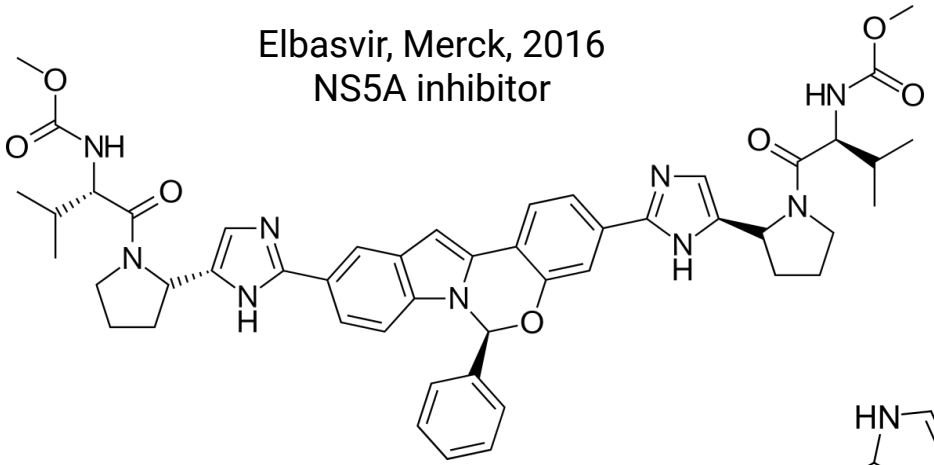
How to achieve the greater realism in predicting the structures of bioactive molecules *in silico*?

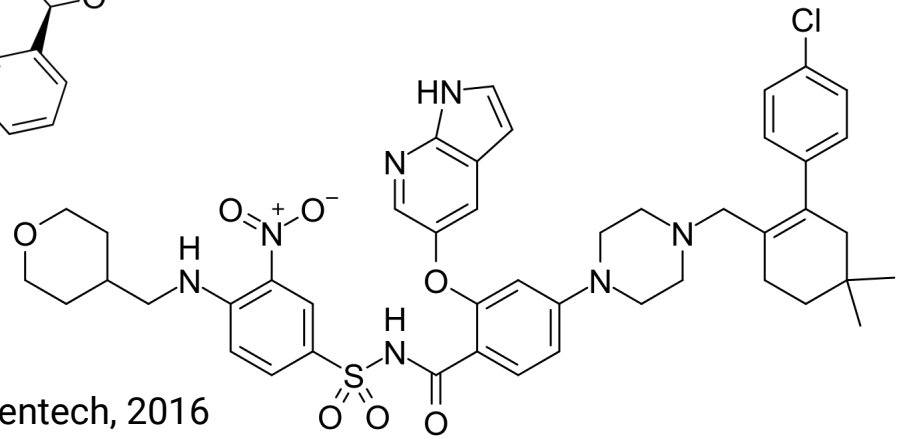
Created by:
Yaroslav Solovev, Ph.D. student, IBCh RAS,
Pavel Kulyaev, MSc, ITMO university,
Yuri Porozov, ass. prof., First MSMU Sechenov,
Mikhail Polynski, jun. res. assoc., ZIOC RAS & SPbU,
March 2020

Соловьев Ярослав Владимирович
E-mail: yaroslavsolovev78@gmail.com

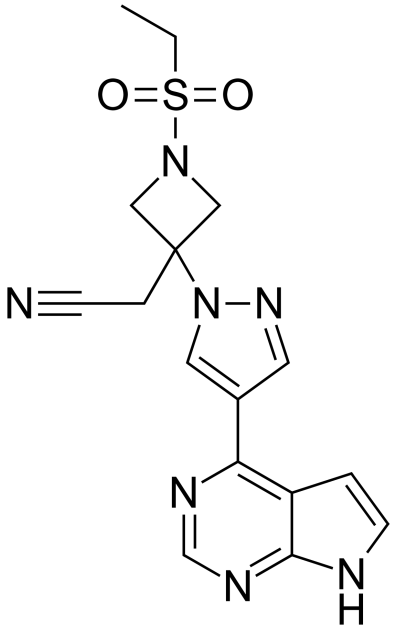
Challenges of modern pharm chemistry

Elbasvir, Merck, 2016
NS5A inhibitor





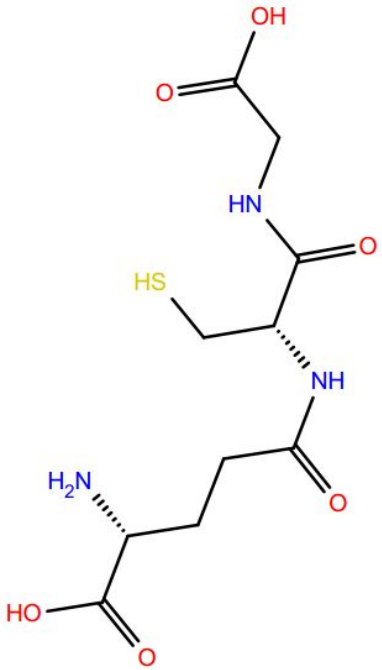
Venetoclax, AbbVie & Genentech, 2016
Bcl-2 inhibitor



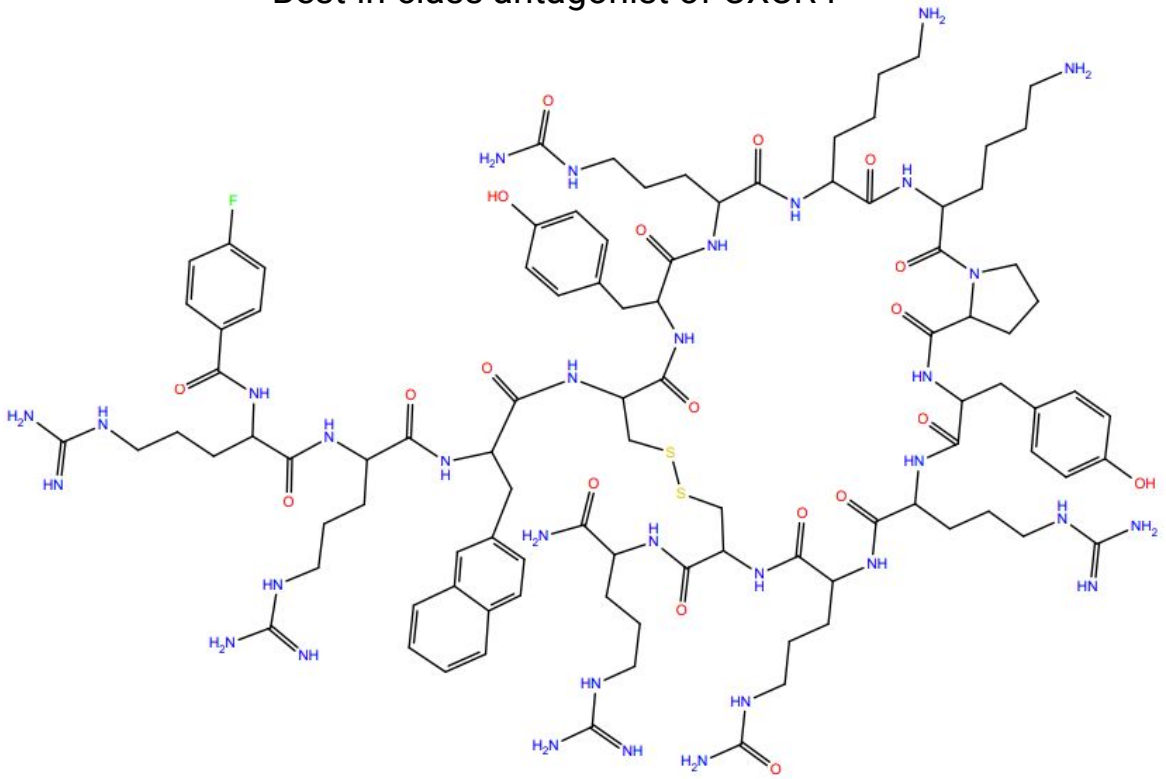
Baricitinib, Lilly, 2017
Jak 1/2 inhibitor

