Application of FEP and QSAR methods to medicinal chemistry projects

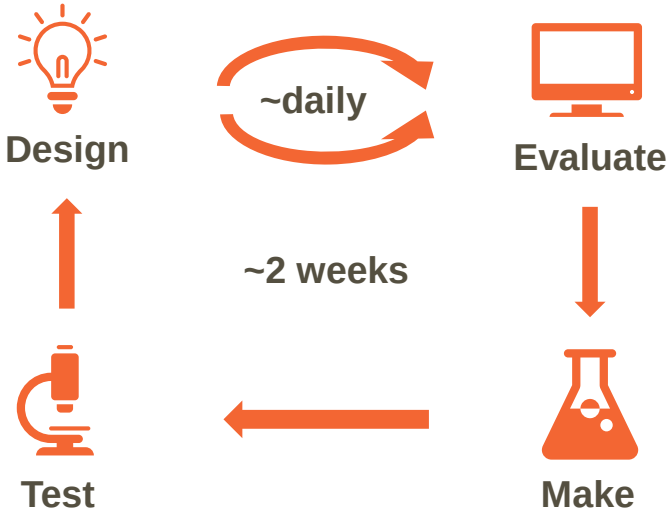
Richard Lonsdale

UK-QSAR Nottingham

26th September 2019

Computational chemistry in drug discovery

- Use atomic resolution simulations to guide new designs
- Evaluate 100s of compound ideas/day
- Design-evaluate step must keep pace with rest of medicinal chemistry cycle

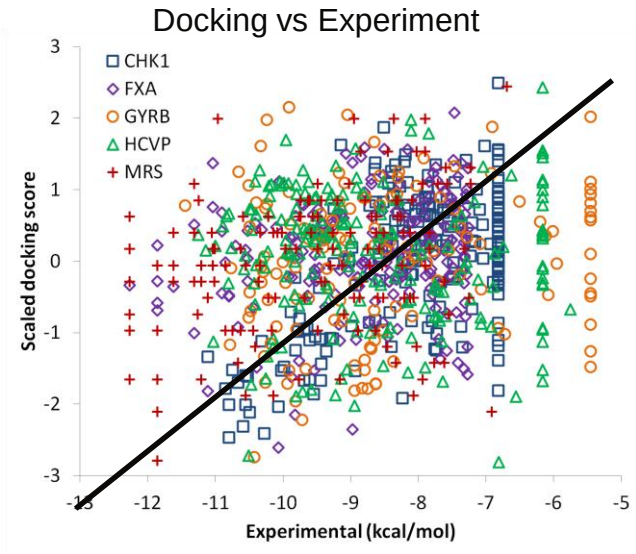


Accurate predictions will help to reduce the number of cycles

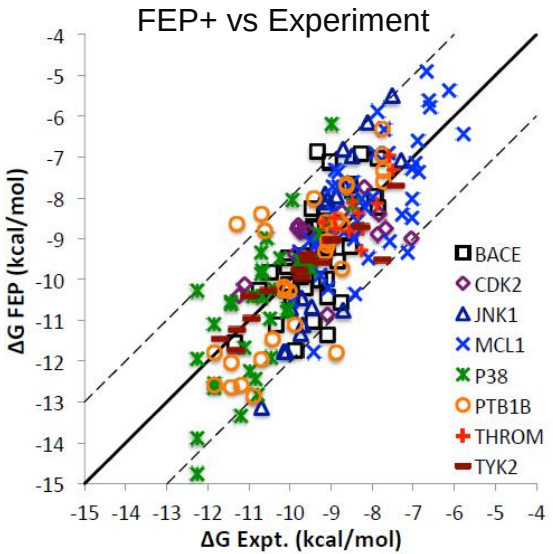
Structure-based potency prediction

Progress over the past ~10 years

- To keep pace with chemistry we had to rely on less rigorous methods
- Now more rigorous methods can be used at scale and for more diverse molecules
- Calculations run on a time scale that can keep up with chemistry (wks→hrs)



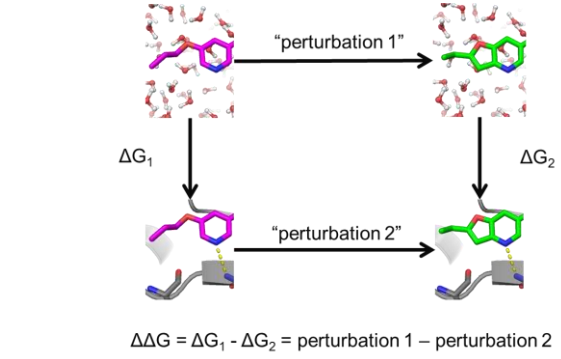
Adapted from Warren (2006) *J. Med. Chem.*

