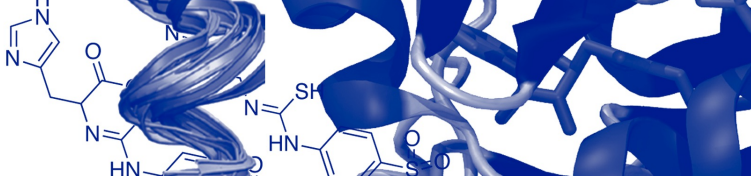
Computational Chemistry Group

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich



Combining ETKDG and NOE-derived Distance Bounds

Prof. Sereina Riniker

10th RDKit UGM, October 15, 2021

What is ETKDG?

ETKDG = Experimental torsions – knowledge (terms) – distance geometry

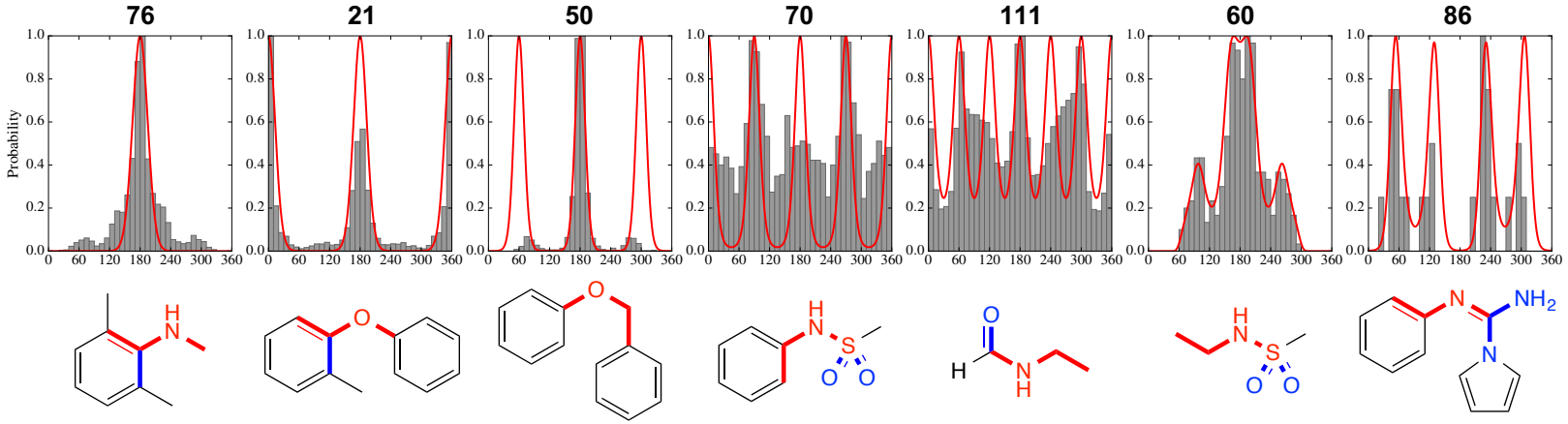
Experimental torsional-angle preferences of *acyclic* bonds from CSD

Examples of fitted torsion potentials:

$$V_{tors}(x) = \sum_{i=1}^{6} K_i \left(1 + \cos(\delta_i)\cos(m_i x)\right)$$

387 SMARTS patterns from:

- **Version 1**: Schärfer et al., *J. Med. Chem.*, 56, 2016 (2013).
- **Version 2**: Guba et al., *J. Chem. Inf. Model.*, 56, 1 (2016).



