Department of Chemistry,
Lomonosov Moscow State University,
Moscow, Russia

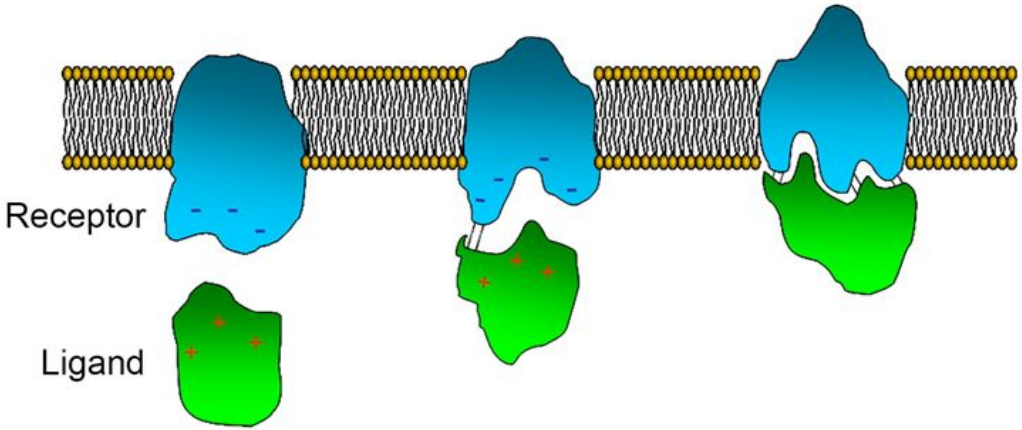
DENR+POL:
Theoretically consistent polarizable empirical charges for drug-like and biological molecules

A.R. Shaimardanov, V.S. Frolov, D.A. Shulga, V.A. Palyulin

XXXI Symposium on Bioinformatics and Computer-Aided Drug Discovery (BCADD-2025)

Empirical atomic charge models

- Atomic charges (q) – a popular approach to evaluate electrostatic interactions.
- No unambiguous definition => many charge calculation methods.
- For much medicinal chemistry goals - empirical atomic charge models.
- Advantages: high speed, moderate accuracy, interpretability.

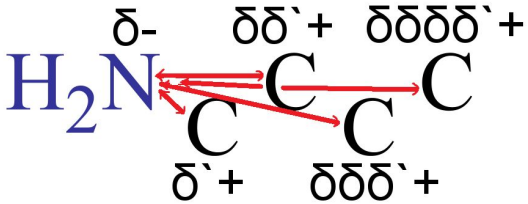


- MM force fields
- Scoring functions
- QSAR/QSPR
- QM/MM

Electronic effects: an Essence of empirical charge models

- Electronic effects concept is a physical basis of empirical models.
- **An electronic effect** - hypothetic shift of electronic density due to mutual influence of some molecular fragments.
- An generalization of organic chemistry experience.
- Examples: inductive effect, polarization.
- The more correct the physics = the more accurate the method.
- Consistent description - the cornerstone of the effects world.

influence through σ-bonds



through electric field

Account for effects: Current situation

Charge model	Inductive effect	Polarization effect
EEM	-	+/-
AACT	?	+
SQE	?	+
SQE+Q0	?	+
ACKS2	?	+
PEOE	+	-
MMFF94 BCI	-	-
DENR	+	-
	+	+

Aim and objectives of the research

The aim:
To develop empirical charge calculation method, DENR+POL, that consistently describes inductive and polarization (Pol) effects in the drug-like molecules.

The objectives:

1. To study the polarization of intra- and intermolecular hydrogen bond (HB) in the drug-like molecules.
2. To determine the consistency of the effects description.
3. To identify the applicability of the model to drug-like molecules.
4. To reproduce QC molecular electrostatic potential (MEP) of the molecules.

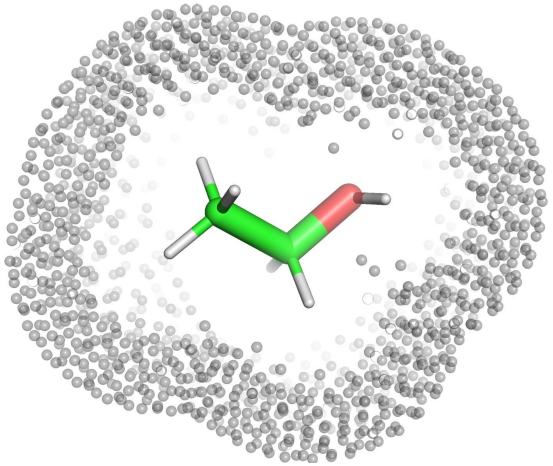
Methods & Materials

Molecular set: 7K drug-likes -
PDBbind refined set v2020 +
subset of the MMFF94 molecule training set.

QC MEP as the reference.

Quality metric of a model:

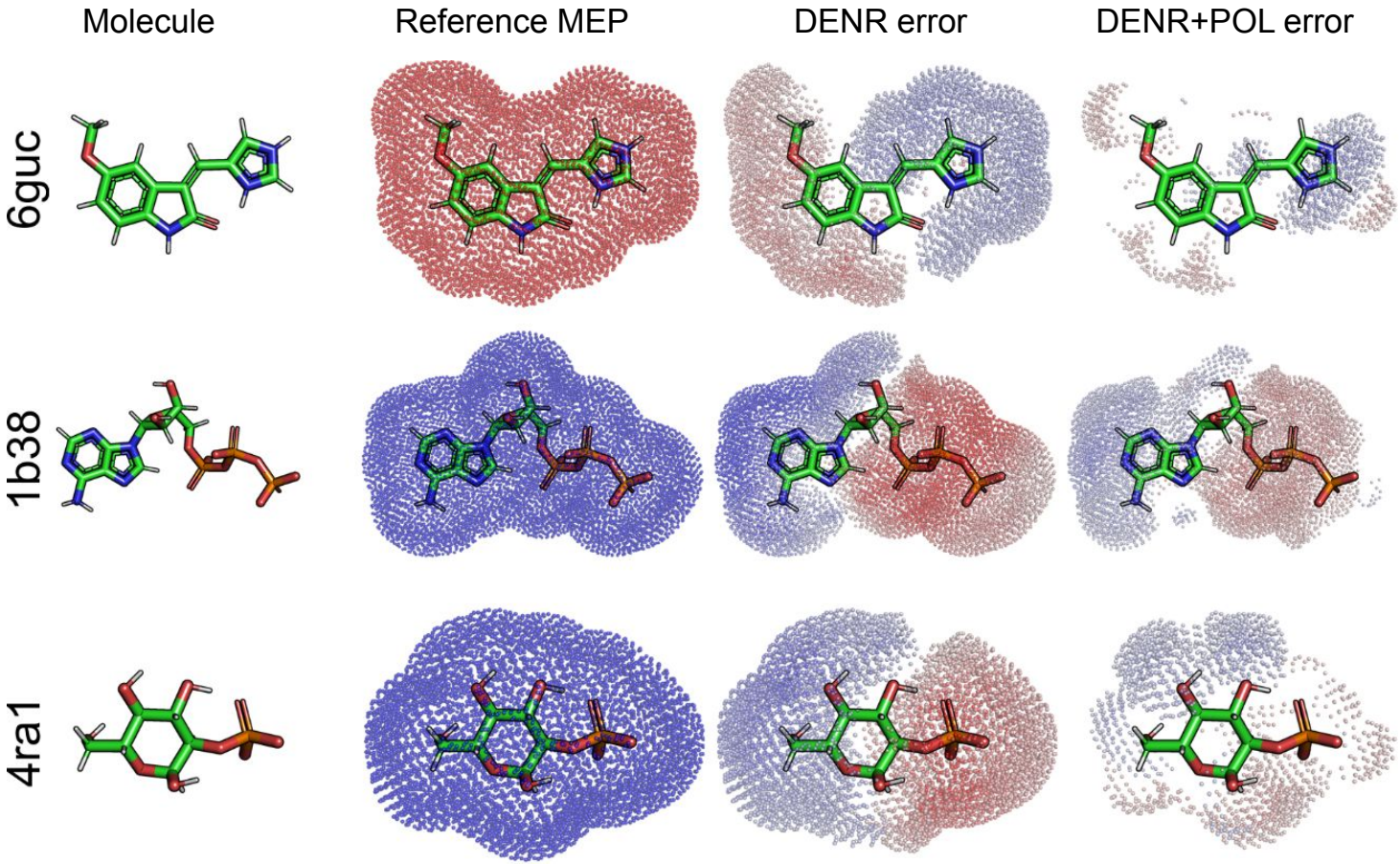
$$residual\,error(m) = MEP_m^{ref} - MEP_m^{emp} = MEP_m^{ref} - \sum_i^{N_{atoms}} \frac{q_i}{R_{i,m}}, \forall_m^{N_{MEPpoints}}$$



MEP on the grid

Selected molecules for analysis:
1) drug-like with intramolecular HB;
2) HB complex of two drug-like molecules.