Comparative size of cyclic peptide drugs

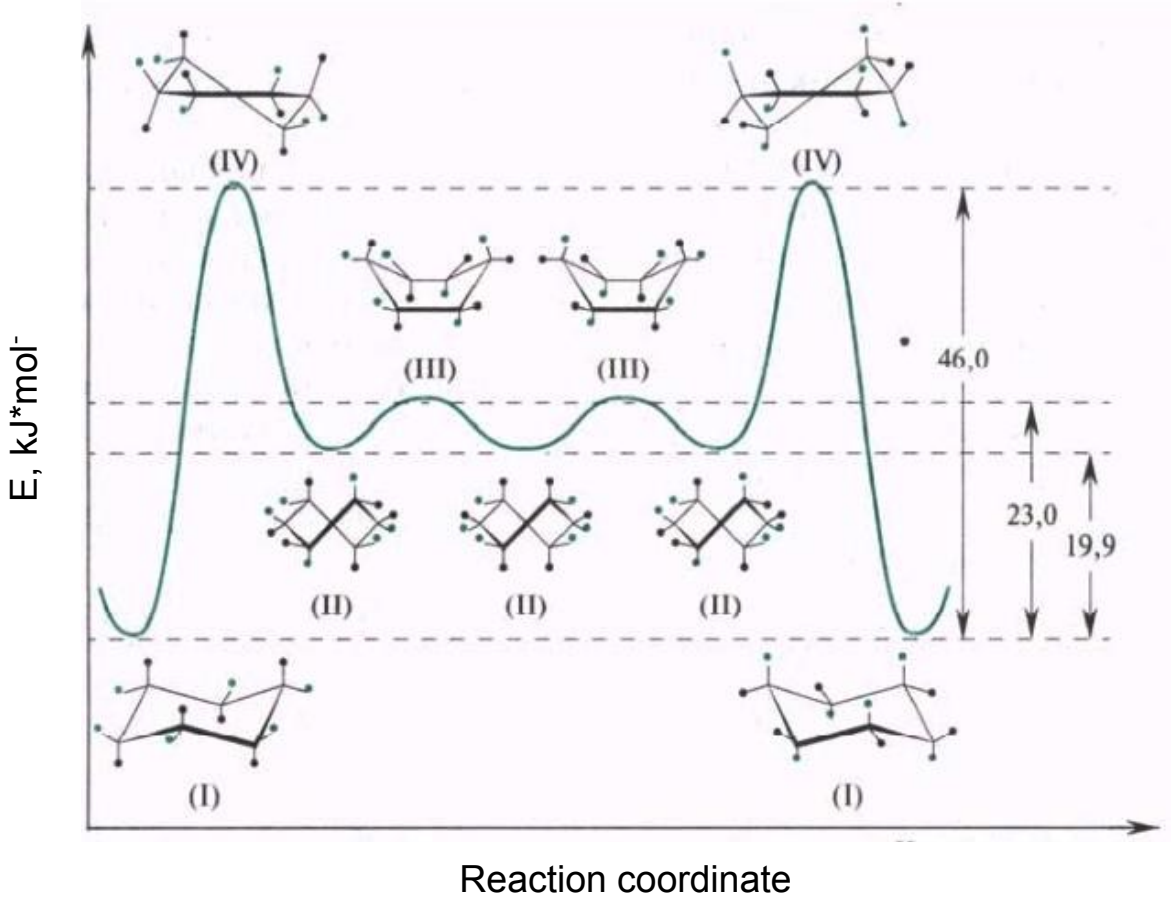
Glutathione-SH



Motixafortide (BL-8040), 2017
Best-in-class antagonist of CXCR4



Conformer distribution



Maxwell-Boltzmann statistics equation

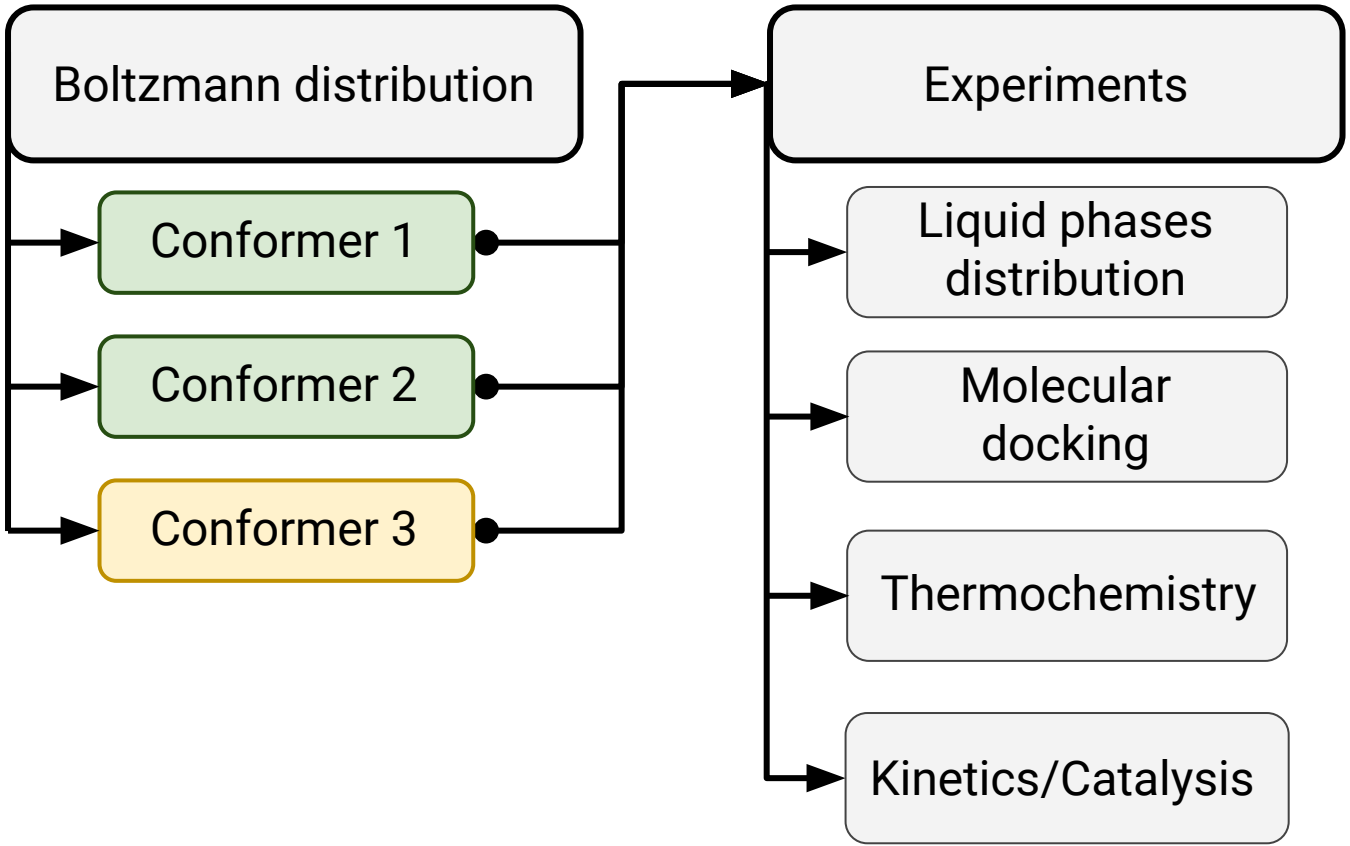
$$P(s) = \frac{1}{Z} e^{-E(s)/kT},$$

where $$Z = \sum_s e^{-E(s)/kT}$$

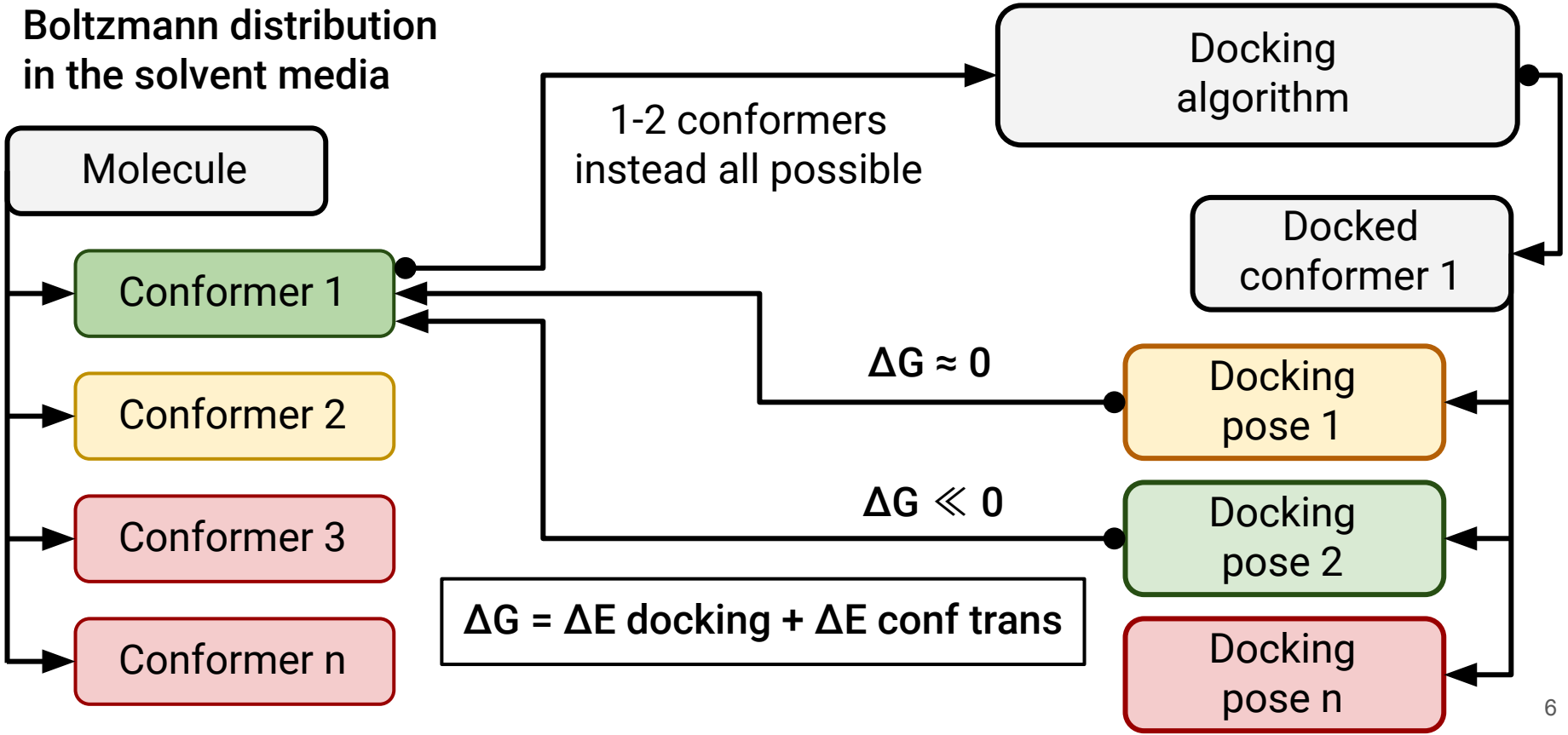
P(II) = 0,05%

P(I) = 99,95%

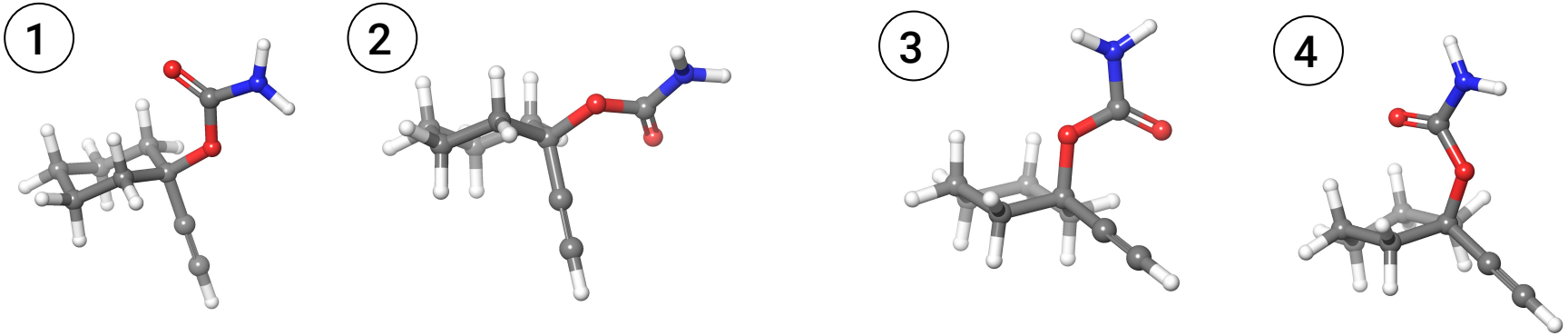
Application of top conformers

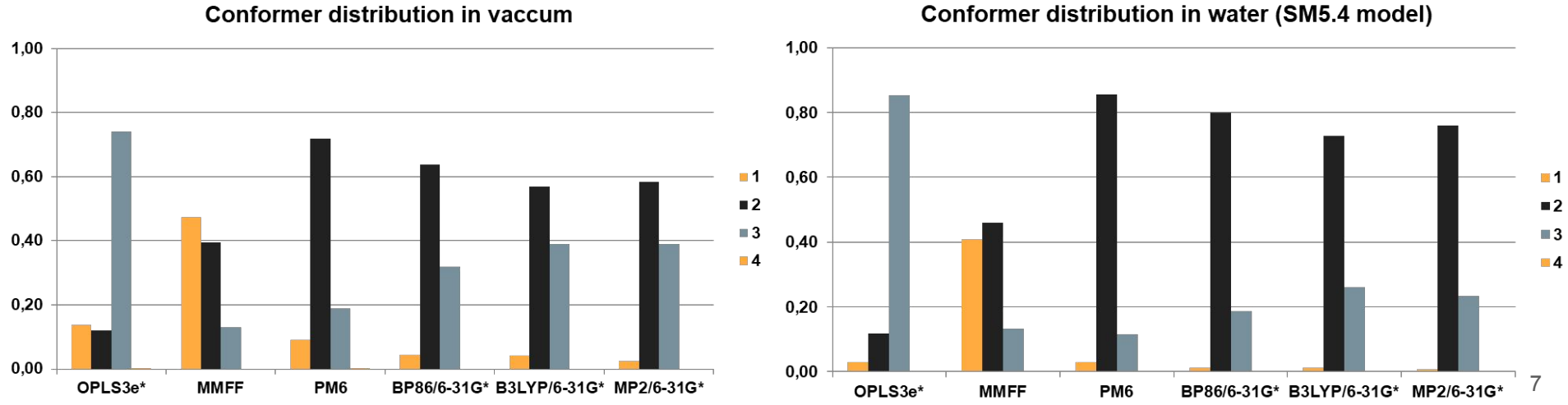


Additional corrections to docking

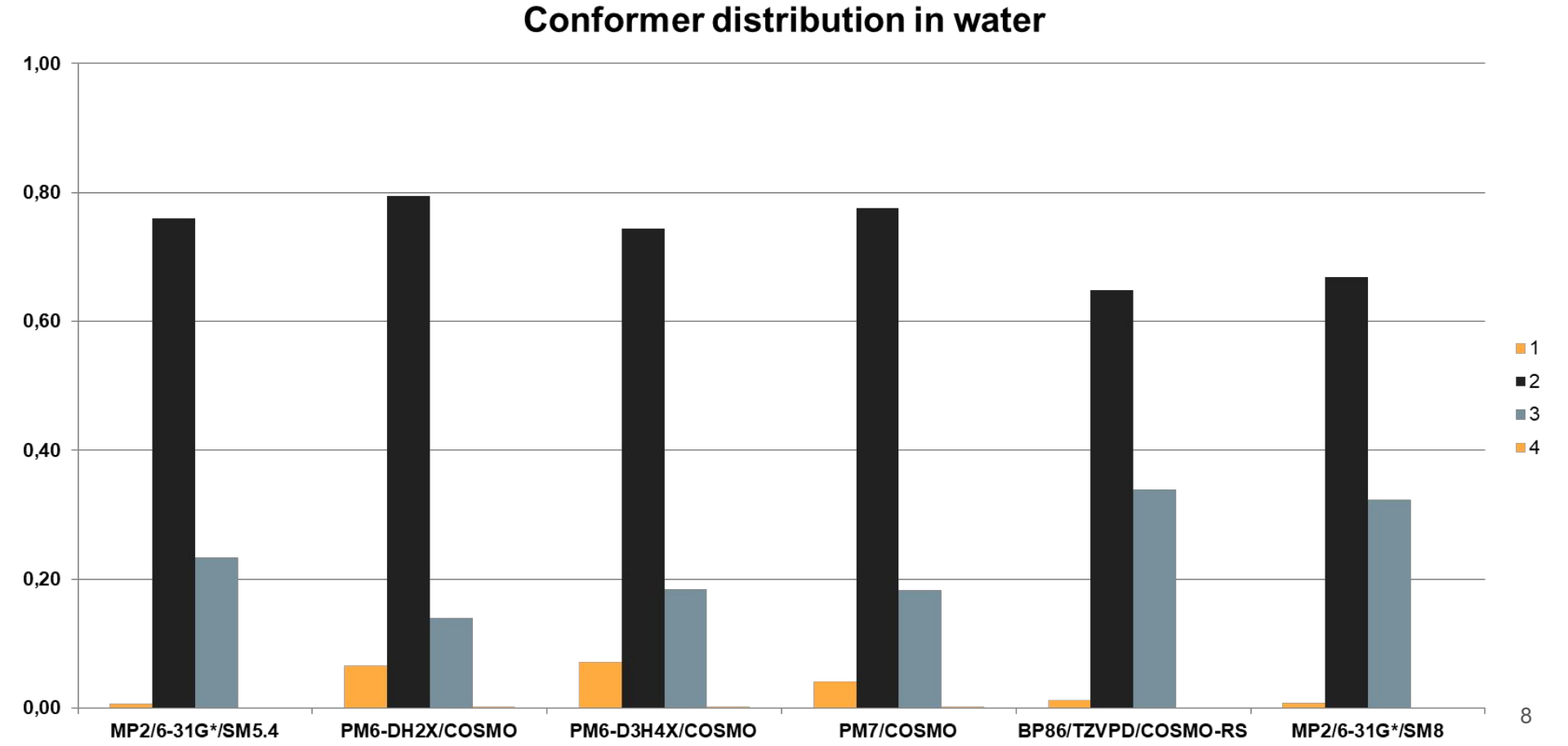


Ethinamate

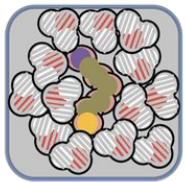

1
2
3
4

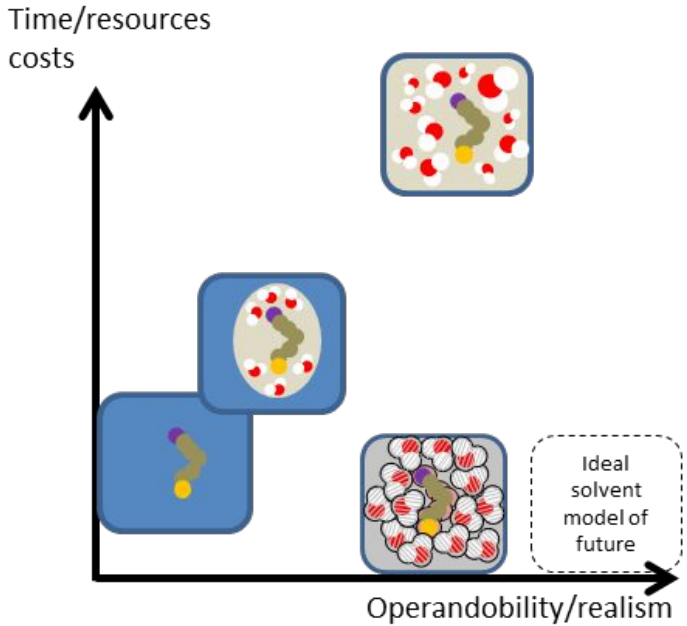

Conformer distribution in vaccum
1,00
0,80
0,60
0,40
0,20
0,00
OPLS3e*
MMFF
PM6
BP86/6-31G*
B3LYP/6-31G*
MP2/6-31G*
1
2
3
4
Conformer distribution in water (SM5.4 model)
1,00
0,80
0,60
0,40
0,20
0,00
OPLS3e*
MMFF
