

Chemie v počítači

Tomáš Bouchal

Laboratoř Výpočetní Chemie (LCC)

Národní Centrum pro výzkum biomolekul (NCBR)

Středoevropský technologický institut (CEITEC)

Přírodovědecká Fakulta, Masarykova Univerzita, Brno

MUNI



CEITEC

MUNI
SCI

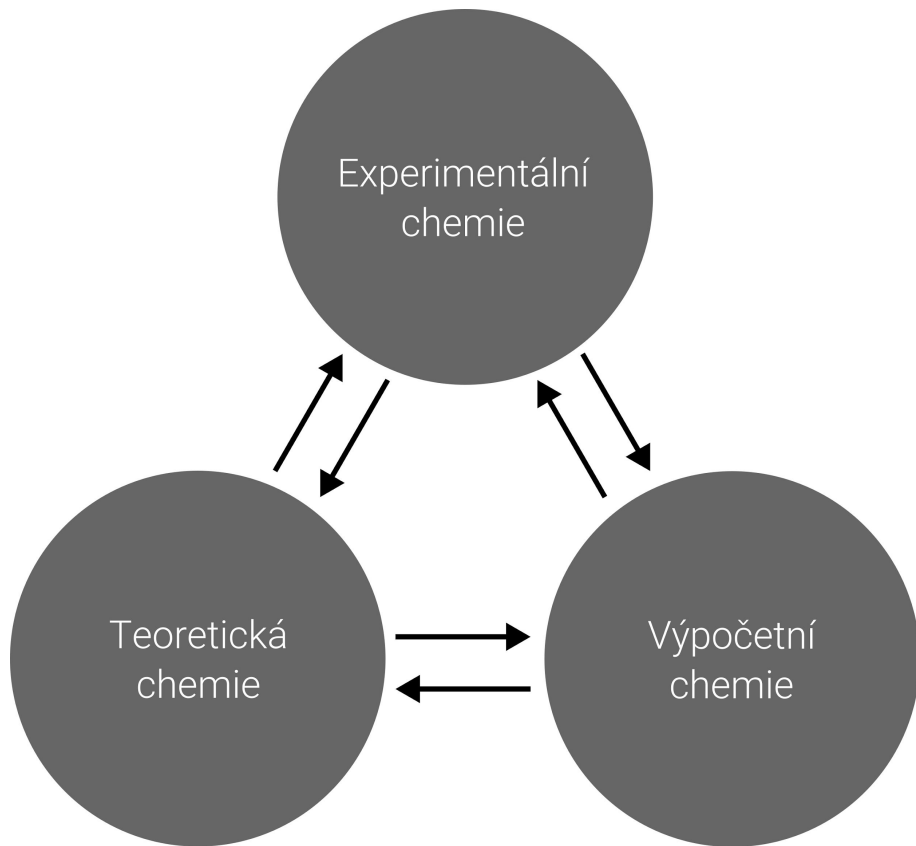
National Centre
for Biomolecular
Research

Co nás čeká?

- Co je to výpočetní chemie?
- Je možné simulovat realitu?
- Kde se počítá?
- Metody výpočetní chemie
- Ukázka reálného projektu

Výpočetní chemie

Vztah k chemii

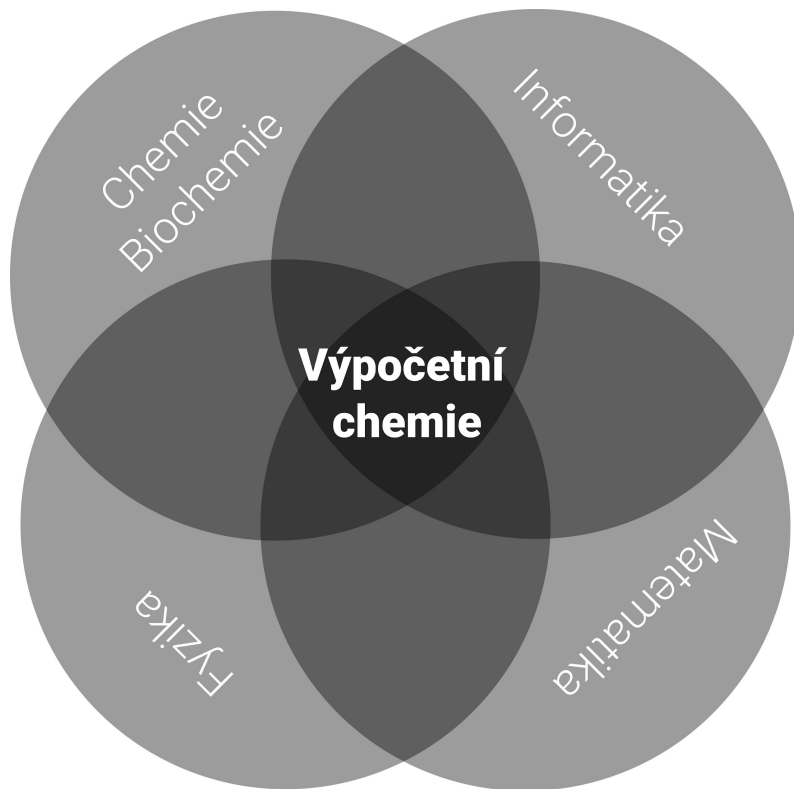


“Výpočetní chemie je obor chemie, který využívá poznatky teoretické chemie při vytváření počítačových programů, které umožňují počítat struktury a vlastnosti molekul.”

https://cs.wikipedia.org/wiki/Výpočetní_chemie

Výpočetní chemie

(bio)chemické problémy,
syntézy, ověřování



algoritmy, CPU/GPU výpočty,
cloud/grid computing,
software

teorie, aproximace.

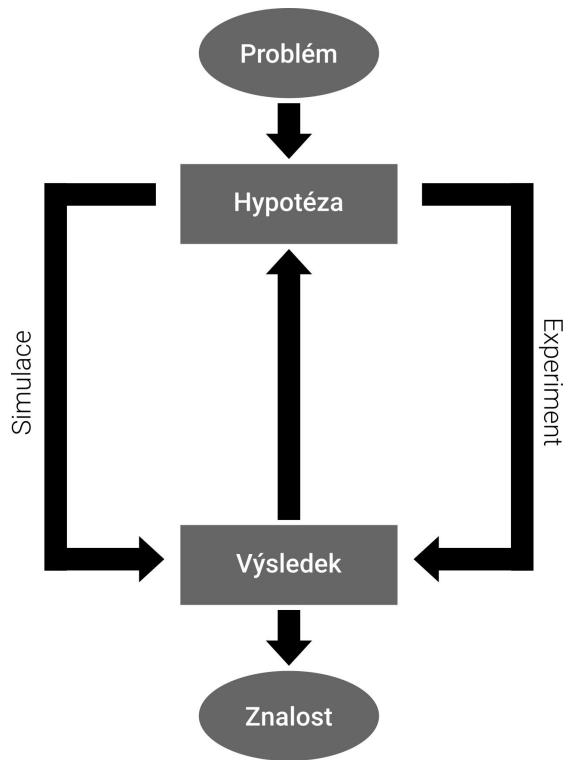
analytické řešení,
numerické řešení,
aproximace, statistika

Výpočty vs Experiment



- + téměř neomezené možnosti
- + rozlišení
- aproximace

1) Photo: Carlos Jones/Oak Ridge National Lab
2) Photo: xfel.eu



- + realita
- rozlišení
- drahé, nebezpečné, někdy nemožné

Simulace reality

Simulace reality?

Kolik molekul obsahuje 180 ml vody při pokojové teplotě?

Kolik počítačové paměti bude zapotřebí pro uložení informací o poloze všech atomů včetně jejich rychlostí za použití reálných čísel s jednoduchou přesností?

Simulace reality?

Kolik molekul obsahuje 180 ml vody při pokojové teplotě?

$$\begin{aligned} \rho &= 1000 \text{ kg/m}^3 = 1 \text{ g/ml} & m &= \rho V \\ N_A &= 6,022 \cdot 10^{23} \text{ mol}^{-1} \approx 6,0 \text{ mol}^{-1} & n &= m/M \end{aligned}$$

$$N = \mathbf{6,0 \cdot 10^{24}} \text{ částic}$$

Kolik počítačové paměti bude zapotřebí pro uložení informací o poloze všech atomů včetně jejich rychlostí za použití reálných čísel s jednoduchou přesností?

