Using FreeForm : A Better and Faster Estimate of Ligand Strain

Varsha Jain
Application Scientist
05/18/2021

OpenEye SCIENTIFIC

FreeForm: Free Energy of the Unbound Ligand

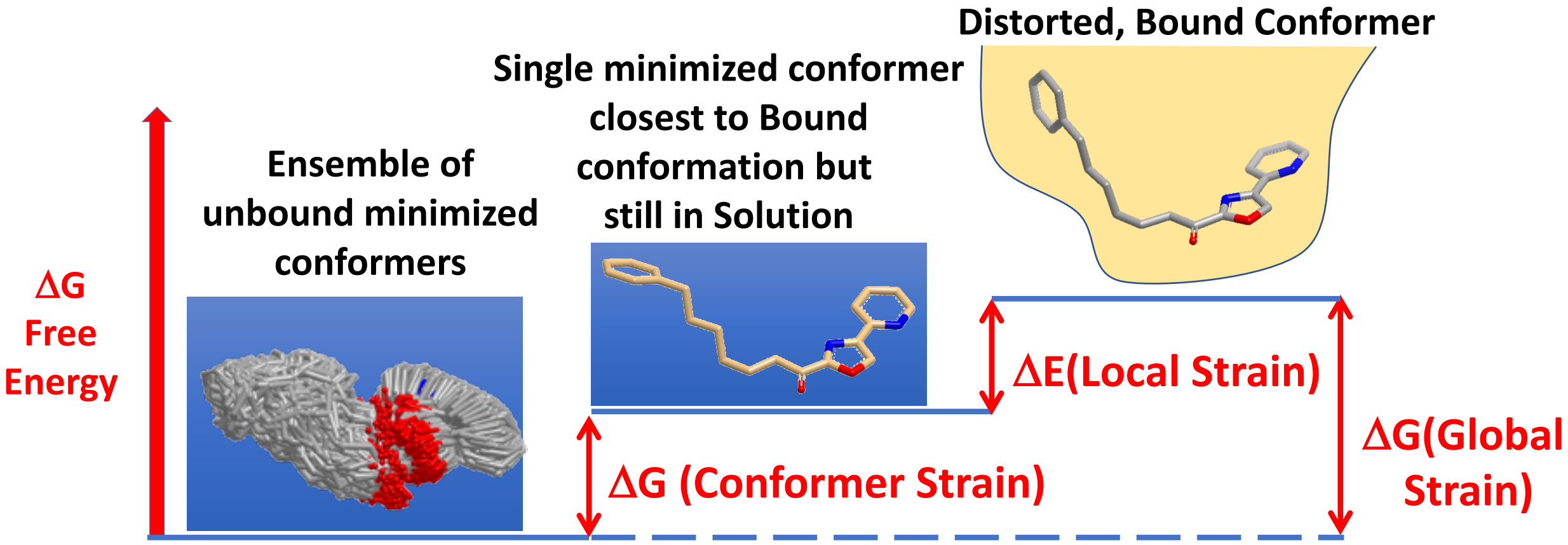
What does it do?

- Estimates the energetic cost of ligand *conformation restriction* upon binding to a protein
- With forcefield, estimates the protein-induced “local” strain

Why would you use it?

- Predicting the effects of constraining analogs in lead optimization
- Selecting low (free) energy conformers for model building

FreeForm calculates *Strain* (Free) Energies



N.B. These are ONLY ligand energies, no protein or protein-ligand energies

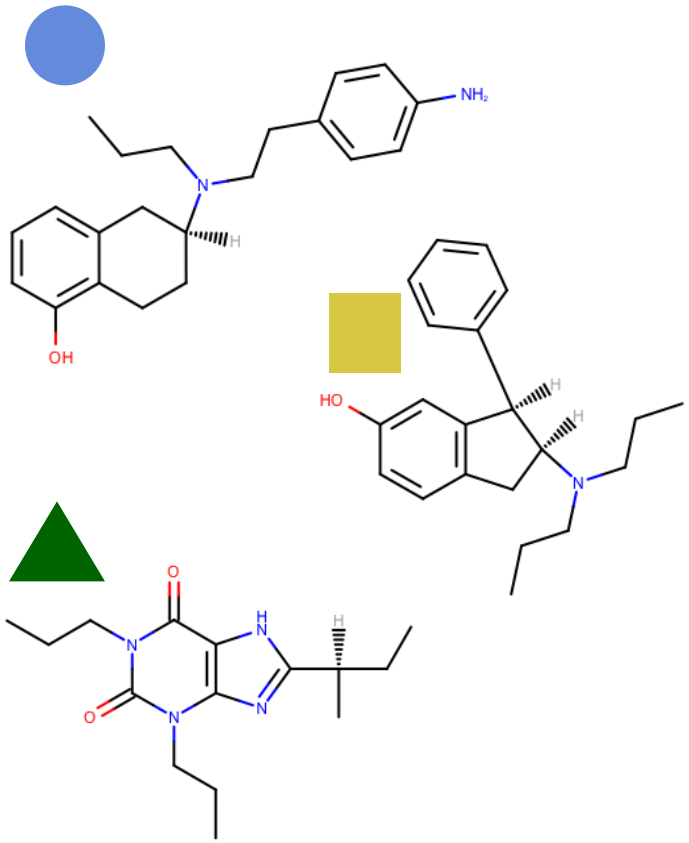
OpenEye SCIENTIFIC

Is Ligand Strain useful? I want Affinity!

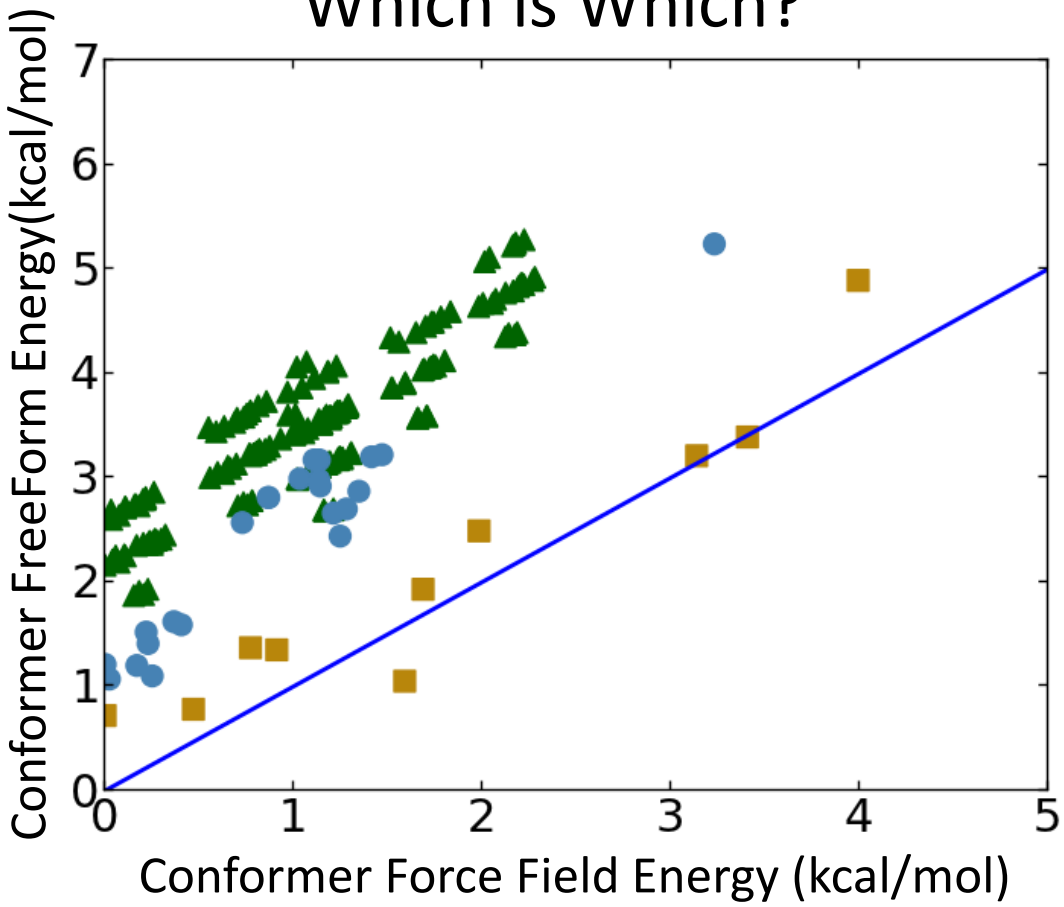
- Affinity (FEP) requires a lot of computation (typically days)
  - Ligand, Protein, Protein-ligand ΔGs required
  - Even then may not be accurate (e.g. large-scale motions on binding)
- **FreeForm** takes minutes and ligand strain may dominate differences in Affinity
  - E.g. Conformational restriction
  - E.g. Changes in protein-induced ligand strain energy
- Complements less rigorous (but faster) affinity estimates (MMPB)
- You may not have the structure
  - Use FreeForm energy in ML, QSAR, i.e. model building

This concept has a long history of being used intuitively -e.g. Rotor Counting

Number of rotatable bonds = 6



Which is Which?



Counting rotors is a poor approximation to Conformer Free Energy

OpenEye SCIENTIFIC

Can we calculate ligand strain *accurately*?

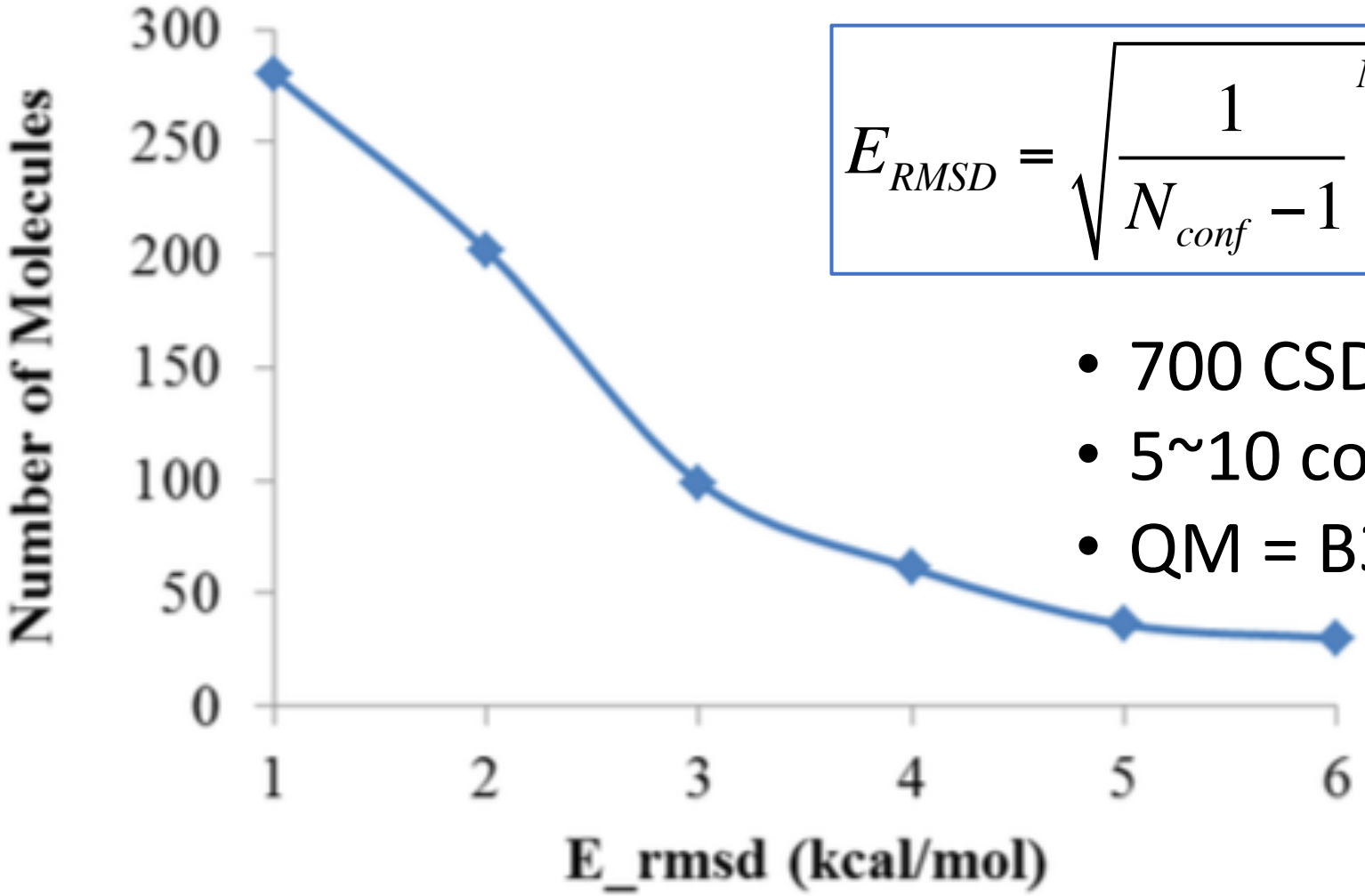
Three necessary components:

