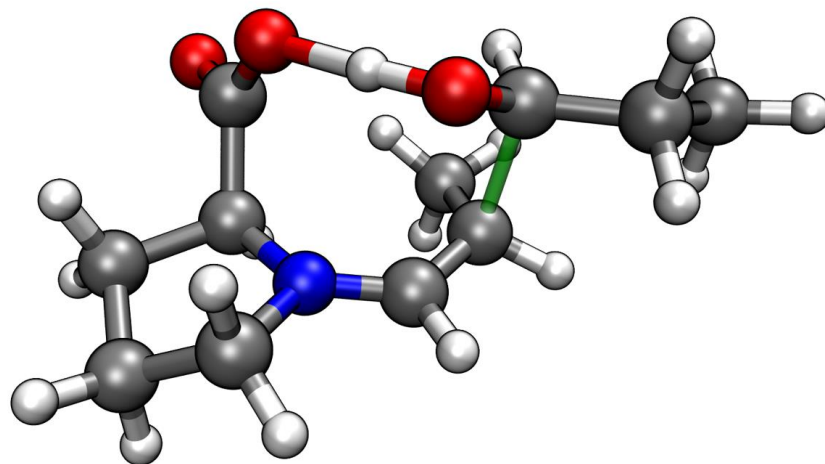# Computing Liquid Phase Free Energy Profiles from *ab initio* Molecular Dynamics Simulations

Martin Brehm
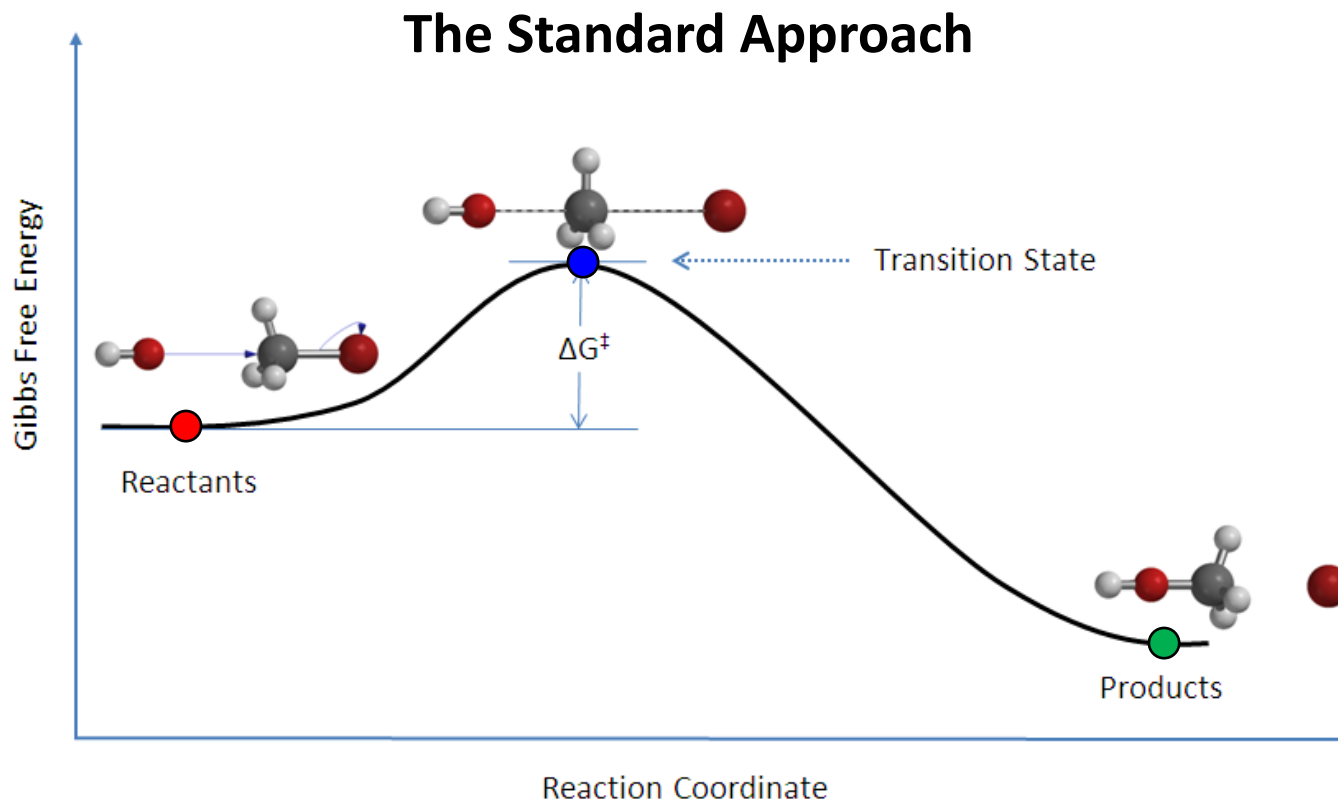
Martin-Luther-Universität Halle–Wittenberg

`https://brehm-research.de/`

Outline

1.) The Standard Approach to Compute Reactions

2.) An Example Reaction

3.) Introduction to Simulation Techniques

4.) Results and Conclusions

Computing Reaction Energy Profiles

The Standard Approach

1.) Optimize **reactand** and **product** complex geometries

2.) Search and optimize **transition state** *(e.g. via Nudged Elastic Band)*

Computing Reaction Energy Profiles
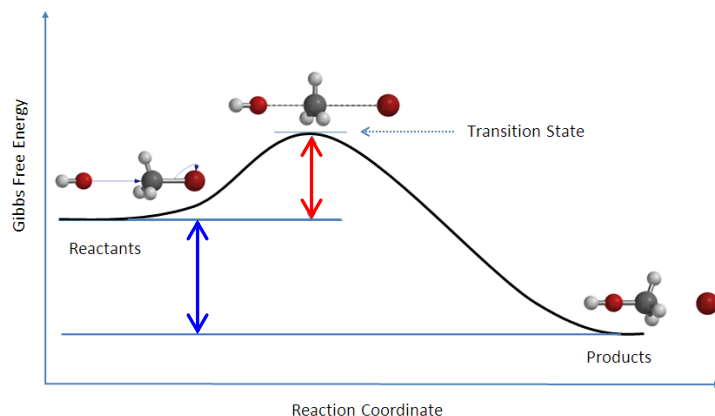
What do we obtain?