Simulace reality?

Kolik molekul obsahuje 180 ml vody při pokojové teplotě?

$$\begin{aligned} \rho &= 1000 \text{ kg/m}^3 = 1 \text{ g/ml} & m &= \rho V \\ N_A &= 6,022 \cdot 10^{23} \text{ mol}^{-1} \approx 6,0 \text{ mol}^{-1} & n &= m/M \end{aligned}$$

$$N = 6,0 \cdot 10^{24} \text{ částic}$$

Kolik počítačové paměti bude zapotřebí pro uložení informací o poloze všech atomů včetně jejich rychlostí za použití reálných čísel s jednoduchou přesností?

$$1 \text{ hodnota} = 4 \text{ (8) byte} \quad 4,32 \cdot 10^{26} \text{ B}$$

432 000 000 000 000 000 000 000 000 B
YB ZB EB PB TB GB MB KB

Simulace reality?

Kolik strojového času zabere simulace 1 s vývoje molekulárního systému?

Nejrychlejším molekulárním pohybem je vibrace O-H vazeb s přibližnou periodou 10 fs, která se bude vzorkovat 10 snímky. Výpočet jednoho snímku trvá přibližně 1 ms strojového času.

Simulace reality?

Kolik strojového času zabere simulace 1 s vývoje molekulárního systému?

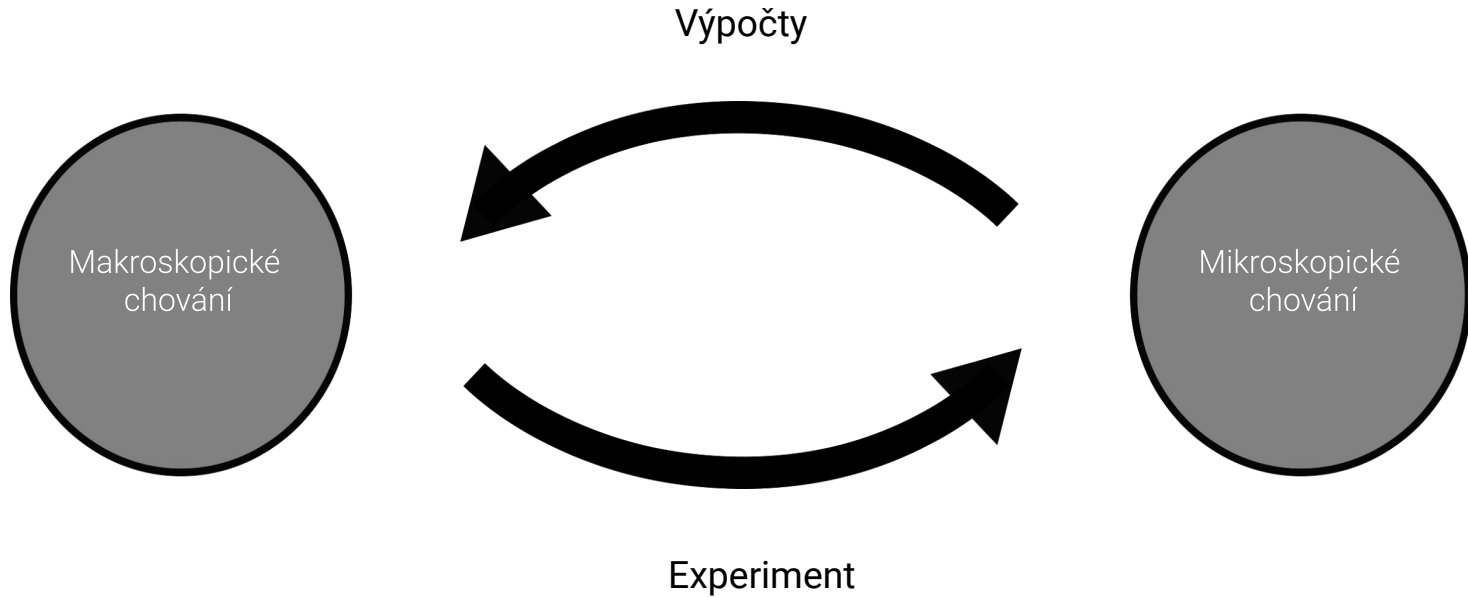
Nejrychlejším molekulárním pohybem je vibrace O-H vazeb s přibližnou periodou 10 fs, která se bude vzorkovat 10 snímky. Výpočet jednoho snímku trvá přibližně 1 ms strojového času.

$$1 \text{ s} = 1\,000\,000\,000\,000\,000 \text{ fs} = 10^{15} \text{ fs}$$

$$1 \text{ ms} = 0,001 \text{ s}$$

$$t = 10^{15} \cdot 0,001 = 10^{12} \text{ s} \Rightarrow 10^{12} / (60 \cdot 60 \cdot 24 \cdot 365) = \mathbf{31\,709 \text{ let}}$$

Simulace reality?



Superpočítače

Co je superpočítač?



Herní PC:
10 jader, 68 GB RAM



Superpočítač

- 1) Photo: NASA Ames Research Center/Tom Trower
- 2) Acer Predator Orion 9000

Superpočítače: Lokální clustery

WOLF: 11 uzlů, 88 CPU, 1 GPU, 245 GB RAM, ~4 TB storage

SOKAR: 16 uzlů, 544 CPU, ~4 TB RAM, 76 TB storage

PIP: 1 uzel, 192 CPU, 3 TB RAM, 52 TB storage



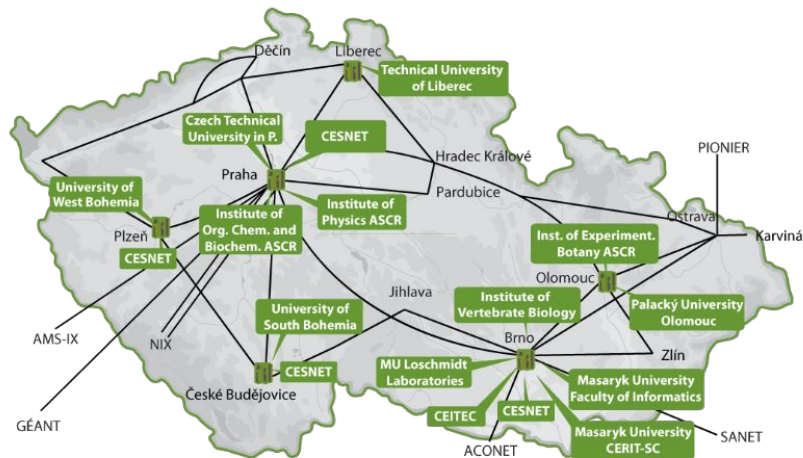
Superpočítače: Česká republika

Zdroje metacentra jsou pro veřejné výzkumné účely zdarma.

