Yale

Collective curvilinear coordinates in materials structure search and beyond

Reinhard J. Maurer

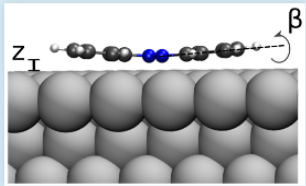
Department of Chemistry, Yale University

Machine Learning Meets Many-Particle Problems, IPAM 2016

Ab Initio simulation of surface chemistry

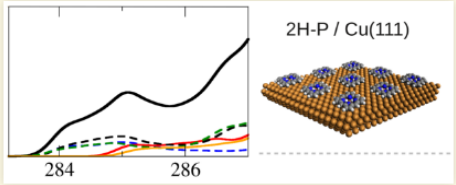
Accurate structure and energetics

- Density Functional Theory+beyond
- dispersion interactions
- finite-temperature effects
- Level alignment at interfaces



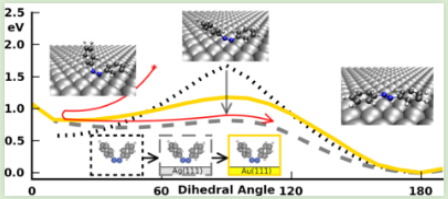
Excited States and Spectroscopy

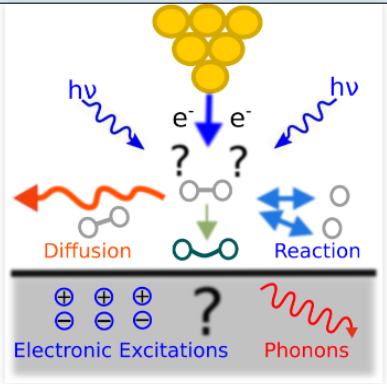
- excited states and couplings
- surface spectroscopy
- XPS,XAS,2PPE,SFG,TERS



Mechanistic Details / Reaction Dynamics

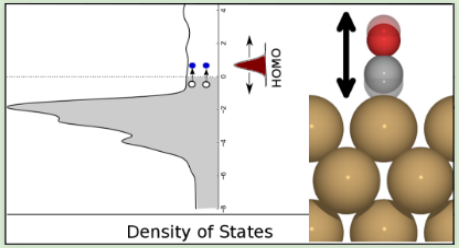
- general reaction mechanisms
- key design parameters





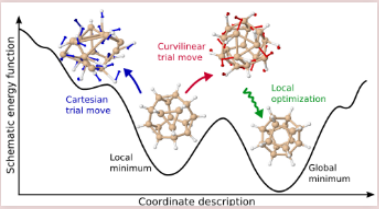
Energy Dissipation Nonadiabatic Dynamics

- role of nonadiabatic effects in dynamics on surfaces
- electron-phonon coupling



Configurational Complexity / Computational Scaling

- enable treatment of larger systems
- address high dimensional systems
- identify structures/pathways in reaction networks

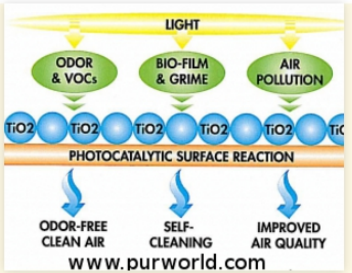


Controlled Surface Chemistry in Science and Technology

Yale

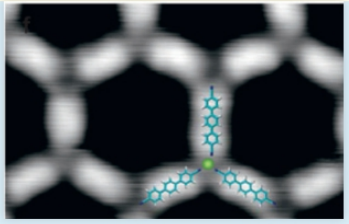
Surface Functionalization/Thin Films

- Smart Coated Surfaces e.g. Self-Cleaning Surfaces
- Photocontrolled Surface Function / Antimicrobial Surfaces
- Biocompatibility of Implants
- Device Physics of Hybrid Thin-Film Interfaces / Solar Cells



Energy Materials

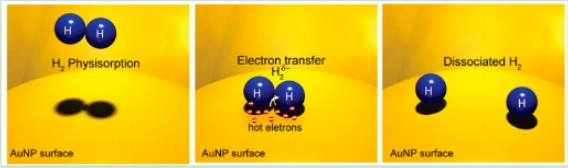
- Photonic Materials
- Conducting Polymers / Organic Crystals
- Surface Electrochemistry / Electron Transport
- Phase Change Materials
- Surface Self-Assembly in Hybrid Materials



Annu. Rev. Phys. Chem. 58, 375-407 (2007)

Catalysis and Surface Dynamics

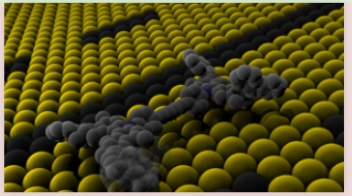
- Surface-Catalysed Photodegradation
- Photo- and Electron enhancement in Catalysis
- Energy Dissipation and Heat Transport
- Hot-electron Chemistry



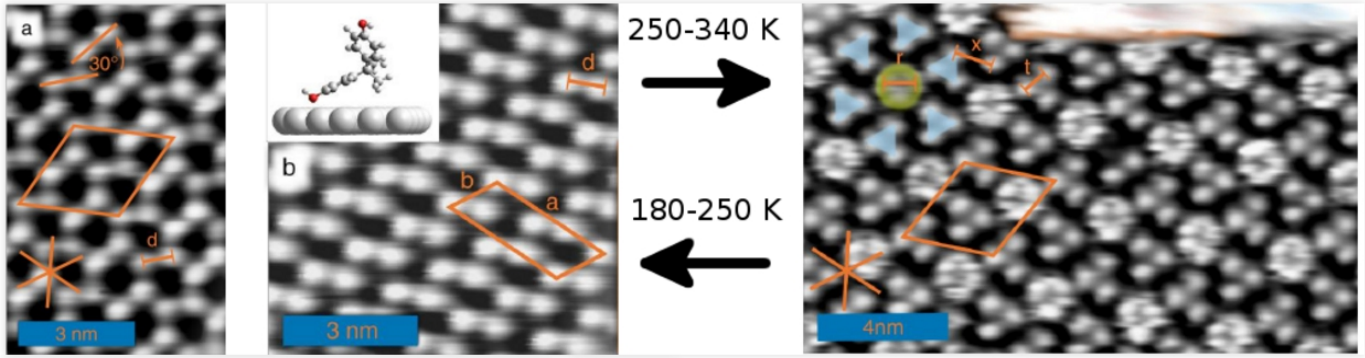
Nano Lett., 13, 240-247 (2013)

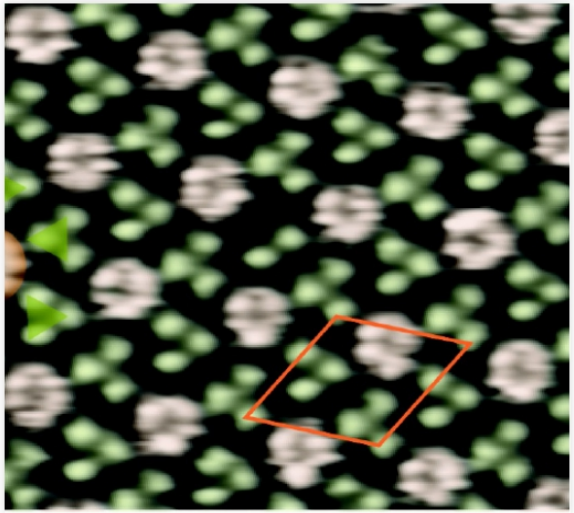
Nanotechnology

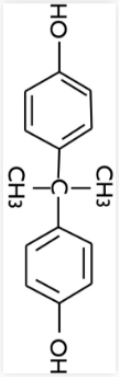
- Metallic Clusters attached to Proteins/Viruses
- Metal Nanoclusters for field-enhanced radiation treatment
- Nanomotors and Molecular Switches



Example 1: Bisphenol A on Ag(111)







Interesting questions for Ab Initio simulation

- structural models for different phases
- role of finite temperature on phase stability
- mechanistic model for high T phase