1.) Equilibrium Constant: $K = \exp(-\Delta G / RT)$

2.) Rates of forward / backward reactions: $k = A \cdot \exp(-\Delta G^{\ddagger} / RT)$

If there exist multiple reactions *(e.g. different stereo-isomers)*:

Can find out which reaction is <u>thermodynamically / kinetically favored</u>!

Ratio of K or *k* can even give <u>quantitative selectivity</u> *(e.g., ee% or cis:trans)*!

Computing Reaction Energy Profiles

How to describe Free Enthalpy?

Quantum chemical calculation gives only enthalpy differences ΔH

But we need ΔG ...

$\Delta G = \Delta H - T \cdot \boxed{\Delta S}$

Ok, how to obtain ΔS?

Very hard to obtain accurate value *(sampling of **full** phase space!)*...

Use **very crude** approximations for partition sum instead:

Harmonic oscillator / Rigid rotator / Particle in a box

Breaks down for hydrogen bonds / hindered rotations

Breaks down for flexible molecules

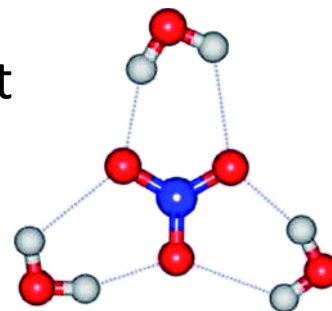Point mass approximation; neglects any solvent

Computing Reaction Energy Profiles

How to describe Solvent Influence?

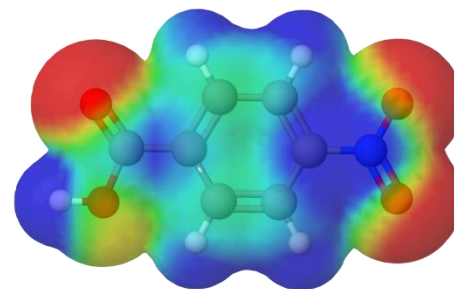The whole approach only works for molecules in vacuum!

a) Microsolvation

Put an *(arbitrary)* small number of explicit solvent
molecules to places that should be solvated.

b) Implicit Continuum Solvation Models

Put the whole molecule into a cavity with the
desired dielectricity constant ε.
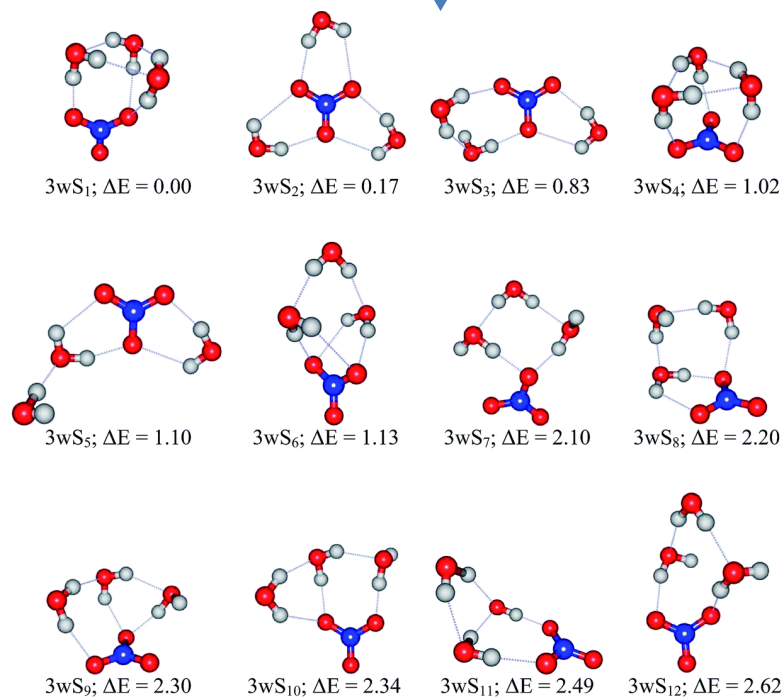
Computing Reaction Energy Profiles

How to describe Solvent Influence?

a) Microsolvation

- Number of required solvent molecules not known

- Even for fixed number: Many different geometries...

- Different solvation of educt, TS, product often <u>breaks the results</u>

→ Even if it works: Requires a lot of experience and manual fine-tuning

Often not an option :-/

3wS$_1$; ΔE = 0.00   3wS$_2$; ΔE = 0.17   3wS$_3$; ΔE = 0.83   3wS$_4$; ΔE = 1.02

3wS$_5$; ΔE = 1.10   3wS$_6$; ΔE = 1.13   3wS$_7$; ΔE = 2.10   3wS$_8$; ΔE = 2.20

3wS$_9$; ΔE = 2.30   3wS$_{10}$; ΔE = 2.34   3wS$_{11}$; ΔE = 2.49   3wS$_{12}$; ΔE = 2.62

Computing Reaction Energy Profiles

How to describe Solvent Influence?

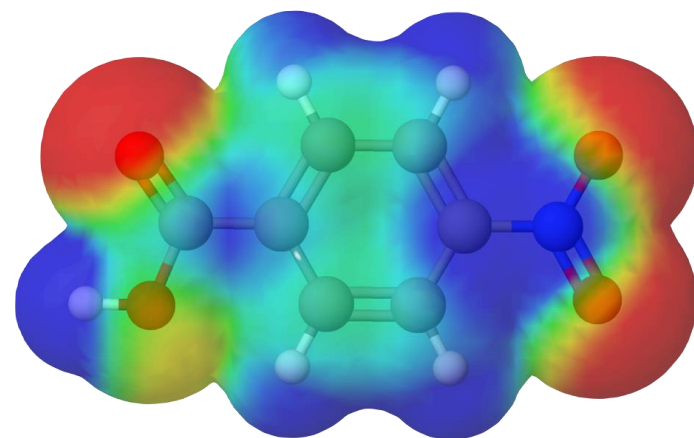**b) Implicit Continuum Solvation Models**

Examples: COSMO, PCM

Put the whole molecule into a cavity
with the desired dielectricity constant ε.

Very crude approximation, neglects
all **directed** solvent effects *(hydrogen bonds,
salt bridges, solvent entropy changes, ...)*

Very easy to use *(only requires ε of solvent)*.

$\rightarrow$ Most studies use it.