Metacentrum:

- koordinátor Národní Gridové Infrastruktury
- integrace výpočetních zdrojů v ČR, jednotné pracovní prostředí

18908 CPU (CEITEC/NCBR 1080 CPU), 28 PB úložných prostorů



1) metacentrum.cz

Superpočítače: Česká republika

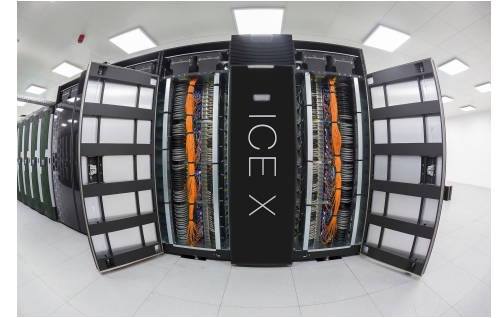
IT4Innovation

- superpočítačové centrum v Ostravě
- budování národního centra excelentního výzkumu v oblasti informačních technologií

Anselm 209 uzlů, 3344 CPU, 15 TB RAM, **94 TFLOP/s**

Salomon 1010 uzlů, 24192 CPU, 129 TB RAM, 2 PFLOP/s

Barbora 201 uzlů, 7232 CPU, 44 TB RAM, **848 TFLOP/s**



1) Photo: IT4Innovation
2) Photo: Jaroslav Ožana, ČTK
3) it4i.cz

Superpočítače: Svět

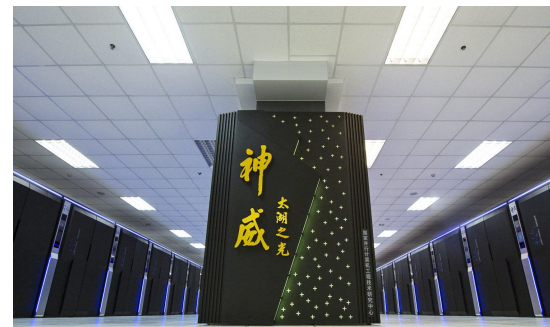
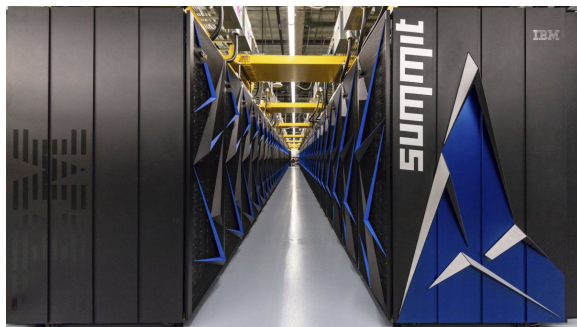
Žebříček nejvýkonnějších počítačů: **TOP500**

Salomon:

11/2019 375. místo

06/2015 39. místo

1. Summit: 2 414 592 CPU, 2 801 TB RAM, **149 PFlop/s**
2. Sierra: 1 572 480 CPU, 1 382 TB RAM, **95 PFlop/s**
3. Sunway TaihuLight: 10 649 600 CPU, 1 310 TB RAM, **93 PFlop/s**

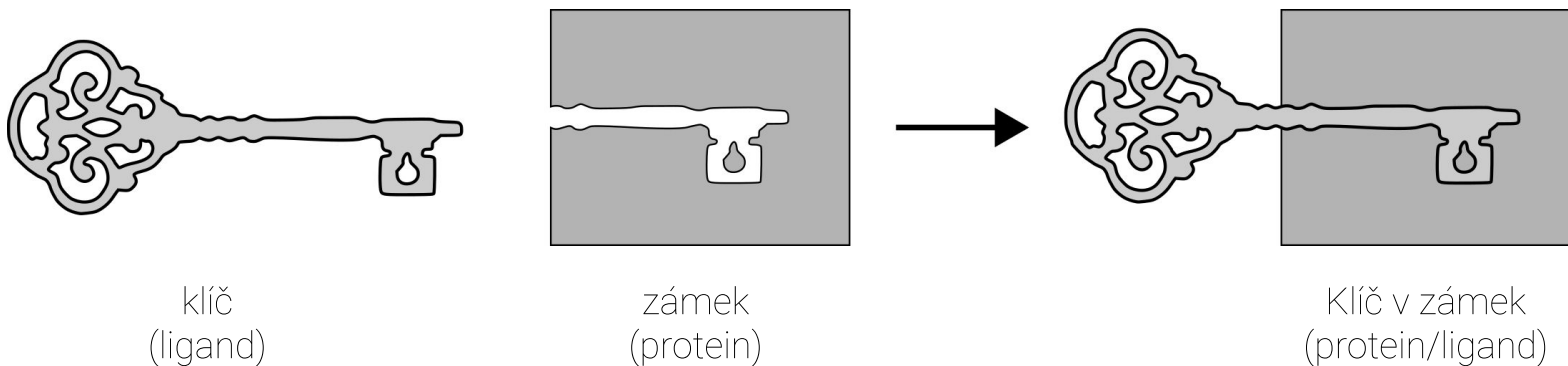


- 1) Photo: IBM
- 2) Photo: Li Xiang/Xinhua via AP
- 3) top500.org

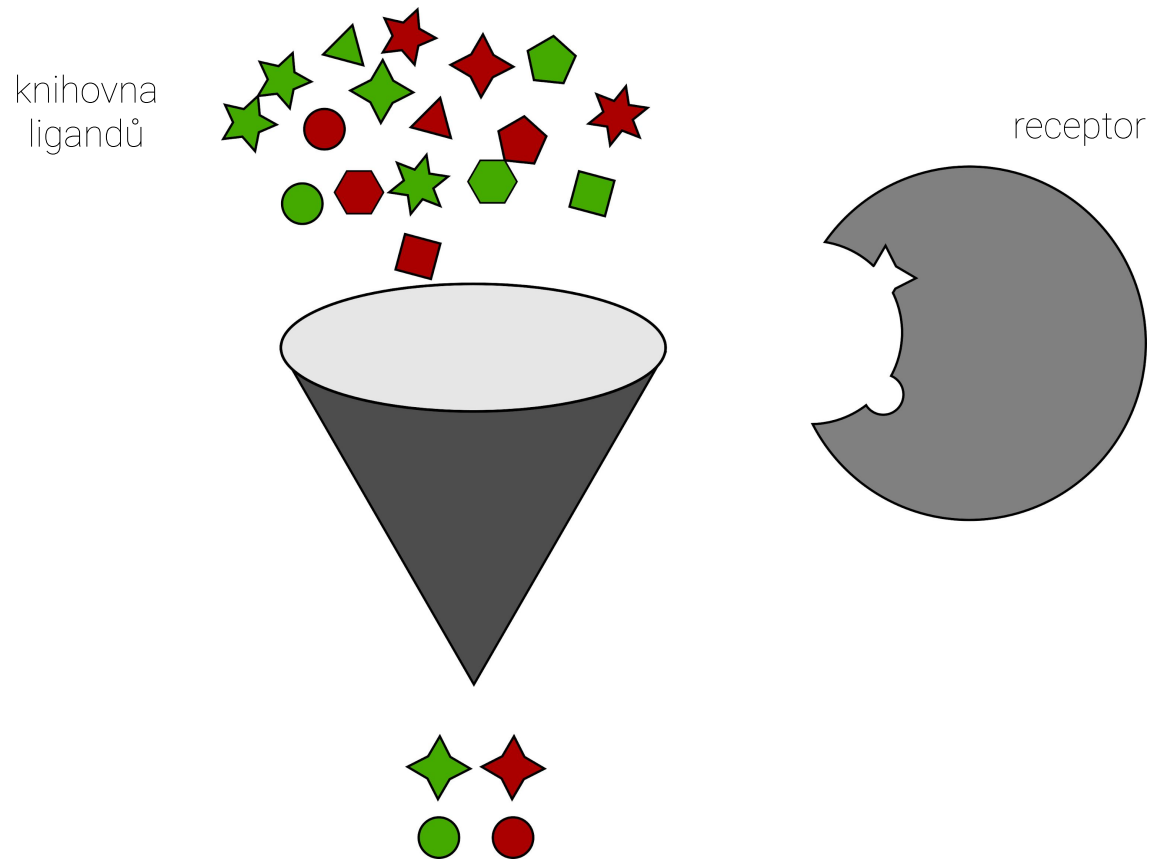
Základní koncepty výpočetní chemie

Jak získat počáteční strukturu?