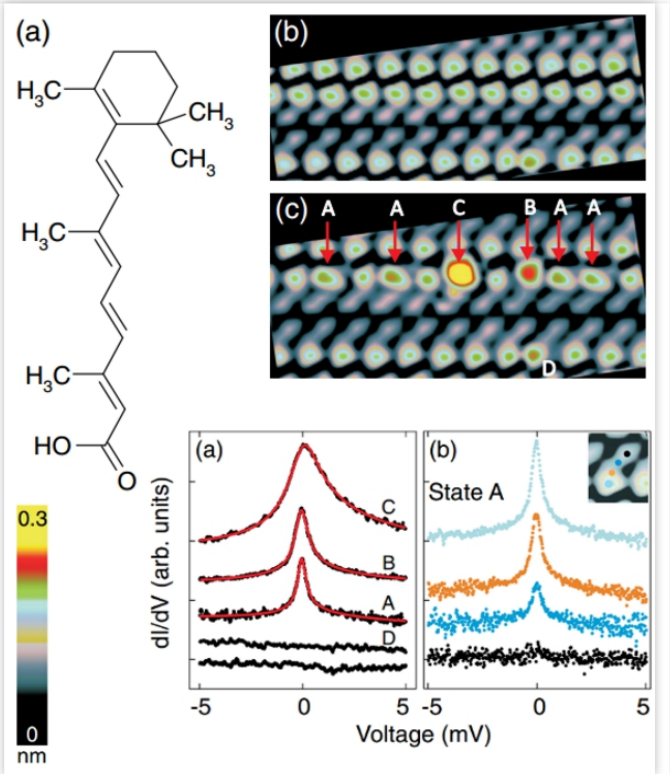
Challenges to Ab Initio simulation

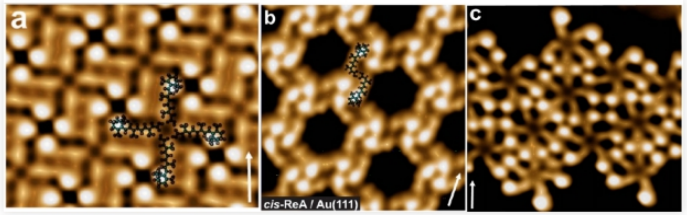
- computational cost, size of system
- covalent and vdW interactions
- incorporate finite-temperature effects
- tough with DFT and co

Characterization by STM, XPS, NEXAFS: Group of Johannes Barth, TU Munich

Lloyd et al., "Dynamics of Spatially Confined Bisphenol A Trimers in a Unimolecular Network on Ag(111)", Nano Lett., 16, 1884-1889 (2016)

Example 2: Molecular Switching of Retinoic Acid on Au(111)





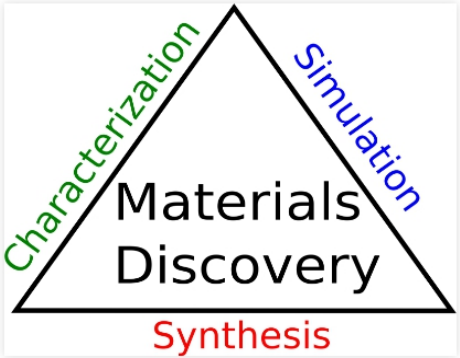
- switching: changes in geometry and electronic structure
- A,B,C are magnetic states
- spin impurity gives rise to Kondo resonance
- magnetic states are stable over many hours

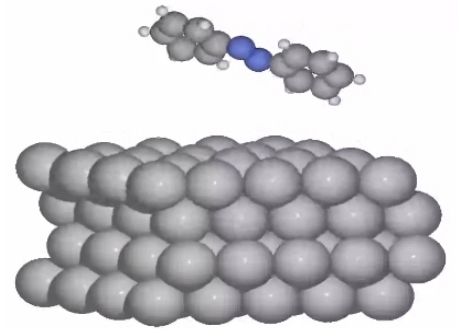
experiment by group of Richard Berndt, University of Kiel (Germany)

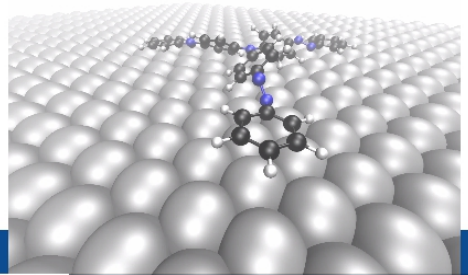
Karan, Wang, Robles, Lorente, Berndt, JACS 135, 14004-14007 (2013)

Karan, Li, Zhang, He, Hong, Song, Lü, Wang, Peng, Wu, Michelitsch, Maurer, Diller, Reuter, Weismann, Berndt , Phys. Rev. Lett. 116, 027201 (2016)

Challenges in Computational Materials Chemistry







Physical / Electronic structure description

- accurate intermolecular interactions, bond breaking/forming
- long-range dispersion interactions
- spectral/response properties
- highly computationally efficient methods needed

! more efficient electronic structure methods based on ML ?

! ML of spectral/response properties ?

Description of molecular (reaction) dynamics

- describe dynamics on picosecond time and on nanometer length scales
- statistical/thermal averaging of accessible properties
- coupled electron-nuclear quantum dynamics

! efficient and accurate many-body potentials from ML

! active learning: starting with ab-initio -> ML potential

Structure and composition of complex materials

- many different atomic species involved
- mixtures of isotropic/anisotropic bonding motifs
- different scales of structural deformations/changes

! compositional screening/prediction of complex materials

! configurational screening/prediction of complex materials/interfaces **!**

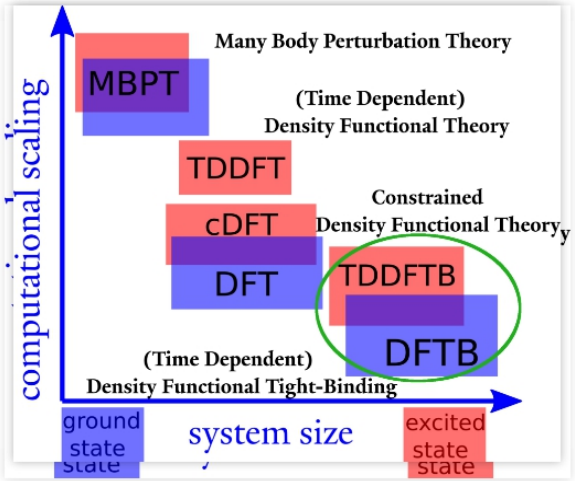
screening/prediction of chemical reaction space !

Density Functional Tight-Binding

efficient electronic structure using tight-binding (TB)

- 'tight-binding': precalculate interaction integrals for speed-up
- modern TB approaches: DFTB[1] and FIREBALL[2] methods
- factor 100 increase in length/time scales
- DFT+vdW^{surf} [3] $\rightarrow$ DFTB+vdW^{surf}

1: Elstner et al., Phys. Rev. B 58, 7260 (1998)
2: Lewis et al., Phys. Rev. B 64, 195103 (2001)
3: Tkatchenko, Scheffler, Phys. Rev. Lett. 102, 73005 (2009)



Density-Functional Tight-Binding (DFTB)

assume "near-sightedness" of interactions

locally confined wavefunctions

expand around superposition of atomic densities:

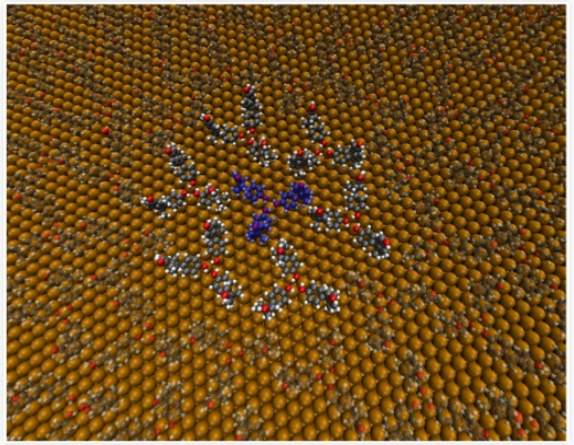
$n(\vec{r}) = n_0(\vec{r}) + \delta n(\vec{r})$

$E^{\mathrm{DFTB}} = H_0 + \frac{1}{2}\sum_{A,B}^{\mathrm{atoms}} \Delta q_a \Delta q_b + E_{\mathrm{rep}}(|R_A - R_B|)$

