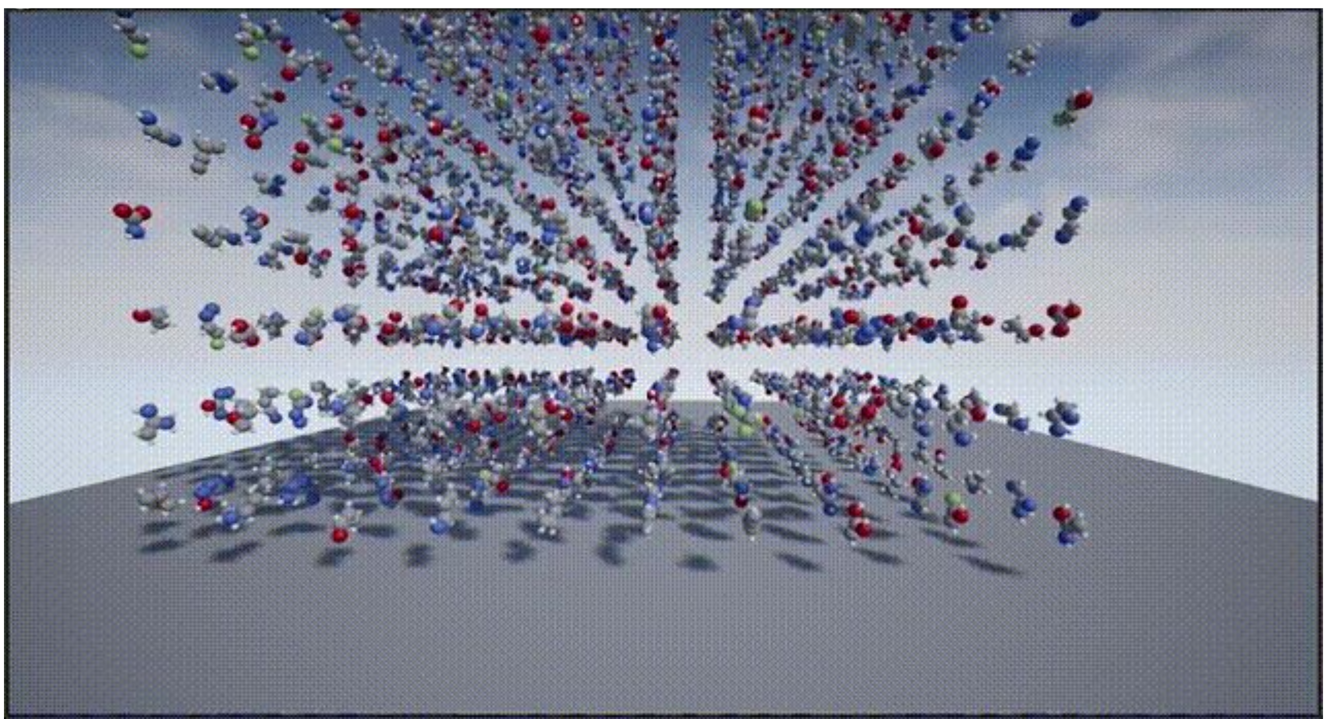
+

+

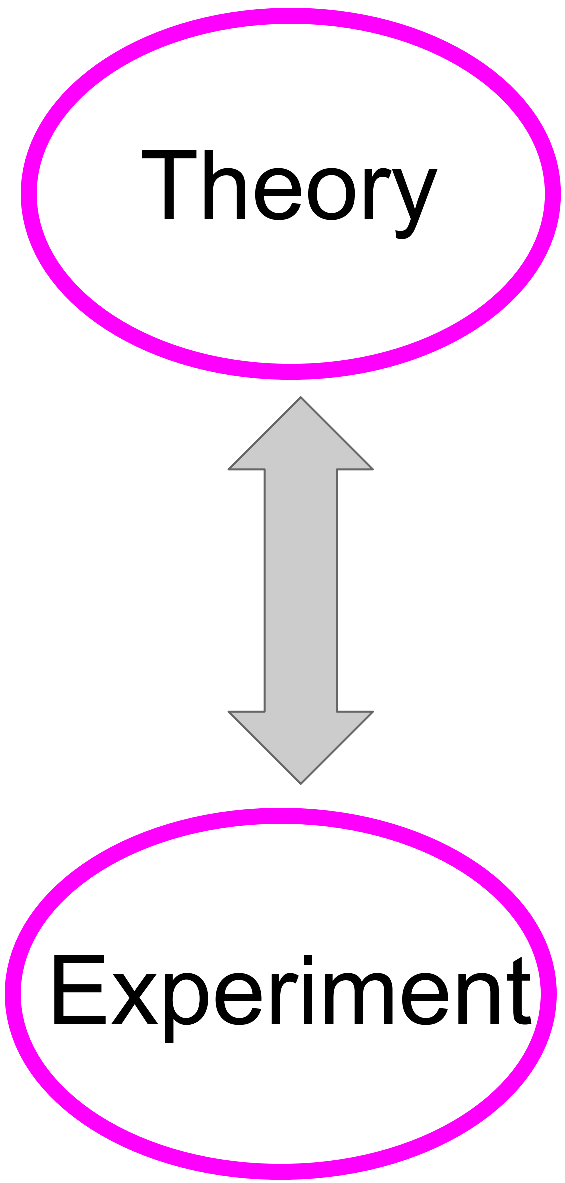


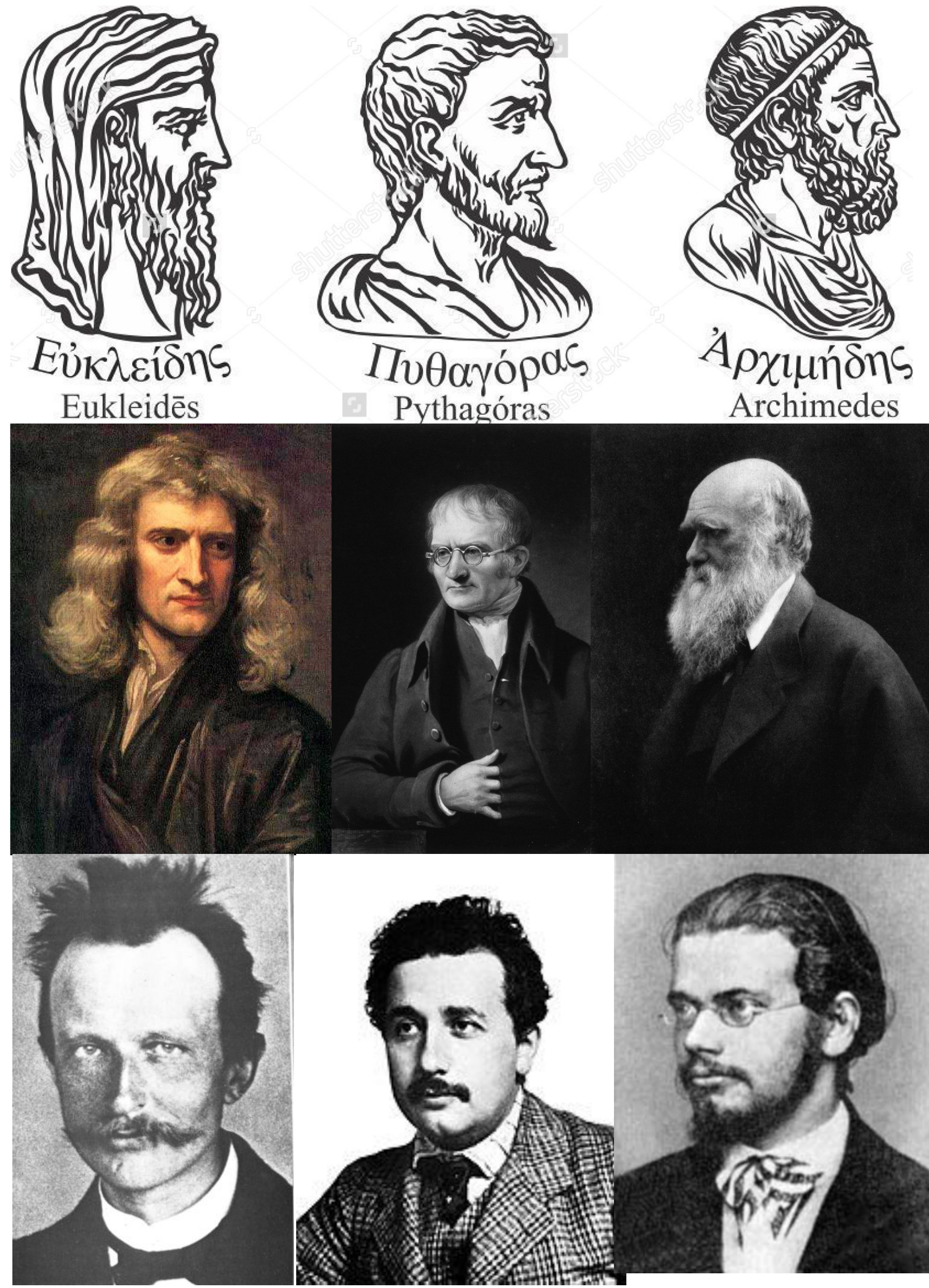
Quantum Machine Learning

‘Quantum machine learning using atom-in-molecule-based fragments selected on the fly’, Huang, von Lilienfeld, *Nature Chemistry* (2020)

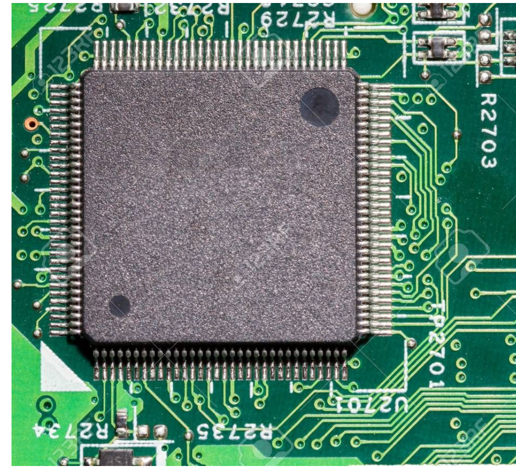


The pillars of science





The pillars of science

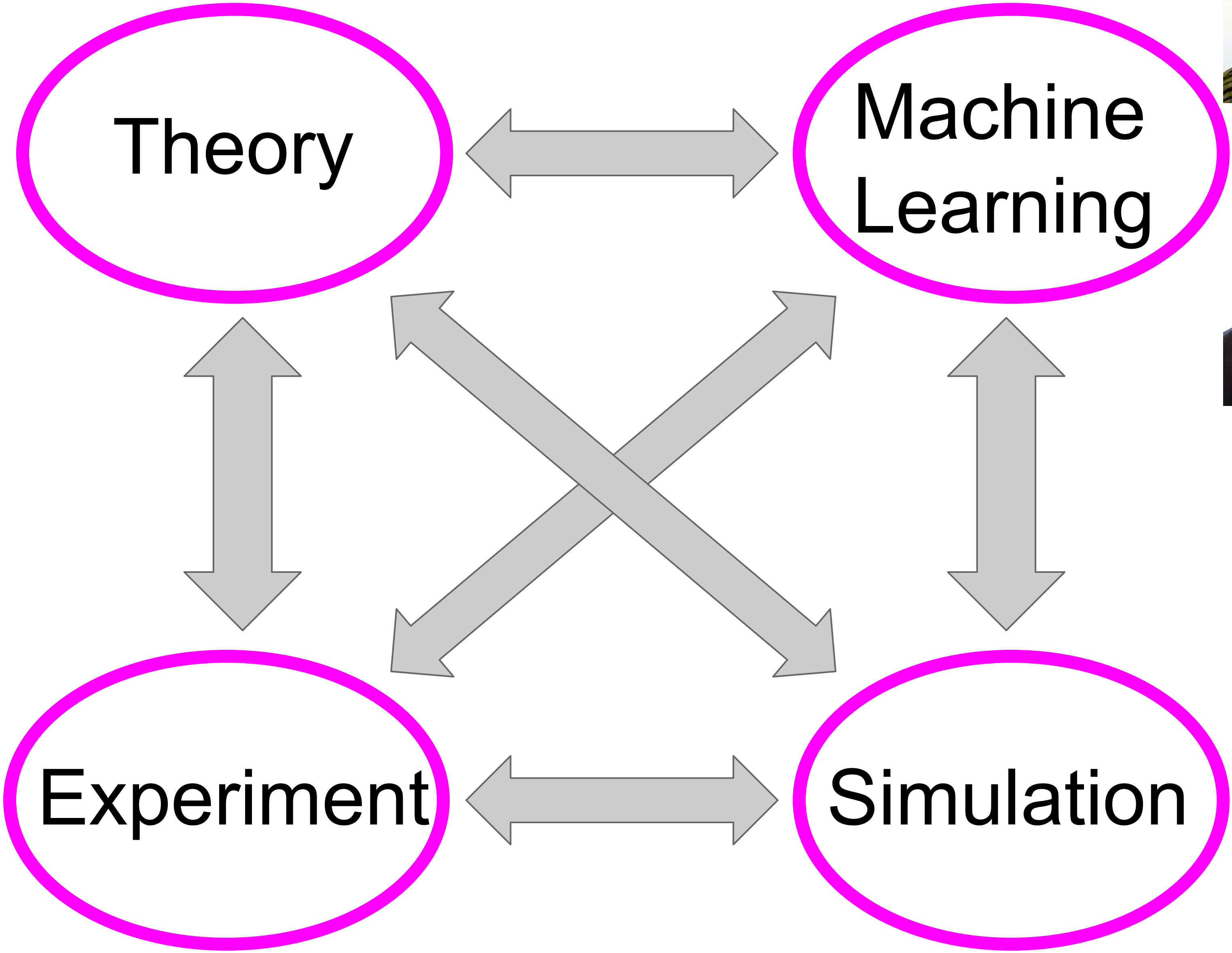


Theory

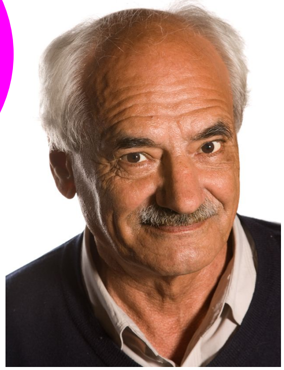
Experiment

Simulation

The pillars of science



Vapnik



Chervonenkis

Stoichiometry? → Mendeleev's table, est. in 1879: Well alive and kicking

hydrogen 1 H 1.0079																		helium 2 He 4.0026
lithium 3 Li 6.941	beryllium 4 Be 9.0122												boron 5 B 10.811	carbon 6 C 12.011	nitrogen 7 N 14.007	oxygen 8 O 15.999	fluorine 9 F 18.998	neon 10 Ne 20.180
sodium 11 Na 22.990	magnesium 12 Mg 24.305												aluminium 13 Al 26.982	silicon 14 Si 28.086	phosphorus 15 P 30.974	sulfur 16 S 32.065	chlorine 17 Cl 35.453	argon 18 Ar 39.948
potassium 19 K	calcium 20 Ca		scandium 21 Sc	titanium 22 Ti	vanadium 23 V	chromium 24 Cr	manganese 25 Mn	iron 26 Fe 55.845	cobalt 27 Co 58.933	nickel 28 Ni 58.693	copper 29 Cu 63.546	zinc 30 Zn 65.39	gallium 31 Ga 69.723	germanium 32 Ge 72.61	arsenic 33 As 74.922	selenium 34 Se 78.96	bromine 35 Br 79.904	krypton 36 Kr 83.80
								ruthenium 44 Ru 101.07	rhodium 45 Rh 102.91	palladium 46 Pd 106.42	silver 47 Ag 107.87	cadmium 48 Cd 112.41	indium 49 In 114.82	tin 50 Sn 118.71	antimony 51 Sb 121.76	tellurium 52 Te 127.60	iodine 53 I 126.90	xenon 54 Xe 131.29
								osmium 76 Os 190.23	iridium 77 Ir 192.22	platinum 78 Pt 195.08	gold 79 Au 196.97	mercury 80 Hg 200.59	thallium 81 Tl 204.38	lead 82 Pb 207.2	bismuth 83 Bi 208.98	polonium 84 Po [209]	astatine 85 At [210]	radon 86 Rn [222]
								hassium 108 Hs [269]	meitnerium 109 Mt [268]	ununnilium 110 Uun [271]	unununium 111 Uuu [272]	ununbium 112 Uub [277]		ununquadium 114 Uuq [289]				
								samarium 62 Sm 150.36	europium 63 Eu 151.96	gadolinium 64 Gd 157.25	terbium 65 Tb 158.93	dysprosium 66 Dy 162.50	holmium 67 Ho 164.93	erbium 68 Er 167.26	thulium 69 Tm 168.93	ytterbium 70 Yb 173.04		
								plutonium 94 Pu [244]	americium 95 Am [243]	curium 96 Cm [247]	berkelium 97 Bk [247]	californium 98 Cf [251]	einsteinium 99 Es [252]	fermium 100 Fm [257]	mendelevium 101 Md [258]	nobelium 102 No [259]		

THE NEW YORKER

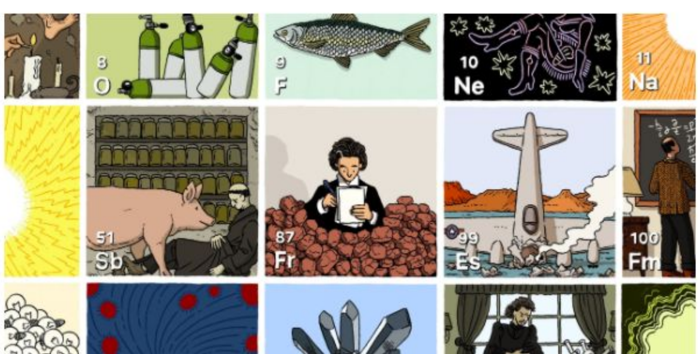
News Books & Culture Fiction & Poetry Humor & Cartoons Magazine Crossword Video Pc

ELEMENTS

THE HISTORIES HIDDEN IN THE PERIODIC TABLE

From poisoned monks and nuclear bombs to the "transfermium wars," mapping the atomic world hasn't been easy.

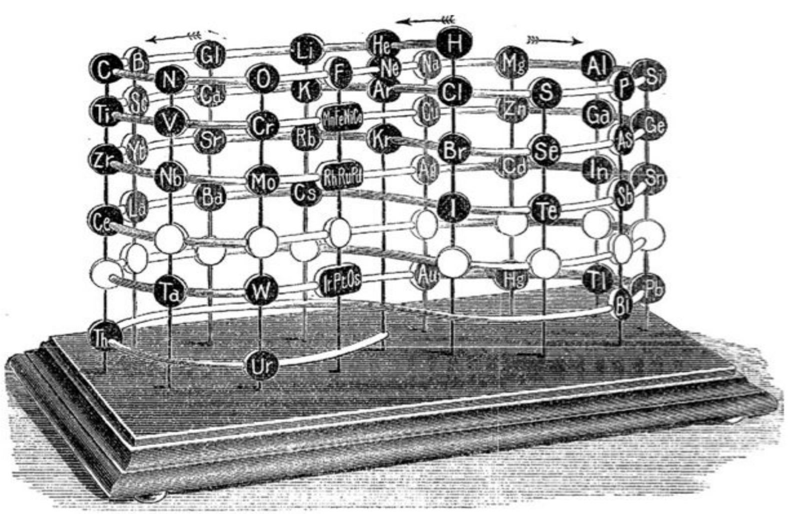
By Neima Jahromi
December 27, 2019

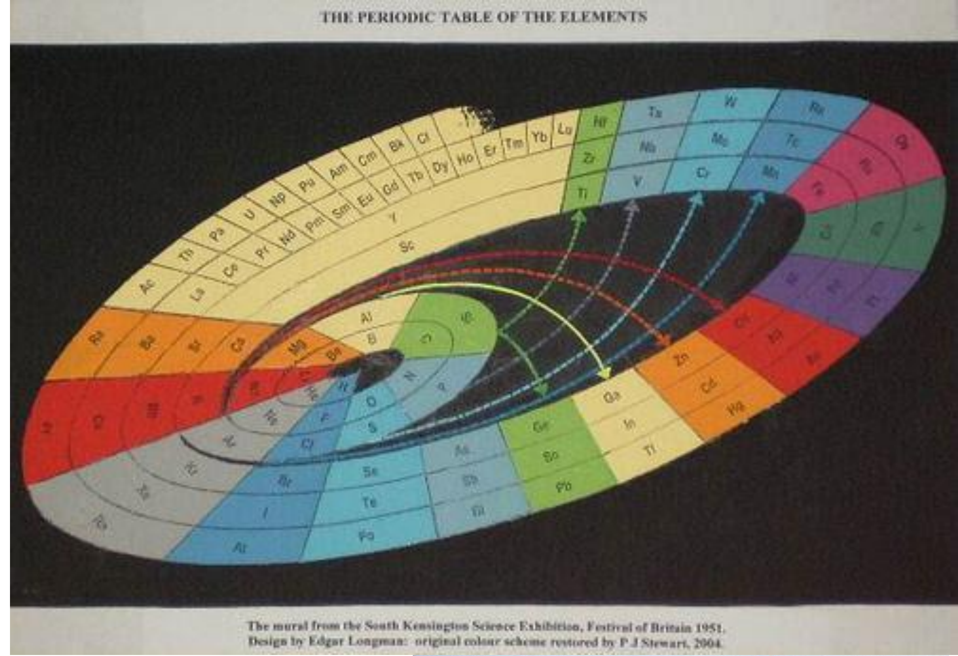


Kahneman (1973))?

Alternatives?

William Crookes (1832-1919)





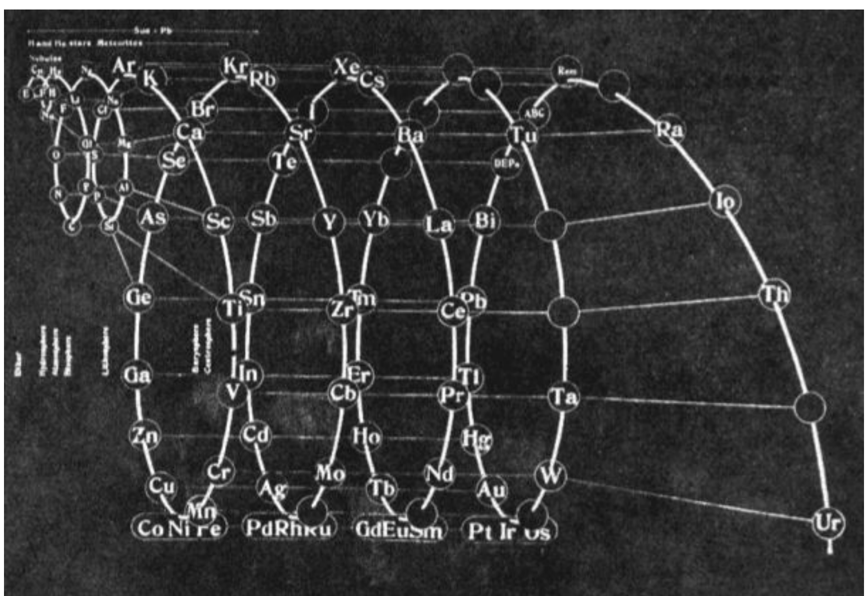
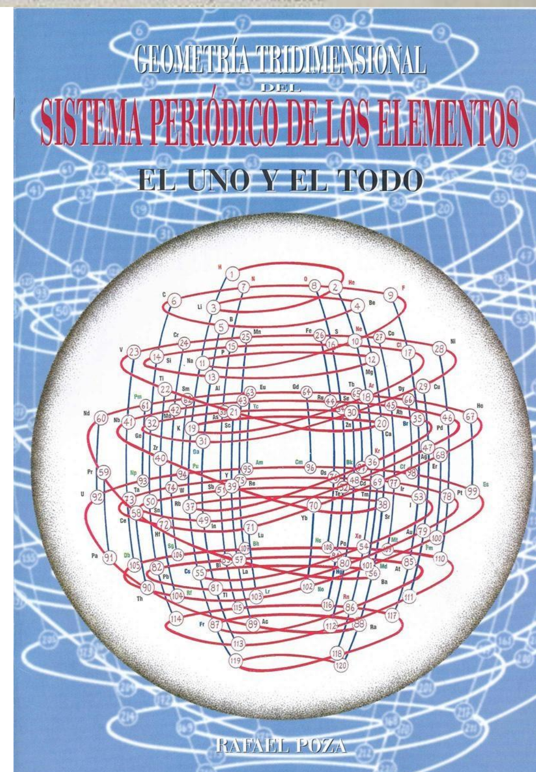
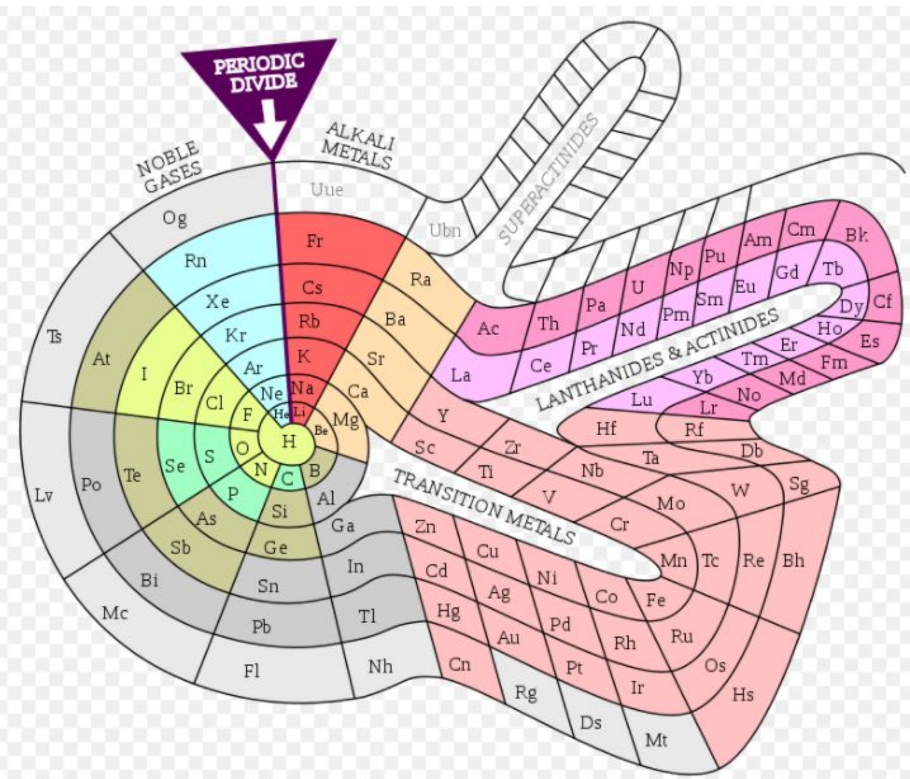


FIGURE 21.—EMERSON'S HELIX



Alternatives?



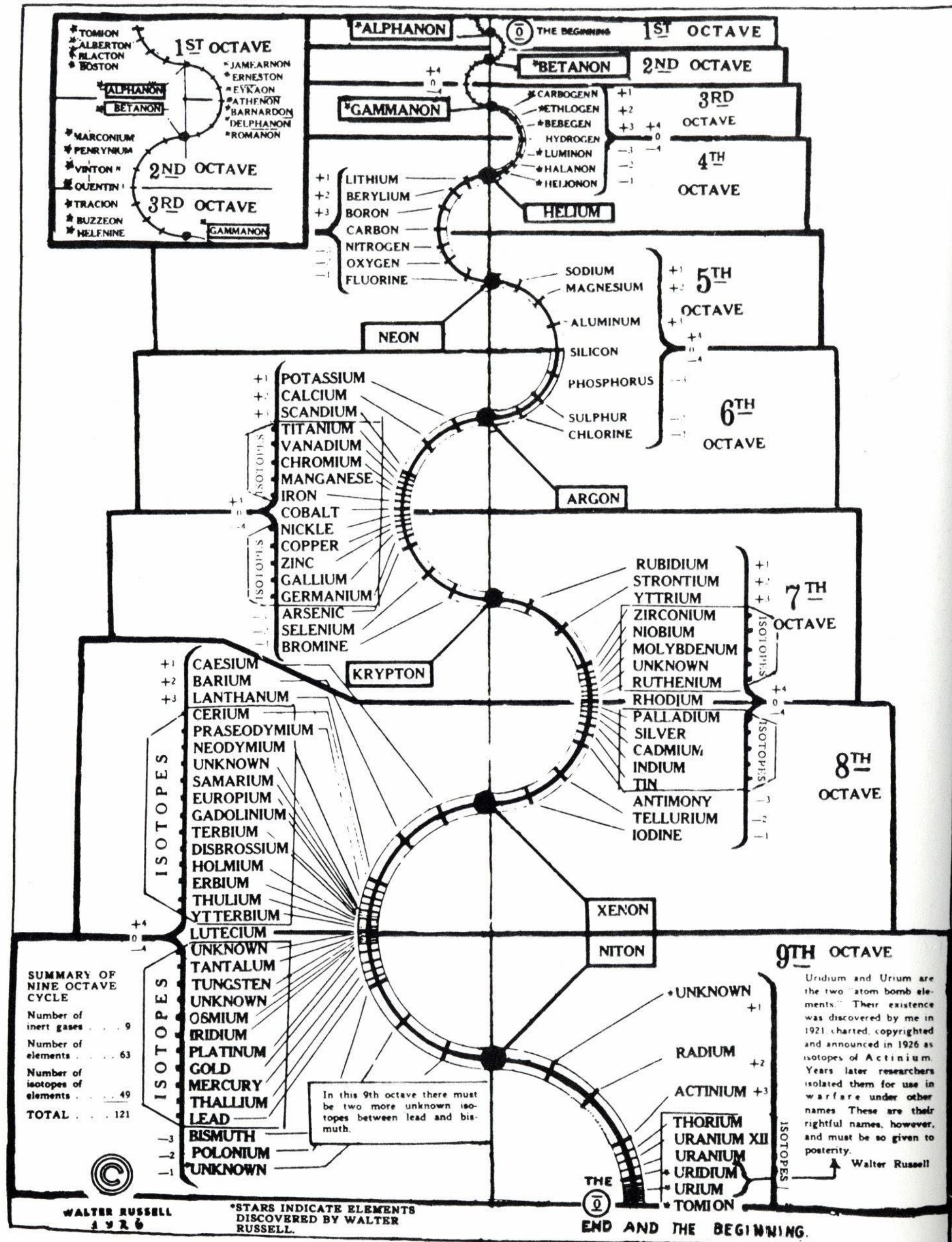
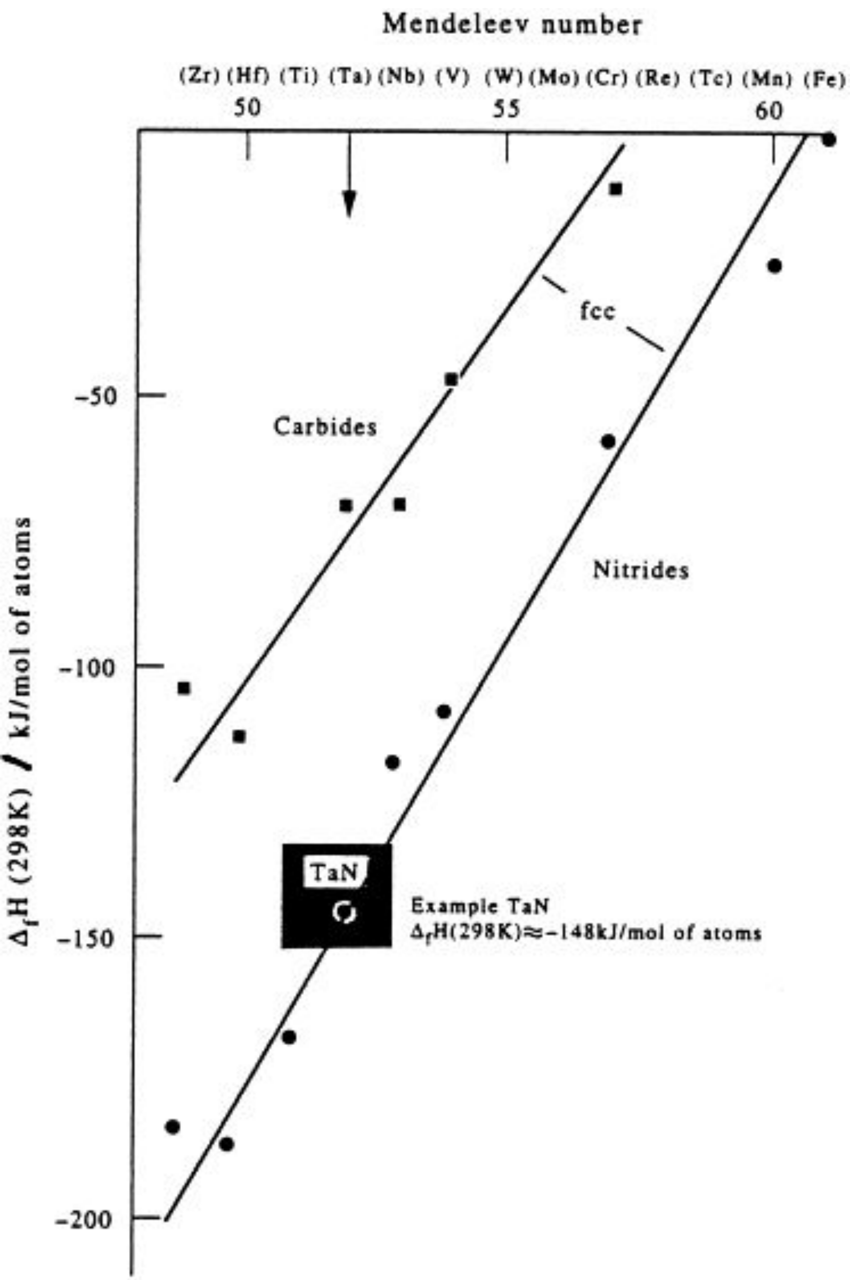


Figure 176. The Russell Periodic Chart of the Elements, No. I

Alternatives?



