Ensemble-based conformational modelling for macrocycles

From X-ray refinement to lead optimisation

Himani Tandon, Kyle G. Kroeck, Ajay N. Jain, Ann E. Cleves, Mario Öeren, Matthew D. Segall

Optibrium Ltd, Cambridge, United Kingdom, ***email: info@optibrium.com***

Introduction

Conventional ligand-fitting and refinement methods in X-ray electron density maps often yield models with unrealistically high conformational strain. This occurs because ligands are typically treated as single conformers using atom-specific B-factors. These parameters often introduce geometric distortions or obscure biologically meaningful pose variations that might misguide lead optimisation.

xGen™ is a real-space refinement method that addresses these problems by balancing how well a ligand fits the electron density with its internal conformational strain. Rather than a single static pose, it produces a realistic ensemble of conformers that collectively explain the experimental data. This approach reveals meaningful structural variations and yields physically plausible models with substantially reduced strain.

The xGen™ method

Step 1. Generate a pool of conformations

Density-aware search where each conformer is expanded into three variants (a trio):

- **Energy-favoured** – strain is minimised
- **Density-weighted** – fit is maximised
- **Compromise** – strain and fit are balanced

⇩

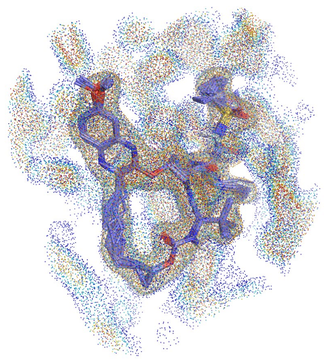
Step 2. Filter for quality

Retain trios whose members are geometrically similar, low in energy, and provide a near-optimal fit to the density.

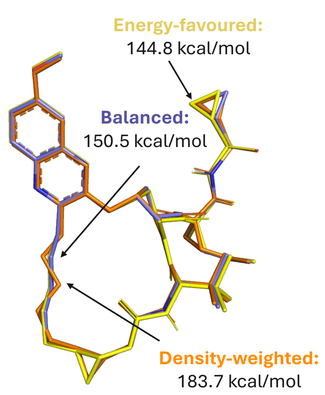
⇩

Step 3. Build the ensemble

Assign occupancy weights to the surviving conformers, minimising the real-space R-factor (RSR).

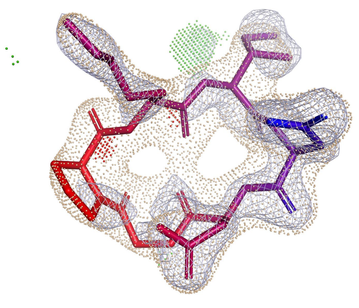


Density-aware conformational search

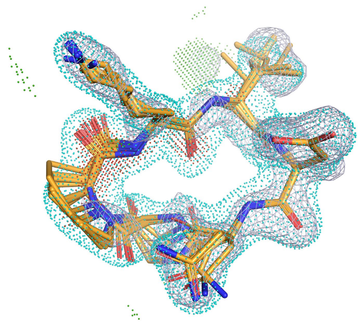


A conformer trio

PDB: RSCC = 0.71 RSR = 0.24



xGen: RSCC = 0.75 RSR = 0.25

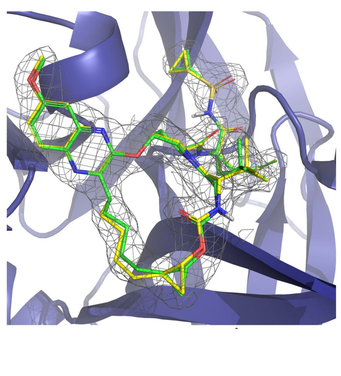


Comparison of the traditional single-conformer model (left) and xGen occupancy-weighted ensemble (right). xGen captures explicit heterogeneity and improved density fit with lower strain.

xGen ensembles reduce ligand strain by half

Validated on ~3,000 complexes, xGen ensembles reduce artificially high PDB ligand strain by ~50% while improving density fits. This occupancy-weighted approach provides physically realistic, low-energy alternatives to traditional single-conformer models.

Average Strain (kcal/mol)	Deposited	xGen
Macrocycles	6.8	3.7
Non-macrocycles	4.2	2.5