Computing Reaction Energy Profiles

Conclusion

- Approximation for ΔG is very crude;
  fails for hydrogen bonds, hindered rotations, flexible molecules, …

- Implicit solvent model is very crude; neglects all directed
  solvent interactions *(hydrogen bonds, salt bridges, …)*

- Solvent entropy *(hydrophobic effect, …)* is completely neglected

- Approach works only for a single conformer at a time,
  no conformer sampling

→ **It is surprising that it works at all!**
  (as it does in 1000s of publications…)
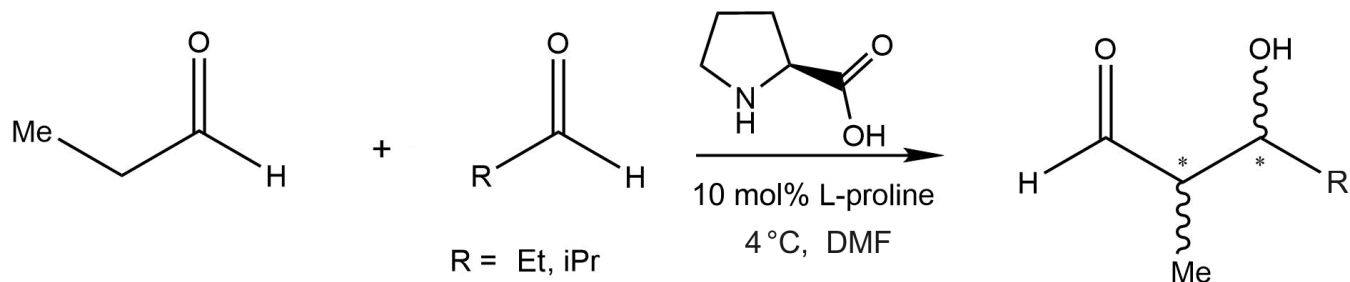
Proline-Catalyzed Cross-Aldol Reaction
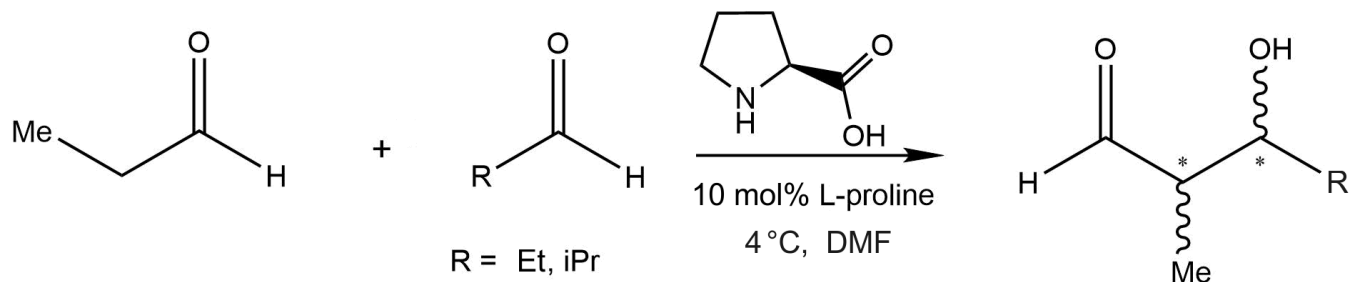
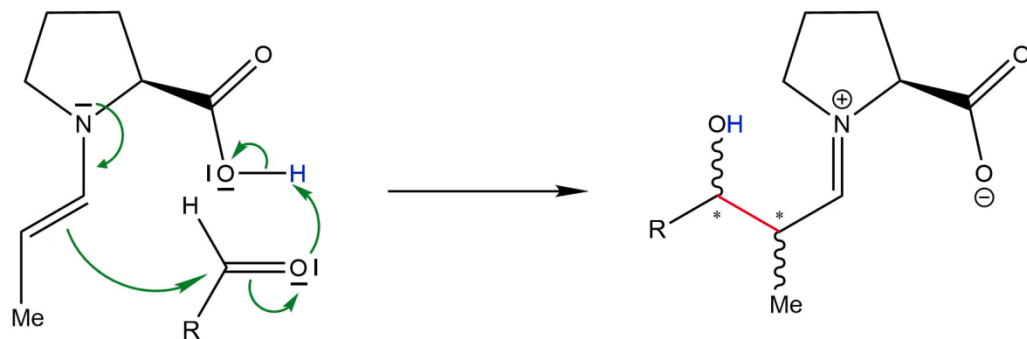| R	%ee	syn : anti	yield / %
Et	99	1 : 4	80
iPr	> 99	1 : 24	82
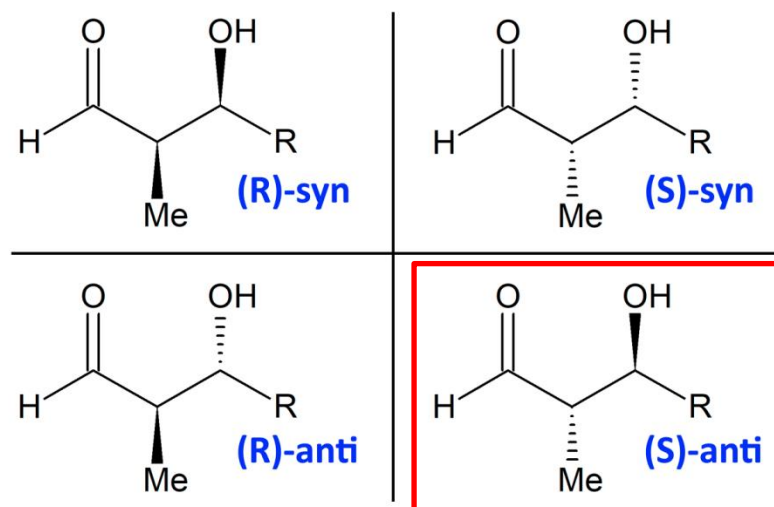
A. B. Northrup, D.W.C. MacMillan: "The First Direct and Enantioselective Cross-Aldol Reaction of Aldehydes", *J. Am. Chem. Soc.* **2002**, *124*, 6798–6799.
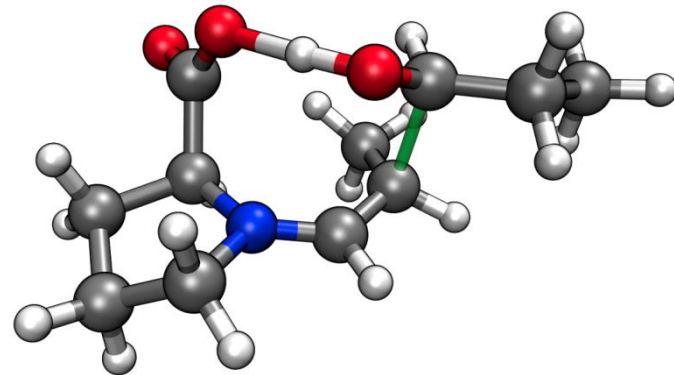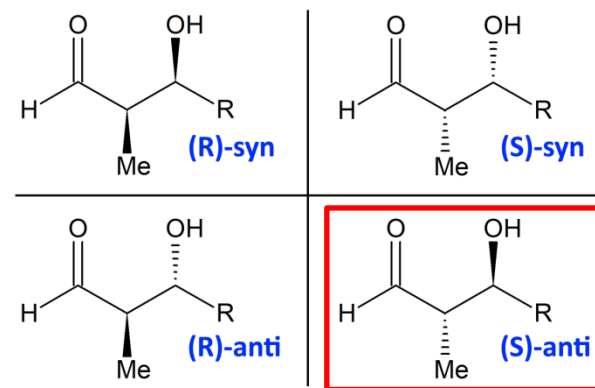
Proline-Catalyzed Cross-Aldol Reaction