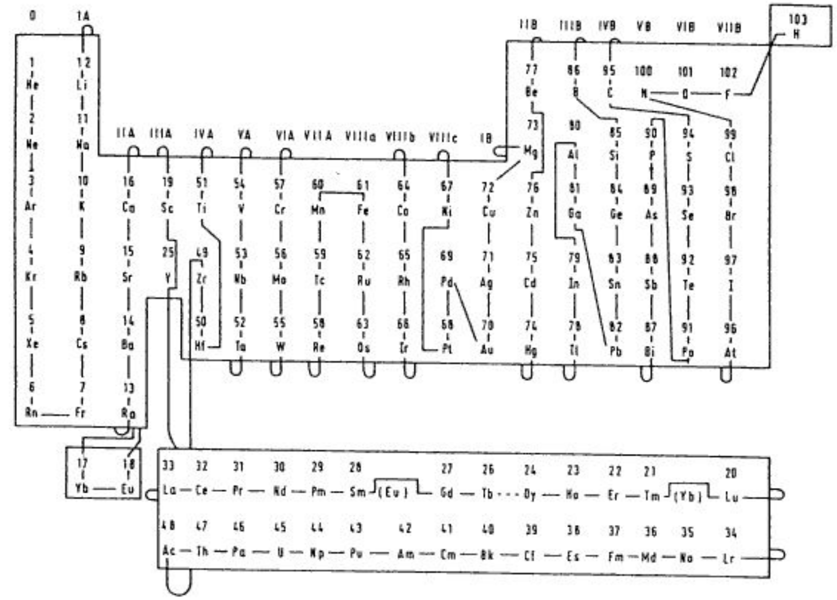
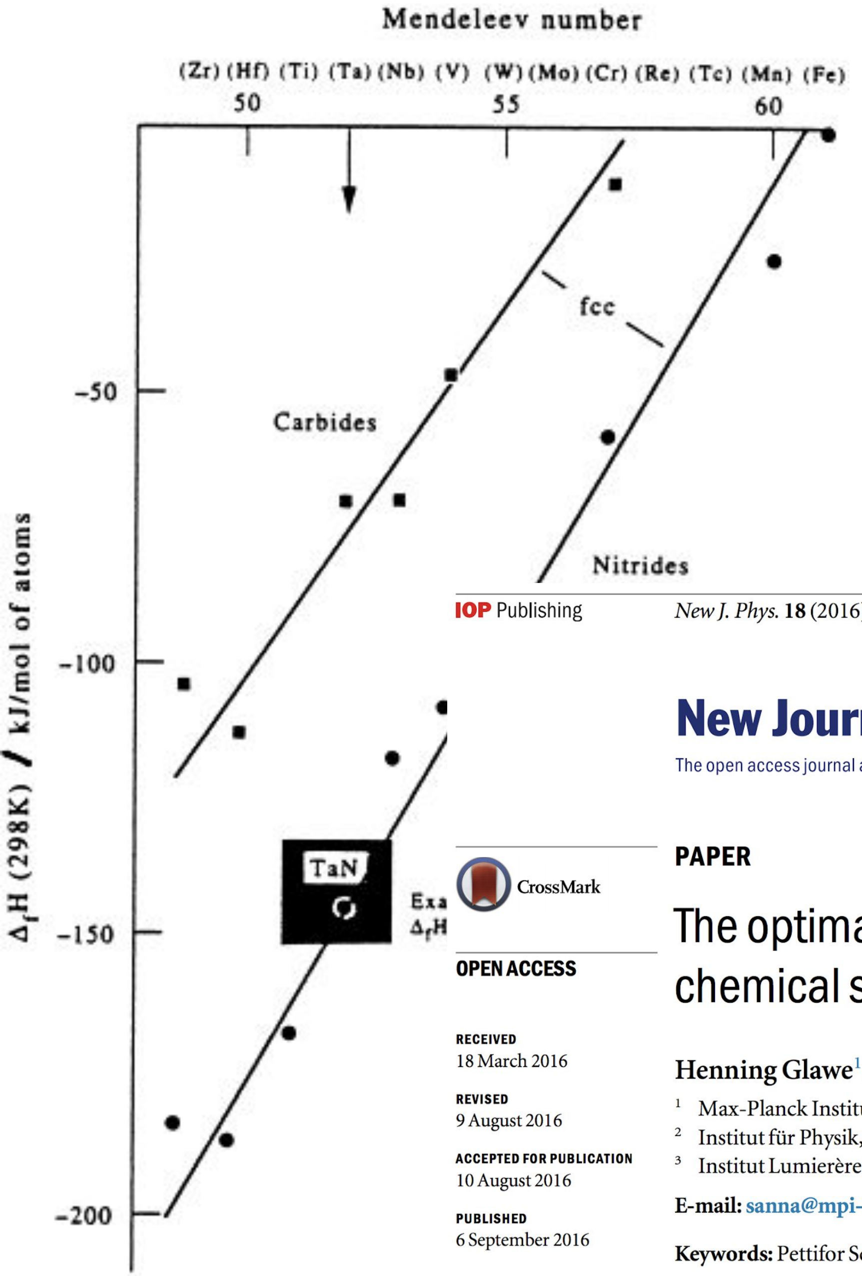


Fig. 1.8 The string running through this modified periodic table puts all the elements in sequential order, given by the relative ordering number $\mathscr{M}$. From Pettifor (1988).

Alternatives?



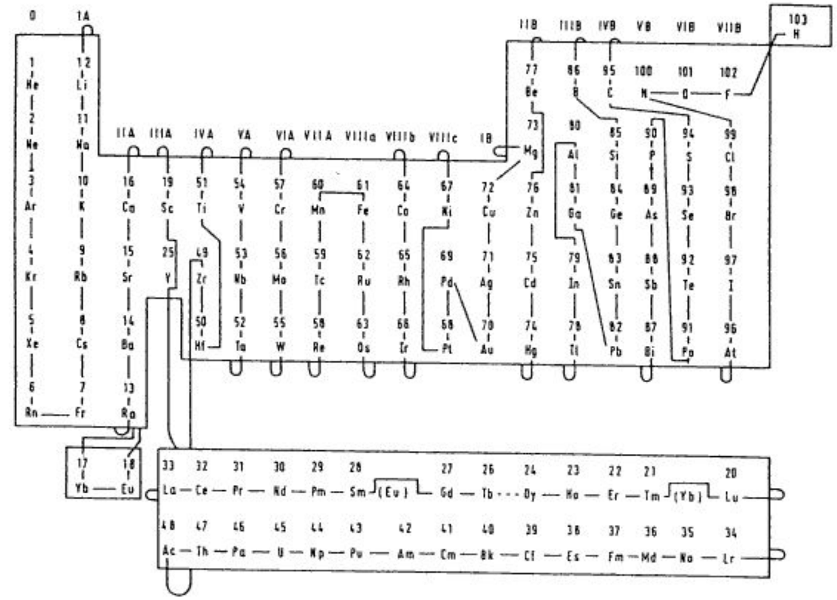


Fig. 1.8 The string running through this modified periodic table puts all the elements in sequential order, given by the relative ordering number $\mathscr{M}$. From Pettifor (1988).

IOP Publishing *New J. Phys.* **18** (2016) 093011 doi:10.1088/1367-2630/18/9/093011

New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
18 March 2016

REVISED
9 August 2016

ACCEPTED FOR PUBLICATION
10 August 2016

PUBLISHED
6 September 2016

PAPER

The optimal one dimensional periodic table: a modified Pettifor chemical scale from data mining

Henning Glawe[1], Antonio Sanna[1], E K U Gross[1] and Miguel A L Marques[2,3]

[1] Max-Planck Institut für Microstrukture Physics, Weinberg 2, 06120 Halle, Germany
[2] Institut für Physik, Martin-Luther-Universität Halle-Wittenberg, D-06099 Halle, Germany
[3] Institut Lumierère Matière, UMR5306 Université Lyon 1-CNRS, Université Lyon, F-69622 Villeurbanne Cedex, France

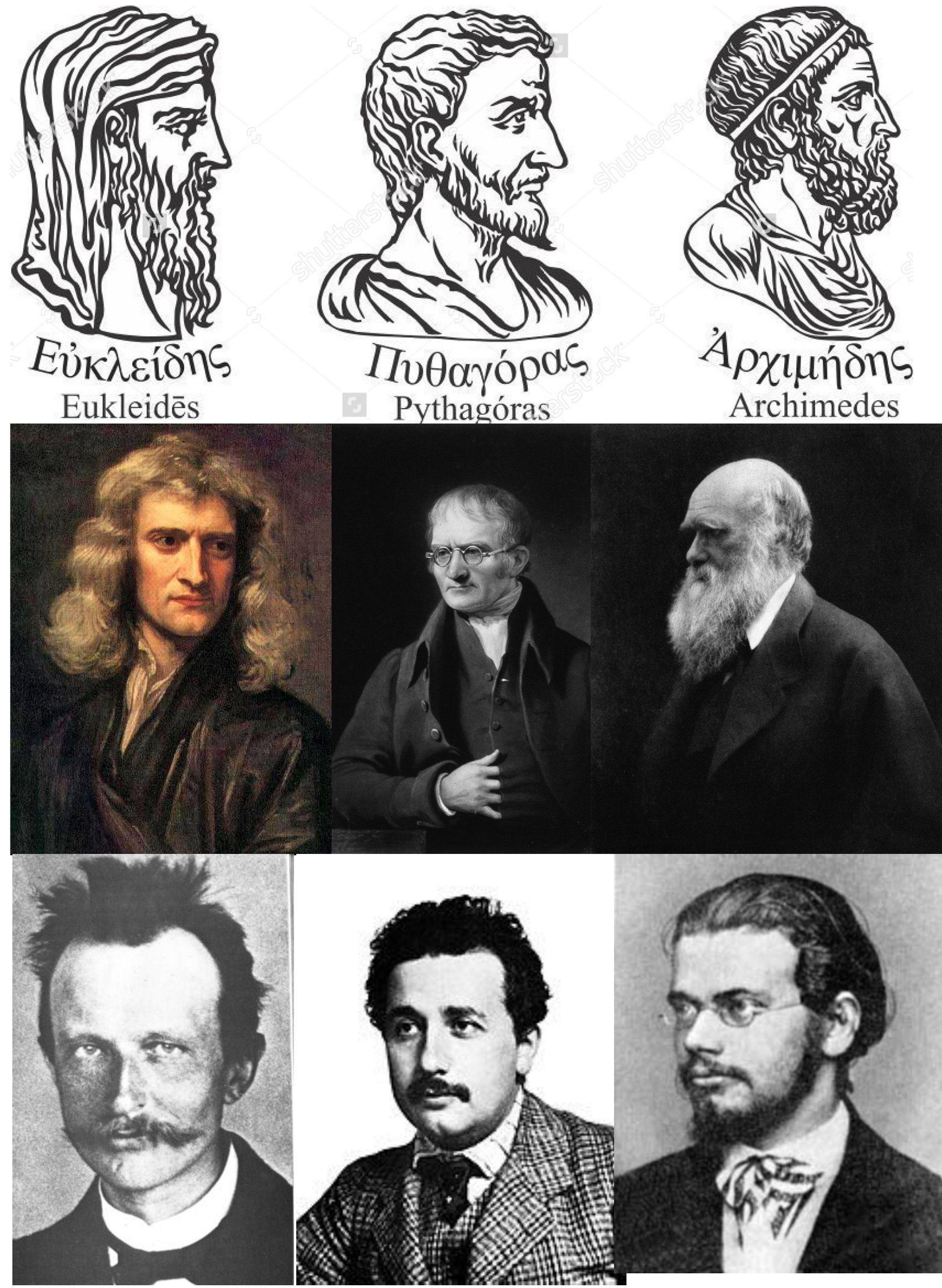
E-mail: sanna@mpi-halle.mpg.de and miguel.marques@physik.uni-halle.de

Keywords: Pettifor Scale, chemical similarity, atomic substitution, crystal structures

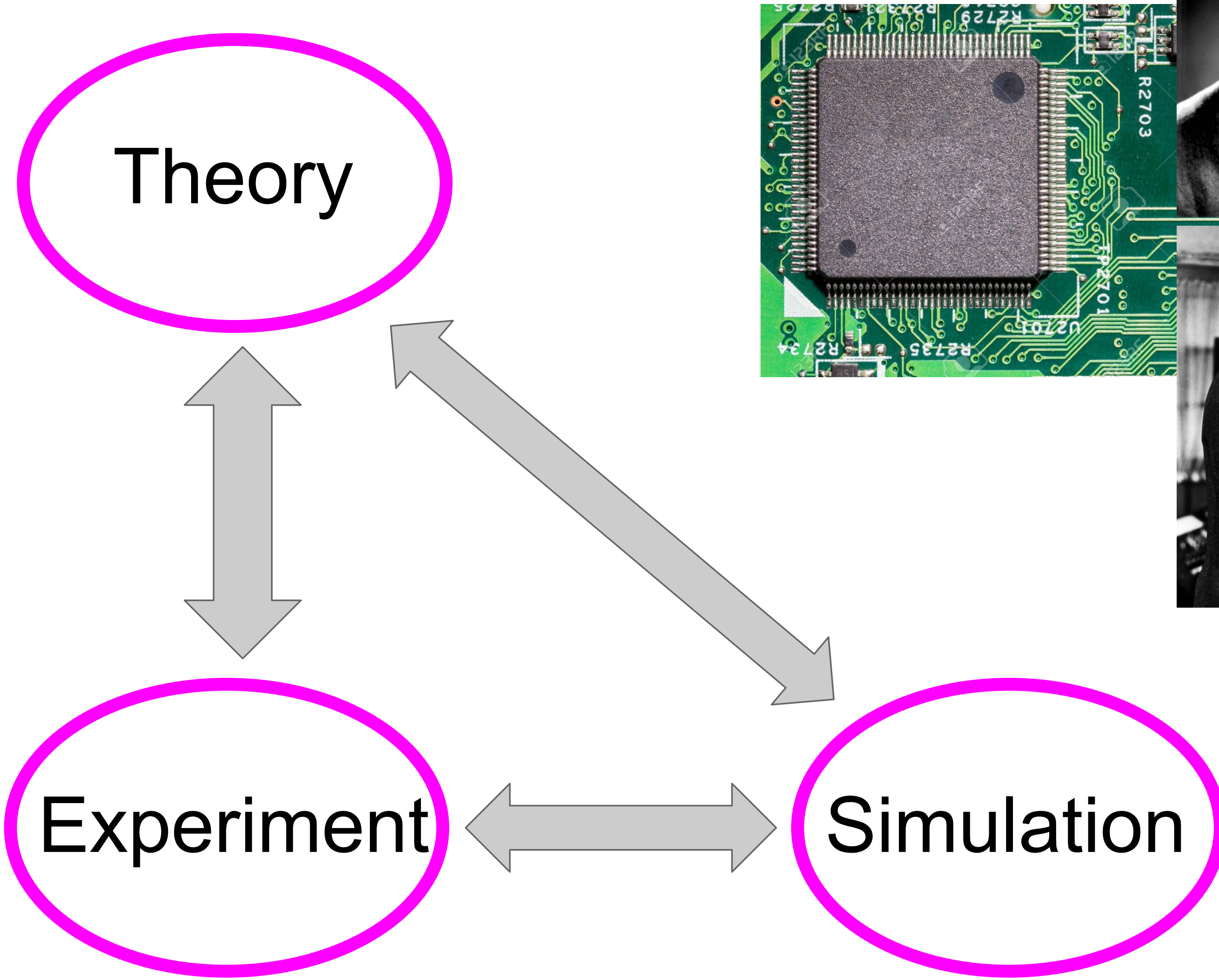
The pillars of science

Theory

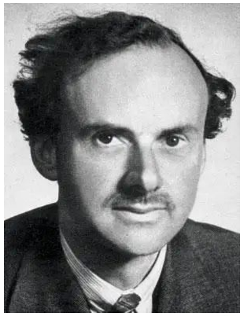
Experiment

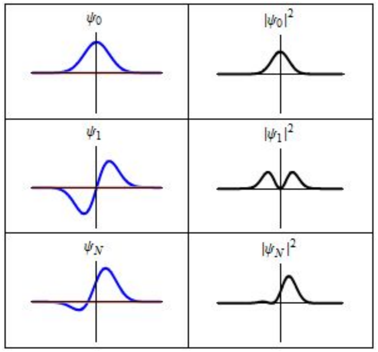


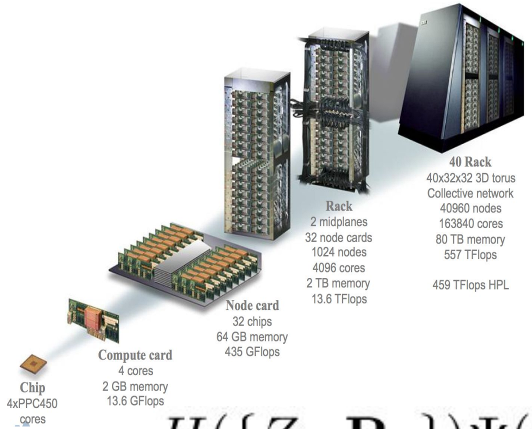
The pillars of science



The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.
Dirac

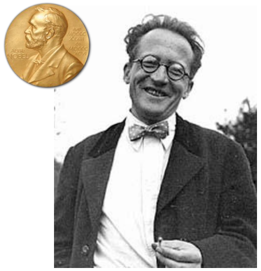






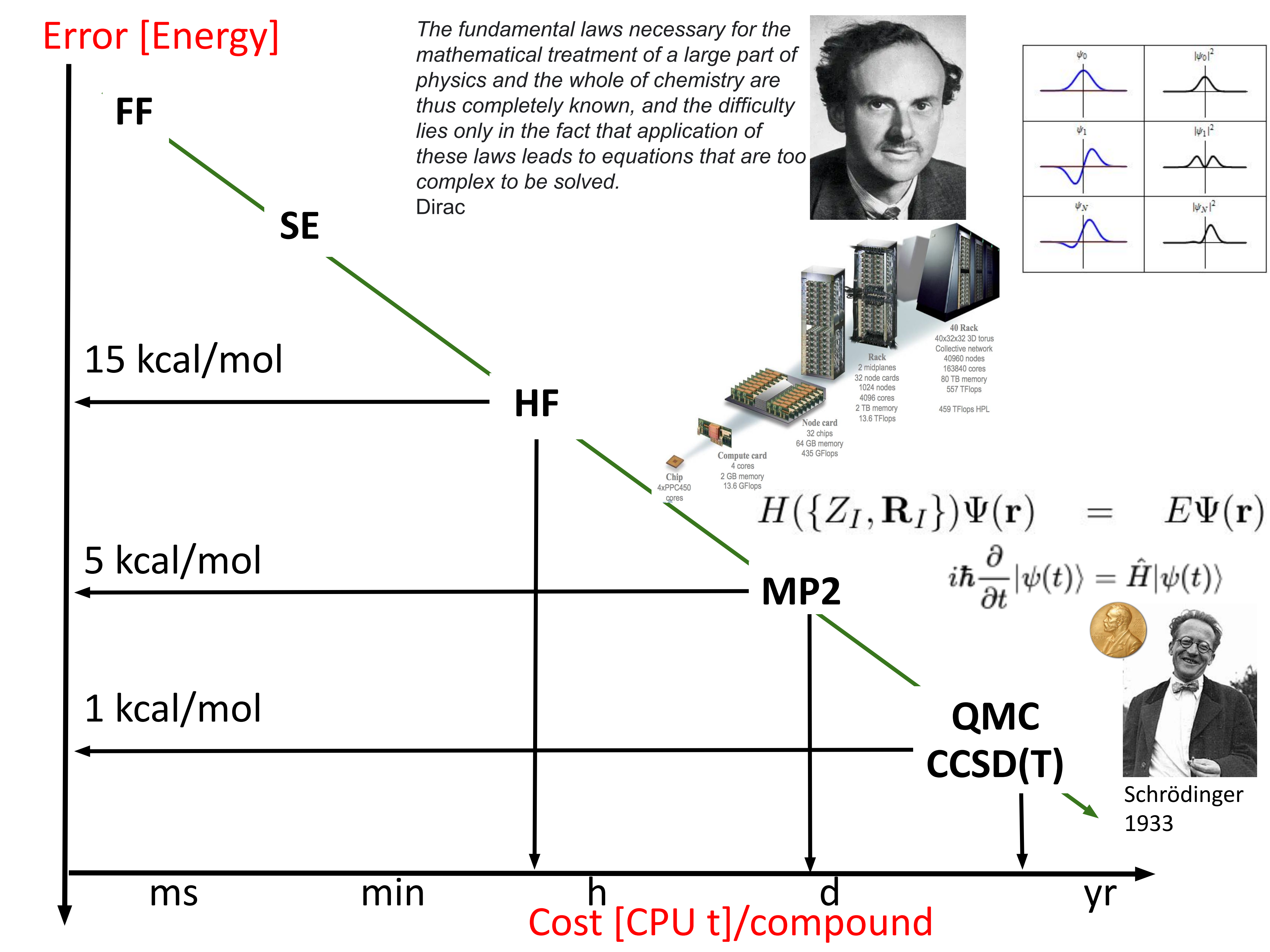
$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) \quad = \quad E\Psi(\mathbf{r})$$

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

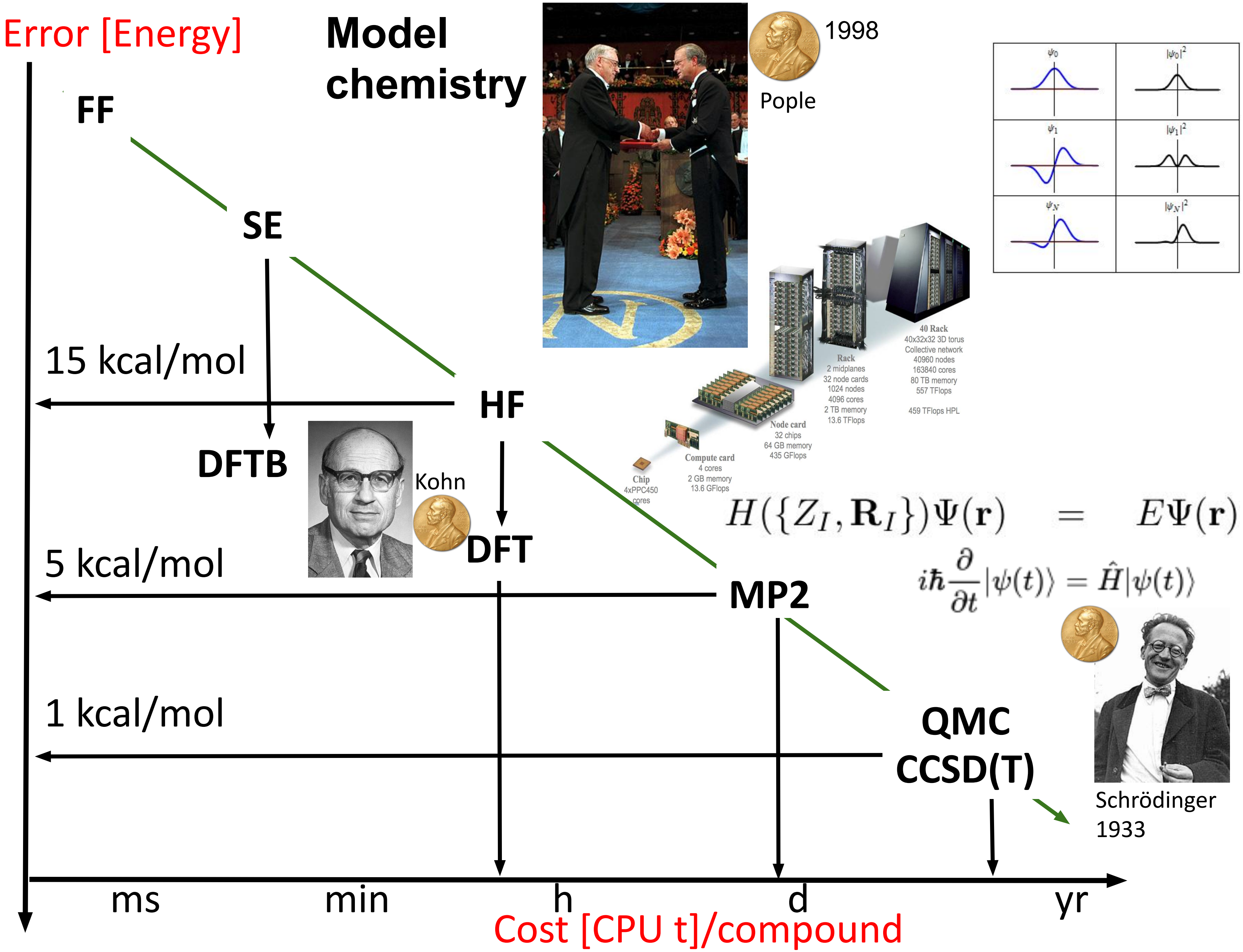


Schrödinger
1933

QML???