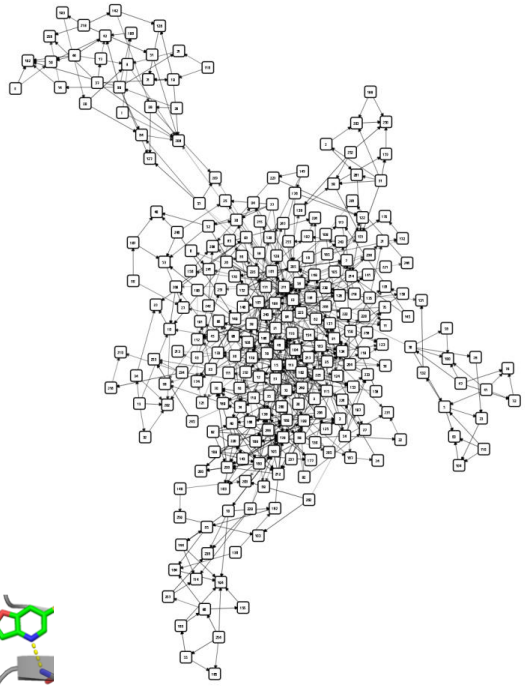


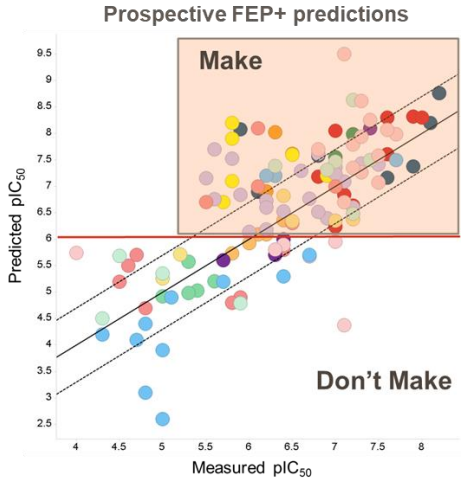
Wang (2015) *JACS*

Free Energy Perturbation (FEP)

$\Delta\Delta G_{bind}$ via explicit solvent molecular dynamics simulations







Generation of input for FEP

Maximising the probability of success

- High resolution crystal structure
  - High confidence in binding pose
  - Binding mode must be conserved across series
  - No chain breaks near to the binding pocket
  - Assay data and crystal structures from same construct
- When generating FEP maps:
  - Only series with same binding mode should be modelled
  - Charge of ligand *should* be the same
  - Only ligands with similar protein conformation for the holo complex should be modelled
  - Protonation and tautomeric state of ligands important
  - Ligands should be aligned to crystal ligand as tightly as possible