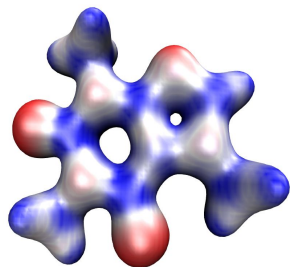
- Stavba molekuly
 - ruční, automatická
- Database:
 - CSD (Cambridge Structural Database): přes **1 000 000** struktur
 - PDB (Protein Data Bank): **153 601** struktur
- Molecular Docking
 - Odhad struktury
 - Odhad energie ("scoring function")



Virtual screening



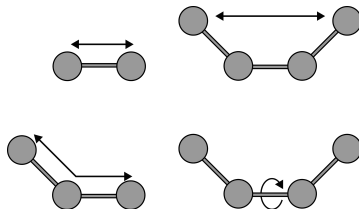
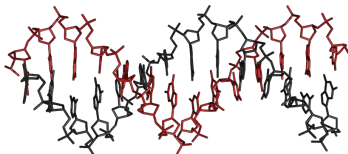
Porovnání QM vs MM vs CG



$$\hat{H}\Psi = E\Psi$$

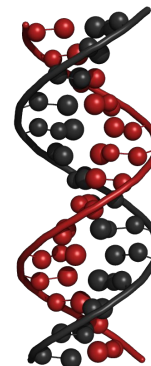
Kvantová mechanika

Atomové rozlišení
Reaktivita
~ 1000 atomů
~ 100 ps



Molekulová mechanika

Atomové rozlišení
Konformační pohyby
~ 1 000 000 atomů
~ μ s



Coarse-grained
mechanika

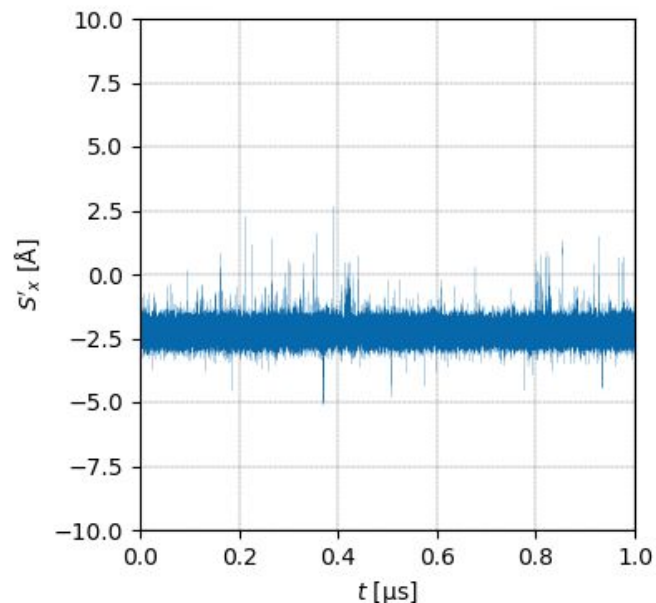
Beads
Pohyby domén, folding
~ 1 000 000 bead
~ ms

aproximace, délka simulace

rozlišení, přesnost

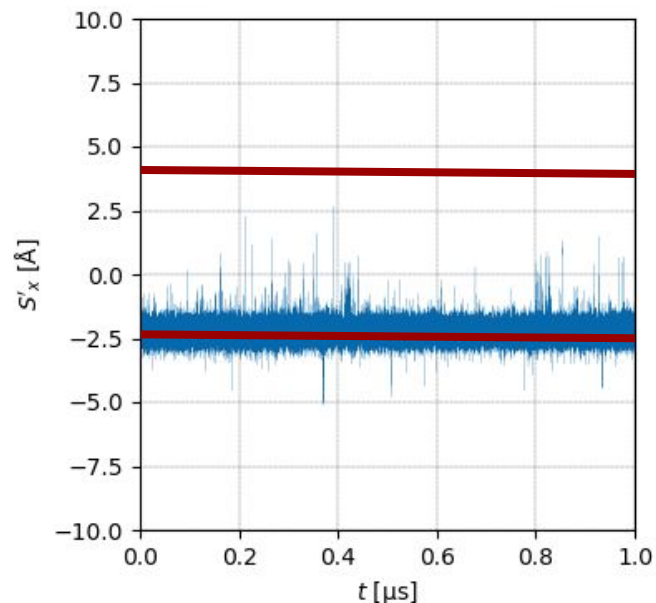
Molekulová dynamika: pohyb molekul

- Aplikace newtonových pohybových rovnic
- Krok pohybu 1/10 nejrychlejšího pohybu
- v MM: 10 fs vibrace R-H vazby



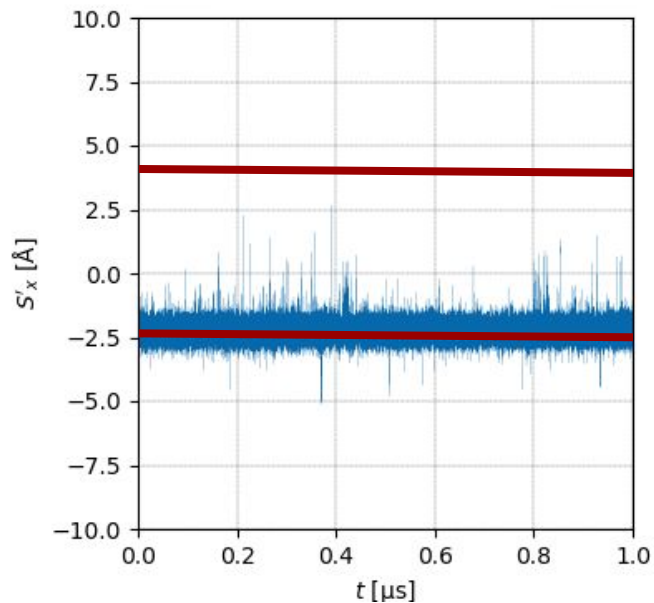
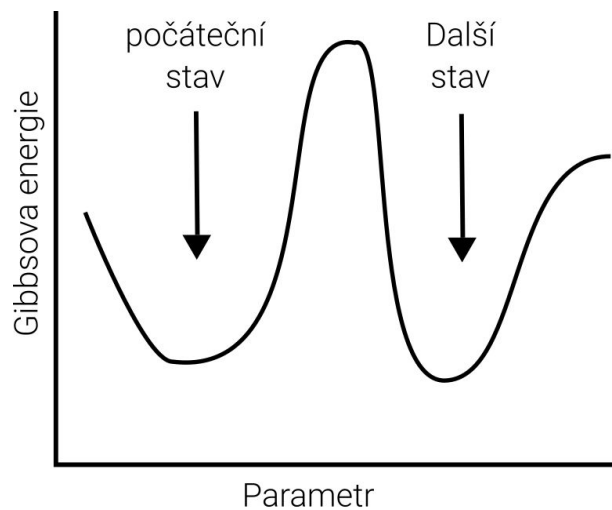
Molekulová dynamika: pohyb molekul

- Aplikace newtonových pohybových rovnic
- Krok pohybu 1/10 nejrychlejšího pohybu
- v MM: 10 fs vibrace R-H vazby