1. Internal conformer energy –in solution + active site

2. Solvation of each conformer

3. Vibration/rotational entropy of each conformer

1) Internal Conformer Energies- MMFF94s

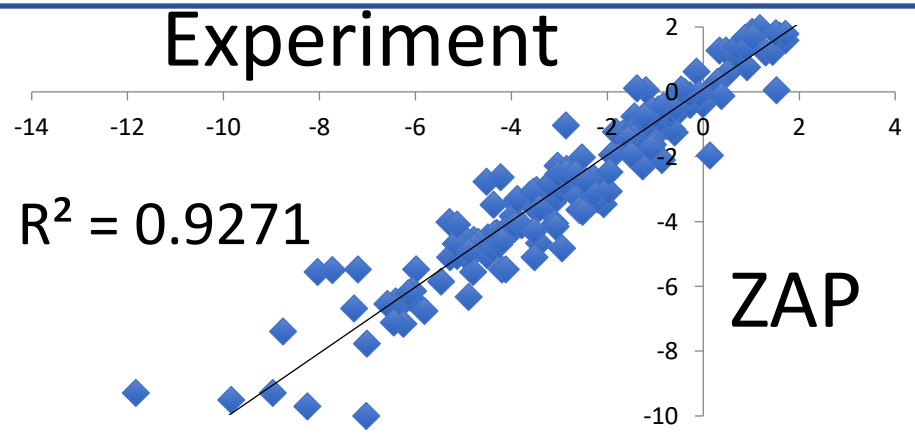


$$E_{RMSD} = \sqrt{\frac{1}{N_{conf} - 1} \sum_{i=2}^{N_{conf}} \left(\Delta E_i^{QM} - \Delta E_i^{MMFF} \right)^2}$$

- 700 CSD molecules
- 5~10 conformers
- QM = B3LYP-D/6-311G**

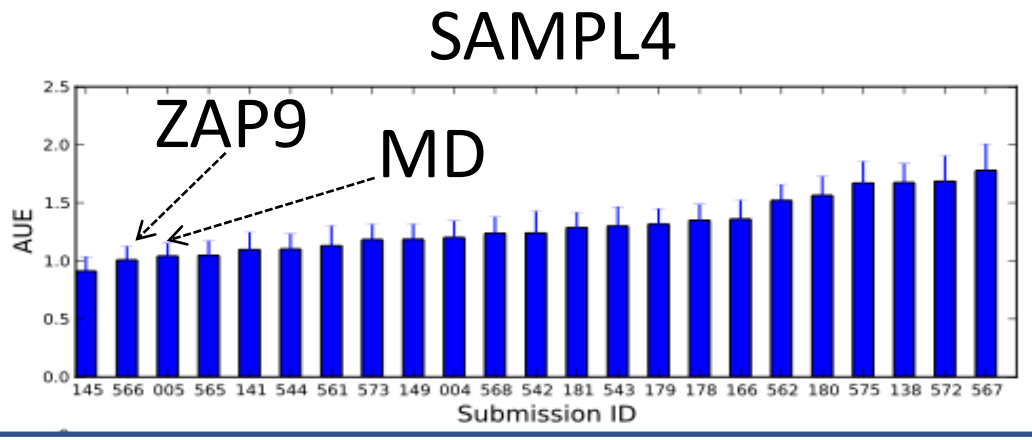
OpenEye SCIENTIFIC

2) Solvation (vacuum to water transfer)



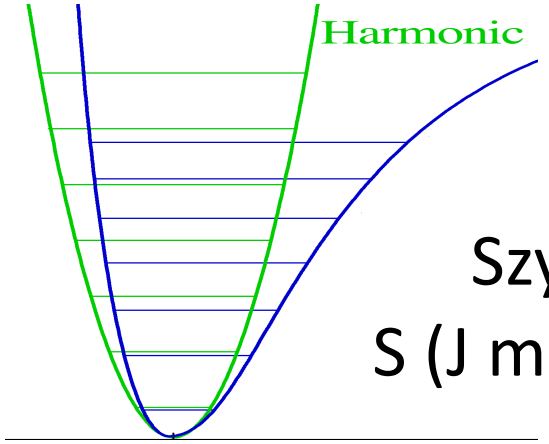
- 200 small molecule solvation ΔG
- ZAP PB Solver
- AM1BCC Charges
- RMS Error = 0.76 kcals/mol

- SAMPL4
  - RMS Error ~ 1.5 kcal/mol
  - <1.0 kcal/mol if >-10 kcal/mol



OpenEye SCIENTIFIC

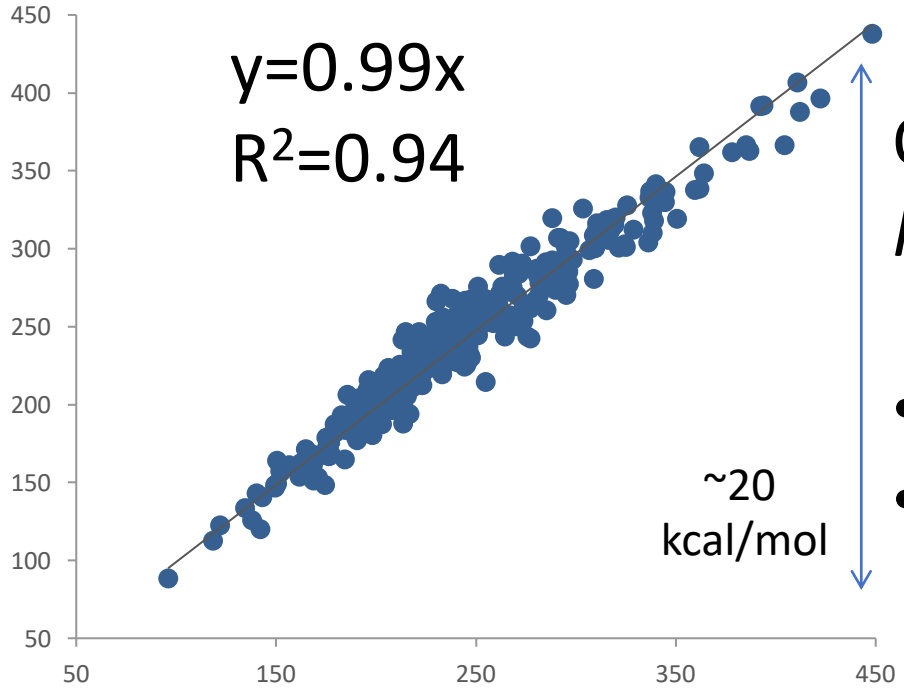
3) Vibrational Entropy



Szybki
S (J mol^{-1} K^{-1})

RMSE~ 1 kcal/mol

Wlodek, Skillman & Nicholls
J. Chem. Theory Comput, 6 (7), p2140



Experimental
S (J mol^{-1} K^{-1})

ORCHYD Dataset
Int. J. Thermo., 25, p352

- Experimental S^{o}
- 290 small organics

OpenEye SCIENTIFIC

FreeForm: Free Energy of the Unbound Ligand

What does it do?

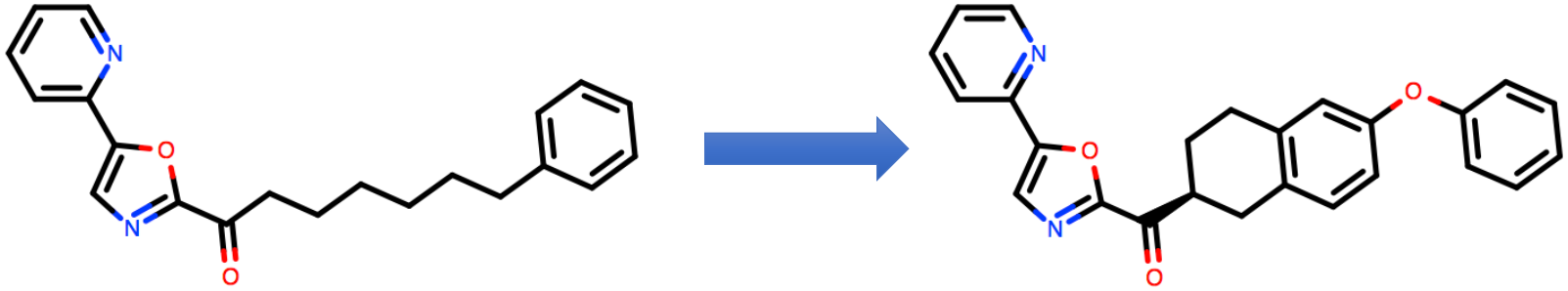
- Estimates the energetic cost of ligand *conformation restriction* upon binding to a protein
- With forcefield, estimates the protein-induced “local” strain

Why would you use it?