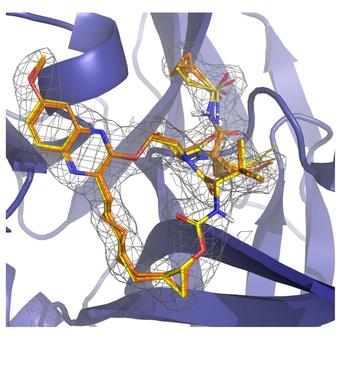


PDB 3SUE, Deposited
RSCC=0.95, RSR =0.12

E_{PDB} = 316.8,
E_{surr} = 154.7,
RMSD = 0.3 Å

E_{gmin} = 138.6,

Strain = 16.1 kcal/mol



PDB 3SUE, xGen Ensemble
RSCC=0.96, RSR =0.11

E_{xGen} = 154.3-183.7,
E_{surr} = 142.5-144.6,
RMSD = 0.3-0.4 Å

E_{gmin} = 138.6,

Strain = 3.9 kcal/mol

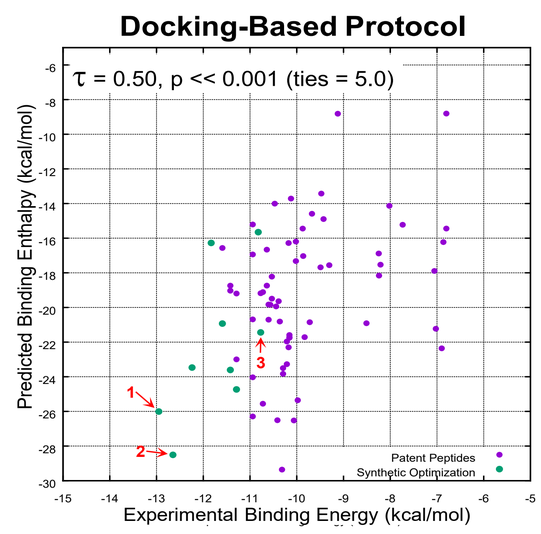
Grazoprevir-NS3/4A protease (3SUE). PDB ligand (green) fits the electron density well (left) but shows high strain (16.1 kcal/mol), calculated as the difference between the surrogate conformer (yellow) energy (E_{surr}) and the global minimum conformer energy (E_{gmin}). xGen ensemble (orange) maintains fit quality (improved RSCC/RSR) while reducing strain by 75% to 3.9 kcal/mol (right).

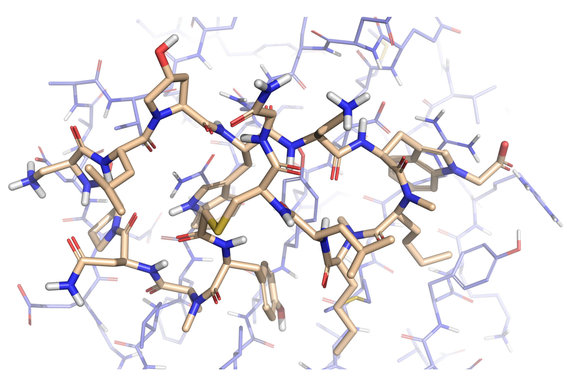
Reduce synthetic effort by up to 90%

In an integrated workflow, xGen-generated low-strain ensembles of the lead peptide were used to derive conformational restraints. Combined with docking and scoring, this approach:

- accurately ranked compounds by predicted binding energies,
- generated correct poses for docked macrocycles.

The conformational strain emerged as the dominant predictor of ligand binding, and the clinical candidate was identified in the top-ranked 10% of synthesised analogues.



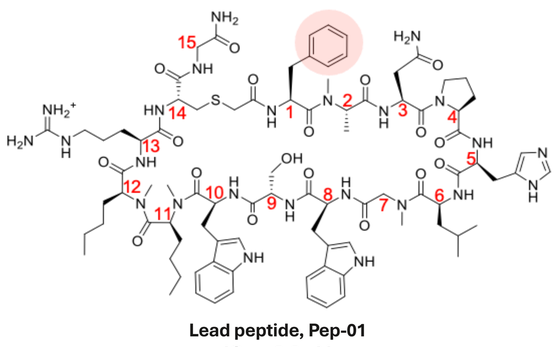


Docked clinical candidate BMT-174900
Intermolecular score: -32.9 kcal/mol, strain: 6.9 kcal/mol, overall enthalpy: -26.0 kcal/mol

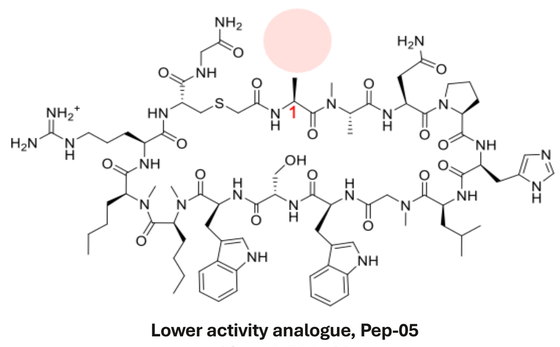
Accurate ranking of PD-L1 macrocyclic peptides. The correlation (left) between experimental and predicted binding energy (calculated as the sum of intermolecular enthalpy and conformational strain) successfully identified the clinical candidate BMT-174900 (right) as one of the top 10% of synthesised analogues. The points 1, 2, and 3 correspond to BMT-174900, BMT-153099, and BMT-139699, respectively.

The importance of strain

1000x drop in activity



Lead peptide, Pep-01
IC_{50} = 9.0 nM



Lower activity analogue, Pep-05
IC_{50} = 6,500 nM
Docking: Δ+0.4 kcal/mol
Strain: Δ+8.6 kcal/mol

Strain as a predictor of binding. Substituting phenylalanine with alanine at position 1 in Pep-05 reduces activity by 3 log units despite a minor loss in intermolecular binding energy (<0.5 kcal/mol). However, this change increases macrocycle strain by ~9 kcal/mol.

Conclusions

xGen offers a paradigm shift for ligand modelling, producing physically realistic conformer ensembles for ligands with lower strain

Ligand strain is a predictive factor for macrocyclic peptide optimisation: If the ligand has high strain, aim to optimise its geometry, and if it already has low strain, improve protein-ligand interaction footprint

Accurate pose prediction enables effective CADD-driven macrocycle design, boosting efficiency and productivity

References

Jain, A. N. *et al.* xGen: Real-Space Fitting of Complex Ligand Conformational Ensembles to X-ray Electron Density Maps. J. Med. Chem. **2020**, *63*, 10509-10528

Jain, A. N. *et al.* Complex peptide macrocycle optimization: combining NMR restraints with conformational analysis to guide structure-based and ligand-based design. JCAMD. **2023**, *37*, 519-535

Acknowledgements

Merck: Alexander C. Brueckner, Mikhail Reibarkh, and Edward C. Sherer

Optibrium: James Halle and Nathan Brown

Read the full paper