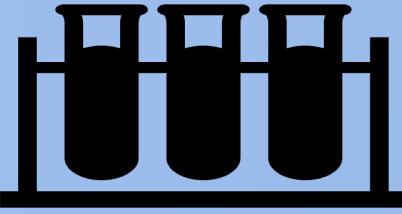


Pipeline to identify stable conformers of APIs

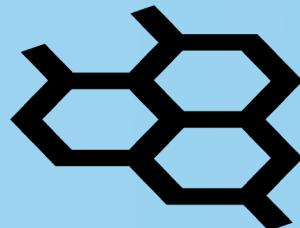
A case study with Rotigotine,
Diflorasone diacetate &
Ritonavir

Context: Why low energy conformers of SMOLs are of interest?



Property modelling

- The physical properties of a molecule, such as melting point and solubility, are often dependent on the stability of its conformers.



Drug discovery

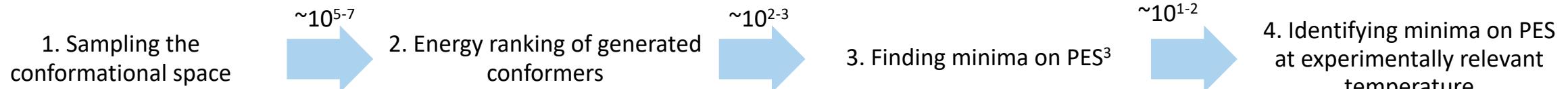
- 3D structure of SMOLs influence their interaction with target biomolecules such as protein, RNA, etc.
- Knowing which molecular conformations are most likely to interact with a specific receptor helps in designing drugs with improved efficiency and reduced side-effects e.g., unwanted off-target binding



ADMET profiling

- Conformers with lower energy are more likely to be present in solution, and thus are more likely to participate in chemical reactions.
- Solid state properties such as tabletability also depend on the ensemble of stable conformers

Limitations of conventional methods in identifying stable molecular conformations



Challenges for *in silico* methods

- Cheminformatics methods¹ have limitations for molecules with large #rotatable bonds
- Challenging to obtain reliable FF parameters for unorthodox dihedrals
- Recent AI based methods² provide only partial solution

- Accurate *in silico* methods like CCSD(T)³ or FCI⁴ are computationally expensive
- Applying high accuracy methods, e.g., DFT on $\sim 10^6$ samples, requires formidable computational expenditure
- There is no single method that is appropriate for all systems

- DFT provides the best compromise between computational cost and accuracy
- But, formidably expensive to perform DFT based geometry optimizations of large number of low energy conformers ($\sim 100s$)

- Involves computationally expensive Hessian calculation

Proposed alternatives

Sampling using *generative AI* based options

Leveraging *DFT-B* for fast, yet accurate energy ranking of APIs
(Supporting literature - Next slide)

Cluster conformers with high geometrical similarity and identify representative structures from each cluster

Inevitable, but GPU based acceleration is possible

¹ [Balloon GA](#), [ETKDG](#), [Frog2](#), [MC-Dock](#), [OMEGA](#), [Confab](#); [GeoDiff](#), [EDM](#), [ConfGF](#) and [ConfVAE](#);

² CCSD(T) = Coupled Cluster with Single, Double, and Perturbative Triple excitations;

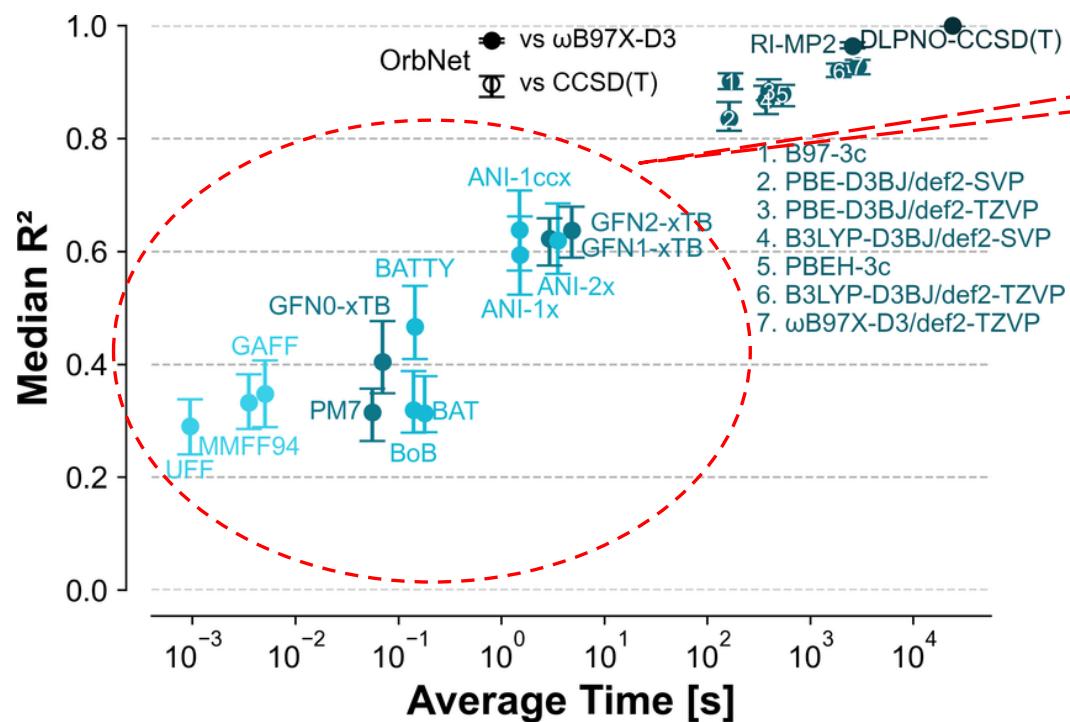
³ FCI = Full Configuration Interaction;

⁴ PES = Potential Energy Scan

Literature review suggesting DFT-B as a fair compromise between computational cost & accuracy for energy ranking

Computational cost vs. accuracy of methods

Hutchison conformer benchmark set ([GitHub](#)).



