



NOE for ETKDG

Kajo Krummenacher

16. December 2020, Zürich



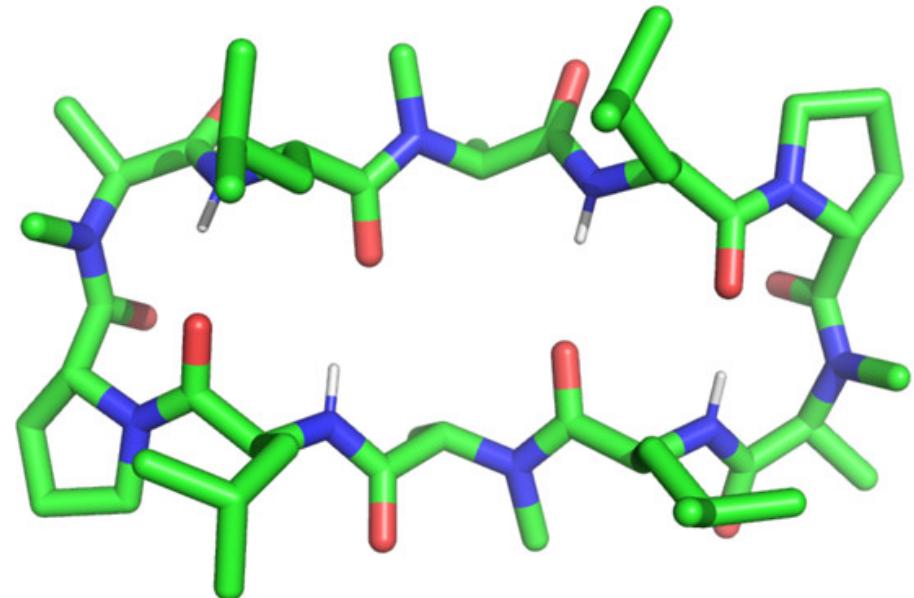
Agenda

1. Cyclic Peptides
2. ETKDG
3. NOE-ETKDG
4. Results
5. Conclusion & Outlook

Cyclic Peptides

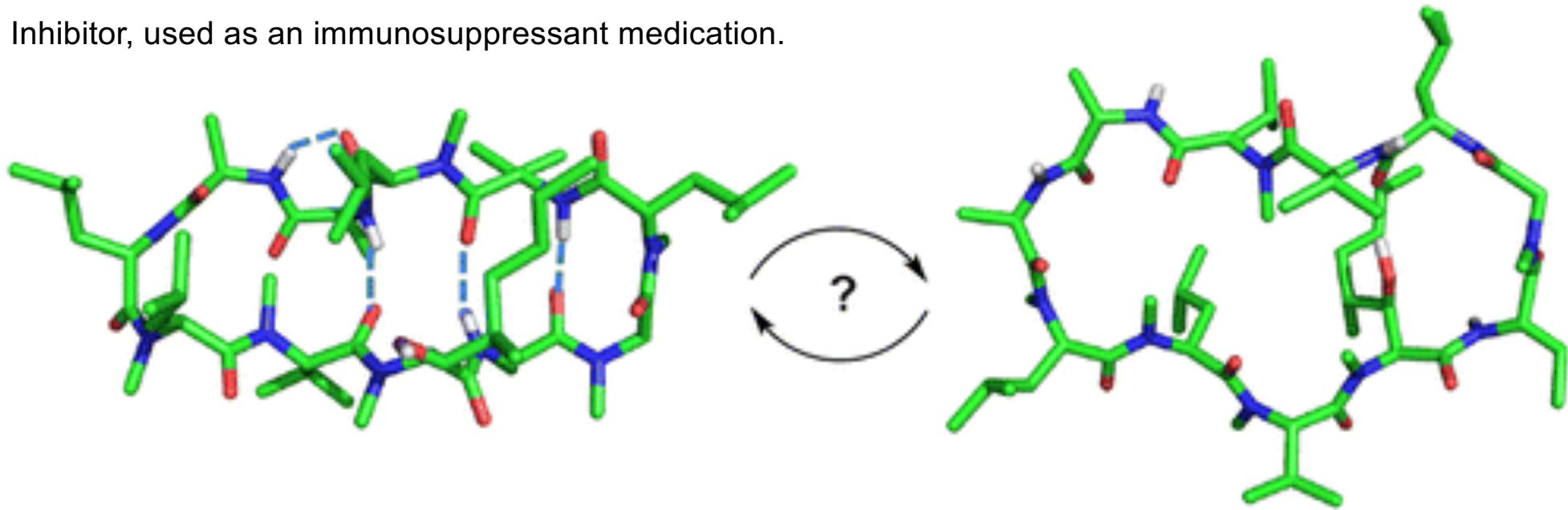
- Range of 10 – 20 residues
- Compared to classical small molecule drugs, they tend to have [1]:
 - higher bioavailability
 - higher stability in-situ
 - lower toxicity
 - more restricted binding options

→ of great interest for drug development



Example: Cyclosporin A [1]

Inhibitor, used as an immunosuppressant medication.



Closed: twisted sheet conformation with 4 H-bonds

Observed in crystal structure and CHCl_3

Allows membrane permeability

Open: conformation without H-bond

Preferred in H_2O and at the receptor

ETKDG: Experimental-Torsion Knowledge Distance Geometry [1]

- Distance geometry **DG**: Geometry theory assuming that all conformations of a molecule can be defined by geometric constraints.
- Initial constraints from:
 - ideal bond lengths
 - ideal bond angles
 - ideal torsional angles → Experimental-Torsions **ET**
- Basic knowledge **K**:
 - flatness of sp^2 -centers and aromatic rings
 - linear triple bonds

ETKDG: Working Principle

Initialize bounds matrix

		Upper bounds: arbitrarily large					
		→ atom index					
↓ atom index	→ atom index	0	100	100	100	100	100
0	2.7	0	100	100	100	100	100
2.7	1.8	2.4	0	100	100	100	100
1.8	3.7	2.7	1.4	0	100	100	100
3.7	2.4	2.7	1.4	0	100	100	100
2.4	1.5	2.7	3.7	0	100	100	100
1.5	2.7	1.9	1.8	2.7	0	100	100

Lower bounds: sum of VdW radii

Set topological bounds

Use

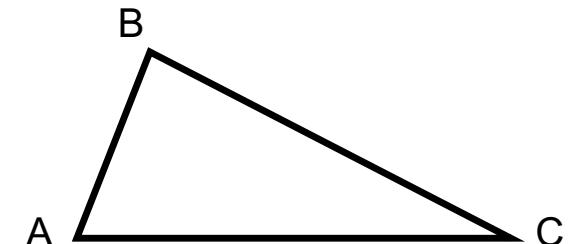
- ideal bond lengths
- bond angles
- torsional angles
- basic knowledge

to
decrease upper bounds
and
increase lower bounds

Remark: Torsions from
substructure search with
torsional SMARTS patterns

Smoothing of bounds

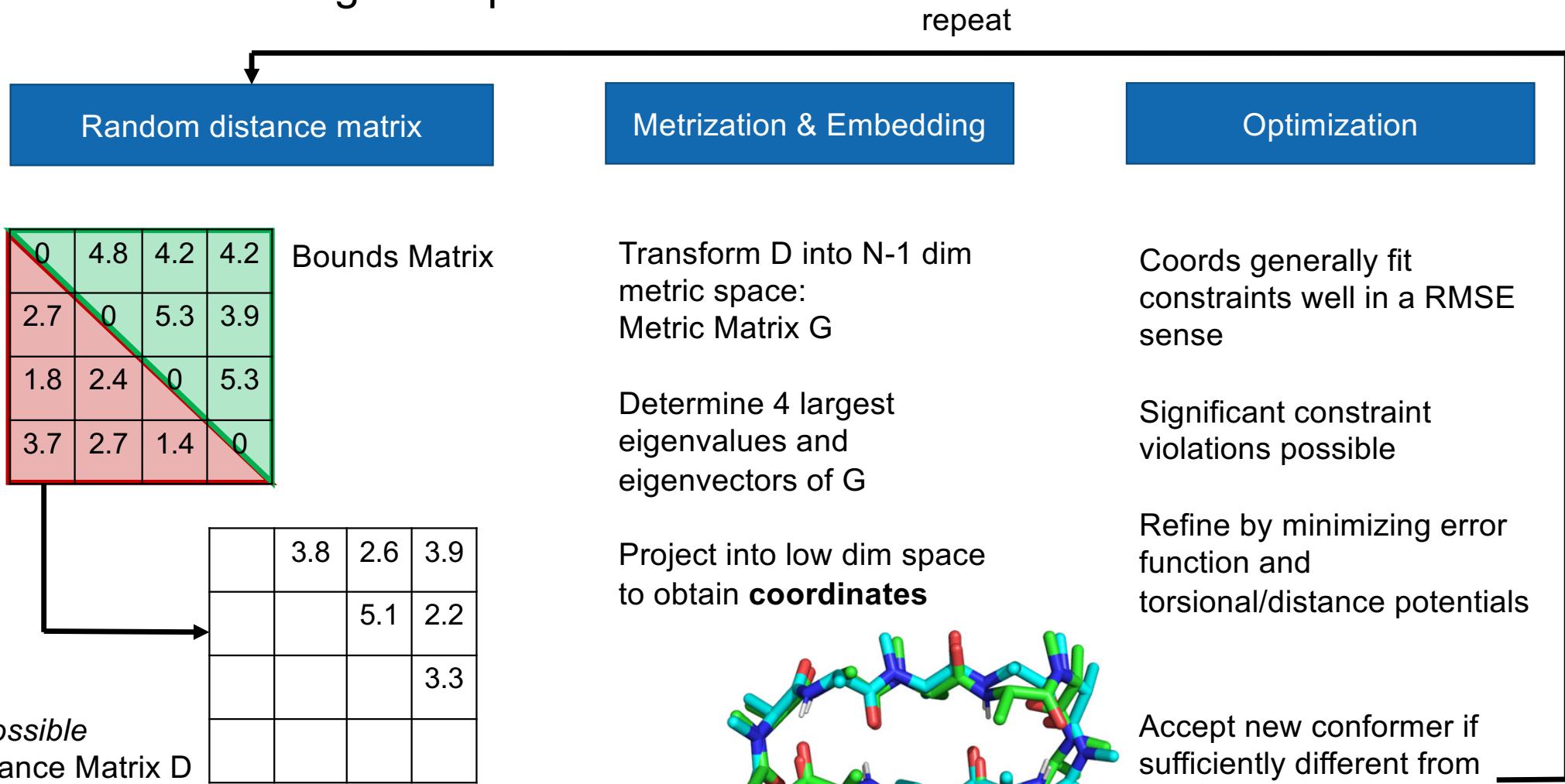
Triangle Inequality:



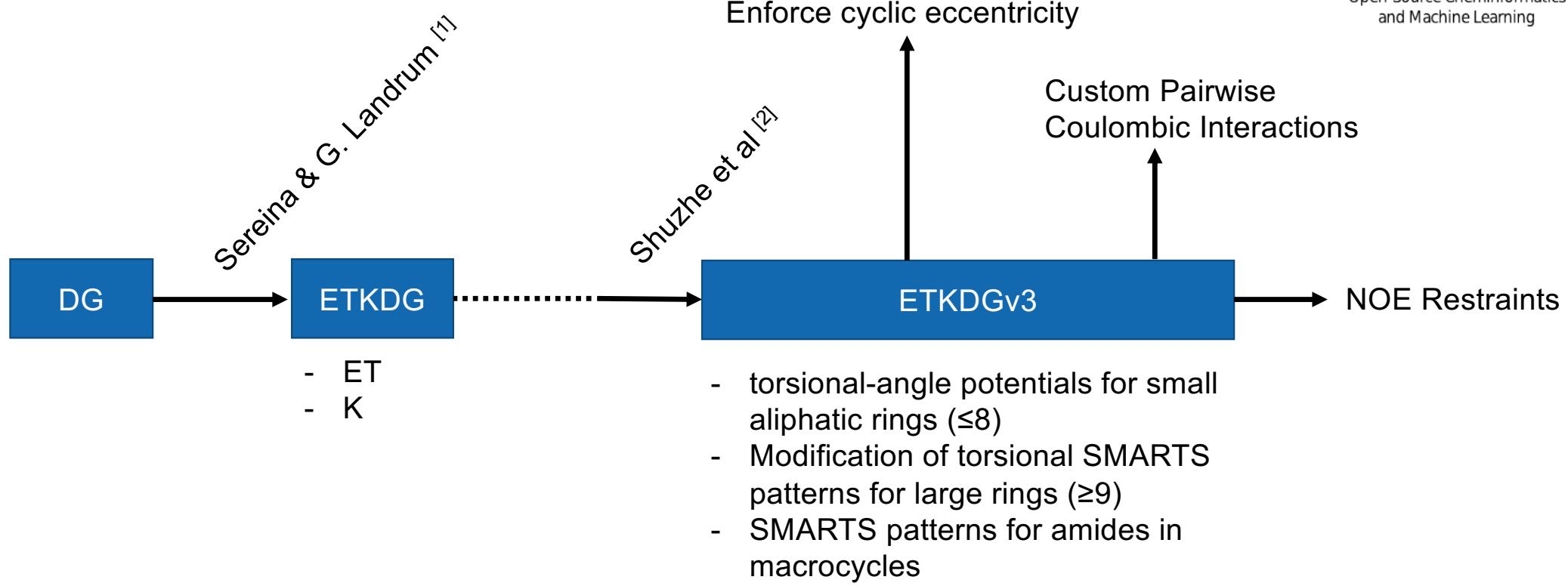
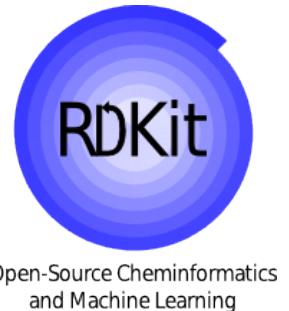
$$\overline{AC} \leq \overline{AB} + \overline{BC}$$

$$\overline{AC} \geq |\overline{AB} - \overline{BC}|$$

ETKDG: Working Principle



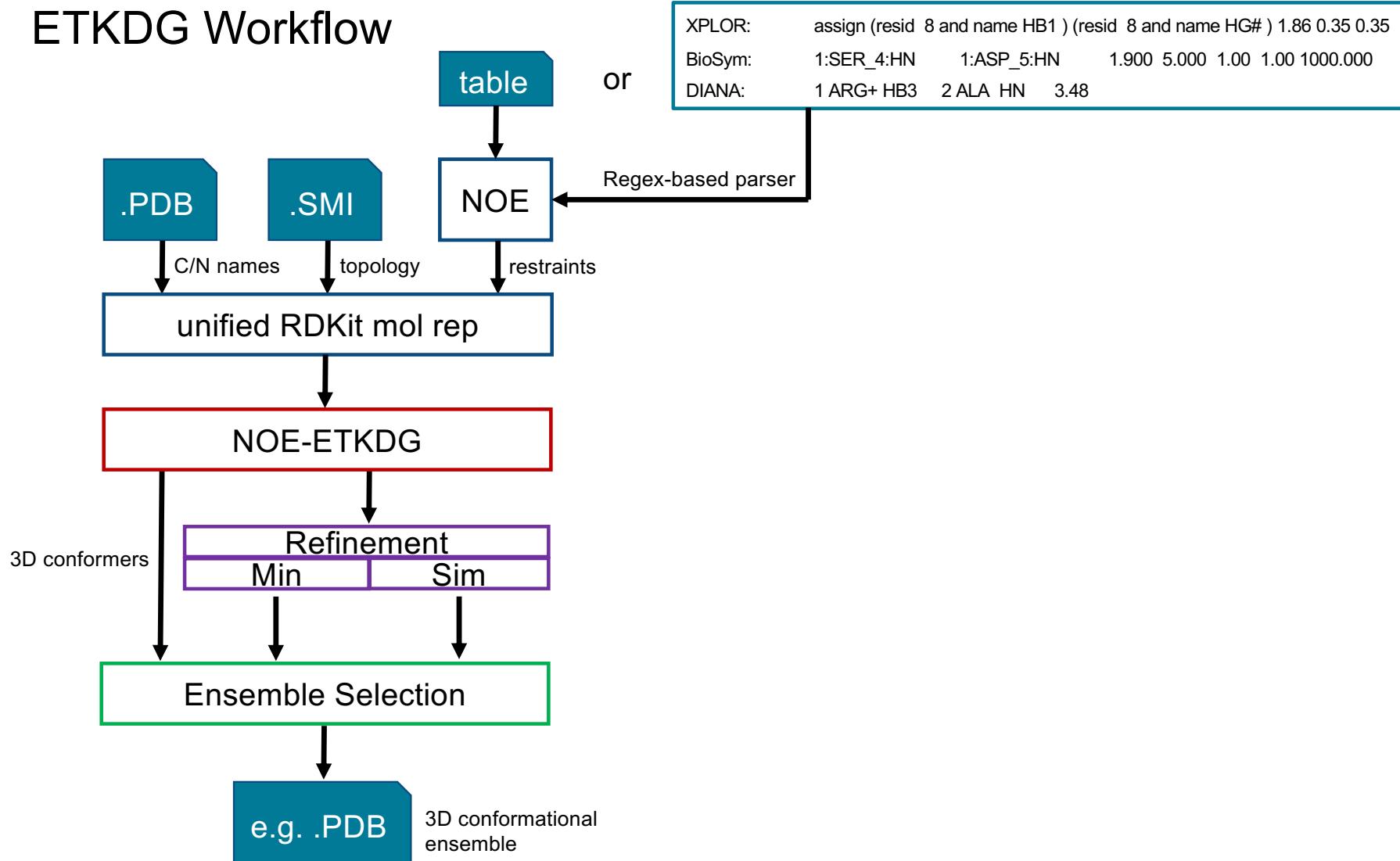
ETKDG: RDKit Developments Concerning Cyclic Peptides