PM6
BP86/6-31G*
B3LYP/6-31G*
MP2/6-31G*
1
2
3
4

Ethinamate


Conformer distribution in water
1,00
0,80
0,60
0,40
0,20
0,00
MP2/6-31G*/SM5.4
PM6-DH2X/COSMO
PM6-D3H4X/COSMO
PM7/COSMO
BP86/TZVPD/COSMO-RS
MP2/6-31G*/SM8
1
2
3
4

Solvent models

		Applications/features	Key limitations
Implicit solvent model		• Solvation energy of exact structure • Fast low demanding simulations • Solute-continuum interactions are considered	• Pure solvents, high-diluted solutions • Applicability depends on the method • Solvent-solute interactions are not considered explicitly
QM/MM hybrid solvent model		• Solvation energy of exact structure • Fast low demanding simulations • Key solvent-solute interactions are considered explicitly	• The result depends on the experience of a researcher • Applicability depends on the method • Each simulation need to be specially tuned by researcher
Statistical hybrid solvent model		• Solvation energy of exact structure • Fast low demanding simulations • All interactions are considered according to their statistical weights	• To increase accuracy, some parameters need to be taken from other simulations/experiments • Applicability depends on parameterization
MD explicit solvent model		• Energy of the whole system • Demanding simulations • All interactions are considered	• Computational costs • Applicability depends on the force field • Recalculation needs restart of simulation

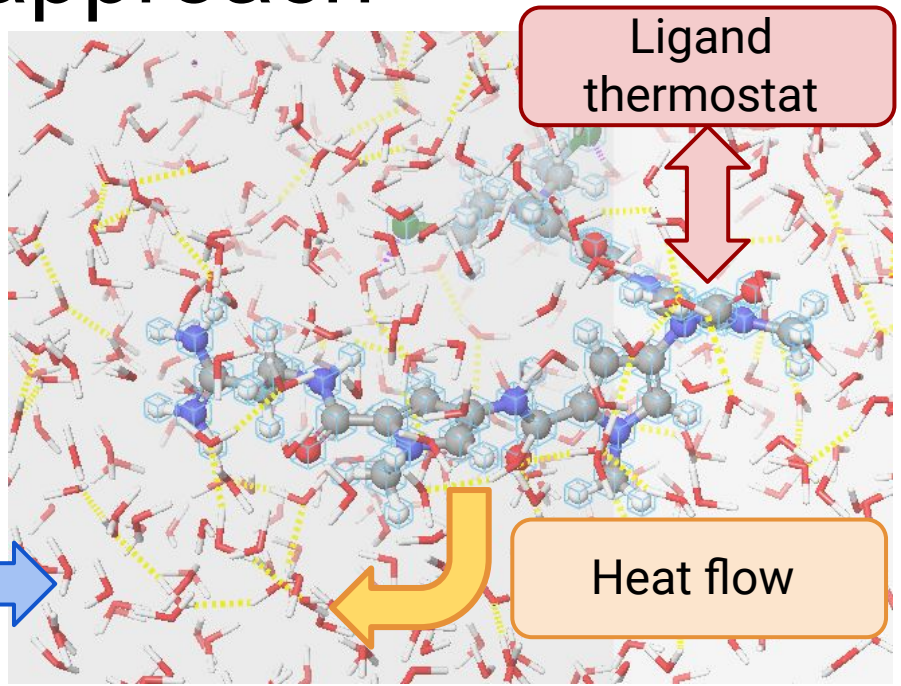


Kuliaev, Pavel O., et al. "Operando solvation thermodynamics simulations examples and perspectives" preprint

Hybrid MD approach

Separate thermostats strategy:

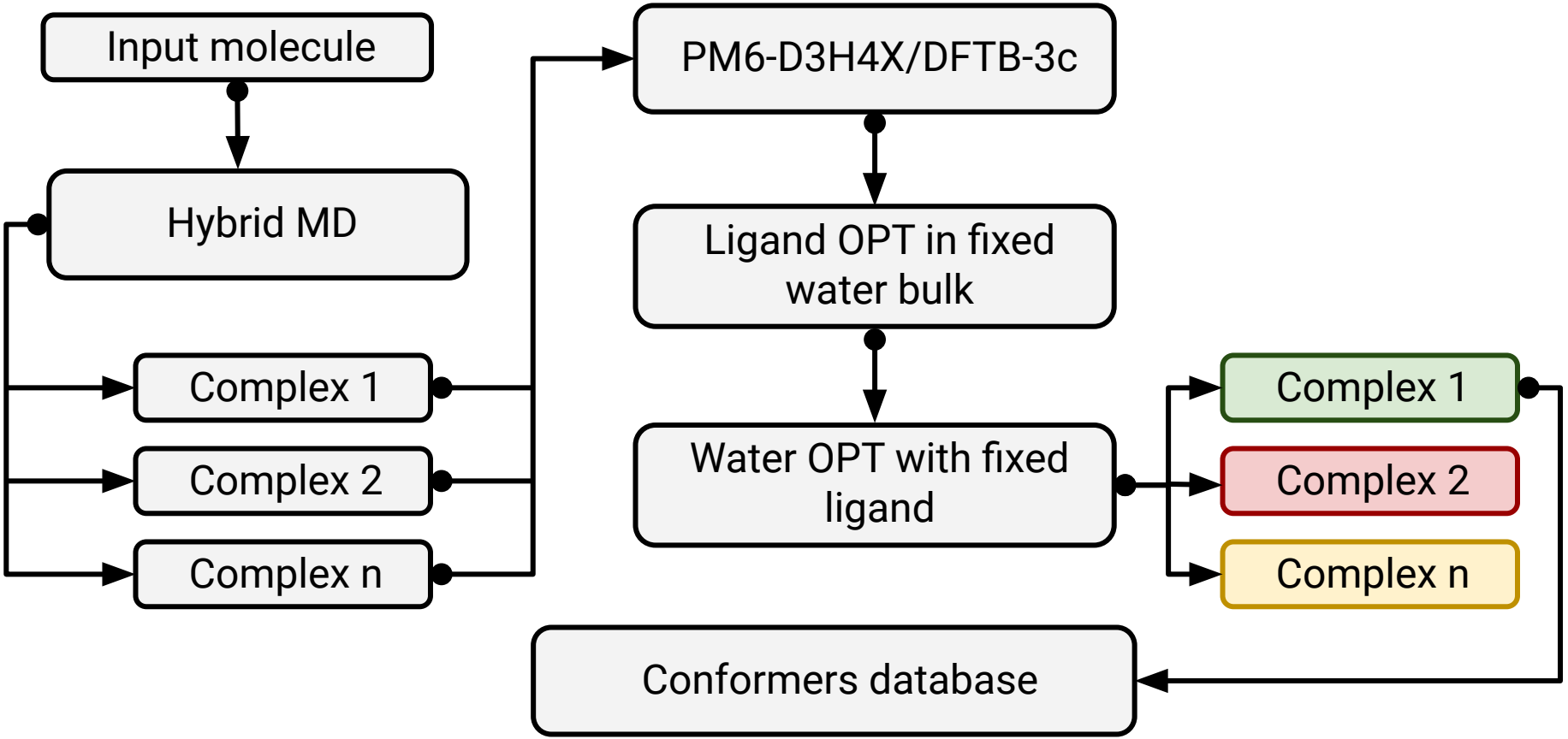
- NVT ensemble
- Hose-Hoover chain
- Coupling strength = 20ps
- Ligand temperature = 700K~1000K
- Solvent (water) and ions temperature = 300K
- Water molecules: TIP3P (TIP4P)
- Box size: 60*60*60Å (min)