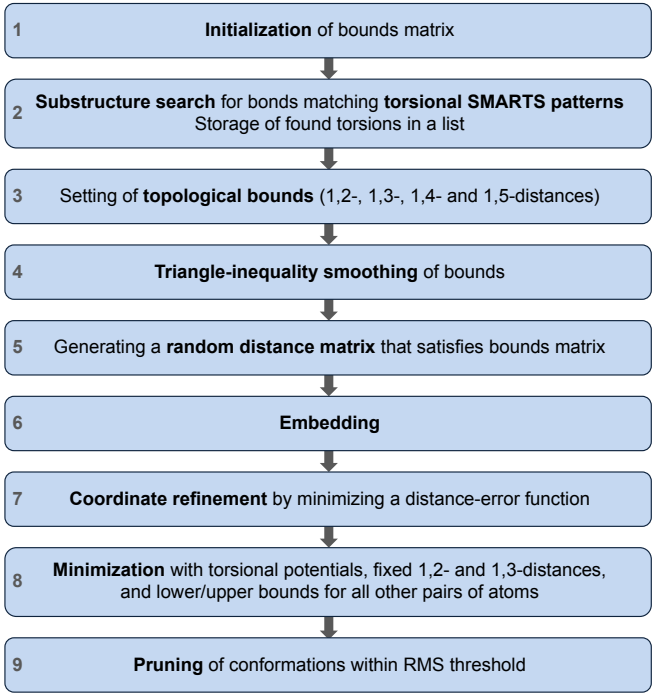
Knowledge terms: Flat sp2 carbons, straight triple bonds

Riniker, Landrum, *J. Chem. Inf. Model.*, **55**, 2562 (2015).

What is ETKDG?

ETKDG = Experimental torsions – knowledge (terms) – distance geometry

Conformer generation workflow



Riniker, Landrum, *J. Chem. Inf. Model.*, **55**, 2562 (2015).

ETKDG version 3

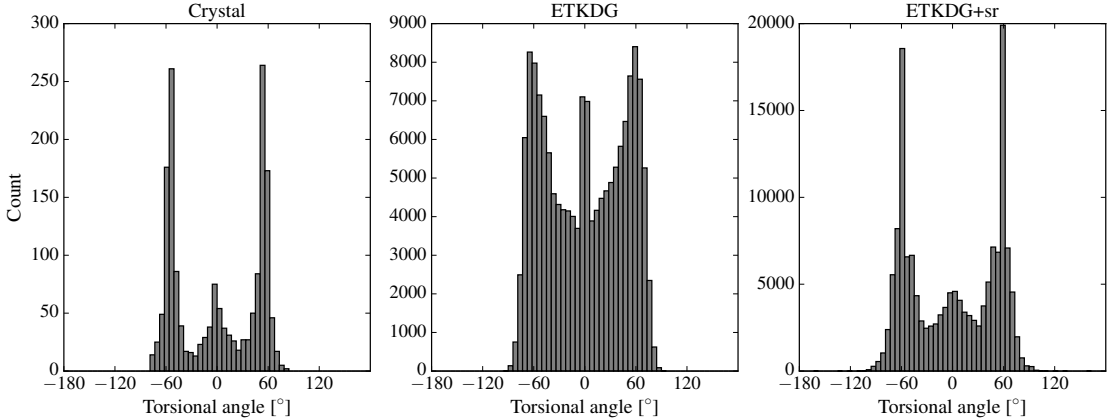
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

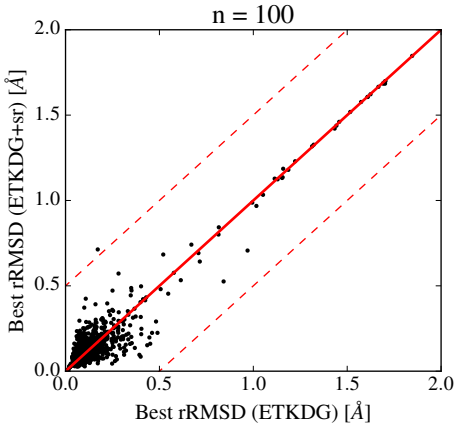
Small rings:

- New torsion SMARTS pattern derived (105 in total)
- Maximum ring size = 8
- Bridged systems: Bonds in >1 SSSR ring excluded, if ring shares >2 bonds with another ring → excluded