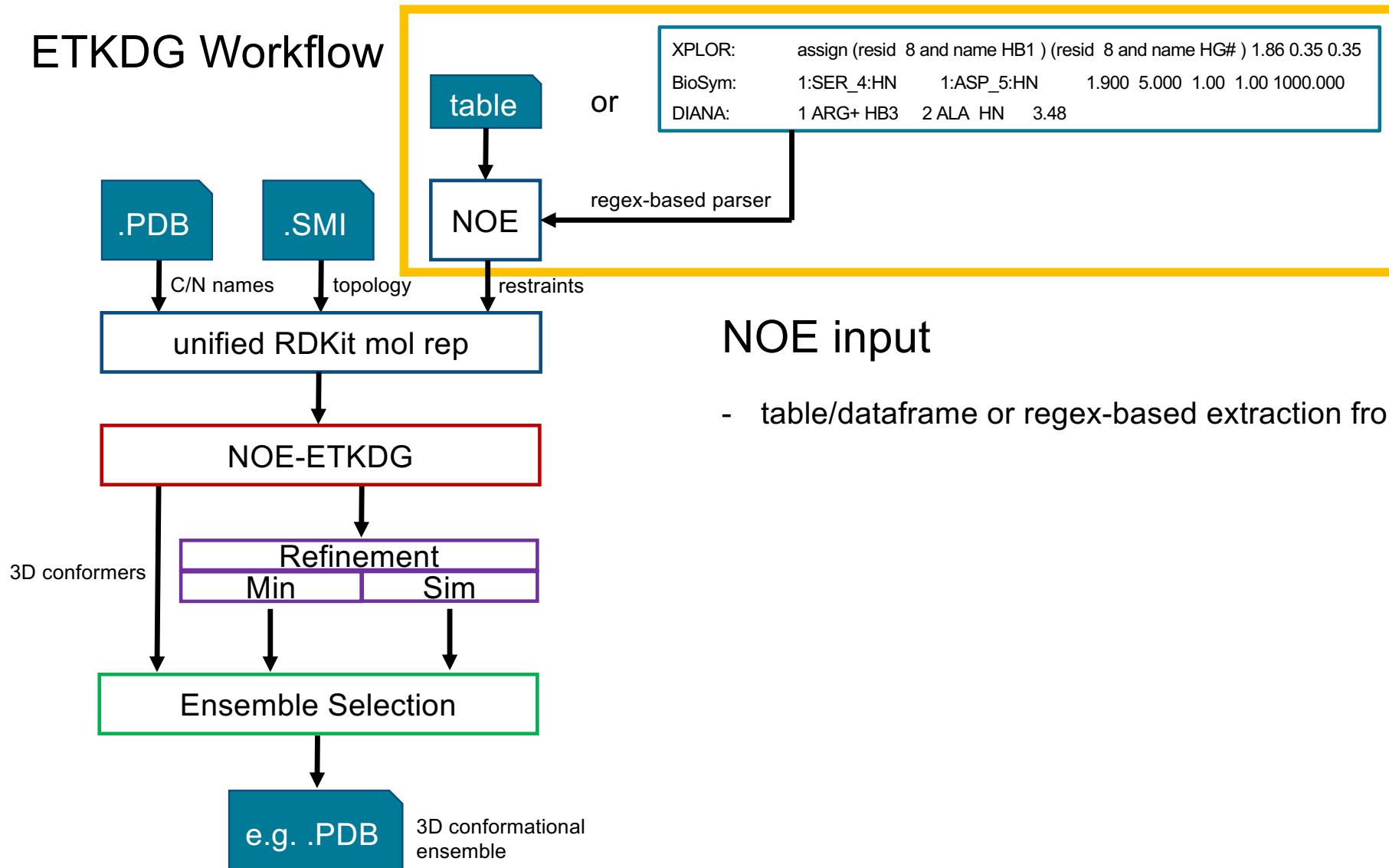
NOE-ETKDG

- Goals:
 - Use of NOE restraints to reduce search space and enable targeted conformer generation
 - Enable conformational analyses in water (high ϵ_r) and membrane (low ϵ_r) to assess permeability
 - Support non-standard amino acids (no residue templating)
- Framework:
 - Open-source solution to be added to RDKit
 - Workflow should include other open-source software only (OpenMM, Open Forcefield...)