Faster methods provide poor estimate of the total energy of the molecule.

- However, good spearman rank correlation coefficient (>0.7) are reported for DFT-B => can be used for ranking



FULL PAPER | Open Access | CC BY

Assessing conformer energies using electronic structure and machine learning methods

Dakota Folmsbee, Geoffrey Hutchison

First published: 09 July 2020 | <https://doi.org/10.1002/qua.26381> | Citations: 27

Source: OrbNet Denali @ [J. Chem. Phys. 155, 204103 \(2021\)](#)

Proposed pipeline for searching stable conformers of SMOLs

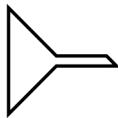
A. Conformer generation and screening

- Torsional diffusion⁰: Leveraging Gen AI
- Screen out conformers with chemically infeasible geometries¹
- 7-8 hrs for 100K conformers²



B. Preliminary energy screening

- DFT-B: Good compromise between computational cost vs. accuracy (Prev. slide)
- 1 – 1.5 days for 100K conformers³
- Select N lowest energy conformers (N=1K in this study)



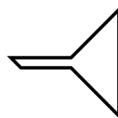
C. Clustering of low energy conformers

- Based on pairwise RMSD
- Optimal #clusters decided based on elbow plot
- ~ 1 hr³



E. Incorporate temperature effects

- Hessian calculation at room temperature
- ~1 day⁴



D. Geometry optimization of the cluster centroids

- Using DFT with appropriate functional and basis set
- ~1 day⁴
- Select conformers with energy < ΔE_{cutoff}
- $\Delta E_{\text{cutoff}} = 20 \text{ kJ mol}^{-1}$ for this study
- ΔE_{cutoff} based on the energy distribution of the geometry optimized cluster centroids

0. <https://arxiv.org/abs/2206.01729>

1. Protocol to identify conformers with chemically infeasible geometries: Convert the generated conformer into SMILES, Calculate Tanimoto Similarity between the original SMILES & converted SMILES. If the value is ≠ 1, remove conformer

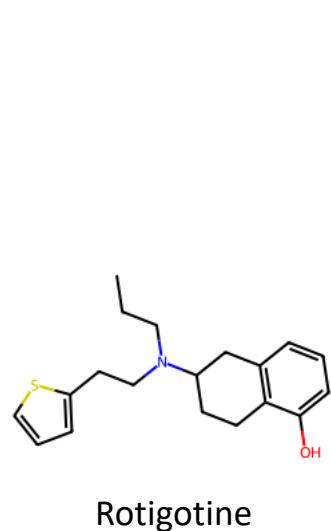
2. Performed on GPU with 11 GB memory

3. Performed on Mac M1 with 8 GB memory

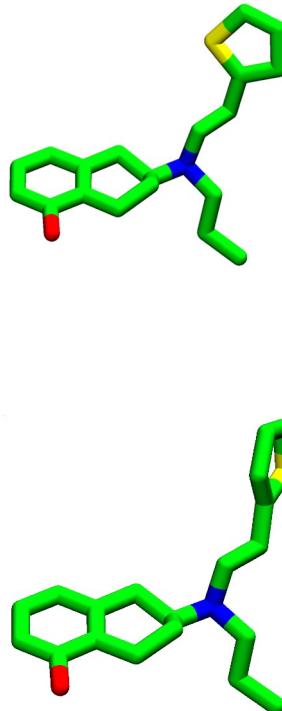
4. Performed on GPU with 24 GB memory

Case study: APIs showing conformational polymorphism leading to different physicochemical properties

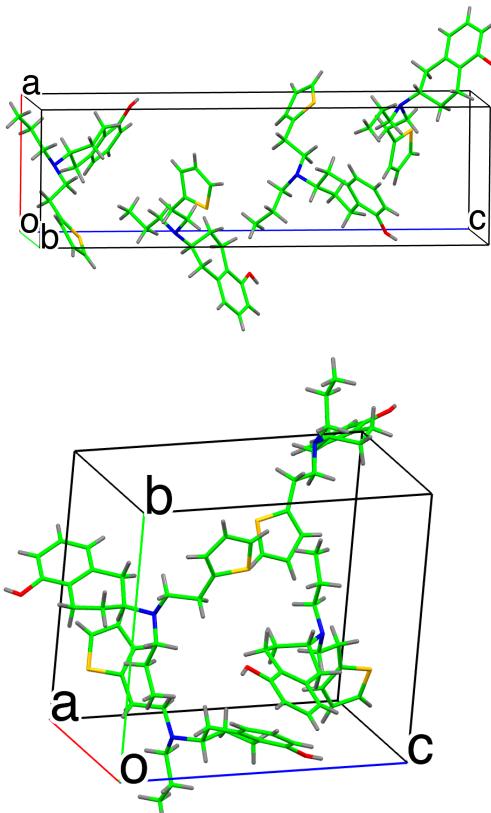
Molecule



Conformations



Crystal structure



APIs showing conformational polymorphism

| | API | MW (g/mol) | #rot. bonds |
|----|---|------------|-------------|
| 1 | Tolbutamide | 270.35 | 5 |
| 2 | Chlorpropamide | 276.74 | 4 |
| 3 | Rotigotine | 315.5 | 6 |
| 4 | N-(4'-methoxyphenyl)-3-bromothiobenzamide | 322.22 | 3 |
| 5 | Acitretin | 326.4 | 6 |
| 6 | Furosemide | 330.74 | 5 |
| 7 | Axitinib | 386.5 | 5 |
| 8 | Aripiprazole | 448.4 | 7 |
| 9 | Diflorasone diacetate | 494.5 | 6 |
| 10 | Ritonavir | 720.9 | 18 |