Mechanism:

Transition State:

A. B. Northrup, D.W.C. MacMillan: "The First Direct and Enantioselective Cross-Aldol Reaction of Aldehydes", *J. Am. Chem. Soc.* **2002**, *124*, 6798–6799.

Proline-Catalyzed Cross-Aldol Reaction

Results from static approach:
(implicit solvent DMF)

Product	$\Delta_R E$ / kJ mol^{-1}	$\Delta_R G$ / kJ mol^{-1}	$\Delta E^{\ddagger}$ / kJ mol^{-1}	$\Delta G^{\ddagger}$ / kJ mol^{-1}
(R)-anti	5.86	23.77	40.79	44.78
(R)-syn	21.50	39.60	78.22	88.63
(S)-anti	24.89	43.23	51.63	62.91
(S)-syn	22.74	36.28	75.67	90.88

→ Fails completely

M. Weiß, M. Brehm: "Exploring Free Energy Profiles of Enantioselective Organocatalytic Aldol Reactions under Full Solvent Influence", *Molecules* **2020**, *25 (24)*, 5861.

Introduction to Simulation Techniques

Molecular Dynamics (MD) Simulations

- Consider a *(small)* chemical system *(up to ≈ 2000 atoms)* in a periodic cell
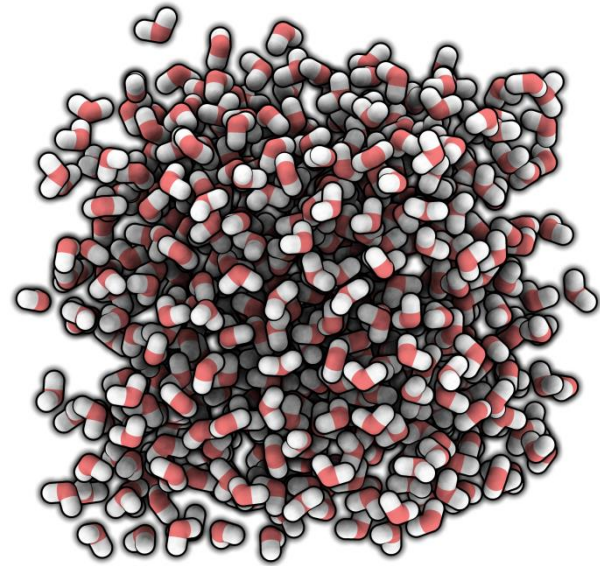
- Compute the forces which act on all the atoms

- Move the atoms a bit by solving Newton's equations of motion:

$$m_i \dot{\mathbf{v}}_i(t) = \mathbf{F}_i(\mathbf{r}_i) = -\nabla V(\mathbf{r}_i)$$
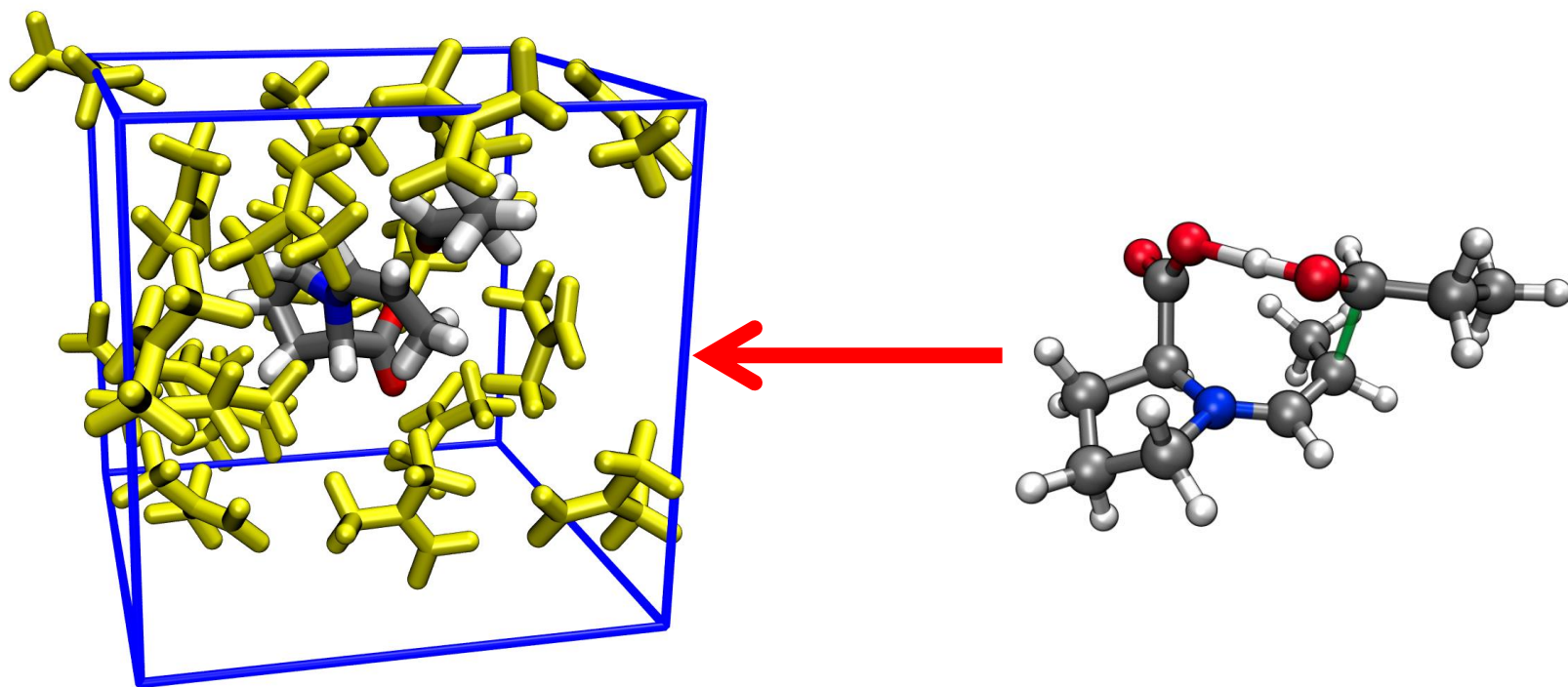
$$\dot{\mathbf{r}}_i(t) = \mathbf{v}_i(t)$$