Example application

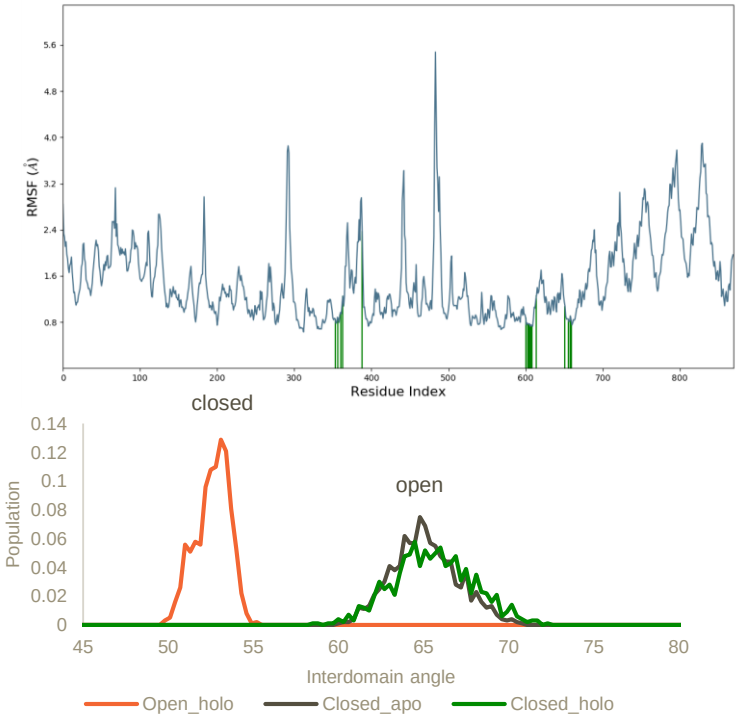
Target A

- **Goal:** increase potency and solubility
- Target exists in distinct open and closed forms
- Four series under consideration at the time
  - Two bind to open form
  - Two bind to closed form
- Crystallography for each series in relevant conformation

Is the target suitable for FEP?

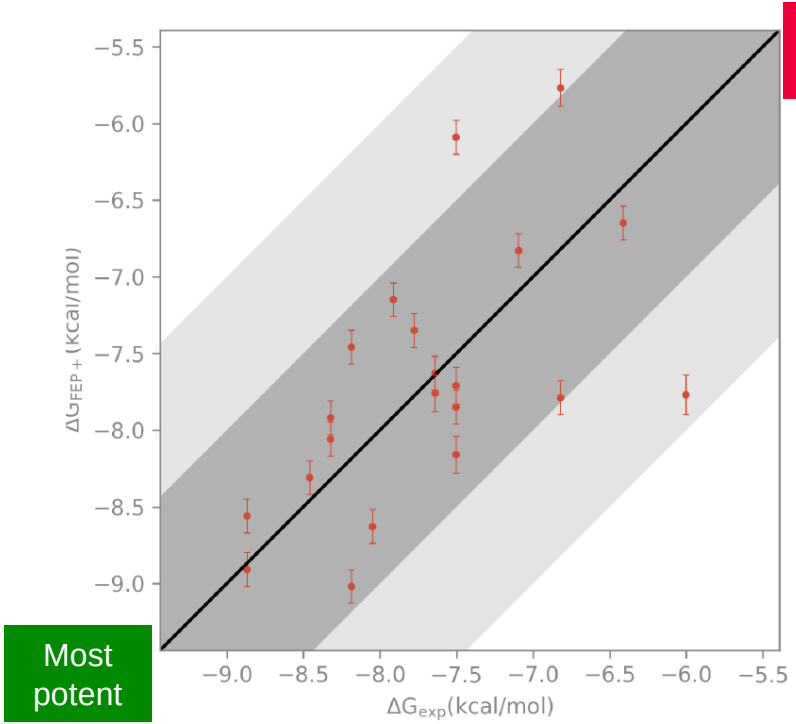
Preliminary MD simulations

- Are the open and closed conformations stable on the FEP timescale (~5 ns)?
- Unrestrained simulations (500 ns) performed on both conformations
- Conformations are retained on this timescale
  - Small change in interdomain angle
- Low fluctuation in region making contact with ligand
- Protein motions beyond FEP+ timescale
- **Suggests system may be amenable to FEP+**

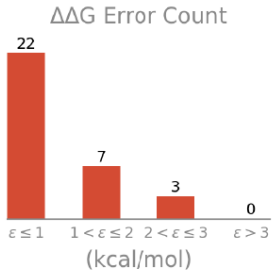


Retrospective benchmarking test

Series 1



R^2 = 0.35
RMSE = 1.03 kcal/mol



- "full" map
- 5 ns
- all ligand atoms in REST region for solvent leg
- GCMC solvation of protein pocket during relaxation stages