-suffers from same lack of dispersion interactions as DFT

+minimal basis, H_0 atomwise integrals are precalculated

-electron density never explicitly constructed



Dispersion-inclusive tight-binding approach

DFT+vdW: atomic polarizability $\propto$ eff. atomic volume

(TS, from elec. density)

$$\frac{\alpha_A}{\alpha_{\text{free}}} \approx \frac{V_A}{V_{\text{free}}}$$

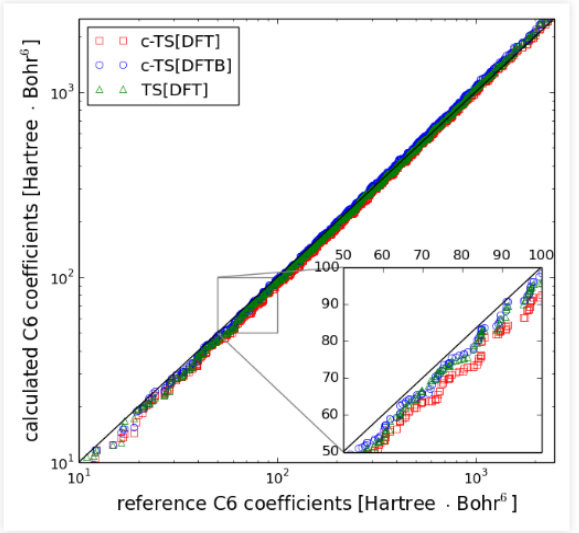
Tkatchenko, Scheffler, Phys. Rev. Lett. 102, 73005 (2009)

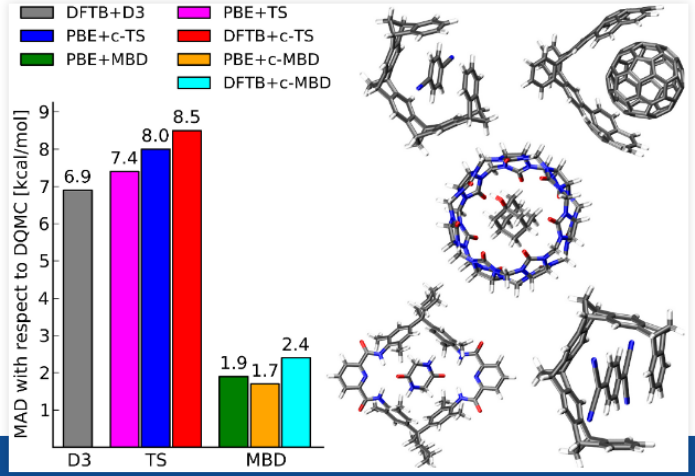
DFTB+vdW: atomic polarizabitlity $\propto$ degree of hybridization

(c-TS, from density matrix)

$$\frac{\alpha_A}{\alpha_{\text{free}}} \approx \frac{\sum_a f_a \sum_{i \in A} |c_i^a|^2}{Z_A}$$

rel. error on 817 intermolecular C_6 coefficients: 6.8% vs. 5.4%



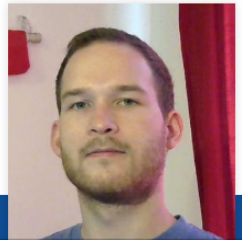


DFTB+vdW / DFTB+MBD

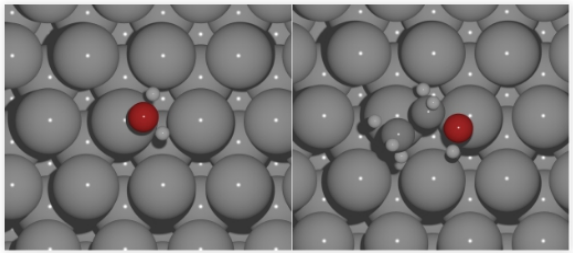
- yields accurate interaction energies for dimers, organic crystals and supramolecular complexes
- independent of DFTB parametrization
- what about molecules on metal surfaces?

MSc project: Martin Stöhr (Yale/TUM)

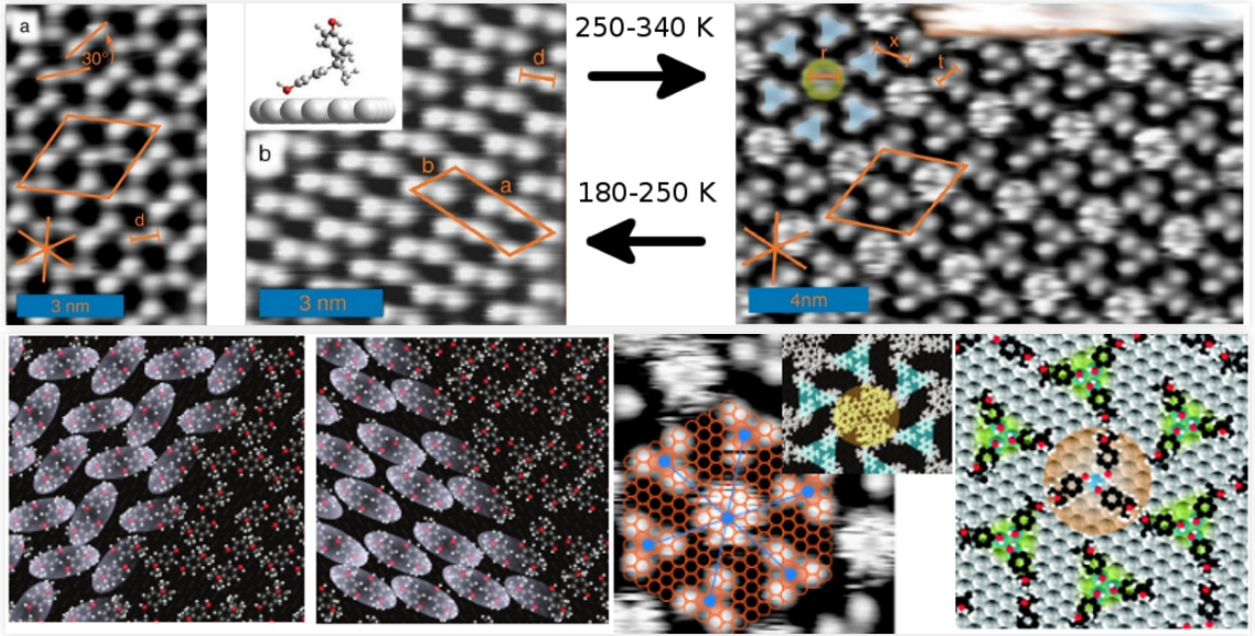
Stöhr, Michelitsch, Tully, Reuter, Maurer, J. Chem. Phys. 144, 151101 (2016)



Proof of Principle: Bisphenol A on Ag(111)

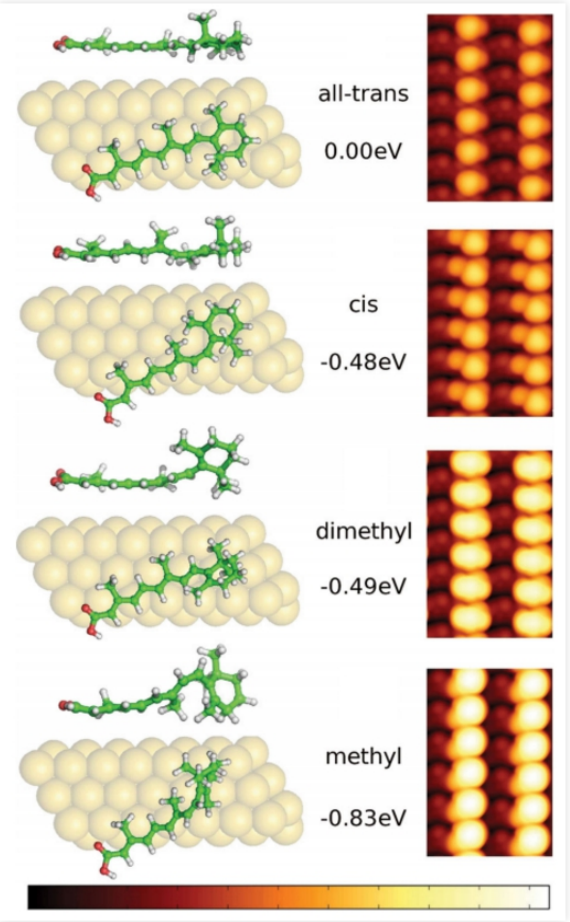


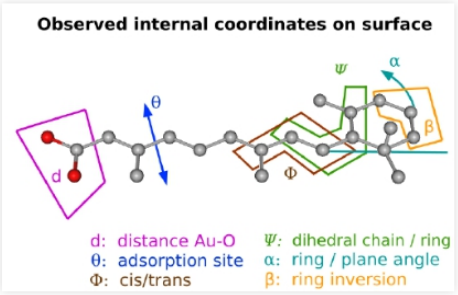
- DFTB parametrization: existing parameters for C,H,O and our own parameters for Ag
- interaction fit to EtOH and H_2O on Ag(111)