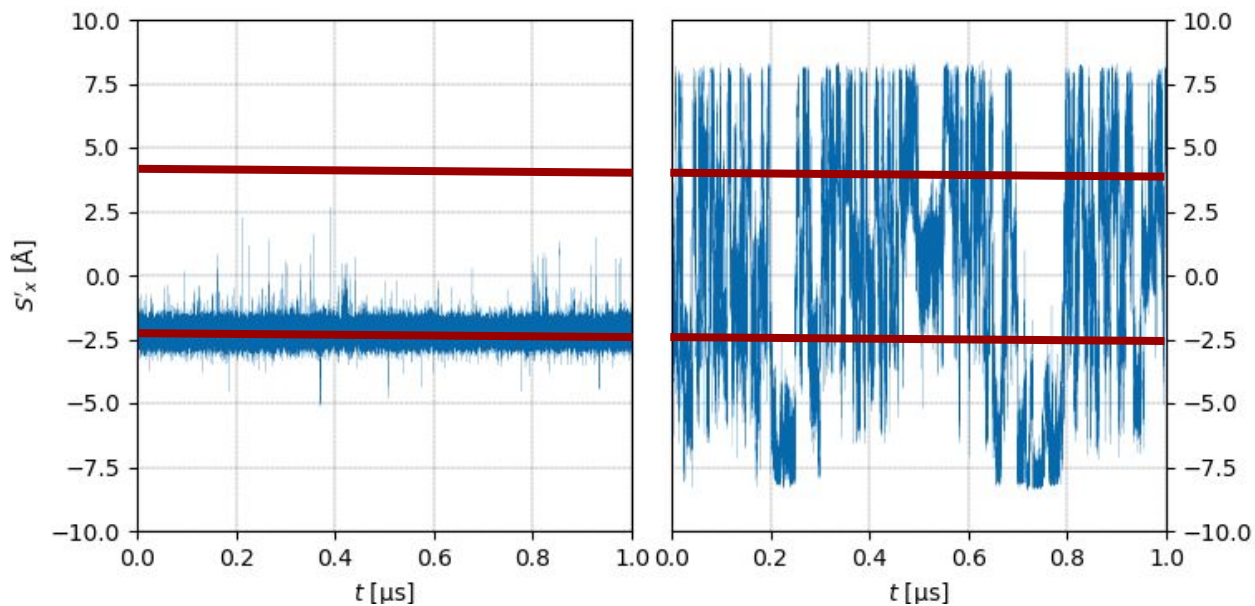
Molekulová dynamika: vzorkování

- Vzorkování (časová škála)
- Termodynamické bariery



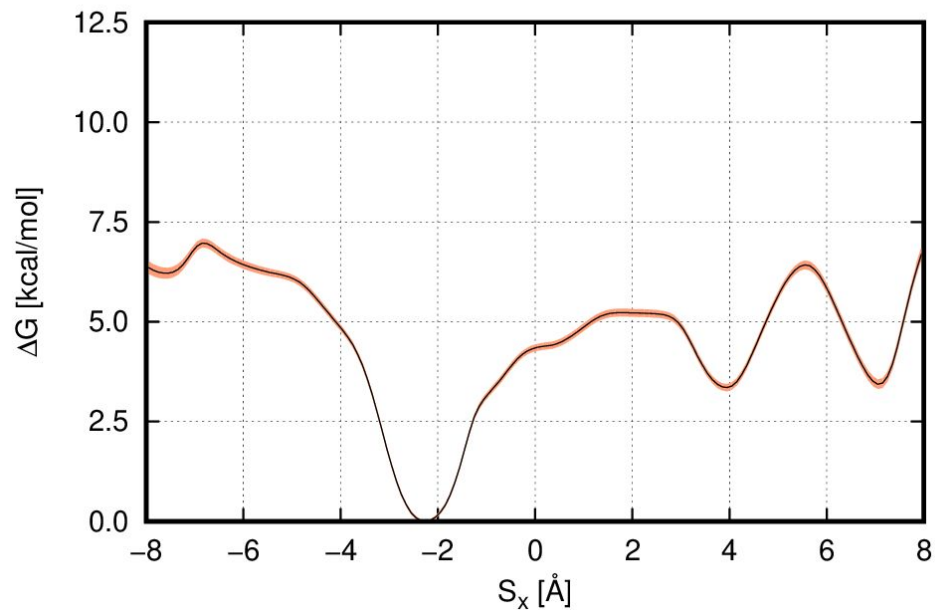
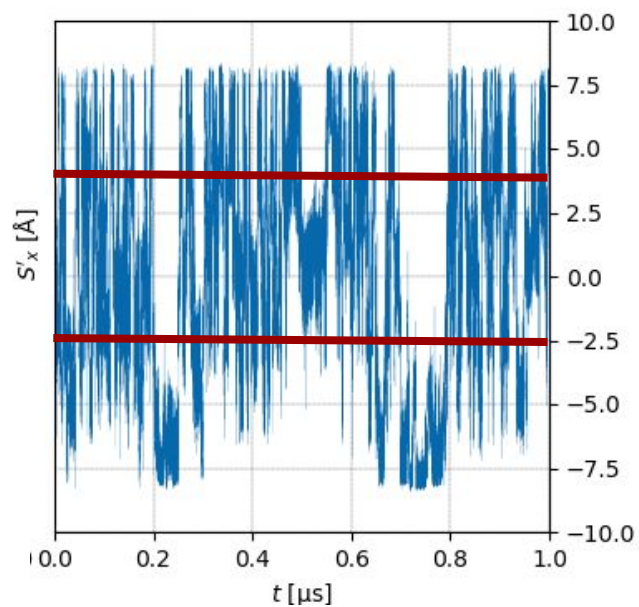
Molekulová dynamika: bias

- Speciální případ molekulové dynamiky
- Přimějeme CV (parametr), aby se hybala volně



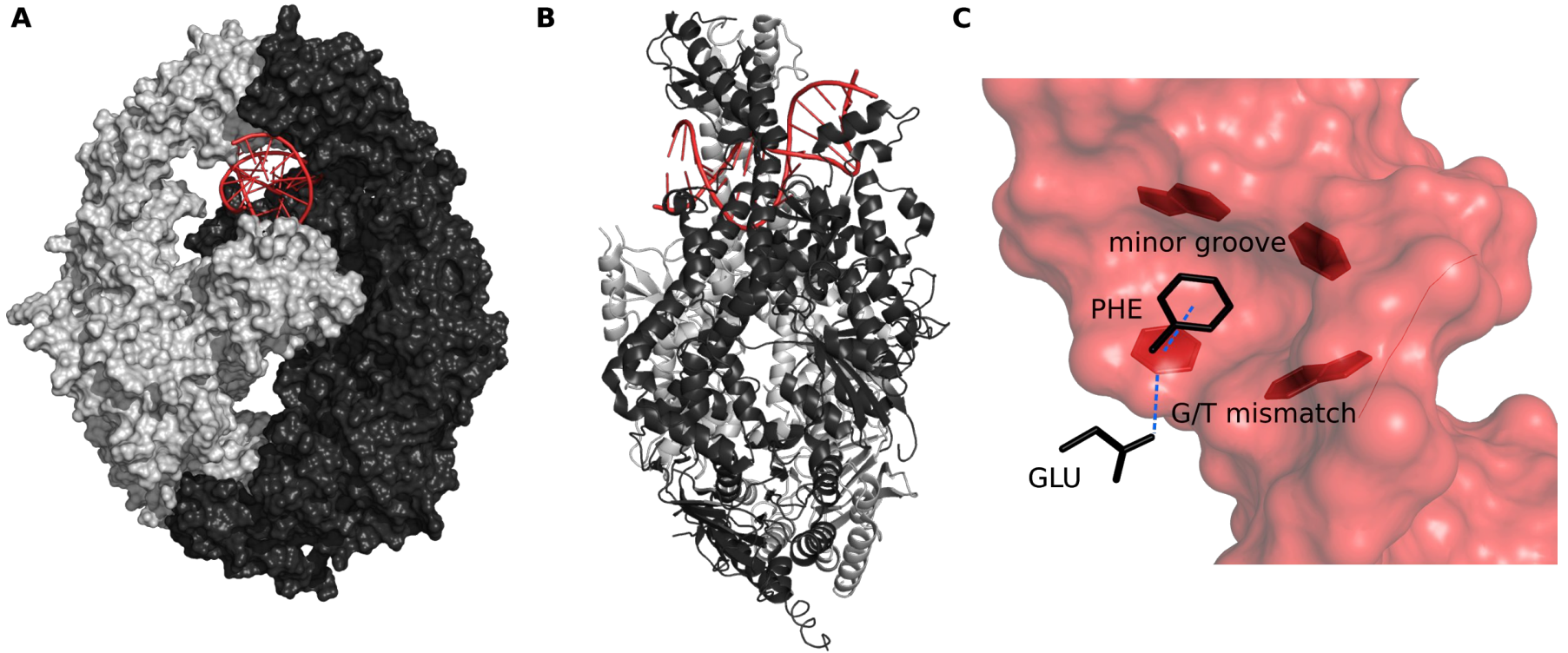
Molekulová dynamika: bias

- Speciální případ molekulové dynamiky
- Přimějeme CV (parametr), aby se hybala volně (biased MD)