Chosen for the case study

CHEMICAL
REVIEWS

Review

pubs.acs.org/CR

Conformational Polymorphism

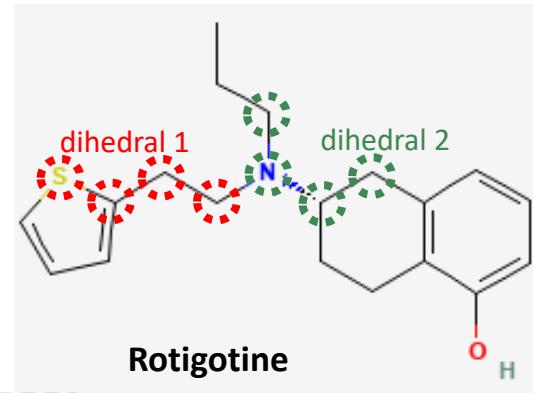
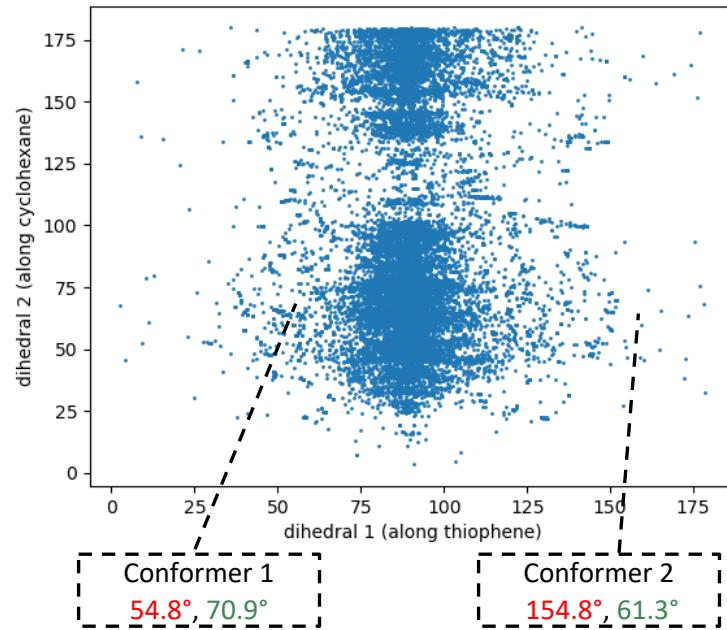
Aurora J. Cruz-Cabeza*† and Joel Bernstein‡§

*Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

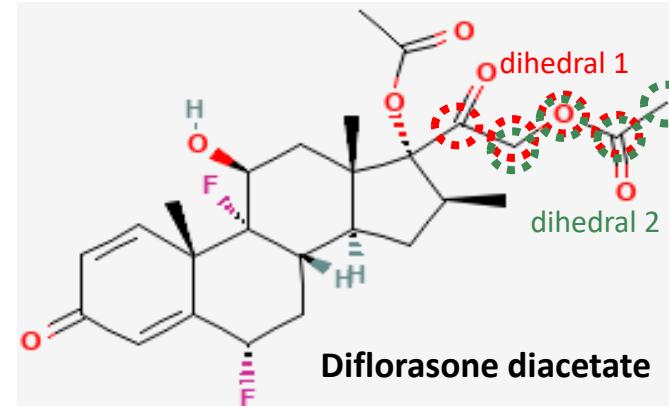
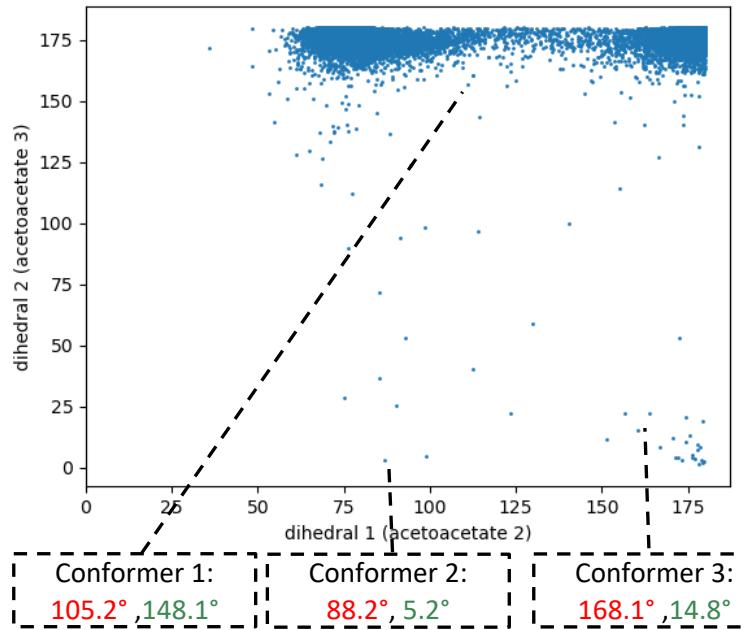
†Faculty of Natural Sciences, New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, United Arab Emirates

‡Department of Chemistry, Ben-Gurion University of the Negev, P.O. Box 653, Beer Sheva, Israel 84120