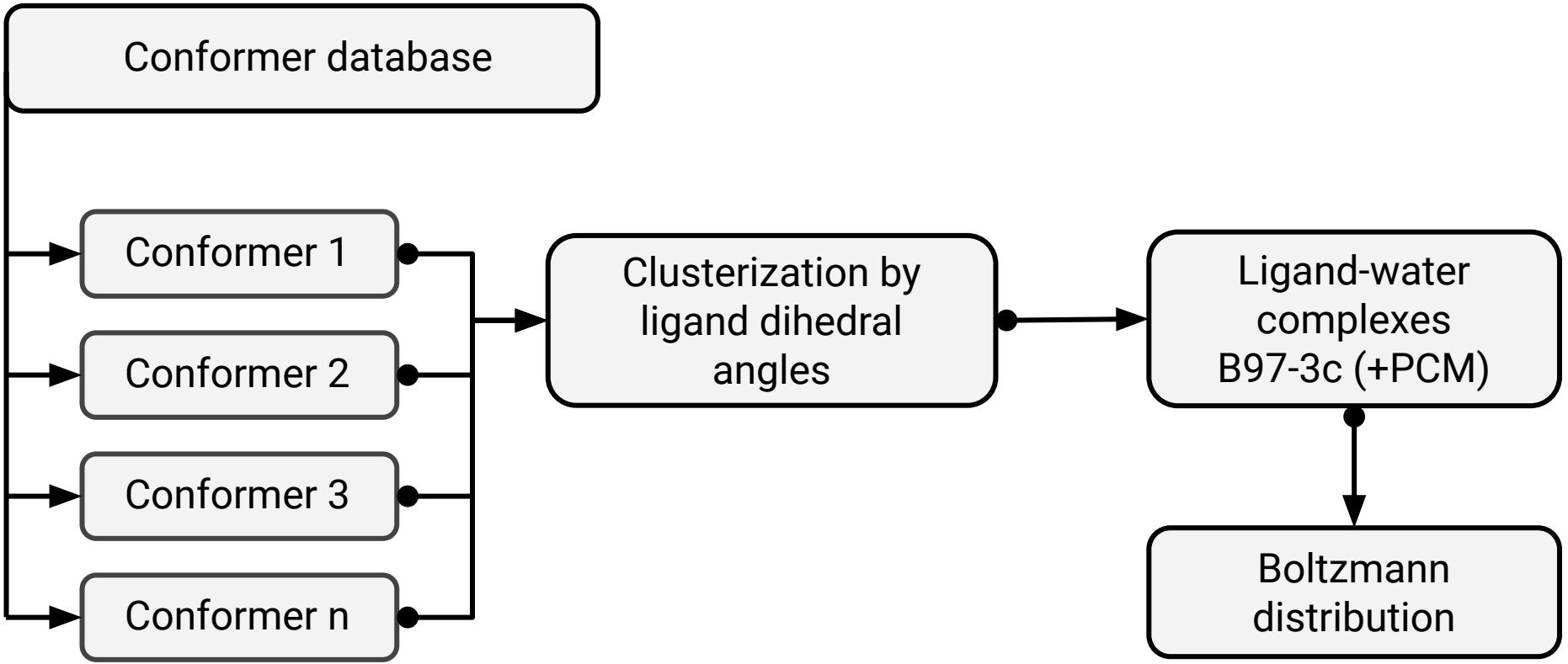
Non-canonical ensemble!
Solvent boiling effects near the ligand

Basconi, Joseph E., and Michael R. Shirts. "Effects of temperature control algorithms on transport properties and kinetics in molecular dynamics simulations." Journal of chemical theory and computation 9.7 (2013): 2887-2899.

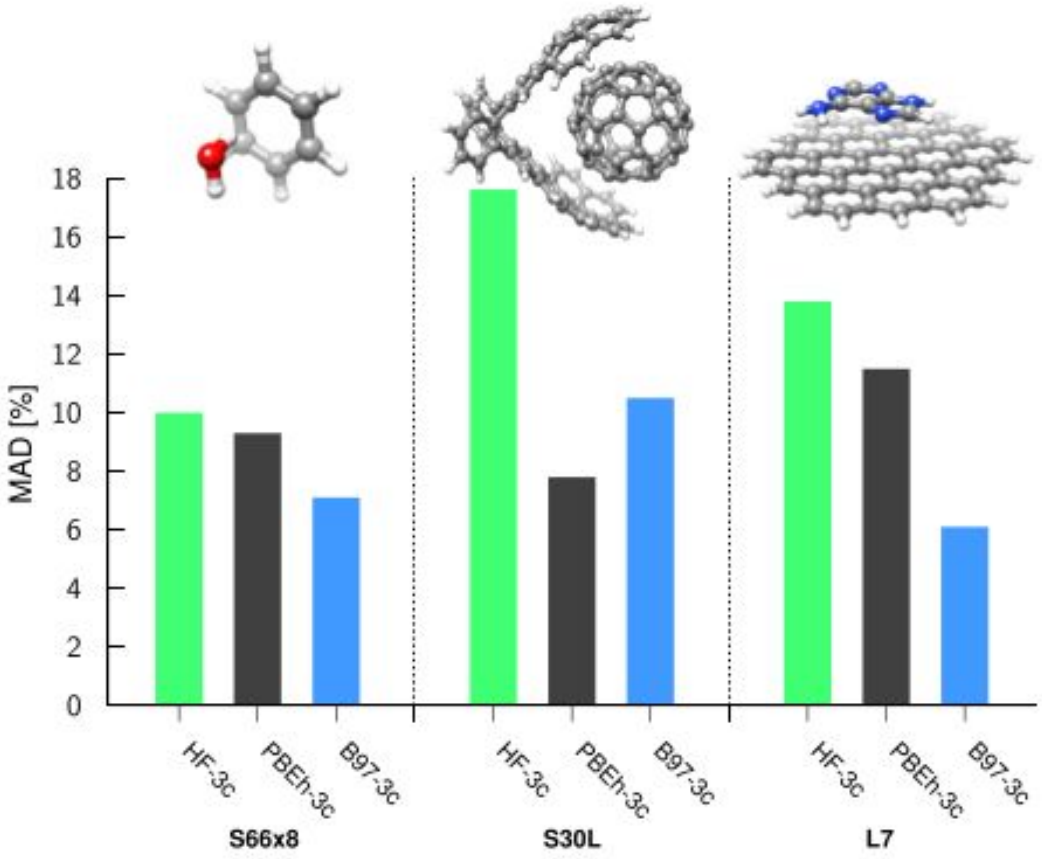
Treating MD ligand-water complexes



Conformer clustering funnel



Why B97-3c?



Mean absolute relative deviations (MADs) of the three composite methods for the noncovalent interaction benchmark sets S66x8 (molecular dimers), S30L (host-guest complexes), and L7 (large molecular complexes).

Why B97-3c?