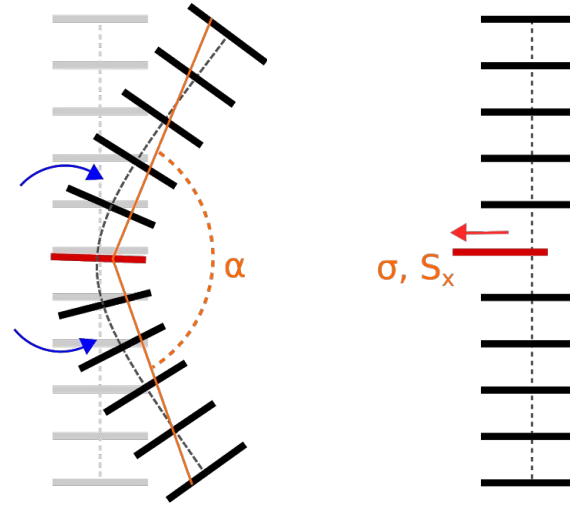
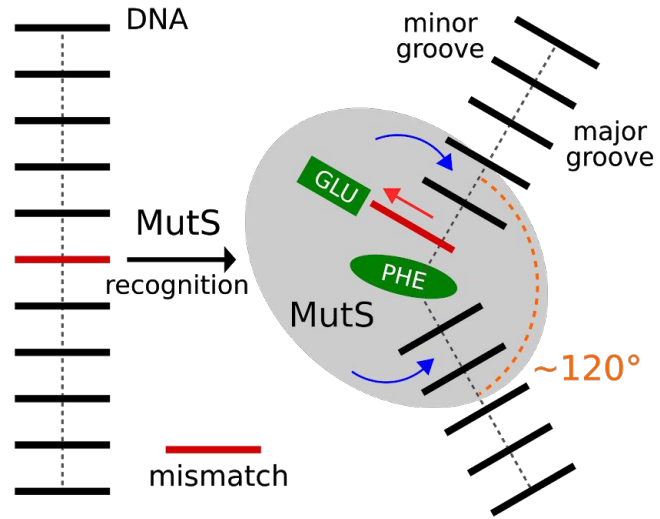


Ukázka projektu

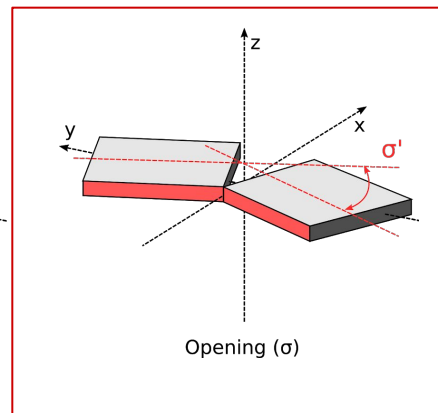
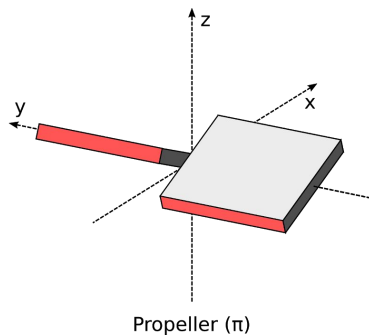
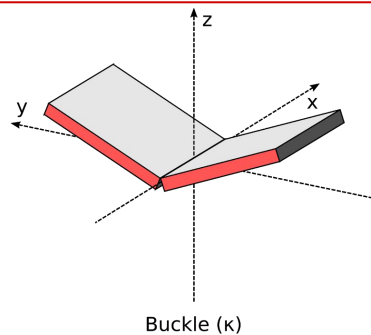
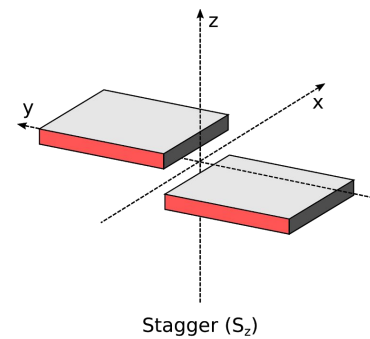
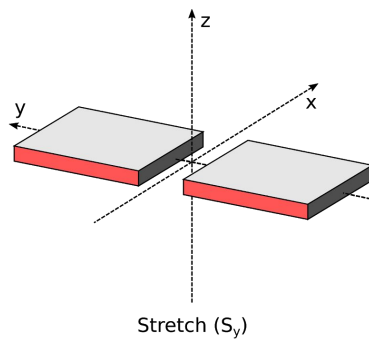
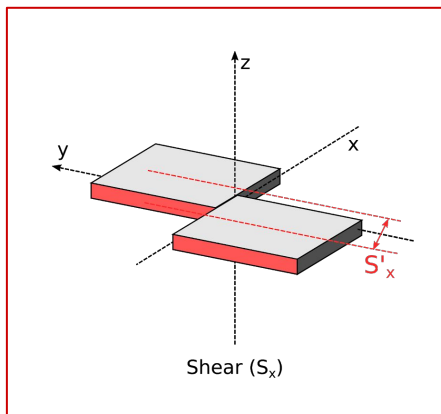
Mechanismus rozpoznání chyb v DNA