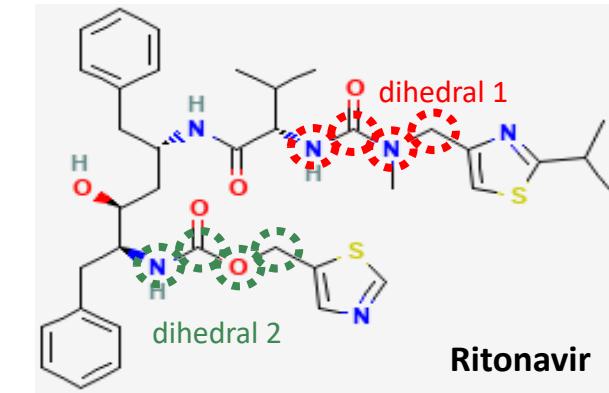
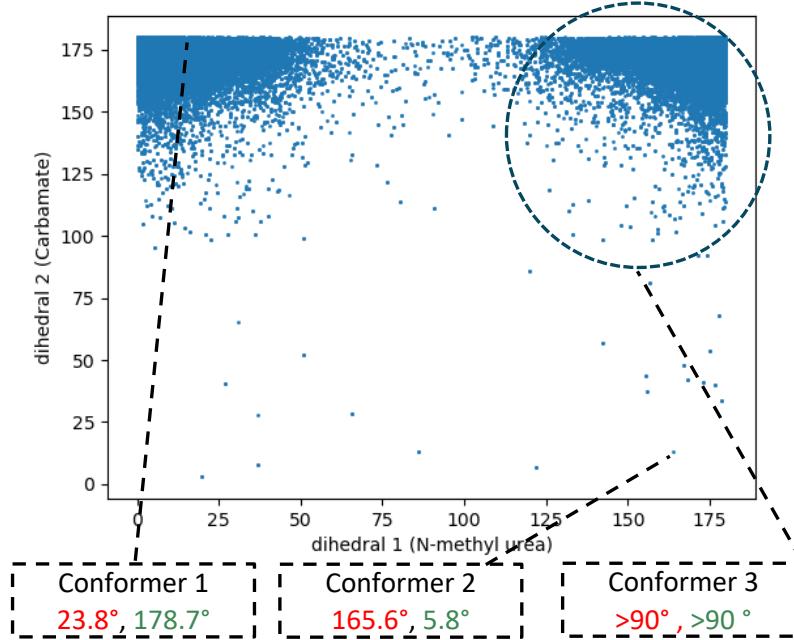
Generative AI helps sample (100K) the conformation space well, including conformers in known crystal structures



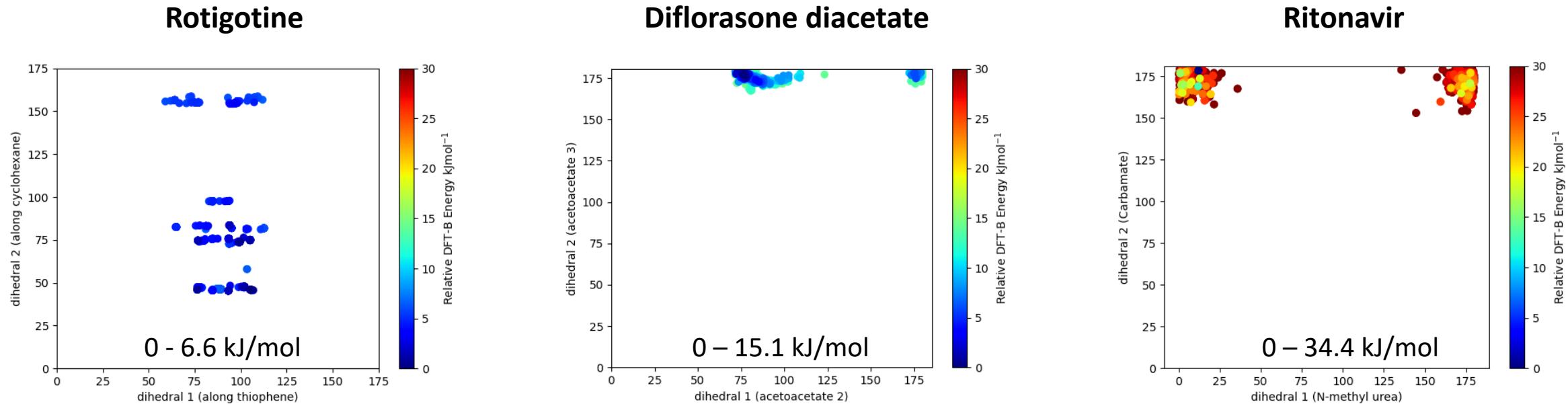
Conformers found in
known crystal structures



Confidential and Proprietary. Copyright © 2017-24 Aganitha



Preliminary energy screening: Top-N conformers to be chosen based on the molecule complexity (N=1000 for this study)

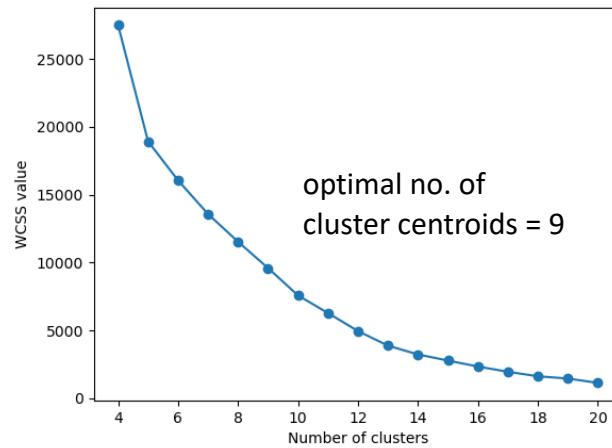


ΔE 1000th conformer: DFT-B energy (kJ/mol) of the 1000th conformer relative to that of the most stable conformer

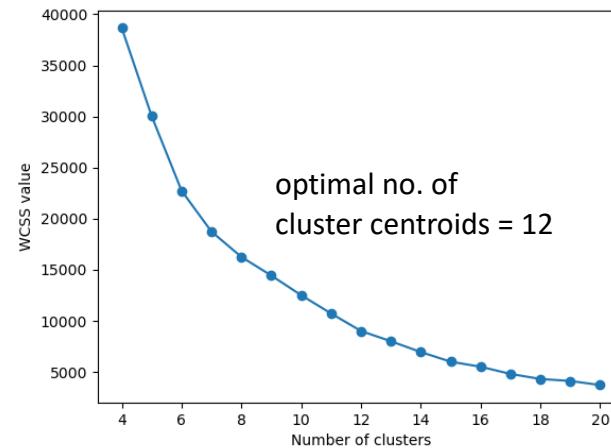
In this case, ΔE (in kJ/mol) 1000th conformer of Ritonavir (34.4) > Diflorasone diacetate (15.1) > Rotigotine (6.6)