- Predicting the effects of constraining analogs in lead optimization
- Selecting low (free) energy conformers for model building

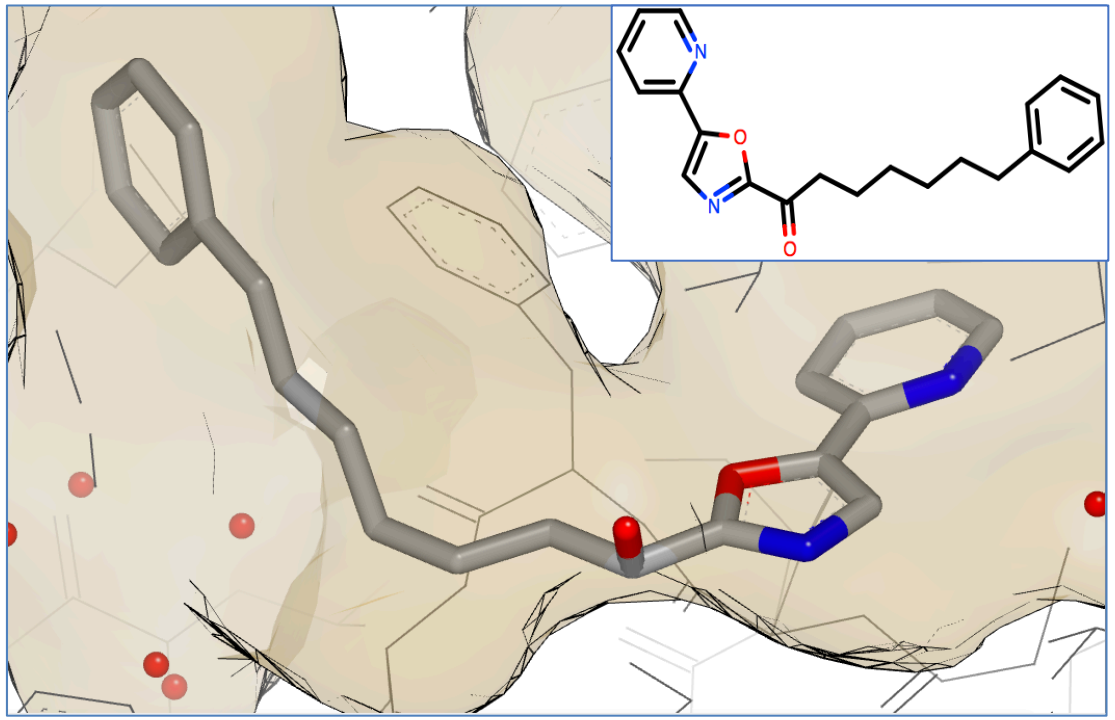
Lead Optimization: Example 1

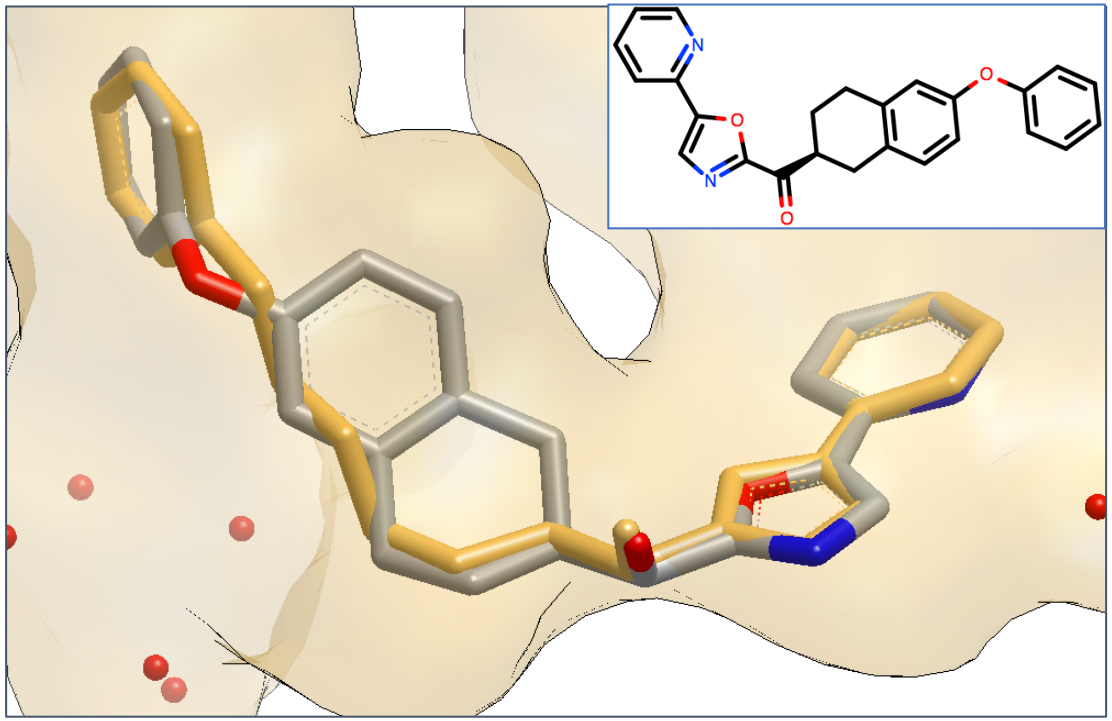
FAAH inhibitors: Ezzili et al., *J Med Chem* **54**, 2805 (2011).



Would constraining the analog increase affinity?

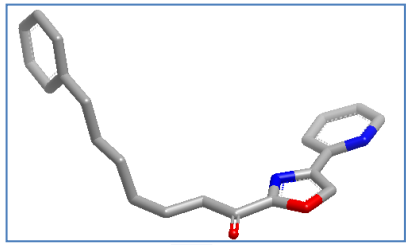
Protein-complex structures available





..therefore, we can also calculate *local strain*

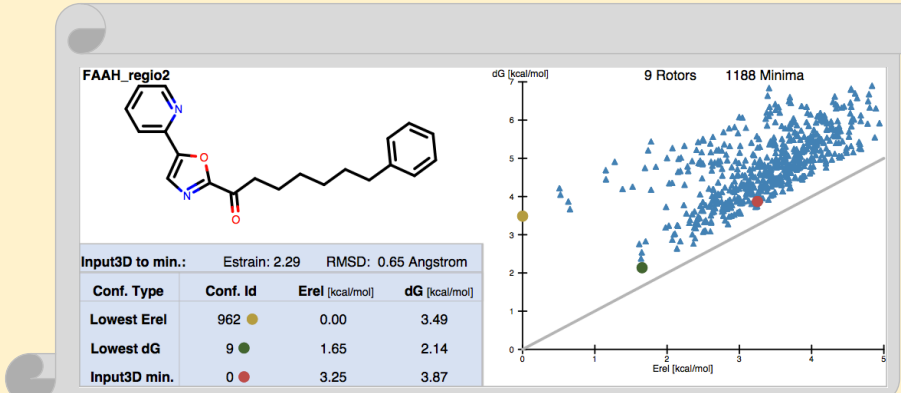
FreeForm: Workflow



FreeForm

Sampling 40000 conformers

Energy 1000 Confs/minute



✓.pdf report

✓.oeb Ensemble of minima

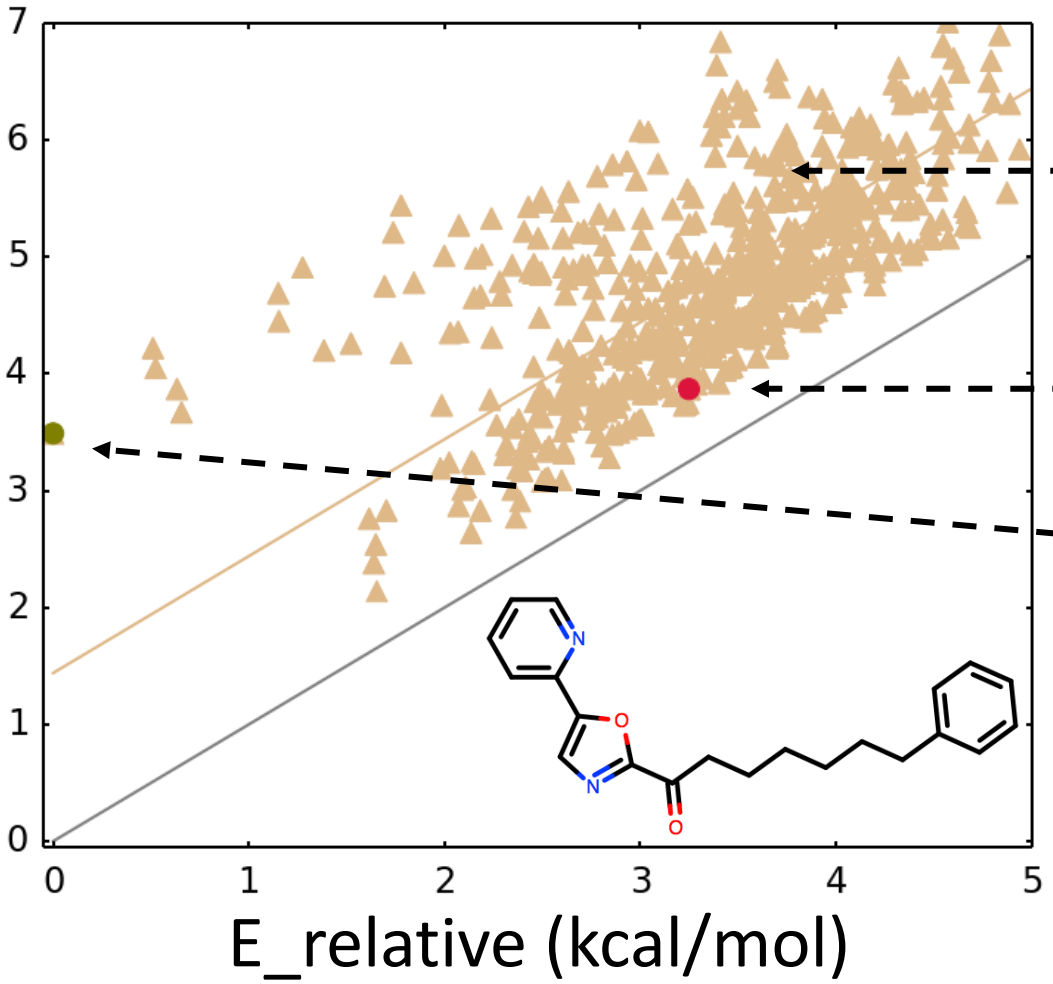
✓.csv key ΔG, ΔH, ΔS quantities

- force field models : MMFF94, MMFF94S, Smirnoff99Frost, PARSLEY_OPENFF1.0.0, PARSLEY_OPENFF1.1.1, PARSLEY_OPENFF1.2.1, PARSLEY_OPENFF1.3.1

OpenEye SCIENTIFIC

FreeForm Free Energy Report

Conformer
ΔG
(kcal/mol)