Results: residual MEP reproduction



Red 🔴 positive potential (+)

Blue 🔵 negative potential (-)

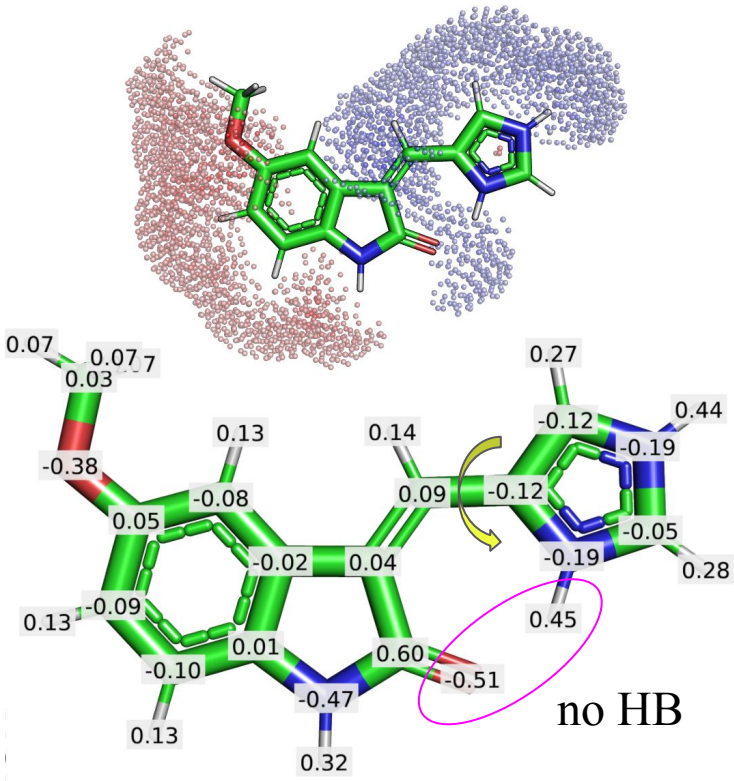
Results: Polarization of intramolecular HB

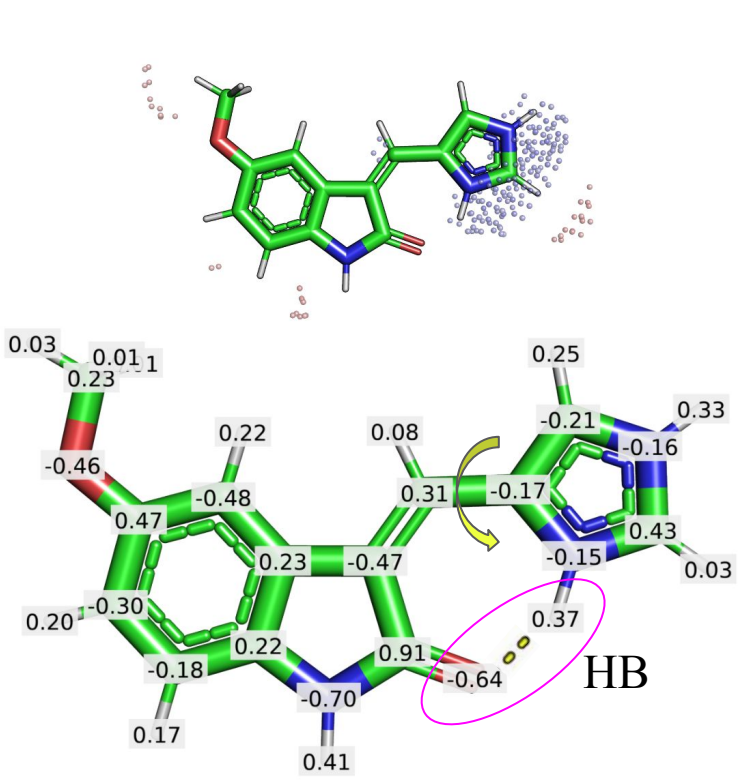
DENR

DENR+POL

0°

C=C–C–N torsion



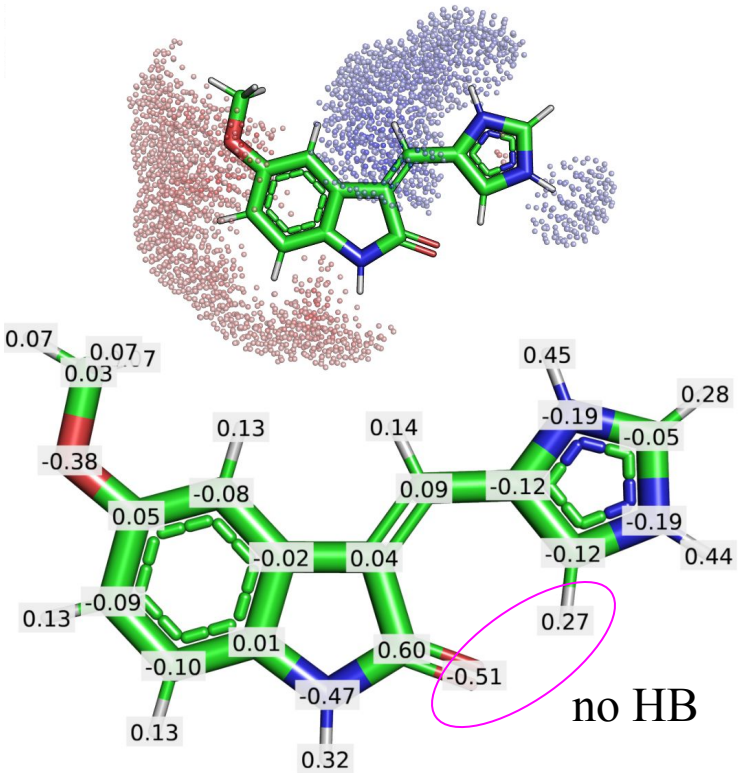


Results: Polarization of intramolecular HB

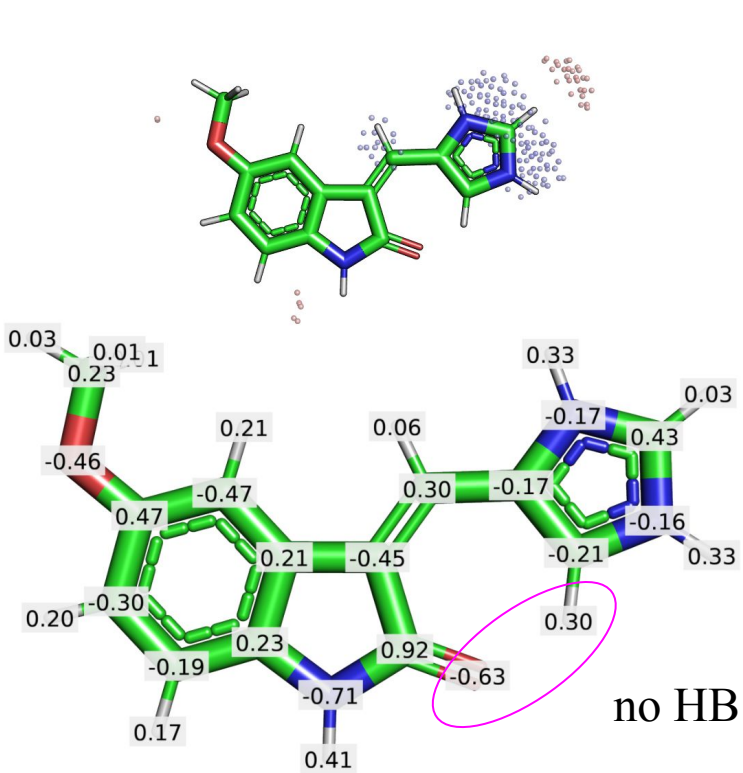
C=C–C–N torsion

180°

DENR



DENR+POL

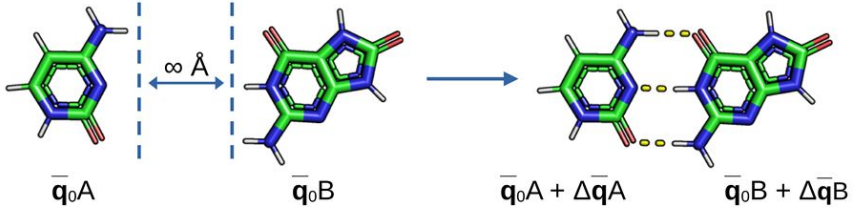


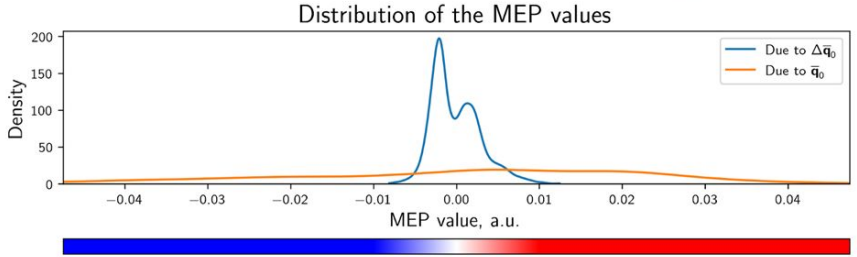
Results: Polarization of intermolecular HB

DENR+POL charges ($\bar{q}$)

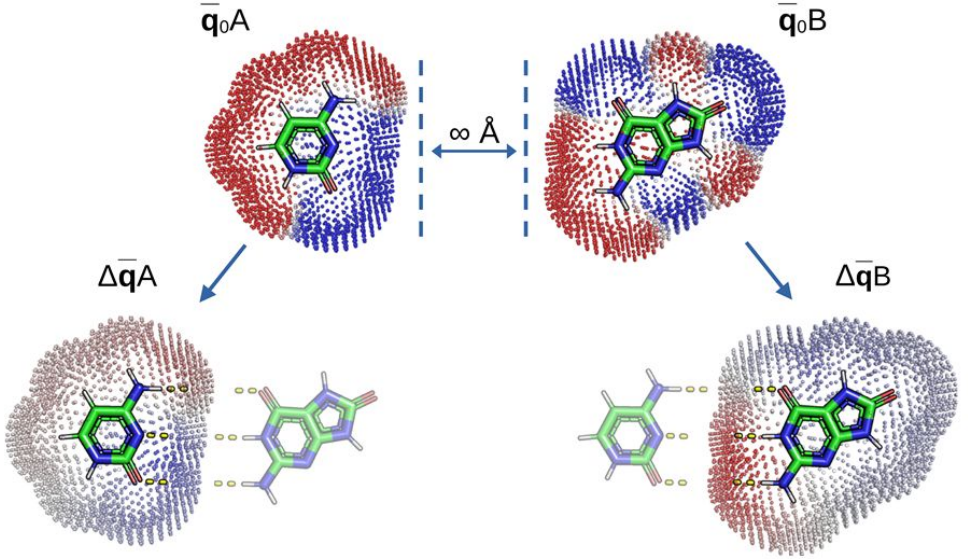
Non-interacting monomers at ∞ Å
intramolecular polarization only

Hydrogen bonded complex
intra- + intermolecular polarization





Molecular Electrostatic Potential by DENR+POL charges



Conclusions

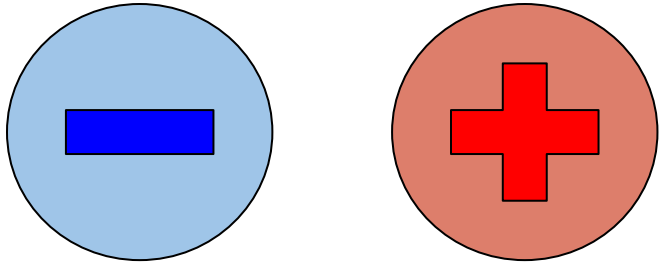
1. We developed new empirical charge model, DENR+POL, and parameterized it using QC MEP.