- Repeat many times

Introduction to Simulation Techniques

How to Describe the Solvent Effect?

Just include all solvent molecules explicitly!   No „cheap trick" as before…

Periodic cell of reactand mixture in 18 DMF molecules:

Introduction to Simulation Techniques

How to Obtain ΔG from Molecular Dynamics?

$\rightarrow$ Free energy sampling methods

$\rightarrow$ **Metadynamics**

A. Laio, M. Parrinello: „Escaping Free-Energy Minima",
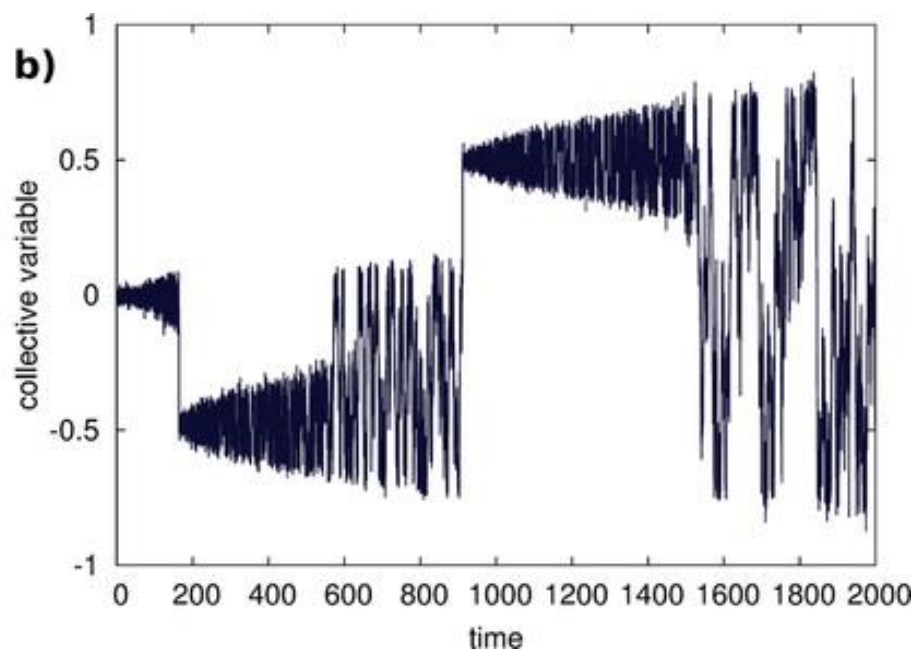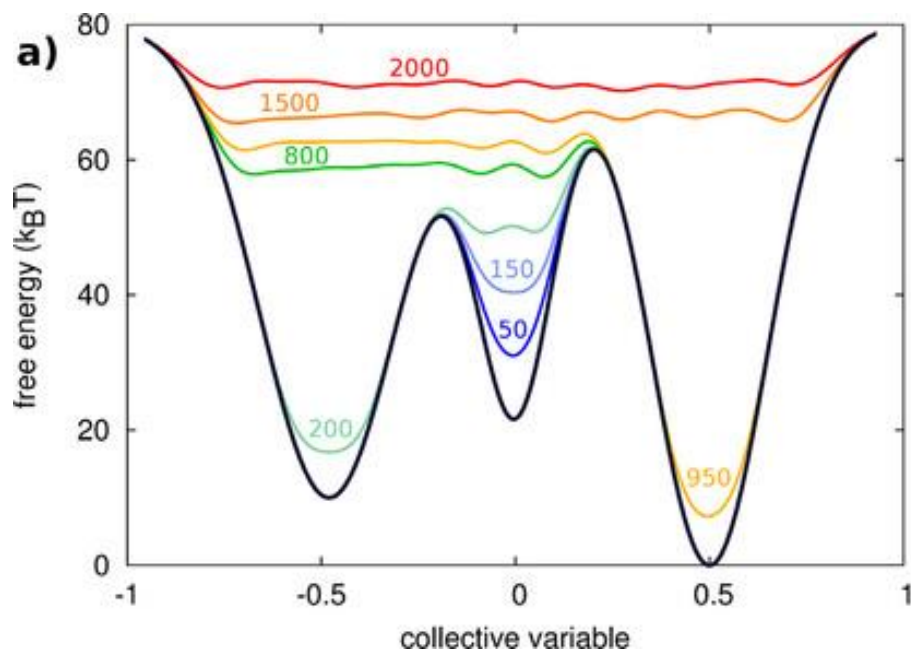*Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99 (20)*, 12562.