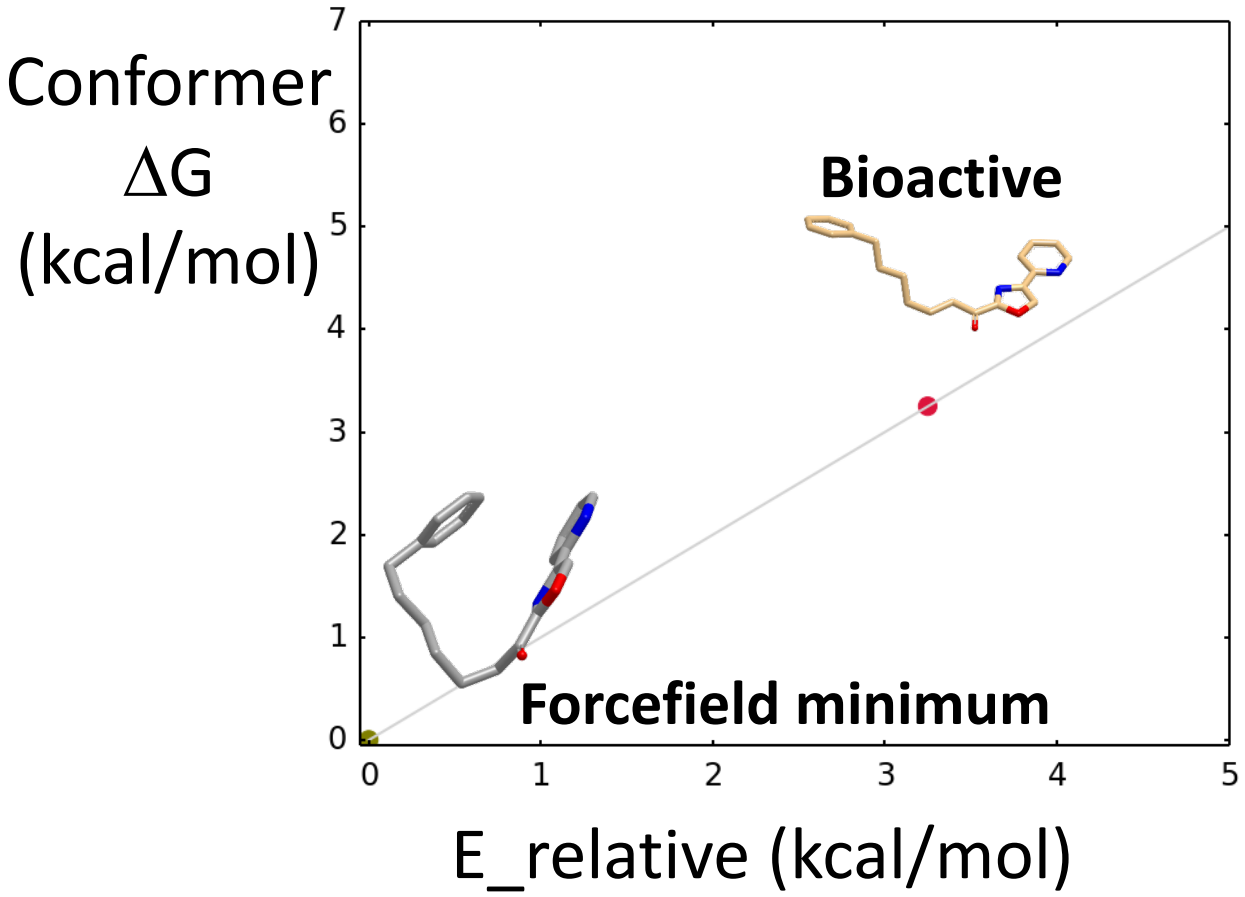
All the conformers from SZYBKI solvent minimization

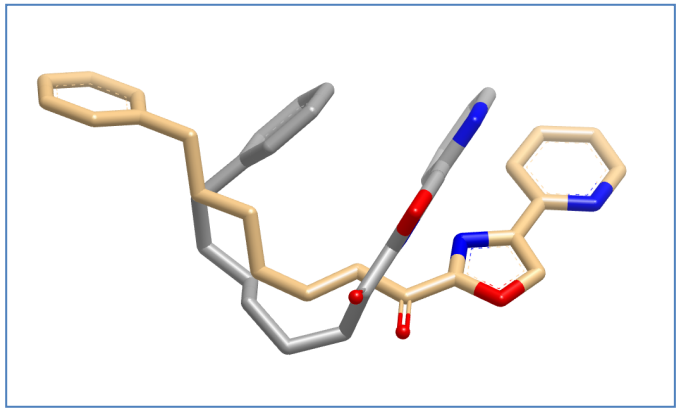
Bioactive conformer

Lowest MMFF94s conformer

OpenEye SCIENTIFIC

The Bioactive and MMFF94 Energy Minima Alone

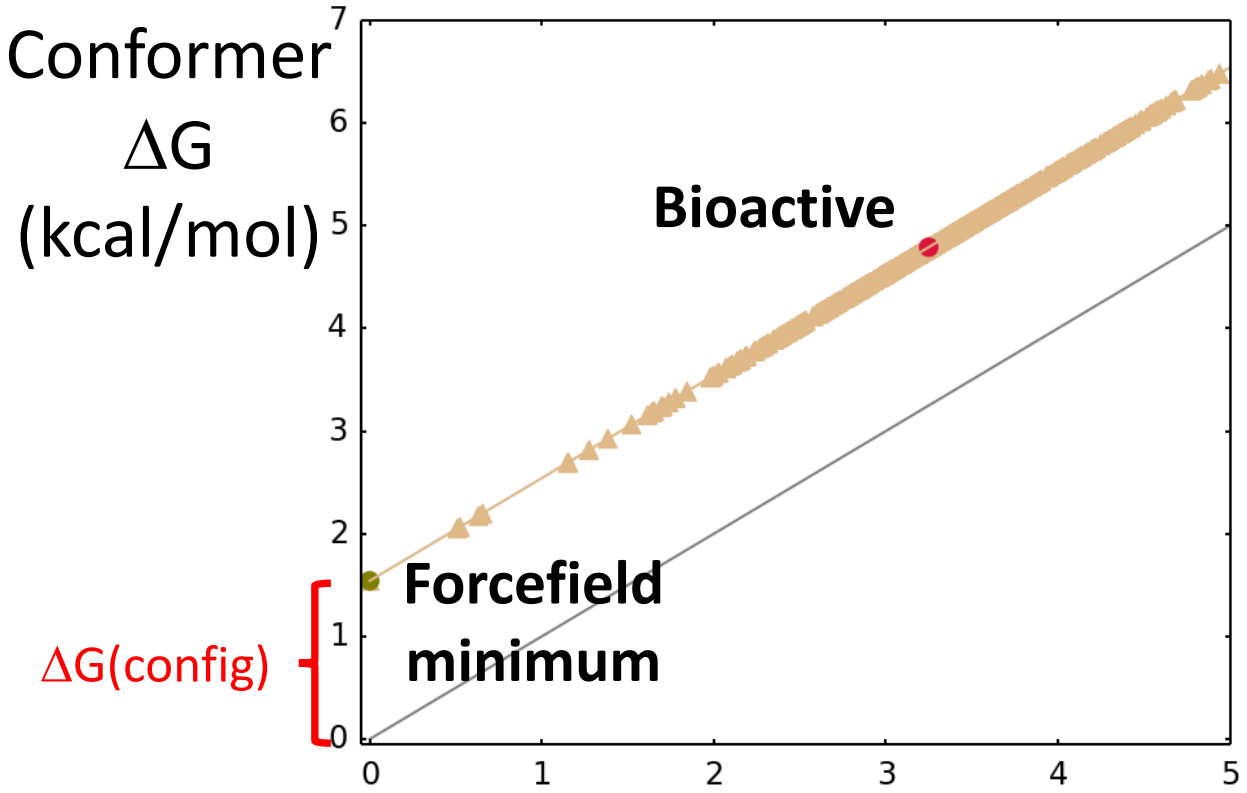




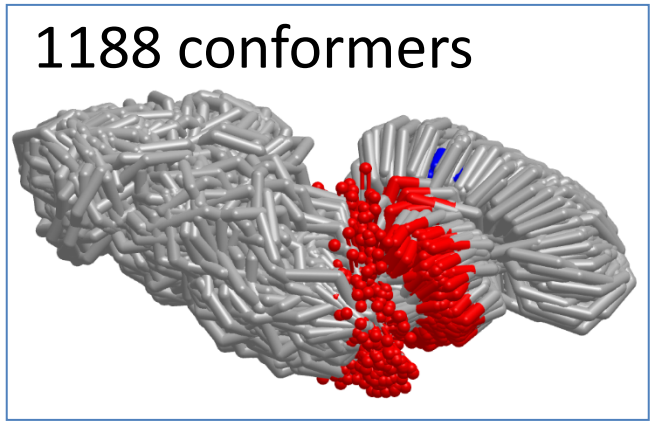
$\Delta G = E\_relative = \Delta E_{ForceField} + \Delta E_{solvation}$

Ignores $T\Delta S$ of ligand and protein

Now add Entropy from all the other Conformations



$$=\Delta E_{ForceField} + \Delta E_{solvation}$$



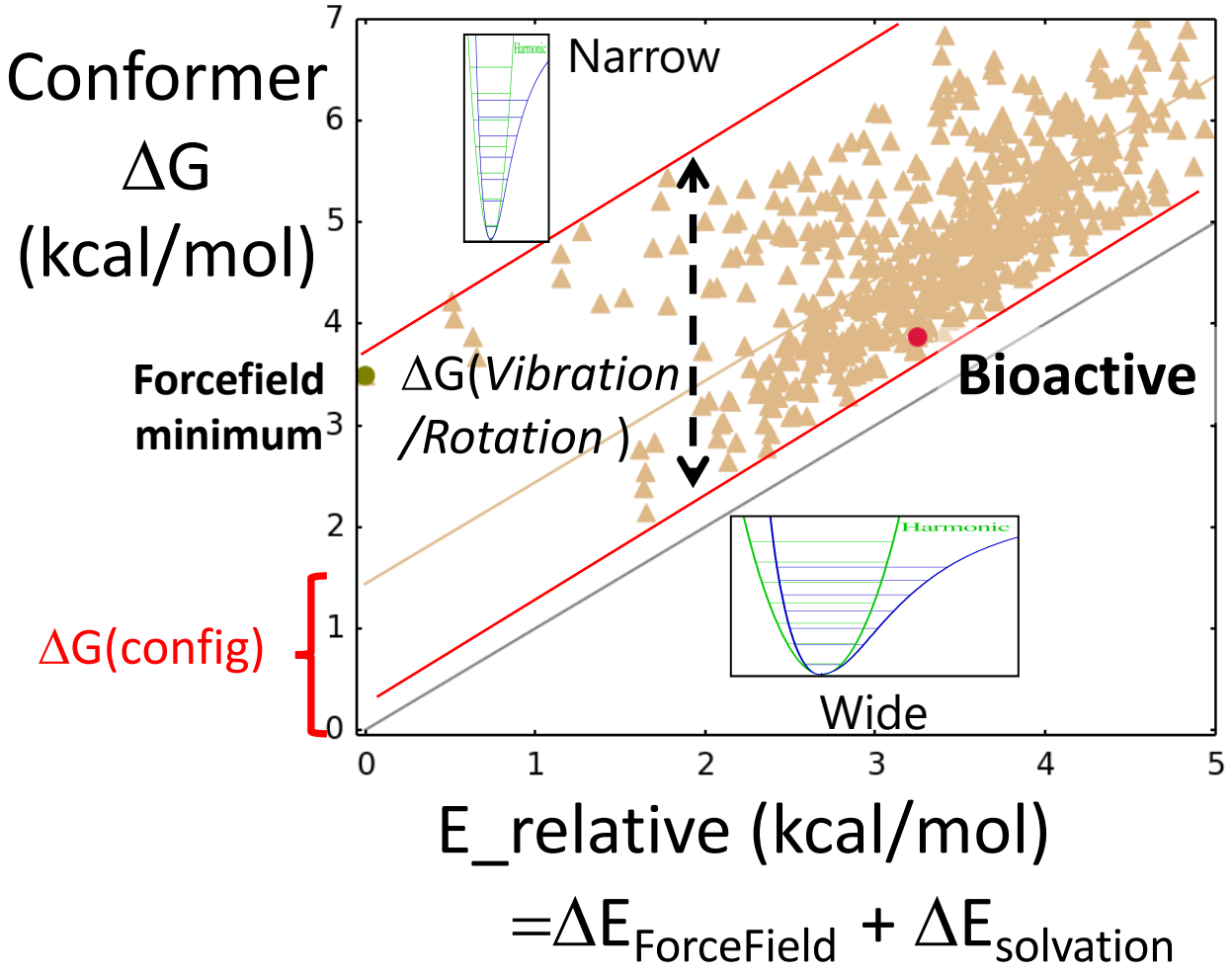
+TΔS(Configurational)=
Constant shift

- Lowers *probability(bioactive)*

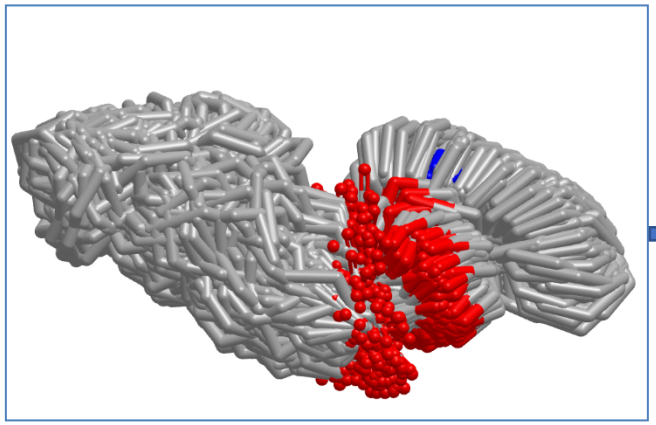
S. Wlodek, A.G. Skillman, and A. Nicholls, *JCTC* **6**, 2140 (2010)