2. The model consistently accounts for both inductive and polarization effects in medicinal chemistry-related molecules.

3. The DENR+POL well describes polarization in both intermolecular and intramolecular HB.

4. Theoretical consistency and low computational cost of the model =>
an appealing first choice method for drug-like molecules.

Thank you

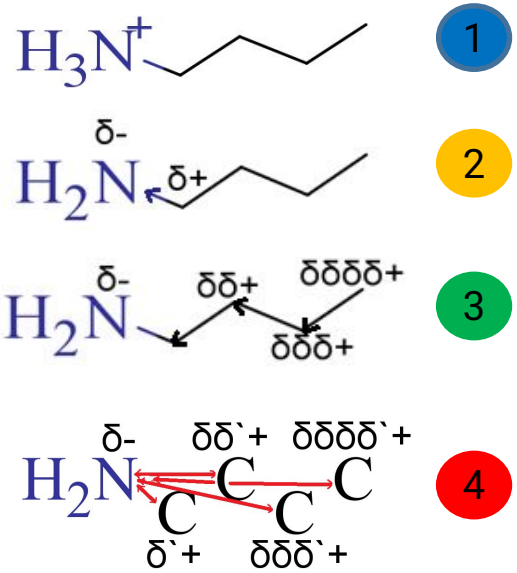


for your attention!

Requirements of good empirical charge model

The important effects:

1. **Ionized (formally charged) groups**
2. **Influence of nearest neighbors – directly bonded atoms**
3. **Inductive effect – remote atoms influence through σ-bonds**
4. **Polarization effect - through electric field**

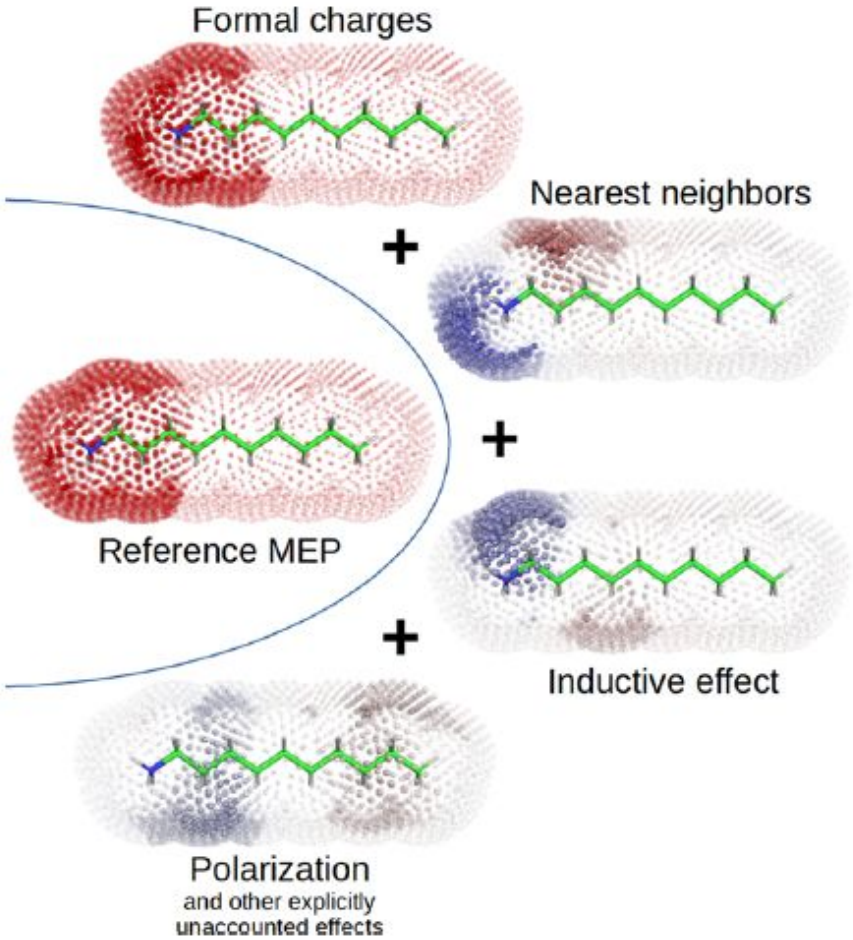


Molecular electrostatic potential (MEP) – unambiguously defined quantity.
MEP reproduction by charges → **correct** description **electrostatic** interactions.

Model X with wide applicability domain:
- ✔ Effects 1-4
- ✔ Correct account for polarization
- ✔ MEP reproduction

Charge Model DENR (it exists):
- ✔ Effects 1-3 (no Pol)
- ✔ Background for Pol treatment
- ✔ MEP reproduction

Concept of consistent description



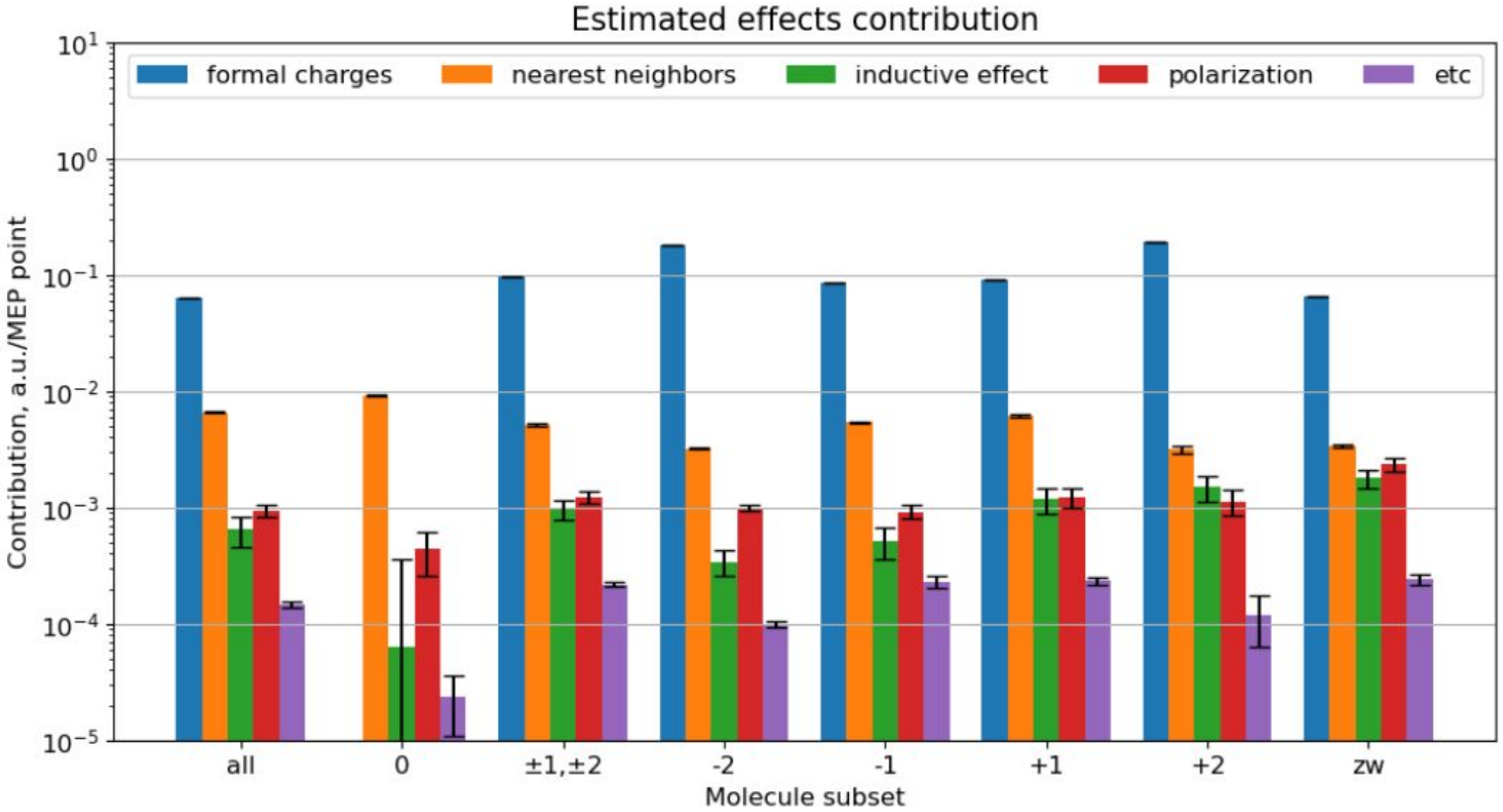
Consistent description =
hierarchy effects[1] **principle**:

the description of more fundamental effect should always precede description of more subtle one.