Aproximace

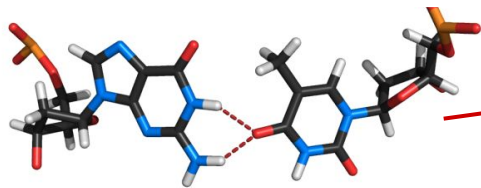


Geometrický popis páru bazí

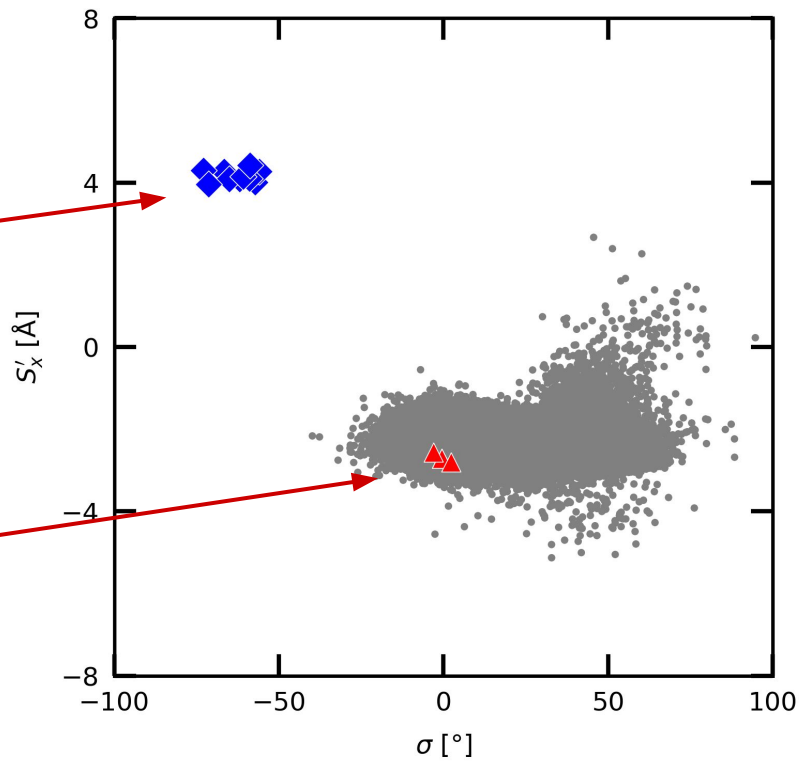


X-ray struktury a MD

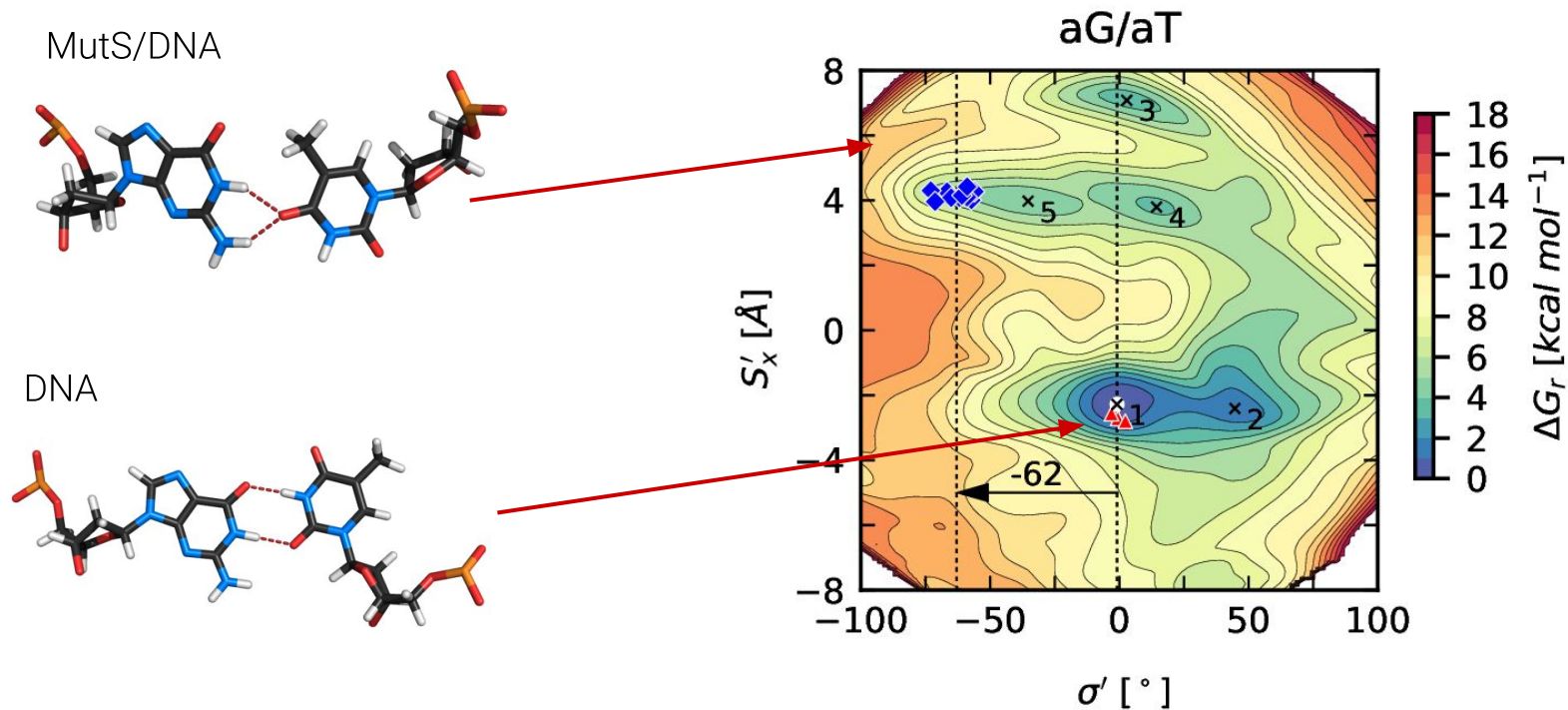
MutS/DNA (GT mismatch)



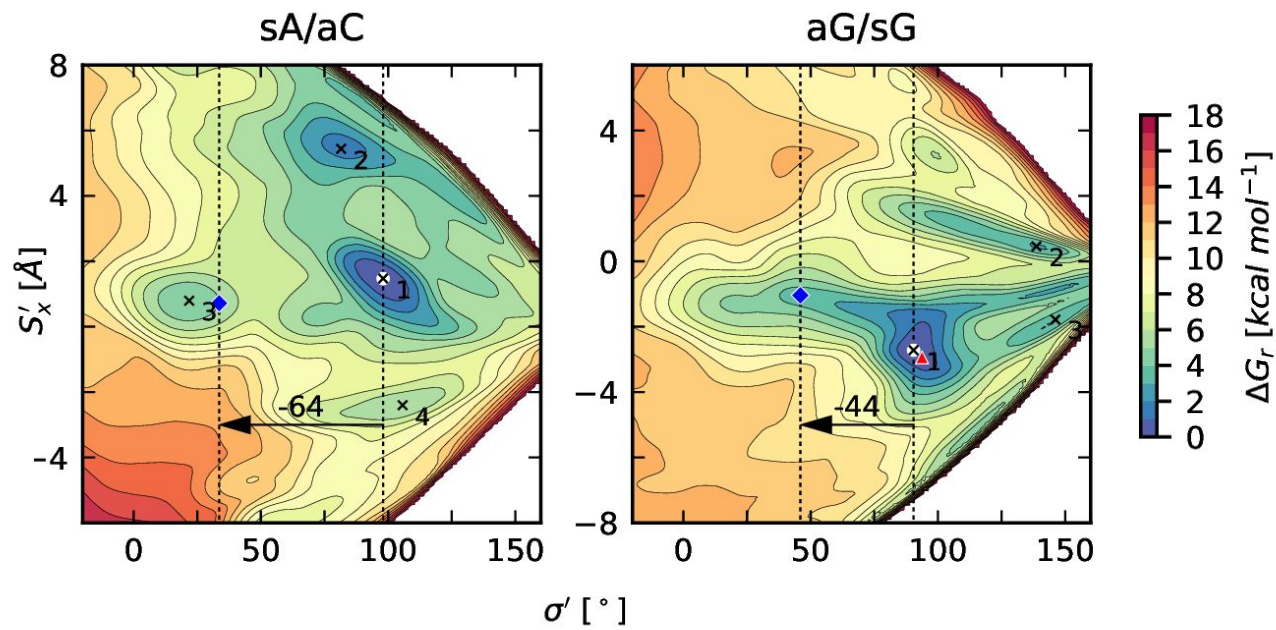
DNA (GT mismatch)



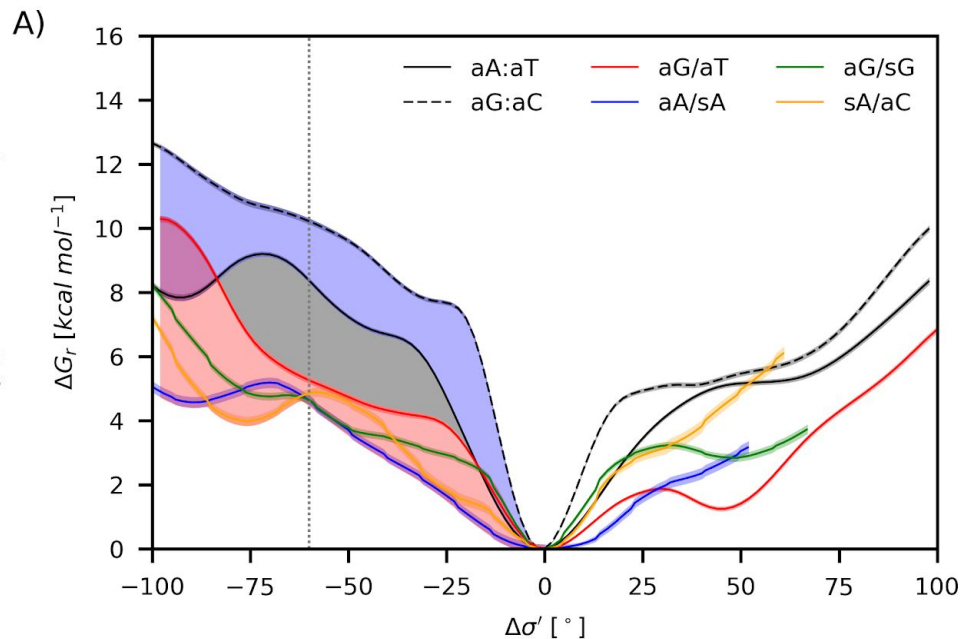
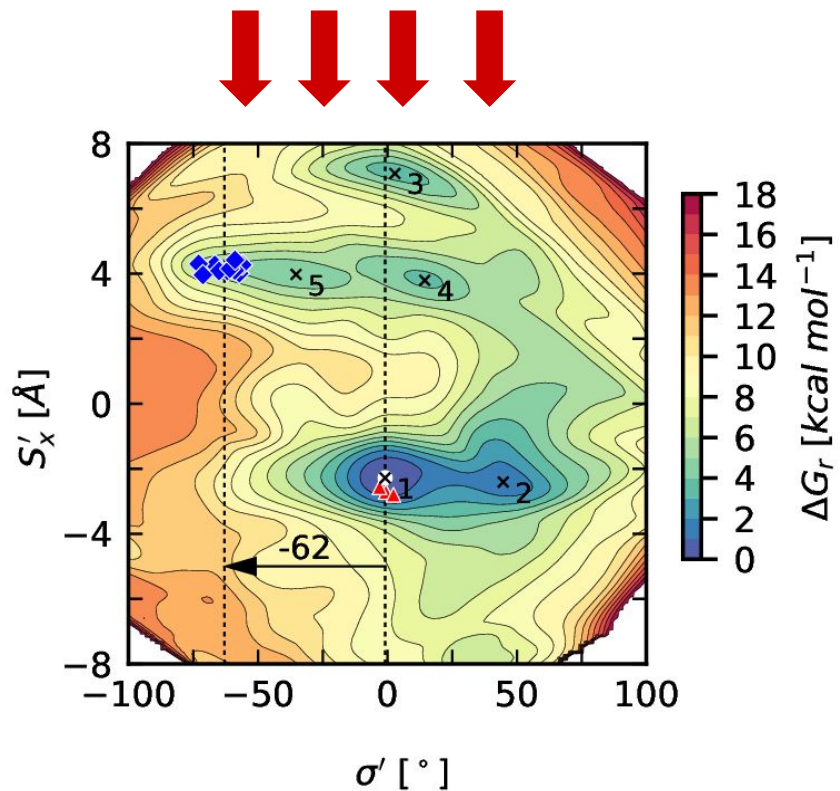
MD bias - Gibbssova energie