Clustering of low energy conformers to identify geometrically diverse low energy conformers

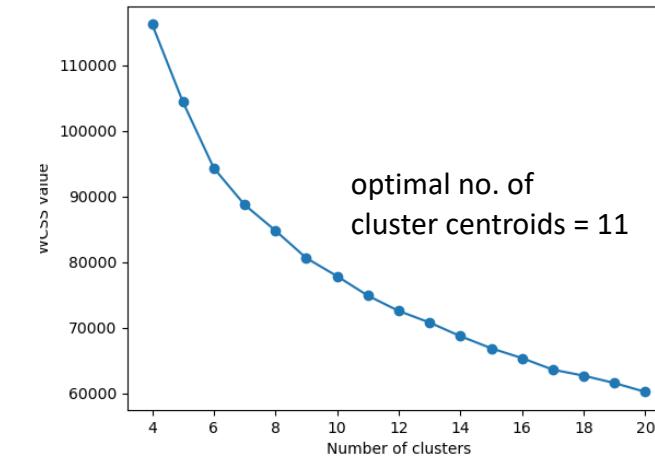
Rotigotine



Diflorasone diacetate



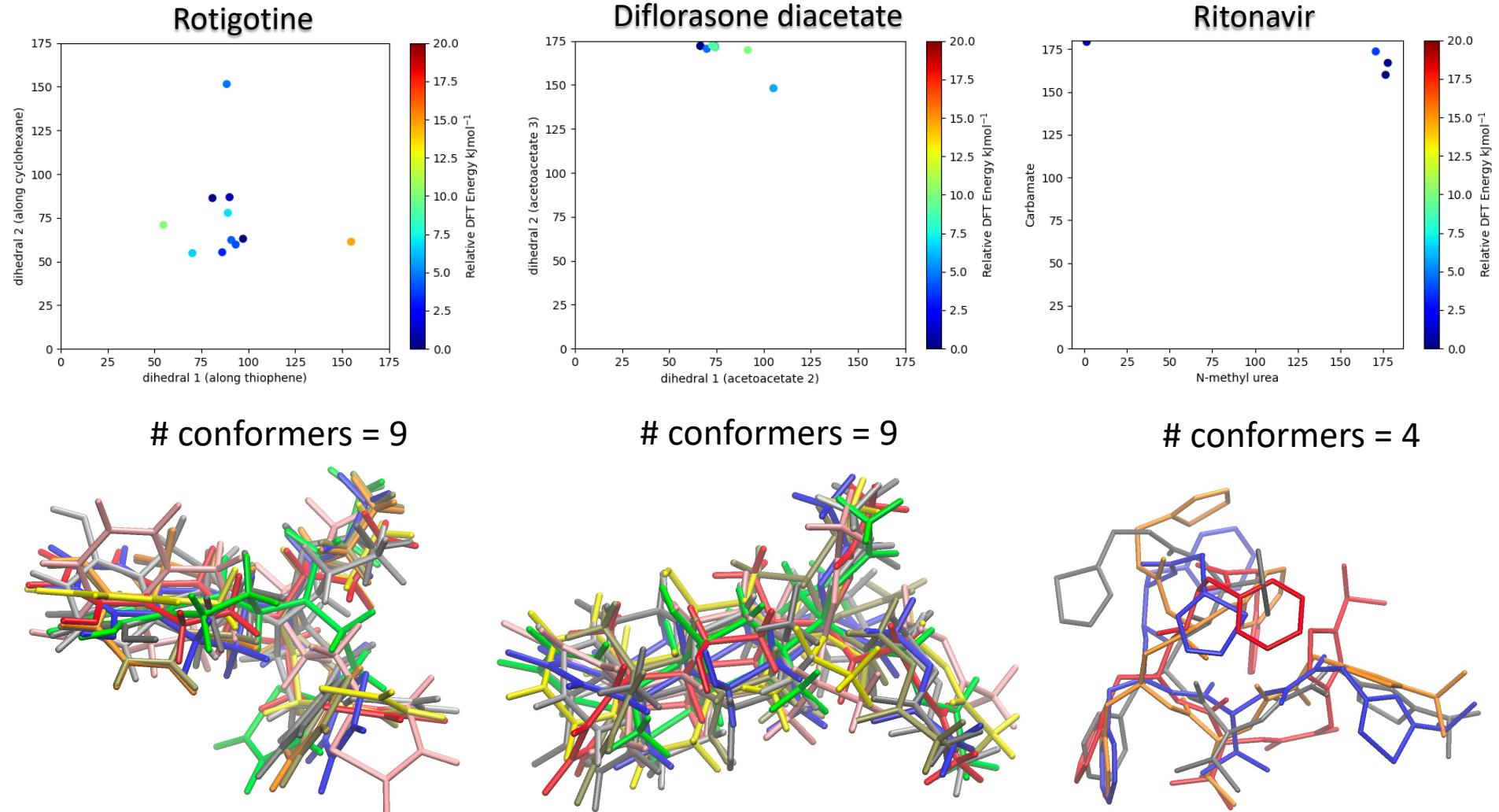
Ritonavir



- K-Means clustering of the Top-1000 lowest energy conformers based on pairwise RMSD.
- RMSD is calculated by aligning two sets of atomic coordinates using the [Kabsch algorithm](#)
- WCSS = Within Cluster Sum of Squares of RMSD w. r. t. centroid
- Optimal number of cluster centroids decided based on the rate of change of slope w. r. t. increase in number of clusters.