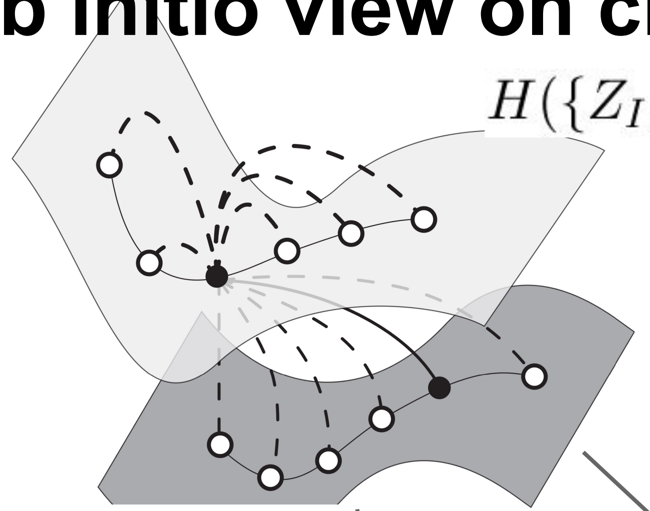


Error [Energy]
FF
SE
15 kcal/mol
HF
5 kcal/mol
MP2
1 kcal/mol
QMC
CCSD(T)
ms
min
h
d
yr
Cost [CPU t]/compound
The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.
Dirac
ψ_0
$|\psi_0|^2$
ψ_1
$|\psi_1|^2$
ψ_N
$|\psi_N|^2$
40 Rack
40x32x32 3D torus
Collective network
40960 nodes
163840 cores
80 TB memory
557 TFlops
459 TFlops HPL
Rack
2 midplanes
32 node cards
1024 nodes
4096 cores
2 TB memory
13.6 TFlops
Node card
32 chips
64 GB memory
435 GFlops
Compute card
4 cores
2 GB memory
13.6 GFlops
Chip
4xPPC450
cores
$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$
$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$
Schrödinger
1933



Error [Energy]
Model chemistry
1998
Pople
FF
SE
15 kcal/mol
HF
DFTB
Kohn
DFT
5 kcal/mol
MP2
1 kcal/mol
QMC
CCSD(T)
Schrödinger
1933
ms
min
h
d
yr
Cost [CPU t]/compound
ψ_0
$|\psi_0|^2$
ψ_1
$|\psi_1|^2$
ψ_N
$|\psi_N|^2$
40 Rack
40x32x32 3D torus
Collective network
40960 nodes
163840 cores
80 TB memory
557 TFlops
459 TFlops HPL
Rack
2 midplanes
32 node cards
1024 nodes
4096 cores
2 TB memory
13.6 TFlops
Node card
32 chips
64 GB memory
435 GFlops
Compute card
4 cores
2 GB memory
13.6 GFlops
Chip
4xPPC450
cores
$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$
$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$

Ab initio view on chemical space

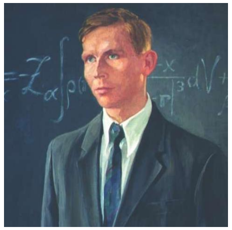


$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) \quad = \quad E\Psi(\mathbf{r})$$

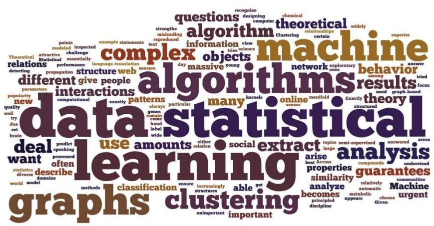
Schrödinger

Feynman

Hellmann



Fukui



correlation (interpolate)

QML

Alchemy

perturbation (extrapolate)

Vapnik

$$\{Z_I, \mathbf{R}_I\} \xmapsto{H\Psi} E$$

supervised learning

$$\{Z_I, \mathbf{R}_I\} \xmapsto{\mathrm{ML}} E$$

$$E(H(\lambda)) \quad = \quad E(H_i + \lambda(H_f - H_i))$$

$$\frac{\partial E[H]}{\partial \lambda} \quad = \quad \left\langle \Psi \left| \frac{\partial H(\lambda)}{\partial \lambda} \right| \Psi \right\rangle$$

$$E_{\mathrm{t}} = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n}{\partial \lambda^n} \left\langle \psi_\lambda \left| \hat{H}(\lambda) \right| \psi_\lambda \right\rangle \Big|_{\lambda=0} = E_{\mathrm{r}} + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial^n E(\lambda)}{\partial \lambda^n} \Big|_{\lambda=0}$$

$$\frac{\partial E[H]}{\partial R_{Ix}} \quad = \quad \left\langle \Psi \left| \frac{\partial H}{\partial R_{Ix}} \right| \Psi \right\rangle \qquad \frac{\partial E[H]}{\partial Z_I} \quad = \quad \left\langle \Psi \left| \frac{\partial H}{\partial Z_I} \right| \Psi \right\rangle$$

$$E_{\mathrm{t}} = E_{\mathrm{r}} + \Delta E^{\mathrm{NN}} + \int_\Omega d\mathbf{r} \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \Delta v \frac{\partial^n \rho_\lambda(\mathbf{r})}{\partial \lambda^n} \Big|_{\lambda=0}$$

von Rudorff, von Lilienfeld, *Phys Rev Res* (2020)

Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

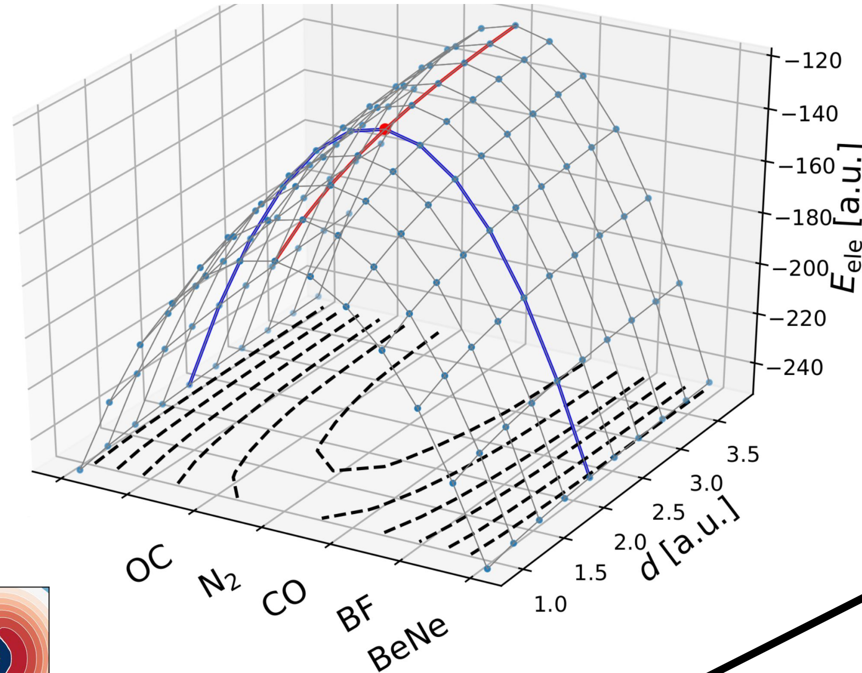

AI
vs.
Physics

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302

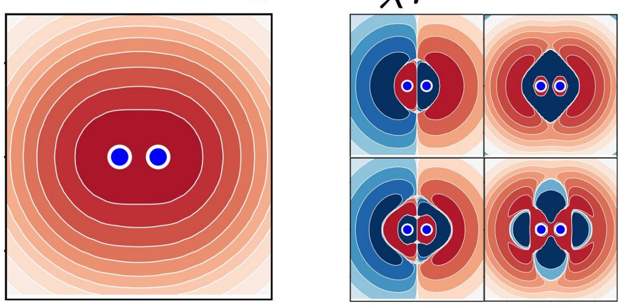
N_e

CCS ~ $O(4N_I + 1)$

$0 < N_I < 10^6$

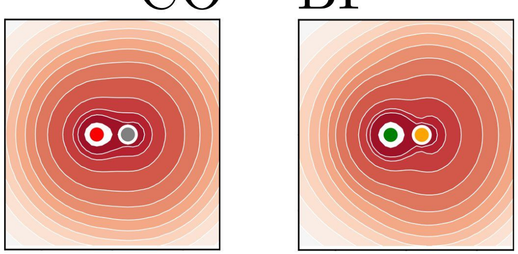


$N_2 + \partial_\lambda^n \rho$



$\Downarrow$

CO BF



R

QM: $M \mapsto f(M)$

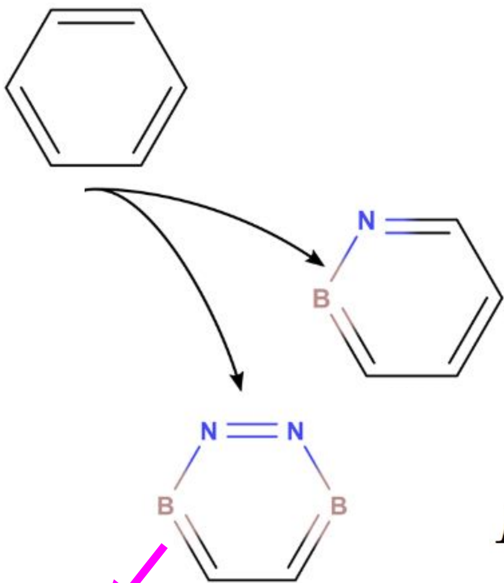
“First principles view on chemical space”, von Lilienfeld, *Int J Quantum Chem* (2013)

von Rudorff, von Lilienfeld, *Phys Rev Research* (2020), *Science Advances* (2021)

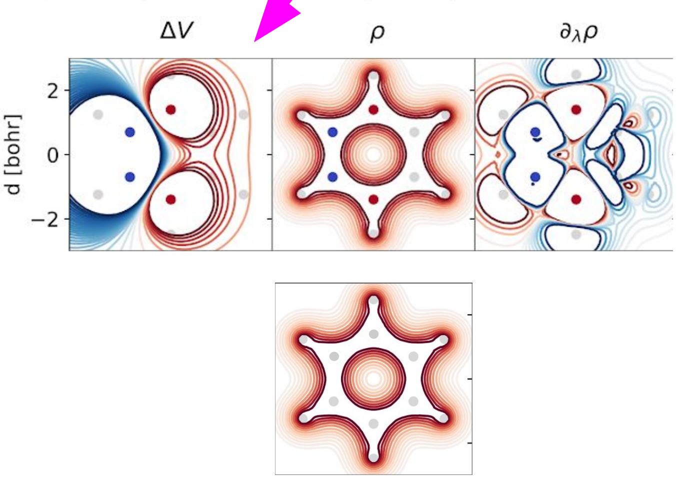
Z

Alchemical Chirality

"Give me a lever and a place to stand and I will move the earth." Archimedes

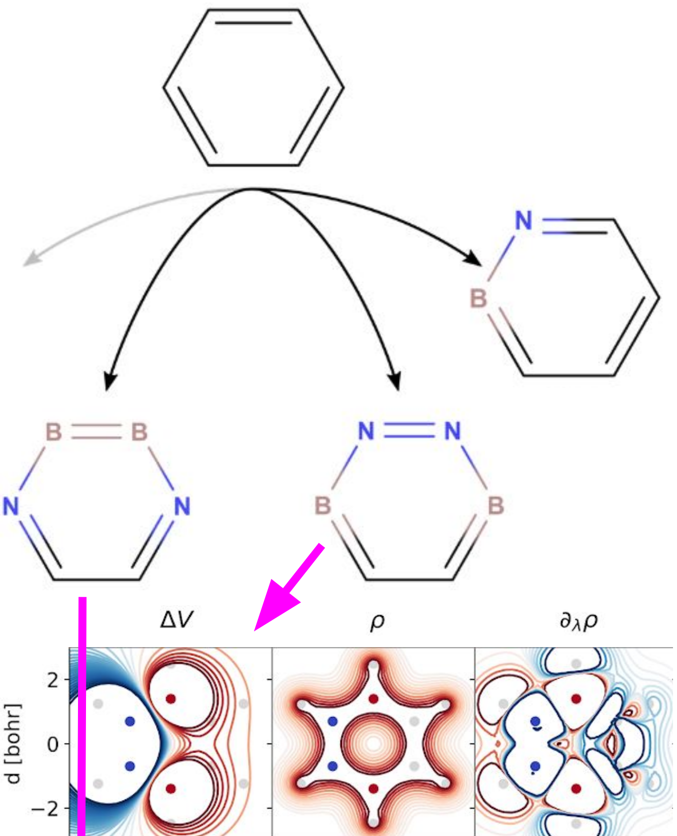


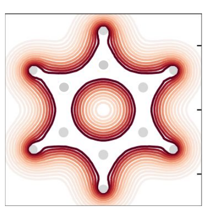
$$E_{\mathrm{t}} - E_{\mathrm{r}} = \int_{\Omega} d\mathbf{r} \Delta v(\mathbf{r}) \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial^{n-1} \rho(\mathbf{r})}{\partial \lambda^{n-1}}$$

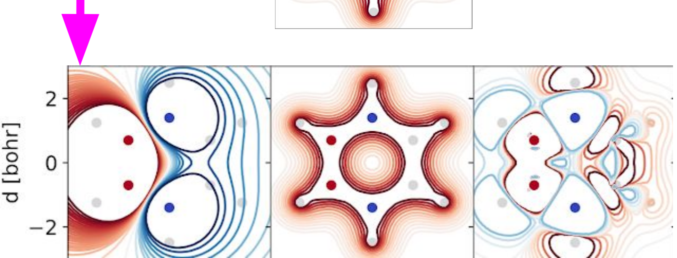


von Rudorff, von Lilienfeld, *Science Advances* (2021)

Alchemical Chirality







$$\Delta E_{ij}^{\mathrm{sym}} \approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\ \Delta v_{ri}(\mathbf{r}) \left(\frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_i^{n-1}} + \frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_j^{n-1}} \right)$$

von Rudorff, von Lilienfeld, *Science Advances* (2021)

"Give me a lever and a place to stand and I will move the earth." Archimedes

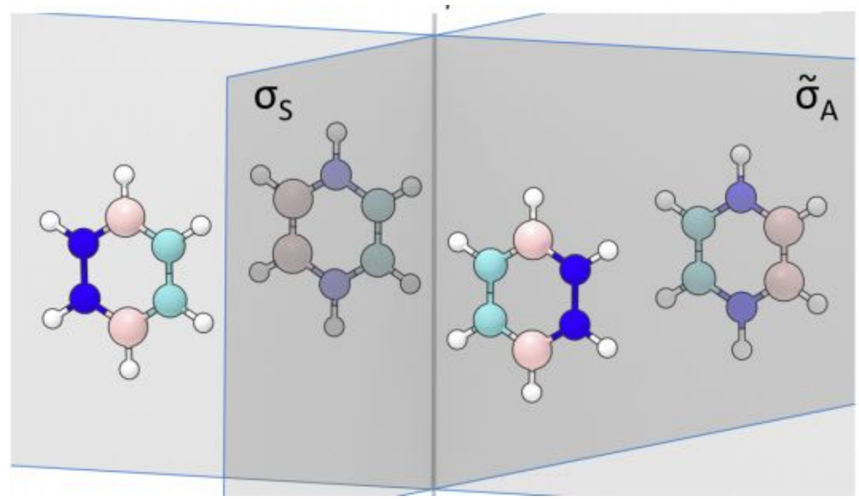
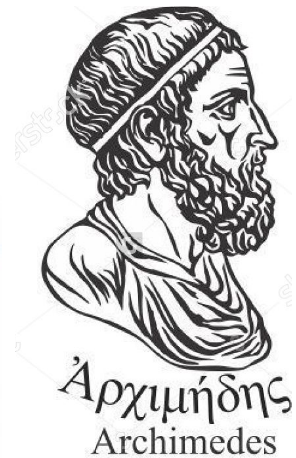
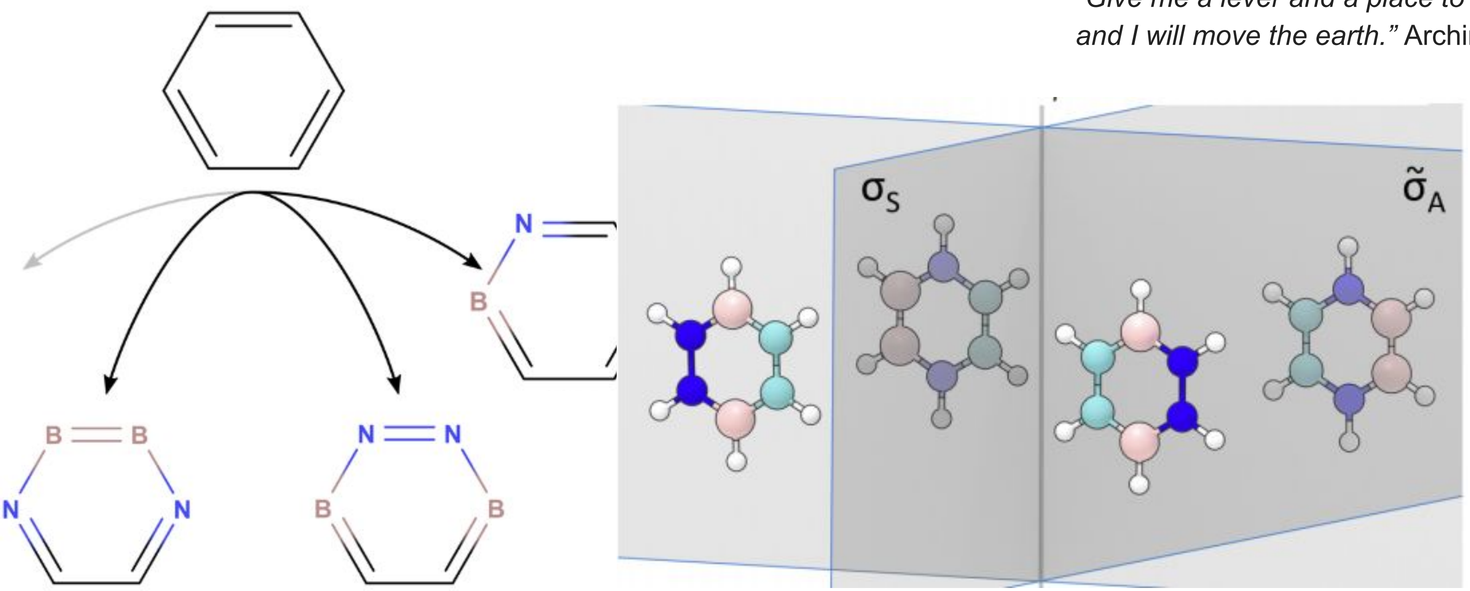


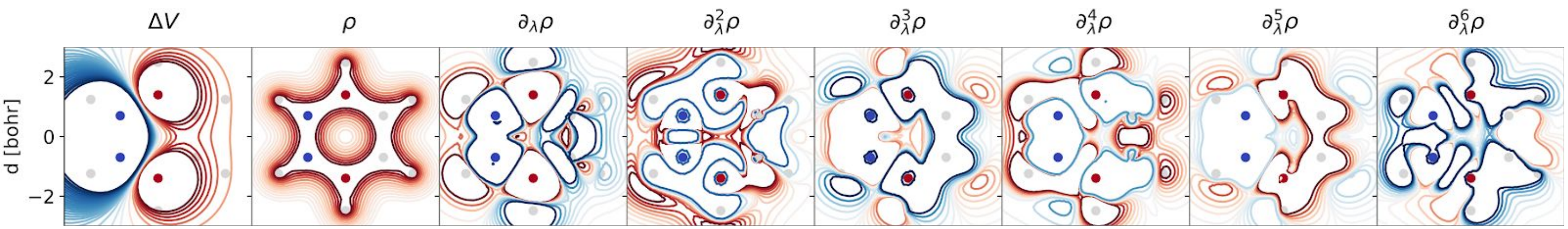
Fig. 1: Illustration of alchemical chirality (from Ref. 7)

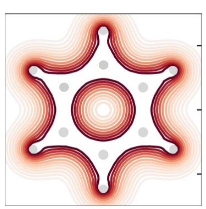


Alchemical Chirality

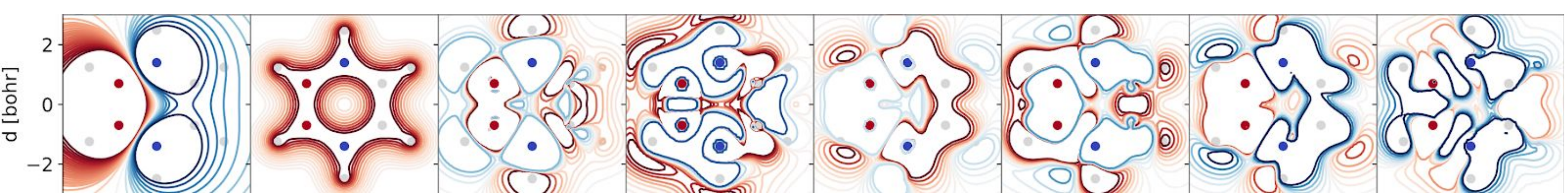
"Give me a lever and a place to stand and I will move the earth." Archimedes





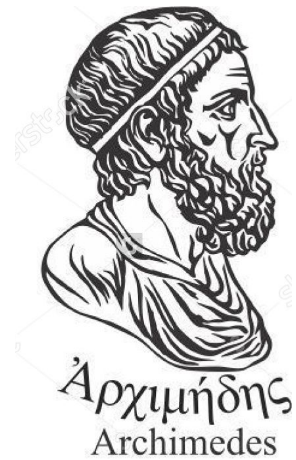


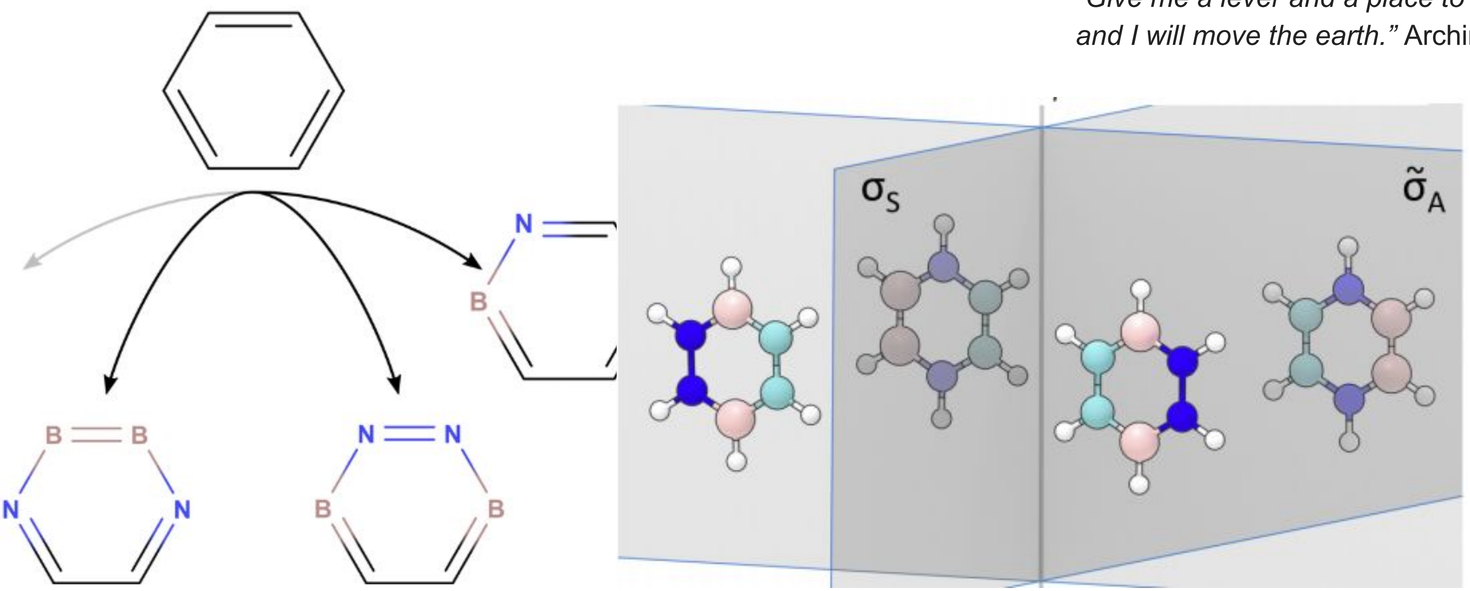
$$\Delta E_{ij}^{\mathrm{sym}} \approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\ \Delta v_{ri}(\mathbf{r}) \left(\frac{\partial^{n-1}\rho_r(\mathbf{r})}{\partial\lambda_i^{n-1}} + \frac{\partial^{n-1}\rho_r(\mathbf{r})}{\partial\lambda_j^{n-1}} \right)$$

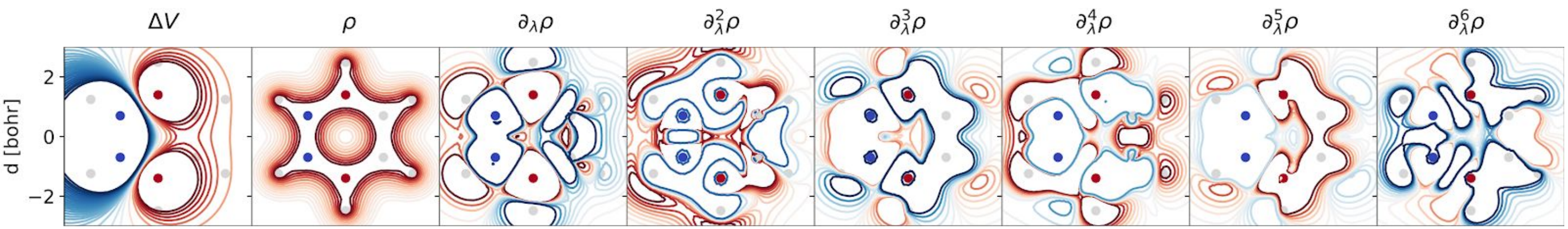


Alchemical Chirality

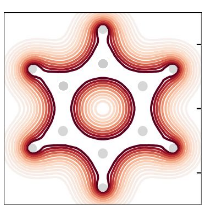
"Give me a lever and a place to stand and I will move the earth." Archimedes







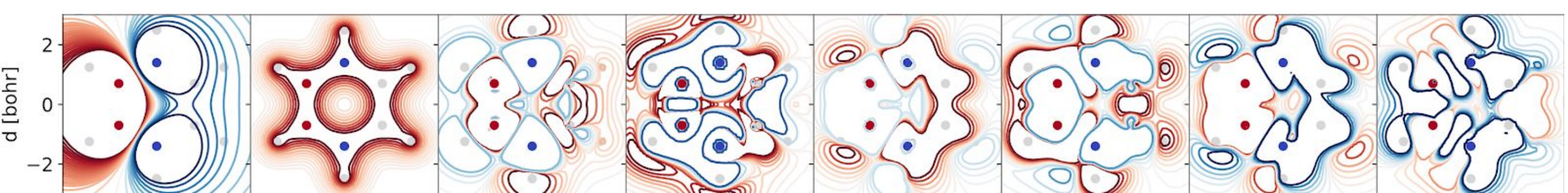
NN+2BC +CC+2NB



$$\Delta E_{ij}^{\mathrm{sym}} \approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\, \Delta v_{ri}(\mathbf{r}) \left(\frac{\partial^{n-1}\rho_r(\mathbf{r})}{\partial \lambda_i^{n-1}} + \frac{\partial^{n-1}\rho_r(\mathbf{r})}{\partial \lambda_j^{n-1}} \right)$$

$\rightarrow$ NC ~ BC +0.5*(NN-BB)

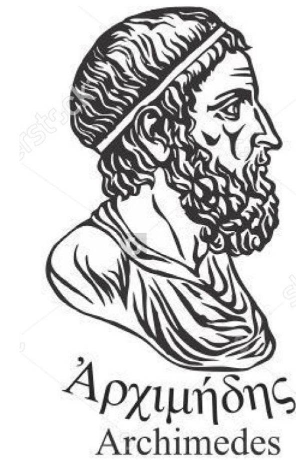
von Rudorff, von Lilienfeld, *Science Advances* (2021)

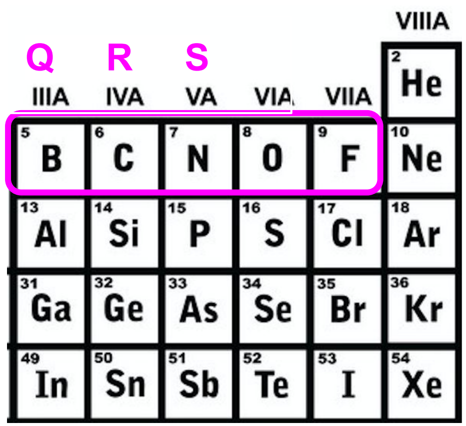


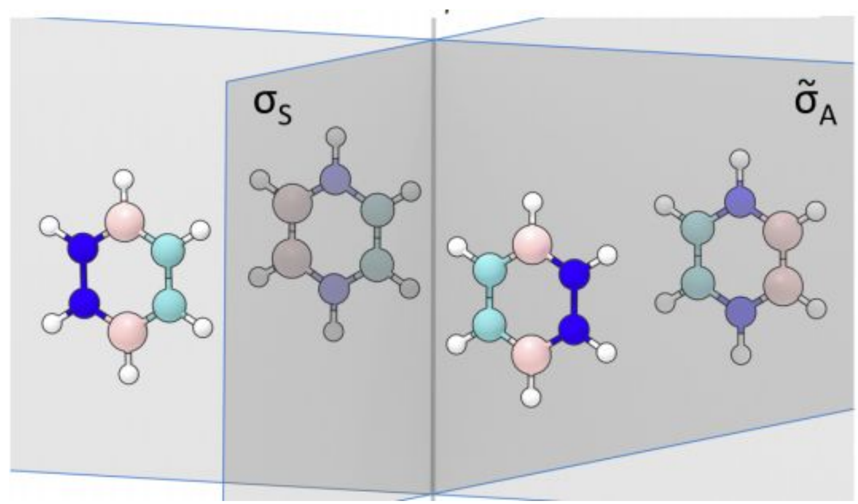
BB+2NC +CC+2NB

Alchemical Chirality

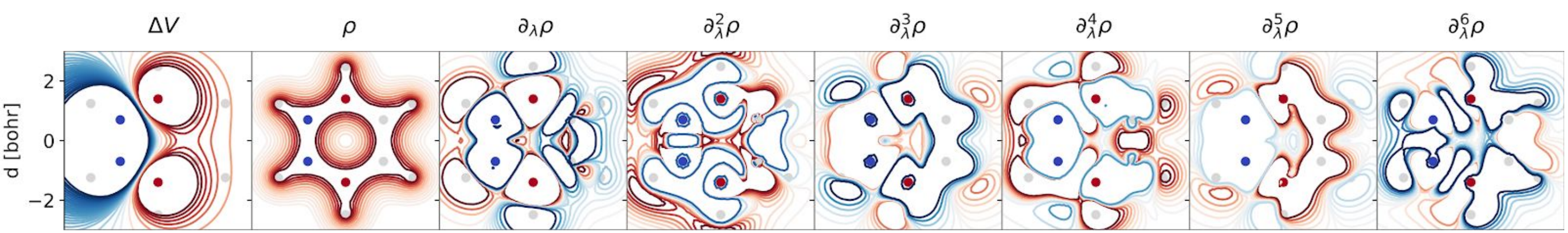
“Give me a lever and a place to stand and I will move the earth.” Archimedes



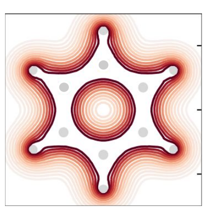




→ SR ~ QR +0.5*(SS-QQ)



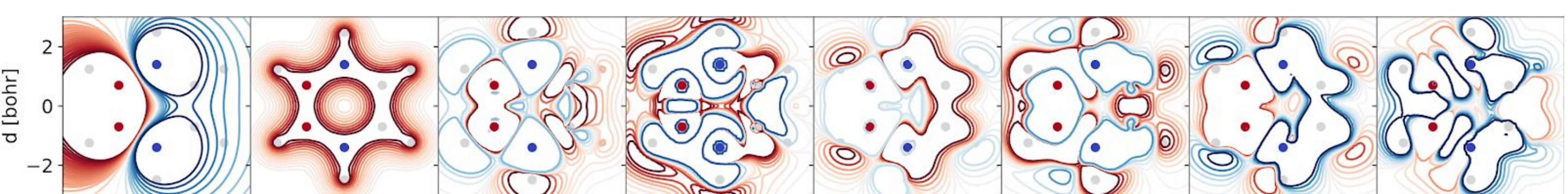
NN+2BC +CC+2NB



$$\Delta E_{ij}^{\mathrm{sym}} \approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\, \Delta v_{ri}(\mathbf{r}) \left(\frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_i^{n-1}} + \frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_j^{n-1}} \right)$$

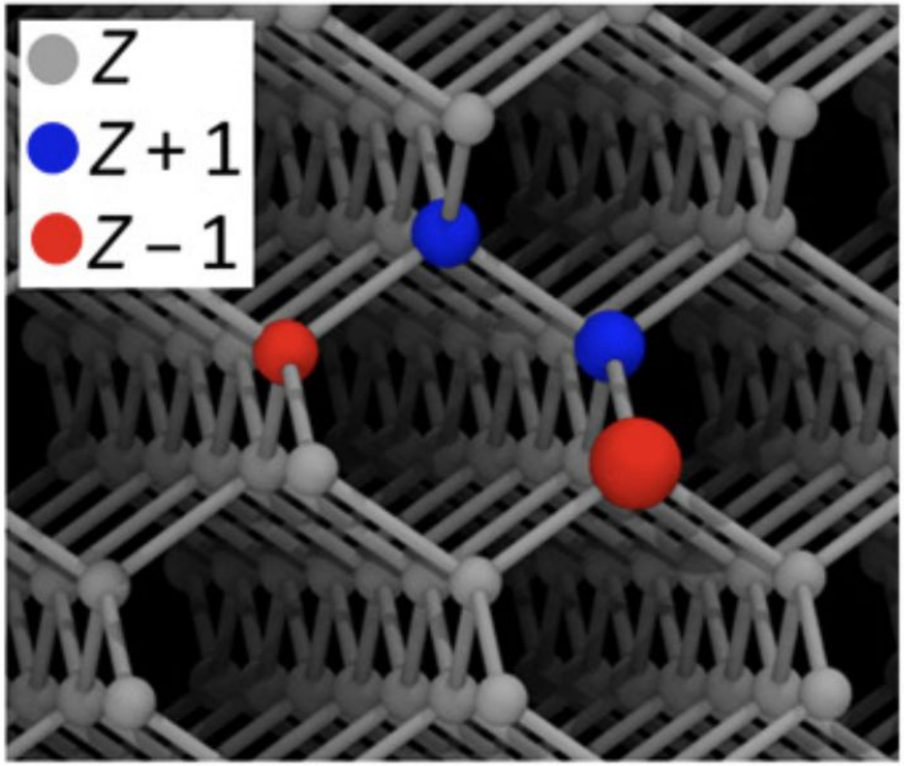
→ NC ~ BC +0.5*(NN-BB)