Lloyd et al., "Dynamics of Spatially Confined Bisphenol A Trimers in a Unimolecular Network on Ag(111)", Nano Lett., 16, 1884-1889 (2016)

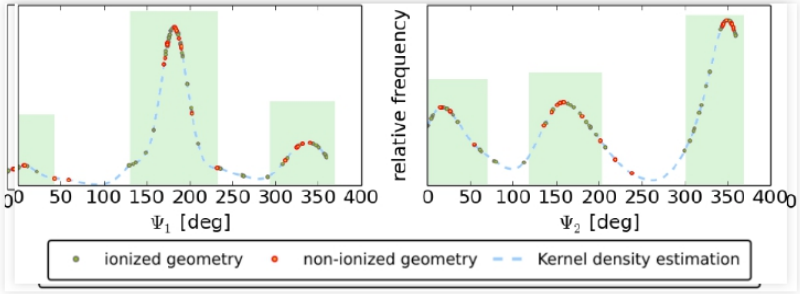
Molecular Switching of Retinoic Acid on Au(111): Simulation





MSc Thesis (U Graz/TUM):
Georg Michelitsch

- screening over 120 possible configurations and STM topographs
- Gaussian Kernel density estimation

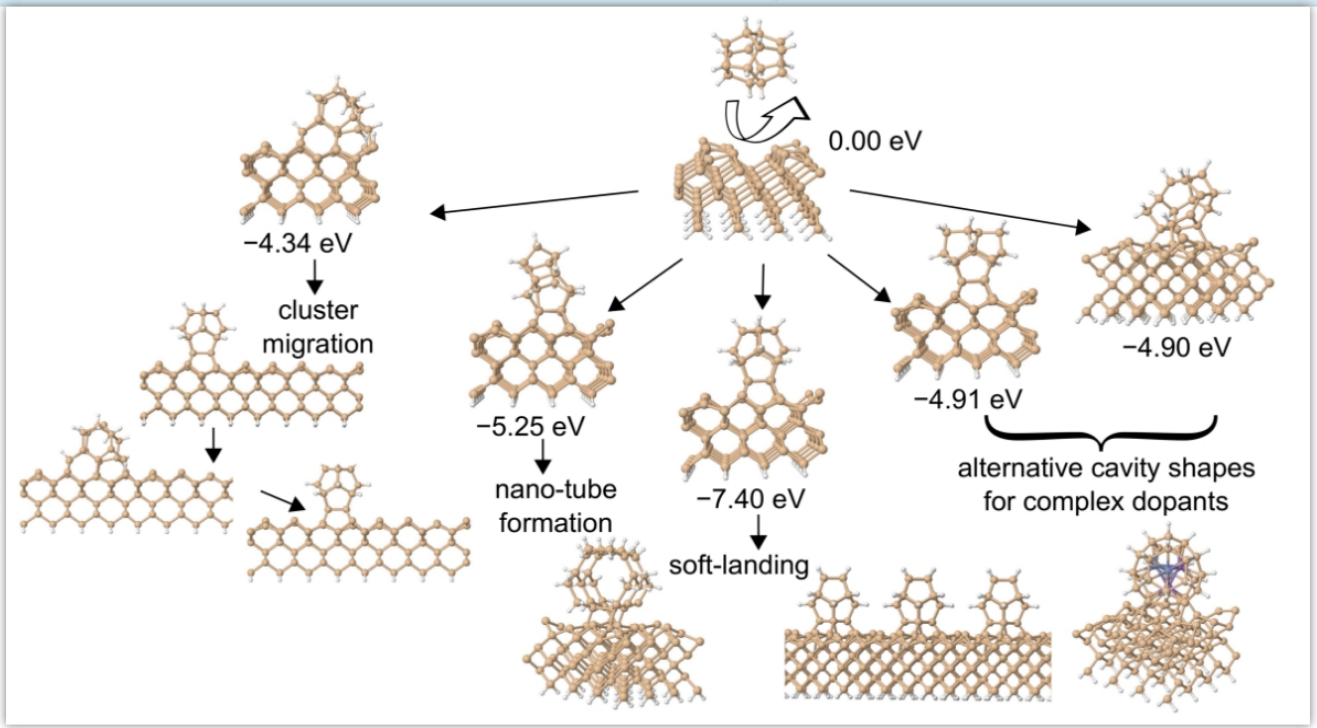


- idenify 4 classes of geometries
- DFT+vdW^{surf}: geometrical change and Kondo screening stabilize positive charge resonance state

Karan, Li, Zhang, He, Hong, Song, Lü, Wang, Peng, Wu, Michelitsch, Maurer, Diller, Reuter, Weismann, Berndt , Phys. Rev. Lett. 116, 027201 (2016)

Chemical reaction discovery: soft-landing $Si_{16}H_{16}$ on a Si(001) surface

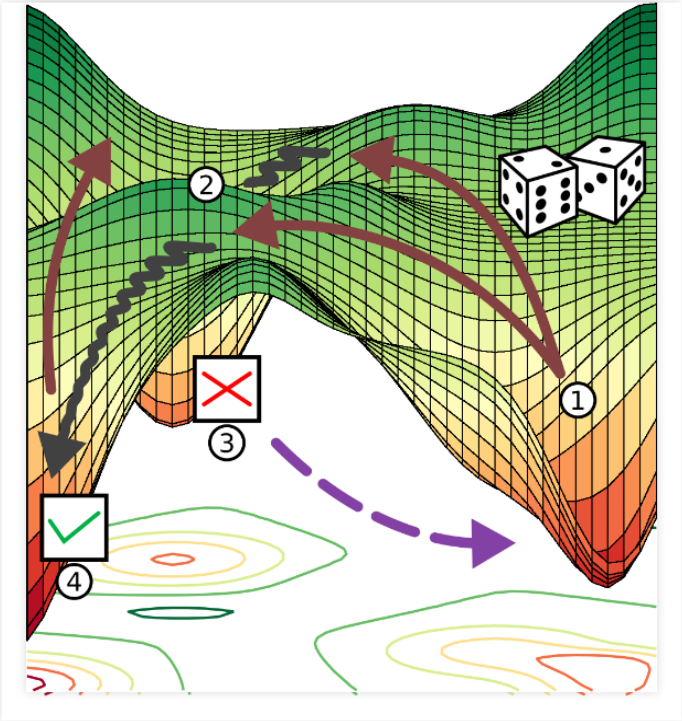
Materials structure search should identify all reaction avenues



Problem: Find all "chemically relevant" equilibrium structures and pathways

global optimization - unbiased structure search?