Introduction to Simulation Techniques

Metadynamics Simulations

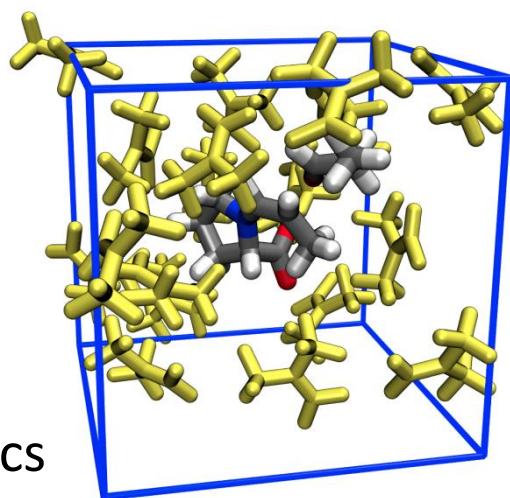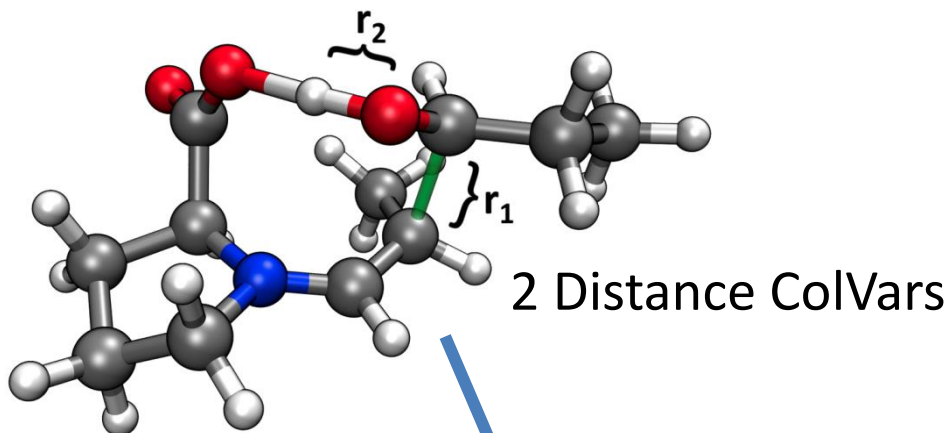Define one or more „Collective Variables" *(reaction coordinates)*

Let the dynamics run; place repulsive bias potentials at current position to push the system away

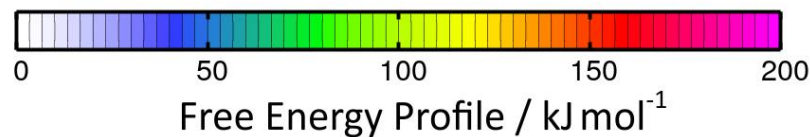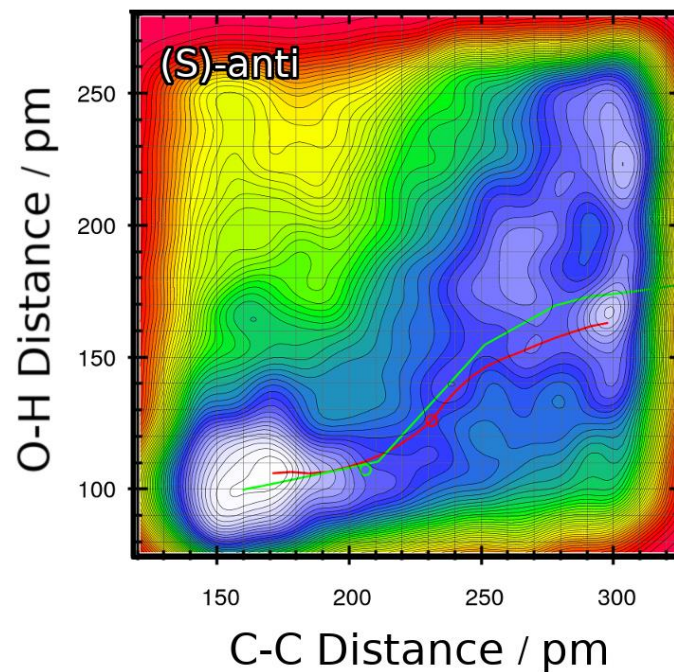→ Efficient sampling of the full relevant configuration space

Introduction to Simulation Techniques

Metadynamics Simulations

2 Distance ColVars

Run
Metadynamics

2D Free Energy Profile!

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Introduction to Simulation Techniques

How to Avoid Unwanted Side Reactions?

Our Metadynamics involves bond breaking.

Resulting intermediates are highly reactive...
→ Either react with solvent or break apart into pieces

We developed the **„Hybrid AIMD" (HyAIMD)** approach to avoid this.

- Place an empirical repulsive potential on all pairs of atoms
  which shall **not** form bonds:
  $$E_{ij}^{\text{LJ}}(r) := 4\epsilon_{ij}\left[\left(\frac{r}{\sigma_{ij}}\right)^{12} - \left(\frac{r}{\sigma_{ij}}\right)^6\right]$$

- Place an empirical spring potential on all bonds
  which shall remain **intact**:
  $$E^{\text{bond}}(r) := \frac{k}{2}(r - r_0)^2$$

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Results

Simulation Protocol

We started Metadynamics simulations for the four different stereoisomers.
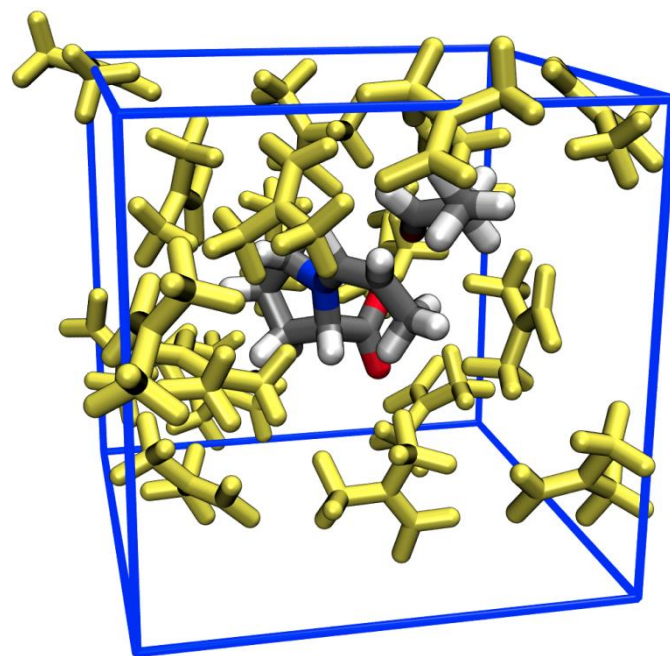
Reactands + 18 DMF molecules,
cell size 14 Angstrom, 350 K.