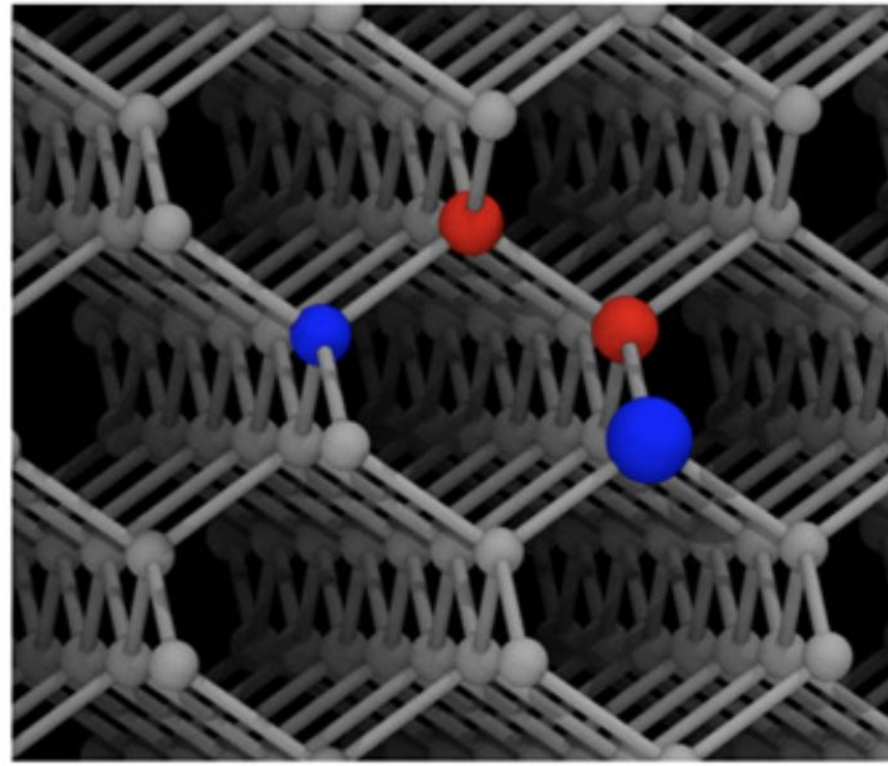
von Rudorff, von Lilienfeld, *Science Advances* (2021)



BB+2NC +CC+2NB

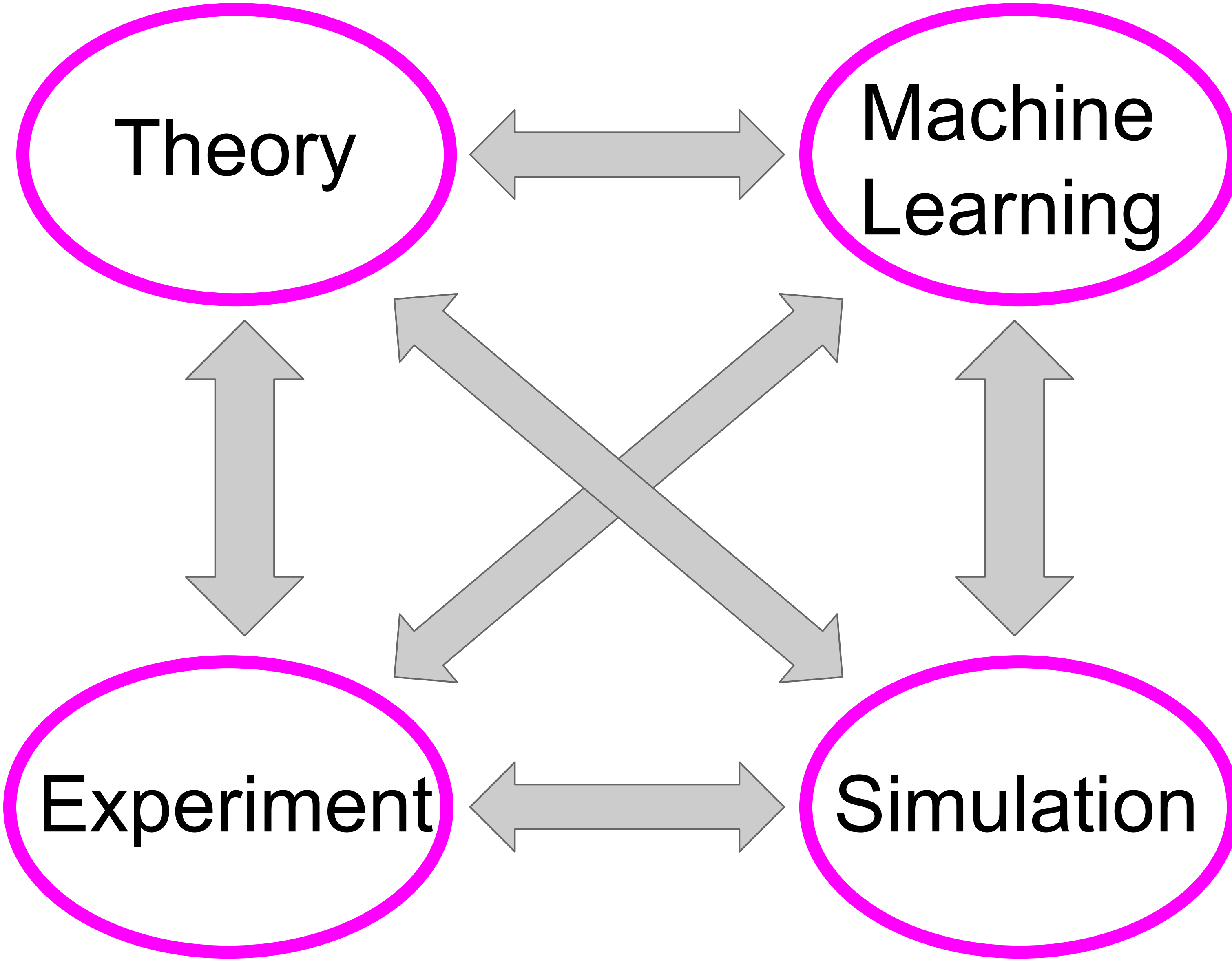
Alchemical Chirality



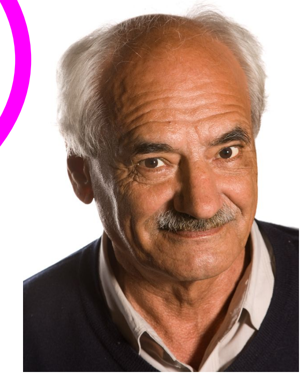


von Rudorff, von Lilienfeld, *Science Advances* (2021)

The pillars of science

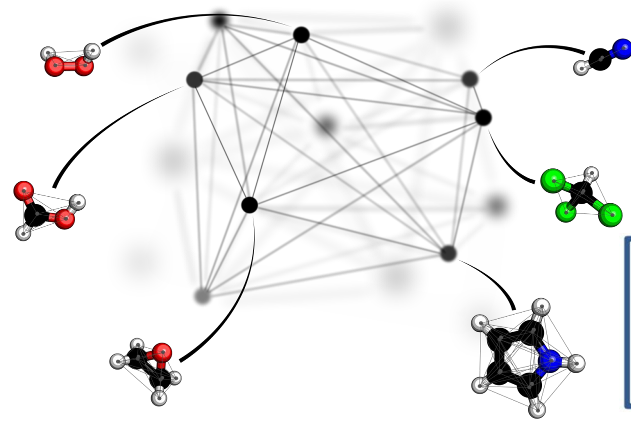


Vapnik

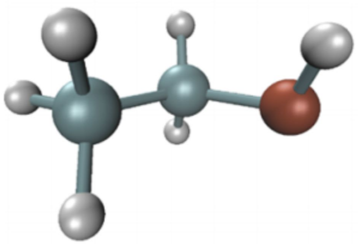


Chervonenkis

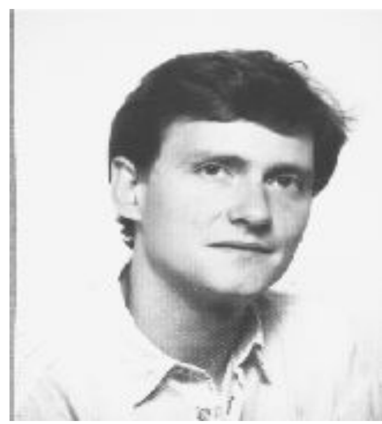

AI
vs.
Physics



$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1}\vec{P}^{\text{ref}}$$



	O	C	C	H	H	H	H	H	H
O	O	OC	OC	OH	OH	OH	OH	OH	OH
C	OC	C	CC	CH	CH	CH	CH	CH	CH
C	OC	CC	C	CH	CH	CH	CH	CH	CH
H	OH	CH	CH	H	HH	HH	HH	HH	HH
H	OH	CH	CH	HH	H	HH	HH	HH	HH
H	OH	CH	CH	HH	HH	H	HH	HH	HH
H	OH	CH	CH	HH	HH	HH	H	HH	HH
H	OH	CH	CH	HH	HH	HH	HH	H	HH
H	OH	CH	CH	HH	HH	HH	HH	HH	H



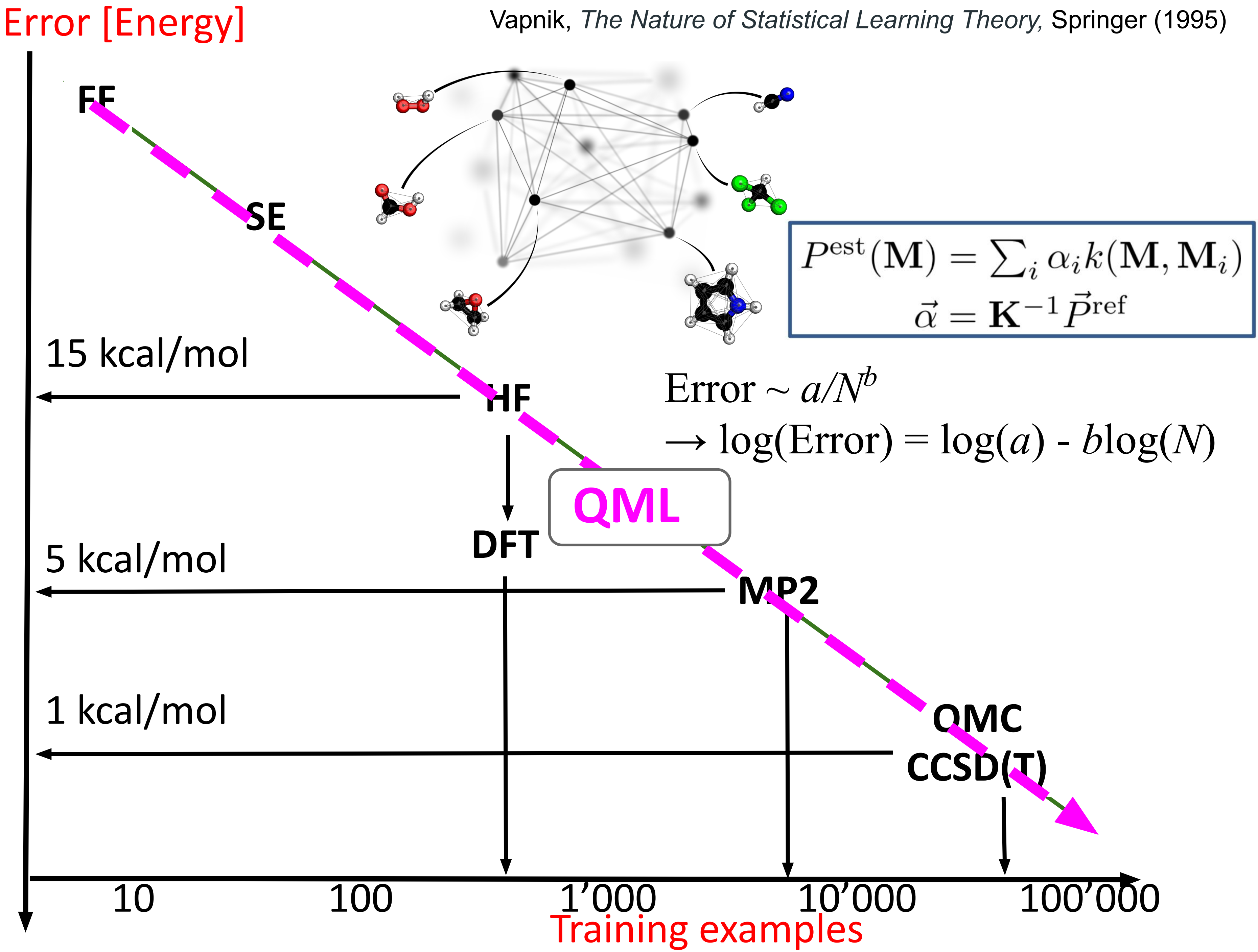
K.-R. Müller
TU Berlin

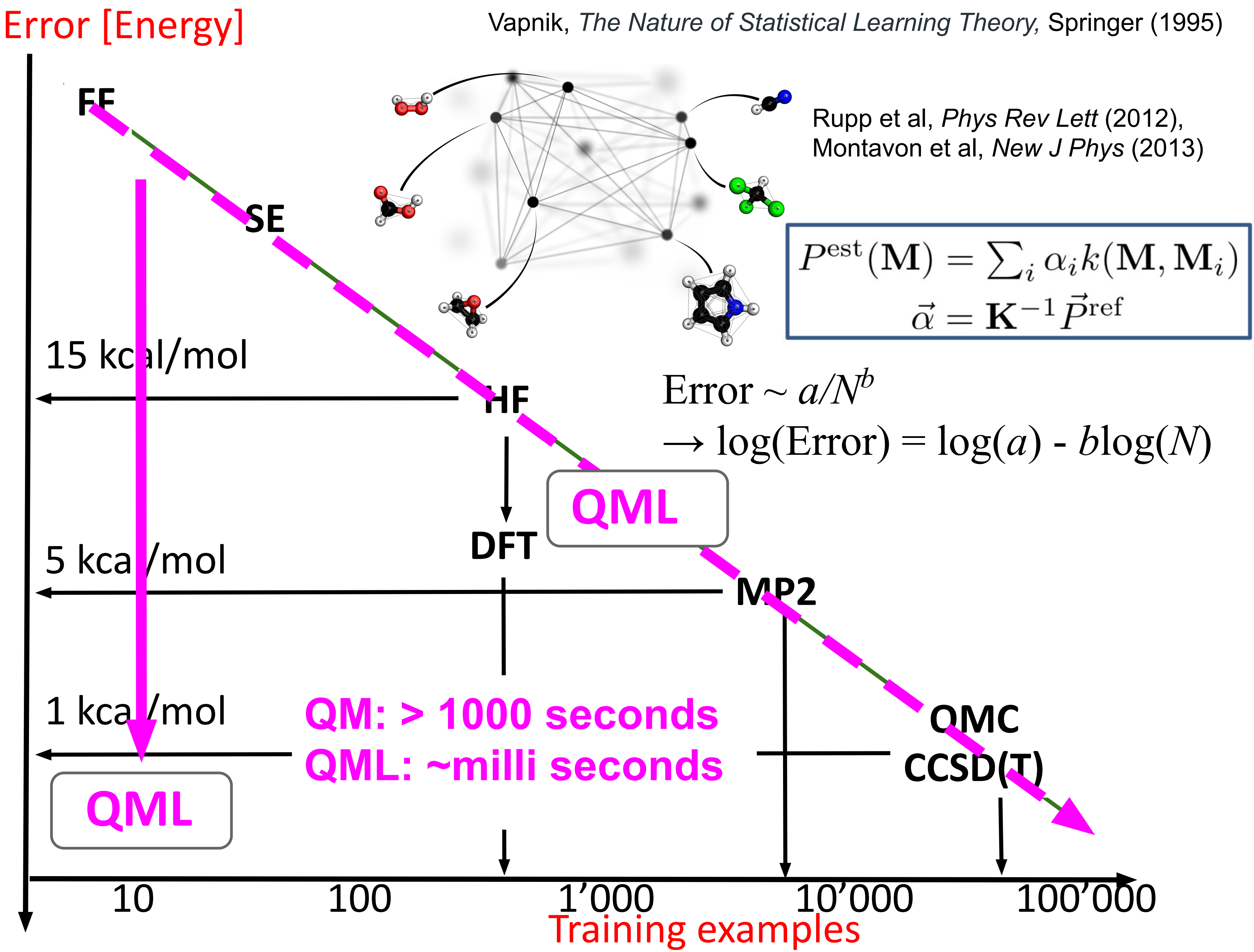
A. Tkatchenko
UofLuxembourg

M. Rupp
UofKonstanz

Institute for Pure & Applied Mathematics
IPAM
University of California, Los Angeles

Vapnik, *The Nature of Statistical Learning Theory*, Springer (1995)

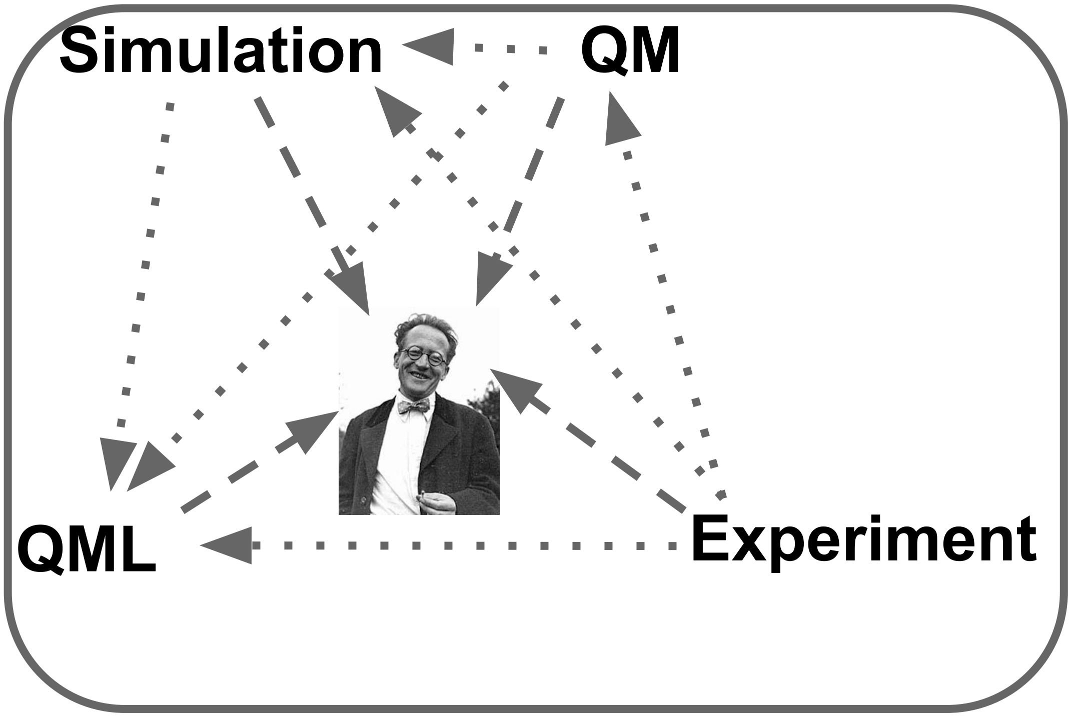


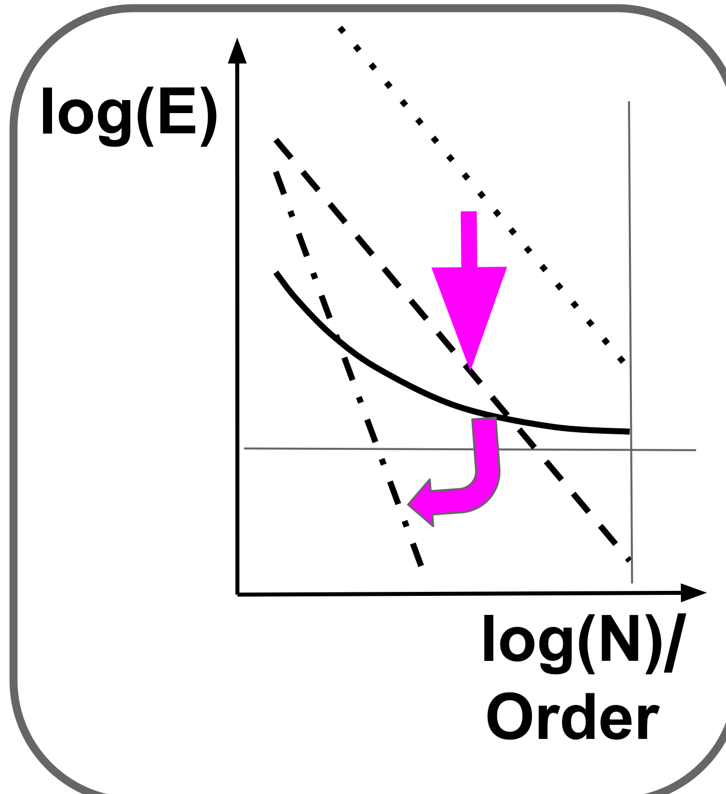


Error [Energy]
Vapnik, The Nature of Statistical Learning Theory, Springer (1995)
Rupp et al, Phys Rev Lett (2012),
Montavon et al, New J Phys (2013)
$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$
$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$
Error ~ a/N^b
$\rightarrow \log(\text{Error}) = \log(a) - b\log(N)$
FF
SE
HF
DFT
MP2
QMC
CCSD(T)
15 kcal/mol
5 kcal/mol
1 kcal/mol
QML
QML
QM: > 1000 seconds
QML: ~milli seconds
10
100
1'000
10'000
100'000
Training examples

Summary

$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$





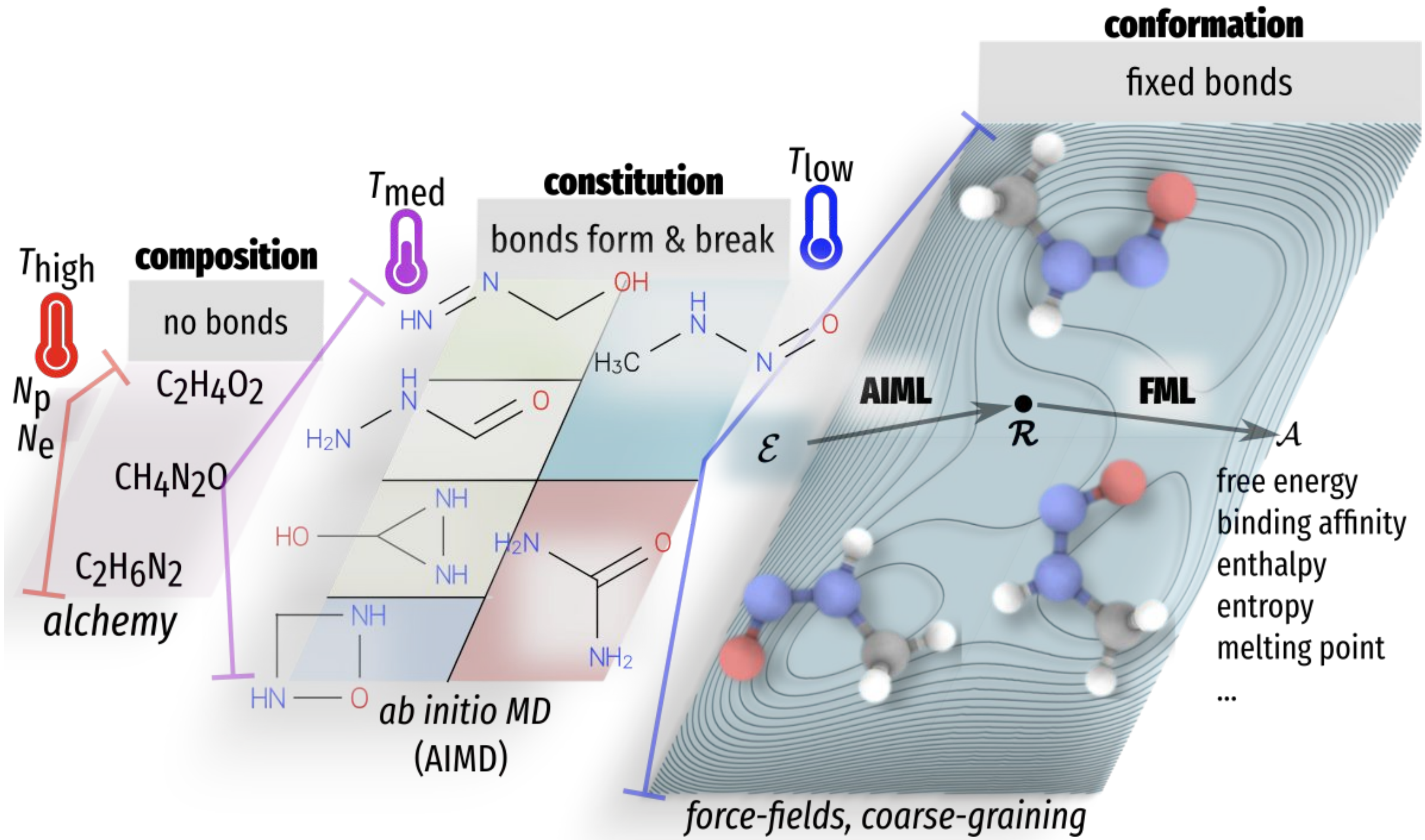
`First principles view on chemical space', *Int. J. Quantum Chem.* (2013)
`Quantum Machine Learning', *Angew. Chem. Int. Ed.* (2018)

Evolution or Revolution???

Control accuracy & speed-up

Ab initio view on chemical space

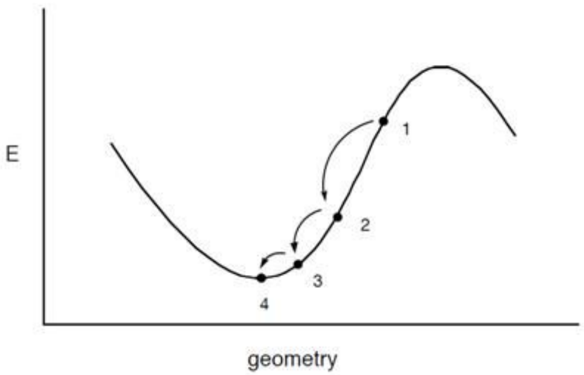
Weinreich et al, *Ab initio machine learning of phase space averages, J Chem Phys* (2022)

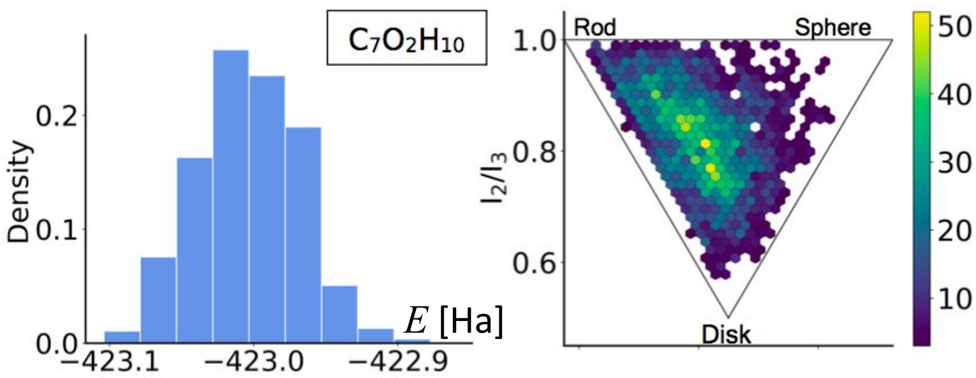


Predicting

1. structure (*Nat Commun* 2021)
2. reactivity (*J Chem Phys* 2021, *J Chem Phys* 2022)
3. Free energy (*J Chem Phys* 2021, *J Chem Phys* 2022)