The example:
formal charges >
nearest neighbors >
inductive effect ≈ polarization

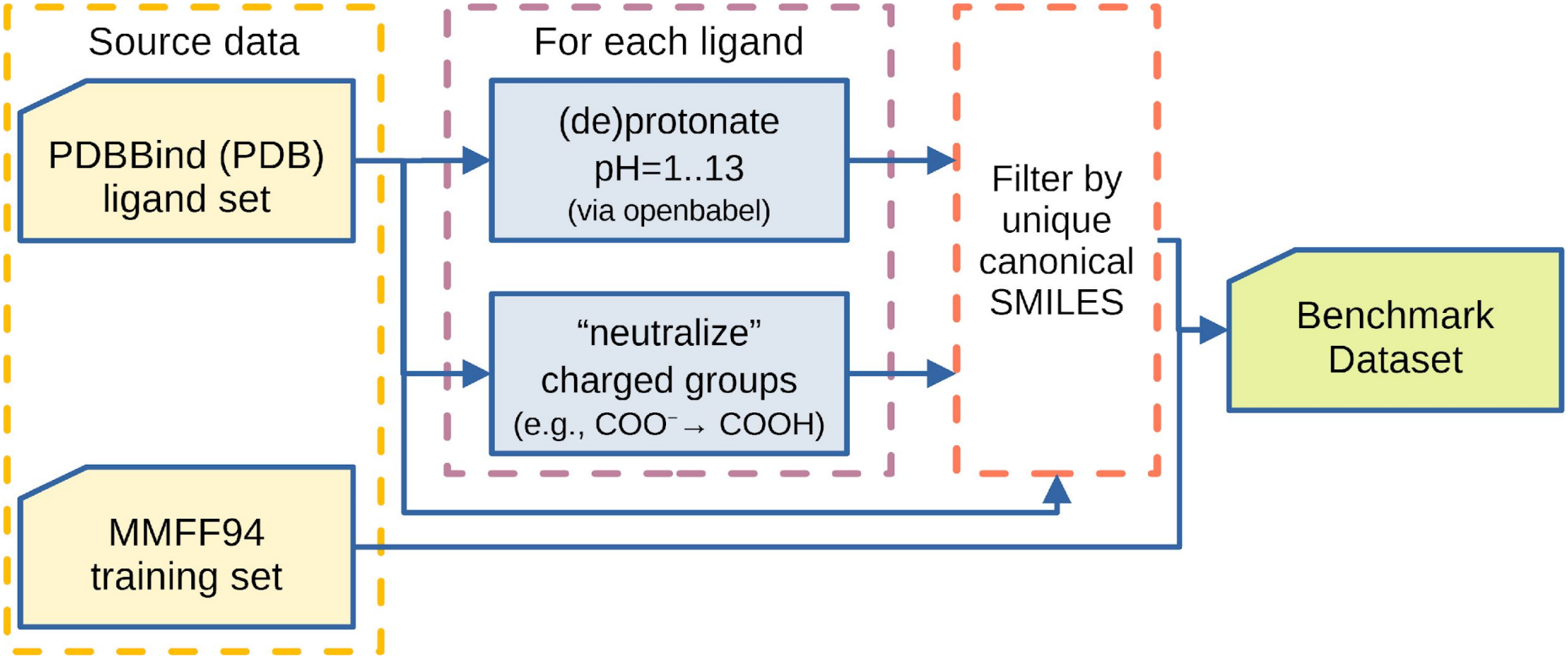
(1) Shaimardanov, A. R. et al. *J. Phys. Chem. A*. **2022**, 126, 6278.

Hierarchy of the effects

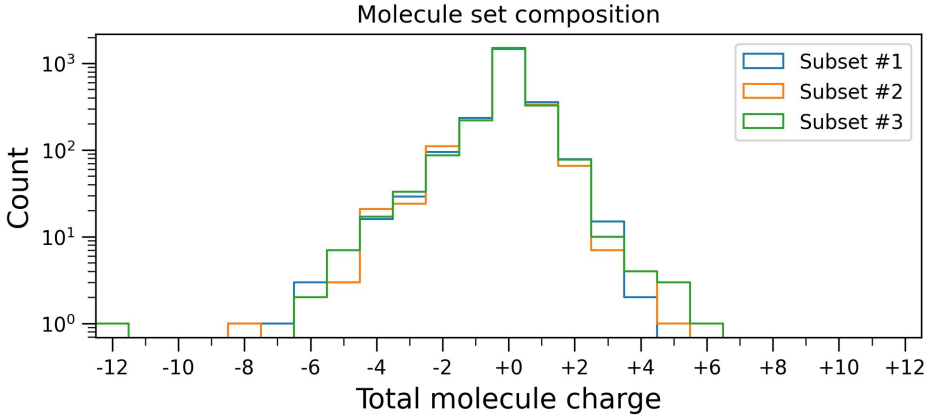


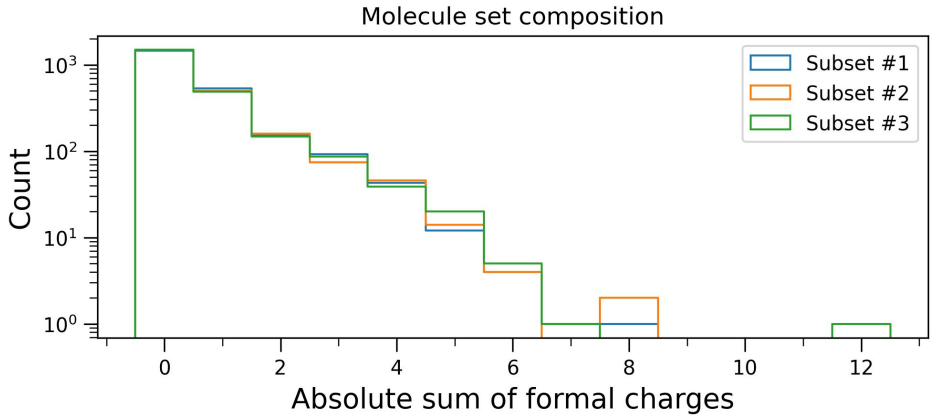
Methodology from (1) Shaimardanov, A. R. et al. *J. Phys. Chem. A*. **2022**, 126, 6278.