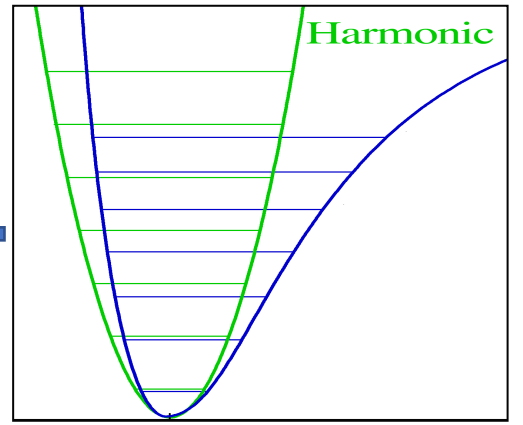
OpenEye SCIENTIFIC

Finally add local vibrational/ rotational entropy



$$=\Delta E_{ForceField} + \Delta E_{solvation}$$





+TΔS(Configurational)

+TΔS(Vibration, Rotation)

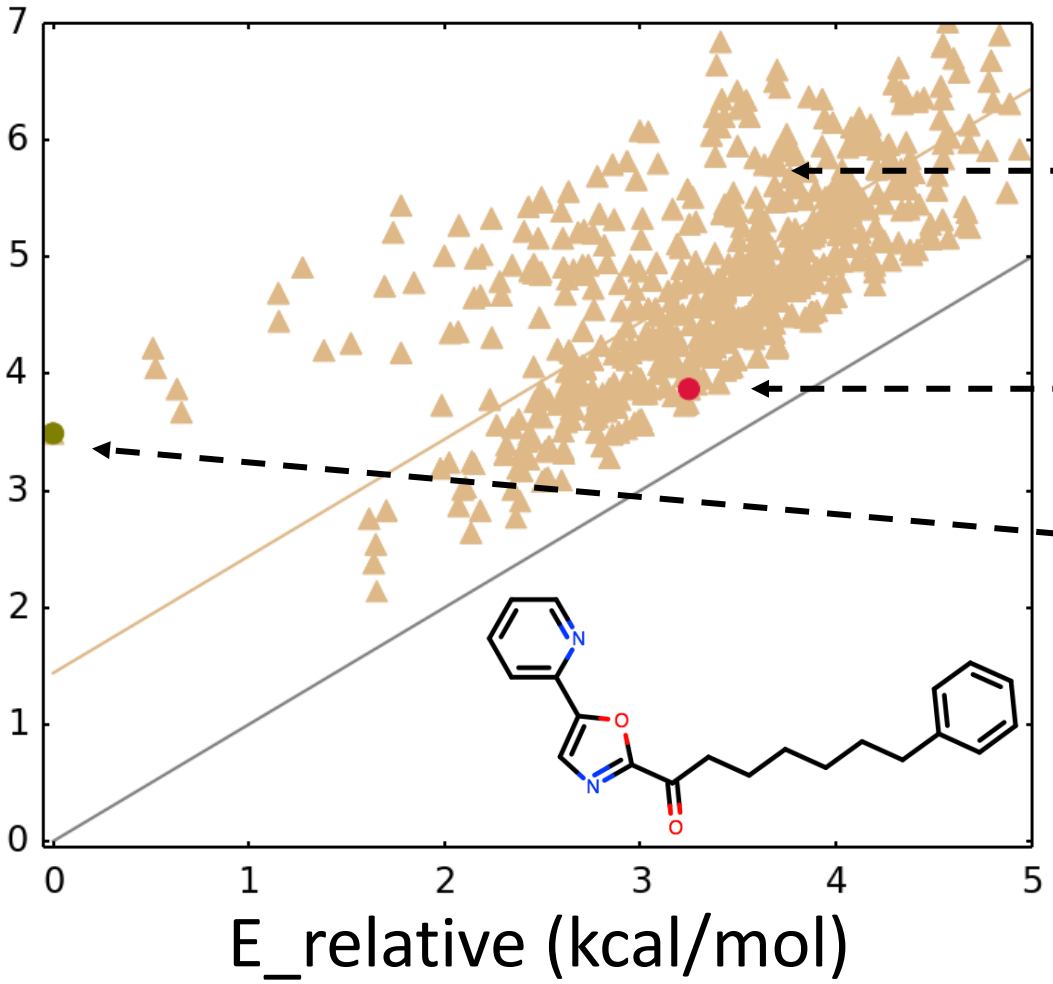
- Wider wells are more favorable
- Narrow wells are less favorable

S. Wlodek, A.G. Skillman, and A. Nicholls, *JCTC* **6**, 2140 (2010)

OpenEye SCIENTIFIC

FreeForm Free Energy Report

Conformer
ΔG
(kcal/mol)



All the minima from SZYBKI solvent minimization

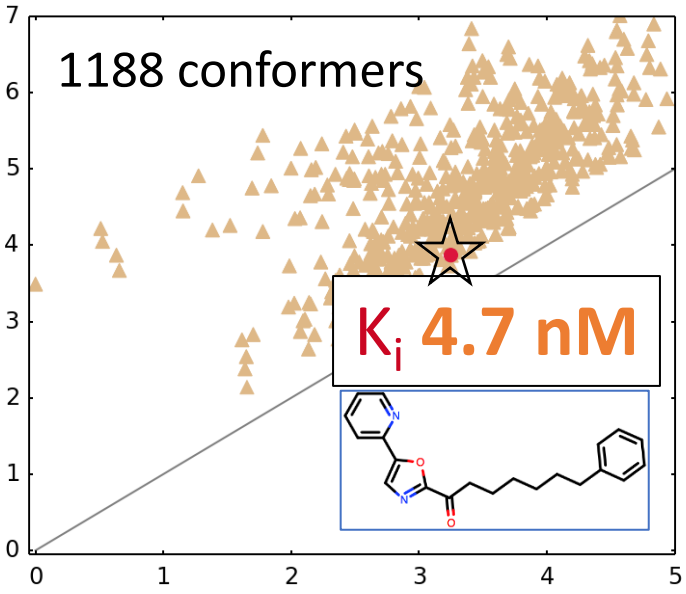
Bound conformer

Lowest MMFF94s conformer

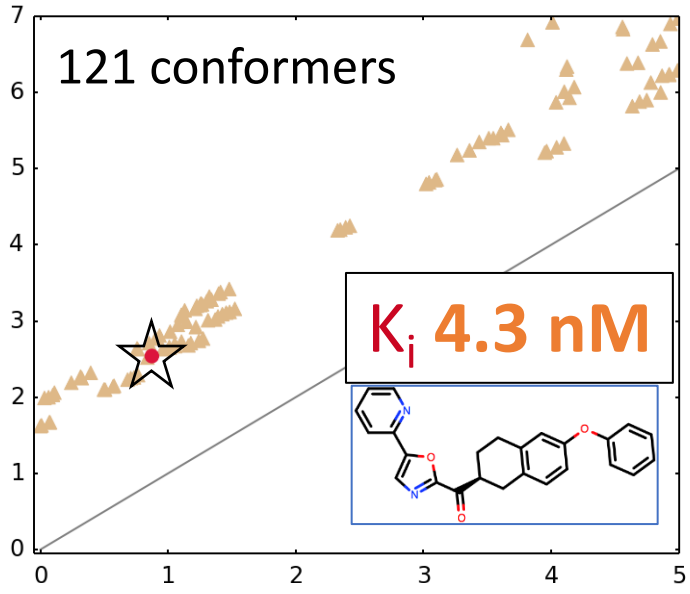
OpenEye SCIENTIFIC

Comparing Lead and Constrained Analog

Conformer ΔG (kcal/mol)



E_relative (kcal/mol)



E_relative (kcal/mol)

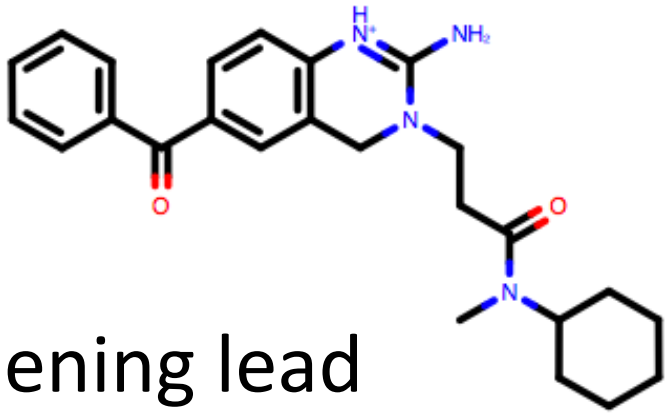
☆ = Bioactive

3.9	Relaxed Bioactive (kcal/mol)	2.5
2.3	Local Strain (kcal/mol)	3.0
6.2	Total (kcal/mol)	5.5

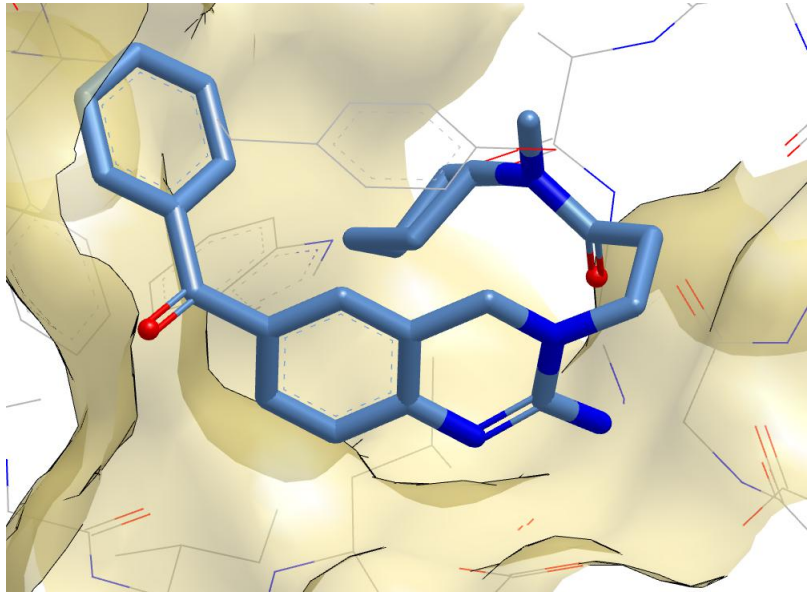
- Constrained analog <1 kcal/mol more stable