Six-membered rings



Best rRMSD on CSD test set



Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

ETKDG version 3

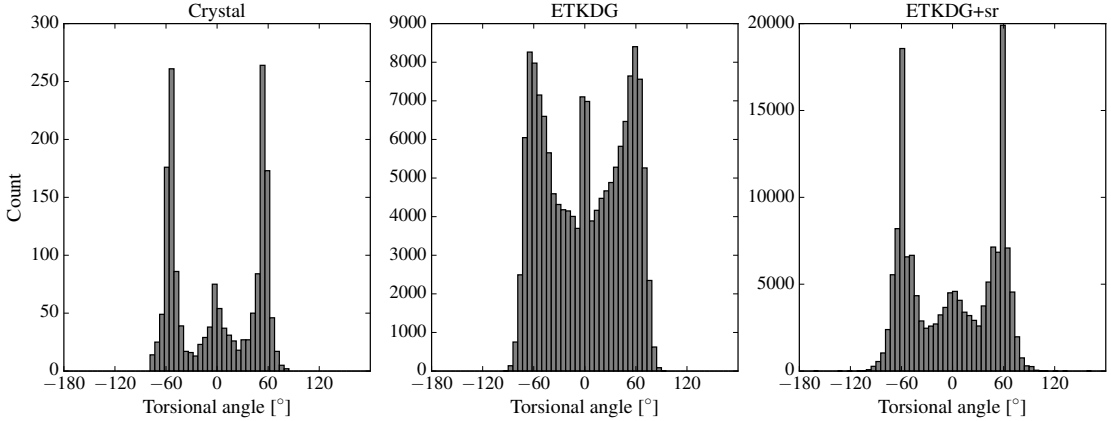
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

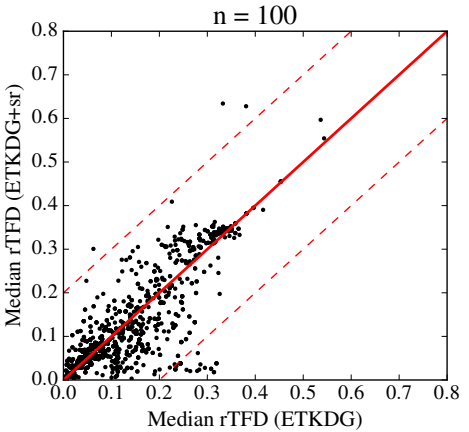
Small rings:

- New torsion SMARTS pattern derived (105 in total)
- Maximum ring size = 8
- Bridged systems: Bonds in >1 SSSR ring excluded, if ring shares >2 bonds with another ring → excluded

Six-membered rings



Median rTFD on CSD test set



Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

ETKDG version 3

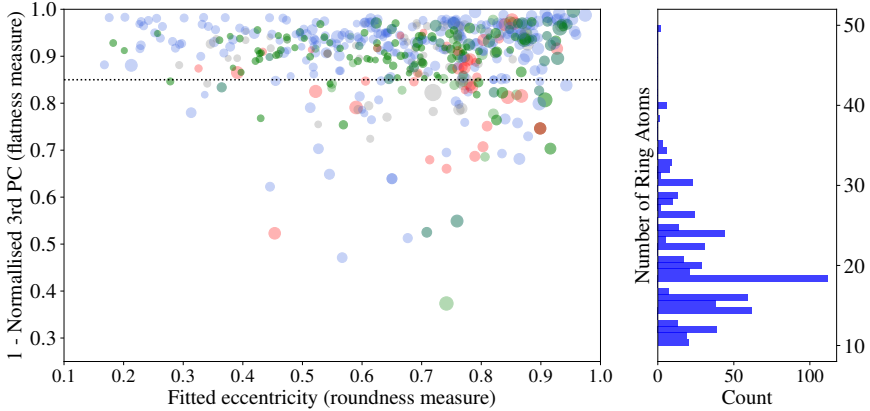
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

Macrocycles:

- Cyclic bonds in rings > 8 behave similar to acyclic bonds (in terms of torsions)
- SMARTS patterns from version 2 adapted for macrocyclic bonds
- Changes to existing 1,4-bounds subroutines to treat macrocyclic bonds as acylic

Flatness and ring size in dataset



Size of markers = macrocyclic ring size
Colours = source of the macrocycles:

blue = CSD
green = Prime
red = BIRD
grey = Mac10

Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

ETKDG version 3

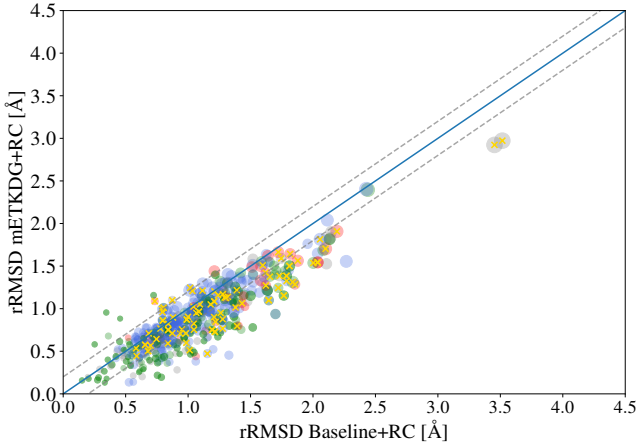
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