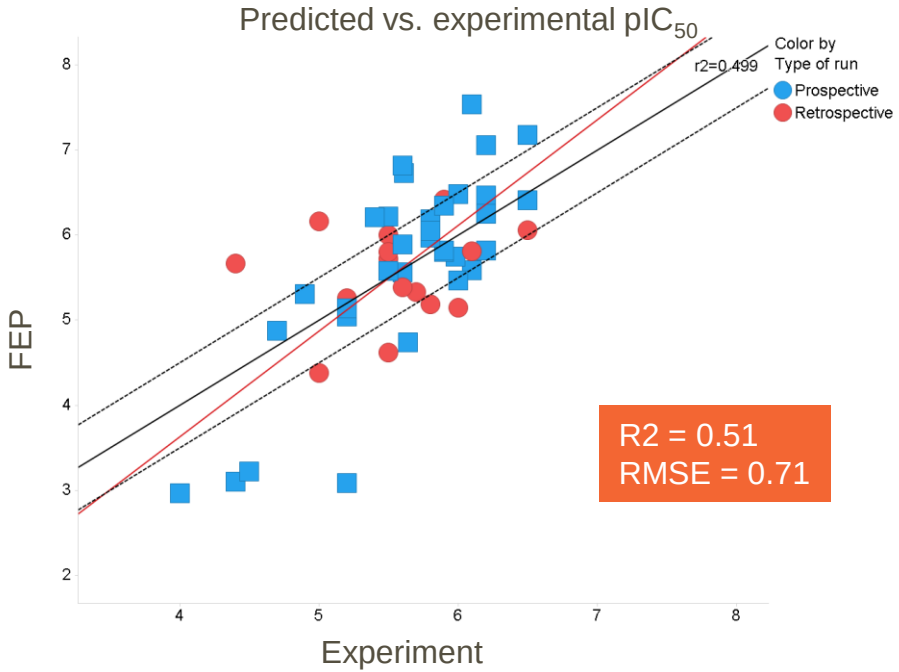
RMSE < 1.3 kcal/mol
➢ suitable for prospective use

- Total time to run simulations: ~24h

Prospective FEP simulations

Series 1

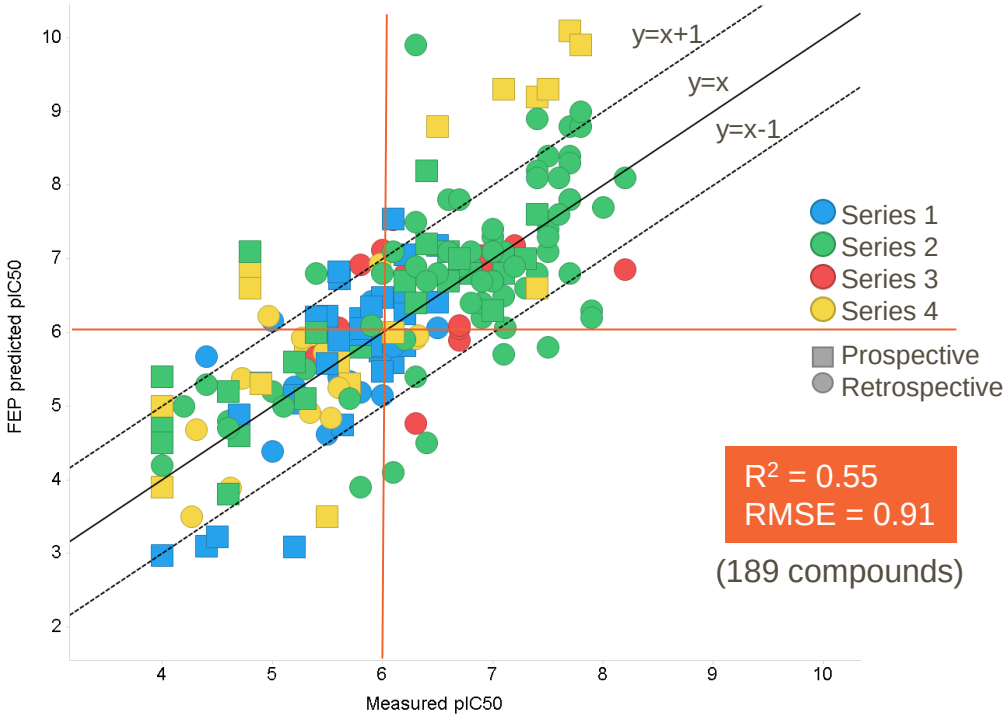
- 127 predictions based on variations at R_1, R_2 and R_3
  - Chemist's ideas
  - Sparse array compounds
  - Compounds generated by molecule generators
- 48 compounds synthesized
  - Most potent compound in series contains R_3 group prioritised with FEP (pIC_{50} = 7)