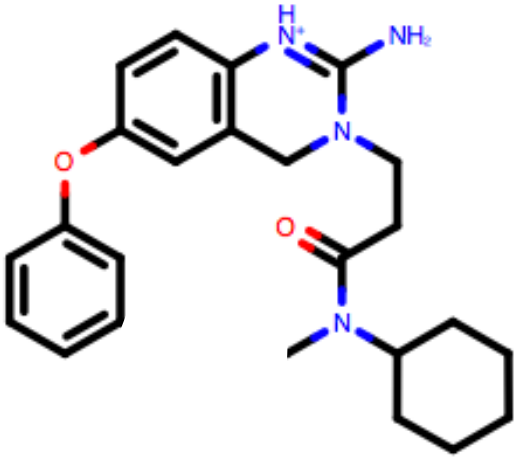
OpenEye SCIENTIFIC

Lead Optimization: Example 2, BACE inhibitor



Screening lead

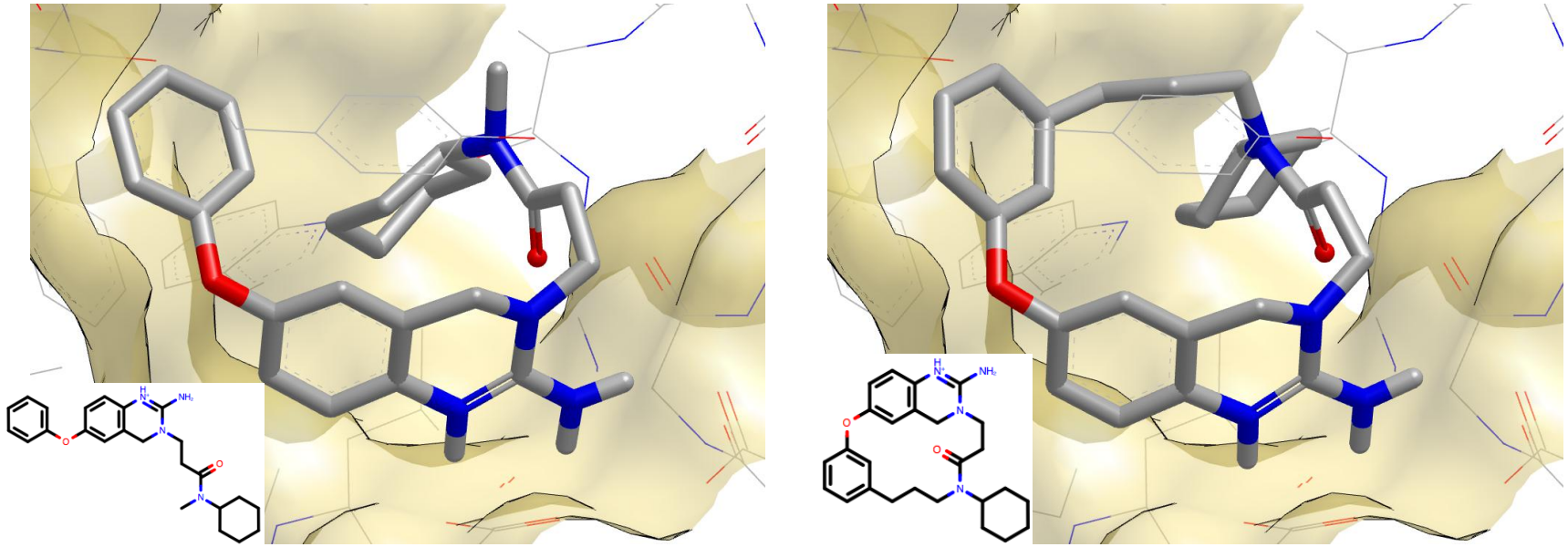




- Lead bound in a hairpin conformation
- Would **macrocyclization** improve potency?

Huang et al., *BMCL* **20,** 3158 (2010)

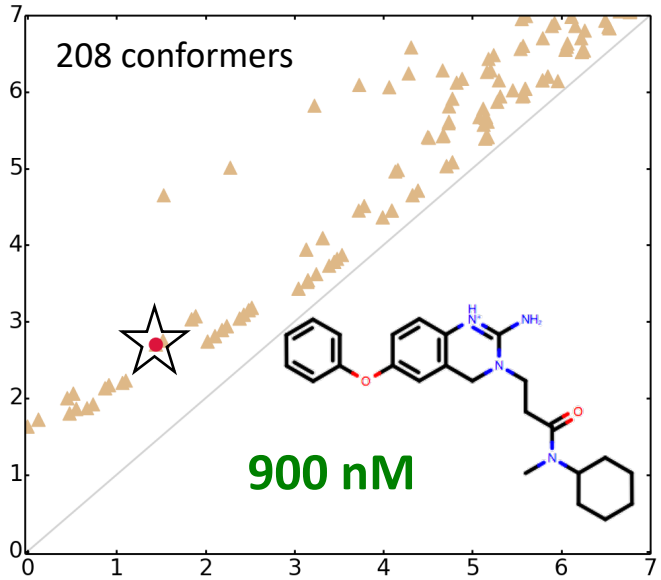
BACE Macrocyclization: Design 1



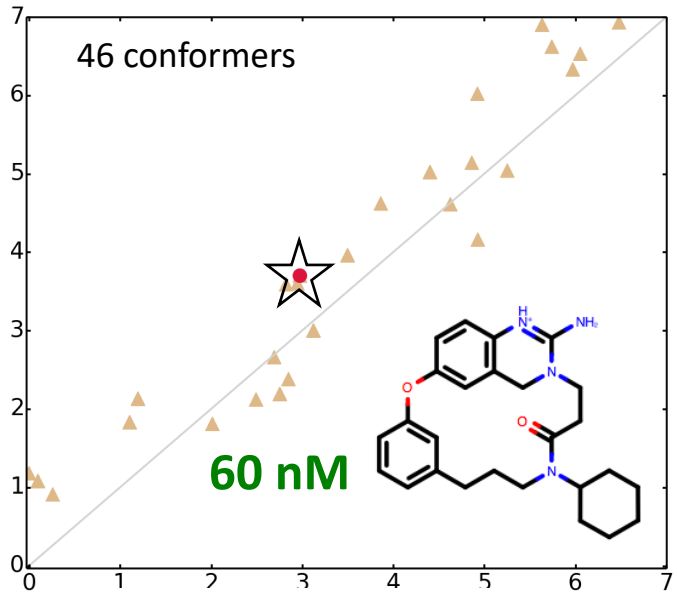
- Macrocyclize via an ethyl linker (model)
  - Assuming ~equivalent binding interactions
- Conformer ΔG on open and macrocyclic analog
  - Decreased Conformer Free Energy?
  - Decreased Global Strain?

BACE Macrocyclization: Design 1

Conformer ΔG (kcal/mol)



E_relative (kcal/mol)



E_relative (kcal/mol)

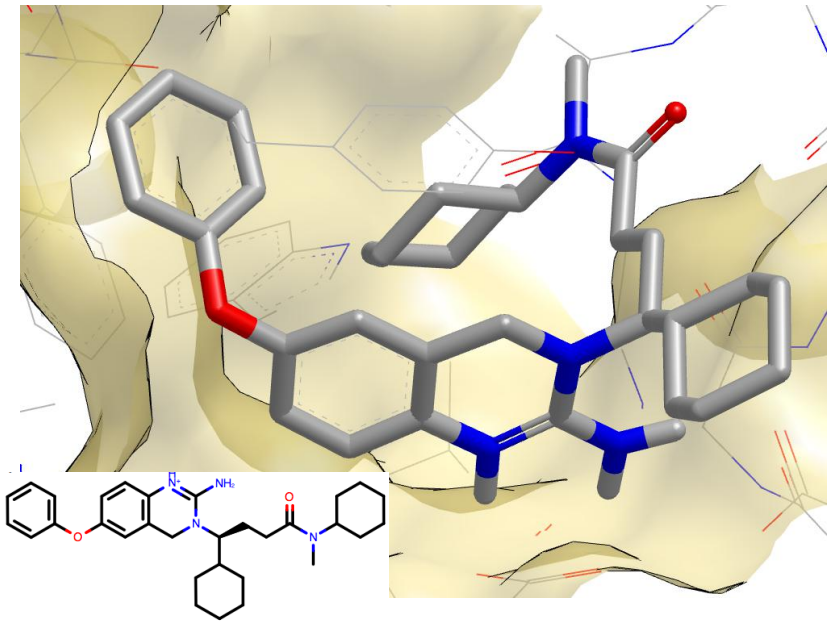
☆ = Bioactive

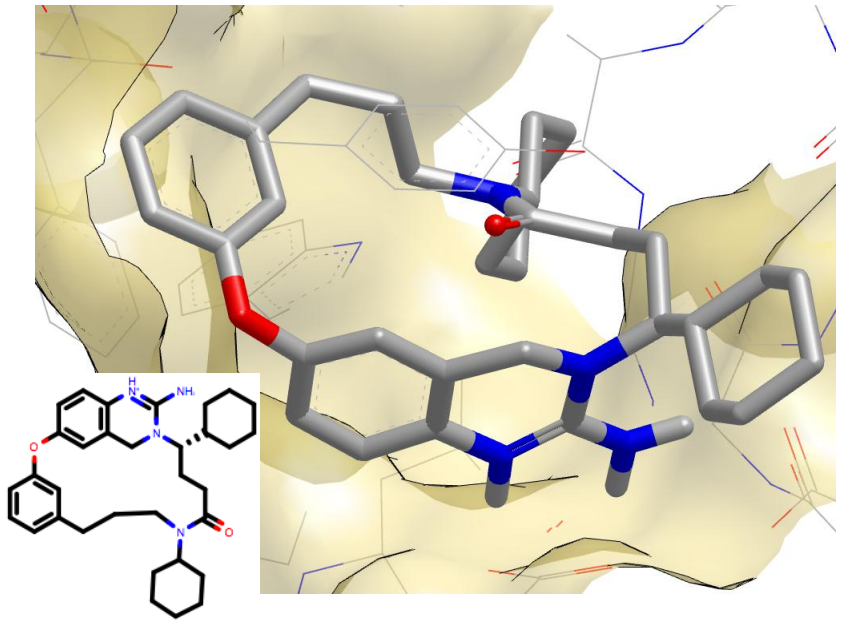
2.7	Relaxed Bioactive (kcal/mol)	3.7
6.5	Local Strain (kcal/mol)	0.8
9.2	Total (kcal/mol)	4.5

- Conformer free energy slightly disfavors macrocycle
- Local strain strongly favors the macrocycle

OpenEye SCIENTIFIC

BACE Macrocyclization: Design 2





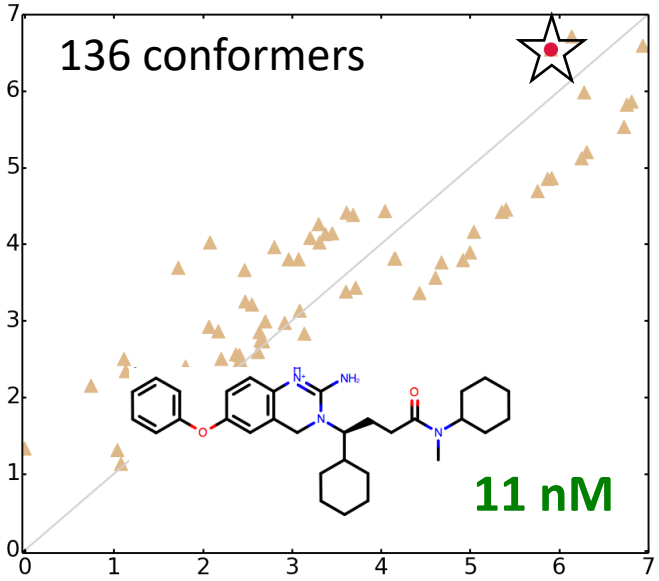
- Add S1’ cyclohexyl
  - Macrocyclize with ethyl linker (model)
- Conformer ΔG on open and macrocyclic analog
  - Decreased Conformer Free Energy?
  - Decreased Global Strain?