Macrocycles:

- Cyclic bonds in rings > 8 behave similar to acyclic bonds (in terms of torsions)
- SMARTS patterns from version 2 adapted for macrocyclic bonds
- Changes to existing 1,4-bounds subroutines to treat macrocyclic bonds as acylic

mETKDG versus ETKDGv2 (random coordinates)



Size of markers = macrocyclic ring size
Colours = source of the macrocycles:
blue = CSD
green = Prime
red = BIRD
grey = Mac10
Blue solid lines: y = x
Gray dotted lines: ±0.20 Å deviation
Golden crosses on a marker = compounds for which the experimental structure has a flatness value of <0.85

Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

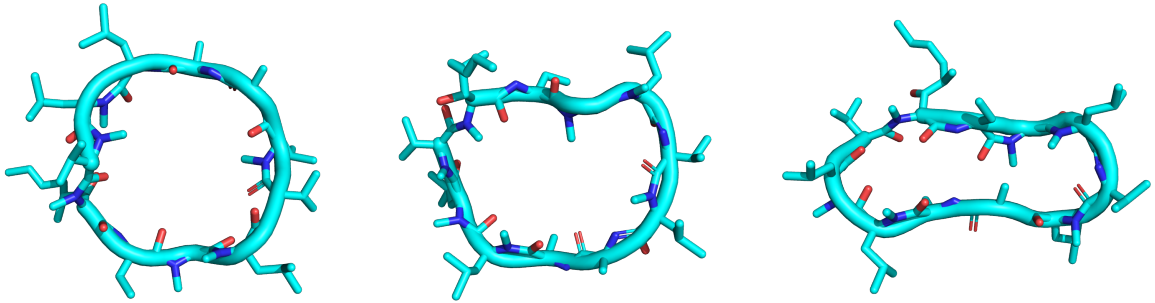
ETKDG version 3

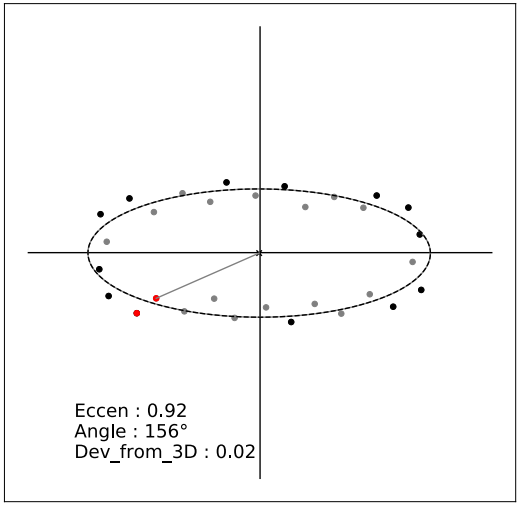
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

Macrocycles: Eccentricity

- Crystal conformations of macrocycle backbone are often ellipsoid





Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

ETKDG version 3

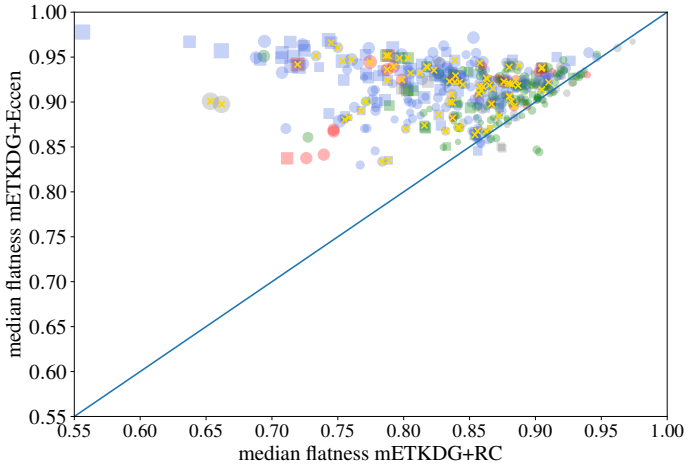
What about aliphatic cyclic bonds?