FEP

Overall performance with Target A

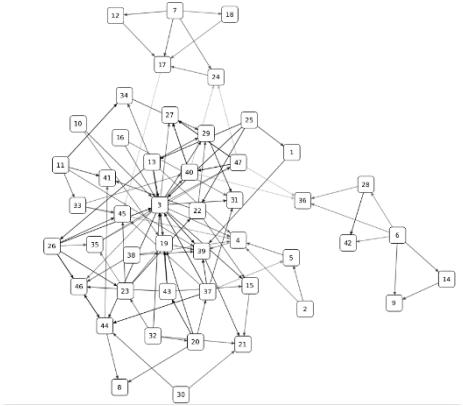
- Retrospective and prospective use
  - Retrospective studies demonstrate suitability of series for FEP
- Performed on 4 series (~800 compounds)
  - Series 1 and 3 bind to closed form
  - Series 2 and 4 bind to open form
- Observing anticipated performance
  - Majority of predictions fall within 1 log unit of measured potency
  - Relatively low number of false positives and negatives



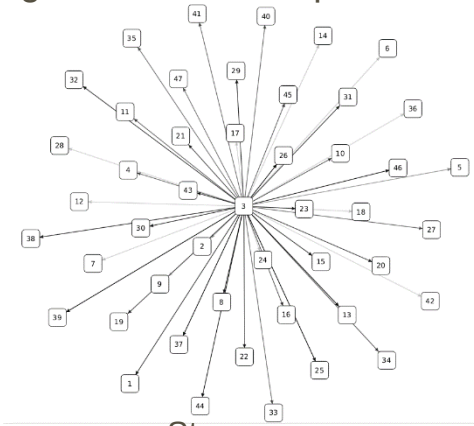
Reducing the overhead for large compound sets

Benchmarking “star” map vs. “full” map

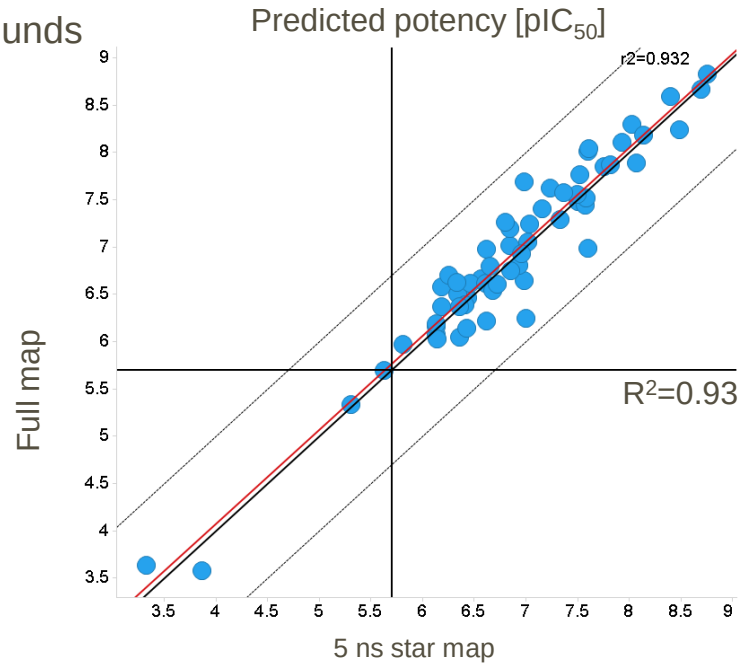
- Number of connections rises sharply with number of compounds
- Full map with 10 ns sampling used as benchmark
  - Star maps tested with 1, 2 and 5 ns sampling
- **Star map with 5 ns adequate for large sets**
- Best compounds can be progressed to full map



Full map
105 connections



Star map
46 connections



QSAR vs. FEP potency prediction

The best of both worlds?

	QSAR	FEP
Throughput	High	Low
Dataset required	Large	Small
Applicability domain	Small	Large

Can we combine the predictive capabilities of FEP with the high throughput of QSAR at an early stage?

- Build QSAR models on FEP predictions
  - Increase chemical space of QSAR without making 100s of compounds
  - Use as a pre-filter for output from molecule generators (1000s of molecules) prior to running FEP

Building a QSAR model on assay data

Series 1

- QSAR model built using experimental pIC_{50} values
  - Constructed with temporal sets of compounds
  - Selection of methods and descriptor sets tested

Compounds in training/test set	Cross-validation R^2	Cross-validation RMSE
35 (28/7)	0.13	0.90
70 (56/14)	0.30	0.61
135 (108/27)	0.28	0.57

Model type	GFA
Split ID	Systematic every 5th
Descriptor subset	Large physchem + fingerprints
Validation	Cross-validation

- > 70 compounds required for QSAR to match predictive capability of FEP (R^2=0.51, RMSE=0.71)
- **FEP is not “fitted”, therefore performance not linked to number of compounds modelled**