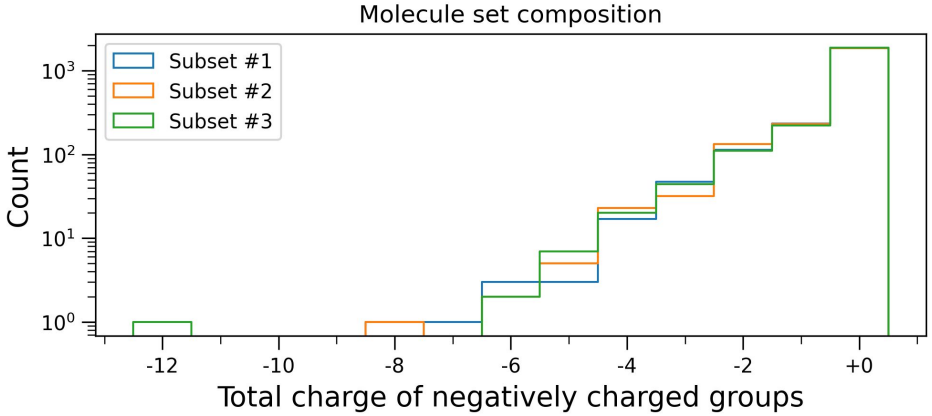
Procedure of dataset preparation

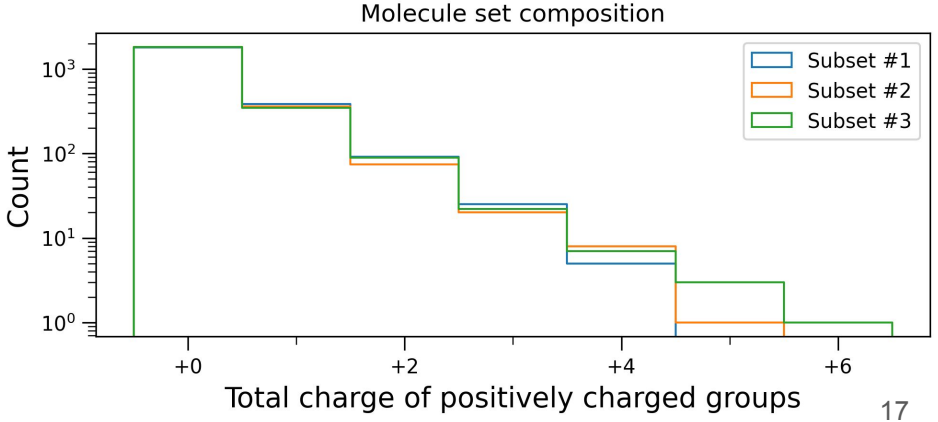


Molecule set composition





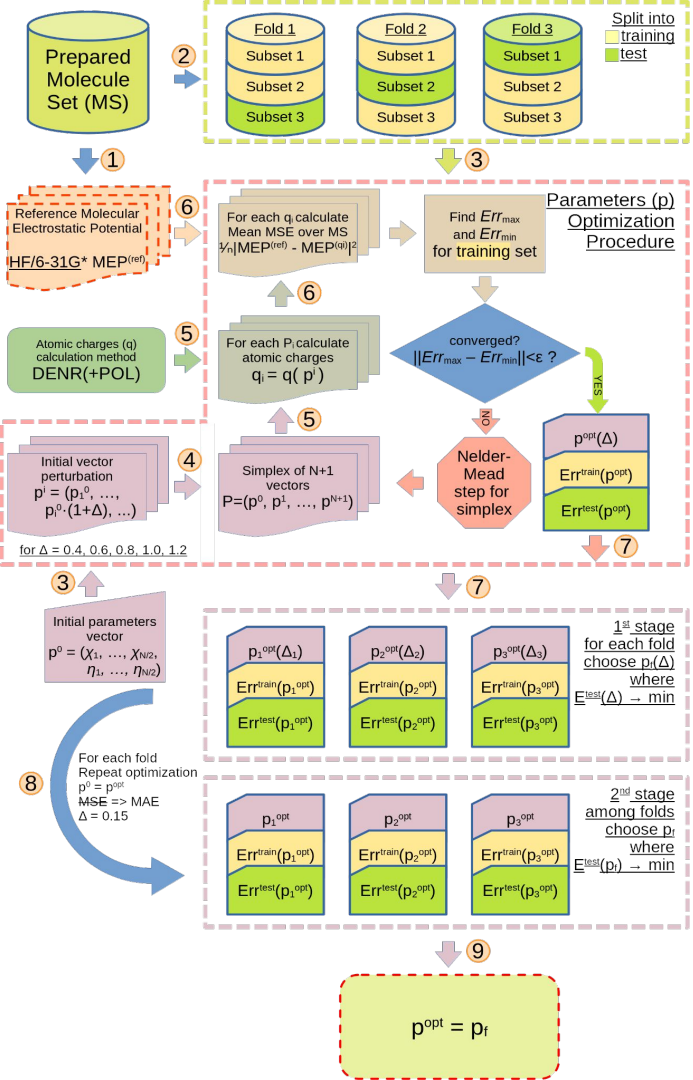


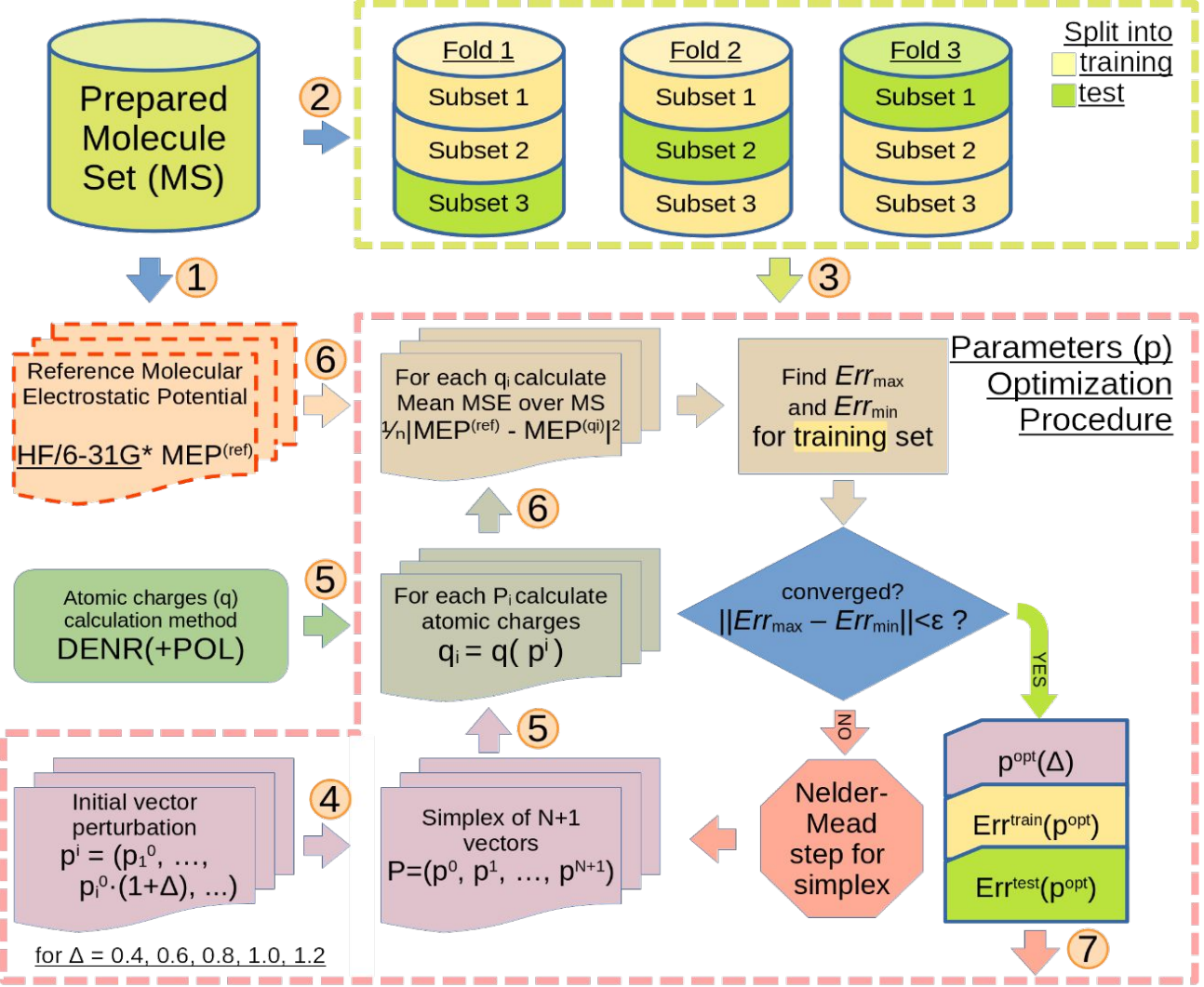


Calculation of the MEP reproduction error

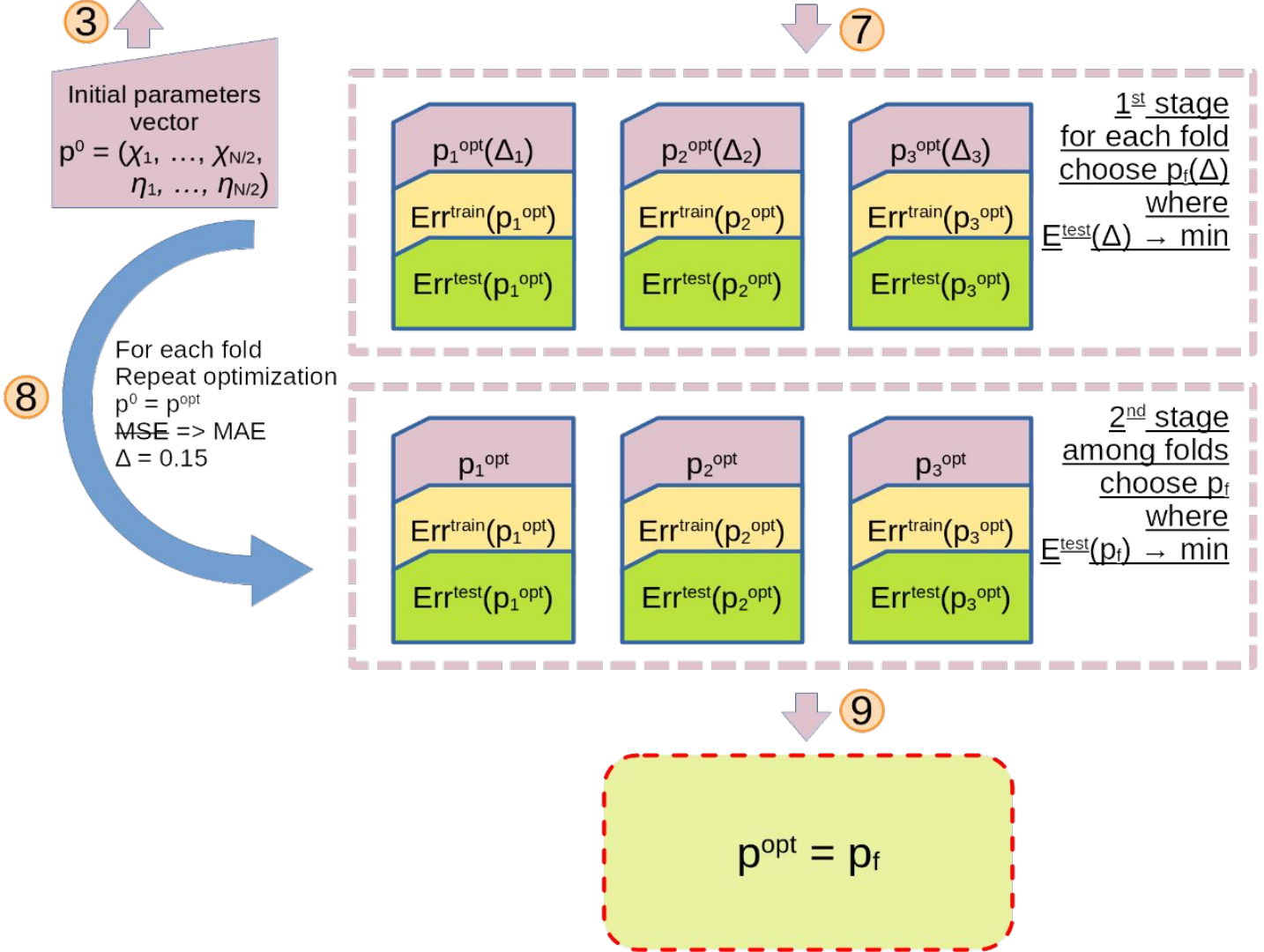
$$MMSE = \frac{1}{Nmols} \sum^{Nmols} \frac{1}{NMEPpoints} \sum_{j}^{NMEPpoints} \left[MEP_j - \sum_{i}^{Natoms} \frac{qi}{Rij} \right]^2$$	(B1)
$$MMAE = \frac{1}{Nmols} \sum^{Nmols} \frac{1}{NMEPpoints} \sum_{j}^{NMEPpoints} \left| MEP_j - \sum_{i}^{Natoms} \frac{qi}{Rij} \right|$$	(B2)
where MMSE stands for Mean (Mean Squared Error), MMAE stands for Mean (Mean Absolute Error), MEP_j is a value of a potential in *j*-th *MEP* point, q_i is charge value on *i*-th atom, R_{ij} is a distance between *i*-th atom and *j*-th MEP point.	