Finding local minima on high dimensional PES



Stochastic or heuristic sampling methods

Heuristic: genetic algorithms

Stochastic: basin hopping

Basin Hopping [1]

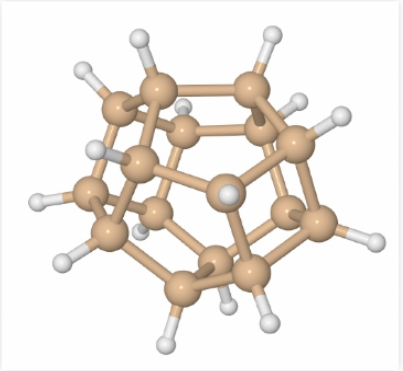
1. displace a geometry randomly by Δx
2. perform a local energy minimization: $\min\{E(\mathbf{X})\}$
3. calculate acceptance probability:
$$P(\Delta E_i) = \exp\left(-\frac{E_i - E_{\min}}{k_b T_{\text{eff}}}\right)$$
4. reject/accept new structure with probability P

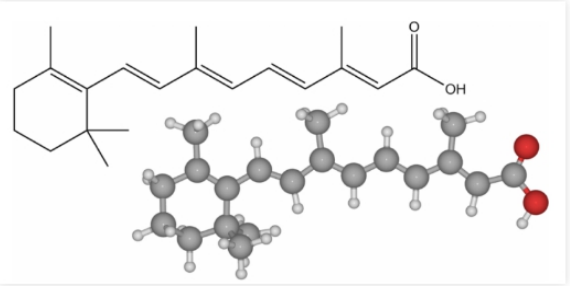
T_{eff} enables acceptance of higher energy structures

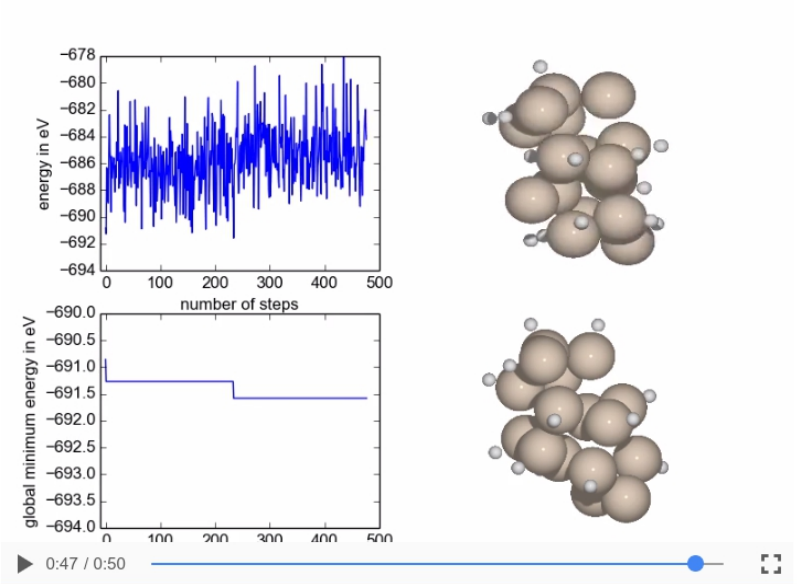
sequentially constructs a global map of "basins of attraction"

[1] Wales, Doye, J. Phys. Chem. A 75, 288-291 (1995)

global optimization of clusters and molecules







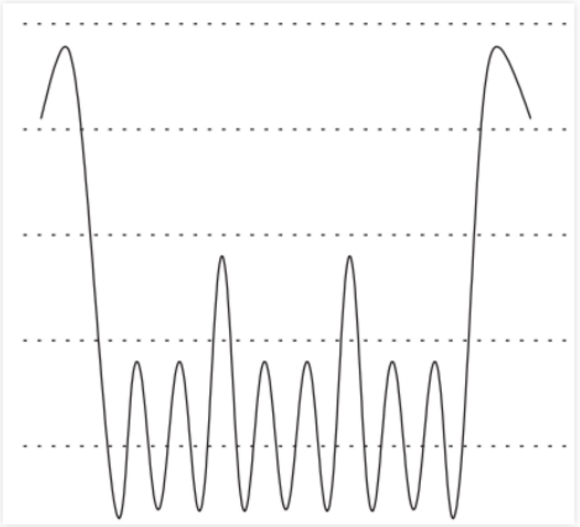
Optimization runs find less than 1%

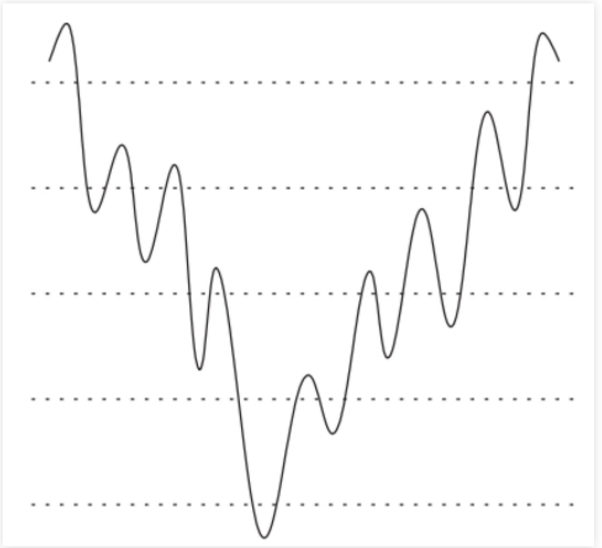
intact stable geometries

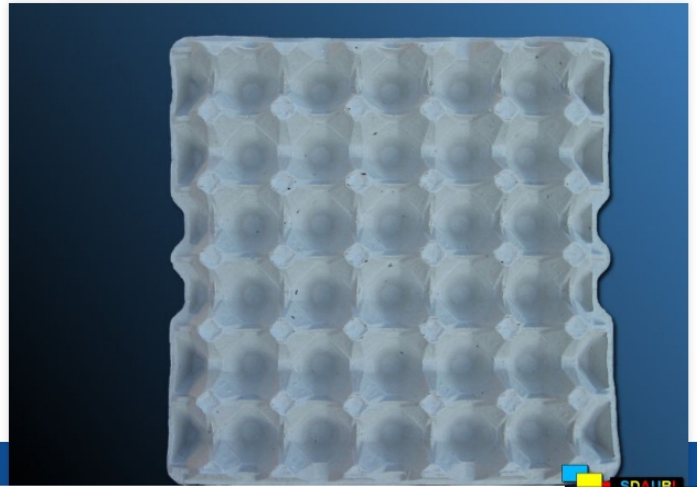
99% of geometries are fragmented

Question: Can we find an optimal choice of trial moves Δx?

global optimization of clusters and molecules



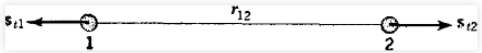




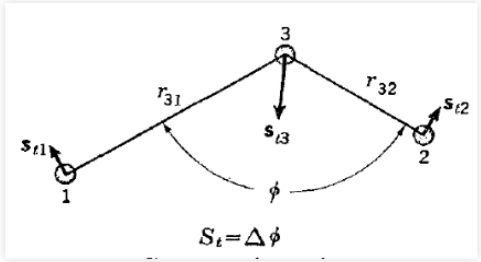
Wales, Miller, Walsh, Nature 394, 758-760 (1998)

What about simple internal coordinates?

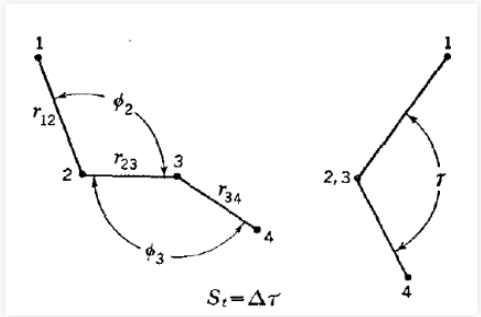
bond stretch



angle bend



dihedral torsion



coordinate	internal coordinates (IC) $\mathbf{q}$	$\mathbb{R}^3$ (CC) $\mathbf{x}$
bond	r_{12}	$\{\vec{s}_{t1}, \vec{s}_{t2}\}$
angle	ϕ_{123}	$\{\vec{s}_{t1}, \vec{s}_{t2}, \vec{s}_{t3}\}$
dihedral	τ_{1234}	$\{\vec{s}_{t1}, \vec{s}_{t2}, \vec{s}_{t3}, \vec{s}_{t4}\}$

Jacobi matrix: $B = \frac{dq}{dx}$ $\quad \mathbf{\Delta q = B \Delta x}$

- rotationally invariant, not permutationally invariant
- capture the local chemistry in covalent systems
- correspond to simple linear graphs with 2, 3, and 4 nodes

- z-matrix construction:
  non-unique, system- and geometry-specific
- redundant set of internal coordinates
- random mixing of ICs for Δx
  -> disconnected local changes

Wilson, Decius, Cross, Molecular Vibrations, Dover Books (2001)

Principal Component Analysis to the rescue!