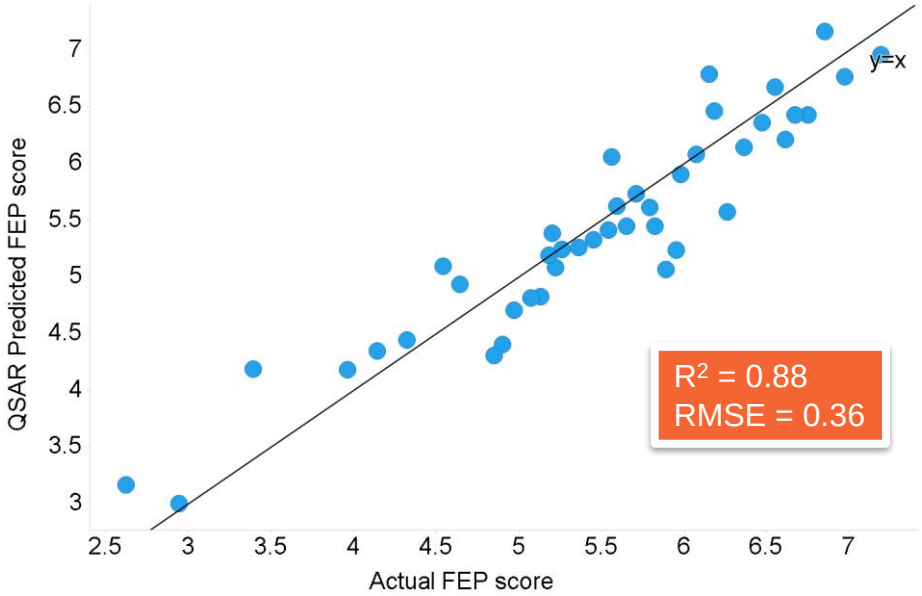
Building a QSAR model on FEP data

Series 4

– Trained/tested on FEP predictions for 215 compounds from Series 4

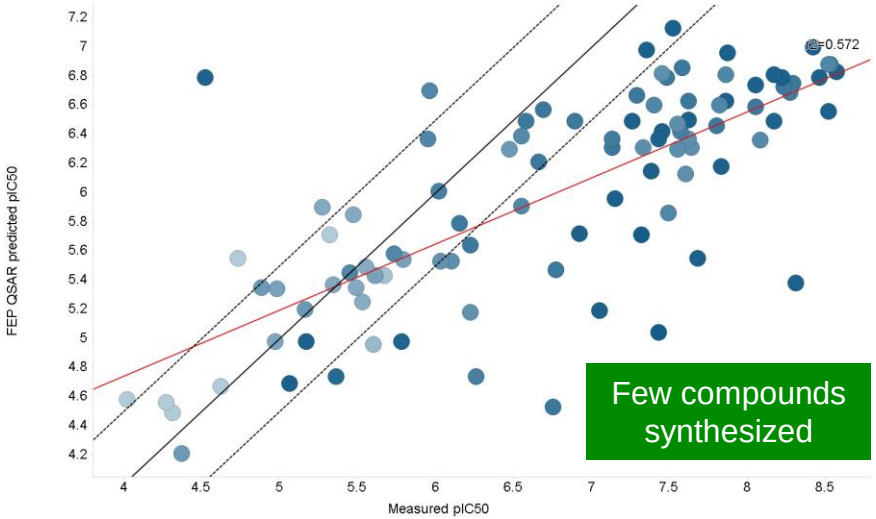
Model type	XG Boost
Split ID	Systematic every 5th
Descriptor subset	Simple physchem with Estates
Validation	Cross-validation

– Should work for similar compounds in this series but not expected to be transferable to other series/diverse compounds