Huang et al., *BMCL* **20,** 3158 (2010)

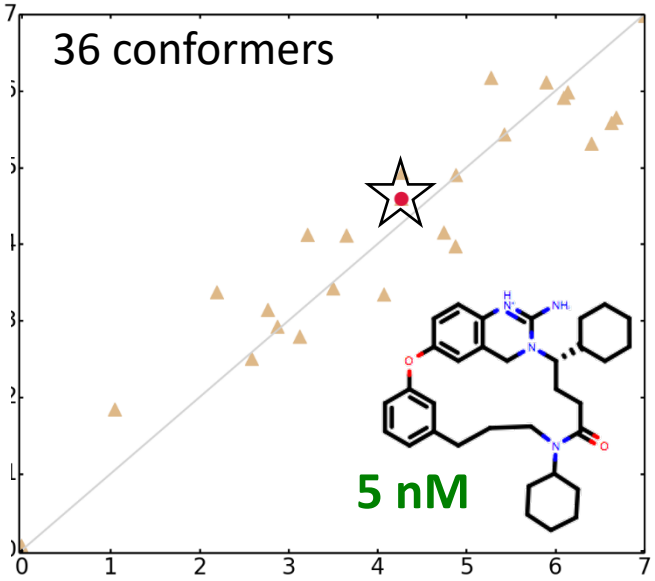
OpenEye SCIENTIFIC

BACE Macrocyclization: Design 2

Conformer ΔG (kcal/mol)



E_relative (kcal/mol)



E_relative (kcal/mol)

☆ = Bioactive

6.6	Relaxed Bioactive (kcal/mol)	4.6
2.1	Local Strain (kcal/mol)	3.9
8.7	Total (kcal/mol)	8.5

- Conformer free energy favors macrocycle
- Local strain disfavors the macrocycle

OpenEye SCIENTIFIC

FreeForm: Free Energy of the Unbound Ligand

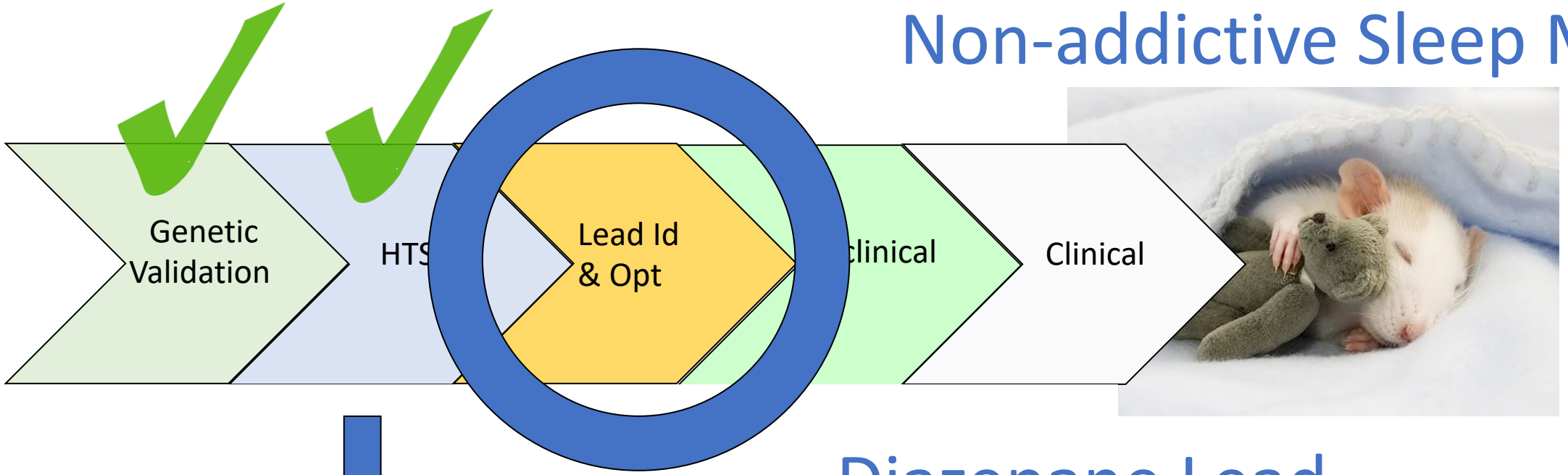
What does it do?

- Estimates the energetic cost of ligand *conformation restriction* upon binding to a protein
- With forcefield, estimates the protein-induced “local” strain

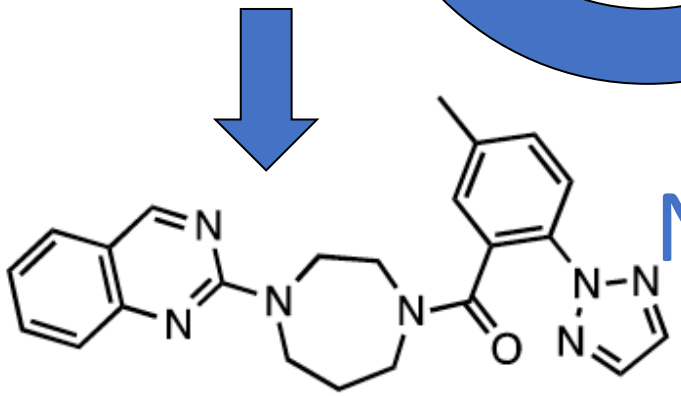
Why would you use it?

- Predicting the effects of constraining analogs in lead optimization
- Selecting low (free) energy conformers for model building

Successful Merck HTS for Orexin Inhibitor



Non-addictive Sleep Med

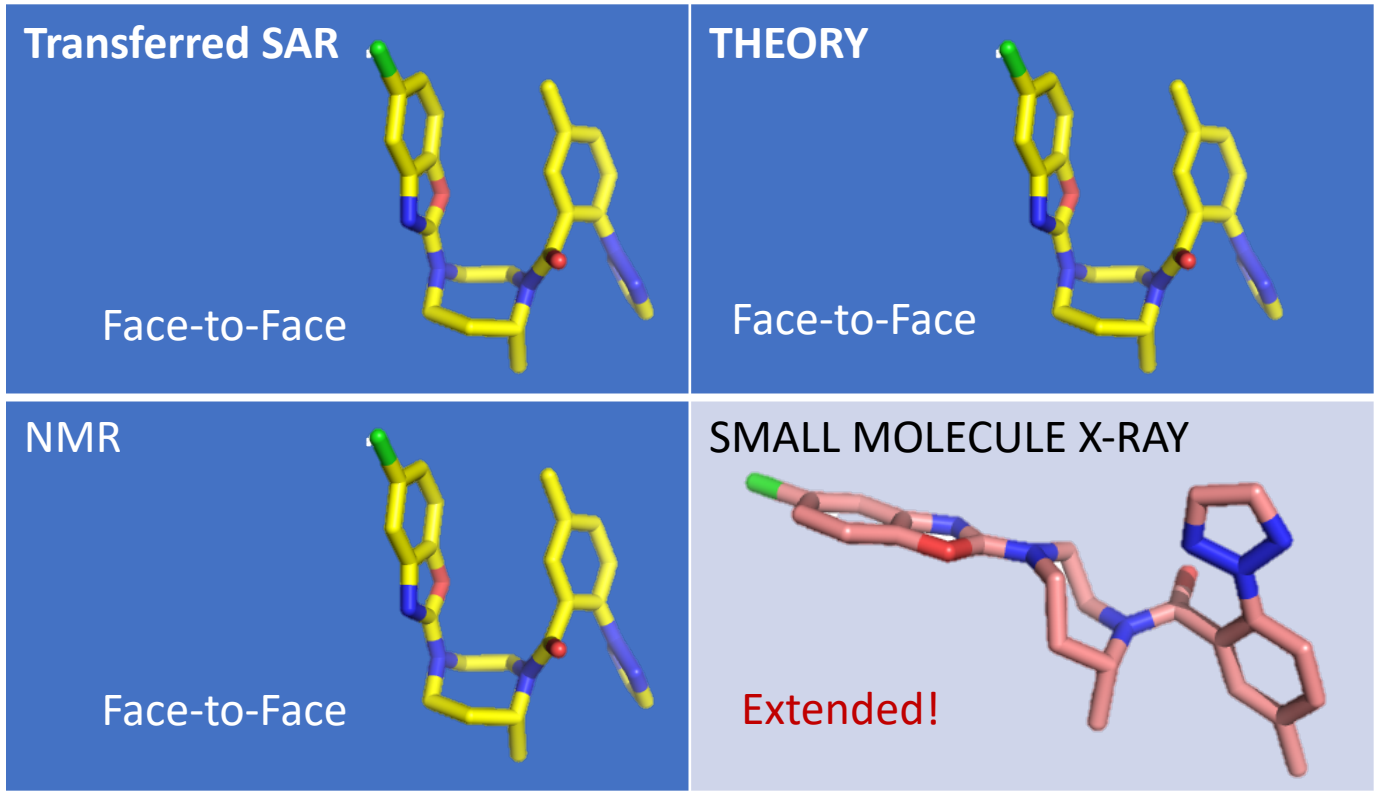


Diazepane Lead

No Bound Crystal Structure!

