Scaling Lead Optimization to Thousands of Compounds: Active Learning with FE-NES and 3D-QSAR

Jingyi Chen, Agnes Huang, and David Hamilton

OpenEye, Cadence Molecular Sciences, 9 Bisbee Court, Santa Fe, NM 87508

Summary:

- Active learning using binding affinity predictions from FE-NES to build 3D-QSAR models that can usefully rank large virtual libraries.
- By training on only a small fraction of the library the workflow identifies high-potency molecules at a fraction of the cost of exhaustive free energy evaluation.
- 3D-QSAR models trained on FE-NES predictions provide reliable prioritization for lead optimization and effectively rank-order held-out compounds with known experimental affinities.

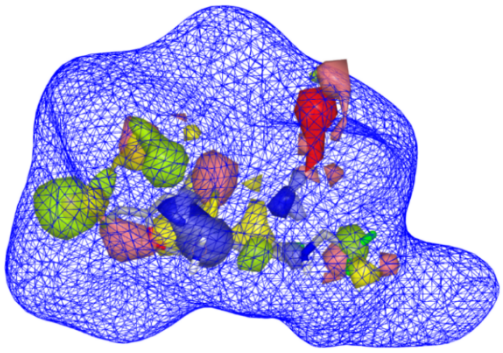
Product Keywords: FE-NES, AL, 3D-QSAR, Orion®

Abstract:

In lead optimization, teams generate virtual libraries of thousands of analogs, but exhaustive free energy calculations on every compound remain cost and time prohibitive. Active learning (AL) addresses this by building a predictive model for affinity from relative binding free energies (RBFE) calculations on a small fraction of the library, enabling prioritization across the full set.

Other active learning frameworks pair free energy calculations with models based on 2D properties. These models can triage large libraries efficiently, but their predictions are hard to interpret in terms of the molecular features driving binding.

OpenEye's approach integrates RBFE predictions from Free Energy Nonequilibrium Switching (FE-NES) with a shape and electrostatics-based 3D-QSAR model, keeping predictions interpretable and tied to specific molecular features rather than abstract fingerprints. Medicinal chemists get a rank-ordered list plus insight into which molecular features drive potency, supporting SAR-driven design decisions. To provide rapid access to the RBFE results the workflow is available on the Orion molecular design platform.



3D-QSAR model showing how molecular features contribute to potency.
Red: acceptor. Blue: donor. Green: hydrophobe

In this prospective study, a virtual library of BACE-1 ligands was generated from known active ligands.[1] Ten AL iterations of 50 FE-NES calculations each (500 total) identified 56 high-potency molecules. The resulting 3D-QSAR model, validated against 36 held-out ligands, achieved a Kendall's Tau of 0.54 (95% CI: 0.35, 0.68), demonstrating useful predictive power on unseen data.[2]

1. *Cumming et al. Bioorg. Med. Chem. Lett.* ***22****, 2444 (2012).*
2. *https://acs.digitellinc.com/live/37/session/597886*

info@eyesopen.com