Idea: Using these local graphs, we build a highly redundant coordinate set and performe a principal component analysis

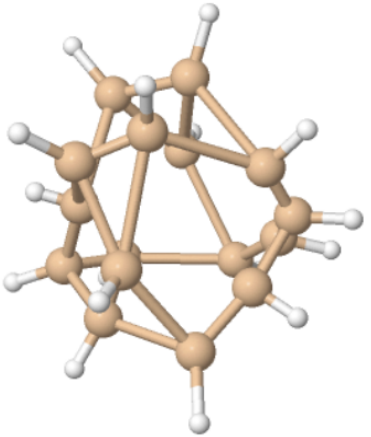
$\mathbf{B}$ is $(M \times 3N)$ matrix

$M/3N \approx 2-40$

Singular Value Decomposition

$$\mathbf{G} = \mathbf{B}^{\dagger}\mathbf{B} = \mathbf{U}\begin{bmatrix} \mathbf{\Lambda} & 0 \\ 0 & 0 \end{bmatrix}\mathbf{U}^{\dagger}$$

- $\mathbf{U}$ defines delocalized internal coordinate (DC) space
- subset of (3N-6) coordinates with singular value >0



JSmol

$$\Delta\mathbf{d} = \mathbf{U}\Delta\mathbf{q} = \mathbf{U}\mathbf{B}\Delta\mathbf{x} = \mathbf{B}'\Delta\mathbf{x}$$

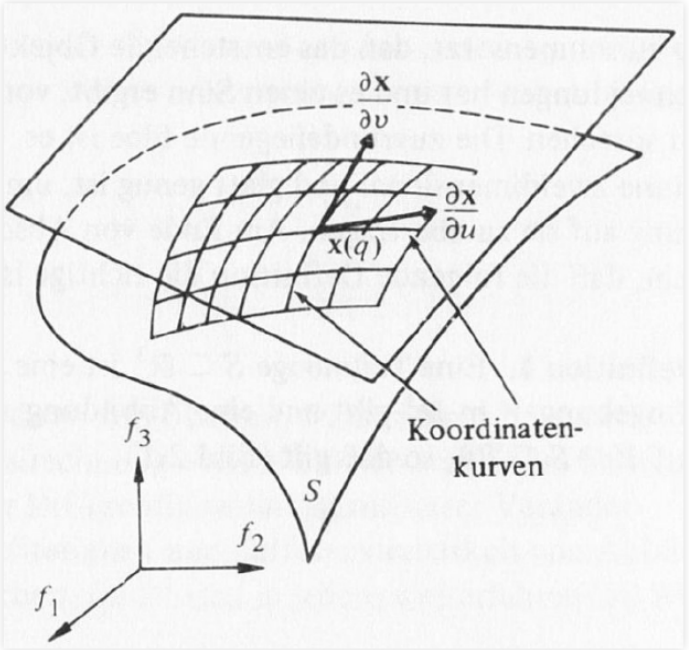
constructing displacement vectors in $\mathbb{R}^3$

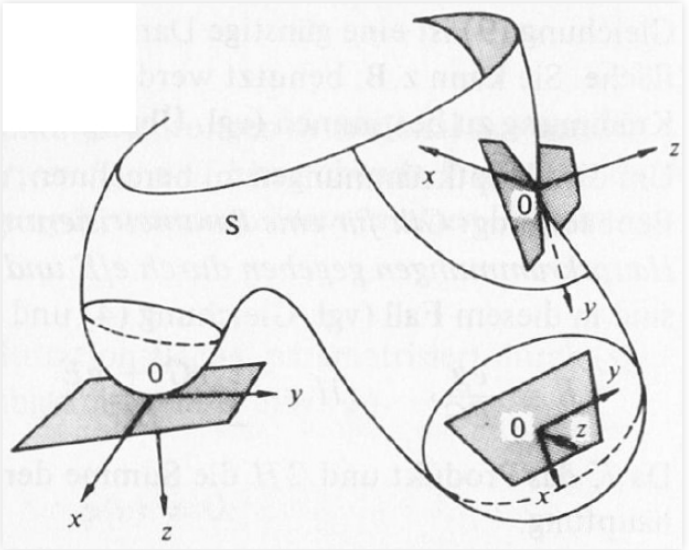
$$\Delta\mathbf{x} = \mathbf{B}'^{-1}\Delta\mathbf{d}$$

generalized inverse: $\mathbf{B}'^{-1} = \mathbf{B}'^{T}\underbrace{(\mathbf{B}'\mathbf{B}'^{T})^{-1}}_{\mathbf{G}'^{-1}}$

Pulay, Fogarasi, J. Chem. Phys. 96, 2856 (1992)
Baker, Kessi, Delley, J. Chem. Phys. 105, 192 (1996)

Interlude: Curvilinear hyperspaces and tangential approximations



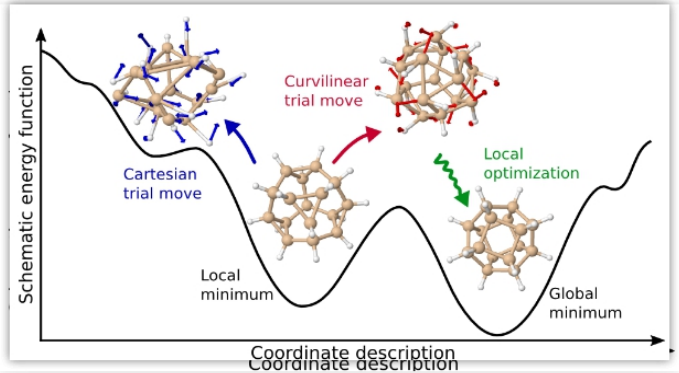


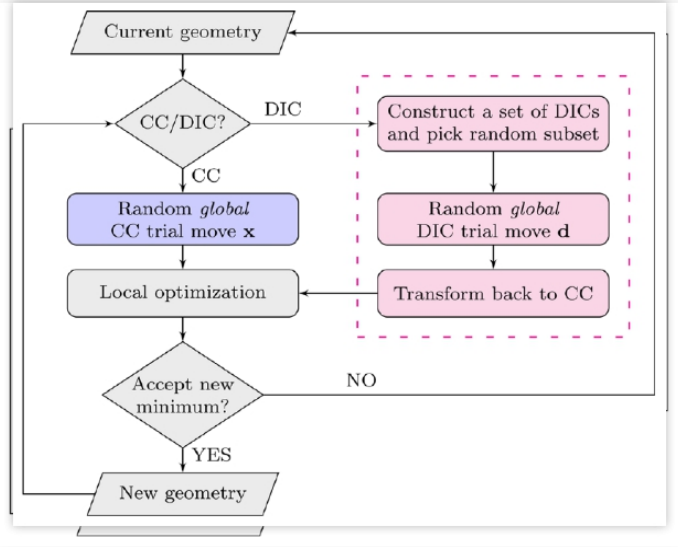
- Jacobi matrix **B** is a tangential approximation to a curvilinear hyperplane
- displacement vectors transform bijectively in this plane
- transformation is not linear bijective for absolute positions, forces, ...
- when generating Cartesian absolute coordinates: we need an iterative backtransformation!

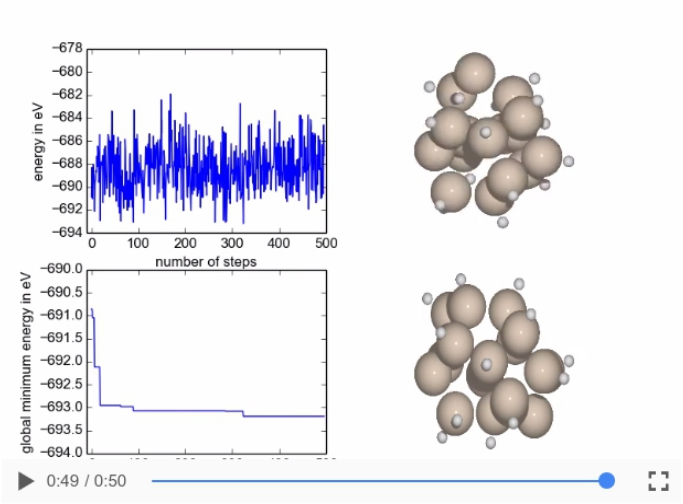
Manfredo do Carmo, "Differentialgeometrie von Kurven und Flächen", Vieweg Verlag (1983)

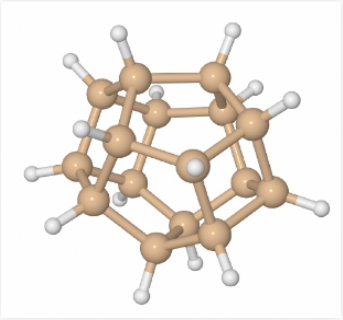
Baker, "Constrained optimization in delocalized internal coordinates." J. Comput. Chem. 18, 1079–1095 (1997).