ETKDG Workflow



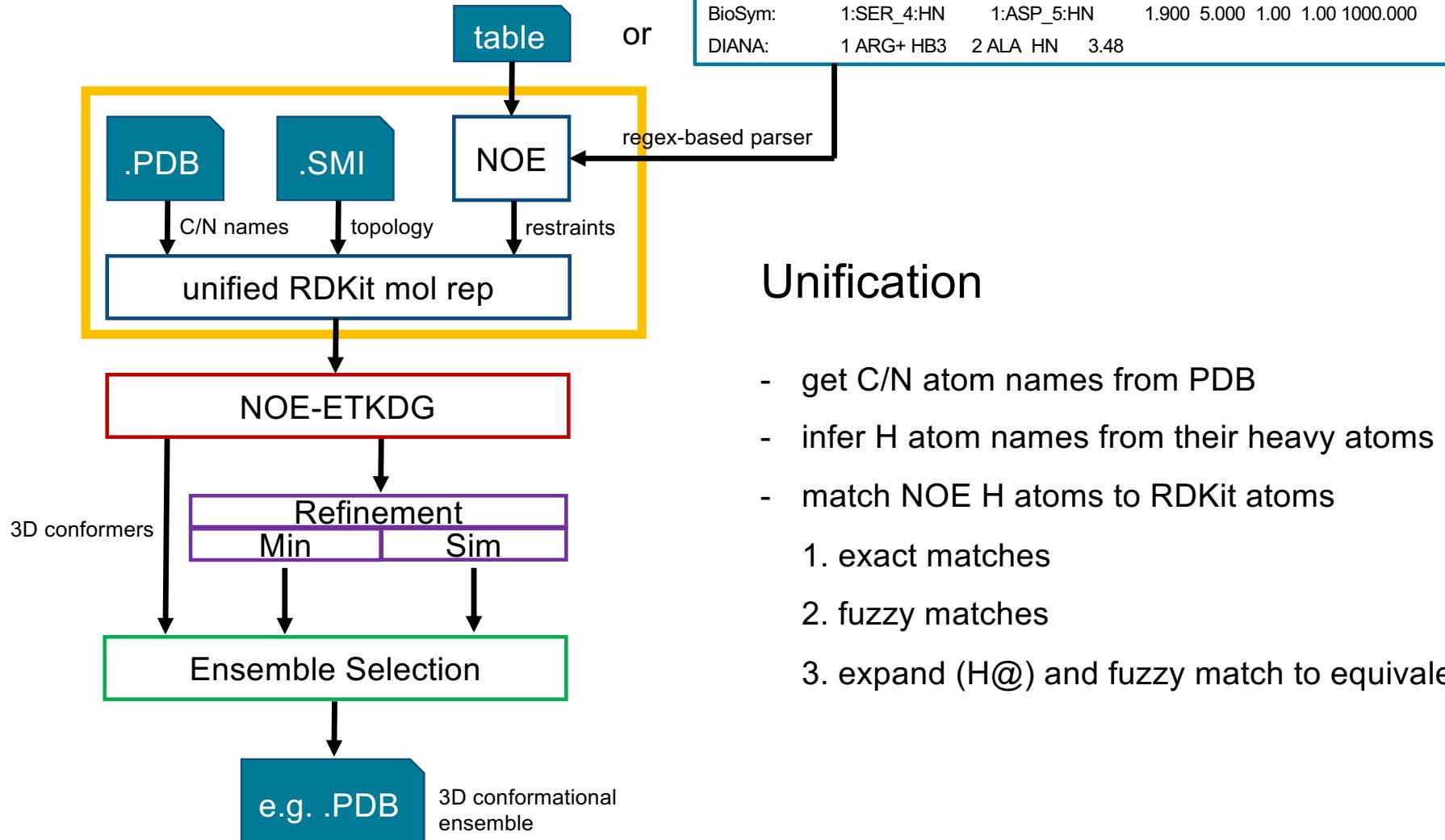
ETKDG Workflow



NOE input

- table/dataframe or regex-based extraction from text

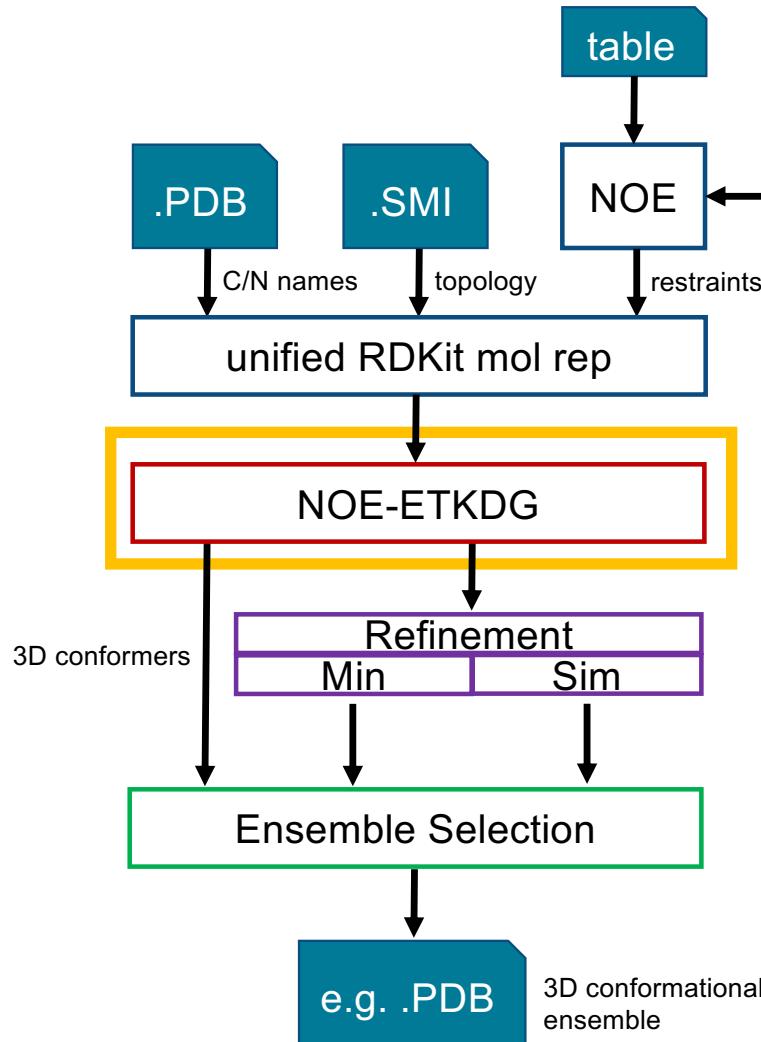
ETKDG Workflow



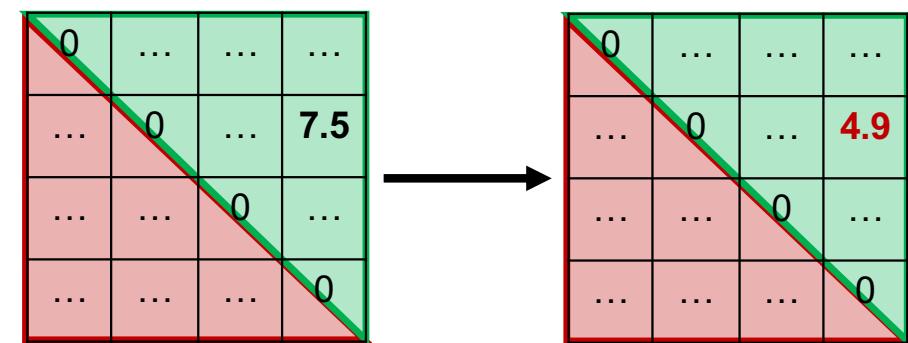
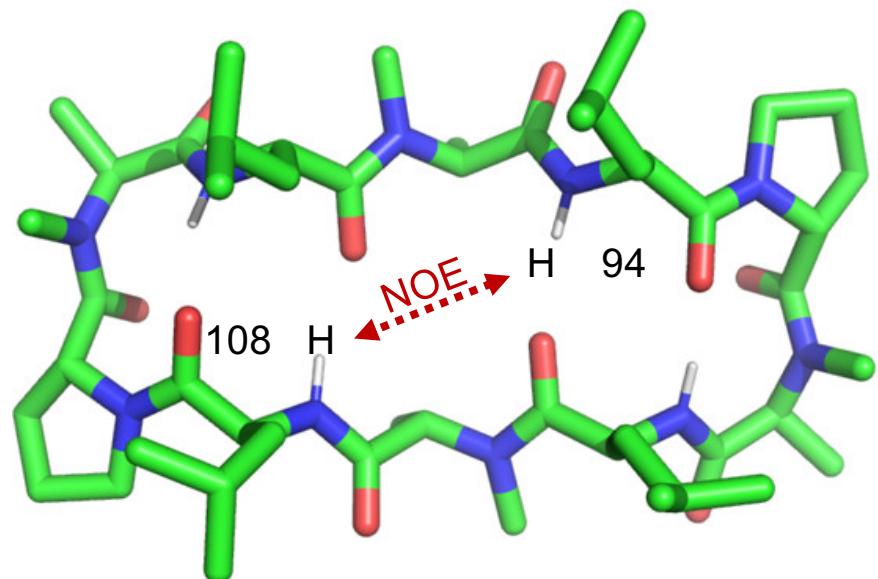
Unification

- get C/N atom names from PDB
- infer H atom names from their heavy atoms
- match NOE H atoms to RDKit atoms
 1. exact matches
 2. fuzzy matches
 3. expand (H@) and fuzzy match to equivalent set

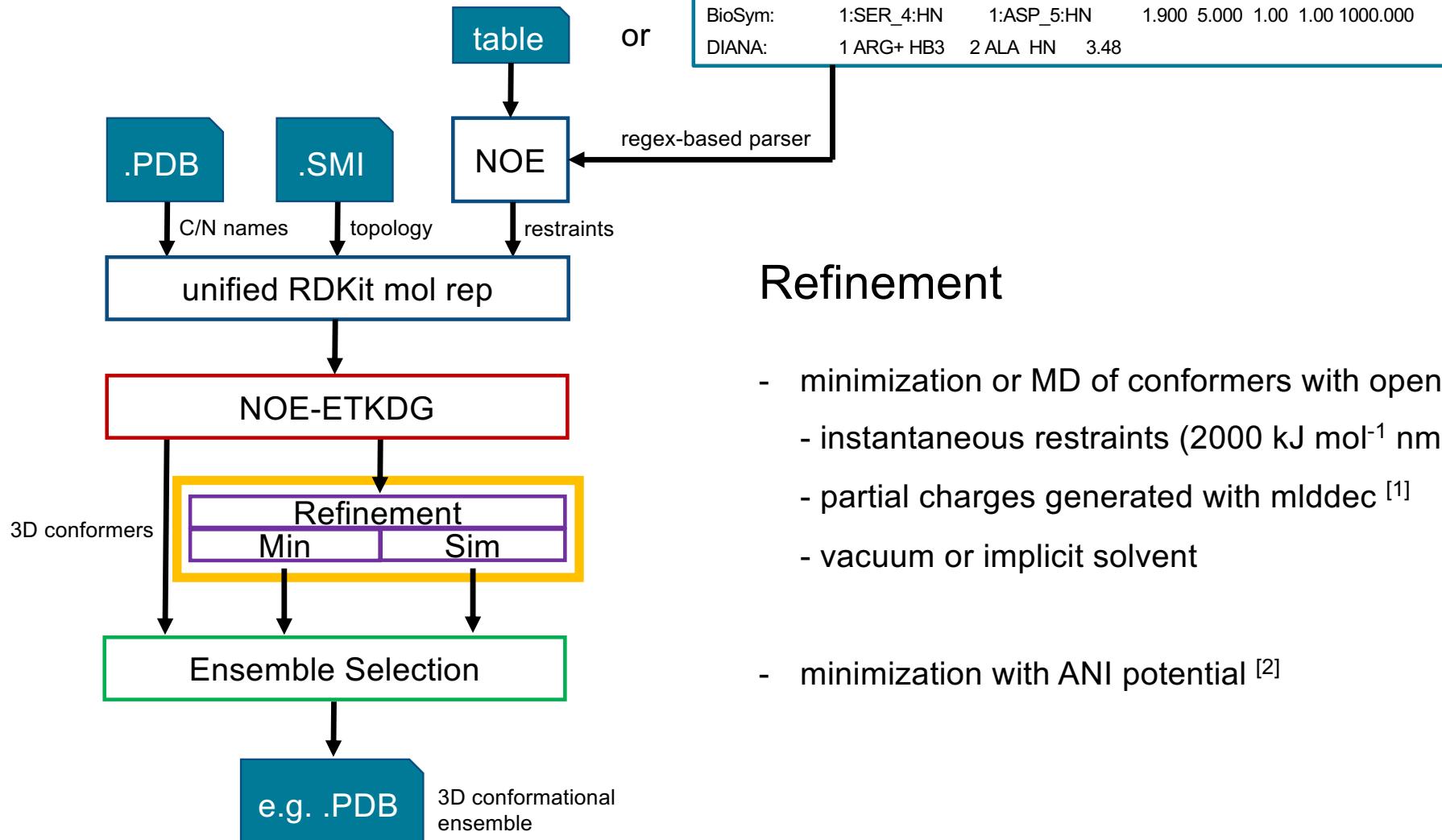
ETKDG Workflow



XPLOR: assign (resid 8 and name HB1) (resid 8 and name HG#) 1.86 0.35 0.35
 BioSym: 1:SER_4:HN 1:ASP_5:HN 1.900 5.000 1.00 1.00 1000.000
 DIANA: 1 ARG+ HB3 2 ALA HN 3.48



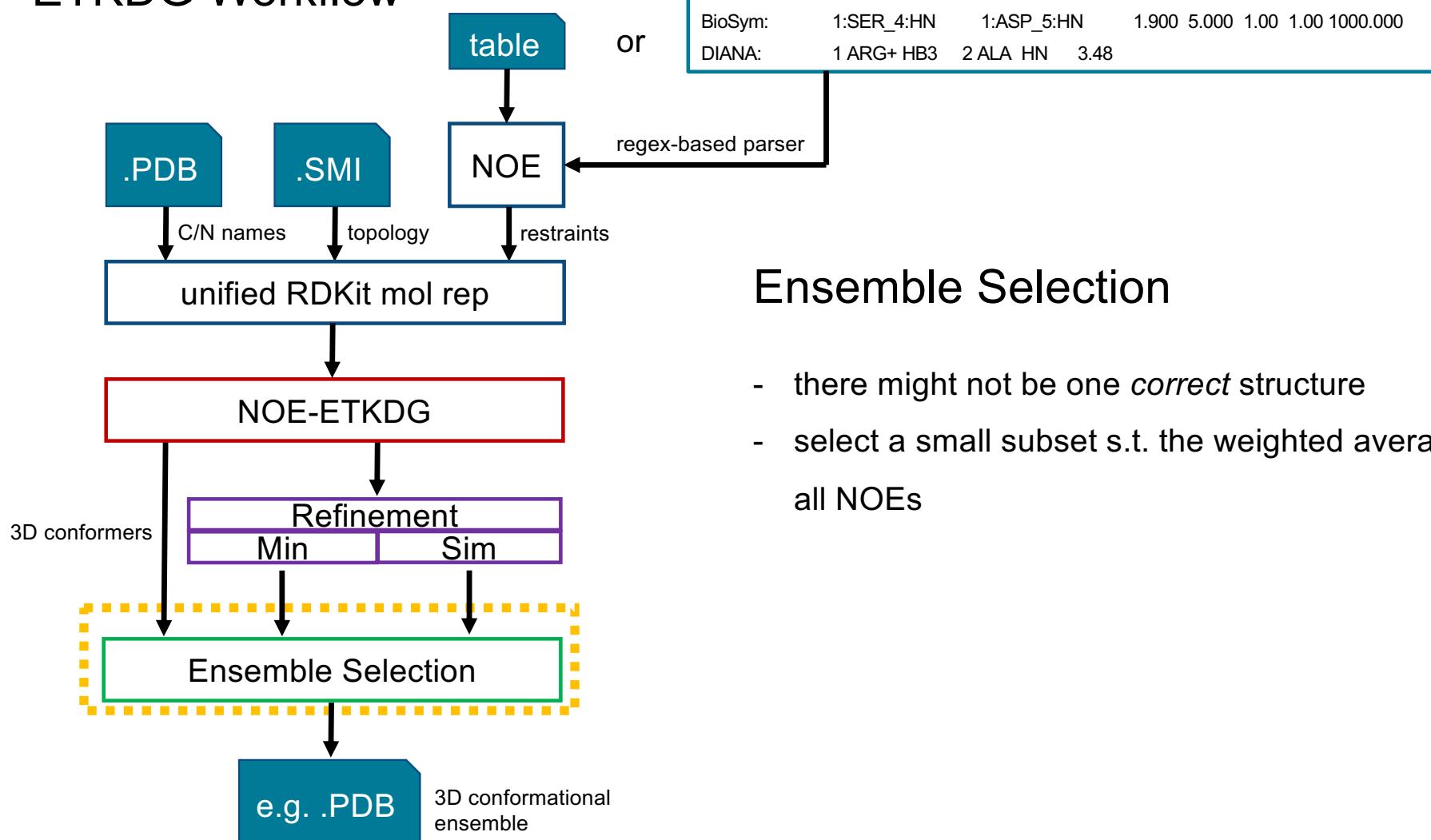
ETKDG Workflow



Refinement

- minimization or MD of conformers with openMM
 - instantaneous restraints ($2000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$)
 - partial charges generated with mlddec [1]
 - vacuum or implicit solvent
- minimization with ANI potential [2]