Parameters minimization scheme



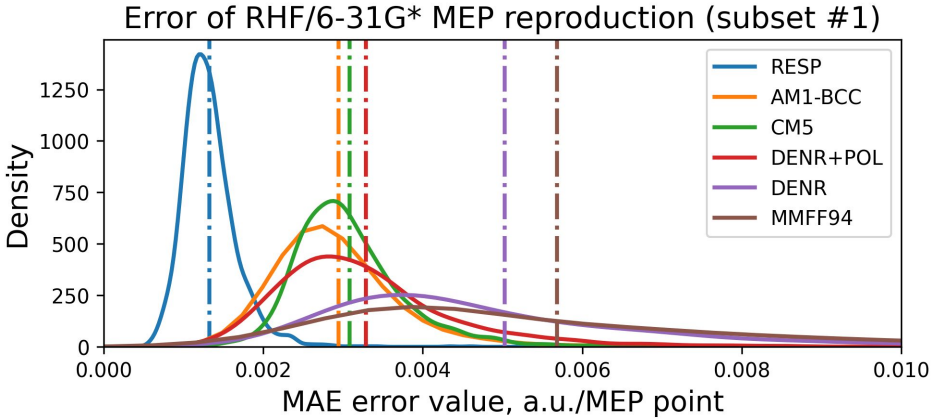


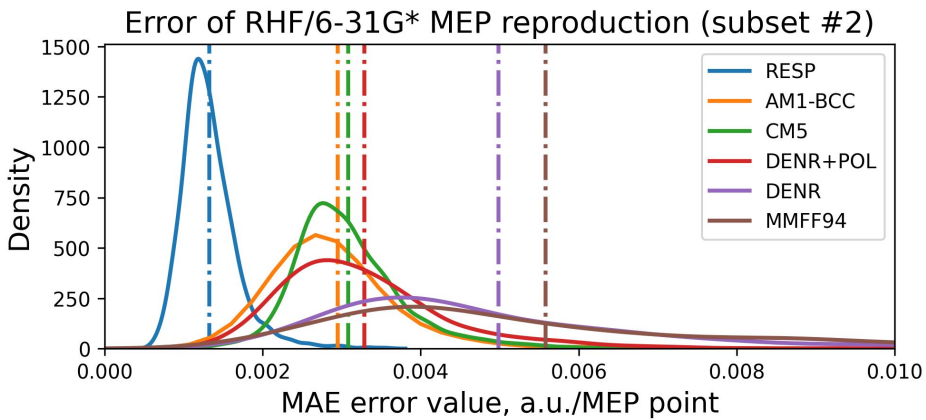
Parameters minimization scheme. Part 1

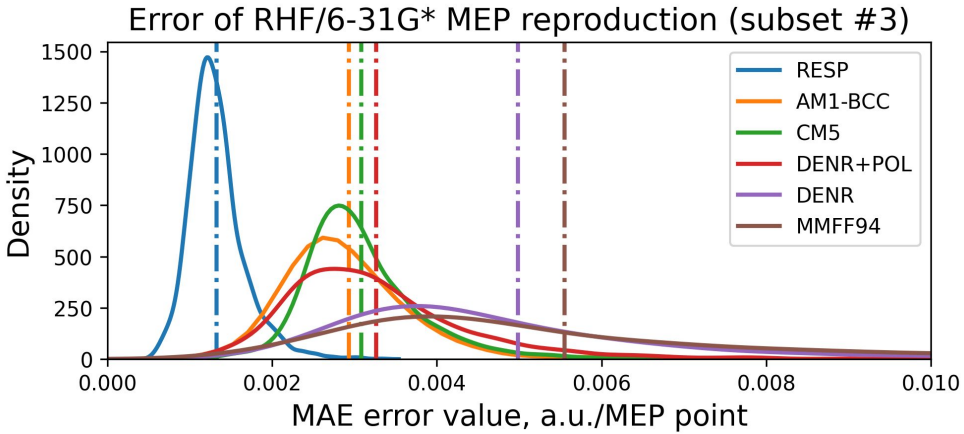


Parameters minimization scheme. Part 2

Cross-validation results







DENR formalism

Molecular energy as a function of atomic charges:

$$E_{mol}(q) = E_0 + \sum_{i}^{N_{atoms}} \left[\chi_i^0 q_i + \eta_i^0 q_i^2 \right]$$

Electronegativity: $\chi_i = \frac{\partial E_{mol}(q)}{\partial q_i}$

Hardness: $\eta_i = \frac{\partial^2 E_{mol}(q)}{\partial q_i^2} = \frac{\partial \chi_i}{\partial q_i}$

Charge flow:

$$\frac{dq_i}{dt} = -\, c \cdot \sum_{j}^{bonds} \left(\chi_j^{eff}\left(q_j^{(n)}\right) - \chi_i^{eff}\left(q_i^{(n)}\right) \right)$$

$$\chi_i^{eff}\left(q_i\right) = \chi_i^0 + \eta_i^0 \cdot q_i$$

$$for\ k = 1..n$$

$$\quad for\ i = 1..N$$

Charge on k-th iteration: $q_i^{(k)} = q_i^{(k-1)} + \frac{dq_i}{dt}\Delta t$

Derivation of formulas for description of polarization

Equation for charge flow:

$$\Delta q_i = -c\Delta t \sum_{\substack{\text{связи} \\ \text{атома } i}} \left(\chi_i^{eff} - \chi_j^{eff}\right)$$

Accounts for directly bonded neighbors and inductive effect

Dependence of **electronegativity** of *i-th* atom on charge in DENR:

$$\chi_i^{eff} = \chi_i^0 + \eta_i^0 q_i$$

Dependence of **electronegativity** of *i-th* atom on charge in DENR+POL:

$$\chi_i^{eff} = \chi_i^0 + \eta_i^0 q_i + \kappa \sum_{j \neq i} \frac{q_j}{R_{ij}}$$

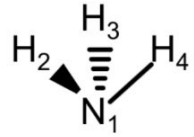
+ polarization

Theoretical background

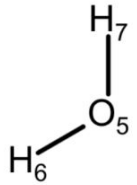
Add correct description of **polarization**

Laplace matrix formalism

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = -c \cdot \sum_{j \in \{i,j\}}^{bonds\ of\ i} \left(\chi_i^{eff}(\mathbf{q}) - \chi_i^{eff}(\mathbf{q}) \right)$$



$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = -c \cdot L\boldsymbol{\chi}^{eff}(\mathbf{q})$$



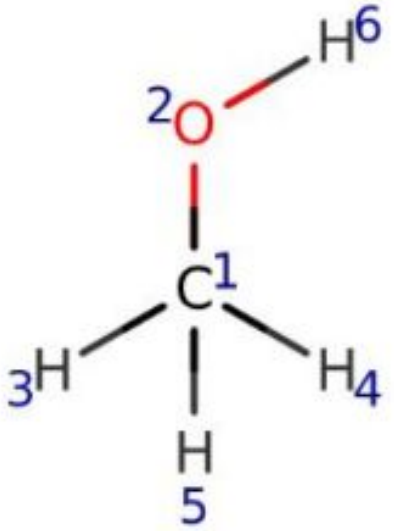
$$L_{ij} = \begin{cases} \deg(i), & \text{if } i = j \\ -1, & \text{if } i \text{ is bonded to } j \\ 0, & \text{otherwise} \end{cases}$$

Laplace matrix

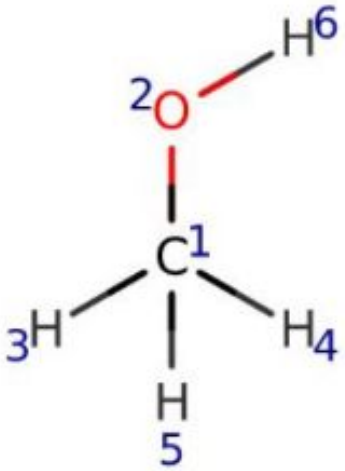
	N_1	H_2	H_3	H_4	O_5	H_6	H_7
$N1$	3	−1	−1	−1	0	0	0
$H2$	−1	1	0	0	0	0	0
$H3$	−1	0	1	0	0	0	0
$H4$	−1	0	0	1	0	0	0
$O5$	0	0	0	0	2	−1	−1
$H6$	0	0	0	0	−1	1	0
$H7$	0	0	0	0	−1	0	1

Hardness matrix for DENR

	C	O	H^3	H^4	H^5	H^6
C	η_C^0	0	0	0	0	0
O	0	η_O^0	0	0	0	0
H^3	0	0	$\eta_{H^3}^0$	0	0	0
H^4	0	0	0	$\eta_{H^3}^0$	0	0
H^5	0	0	0	0	$\eta_{H^3}^0$	0
H^6	0	0	0	0	0	$\eta_{H^6}^0$