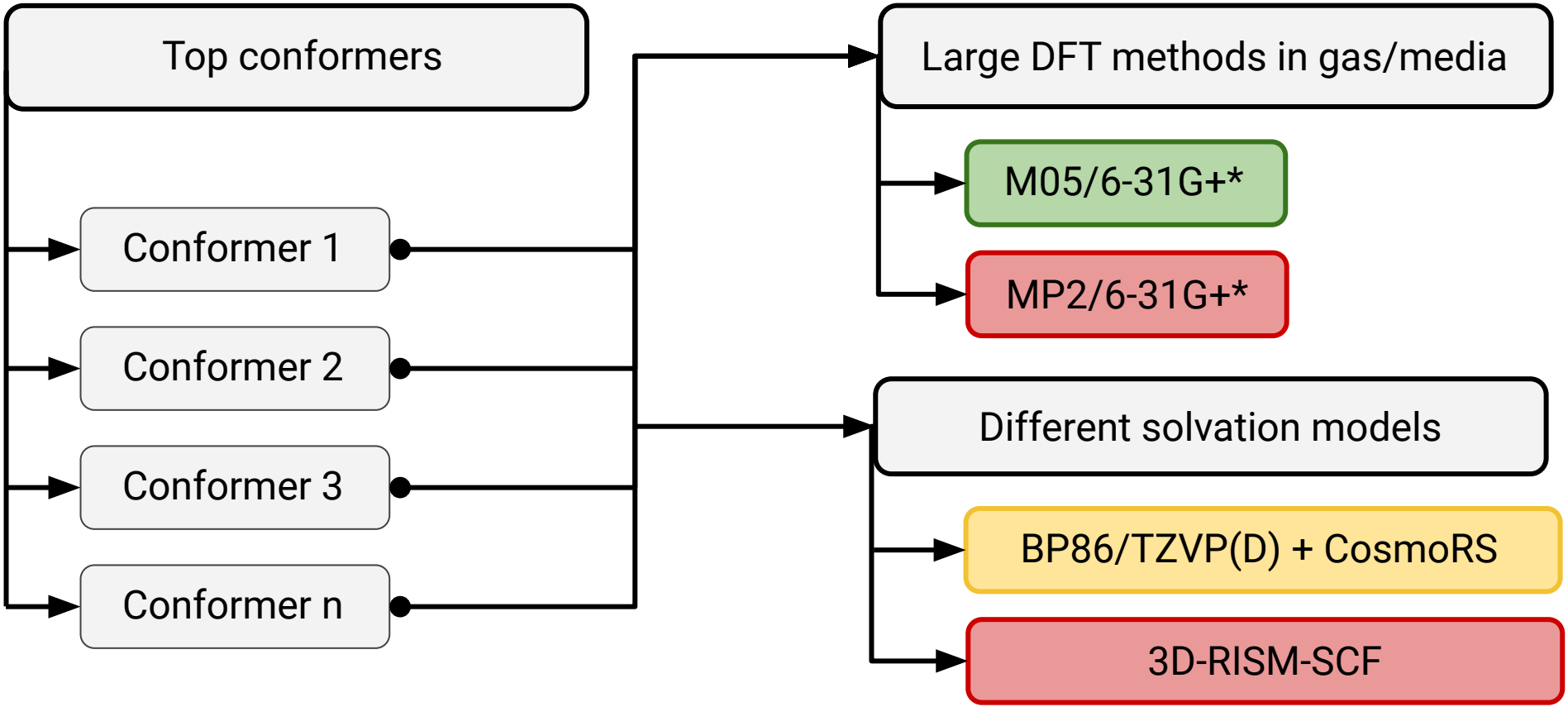
	Phenanthrene		[7]helicene		[16]helicene	
	No. of orbitals	Time	No. of orbitals	Time	No. of orbitals	Time
HF-3c	80	00:00:02	168	00:00:16	366	00:05:02
B97-3c[a,b]	296	00:00:45	624	00:03:18	1362	00:18:26
PBEh-3c[a,b]	230	00:02:05	486	00:13:29	1062	02:04:41
BP86-D3^{atm}/def2-TZVP[a,b]	494	00:01:27	1038	00:07:09	2262	00:42:56
PBE-D3^{atm}/def2-QZVP[a,b]	1098	00:04:11	2250	00:21:10	4842	02:30:44
TPSS-D3^{atm}/def2-QZVP[a,b]	1098	00:06:20	2250	00:30:12	4842	02:50:37
HF/def2-QZVP	1098	04:09:36	2250	36:41:06	4842	401:50:54

[a]The numerical integration grid *grid4* has been used.

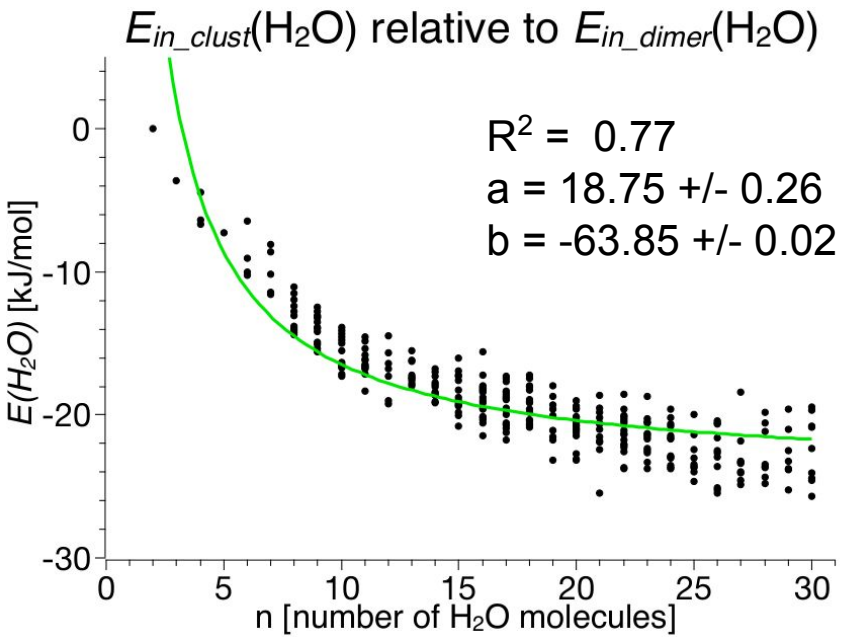
[b]The RI-J approximation has been used.

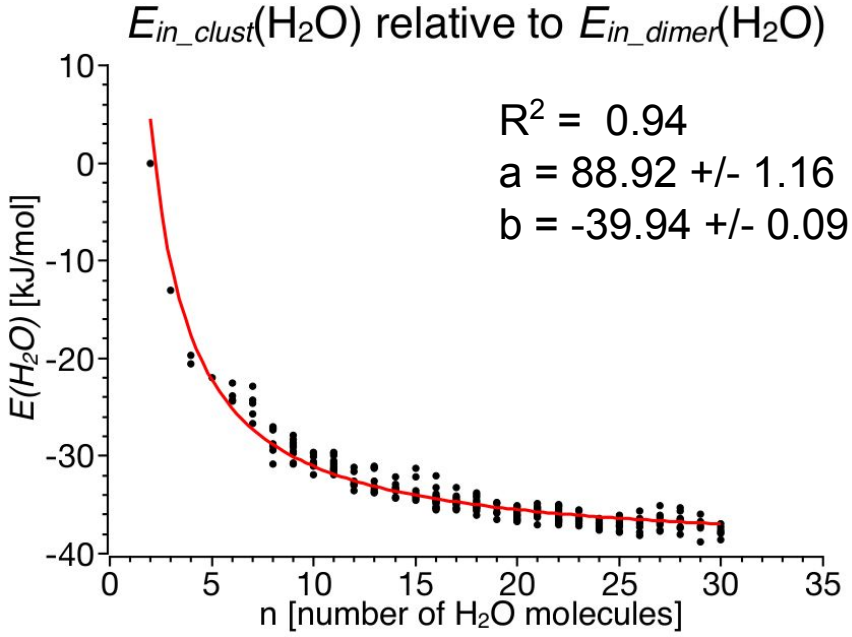
Brandenburg, Jan Gerit, et al. "B97-3c: A revised low-cost variant of the B97-D density functional method." *The Journal of chemical physics* 148.6 (2018): 064104.