OpenEye SCIENTIFIC

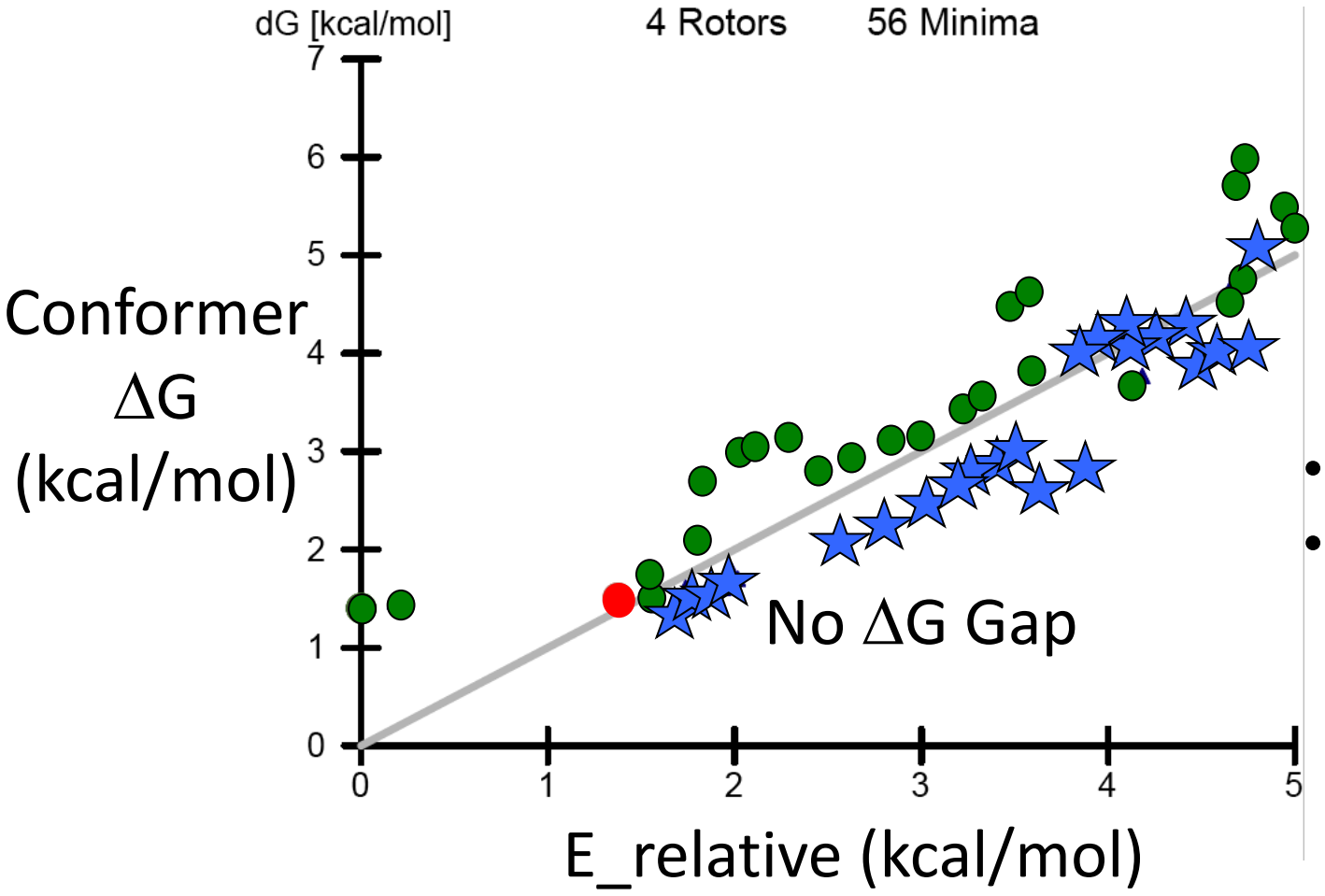
So, what is the Bioactive 3D Form on which to base Design?



QM, Force-Fields

Usually, reliable guide to bound form

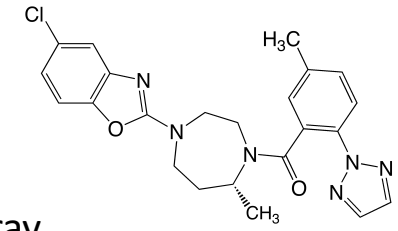
Explanation from FreeForm



Face-to-Face

Extended

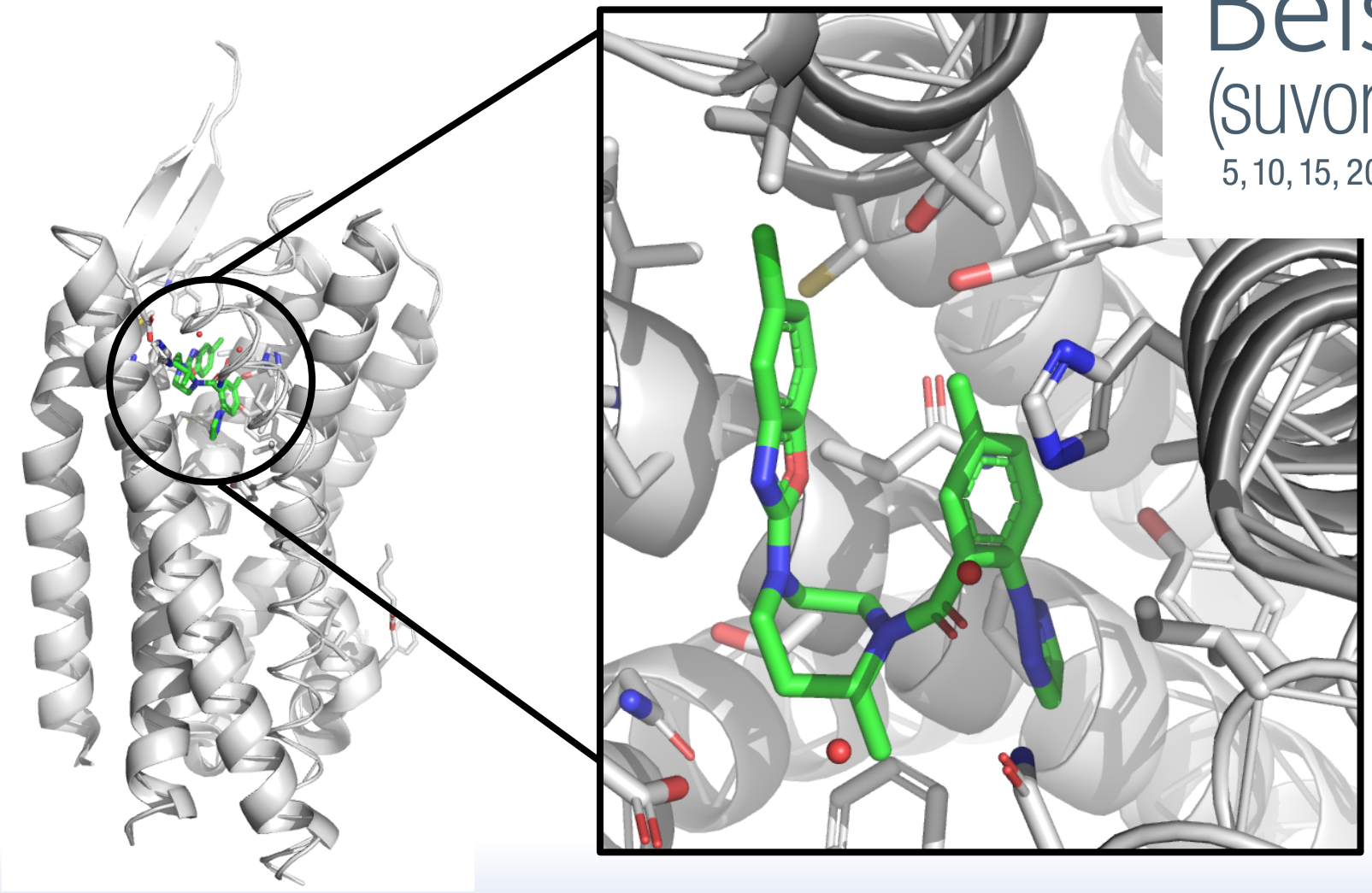
Small molecule X-ray



- Face-to-Face Enthalpy preferred
- But configurations **equivalent** in ΔG

“Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant”

Nature Letter, 2015 & Nature Structural & Molecular Biology, 2016



Belsomra®
(suvorexant) ℞
5, 10, 15, 20 mg tablets

OpenEye SCIENTIFIC

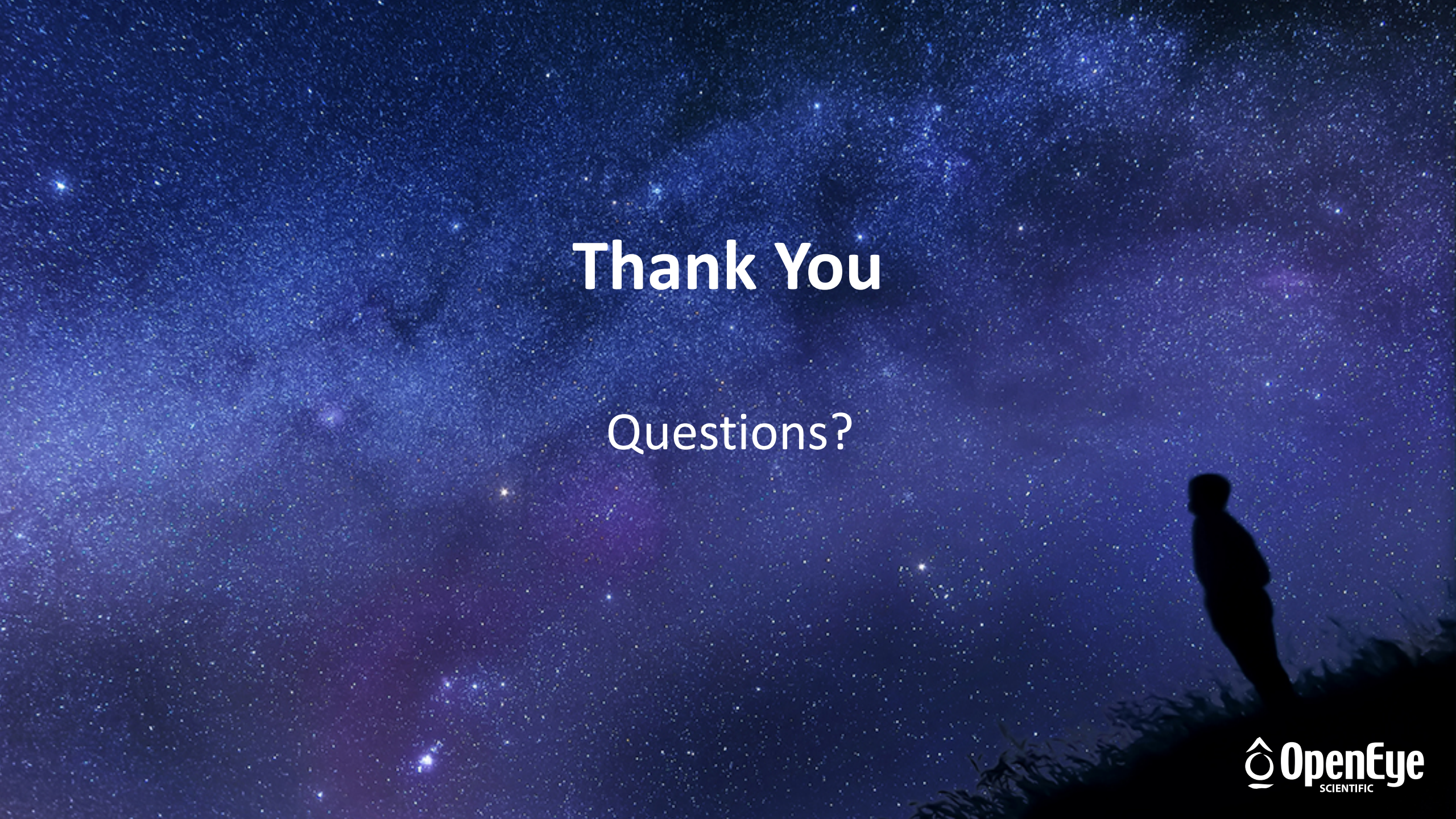
Summary

FreeForm

- Estimates free energies: “entropic” *conformation restriction* + “enthalpic” protein-induced *local strain*

Uses for Design

- SBDD: Assessing analogs in lead optimization
- LBDD: Selecting low (free) energy conformers for model building


Thank You
Questions?
OpenEye
SCIENTIFIC