ETKDG Workflow



Ensemble Selection

- there might not be one *correct* structure
- select a small subset s.t. the weighted average fulfills all NOEs

Test Cases: Overview

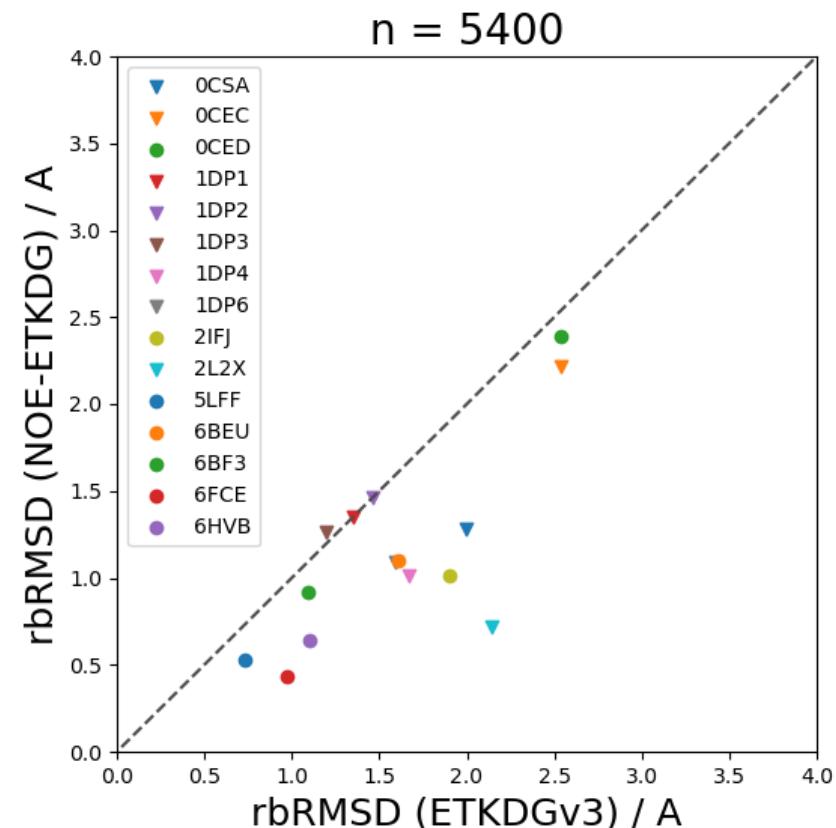
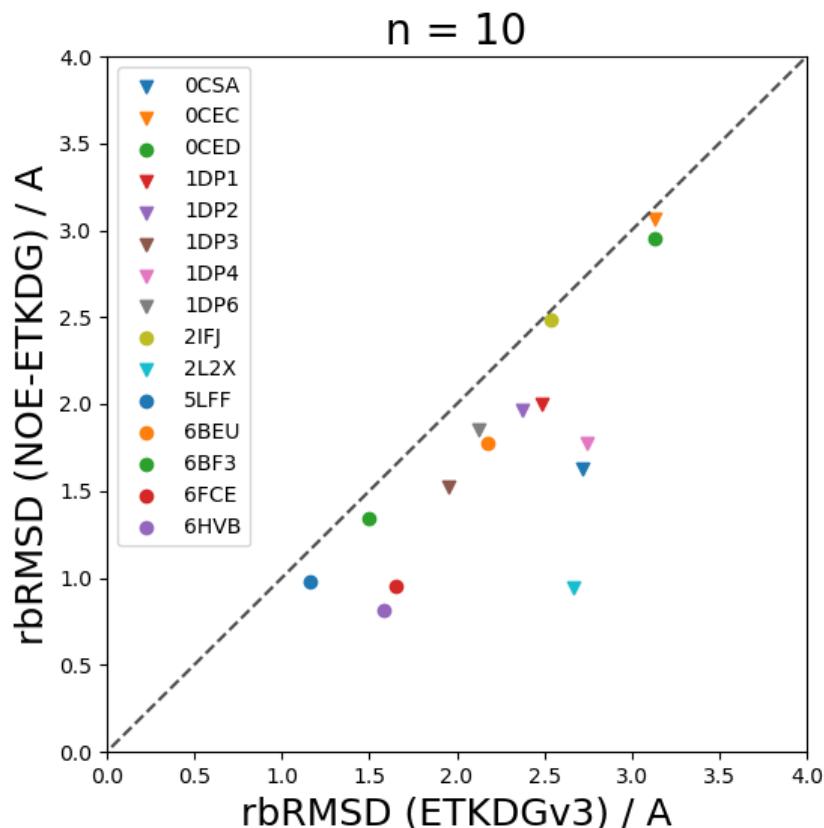
Code	Remark	Solvent	Dielectric
0CSA	Cyclosporin A	CHCl3	4.8
0CEC	Cyclosporin E	CCl4	2.2
0CED	Cyclosporin E	DMSO	48
1DP1	Decapeptide 1	CHCl3	4.8
1DP2	Decapeptide 2	CHCl3	4.8
1DP3	Decapeptide 3	CHCl3	4.8
1DP4	Decapeptide 4	CHCl3	4.8
1DP5	Decapeptide 5	CHCl3	4.8
1DP6	Decapeptide 6	CHCl3	4.8
2IFJ	Genentech / RCSB	D2O/H2O	78
2L2X	Genentech / RCSB	chloroform-d/ethanol-d5 5:1	~10 (*)
5LFF	Genentech / RCSB	D2O/H2O	78
6BEU	Genentech / RCSB	D2O/H2O	78
6BF3	Genentech / RCSB	D2O/H2O	78
6FCE	Genentech / RCSB	D2O/H2O	78
6HVB	Genentech / RCSB	D2O/H2O	78

Clarifications and Terminology

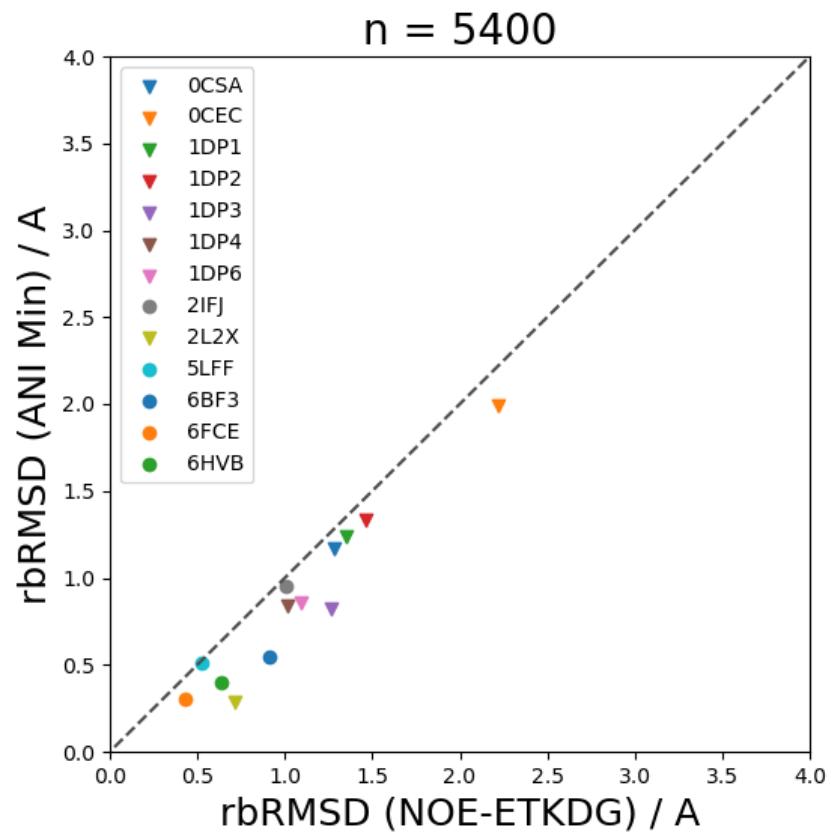
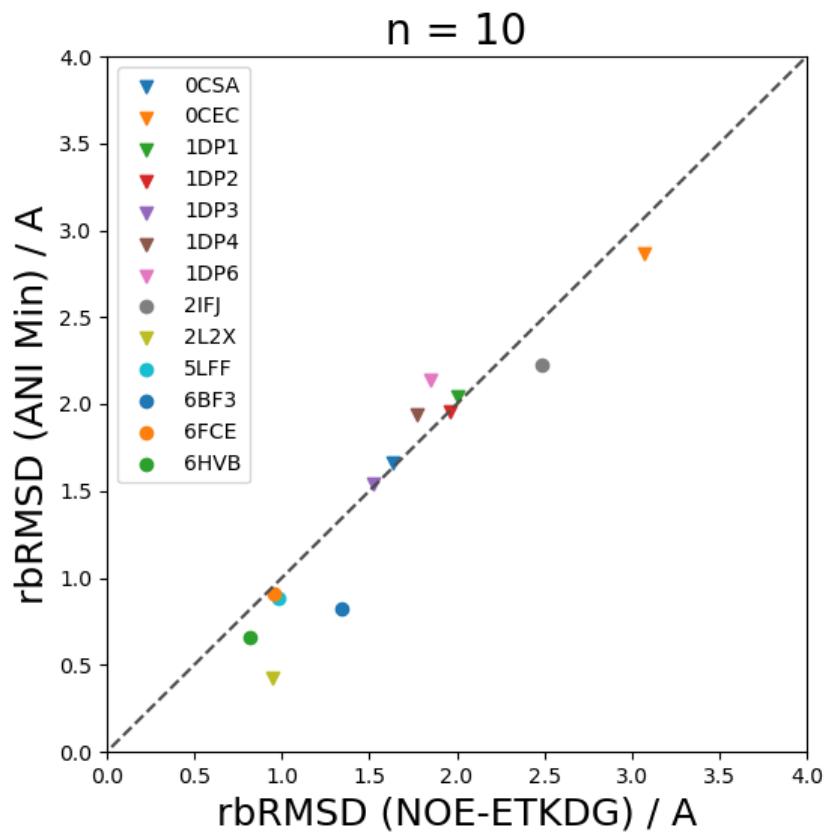
- Number of conformers:
 - 5400 for ETKDG and minimization techniques
 - 540 for 1 ns simulations
 - 54 for 5 ns simulations
- Implicit solvent min/sim according to dielectric of NOE solvent
- All min/sim with constraints (except ANI potential)
- rbRMSD: Root-mean square deviation of ring and beta-atoms between best conformer and first reference