Structure search with delocalized curvilinear coordinates









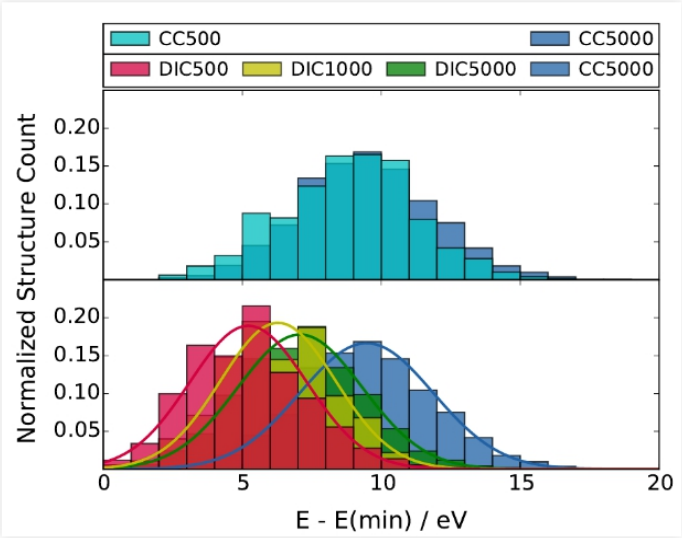
BSc/MSc thesis:
Konstantin Krautgasser

Dr. Chiara Panosetti

Panosetti, Krautgasser, Palagin, Reuter, Maurer, Nano Lett. 15, 8044 (2015)

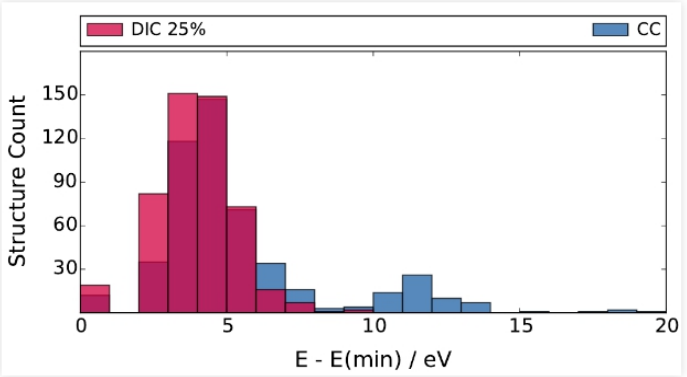
bias or constraint?

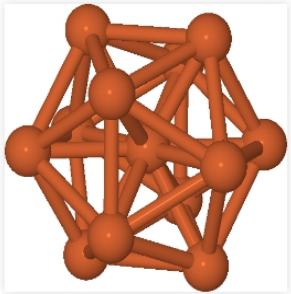
as $N_{step} \rightarrow \infty$



- for large numbers of steps we retreive CC sampling
- DC trial moves do not constrain search
- they bias towards energetically lower lying structures

What about chemically isotropic systems?





Even without this bias, we are more efficient

DC displaced structures are less strained

-> ~ -50% relaxation steps

organic molecules

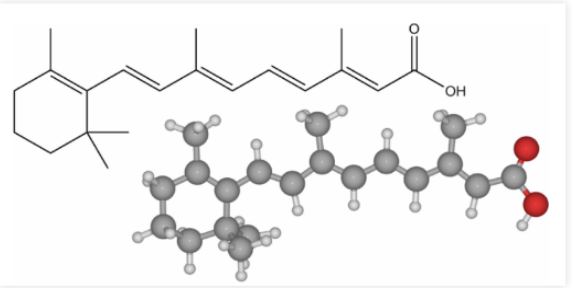
Introducing constraints

defining a constraint vector in DC space

$$\mathbf{C}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

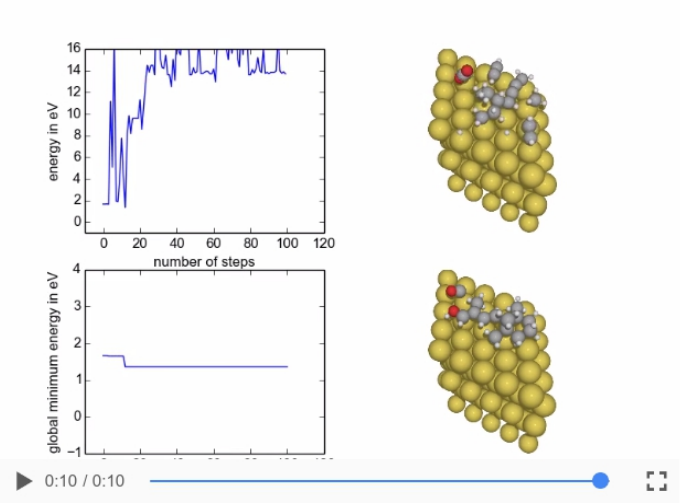
$$\tilde{\mathbf{C}}_i = \sum_j \langle \mathbf{C}_i | \mathbf{U}_j \rangle \mathbf{U}_j$$

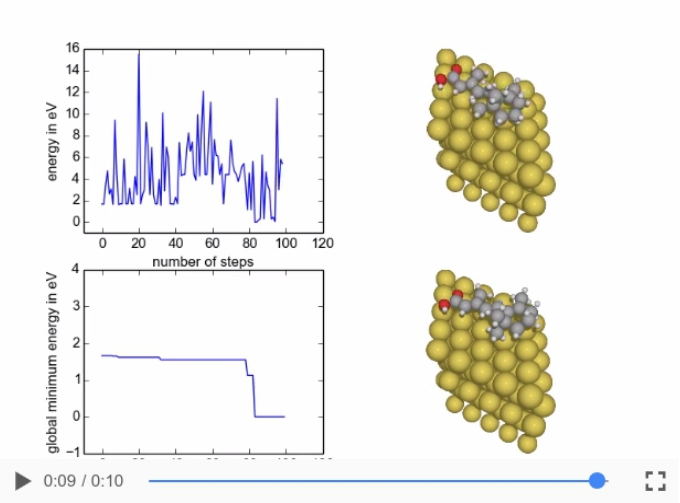
Gram-Schmidt orthogonalization $\rightarrow \mathbf{U}''\mathbf{B}$



method	**# of unique unfragmented minima/100 steps**
CC dr=0.90	-
DC 25% dr=0.90	2
constr. DC 25% dr=0.90	4
constr. DC 25% dr=2.50	9

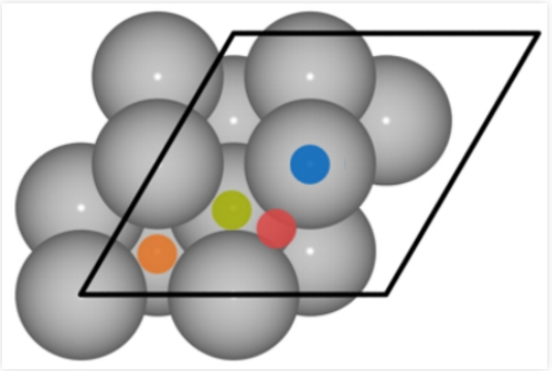
molecules on surfaces: β-acid on Au(111)