CP2k package, HyAIMD, BLYP-D3.
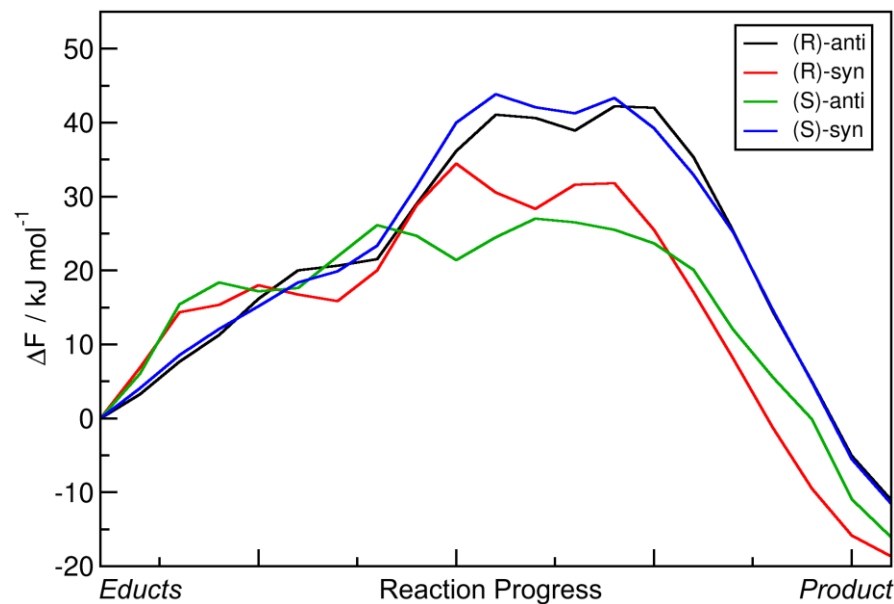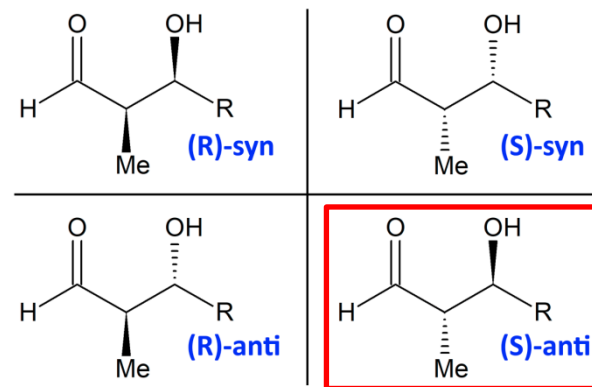
TRAVIS for analyzing the trajectories.

250 ps total simulation time,
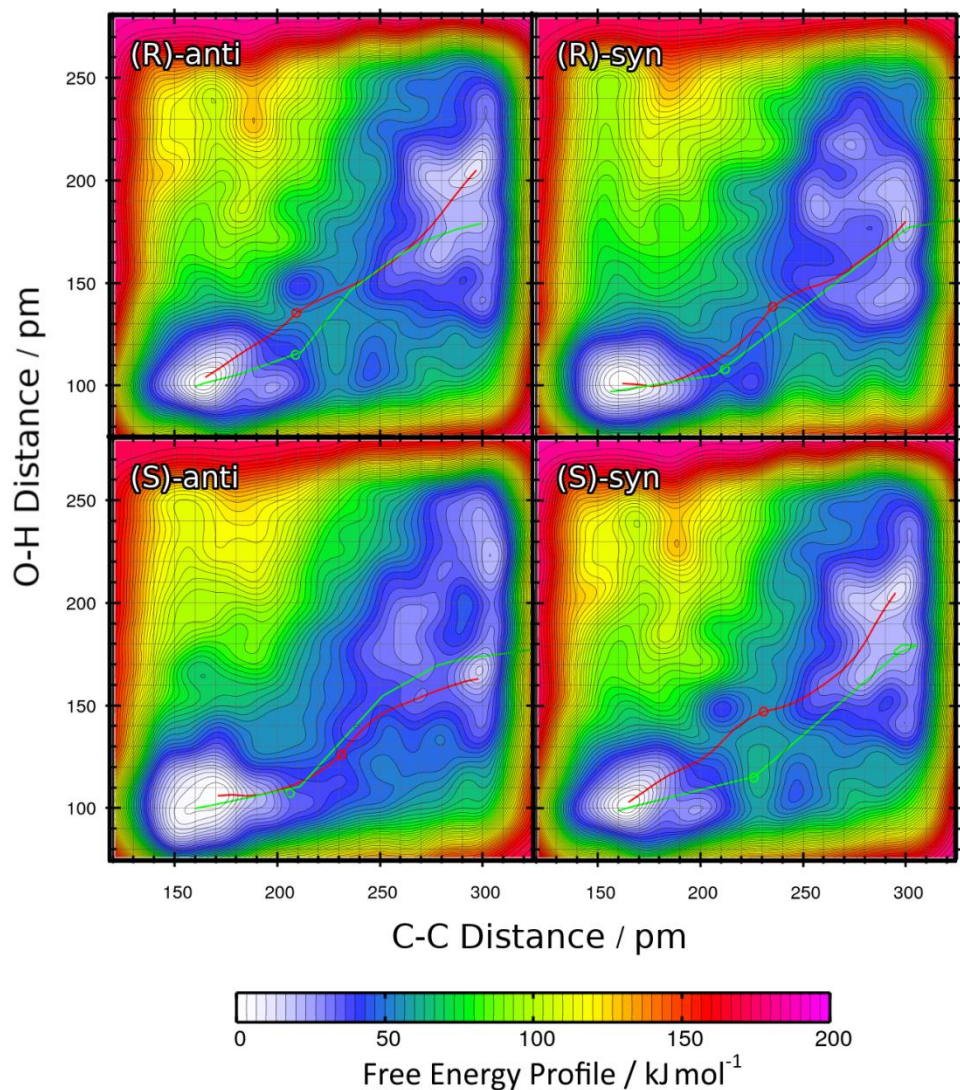30'000 core hours per isomer.