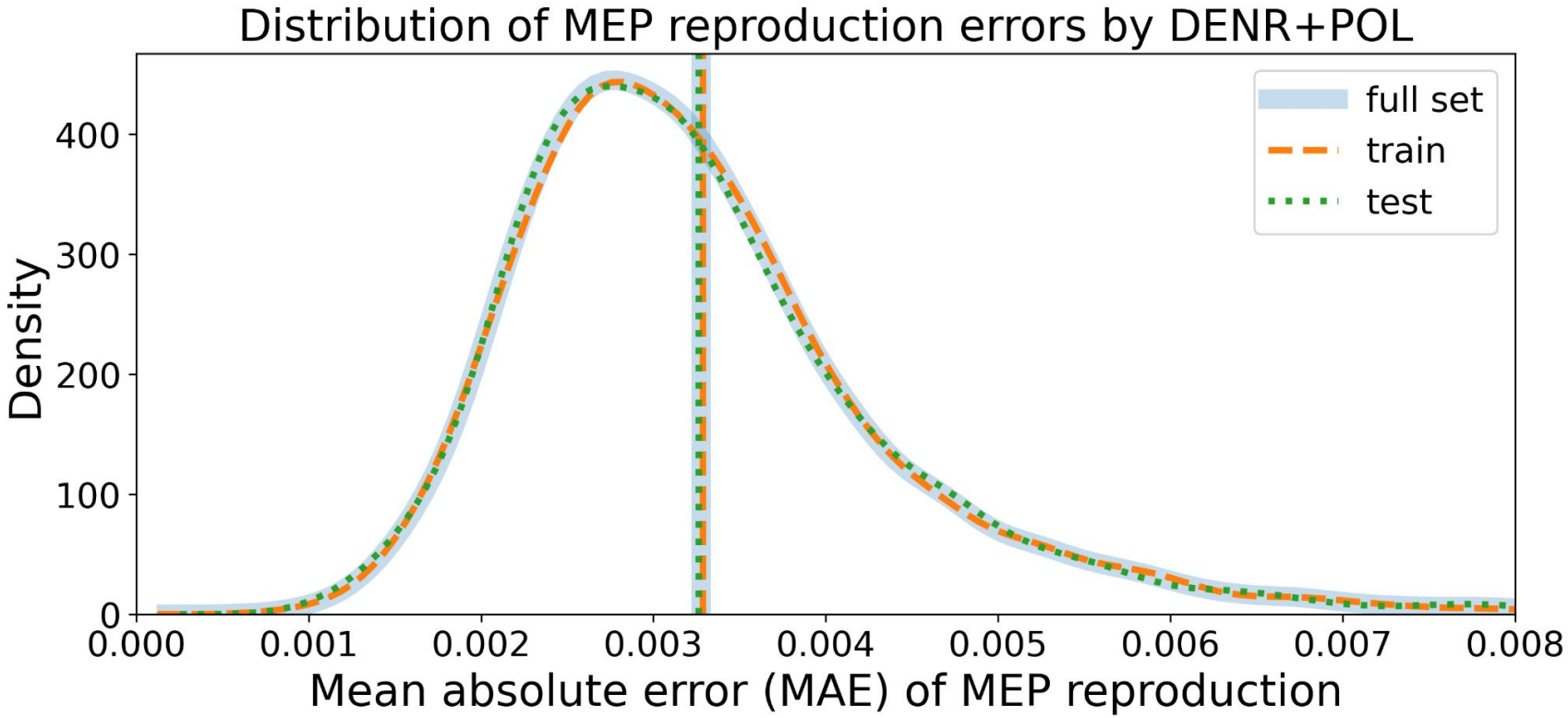


Hardness matrix for DENR+POL

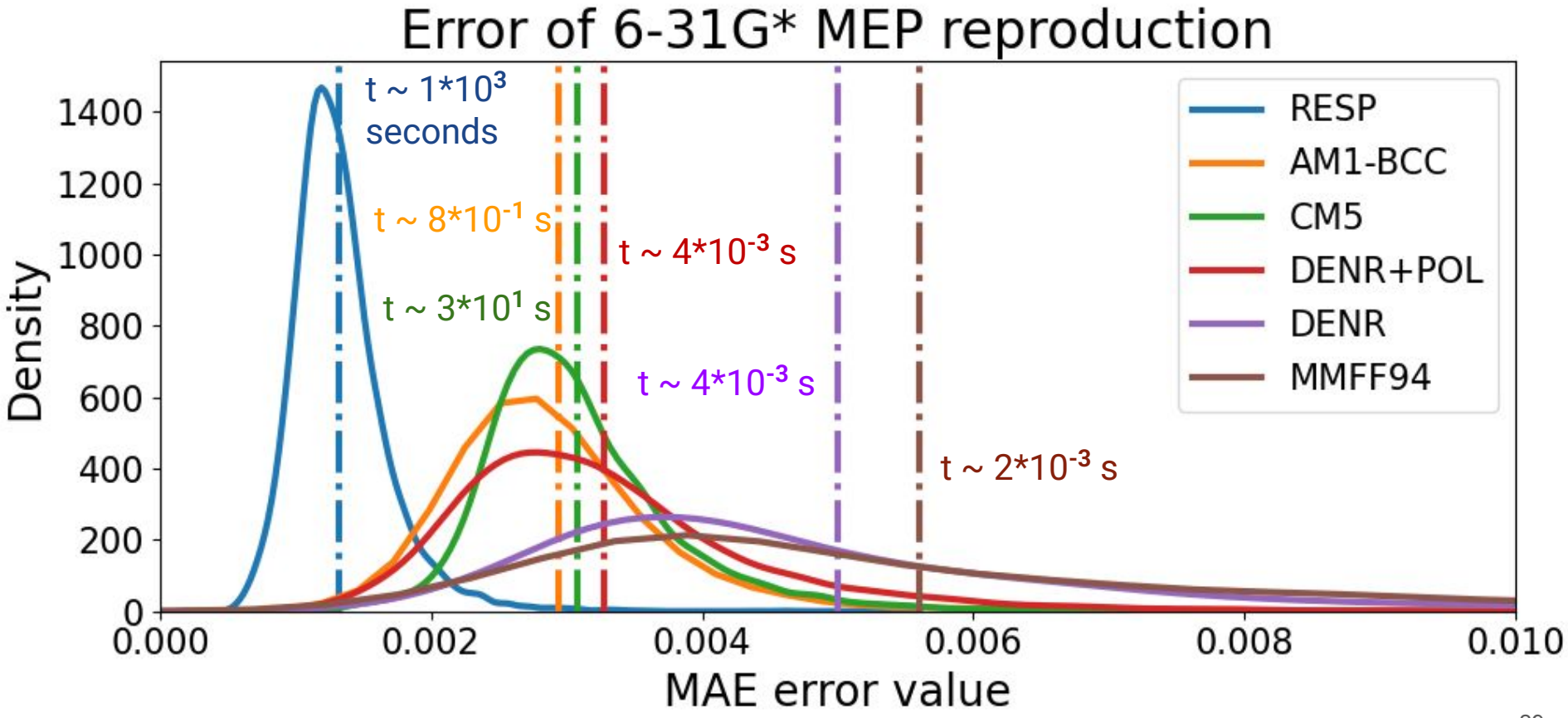


$$
\begin{array}{c}
 \\
\mathbf{C} \\
\mathbf{O} \\
\mathbf{H^3} \\
\mathbf{H^4} \\
\mathbf{H^5} \\
\mathbf{H^6}
\end{array}
\begin{array}{c}
\begin{array}{cccccc}
\mathbf{C} & \mathbf{O} & \mathbf{H^3} & \mathbf{H^4} & \mathbf{H^5} & \mathbf{H^6}
\end{array} \\
\begin{pmatrix}
\eta_C^0 & r_{CO}^{-1} & r_{CH^3}^{-1} & r_{CH^4}^{-1} & r_{CH^5}^{-1} & r_{CH^6}^{-1} \\
r_{CO}^{-1} & \eta_O^0 & r_{OH^3}^{-1} & r_{OH^4}^{-1} & r_{OH^5}^{-1} & r_{OH^6}^{-1} \\
r_{CH^3}^{-1} & r_{OH^3}^{-1} & \eta_{H^3}^0 & r_{H^3H^4}^{-1} & r_{H^3H^5}^{-1} & r_{H^3H^6}^{-1} \\
r_{CH^4}^{-1} & r_{OH^4}^{-1} & r_{H^3H^4}^{-1} & \eta_{H^3}^0 & r_{H^4H^5}^{-1} & r_{H^4H^6}^{-1} \\
r_{CH^5}^{-1} & r_{OH^5}^{-1} & r_{H^3H^5}^{-1} & r_{H^4H^5}^{-1} & \eta_{H^3}^0 & r_{H^5H^6}^{-1} \\
r_{CH^6}^{-1} & r_{OH^6}^{-1} & r_{H^3H^6}^{-1} & r_{H^4H^6}^{-1} & r_{H^5H^6}^{-1} & \eta_{H^6}^0
\end{pmatrix}
\end{array}
$$