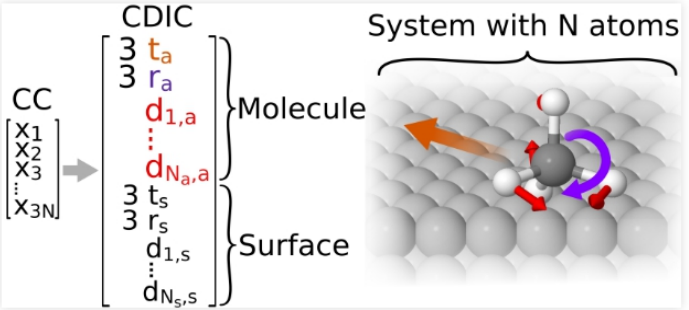


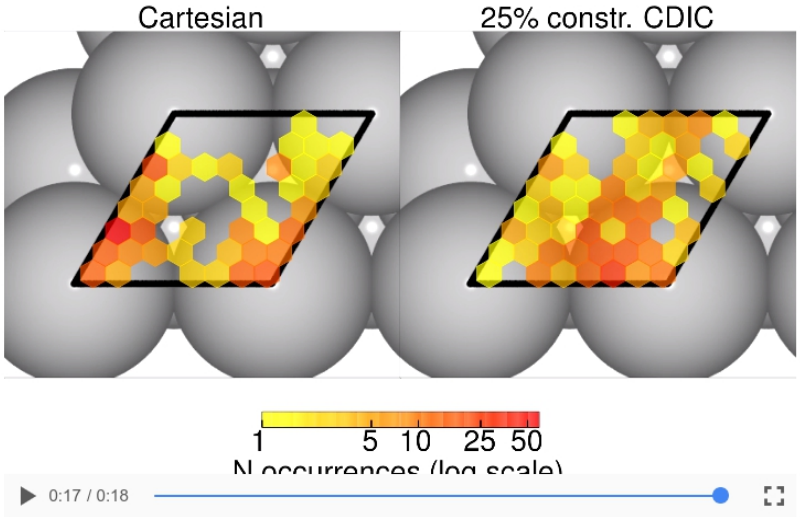


CC dr=0.5	**% of structures**	**25%cDC** dr=1.5	**% of structures**
dissociations	93	dissociations	38
revisits	7	revisits	3
different site	0	different site	48
new structure	0	new structure	11

adding translations and rotations: CH_4/Ag(111) adsorption site sampling

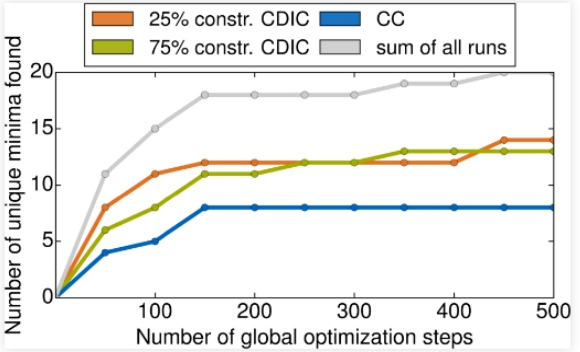






surface-symmetry-adapted translations

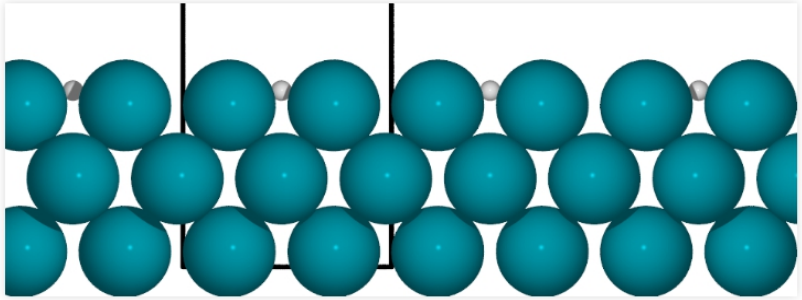
quaternion description of rotation angles

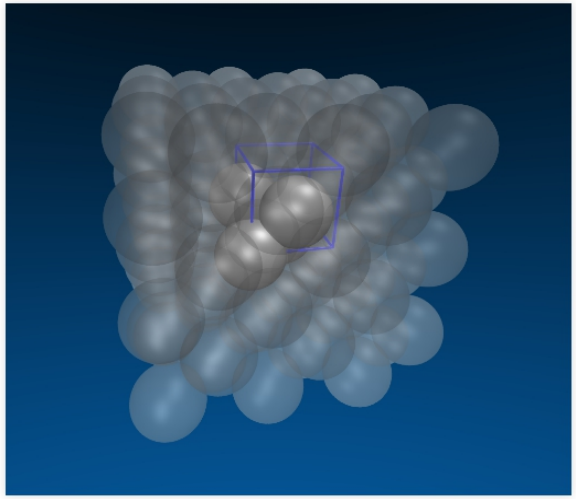


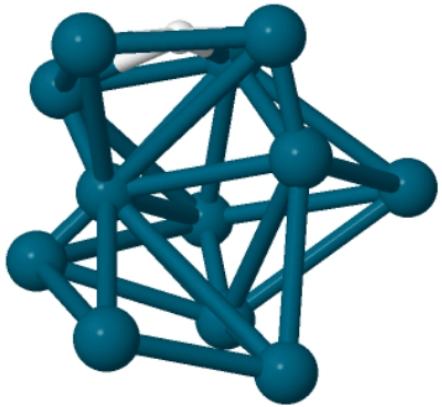
we sample the lateral PES more equally

we sample more unique minima

periodicity and unit cell deformations







1. Extend cell periodcally
2. Calc. redundant internals
3. graphs across unit cell boundaries
4. SVD on $G=B^{\dagger}B$

results in 3N non-zero DCs

- Fully decoupled period. deloc. coords.
- basis for Cluster-expansion interpolation

Kudin, Scuseria, Schlegel, JCP **114**, 2919 (2001)
Bucko, Hafner, Angyán JCP **122**, 124508 (2005)

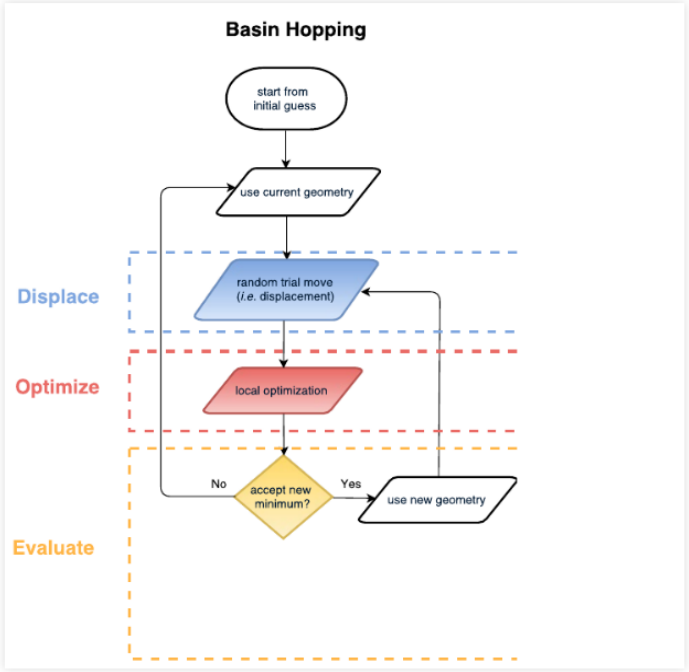
winak: open-source package for materials structure search and beyond

Python framework for curvilinear coordinate construction

Interfaces to QM packages via

Atomic Simulation Environment (ASE)

generic classes: Displace Optimize Evaluate



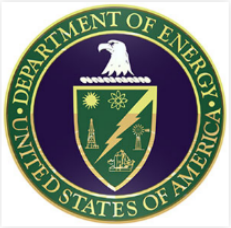
https://gitlab.com/reinimaurer1/winak

- Clear task separation
- Easily customizable
- Basin hopping, minima hopping
  genetic algorithms (in progress), ...
- analysis/clustering tools
- recursive fragment detection

Contributors: Konstantin Krautgasser, Chiara Panosetti, Christoph Scheurer, Daniel Strobusch

… and beyond

- PCA-based selection of optimal collective displacements
- combination with particle insertion, permutational invariance?!
- more complicated molecular graphs for complexes/bulk materials
- organic crystals, hybrid interfaces, …
- collective variable based transition path identification
- ?DC singular value-property correlations?



Thank you for your attention

Acknowledgements

Yale University

Prof. John Tully

Martin Stöhr

TU Munich

Prof. Karsten Reuter

Dr. Chiara Panosetti

Konstantin Krautgasser

Georg Michelitsch

U Oxford

Dr. Dennis Palagin

Funding

DFG

DoE

Computing

Leibnitz Supercomputing Center

Panosetti, Krautgasser, Palagin, Reuter, Maurer, Nano Lett. **15**, 8044-8048 (2015)

Krautgasser, Panosetti, Palagin, Reuter, Maurer, J. Chem. Phys. **145**, 084117 (2016)

Stöhr, Michelitsch, Tully, Reuter, Maurer, J. Chem. Phys. **144**, 151101 (2016)