Around 3 months runtime on a
HPC node with 16 cores per isomer.

https://www.cp2k.org
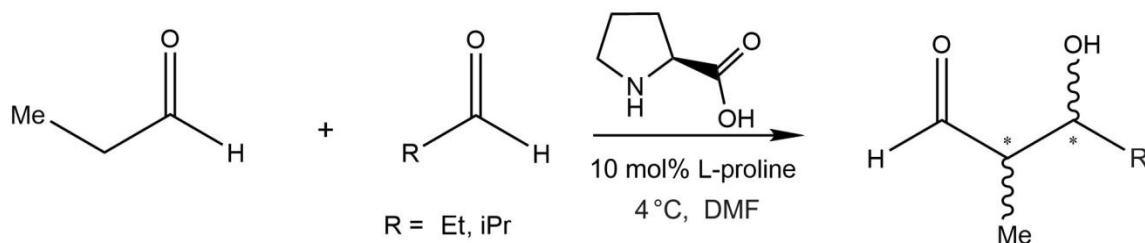http://www.travis-analyzer.de

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Results

→ **Correct stereoisomer is predicted!** 👍

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Results

Can we Capture the Substituent Effect?

R = Et, iPr

10 mol% L-proline
4 °C, DMF

R	%ee	syn : anti	yield / %
Et	99	1 : 4	80
iPr	> 99	1 : 24	82

Experiment

(R)-syn   (S)-syn   (R)-anti   (S)-anti

Simulation R = Et

Product	$\Delta_R F$ / kJ mol^{-1}	$\Delta F^{\ddagger}$ / kJ mol^{-1}
(R)-anti	−11.06	42.24
(R)-syn	−18.66	31.79
(S)-anti	−17.71	25.40
(S)-syn	−11.52	43.84

Simulation R = iPr

Product	$\Delta_R F$ / kJ mol^{-1}	$\Delta F^{\ddagger}$ / kJ mol^{-1}
(R)-anti	−14.33	36.54
(R)-syn	−9.35	39.35
(S)-anti	−35.63	24.29
(S)-syn	−13.70	22.59

Substituent effect is correctly described!

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Results

Article published in 2020:

by ⊗ Moritz Weiß ✉ ⓘ and ⊗ Martin Brehm * ✉ ⓘ

Institut für Chemie, Martin-Luther-Universität Halle-Wittenberg, von-Danckelmann-Platz 4, D-06120 Halle (Saale), Germany
* Author to whom correspondence should be addressed.

View Full-Text   Download PDF   Browse Figures   Citation Export

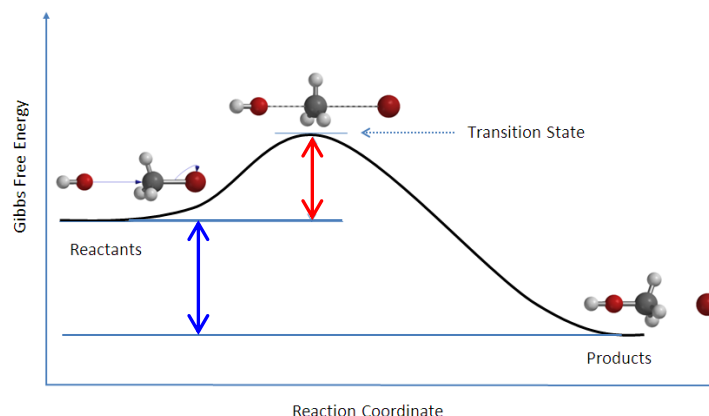M. Weiß, M. Brehm: "Exploring Free Energy Profiles of Enantioselective Organocatalytic Aldol Reactions under Full Solvent Influence", *Molecules* **2020**, *25 (24)*, 5861.

Conclusions I

- The **standard approach** for computing reactions is still
geometry optimization → transition state search → frequency calculation

- Many **crude approximations**:

    - Implicit solvent model → no directed solvent interactions

    - Solvent entropy missing *(hydrophobic effect, …)*

    - Can break down if hydrogen bonds or hindered rotations occur

    - No conformer sampling, only one conformer at a time

→ Surprising that it indeed works so often

Conclusions II

- **Metadynamics** on top of **DFT-based** *ab initio* **MD** (AIMD) can be a good *(but expensive)* alternative for cases where the classical approach fails

- Unwanted side reactions are an issue; we developed the **Hybrid AIMD** (HyAIMD) scheme to avoid them.

- Concerning the proline-catalyzed Aldol reaction, we were able to give an *ab initio* prediction for the stereoselectivity for the first time

- Both the product stereochemistry and the influence of the substituent are correctly predicted
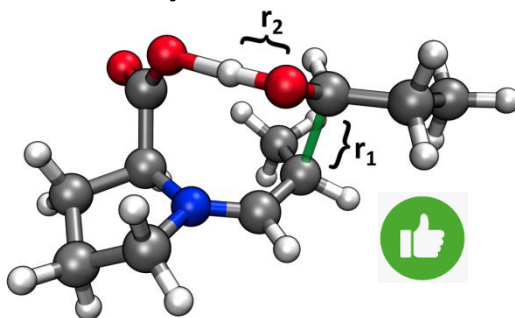
- The approach should work for all kinds of organic reactions in solvent; we are currently studying other examples

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Can it be done with ORCA?

Metadynamics: **yes, since ORCA 5.0**

```
Timestep 0.5_fs
Initvel 350_K
Thermostat NHC 350_K Timecon 100.0_fs
Dump Position Stride 1 Filename "trajectory.xyz"

Manage_Colvar Define 1 Distance Atom 0 Atom 1
Manage_Colvar Define 2 Distance Atom 2 Atom 3

Metadynamics Colvar 1 Scale 1.0_A Wall Lower 3.0 50.0 Upper 10.0 50.0 Range 0.0 15.0 100
Metadynamics Colvar 2 Scale 1.0_A Wall Lower 1.0 50.0 Upper 8.0 50.0 Range 0.0 13.0 100
Metadynamics HillSpawn 40 0.5 0.5 Store 2000
Metadynamics WellTempered 6000_K
Metadynamics Lagrange 100.0 10.0 200.0_K 10.0_fs

Restart IfExists
Run 100000
```

→ **See Manual**

HyAIMD approach: **hopefully soon :-)**

M. Weiß, M. Brehm, *Molecules* **2020**, *25 (24)*, 5861.

Acknowledgment

„AK Brehm"

MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG

Prof. Daniel Sebastiani

`https://brehm-research.de/`