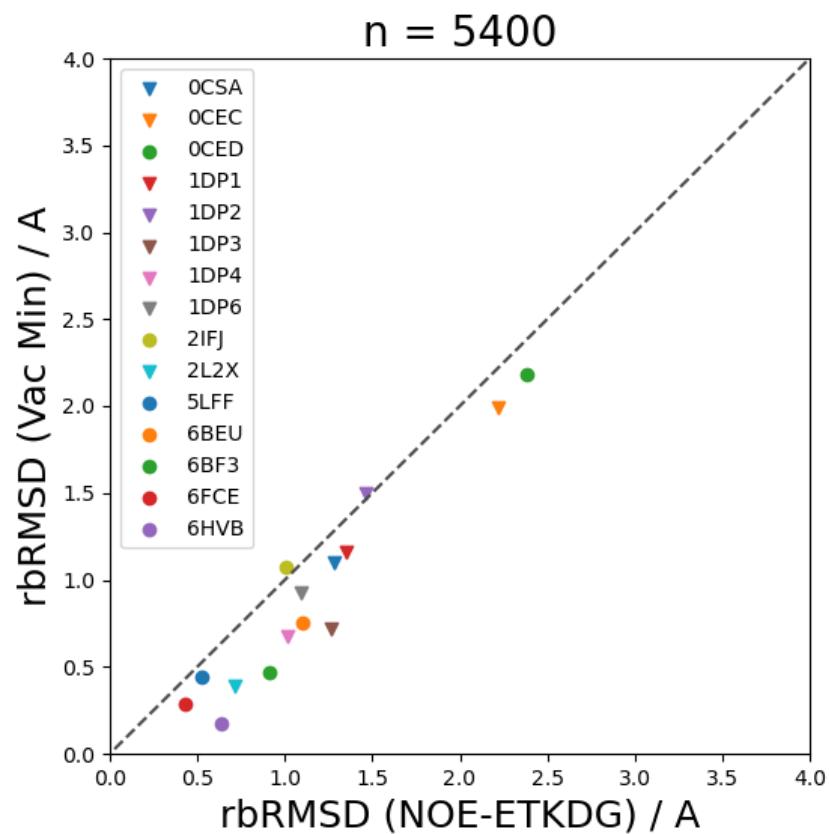
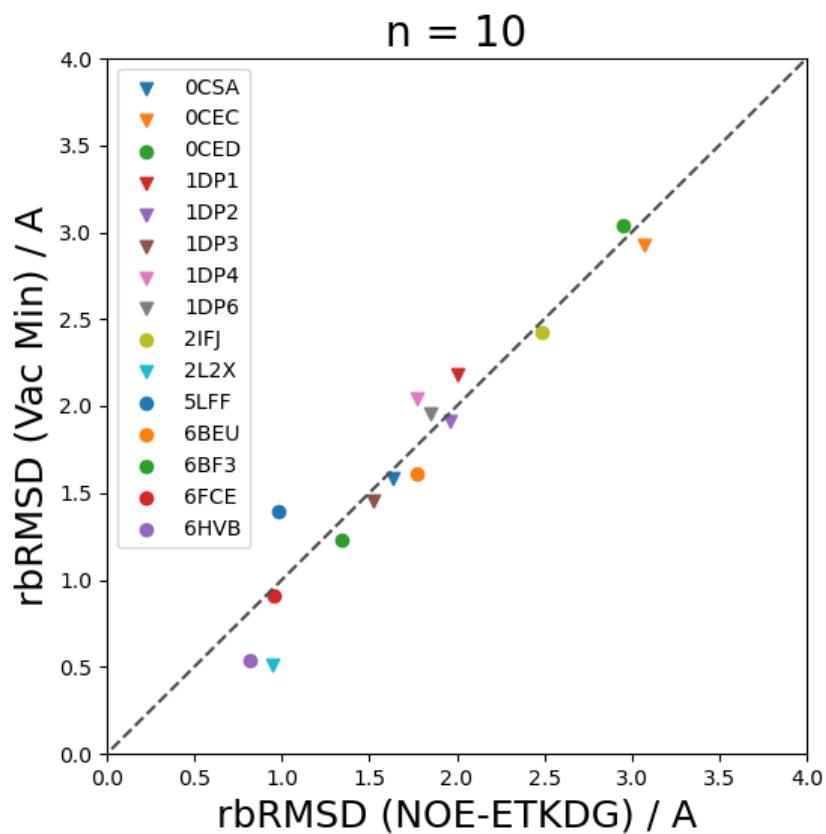
rbRMSD: ETKDGv3 v NOE-ETKDG



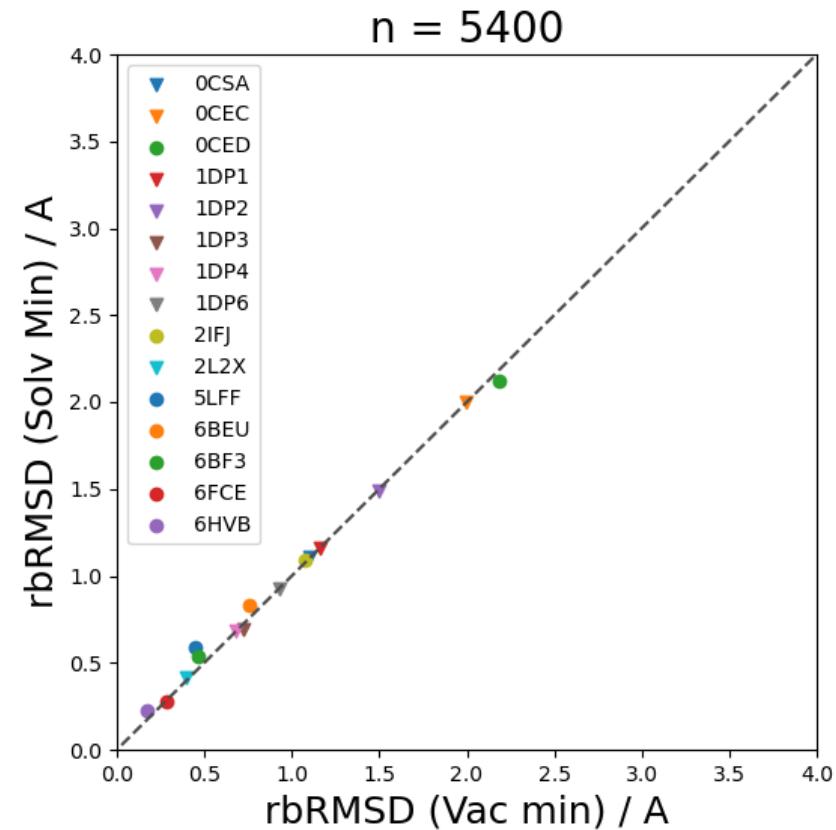
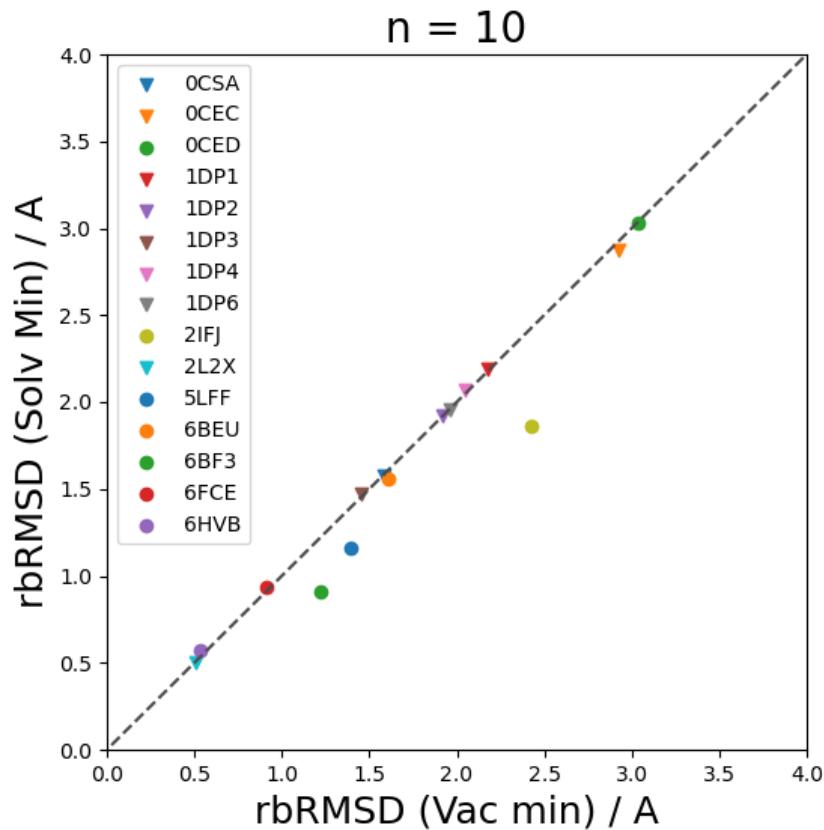
rbRMSD: NOE-ETKDG v ANI Min