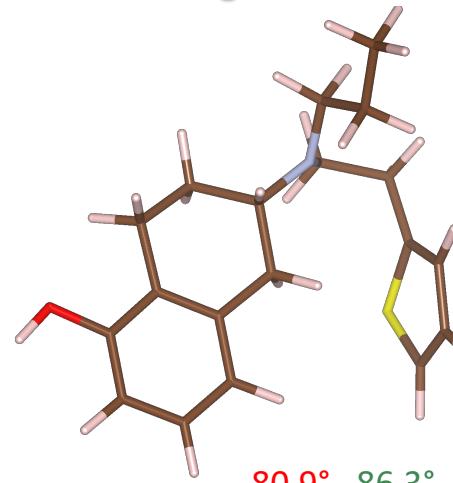
DFT optimization of cluster centroids reveal unique local minima

- Unique low energy conformers identified with $\Delta E < 20 \text{ kJ mol}^{-1}$ for this study
- Represent the local minima on the PES at $T = 0 \text{ K}$
- DFT calculations were carried out at B3LYP/6-31G level of theory



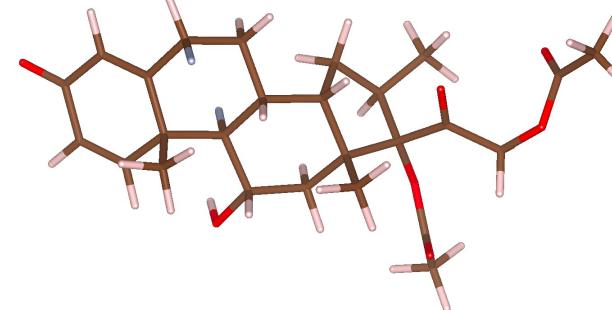
Temperature correction on the local minima to identify the most stable conformer (in vacuum) at room temperature¹

Rotigotine



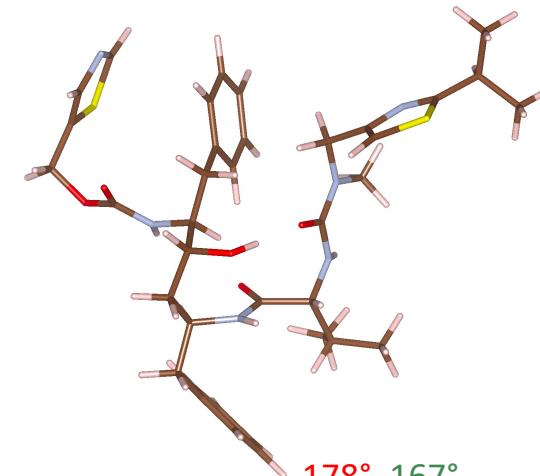
80.9° , 86.3°

Diflorasone diacetate



74.0° , 172.1°

Ritonavir



178° , 167°

Comparison with the in vacuum optimized geometries of the conformers found in known crystal structures

| Expt. known conformer | ΔE_{298} (expt.) in kJ/mol ² | #local minima with $\Delta E_{298} <$ ΔE_{298} (expt.) |
|-----------------------------|--|--|
| 1 | -6.9 | 0/9 |
| 2 | +5.6 | 4/9 |

| Expt. known conformer | ΔE_{298} (expt.) in kJ/mol ² | #local minima with $\Delta E_{298} <$ ΔE_{298} (expt.) |
|-----------------------------|--|--|
| 1 | +1.2 | 6/9 |
| 2 | +66.4 | 9/9 |
| 3 | +93.0 | 9/9 |

| Expt. known conformer | ΔE_{298} (expt.) in kJ/mol ² | #local minima with $\Delta E_{298} <$ ΔE_{298} (expt.) |
|-----------------------------|--|--|
| 1 | +26.0 | 4/4 |
| 2 | +60.5 | 4/4 |
| 3 | NA ³ | - |

1. ΔE_{298} = Relative enthalpy from Hessian calculations at T = 298 K. Some of the conformers have one imaginary frequencies that is $> 100 \text{ cm}^{-1}$.

2. ΔE_{298} (expt.) = Thermally corrected enthalpy_{Most stable conformer identified Insilico} - Thermally corrected enthalpy_{DFT optimized geometry (vacuum) of the conformers in known crystal structures}

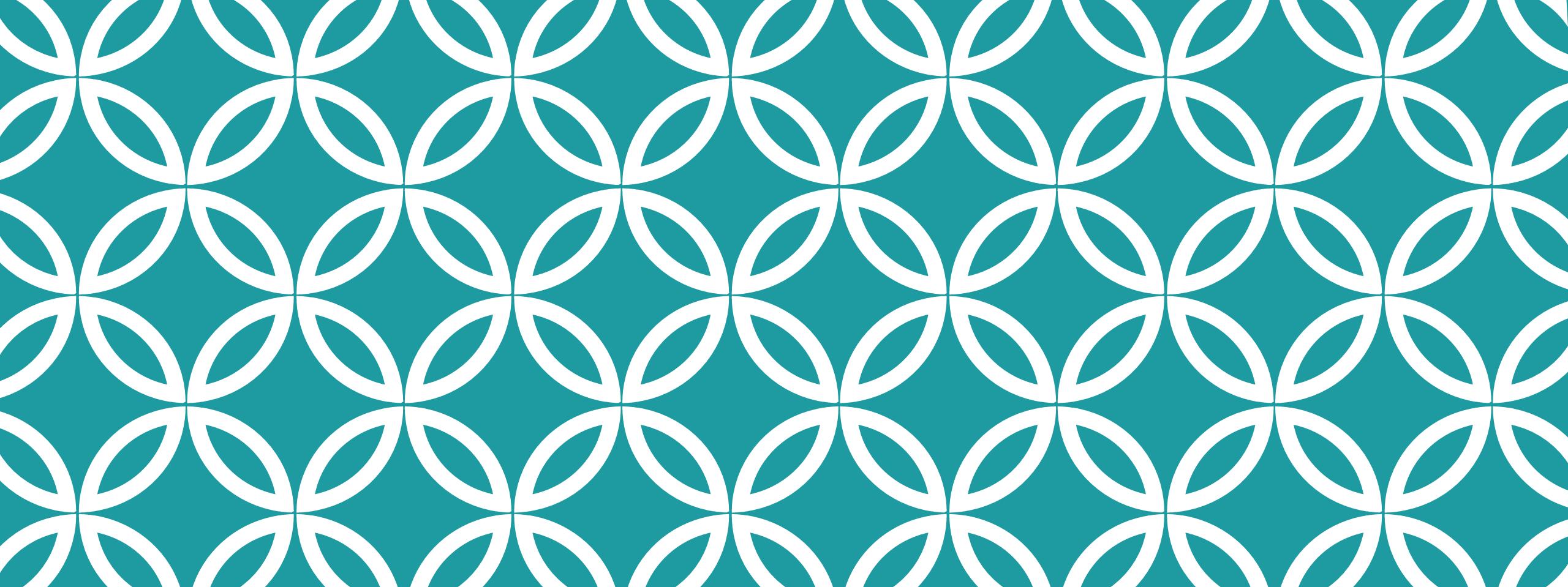
3. Crystal structure corresponding to conformer 3 is not available