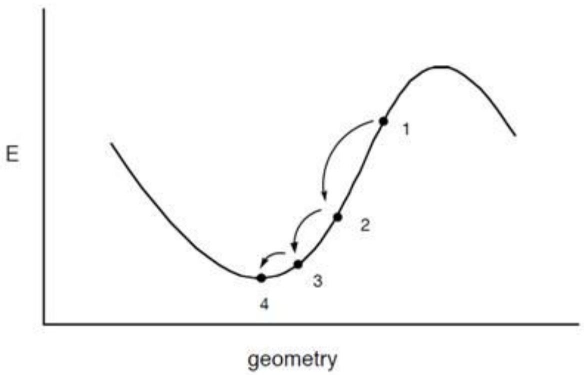
Structure

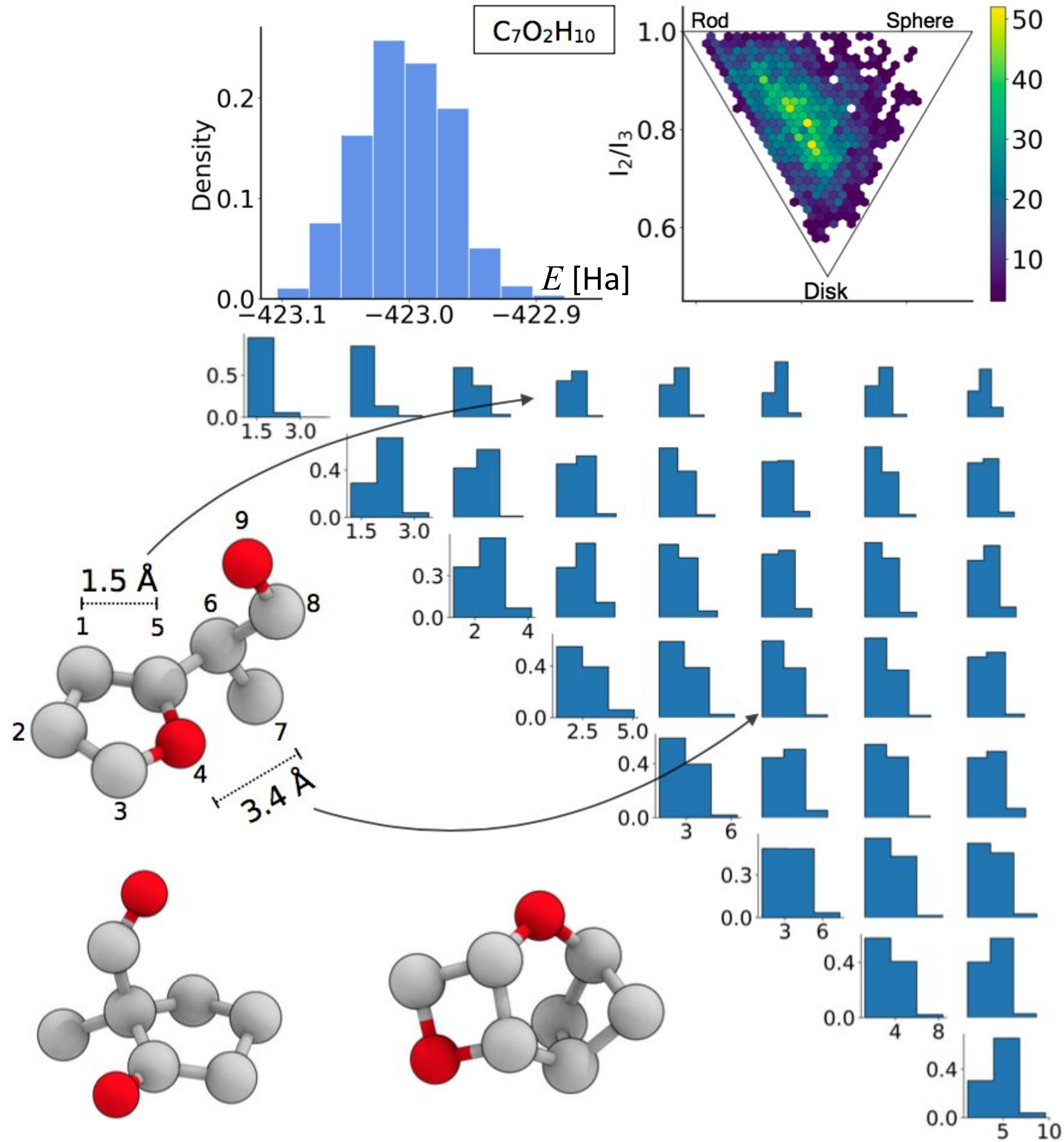




Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

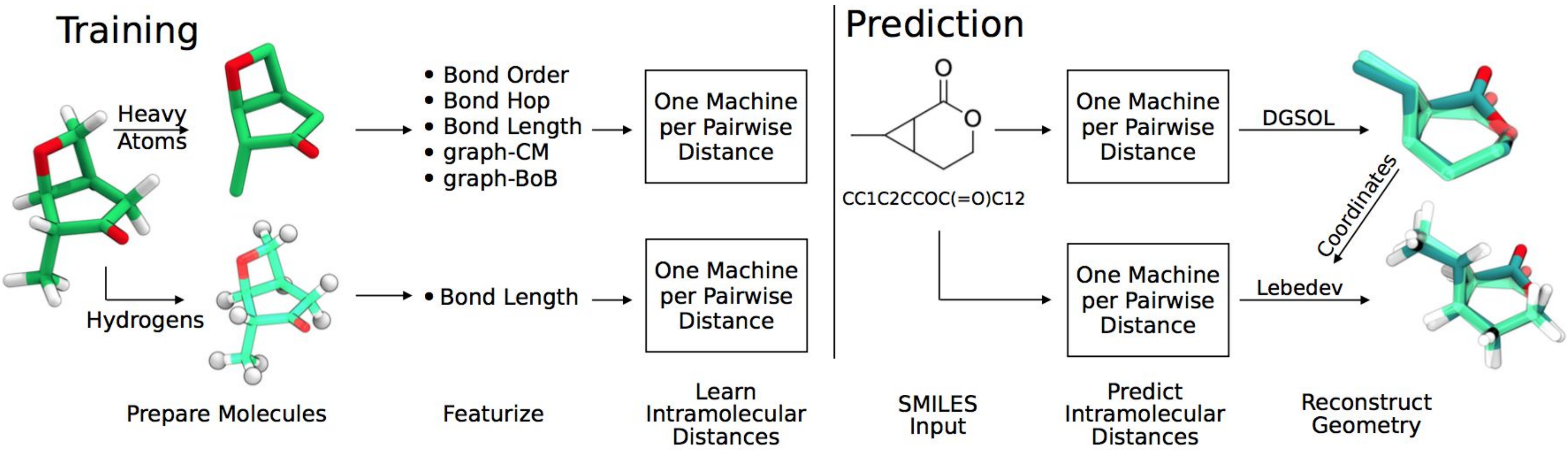
Structure





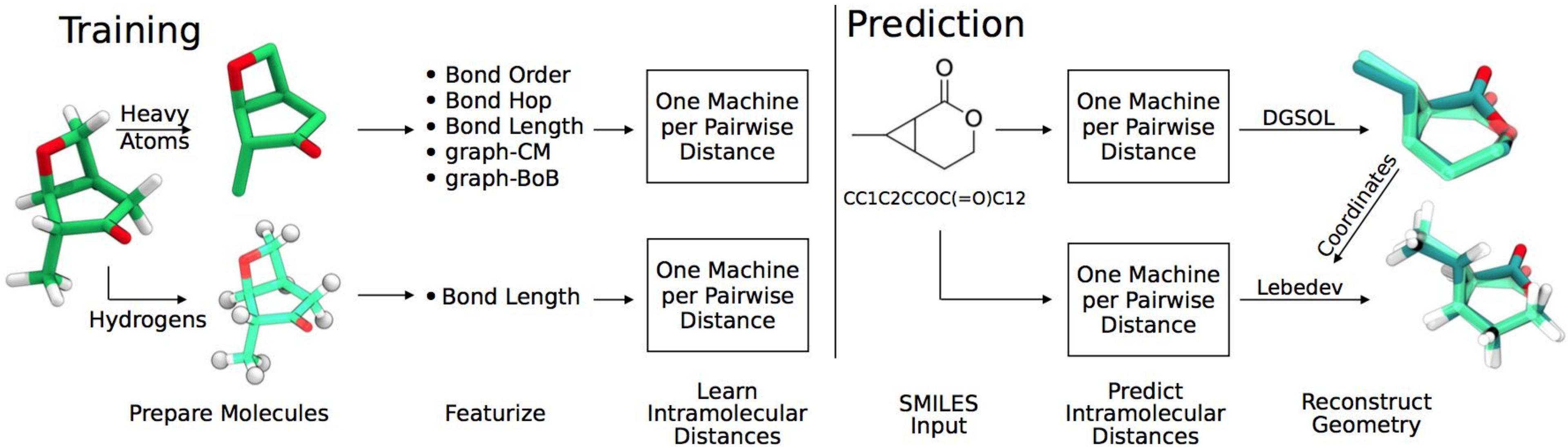
Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

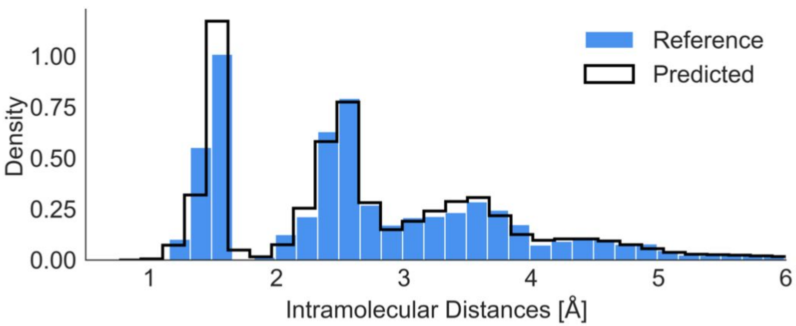
Graph to Structure (G2S)

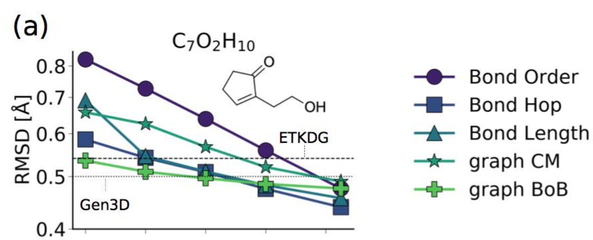


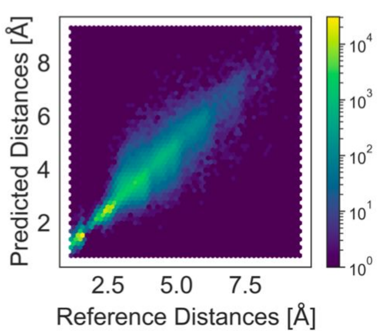
Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

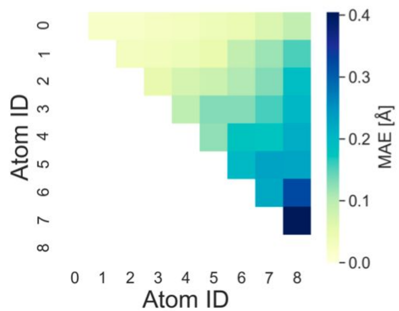
Graph to Structure (G2S)





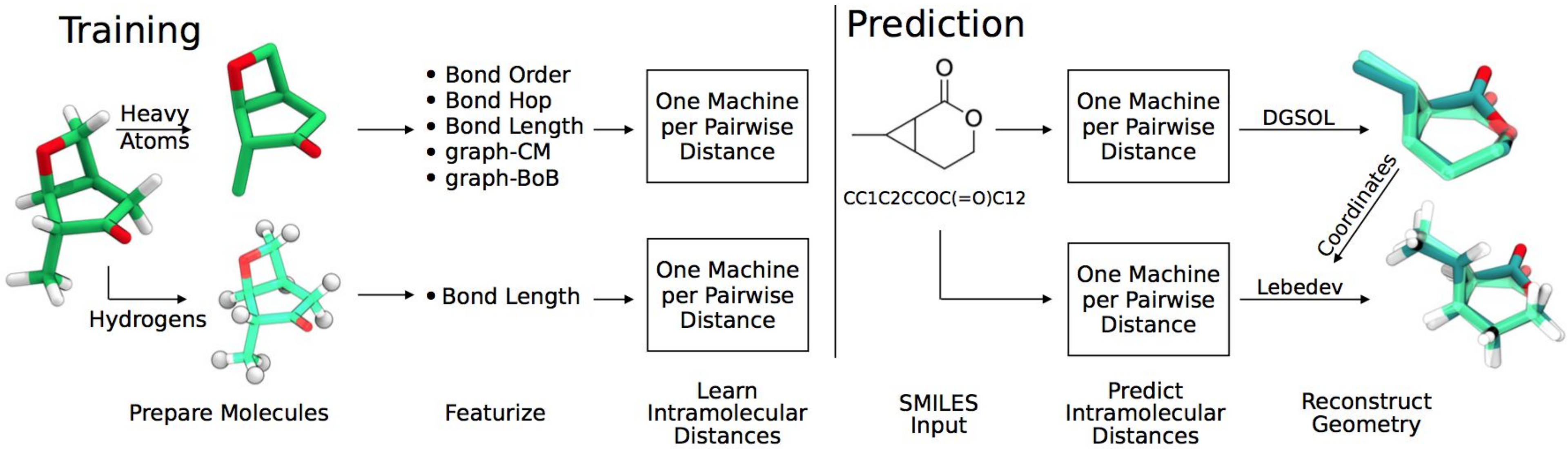


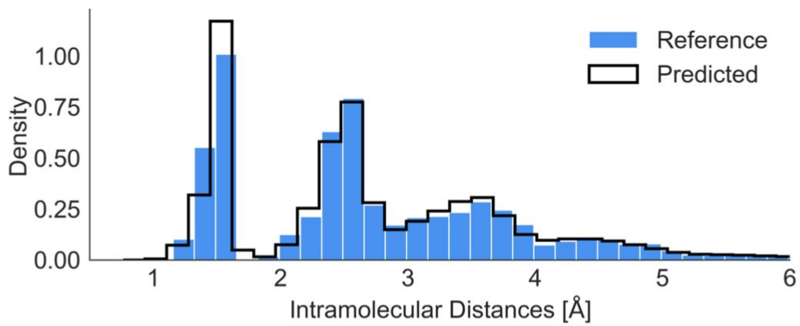


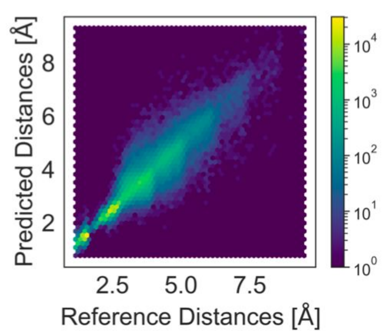


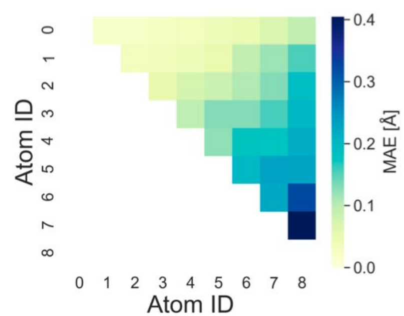
Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

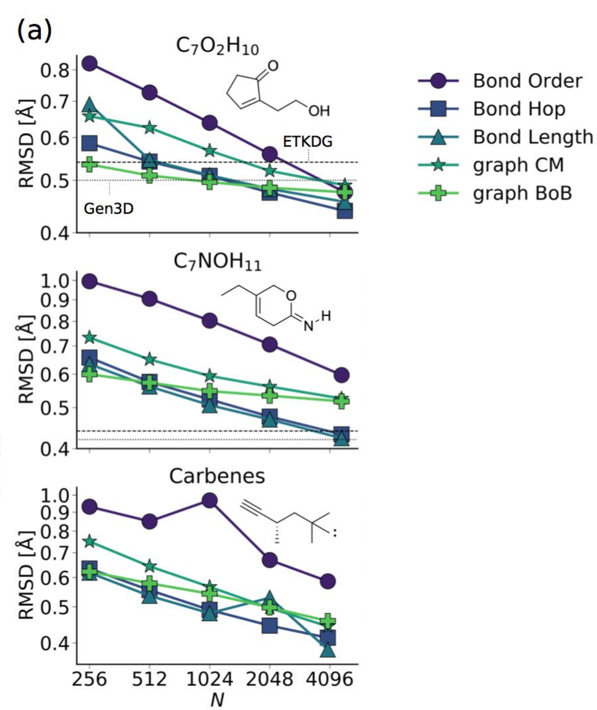
Graph to Structure (G2S)





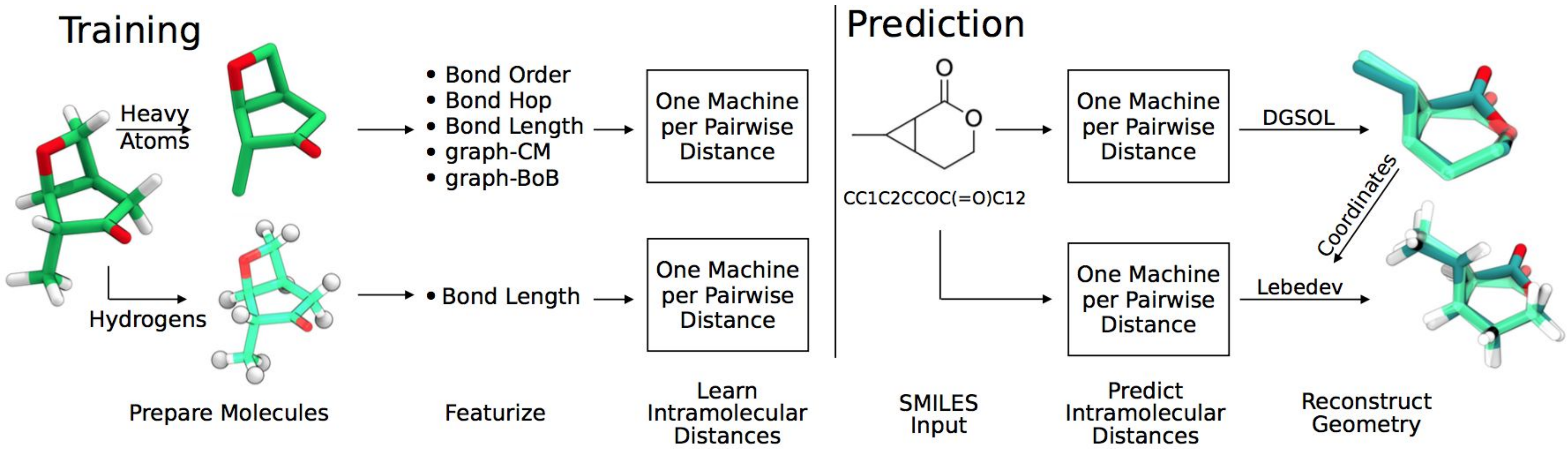


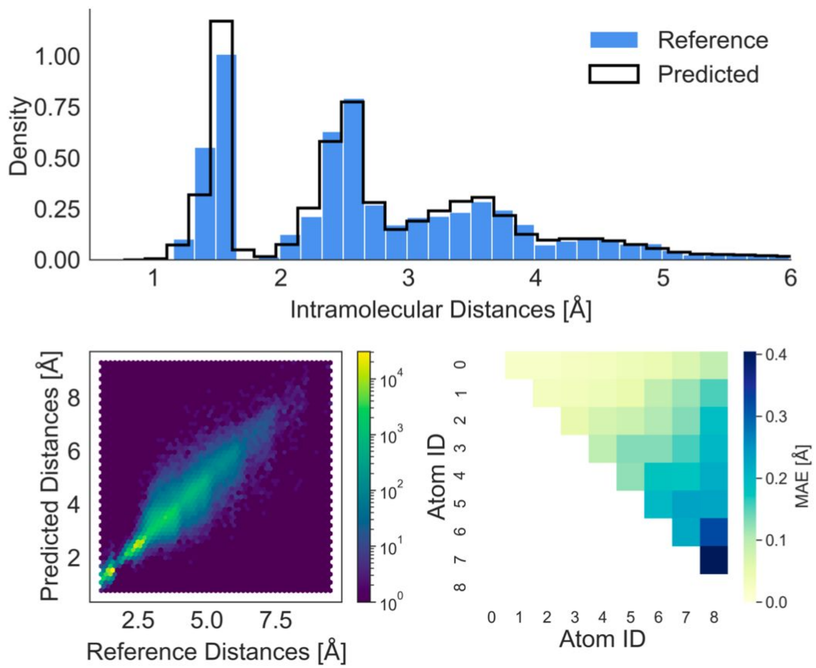


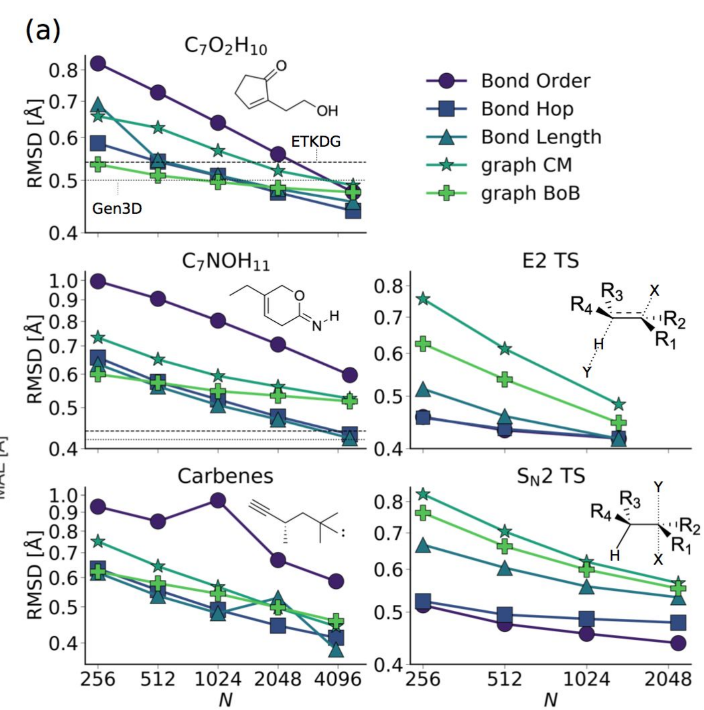


Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

Graph to Structure (G2S)

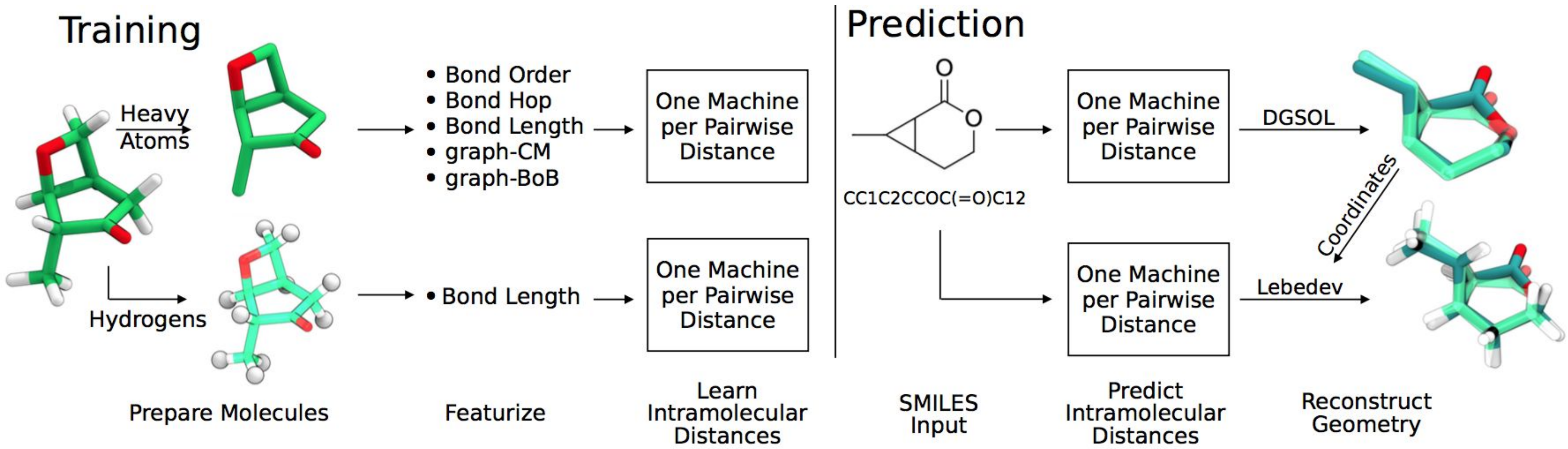


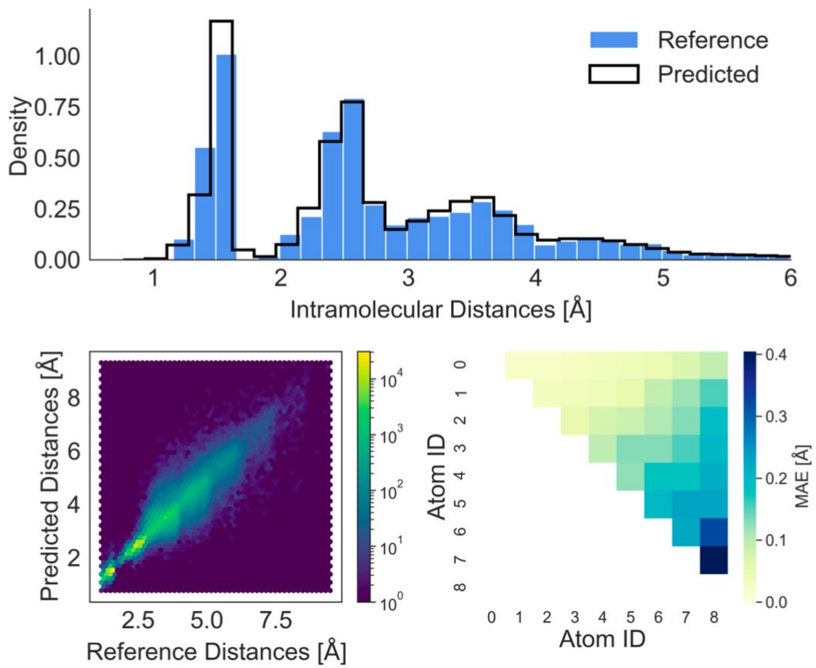


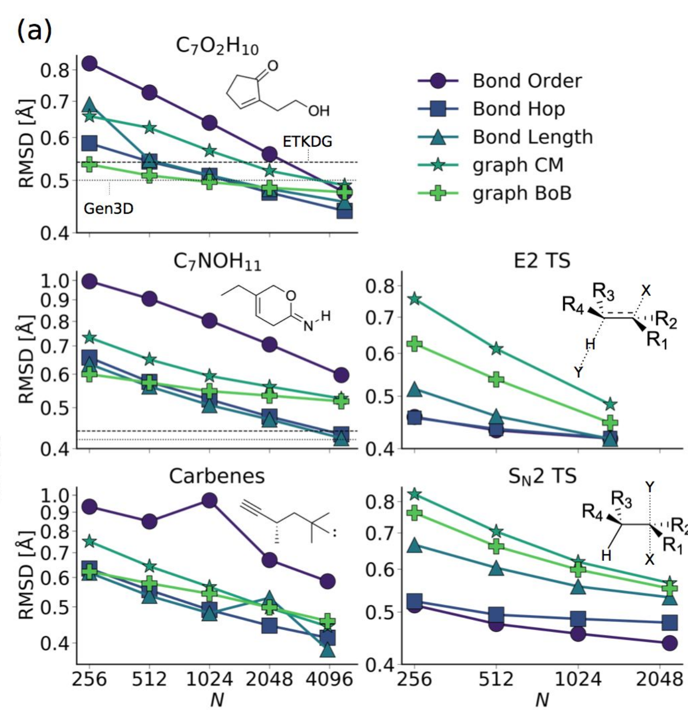


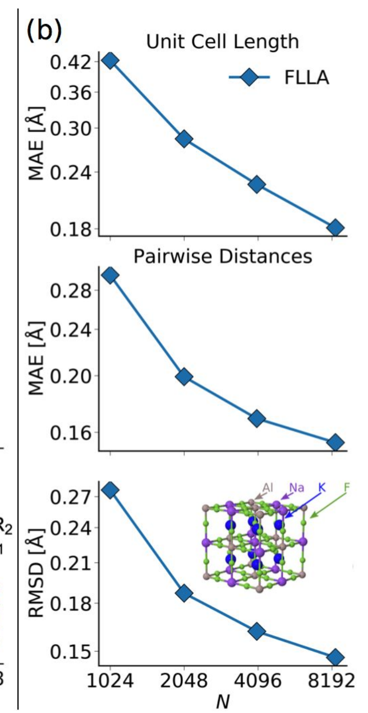
Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

Graph to Structure (G2S)





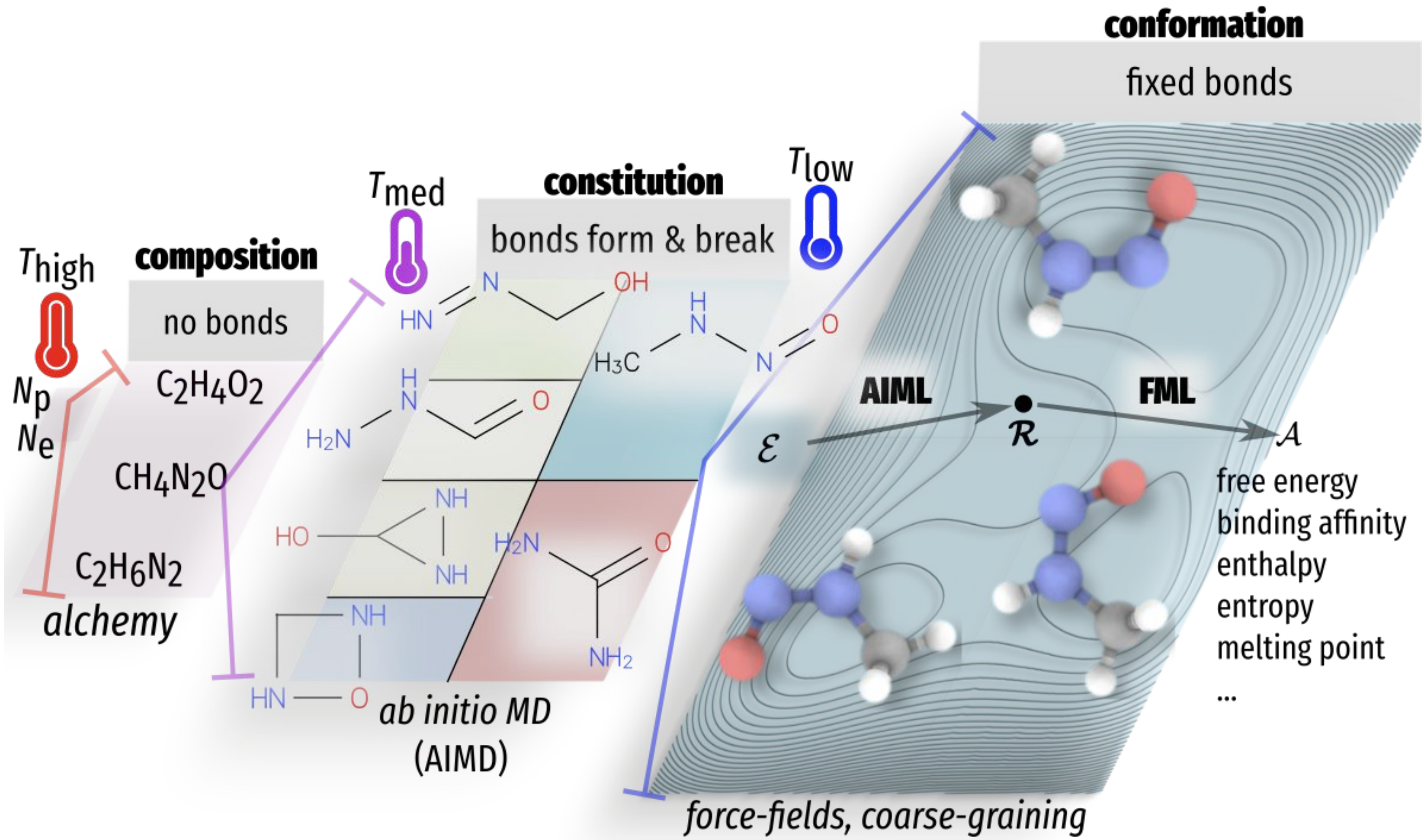




Lemm, von Rudorff, OAvL, *Nat Commun* (2021)