rbRMSD: NOE-ETKDG v Vac Min

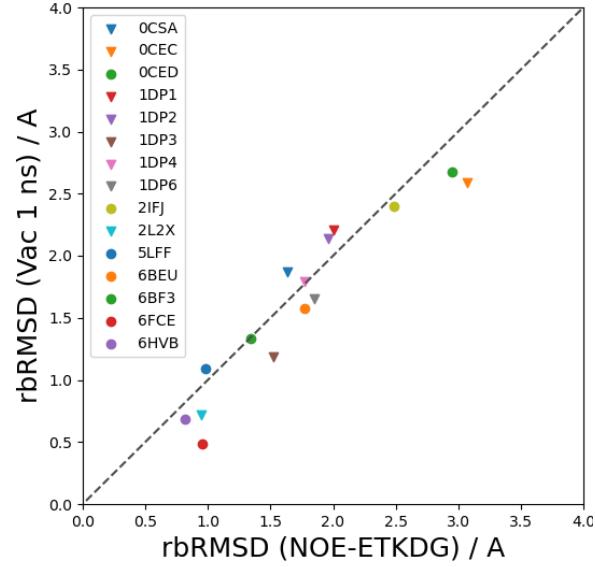


rbRMSD: Vac Min v Solv Min

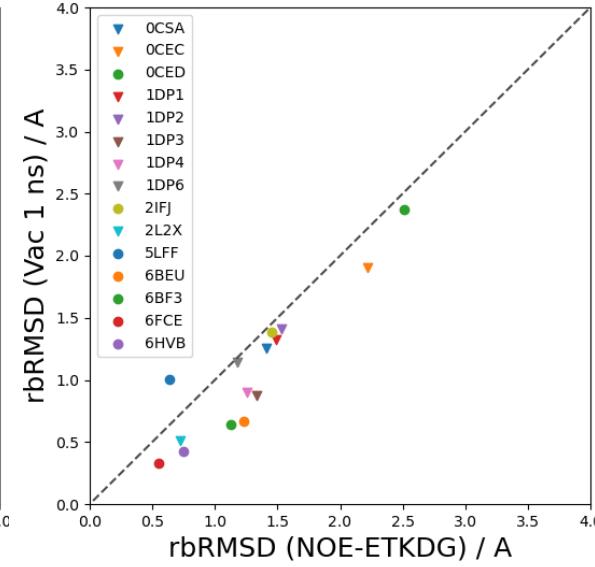


rbRMSD: NOE-ETKDG v Vac Sim

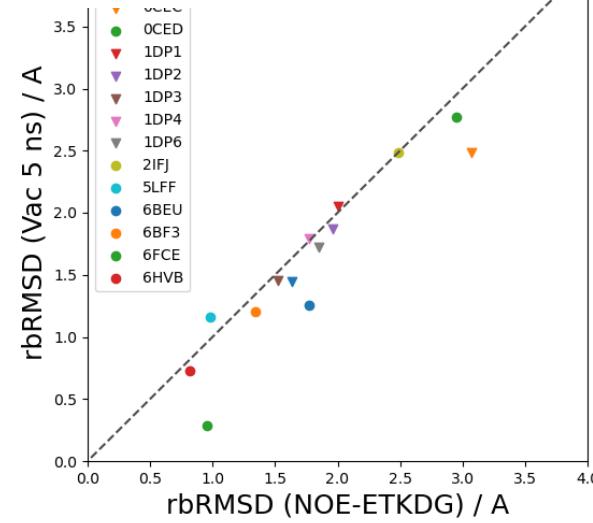
$n = 10$



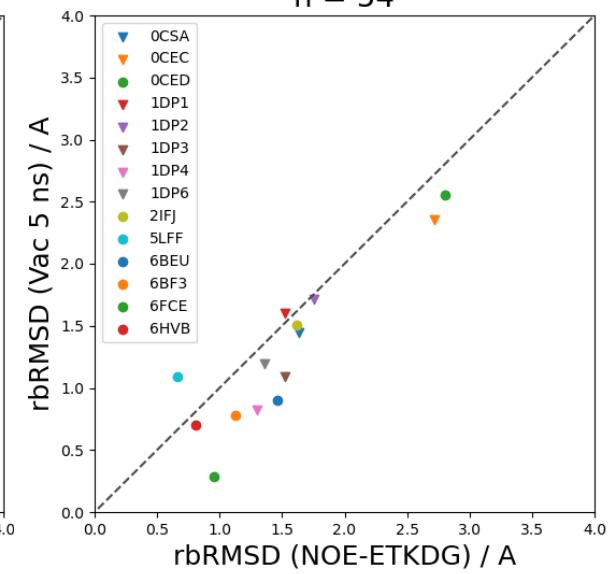
$n = 540$



$n = 10$

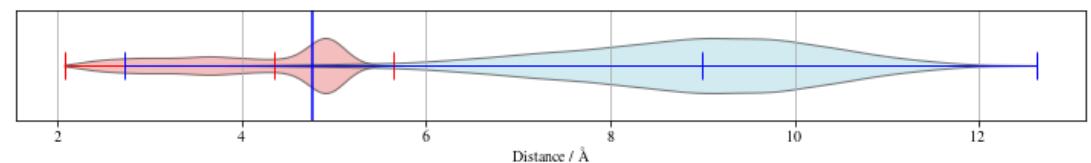
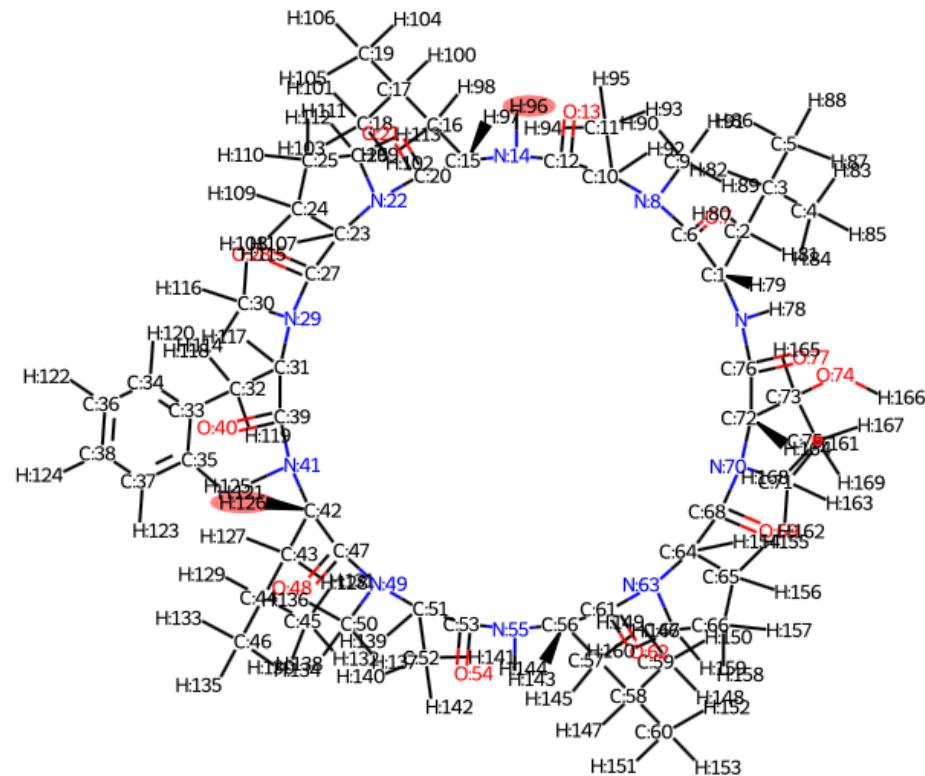