Pipeline identified stable conformers; In some cases, those are more stable than the known conformers

| APIs that display conformational polymorphism | {SC} = Set of <i>in silico</i> identified stable conformers | Conformers from known polymorphs present in {SC} † | Stability of the conformers in {SC} w.r.t. gas phase optimized geometry of conformers found in known polymorphs † |
|---|---|--|--|
| Rotigotine (MW: 315 g mol ⁻¹ , #rot. bond: 6) | 9 | 1/2 | <ul style="list-style-type: none"> • 0/9 are more stable than *conformer 1 • 4/9 are more stable than *conformer 2 |
| Diflorasone diacetate (MW: 494 g mol ⁻¹ , #rot. bond: 6) | 9 | 1/3 | <ul style="list-style-type: none"> • 6/9 are more stable than *conformer 1 • 9/9 are more stable than the *conformer 2 & 3 |
| Ritonavir (MW: 720 g mol ⁻¹ , #rot. bond: 18) | 4 | 2/3 | <ul style="list-style-type: none"> • All the insilico identified conformers are more stable than the known conformers |

† For the purpose of PoC, (a) only 10^5 conformers were generated, (b) only top 1% lowest energy conformers were considered for clustering & (c) the DFT calculations were performed using a relatively small basis set (6-31G) to reduce the computational overhead. Ideally, one should (a) generate 10^{6-7} conformers, (b) select top 10% lowest energy conformers for clustering & (c) perform DFT calculations with larger basis sets augmented with polarization corrections for drug-like molecules

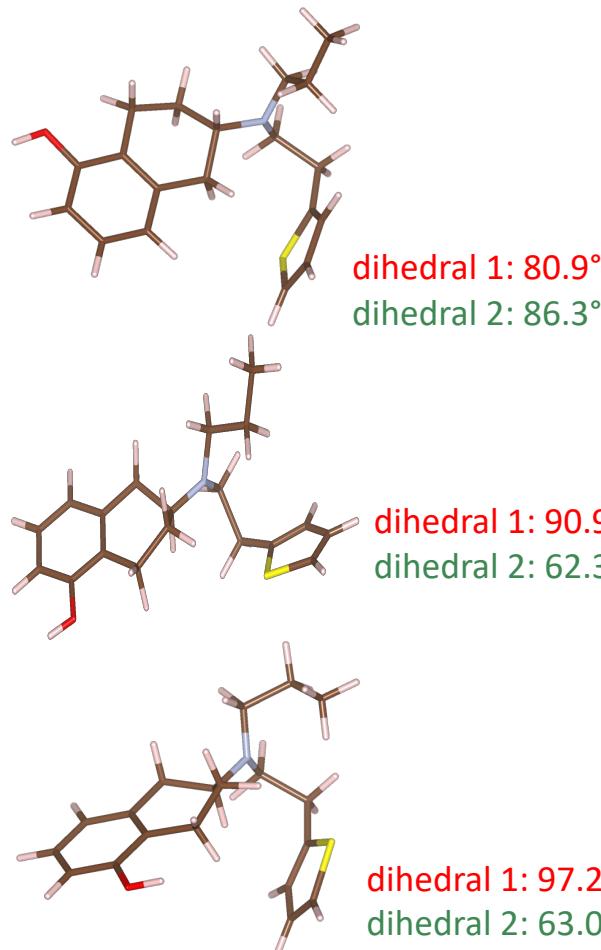
* Conformer 'x' are gas phase optimized geometry of the conformers found in known crystal structures.



Appendix

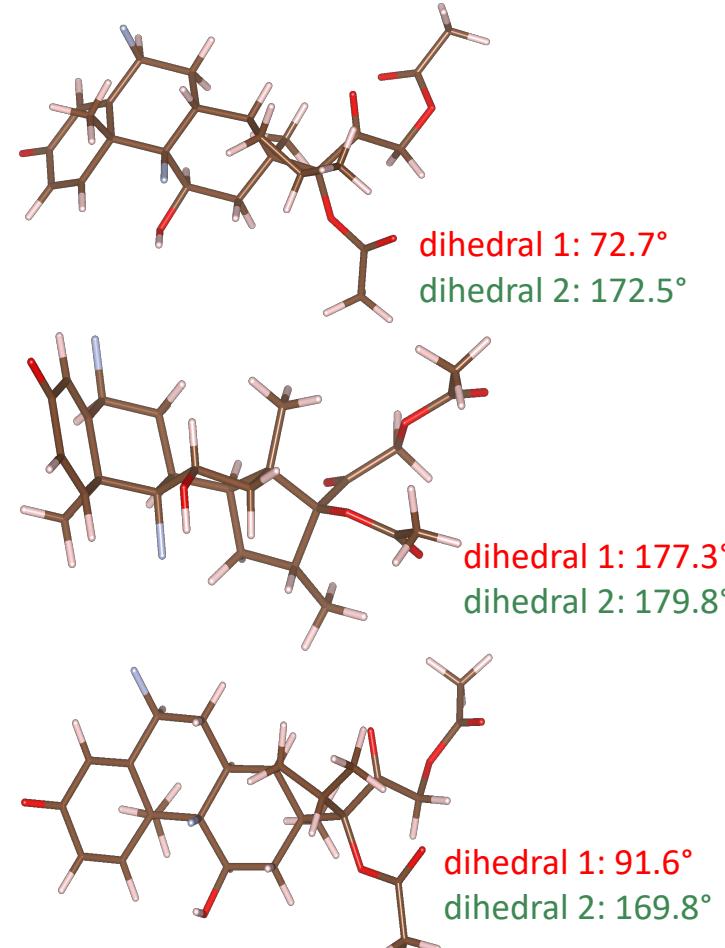
Sample structures of insilico identified stable conformers