Alternative approaches



Energy of H_2O molecule in the bulk

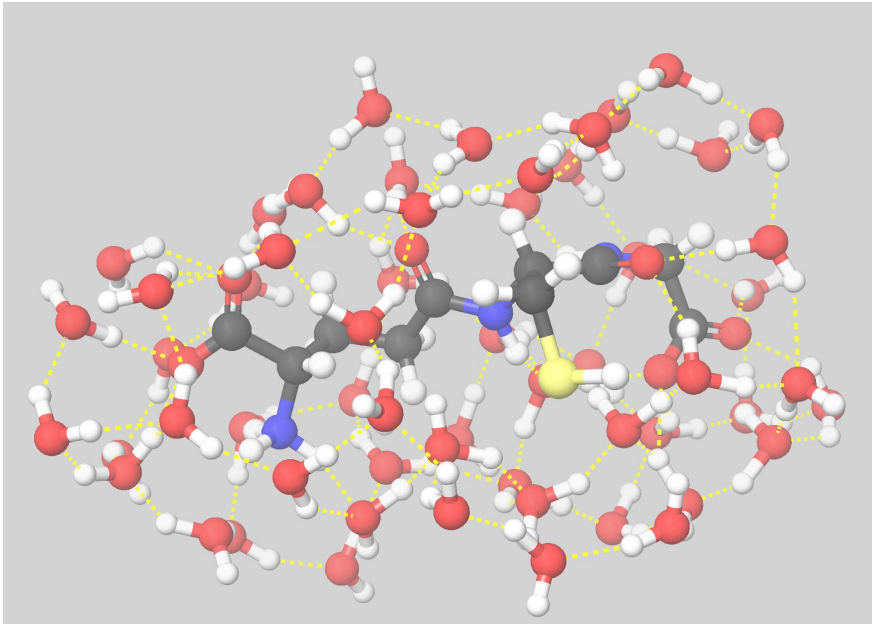


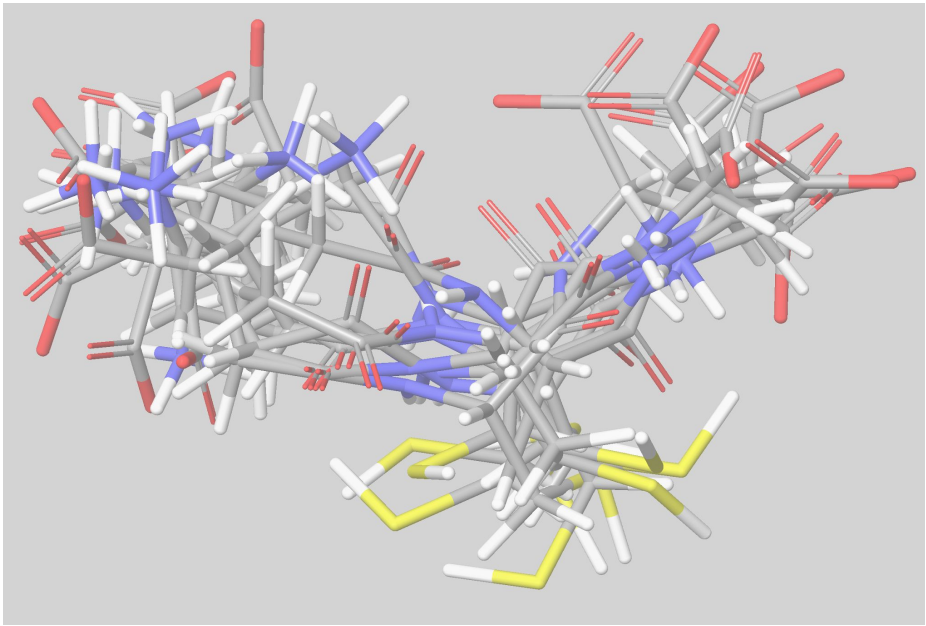


PM6-D3H4 $E/n = a/n + b$ B97-3c

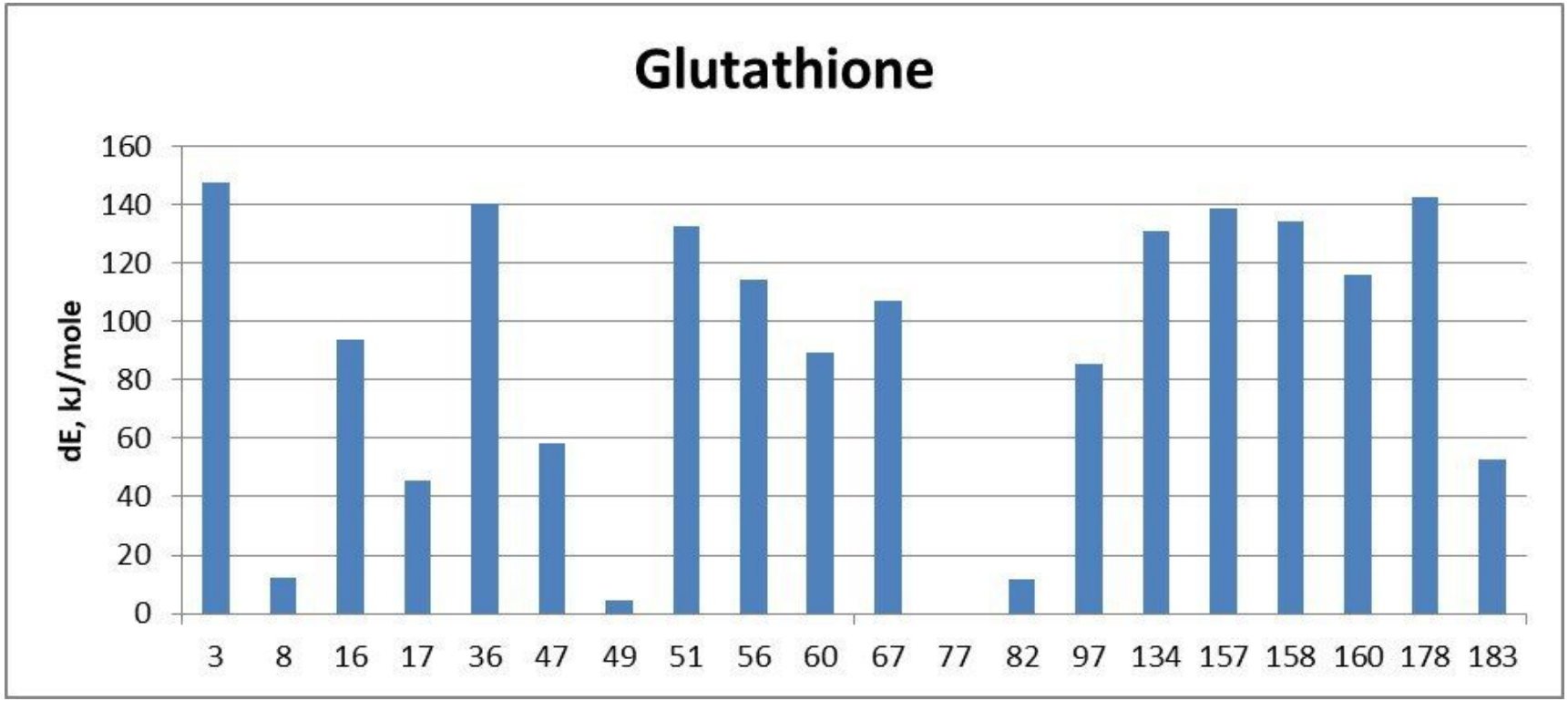
Malloum, Alhadji, et al. "Structures, relative stability and binding energies of neutral water clusters,(H_2O) 2–30." New Journal of Chemistry 43.33 (2019): 13020-13037.