$n = 54$



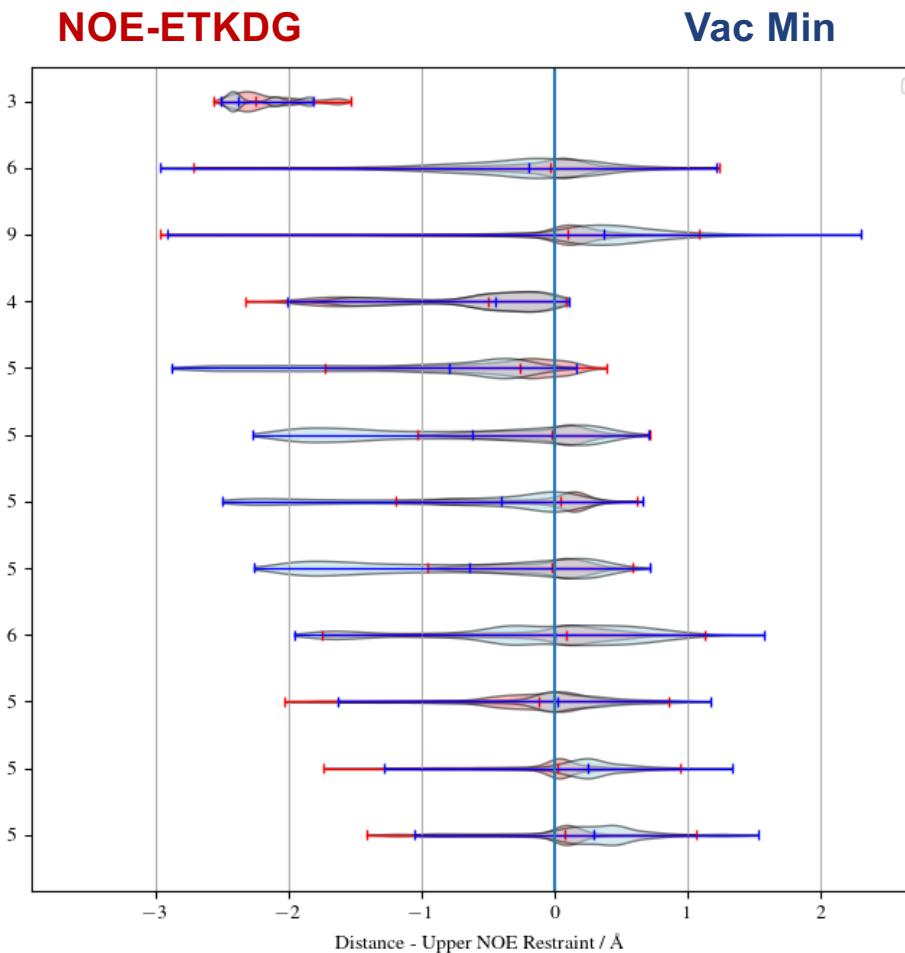
NOE Pairs: Distance Distributions

NOE ————— 22

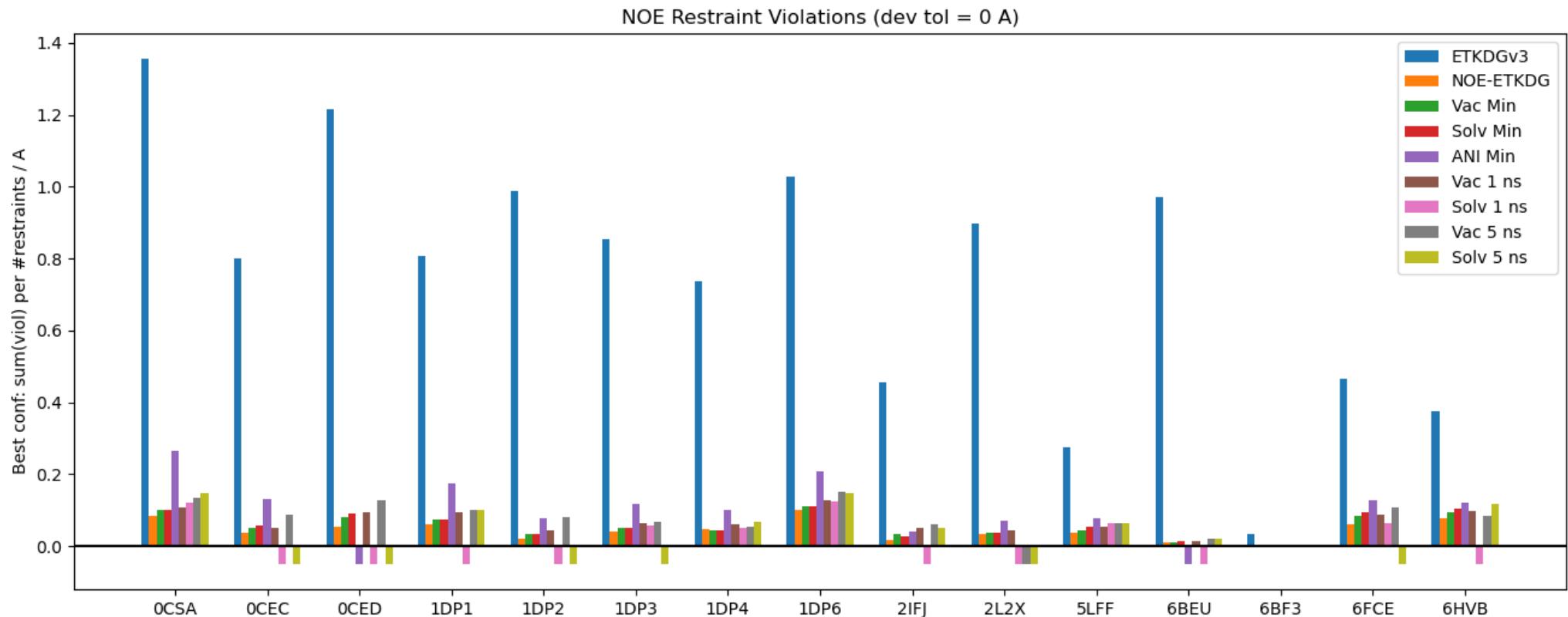


--> Live demo, added pic for archive

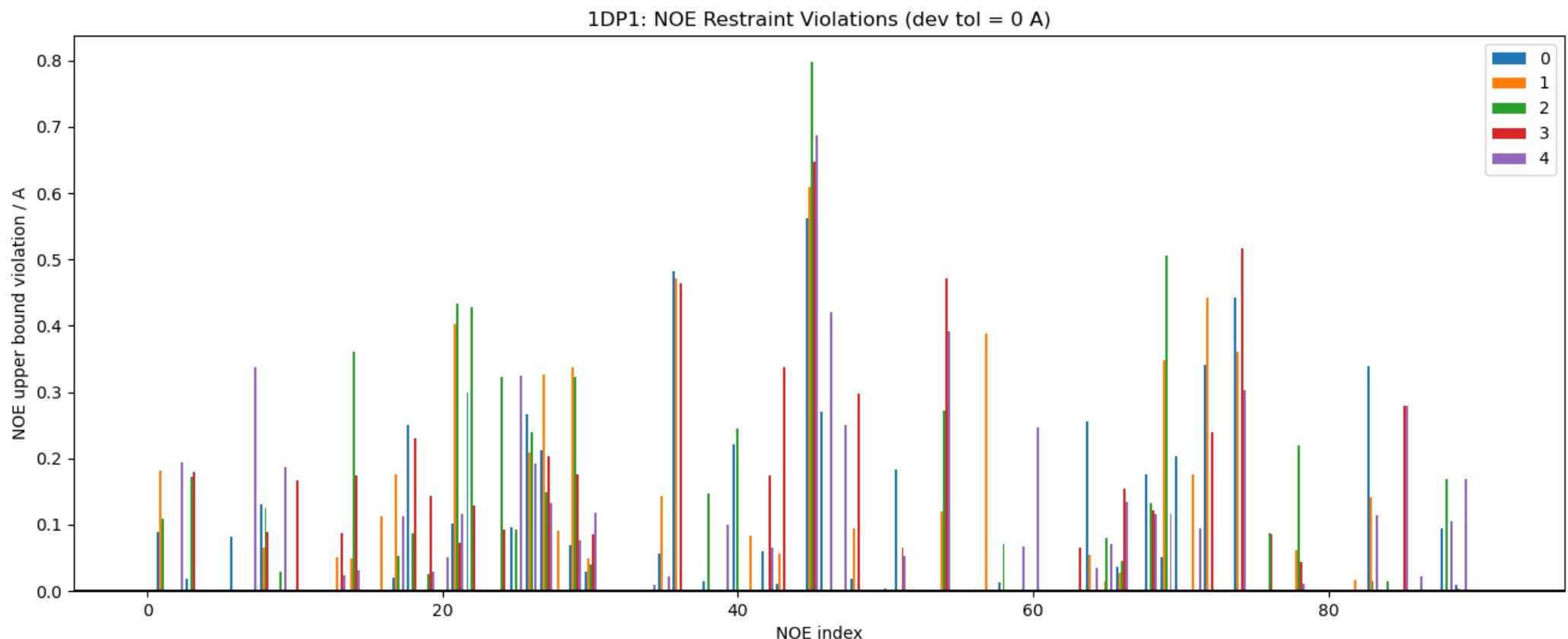
NOE Pairs: Broader Distance Distributions after Minimization



Violations of NOE Restraints



Single Molecule NOE Index graph



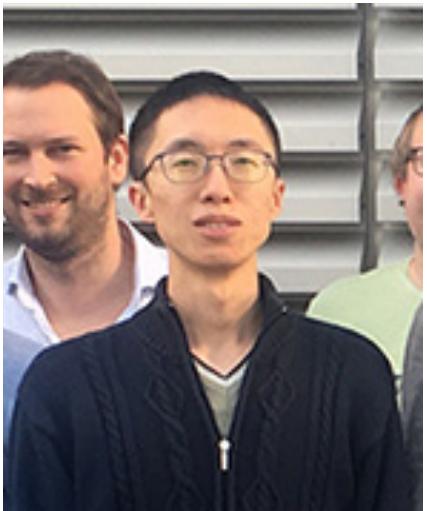
Conclusions

- NOE-ETKDG effectively shrinks the search space to enable targeted conformer generation with restraints.
- Different minimization schemes are a fast way to improve rbRMSD.
- Simulations of such macrocycles is already computationally expensive and adds (as done so far) little value
- Minimization/Simulation (as done so far) add little value w.r.t NOE restraint violations, but might still help due to increased variation.

Outlook

- More test cases to increase robustness
- Time-averaged NOE restraints for openMM simulations
- Find a good and fast charge generator for different dielectrics
- Open-source/RDKit based ensemble selection tool

Final Remarks



Genen
tech

CCG
CSMS