Rotigotine (MW: 315.5 g mol⁻¹)

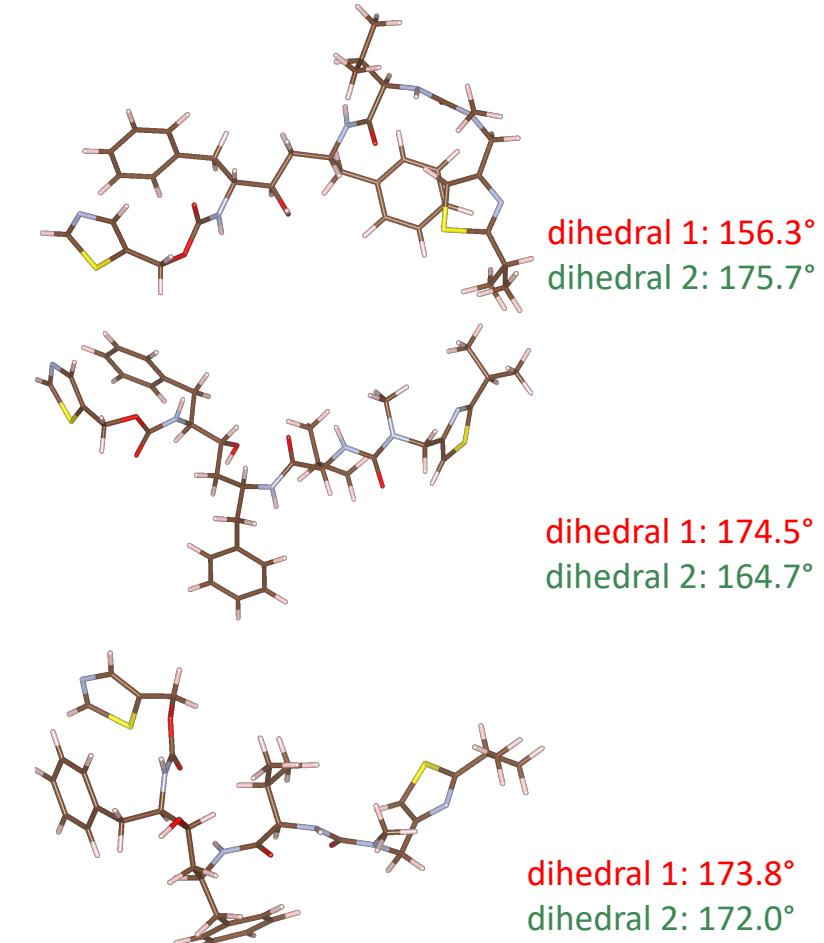


14

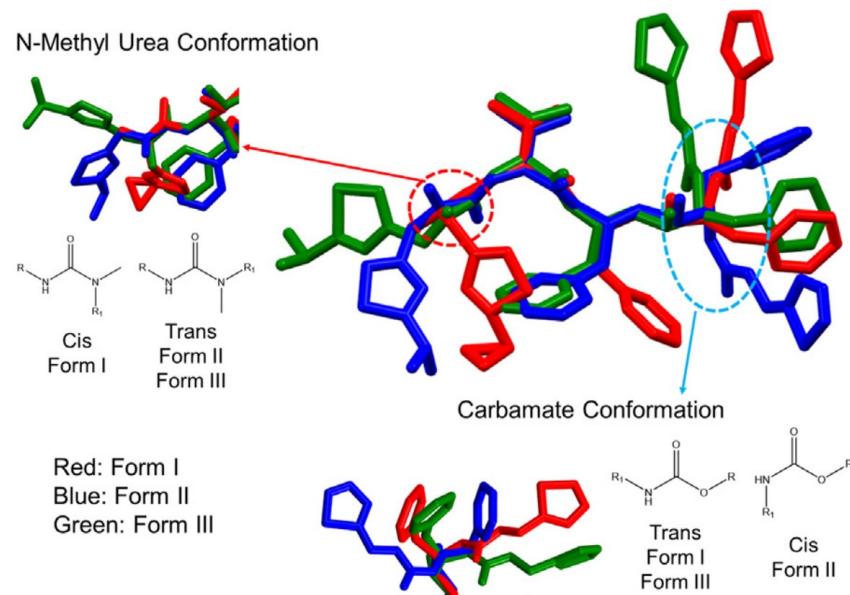
Diflorasone diacetate
(MW: 494.5 g mol⁻¹)



Ritonavir (MW: 720.9 g mol⁻¹)



Increasing the number of generated conformers leads to better sampling of the conformational space



| Number of <i>Ritonavir</i> conformers generated | Number of <i>conformer 2 like</i> structures observed |
|---|---|
| 10K | 3 |
| 20K | 9 |
| 30K | 13 |
| 100K | 45 |

Molecular conformations in Ritonavir polymorphs

Ref: [Zhang et al. Journal of Pharmaceutical Sciences \(2023\)](#)