Ab initio view on chemical space

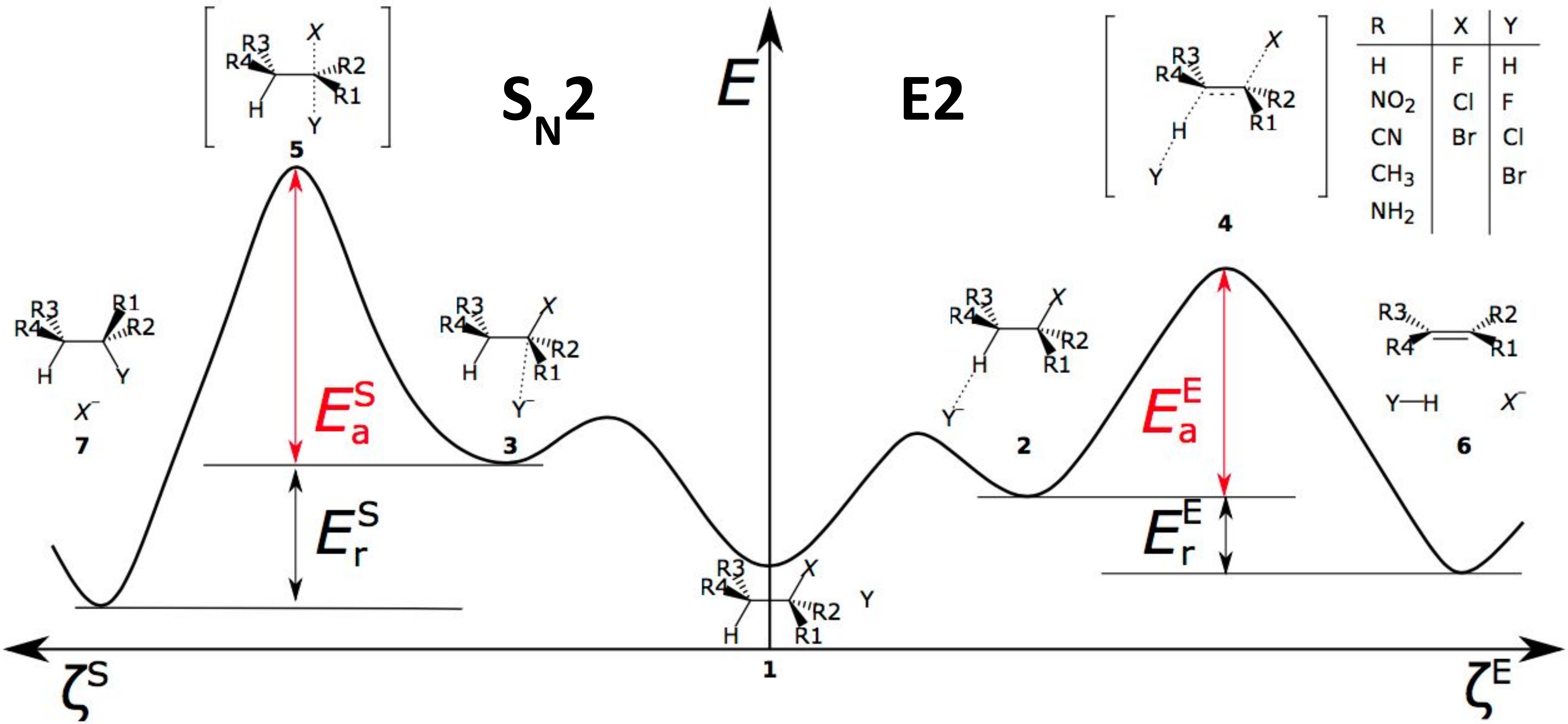
Weinreich et al, *Ab initio machine learning of phase space averages, J Chem Phys* (2022)



Predicting

1. structure (*Nat Commun* 2021)
2. reactivity (*J Chem Phys* 2021, *J Chem Phys* 2022)
3. Free energy (*J Chem Phys* 2021, *J Chem Phys* 2022)

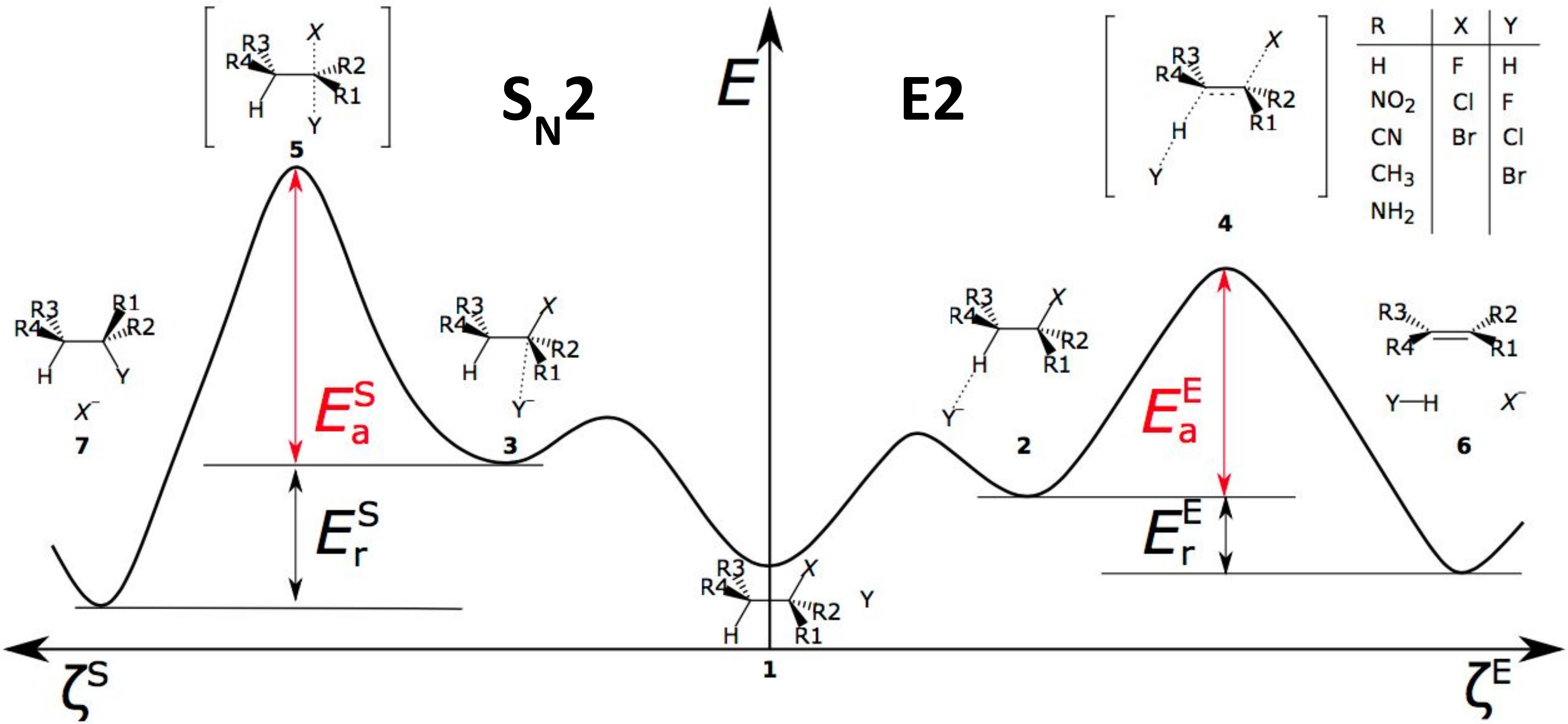
RXN



R	X	Y
H	F	H
NO_2	Cl	F
CN	Br	Cl
CH_3		Br
NH_2		

von Rudorff, Heinen, Bragato, OAvL, *MLST* (2020); Heinen, von Rudorff, OAvL, *J Chem Phys* (2021)*, arxiv* (2020)

RXN



R	X	Y
H	F	H
NO_2	Cl	F
CN	Br	Cl
CH_3		Br
NH_2		

Which is more frequent, E2 or S_N2?
Does Hammond's postulate hold?

von Rudorff, Heinen, Bragato, OAvL, *MLST* (2020); Heinen, von Rudorff, OAvL, *J Chem Phys* (2021), *arxiv* (2020)

Data: QMrxn

IOP Publishing Mach. Learn.: Sci. Technol. 1 (2020) 045026 https://doi.org/10.1088/2632-2153/aba822

MACHINE LEARNING
Science and Technology

CrossMark

OPEN ACCESS

RECEIVED
31 May 2020

REVISED
7 July 2020

ACCEPTED FOR PUBLICATION
21 July 2020

PUBLISHED
28 October 2020

PAPER

Thousands of reactants and transition states for competing E2 and S_N2 reactions

Guido Falk von Rudorff, Stefan N Heinen, Marco Bragato and O Anatole von Lilienfeld

Institute of Physical Chemistry and National Center for Computational Design and Discovery of Novel Materials (MARVEL), Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

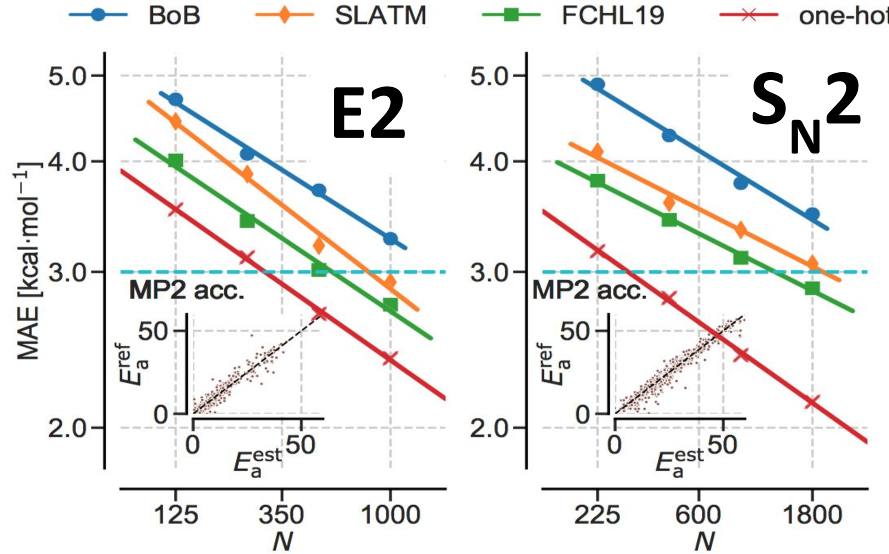
E-mail: anatole.vonlilienfeld@unibas.ch

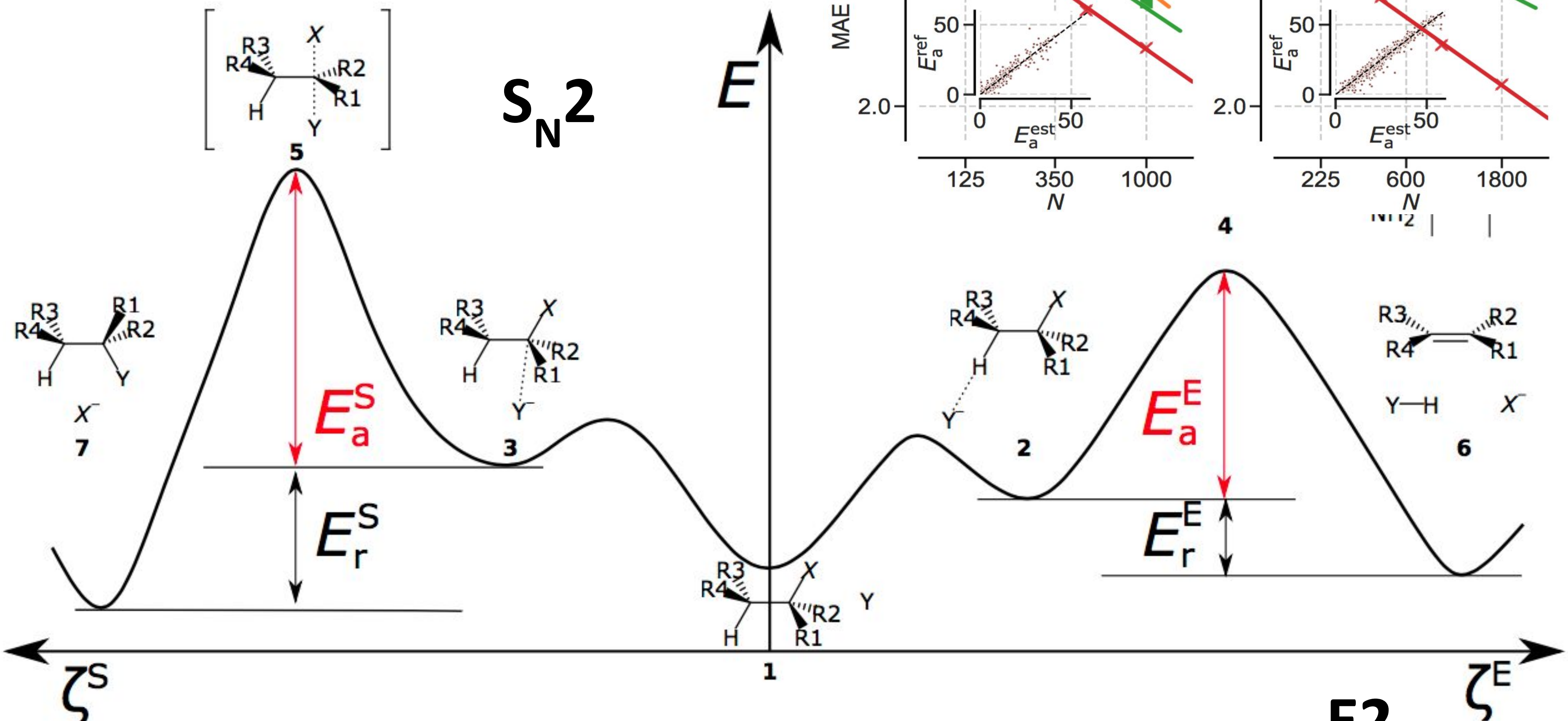
Keywords: reactions, chemical space, competing reaction channels

Abstract

Reaction barriers are a crucial ingredient for first principles based computational retro-synthesis efforts as well as for comprehensive reactivity assessments throughout chemical compound space. While extensive databases of experimental results exist, modern quantum machine learning applications require atomistic details which can only be obtained from quantum chemistry protocols. For competing E2 and S_N2 reaction channels we report 4,466 transition state and 143,200 reactant complex geometries and energies at MP2/6-311G(d) and single point DF-LCCSD/cc-pVTZ level of theory, respectively, covering the chemical compound space spanned by the substituents NO_2, CN, CH_3, and NH_2 and early halogens (F, Cl, Br) and hydrogen as nucleophiles and early halogens as leaving groups. Reactants are chosen such that the activation energy of the competing E2 and S_N2 reactions are of comparable magnitude. The correct concerted motion for each of the one-step reactions has been validated for all transition states. We demonstrate how quantum machine learning models can support data set extension, and discuss the distribution of key internal coordinates of the transition states.

Reactant to barrier (R2B)



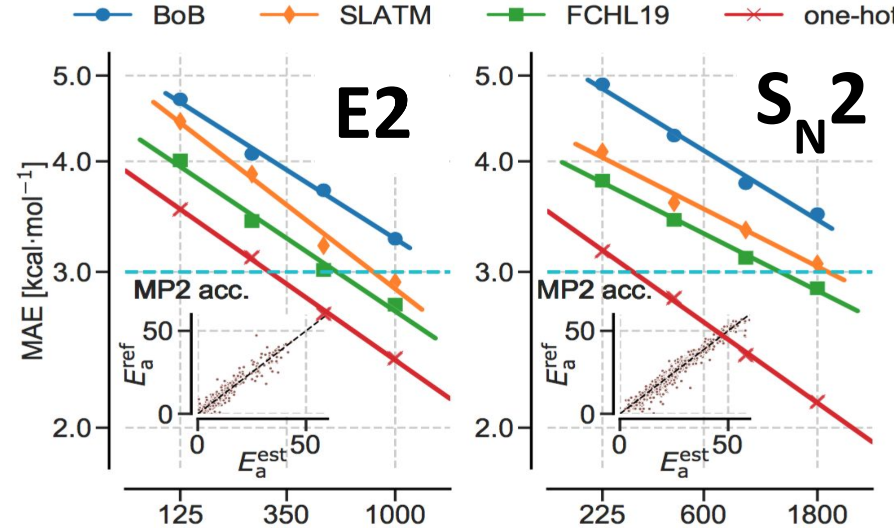


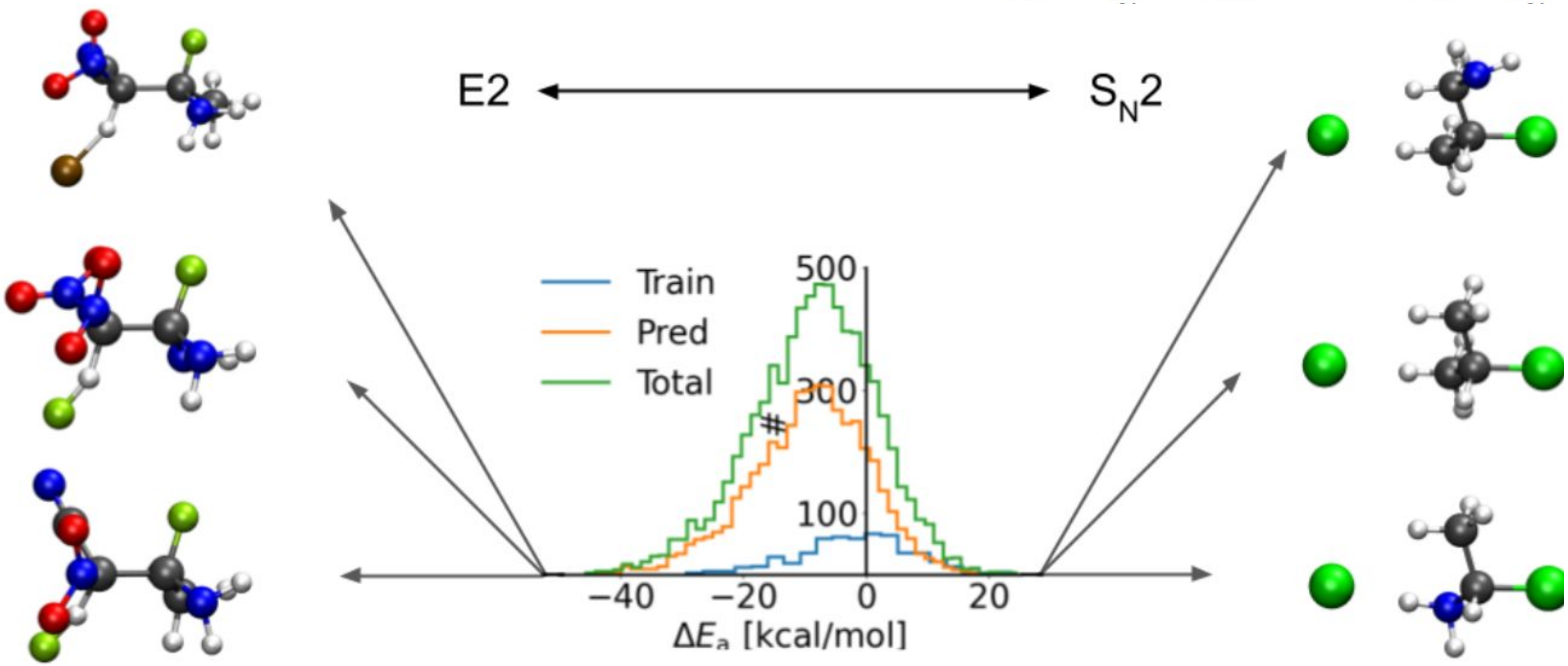
Which is more frequent, E2 or S_N2?

~11k R2B estimates: ¾ E2 vs ¼ S_N2

von Rudorff, Heinen, Bragato, OAvL, *MLST* (2020); Heinen, von Rudorff, OAvL, *J Chem Phys* (2021), *arxiv* (2020)

Reactant to barrier (R2B)

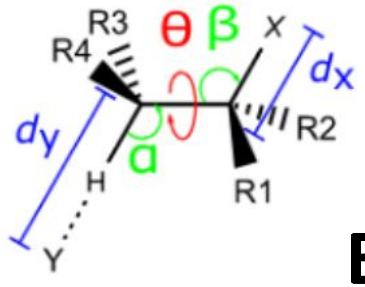




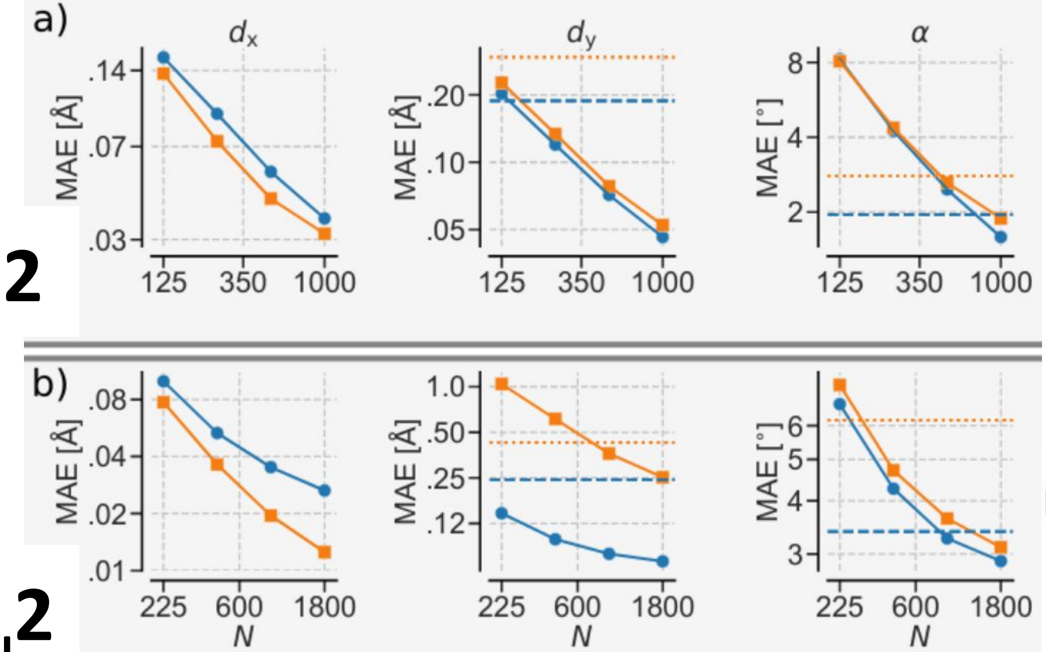
Heinen, von Rudorff, OAvL, *J Chem Phys* (2021), *arxiv* (2020)

Which is more frequent, E2 or S_N2?
~11k R2B estimates: ¾ E2 vs ¼ S_N2

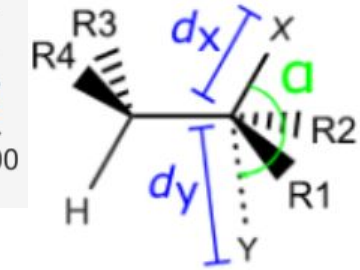
R2B

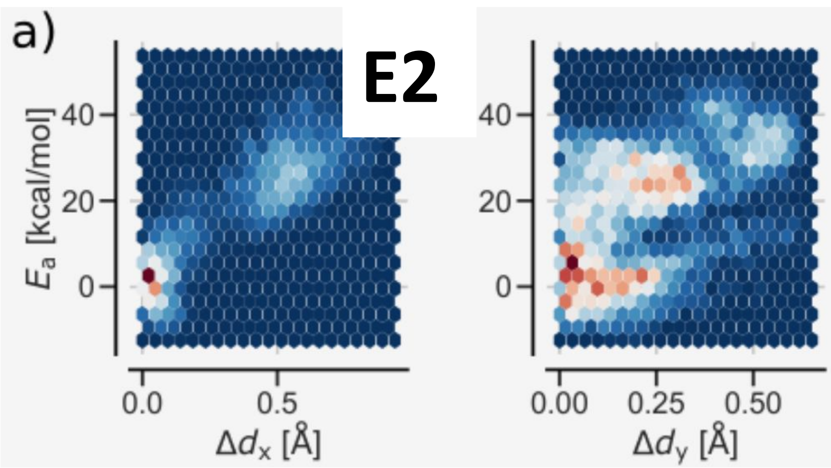


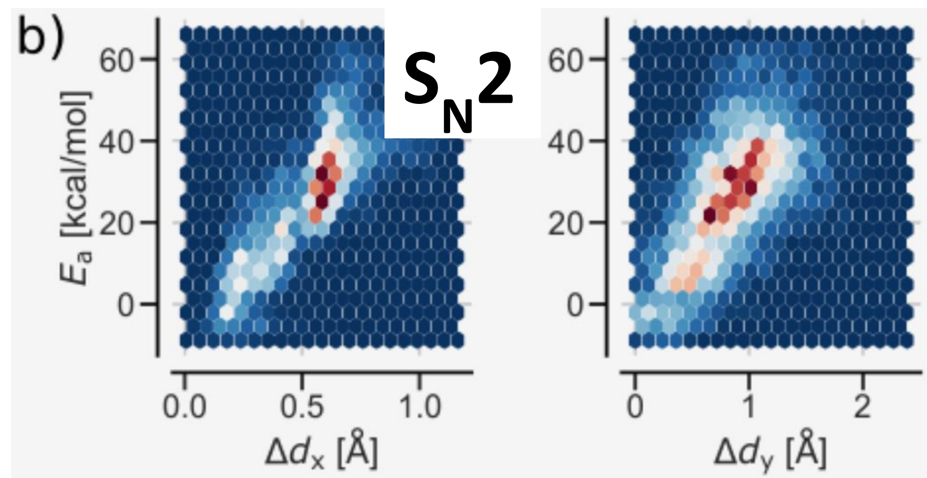
E2



S_N2



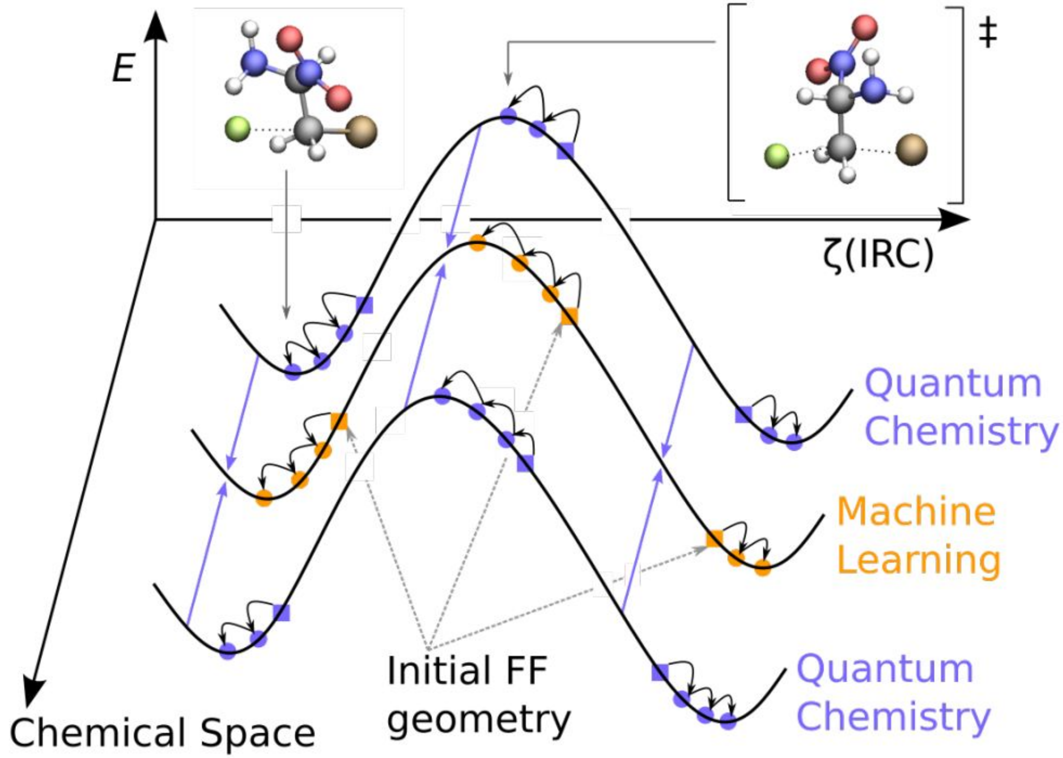


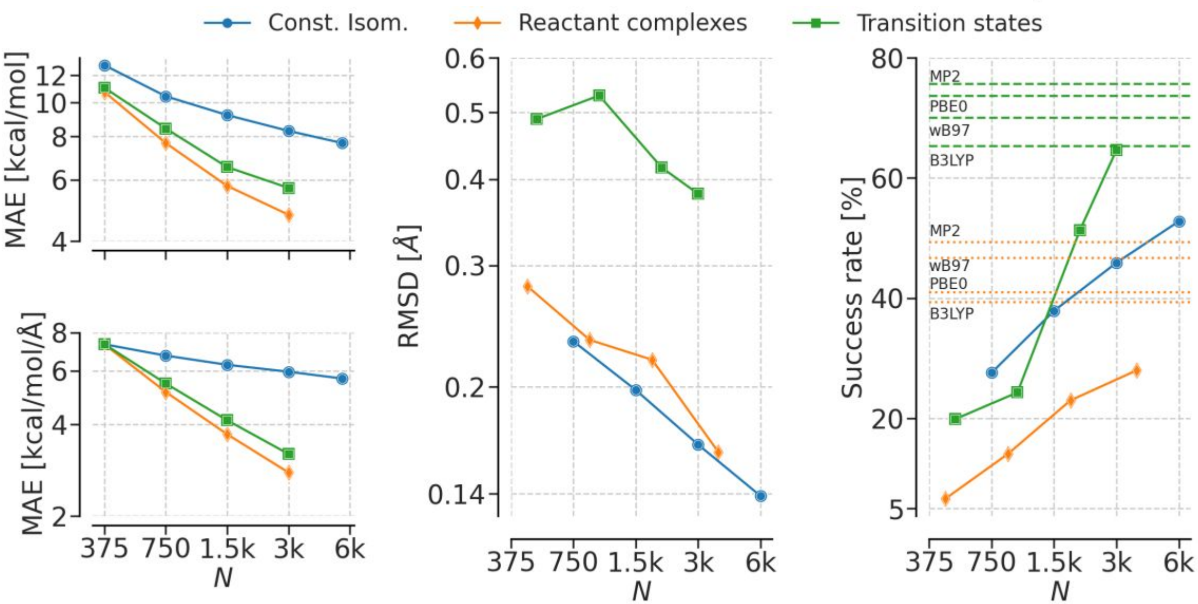


Hammond's postulate?
Yes for S_N2, No for E2

Heinen, von Rudorff, OAvL, *J Chem Phys* (2021), *arxiv* (2020)