Distinguishing between small rings and macrocycles

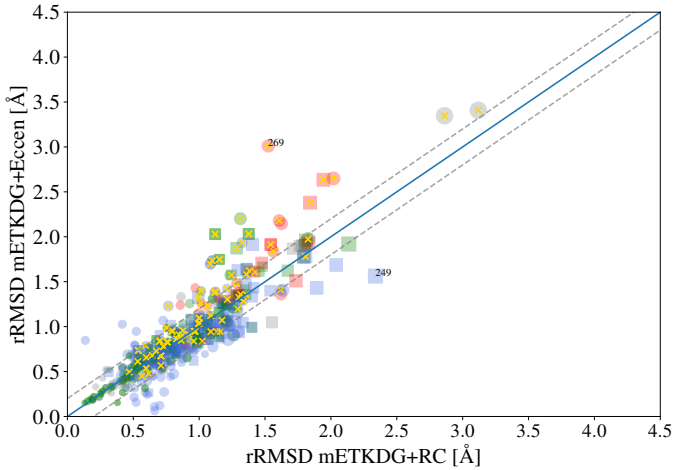
Macrocycles: Eccentricity

- Crystal conformations of macrocycle backbone are often ellipsoid
- Eccentricity constraints to target search in conformational space

Median flatness



Best rRMSD



Wang, Witek, Landrum, Riniker, *J. Chem. Inf. Model.*, **60**, 2044 (2020).

ETKDG + NOE

Can we use NOE-derived distance bounds to direct conformational search?

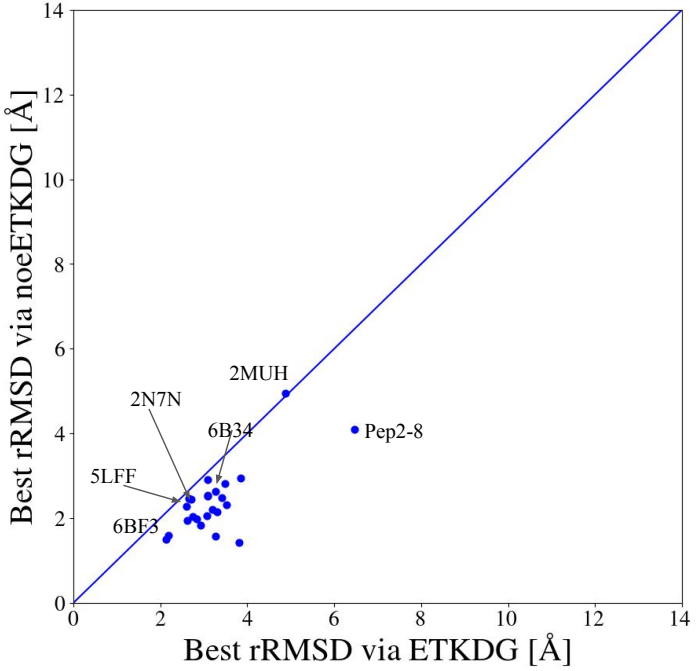
Goal: Conformer bundles fulfilling experimental restraints

ETKDG workflow:

- Modification of upper bounds in the molecular bounds matrix
- Advantage: No force field involved (easy to handle unnatural amino acids, non-peptidic motifs)

Dataset + pipeline:

- 24 peptides (23 cyclic + 1 linear)
- Parsing: XPLOR-type files
- 2000 conformer generated with noeETKDG
- Bundles: Subselection of 9-20 conformers
  - Random
  - Most diverse macrocycle conformations
  - Lowest energy
  - Least NOE violations (LICUV)
  - NAMFIS [JACS, 117, 1027 (1995)]



Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

ETKDG + NOE

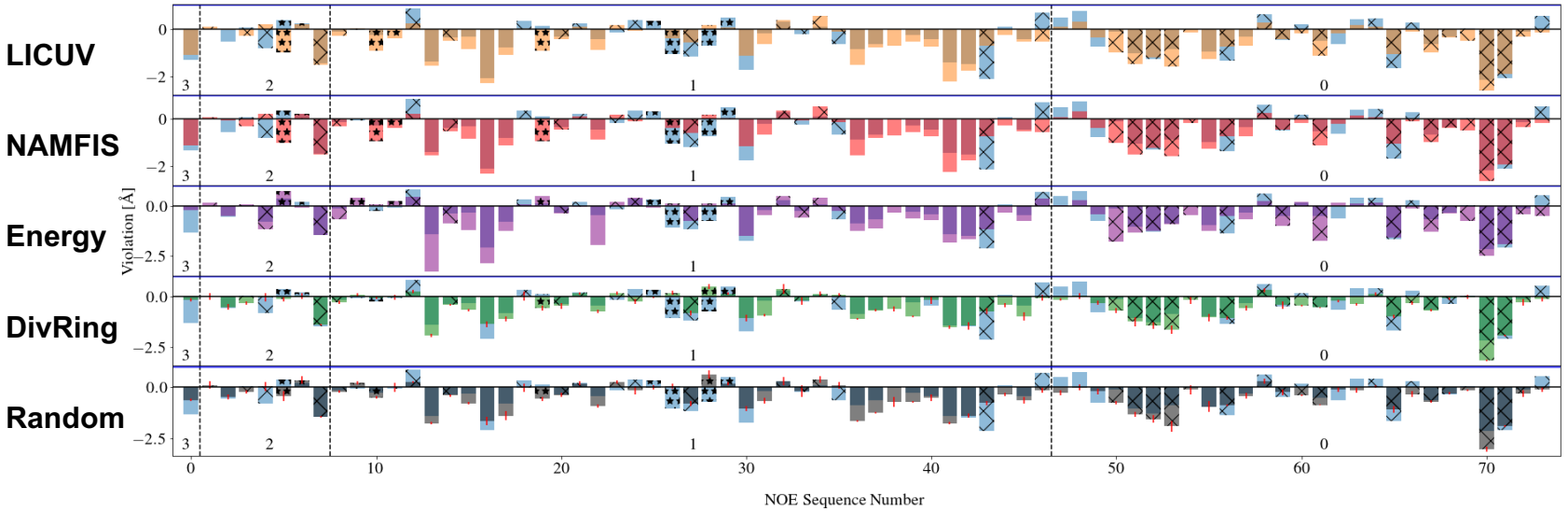
Can we use NOE-derived distance bounds to direct conformational search?

Goal: Conformer bundles fulfilling experimental restraints

Comparison of bundles:

Violations of NOE upper bounds (example 6FCE)

Blue = reference bundle



Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

ETKDG + NOE

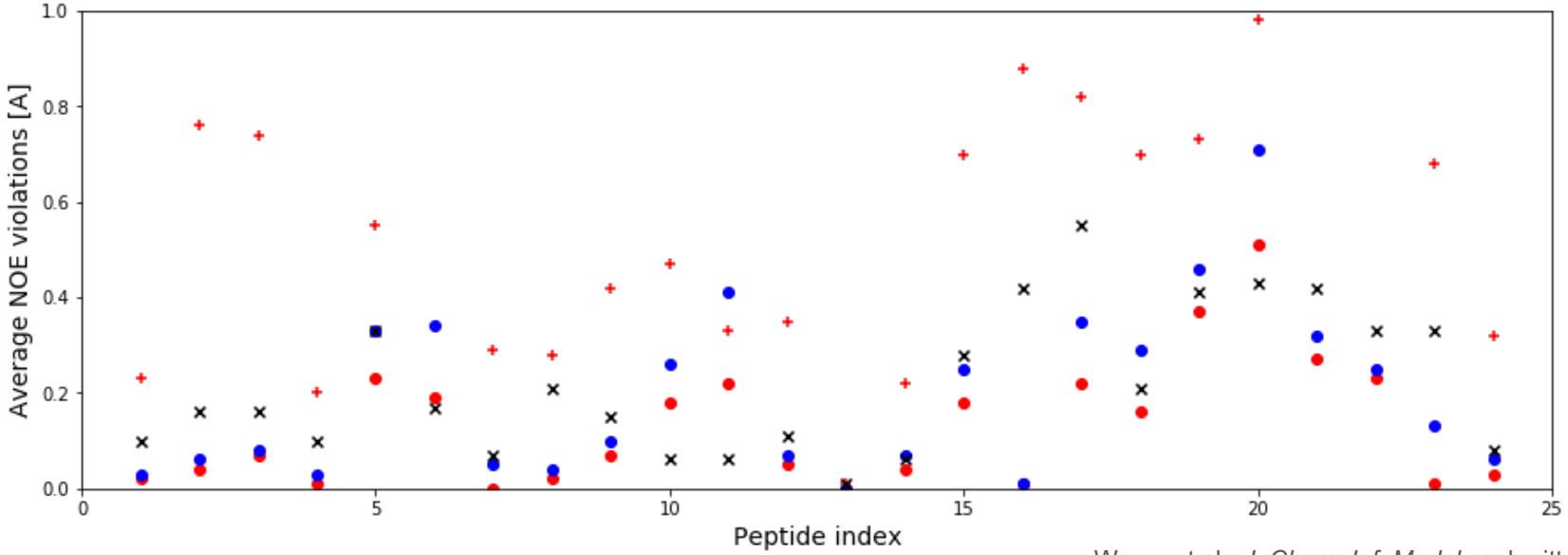
Can we use NOE-derived distance bounds to direct conformational search?