Distribution of the MEP errors reproduction by DENR+POL



Comparison with popular models: MEP reproduction

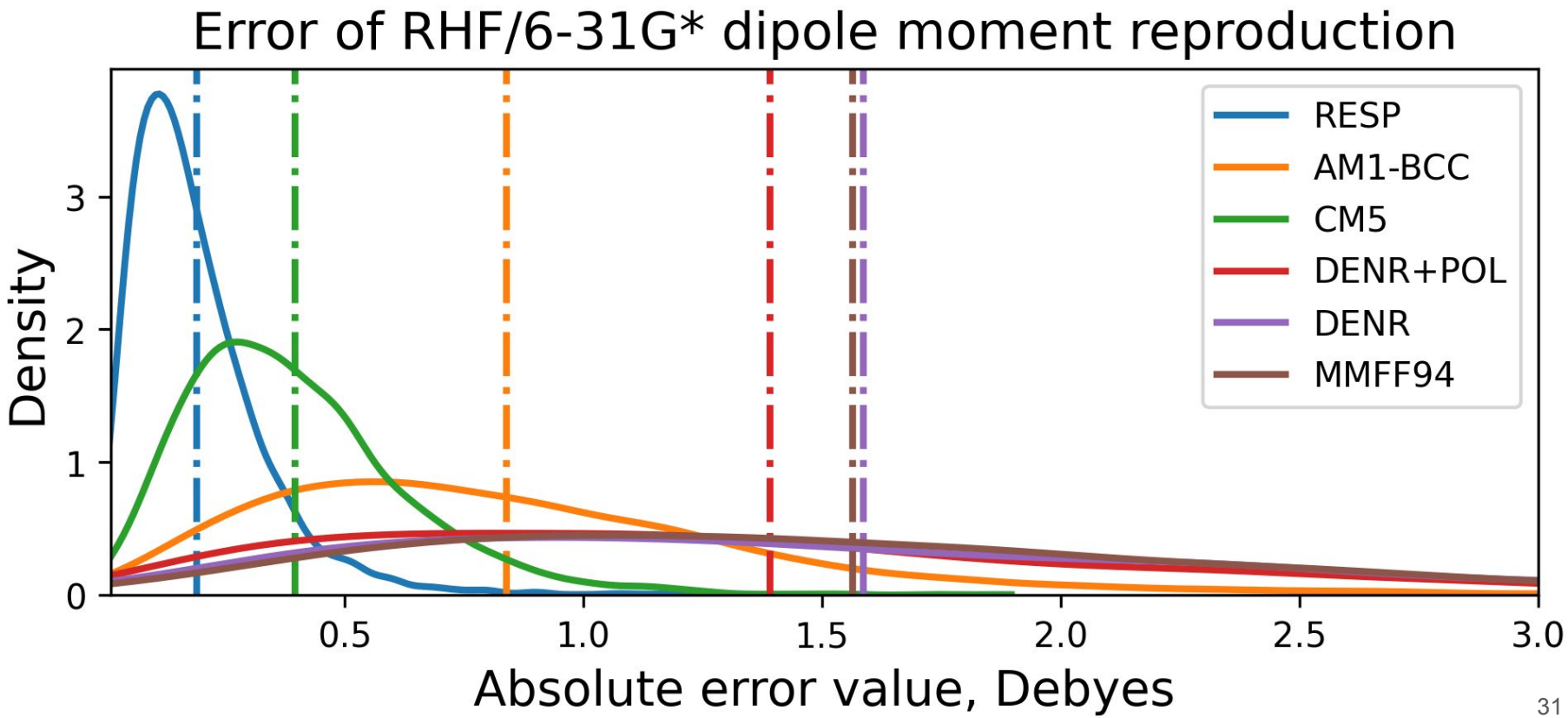


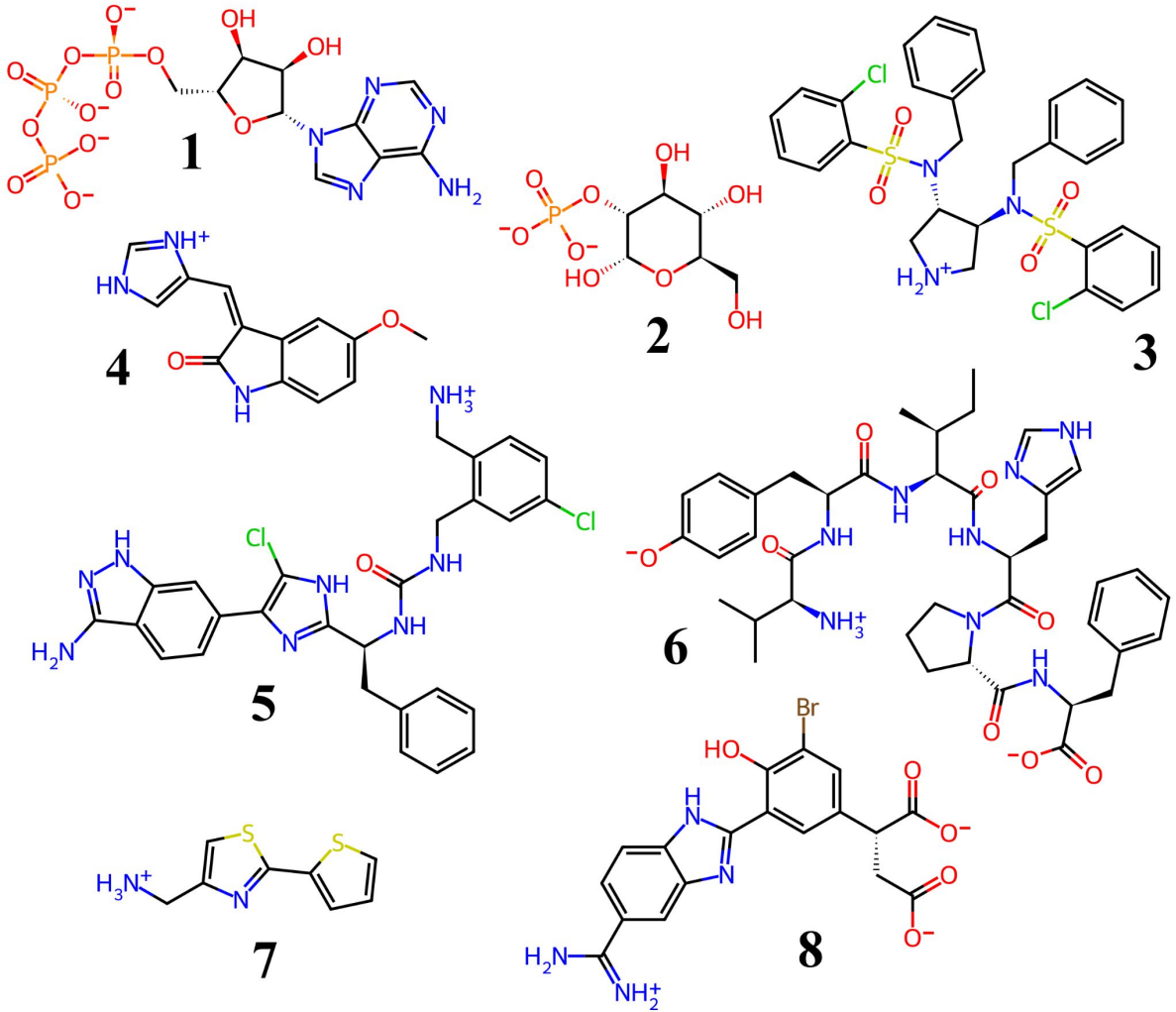
(averaged on all points of grid around molecule)

Comparison with popular models: Runtime

	Runtime, sec		
Method	**mean**	**σ**	**mean runtime relatively to the DENR+POL**
RESP	1.2×10^{3}	1.0×10^{3}	2.3×10^{5}
CM5	30	22	5.9×10^{3}
AM1-BCC	0.8	1.1	1.6×10^{2}
DENR+POL	5.1×10^{-3}	1.8×10^{-3}	1
DENR	4.5×10^{-3}	4.2×10^{-4}	0.88
MMFF94	1.8×10^{-3}	4.3×10^{-5}	0.35

Comparison with popular models: dipole reproduction

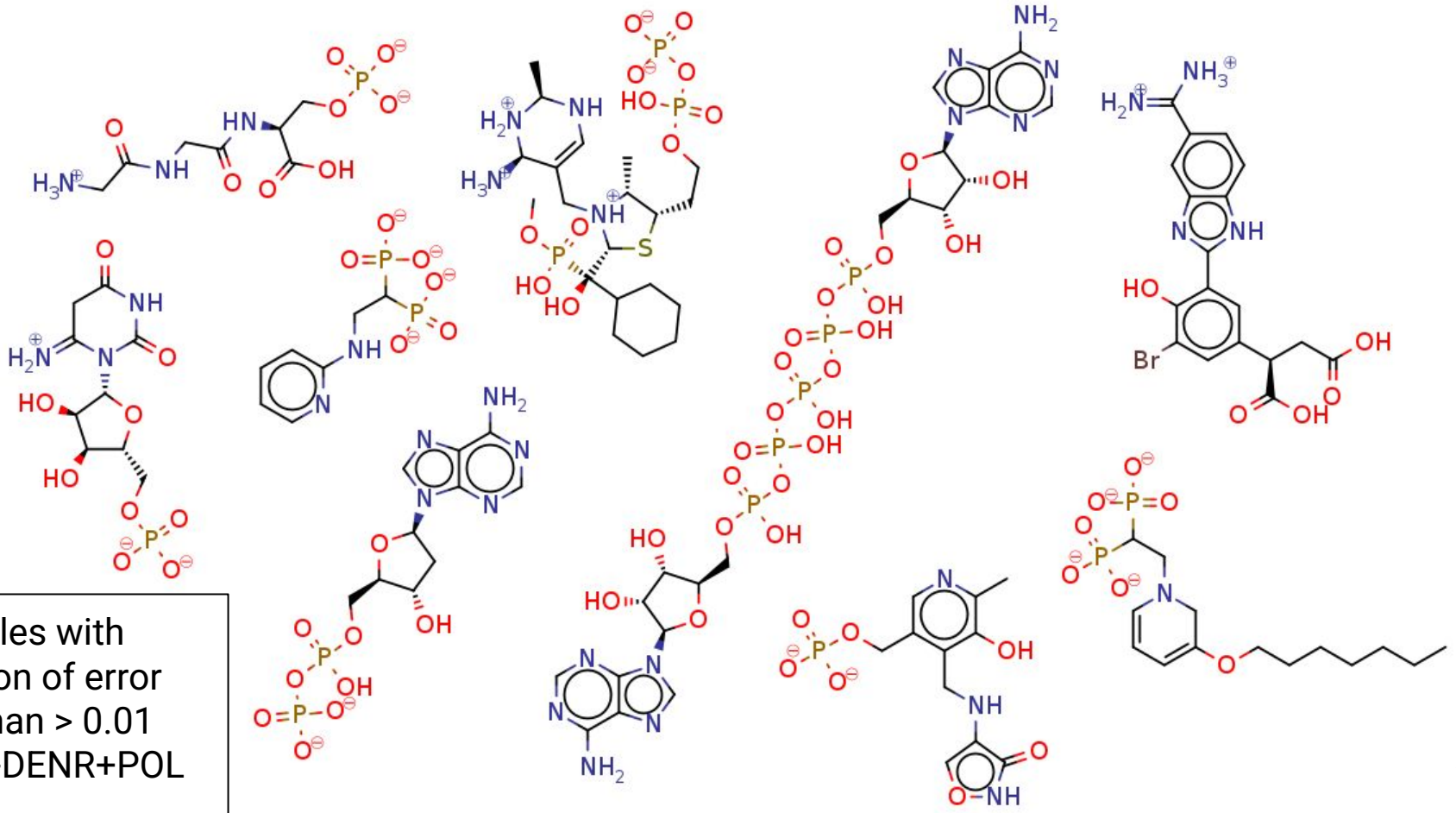




DENR+POL "wins" DENR in MEP reproduction

	MEP reproduction error, a.u. × 10^{-3}	
ID	**DENR +POL**	**DENR**
1	6.7	18.9
2	3.7	11.9
3	3.6	8.6
4	3.2	8.3
5	3.2	8.0
6	6.3	15.1
7	4.9	9.9
8	8.4	18.9

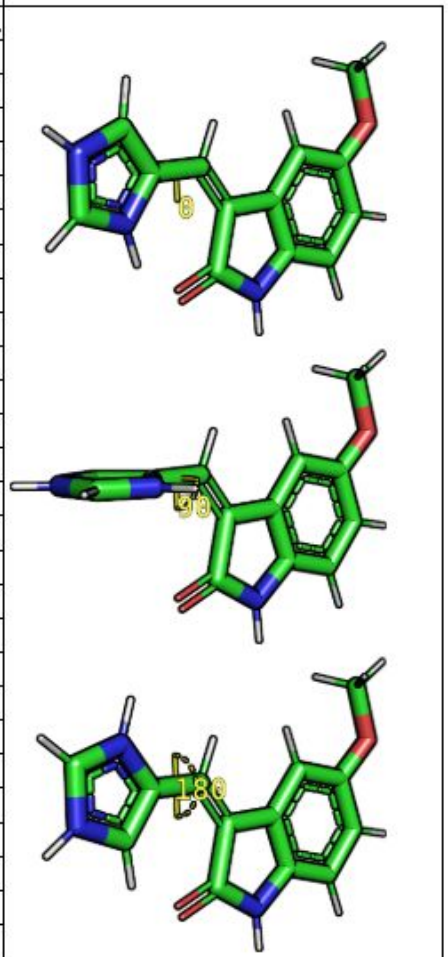
Biomolecules: perspectives



Molecules with reduction of error more than > 0.01 DENR->DENR+POL

Table G1. Dependence of the MEP reproduction error on the molecule conformation.

	MAE, a.u. × 10^{-3}	
deg	**DENR**	**DENR+POL**
0	9.0	3.1
15	8.8	3.1
30	8.5	2.9
45	8.3	2.8
60	8.3	2.9
75	8.3	2.9
90	8.3	2.9
105	8.3	2.9
120	8.2	2.9
135	8.1	2.9
150	8.1	2.9
165	8.3	2.8
180	8.4	2.8
195	8.3	2.8
210	8.1	2.9
225	8.1	2.9
240	8.2	2.9
255	8.3	2.9
270	8.3	2.9
285	8.3	2.9
300	8.3	2.8
315	8.3	2.8
330	8.5	2.9
345	8.8	3.1
360	9.0	3.1
mean	**8.4**	**2.9**
avedev	**0.2**	**0.1**
avedev/mean, %	**2.31%**	**2.17%**



Dependence of the MEP reproduction error on the molecular conformation.