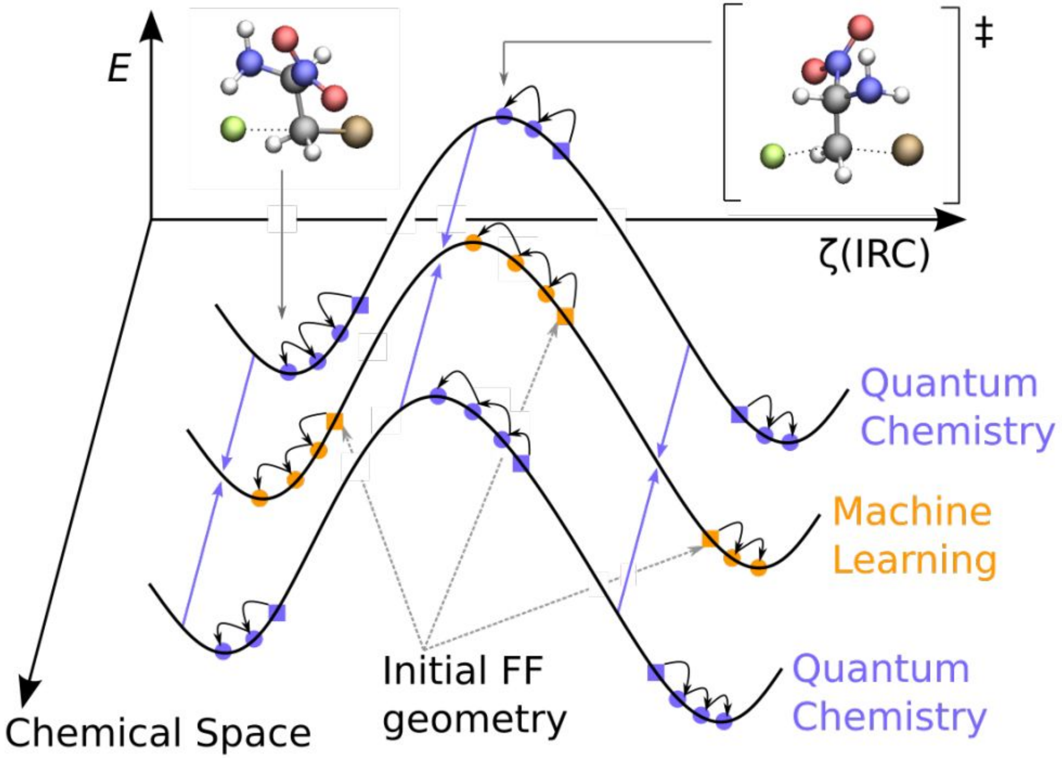
TS search S_N2

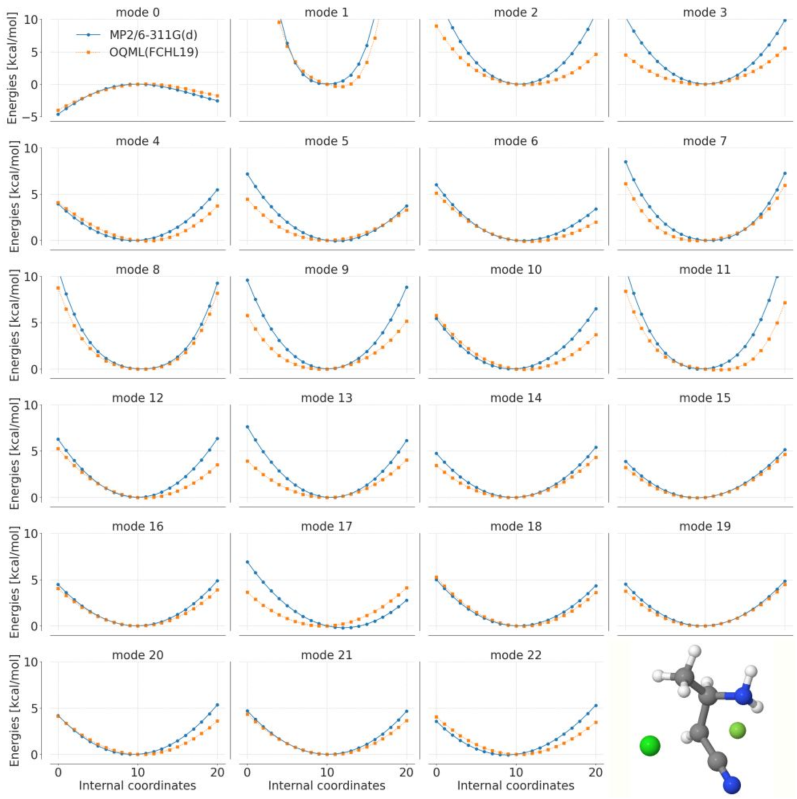




Heinen, von Rudorff, OAvL, *under revision in J Chem Phys* (2022), *arxiv* (2021)

TS search S_N2

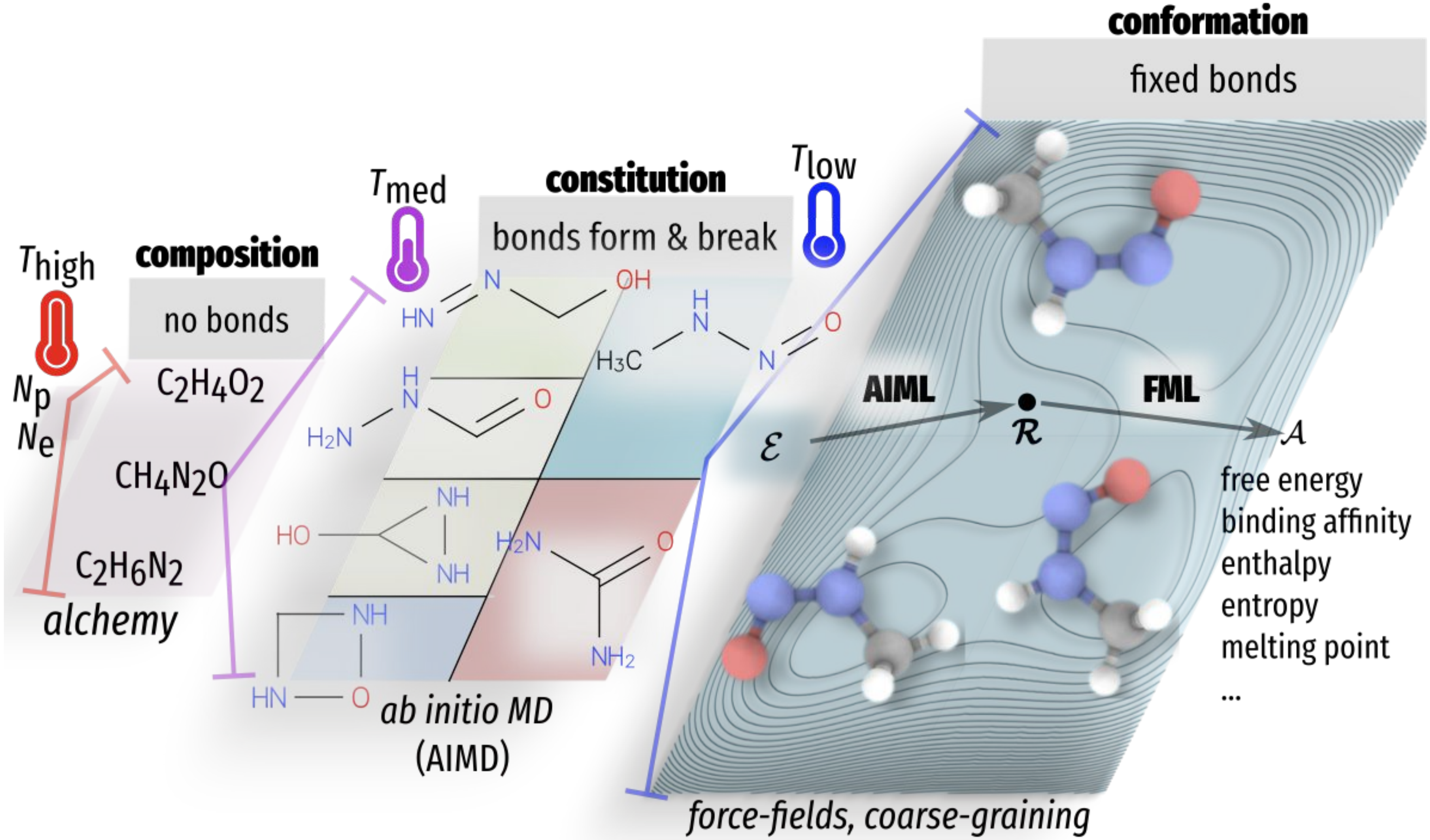




Heinen, von Rudorff, OAvL, *under revision in J Chem Phys* (2022), *arxiv* (2021)

Ab initio view on chemical space

Weinreich et al, *Ab initio machine learning of phase space averages, J Chem Phys* (2022)



Predicting

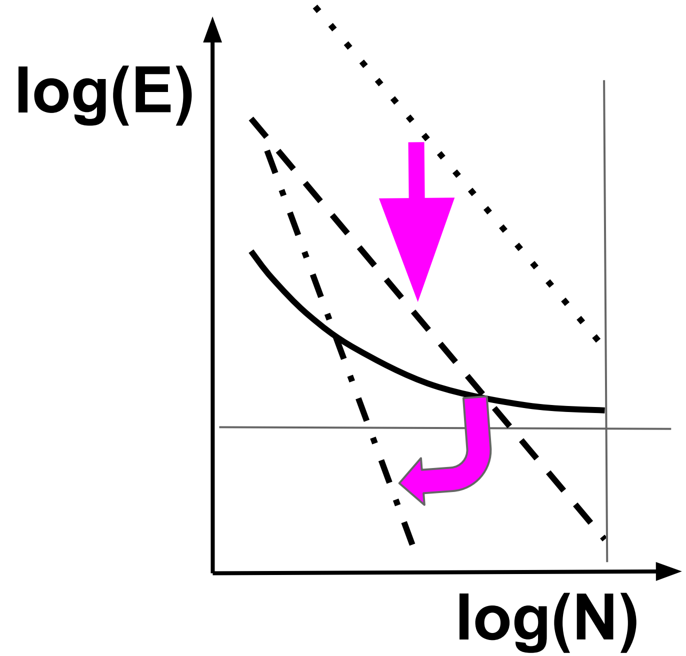
1. structure (*Nat Commun* 2021)
2. reactivity (*J Chem Phys* 2021, *J Chem Phys* 2022)
3. Free energy (*J Chem Phys* 2021, *J Chem Phys* 2022)

Free energy

$$p_q = p(\mathbf{X}_q) = \sum_i^N \alpha_i K(\mathbf{X}_i^{\text{train.}}, \mathbf{X}_q)$$

$$K(\mathbf{X}_i, \mathbf{X}_j) = \exp\left(-\frac{||\mathbf{X}_i - \mathbf{X}_j||_2^2}{2\sigma^2}\right)$$

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \cdot \mathbf{1})^{-1}\mathbf{p}$$



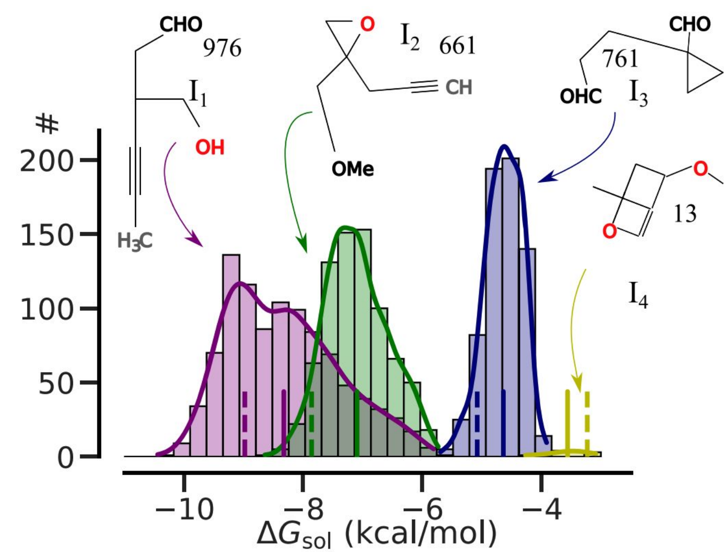


FIG. 4. Distribution of hydration free energies of conformers at 298K estimated by QML for four constitutional isomers ($C_7H_{10}O_2$). Corresponding FML estimates, as well as averages, are denoted by solid and dashed vertical lines, respectively. Insets show molecular graphs of corresponding constitutional isomers and number of conformers.

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

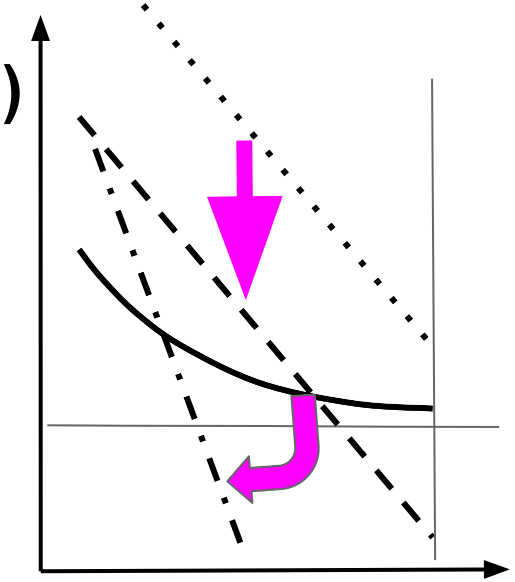
Free energy

$$p_q = p(\mathbf{X}_q) = \sum_i^N \alpha_i K(\mathbf{X}_i^{\text{train.}}, \mathbf{X}_q)$$

$$K(\mathbf{X}_i, \mathbf{X}_j) = \exp\left(-\frac{||\mathbf{X}_i - \mathbf{X}_j||_2^2}{2\sigma^2}\right)$$

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \cdot \mathbf{1})^{-1}\mathbf{p}$$

$$\bar{A} = \frac{\int A e^{-\beta H(q_1, q_2, \ldots q_M, p_1, p_2, \ldots p_N)} d\tau}{\int e^{-\beta H(q_1, q_2, \ldots q_M, p_1, p_2, \ldots p_N)} d\tau}$$



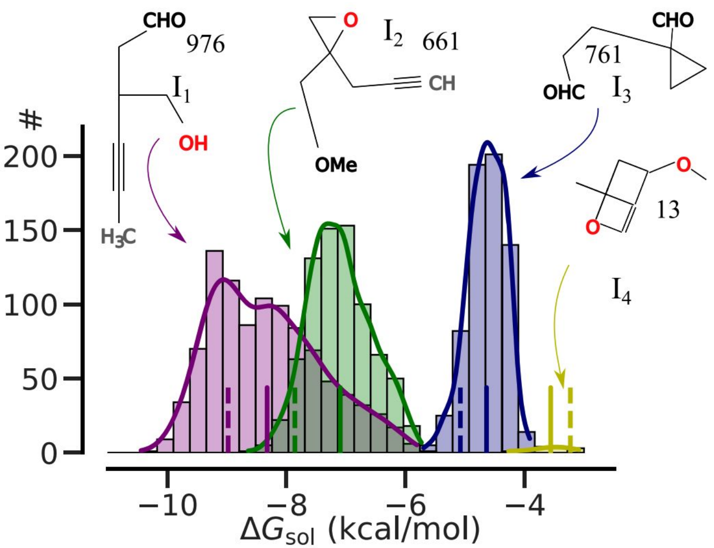


FIG. 4. Distribution of hydration free energies of conformers at 298K estimated by QML for four constitutional isomers ($C_7H_{10}O_2$). Corresponding FML estimates, as well as averages, are denoted by solid and dashed vertical lines, respectively. Insets show molecular graphs of corresponding constitutional isomers and number of conformers.

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

Free energy ML (FML)

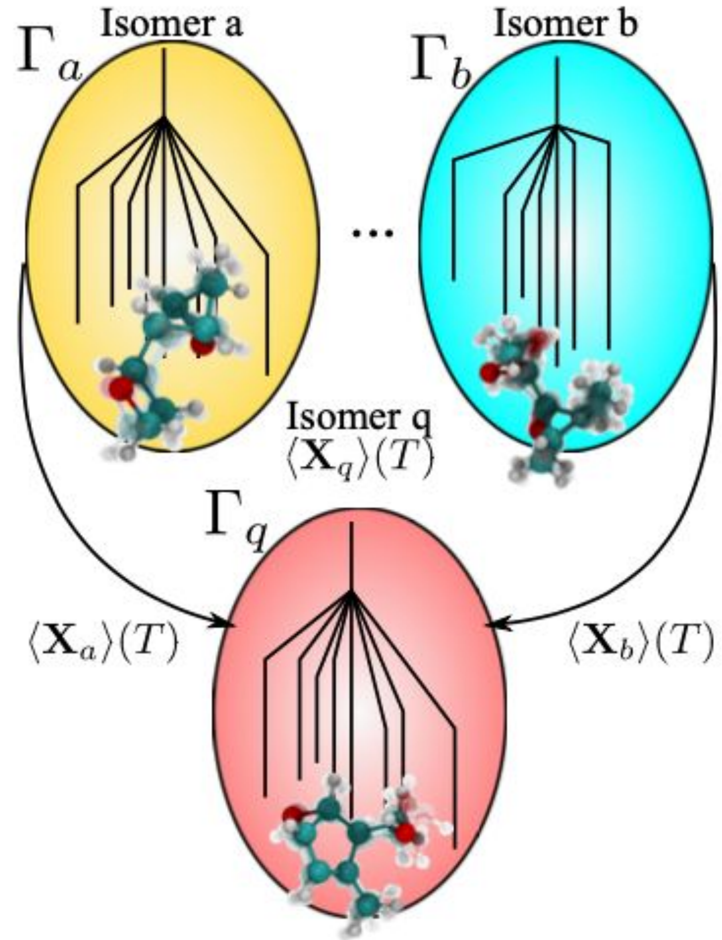
$$p_q = p(\mathbf{X}_q) = \sum_i^N \alpha_i K(\mathbf{X}_i^{\text{train.}}, \mathbf{X}_q)$$

$$K(\mathbf{X}_i, \mathbf{X}_j) = \exp\left(-\frac{||\mathbf{X}_i - \mathbf{X}_j||_2^2}{2\sigma^2}\right)$$

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \cdot \mathbf{1})^{-1}\mathbf{p}$$

$$\bar{A} = \frac{\int A e^{-\beta H(q_1, q_2, \ldots q_M, p_1, p_2, \ldots p_N)}\, d\tau}{\int e^{-\beta H(q_1, q_2, \ldots q_M, p_1, p_2, \ldots p_N)}\, d\tau}$$

$$\langle \mathbf{X} \rangle(T) = \frac{1}{Z} \int_{\mathbf{\Gamma}(T)} \mathbf{X}(\{\mathbf{r}_i\}) e^{-\beta E}\, d\mathbf{\Gamma}$$
$$\approx \frac{1}{N_s} \sum_i^{N_s} \mathbf{X}_i \ ,$$



Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

Training and Test: FreeSolv data base from:
G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, and D. L. Mobley, *J. Chem. Eng. Data* **62**, 1559 (2017).

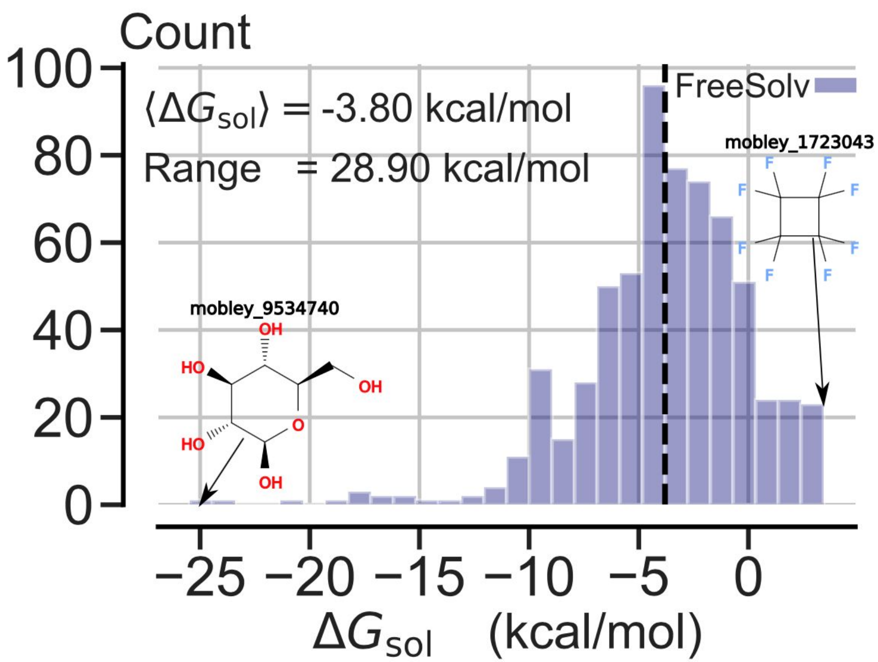


FIG. 3. Distribution of the experimental free energies of solvation provided by the FreeSolv database[1]. The mean value of the free energies is indicated by a vertical line. The most and least soluble molecules and their identifier in the FreeSolv database are also shown.

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

Training and Test: FreeSolv data base from:
G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, and D. L. Mobley, *J. Chem. Eng. Data* **62**, 1559 (2017).

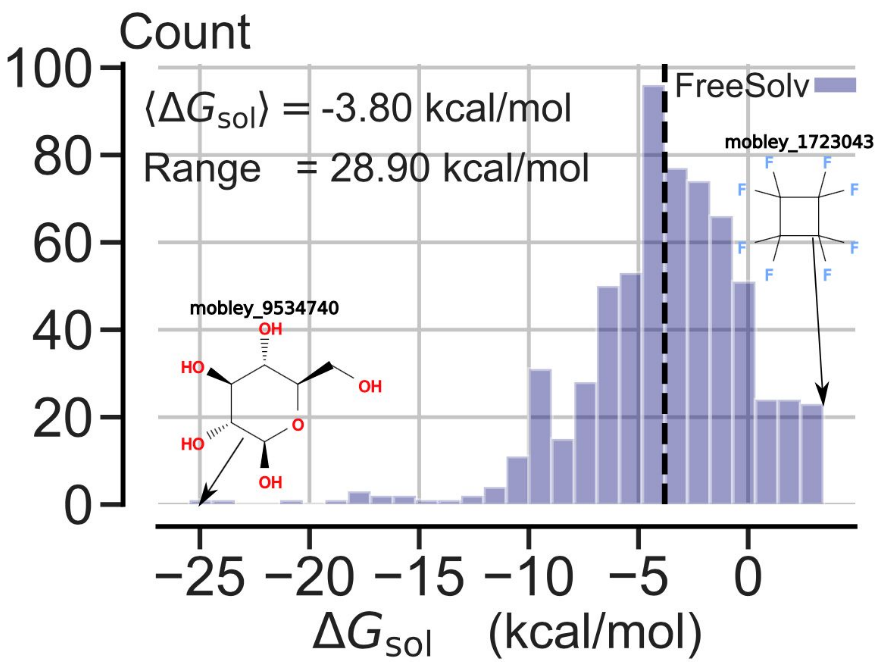
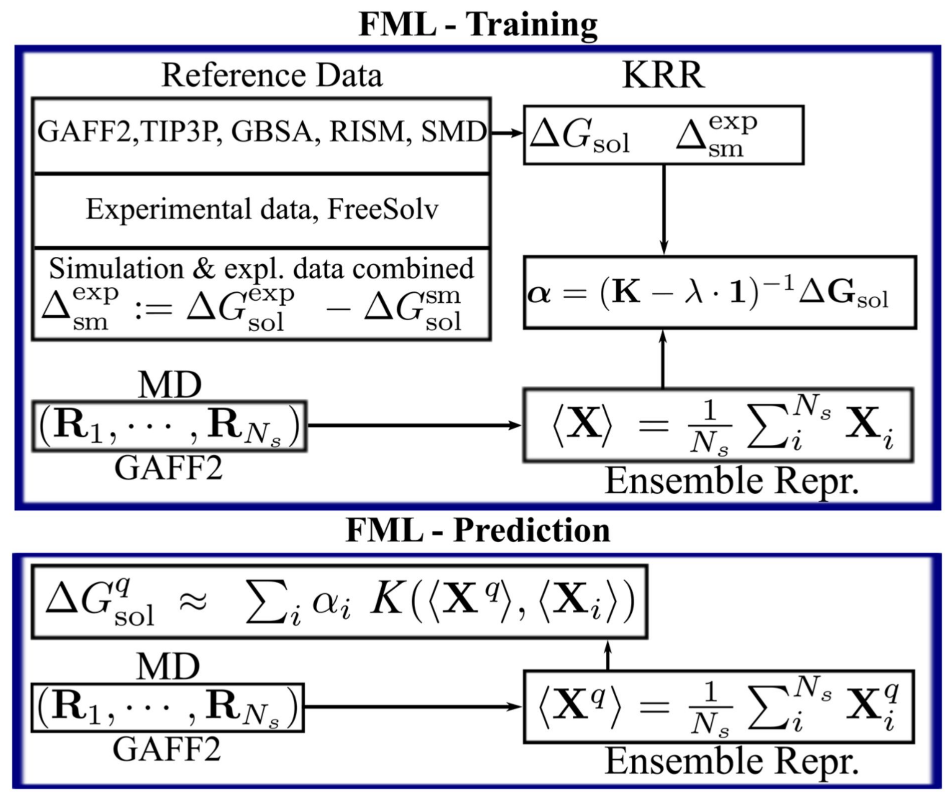


FIG. 3. Distribution of the experimental free energies of solvation provided by the FreeSolv database[1]. The mean value of the free energies is indicated by a vertical line. The most and least soluble molecules and their identifier in the FreeSolv database are also shown.



Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

Training and Test: FreeSolv data base from:
G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, and D. L. Mobley, *J. Chem. Eng. Data* **62**, 1559 (2017).

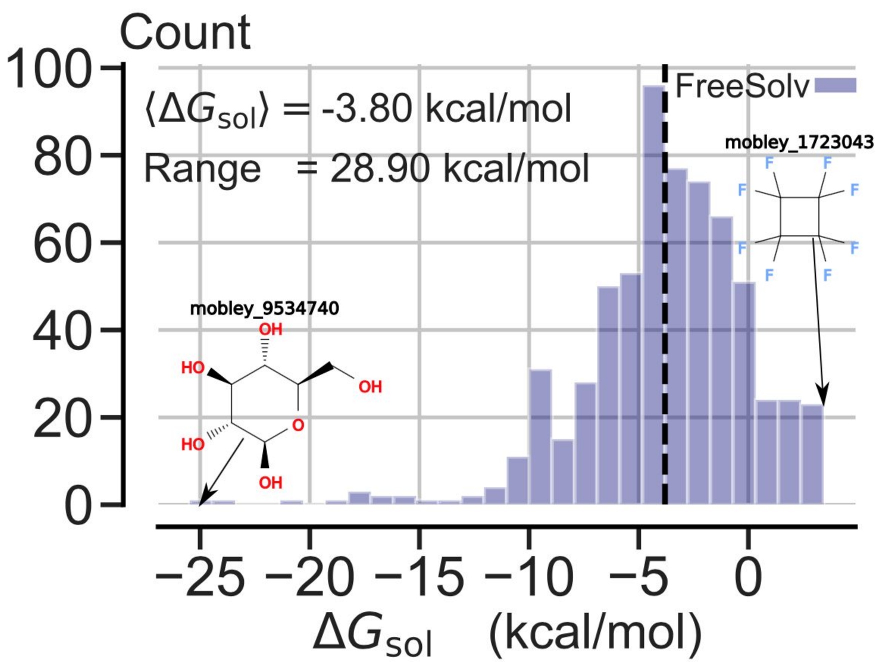
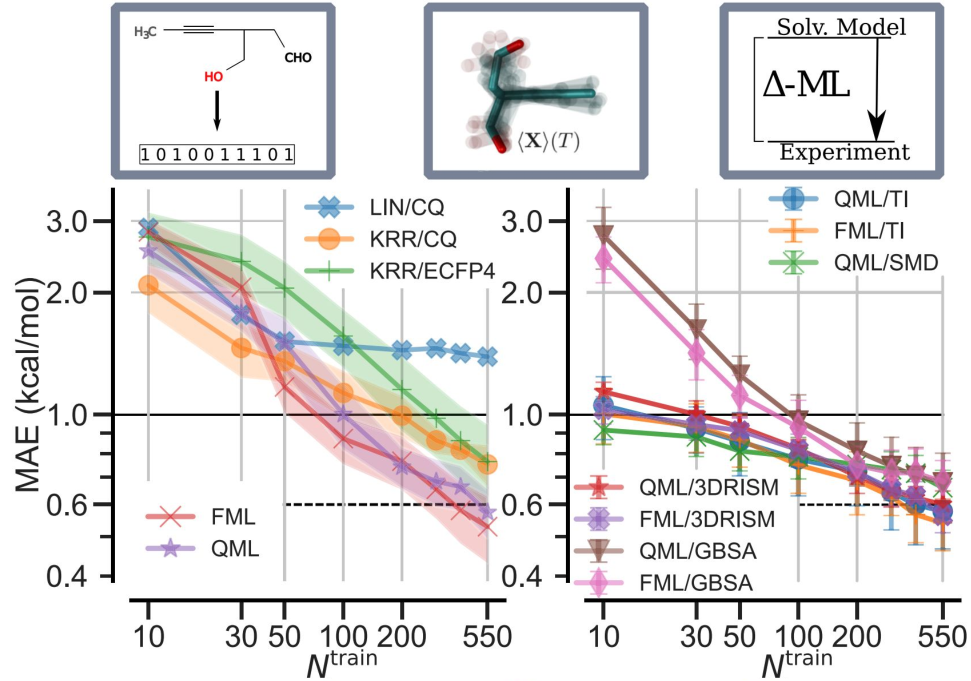
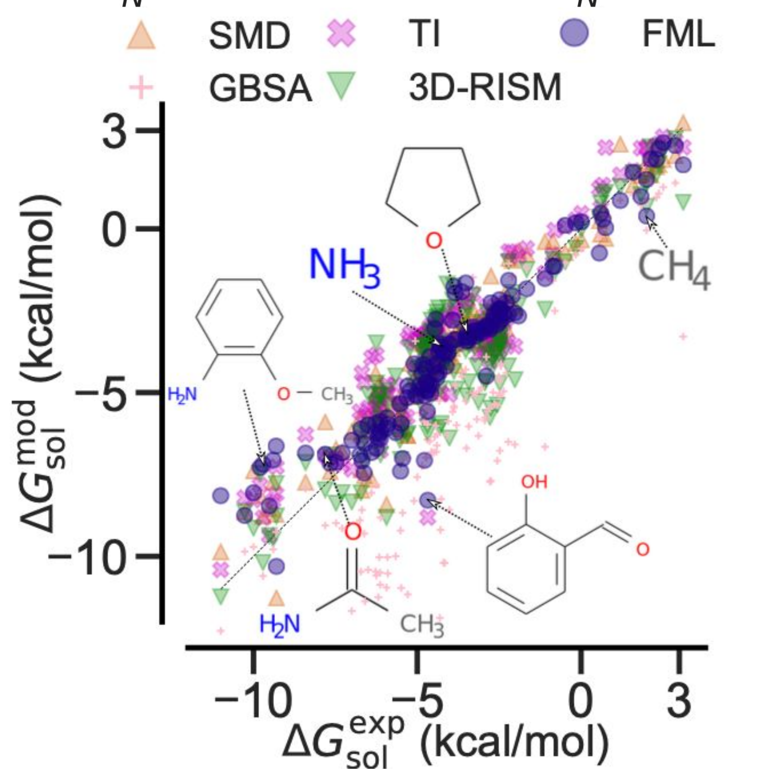


FIG. 3. Distribution of the experimental free energies of solvation provided by the FreeSolv database[1]. The mean value of the free energies is indicated by a vertical line. The most and least soluble molecules and their identifier in the FreeSolv database are also shown.





Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

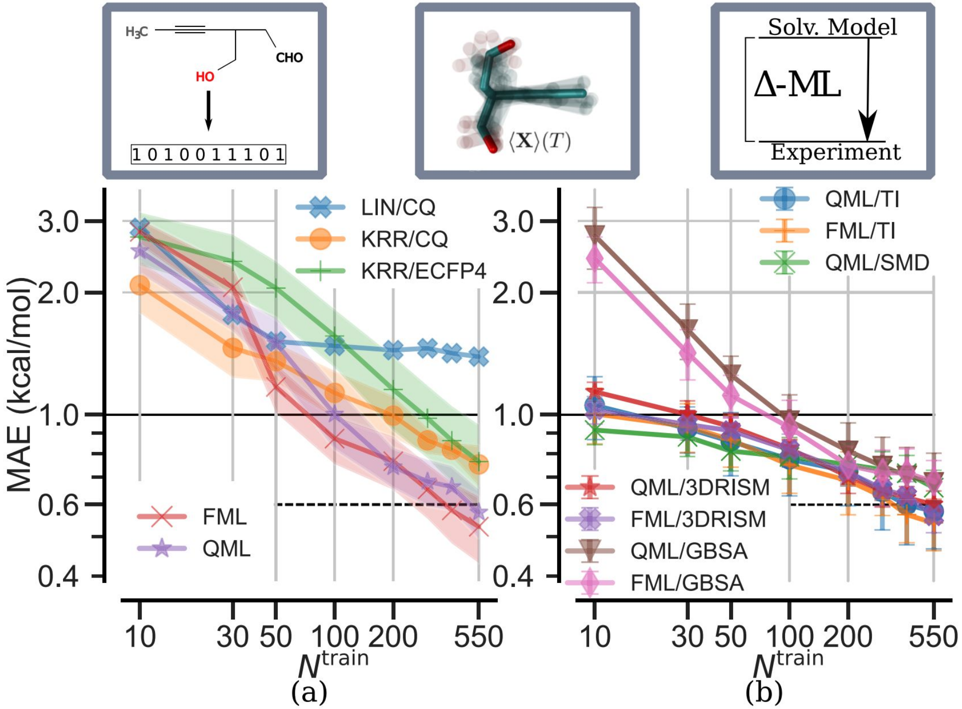


TABLE I. Comparison of MAEs, Pearson's r, and estimated order of CPU time per solute prediction for various solvation models and FML. Approximate number of training molecules N^{train} needed for FML to reach the MAE of each respective method. The conversion factor from GPU to CPU $c = \frac{t_{CPU}}{t_{GPU}}$ may vary substantially between ~10 and 60 depending on hardware/MD code (here OpenMM[58]). g is the number of grid points for TI, typically ~10.

Class	Model	MAE (kcal/mol)	r	N^{train}	~CPU h/solute	References	Year of publication
DFT	SMD	0.61[14]	0.96	400	*High*	8 and 71	2009
	COSMO	1.94	0.90	20	10^{-1}	9, 10, and 72	1993
	COSMO-RS	0.52[78]	0.91	...	10^{-1}	11, 72, 76, and 78	2000
	DCOSMO-RS	0.94[78]	0.87	100	10^{-1}	79	2006
FF	3D-RISM	0.99[14]	0.90	100	10^{-1}	12 and 13	1998
	TI	0.93[55]	0.94	100	$g \times c \times 10^{-1}$	55	2017
	GBSA	2.41	0.84	20	10^{-2}	63 and 64	2004
ML	FML	0.57	0.95	490	10^{-2}	This work	2020

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

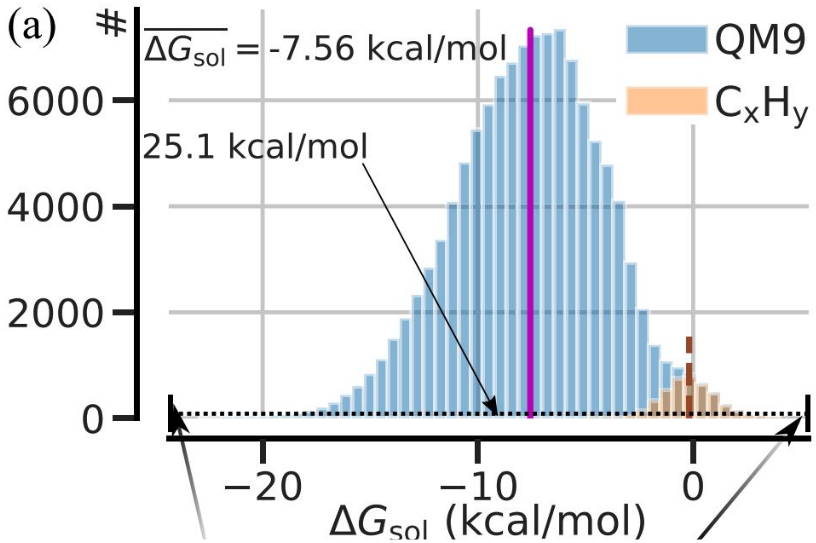


FIG. 7. Free energy distribution of 116k molecules predicted by FML in (a); small histogram corresponds to a subset of 4907 hydrocarbons with averages shown as solid and dashed lines, respectively. Four molecules with most negative (b) and positive (c) ΔG_{sol} in clockwise order. The mean of ΔG_{sol} for simple descriptors such as number of H-bond donors H_d (d), NH/OH groups (e), oxygen atoms O in stoichiometry $C_{n_C}H_{n_H}O_{n_O}$ (f), and heavy atoms h (g).

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

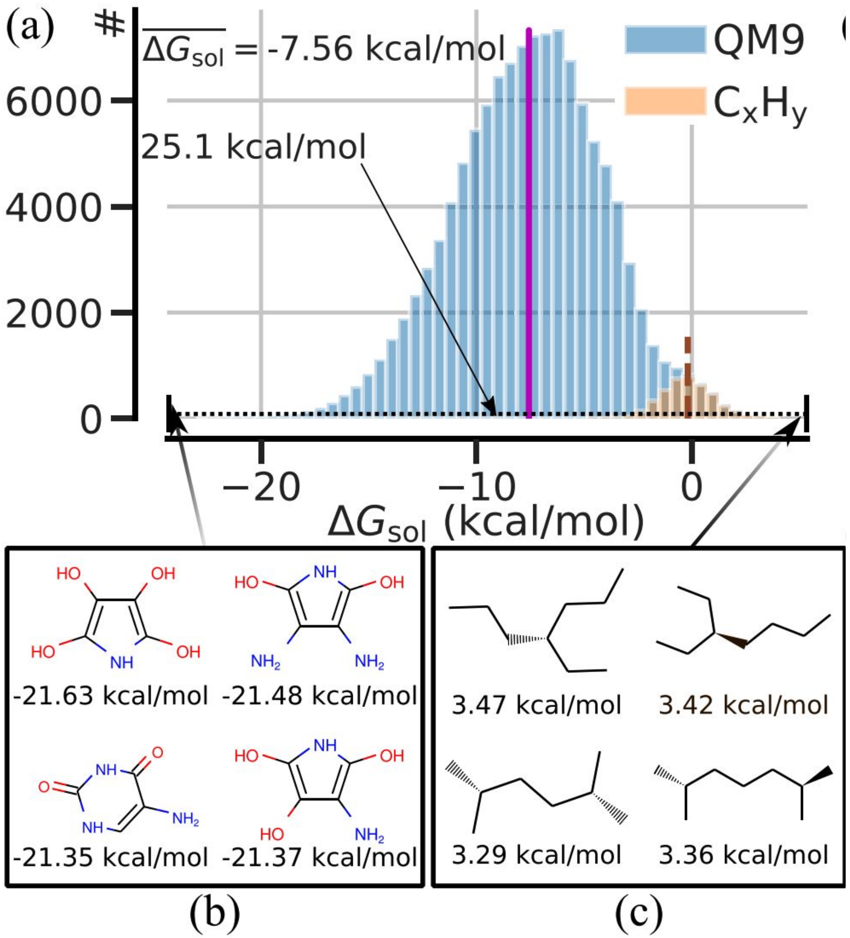


FIG. 7. Free energy distribution of 116k molecules predicted by FML in (a); small histogram corresponds to a subset of 4907 hydrocarbons with averages shown as solid and dashed lines, respectively. Four molecules with most negative (b) and positive (c) ΔG_{sol} in clockwise order. The mean of ΔG_{sol} for simple descriptors such as number of H-bond donors H_d (d), NH/OH groups (e), oxygen atoms O in stoichiometry $C_{n_C} H_{n_H} O_{n_O}$ (f), and heavy atoms h (g).

Weinreich, Browning, OAvL, *J Chem Phys* (2021)

FML

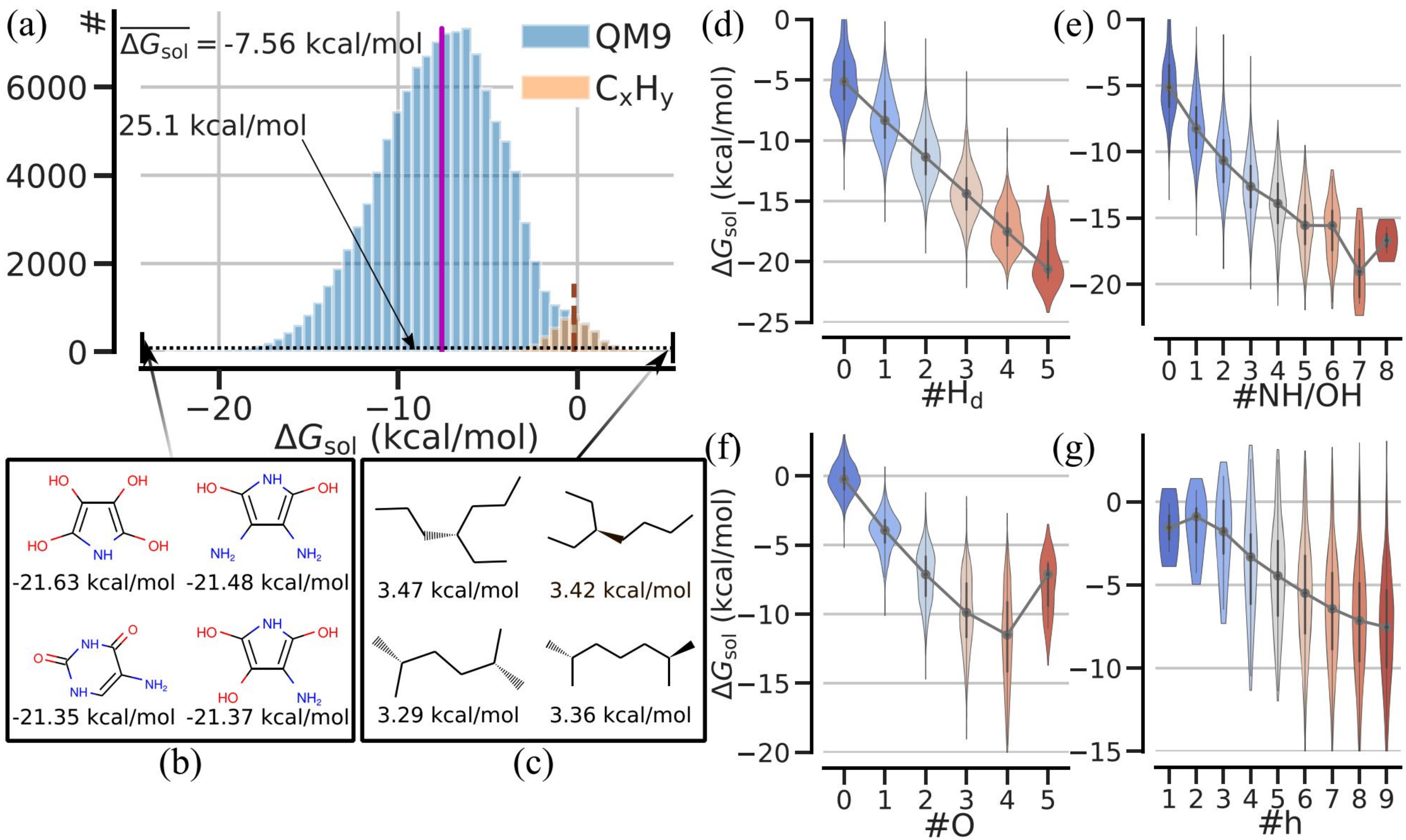


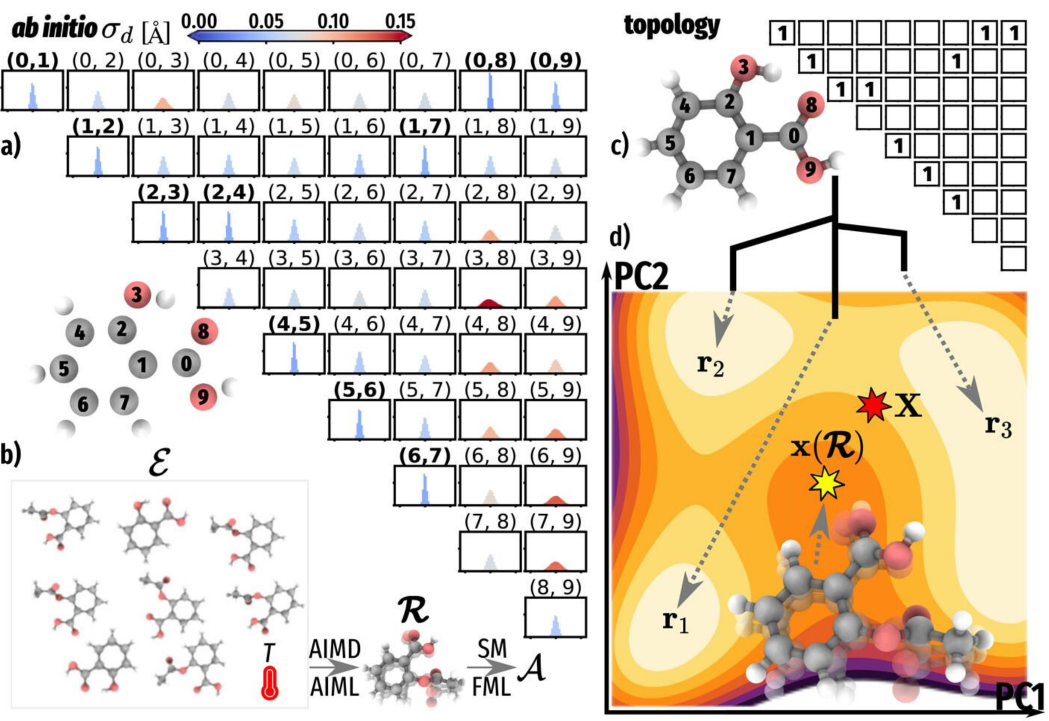
FIG. 7. Free energy distribution of 116k molecules predicted by FML in (a); small histogram corresponds to a subset of 4907 hydrocarbons with averages shown as solid and dashed lines, respectively. Four molecules with most negative (b) and positive (c) ΔG_{sol} in clockwise order. The mean of ΔG_{sol} for simple descriptors such as number of H-bond donors H_d (d), NH/OH groups (e), oxygen atoms O in stoichiometry $C_{n_C}H_{n_H}O_{n_O}$ (f), and heavy atoms h (g).

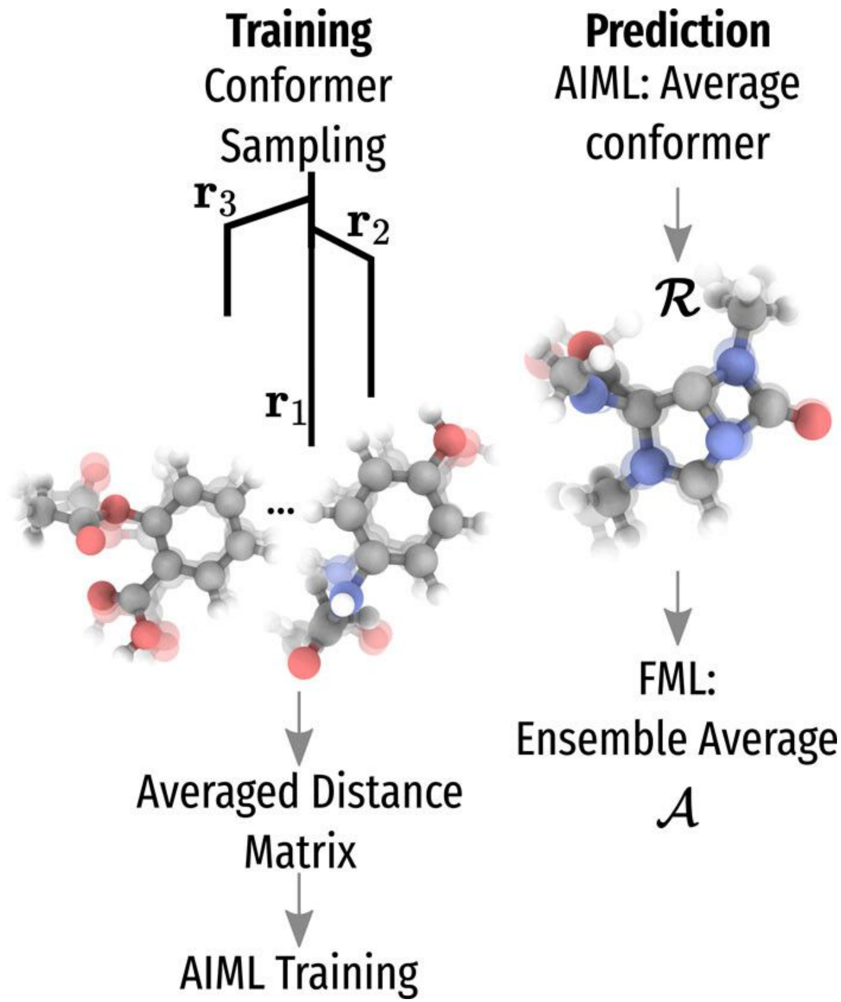
Weinreich, Browning, OAvL, *J Chem Phys* (2021)

AIML of <...>

$$\mathcal{A} = \frac{1}{Z}\int a(\mathbf{r},\mathbf{p})\, e^{-\beta E(\mathbf{r},\mathbf{p})}\, d\mathbf{r}\, d\mathbf{p} \approx \mathcal{A}^{\mathrm{ML}}(\mathcal{R})$$

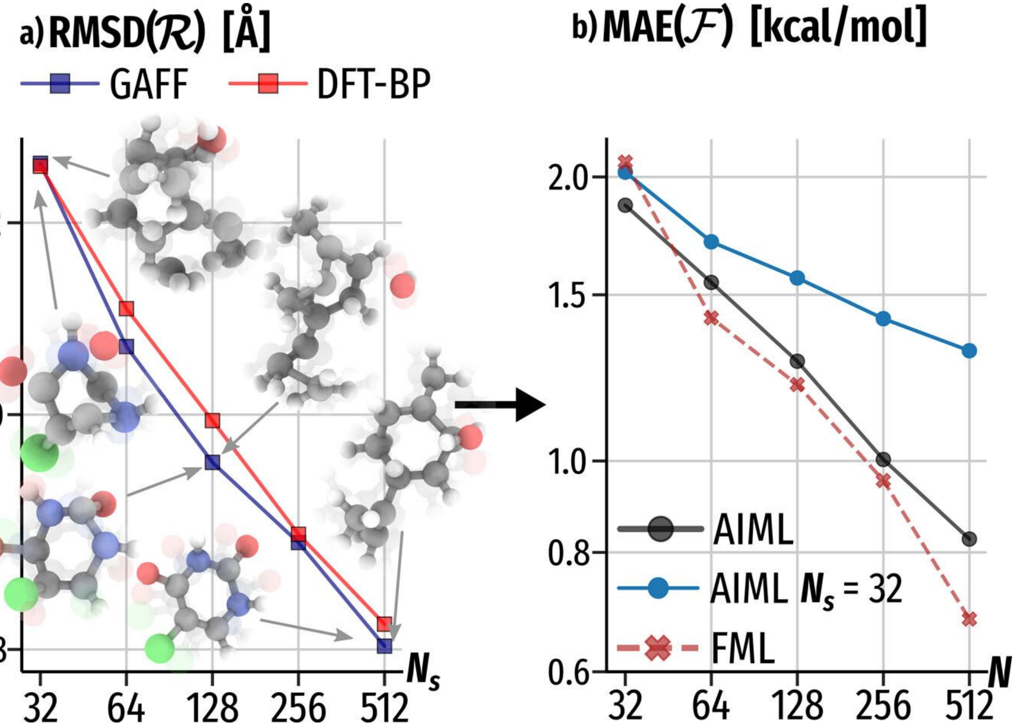
$$\mathcal{E} \rightarrow \mathcal{R} \rightarrow \mathcal{A}$$

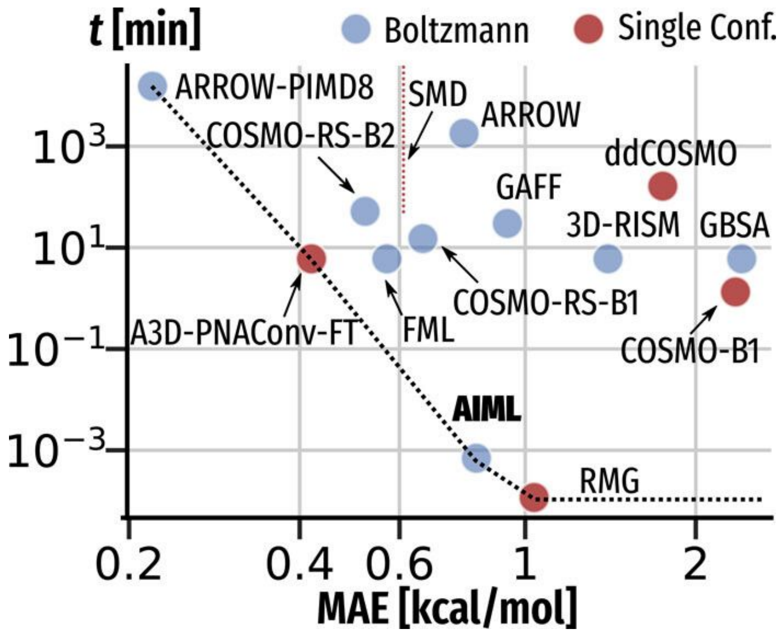




Weinreich, Lemm, von Rudorff, OAvL, *J Chem Phys* (2022)

AIML of <...>





Weinreich, Lemm, von Rudorff, OAvL, *J Chem Phys* (2022) https://www.chemspacelab.org

Twitter: @ProfvLilienfeld

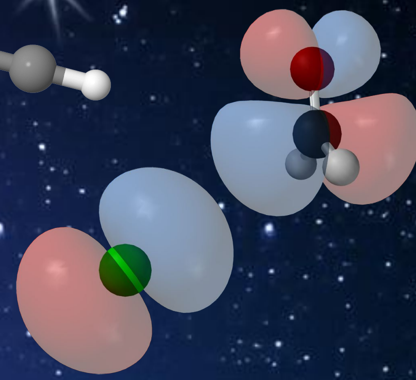
$$P^{\mathrm{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\mathrm{ref}}$$

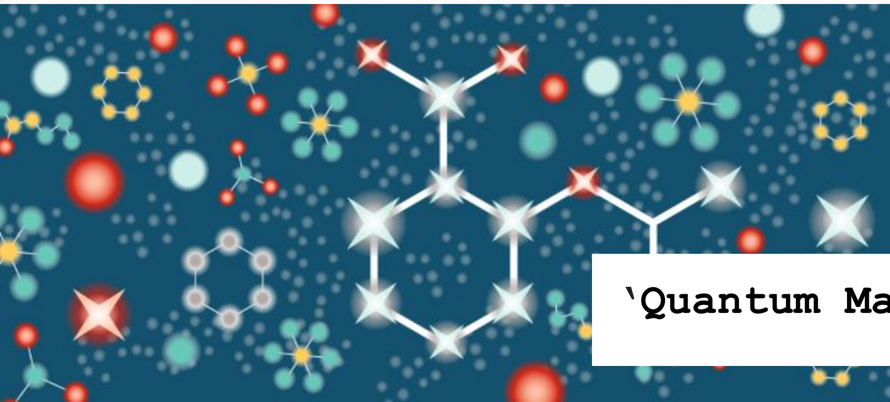
QML

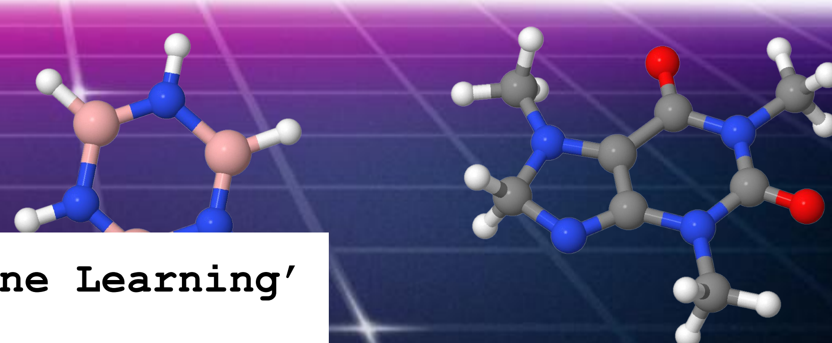
CHEMSPACELAB

$$\frac{dE}{d\lambda} = < \psi | \frac{d\hat{H}}{d\lambda} | \psi >$$

ALCHEMY



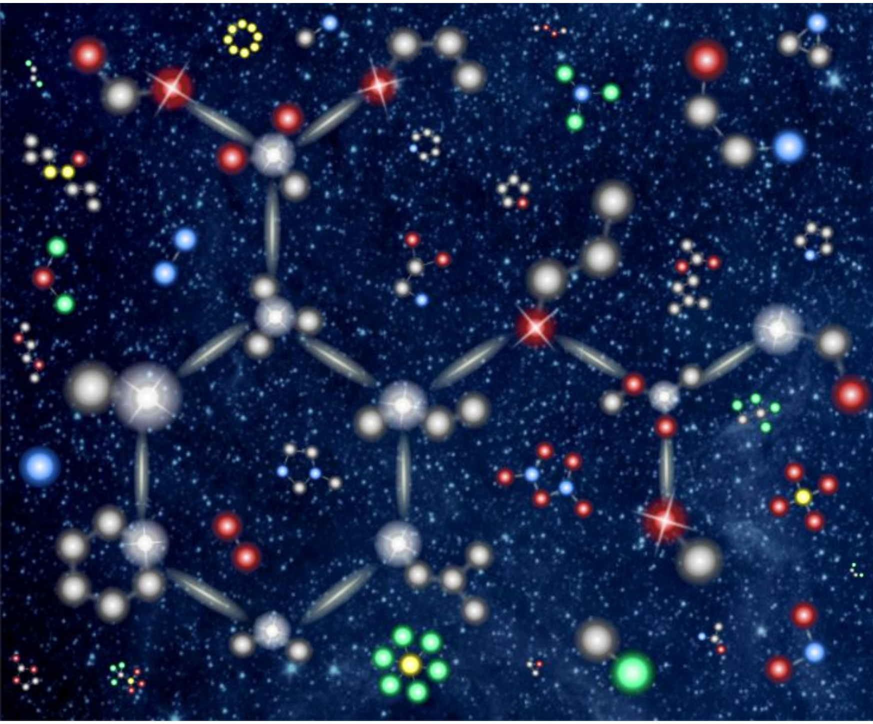


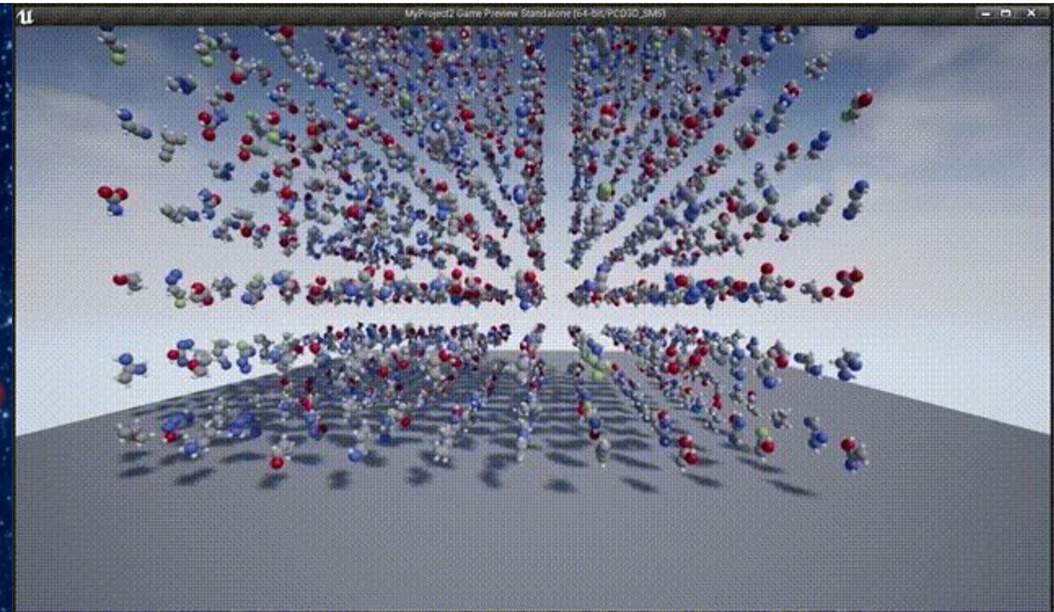


'Quantum Machine Learning'


BIG-MAP
FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION
Institute for Pure & Applied Mathematics
IPAM
University of California, Los Angeles
erc
Research at Google
MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
UNITED STATES AIR FORCE
U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

Ab initio view on chemical space

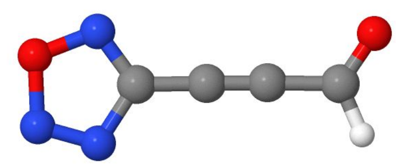


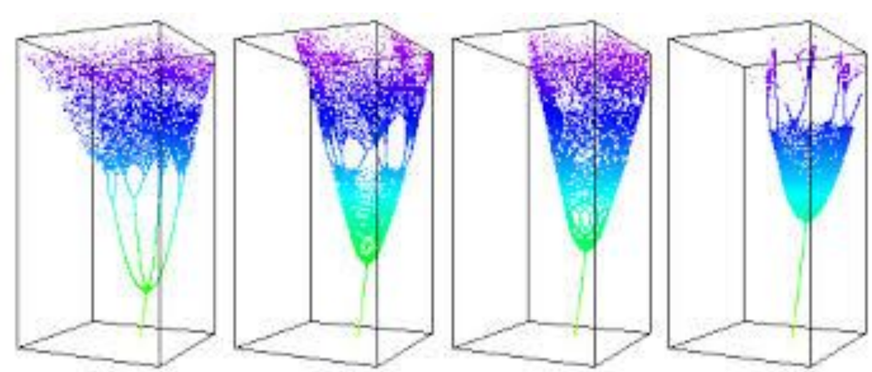


``Quantum machine learning using atom-in-molecule-based fragments selected on the fly'', Huang, von Lilienfeld,
Nature Chemistry (2020)

``Exploring **Chemical Compound Space** with Quantum-Based Machine Learning'', von Lilienfeld, Muller, Tkatchenko,
*Nature Chemistry Review*s (2020)

``Ab initio machine learning in **chemical compound space**", Huang, von Lilienfeld,
Chem Rev (2021)





→ **Quantum Compound Space** inherently
multi-scale: Composition/Constitution/Conformation
(reminiscent of fractal geometry)

``Exploring Chemical Compound Space with Quantum-Based Machine Learning'', von Lilienfeld, Muller, Tkatchenko, *accepted in Nature Chemistry Reviews* (2019)