Goal: Conformer bundles fulfilling experimental restraints

Comparison of bundles:

Average NOE violations

+ ETKDG / LICUV
o noeETKDG / LICUV
o noeETKDG / NAMFIS
x reference bundle



Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

ETKDG + NOE

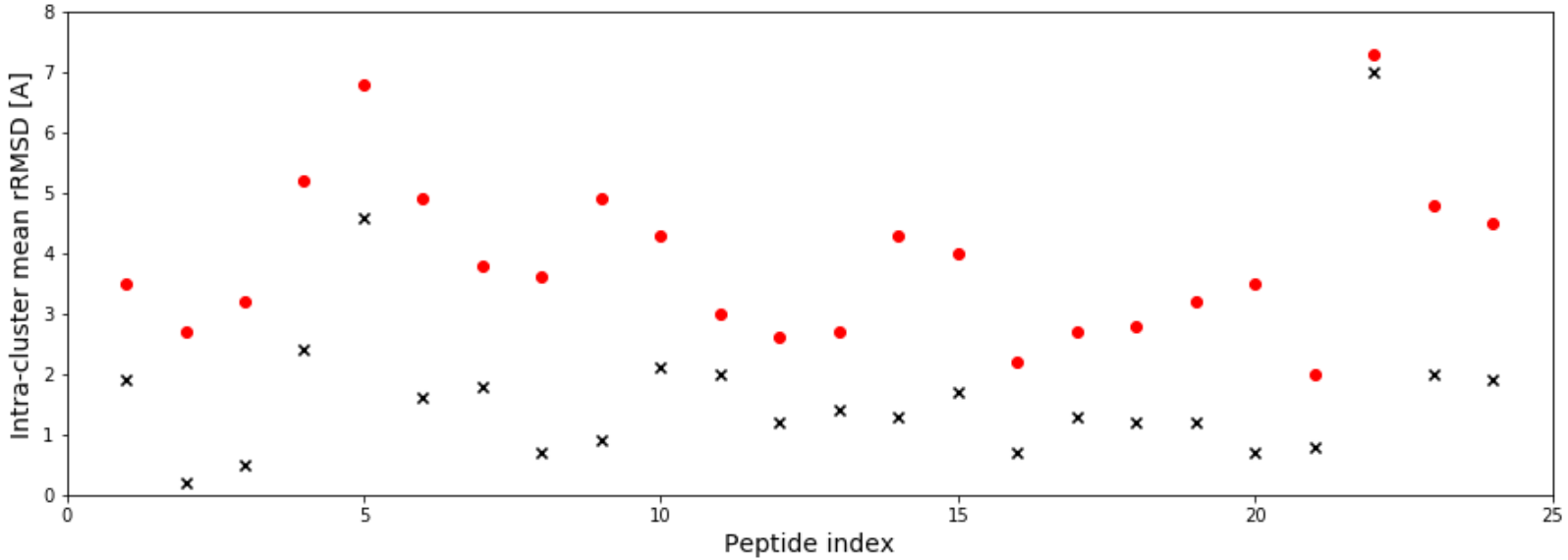
Can we use NOE-derived distance bounds to direct conformational search?