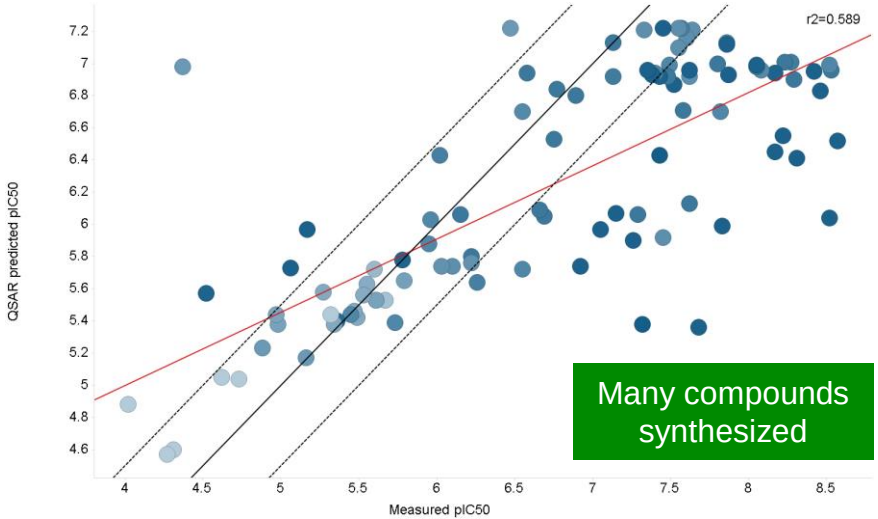
FEP vs. experimental QSAR models

Series 4



FEP QSAR model built on 215 real and virtual compounds

XGBoost – Simple PhysChem descriptor set



QSAR model built on 300 compounds from two “closed form” series

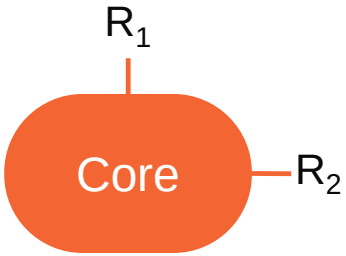
RP Forest, ECFP4 + Large PhysChem descriptor set

Colouring by date compound was tested – white = early, dark blue = late in programme

– **Performance of model degrades with time – potential to apply active learning**