Glutathione rotamers

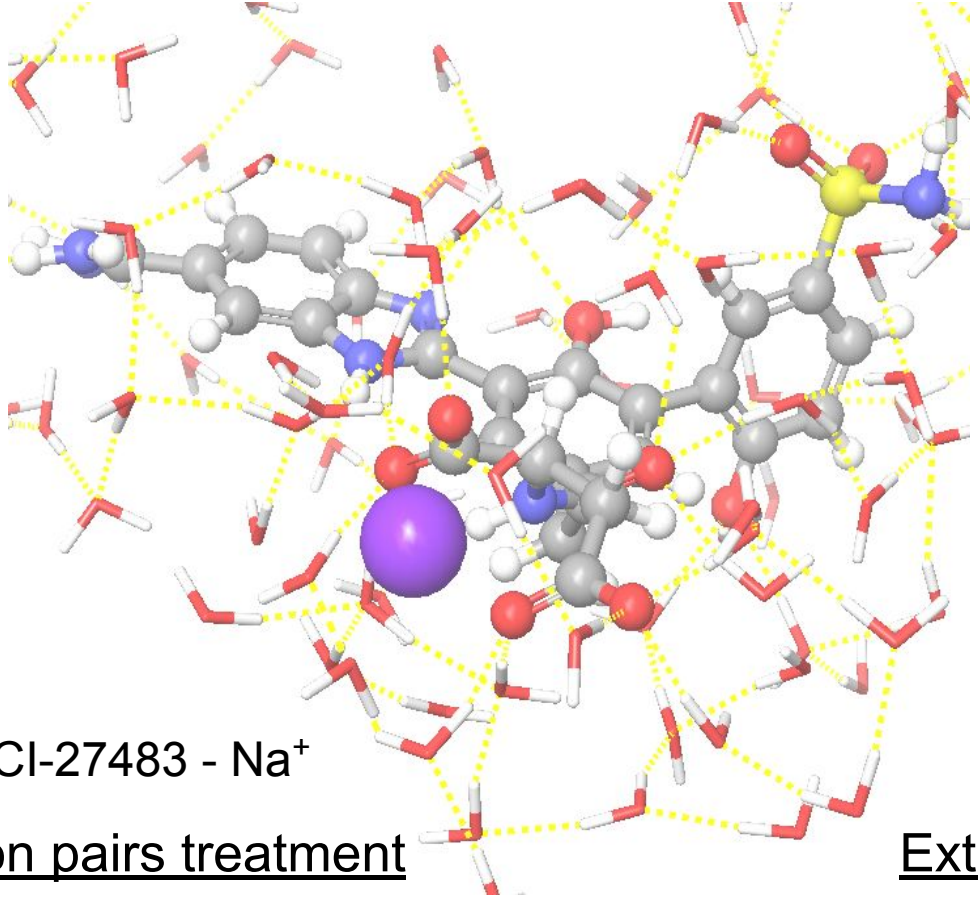




OPLS3e/PM6-D3H4 energy evaluation

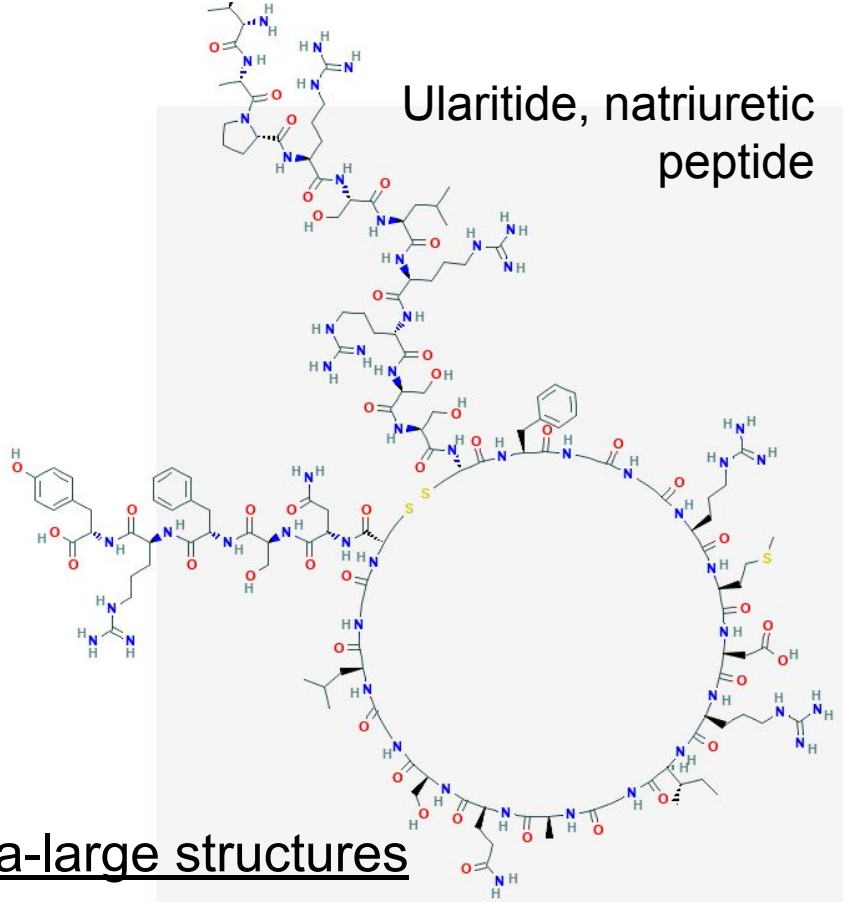


Further challenges



PCI-27483 - Na^+

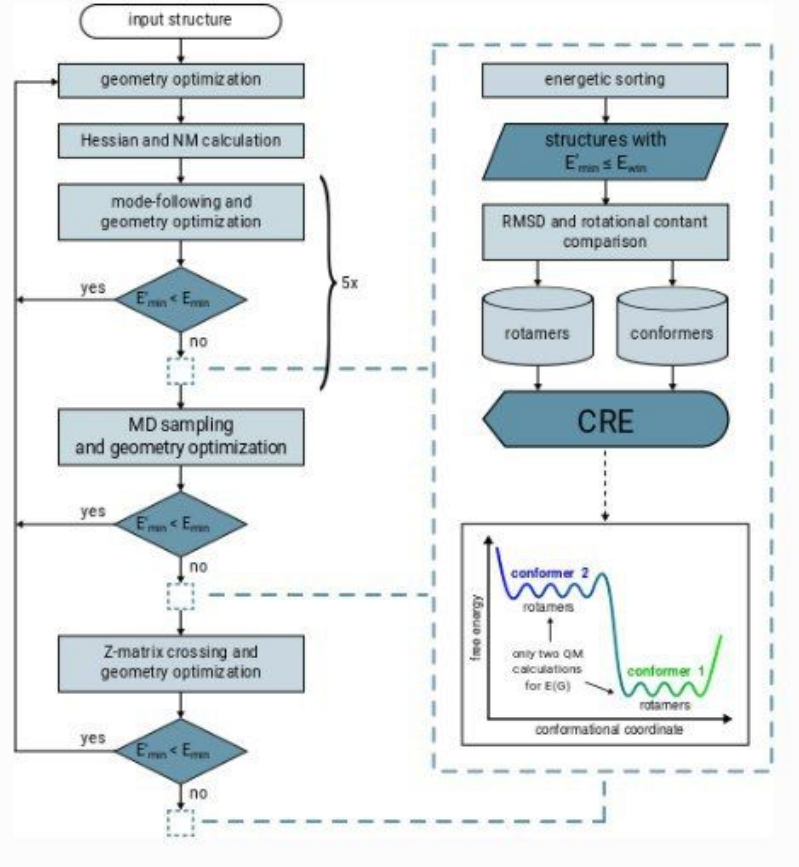
Ion pairs treatment

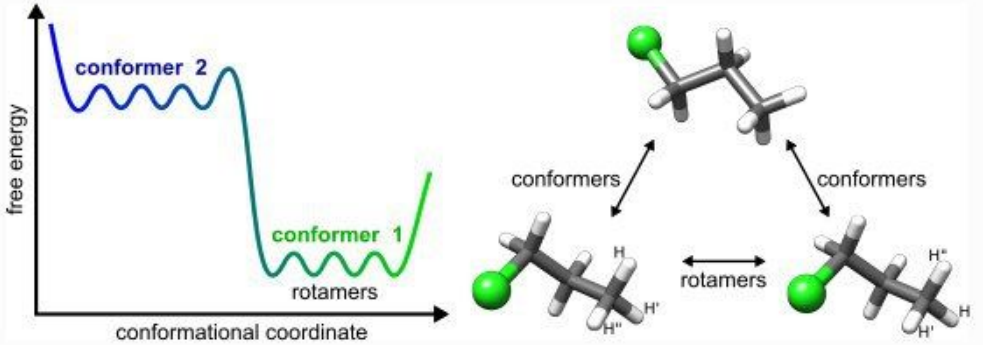


Ularitide, natriuretic peptide

Extra-large structures

Analogs: CREST





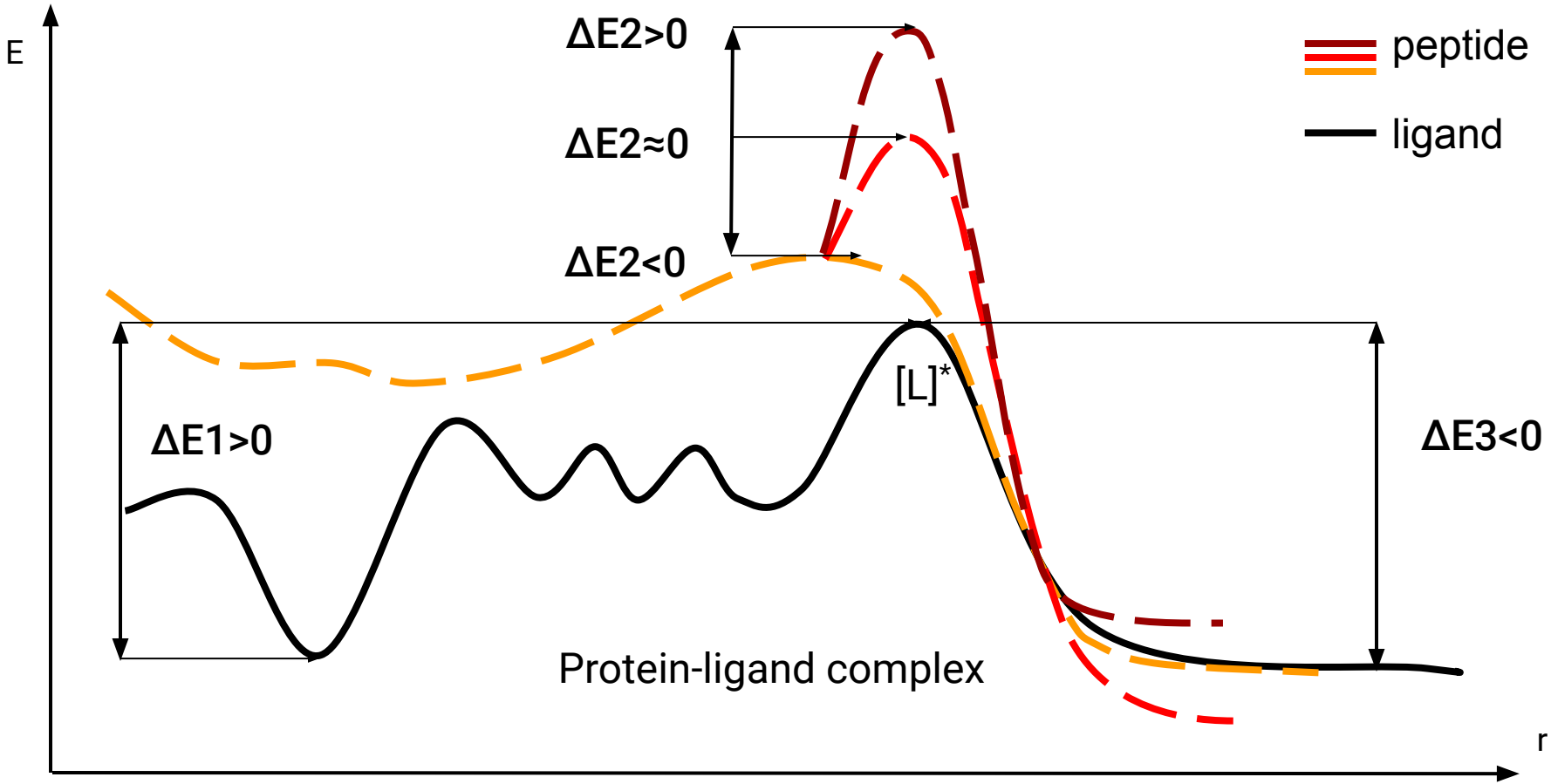
$$S_{CR} = R \sum_{i=1}^{CRE} p_i \log p_i \ ,$$

where R is the molar gas constant and the sum runs over all populations p_i of all species with energy ΔE_i at temperature T,

CREST documentation:
https://xtb-docs.readthedocs.io/en/latest/crest.html

Stefan Grimme research group

Binding site conformation analysis



Thank you for attention!

GitHub: https://github.com/mag-id/Backstreet-Science-Project

Contacts:

- Yaroslav Solovev: yaroslavsolovev78@gmail.com,
- Mikhail Polynski: polynskimikhail@gmail.com,
- Yuri Porozov: yuri.porozov@gmail.com