Goal: Conformer bundles fulfilling experimental restraints

Comparison of bundles:

Intra-bundle macrocycle diversity (average rRMSD)

o noeETKDG / LICUV
x reference bundle



Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

ETKDG + NOE

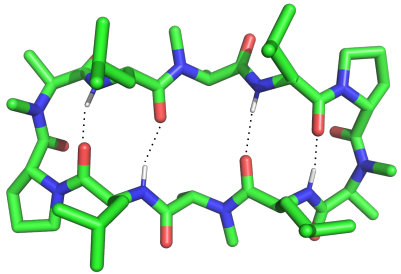
Can we use NOE-derived distance bounds to direct conformational search?

Goal: Conformer bundles fulfilling experimental restraints

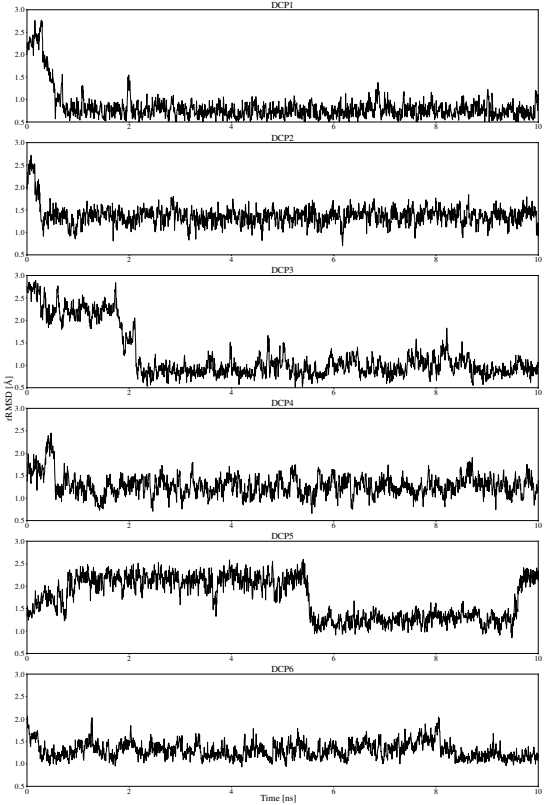
Post-processing with time-averaged NOE-restraining MD:

Use noeETKDG conformers as good starting points

Example: Six cyclic decapeptides with 4 transannular hydrogen bonds



[A]	noeETKDG LICUV		Restraint MD Low-energy frames		Ref. bundle
	Mean NOE viol.	Intra-bundle rRMSD	Mean NOE viol.	Intra-bundle rRMSD	Intra-bundle rRMSD
DCP1	0.2	4.0	0.4	1.5	1.7
DCP2	0.01	2.2	1.1	1.3	0.7
DCP3	0.2	2.7	0.5	1.7	1.3
DCP4	0.2	2.8	0.3	1.5	1.2
DCP5	0.4	3.2	0.6	2.1	1.2
DCP6	0.5	3.5	0.6	1.7	0.7



Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

ETKDG + NOE

Availability

External Python package using RDKit to modify conformer generator

customETKDG on Github:

https://github.com/rinikerlab/customETKDG

Wang, et al., *J. Chem. Inf. Model.*, submitted (2021).

Example Jupyter notebook:

https://github.com/rinikerlab/customETKDG/blob/api/examples/demo_rdkit_ugm_2021.ipynb

Conclusions

Macrocyclic ETKDG (version3)

- Cyclic bonds behave as acyclic ones
- Tools for restricting the conformational search space:
  - Eccentricity constraints
  - Custom pairwise Coulombic interactions (CPCIs)

ETKDG + NOE-derived distance restraints

- Modification of upper distance bounds in ETKDG
- noeETKDG conformers more similar to structures in reference bundles (compared to ETKDG)
- Bundle selection: least NOE violations appears to be a good strategy
- noeETKD bundles contain more diverse ring structures (compared to reference bundles)
- Post-processing with time-averaged NOE-restraining MD simulations