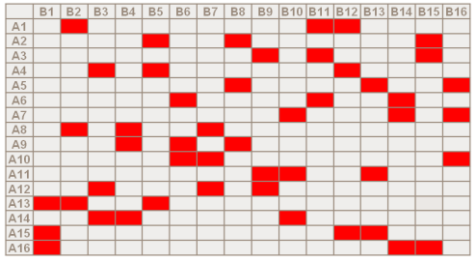
Sparse Array

Incorporating design strategies with FEP+

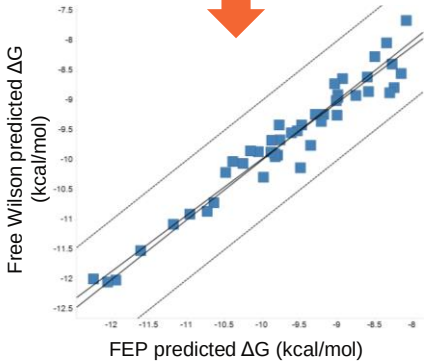


16 different building blocks for R_1 and R_2

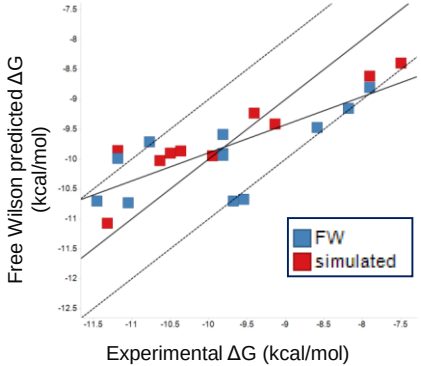
256 possible compounds



Use a sparse array to reduce total number of compounds (48) for FEP+ simulation



Free Wilson analysis used to predict ΔG of remaining compounds

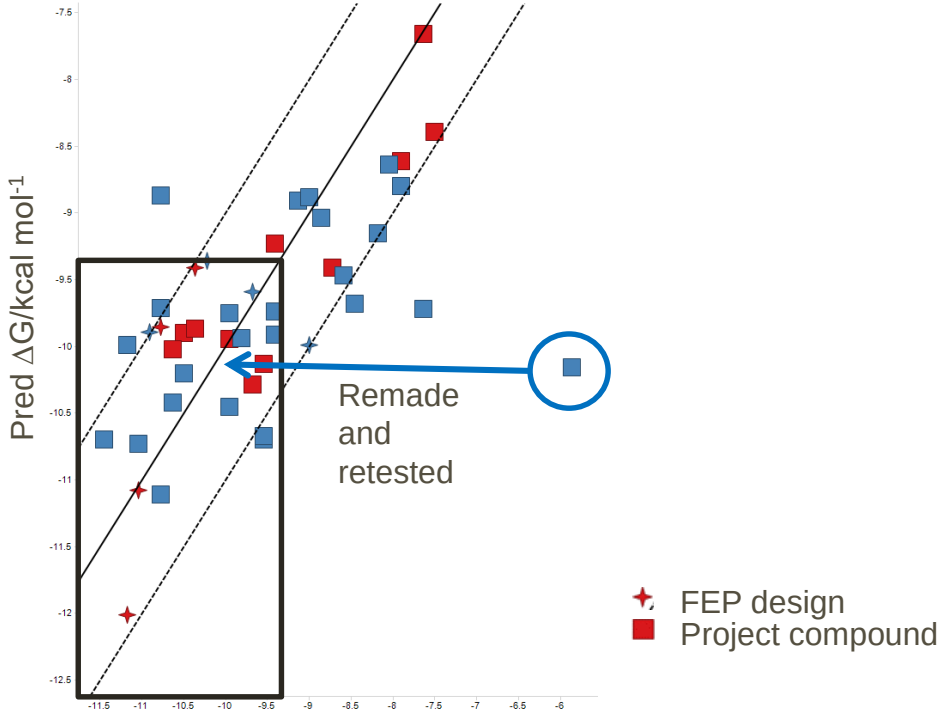


22 compounds synthesized from the Free Wilson and original FEP simulated set

The initial results look promising

Sparse Array

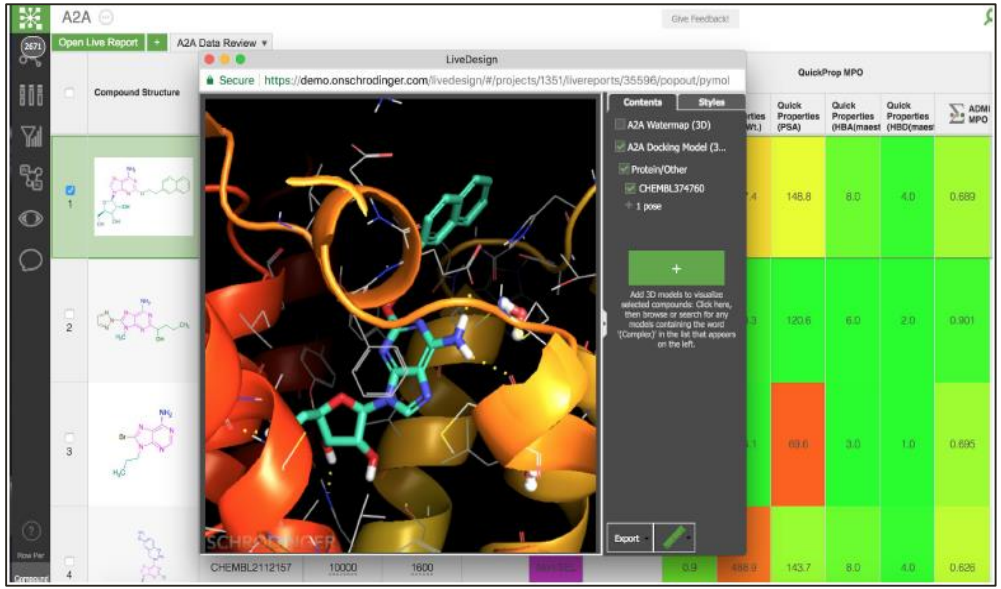
Final results



LiveDesign

Rapid dissemination of FEP+ and QSAR results

- Allow chemistry teams to focus on the ideas
  - Capture, discussion, prioritization
- Curated FEP+ output captured in LD
- Integration with other modeling methods' results
- Deployment of QSAR models
  - Rapidly run on large sets of compounds



Summary

- FEP is the current gold standard for potency prediction
  - But relatively low-throughput
- Building QSAR models on FEP-predicted potency data
  - Increases domain of applicability whilst making fewer compounds
  - Allows fast pre-filtering of large compound sets
  - Potential usage in library design, e.g. sparse-array
- Future usage may benefit from Active Learning

Acknowledgements

GSK:

- Xiao Lewell
- Ian Wall
- Darren Green
- Jacob Bush
- Rob Young
- Chris Luscombe
- Alan Graves
- Molecular Design group

Schrodinger:

- Dan Robinson
- Davide Branduardi