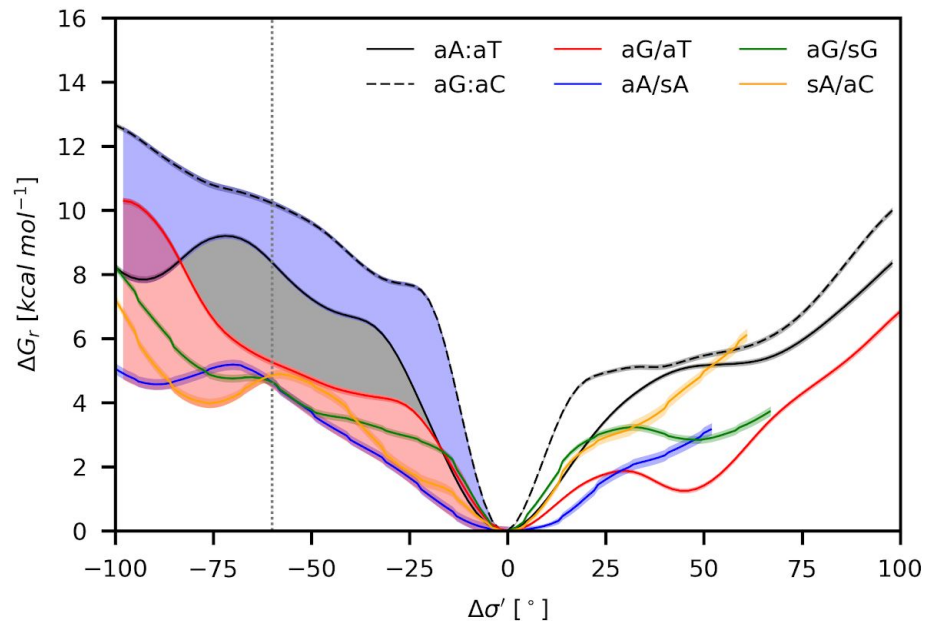
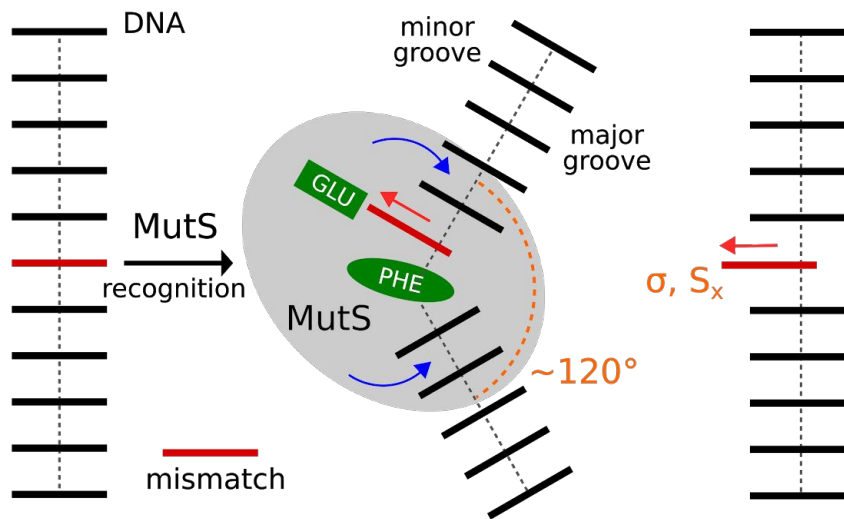
MD bias - Gibbssova energie



Opening



Opening



Co jiného můžeme spočítat?

- Predikce struktur
- Kinetické vlastnosti
- Výpočet vlastností (Chemický posun, spektroskopické vlastnosti)
- Predikce reakčních mechanismů
- Vazebné energie
- ...

Na závěr

Výpočetní chemie a Nobelovy ceny

2013



Martin Karplus



Michael Levitt



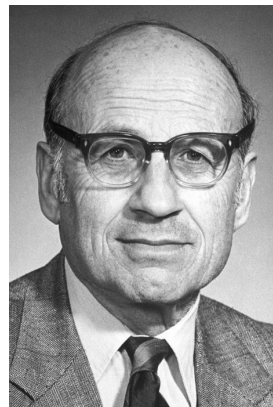
Arieh Warshel

For the development of multiscale models for complex chemical systems.

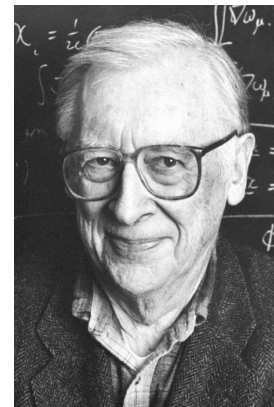
1) Photo: A. Mahmoud, Nobel Media AB

2) Photo: Nobel Foundation archive.

1998



Walter Kohn

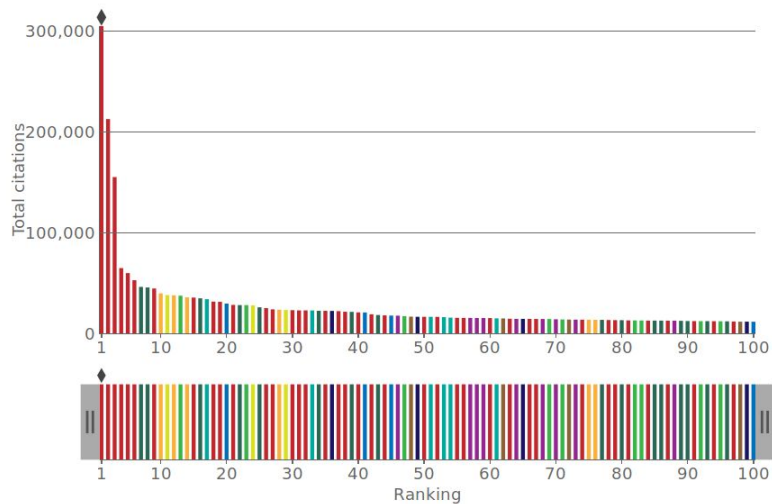


John A. Pople

W. K.: For his development of the density-functional theory.

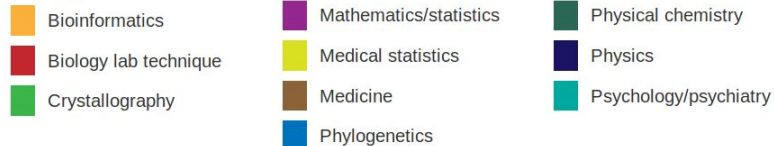
J.A.P.: For his development of computational methods in quantum chemistry.

Výpočty mezi Top 100 články



12/100 (2/10) článku s DFT výpočty

Discipline colour



Bioinformatics (Clustal, Blast)

1) Van Noorden,R., Maher,B. and Nuzzo,R. (2014) The top 100 papers. *Nature*, 514, 550–553.

Nové poznatky

- Výpočetní a experimentální chemie se doplňují.
- Nemůžeme simulovat realitu, musíme použít aproximace.
- Co je to superpočítač
- Nějaké ty metody

Děkuji za pozornost

“A theory is something nobody believes, except the person who made it. An experiment is something everybody believes, except the person who made it.”

Albert Einstein

