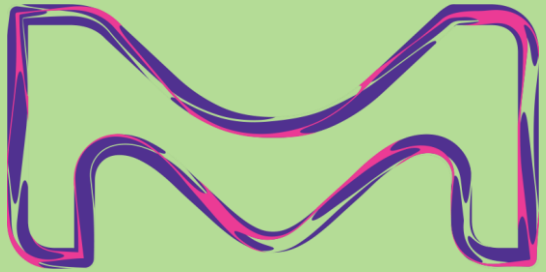
opportunities and challenges for free energy calculations in drug design

Christina Schindler
CompBioMed 2019



MERCK

Acknowledgements

MedChem Boston & Darmstadt

CompChem Boston & Darmstadt

- Theresa Johnson
- Alejandro Crespo
- Liwei Li
- Daniel Kuhn
- Jakub Gunera
- Paul Czodrowski
- Mireille Krier
- Jakub Gunera
- Hannah Baumann
- Robert Schulz
- Merveille Eguida

Schrödinger

- Daniel Cappel
- Thomas Steinbrecher
- Jörg Weiser
- Thijs Beuming
- Olivia Pierce

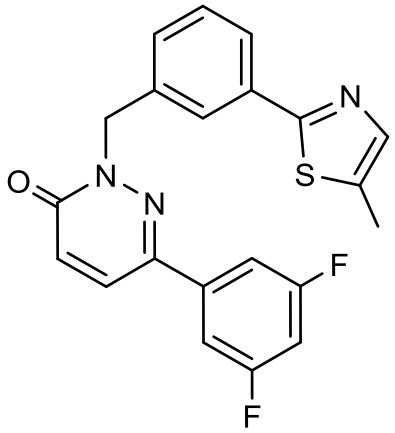
Digitize Merck team

- Friedrich Rippmann
- Thomas Fürst

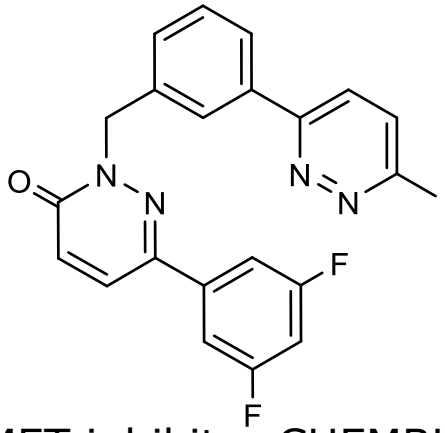
MSKCC

- Levi Naden
- John Chodera

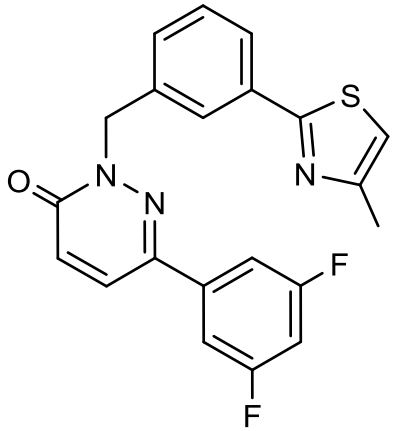
Which compounds to synthesize next?
The holy grail of drug design: Prediction of binding constants



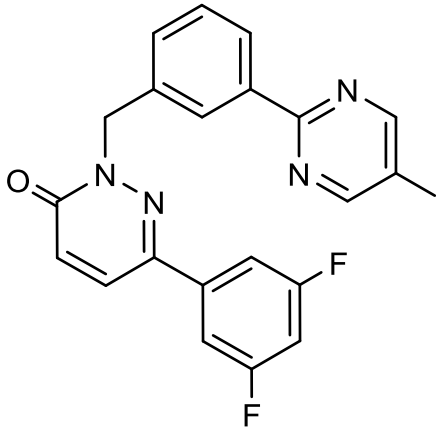
c-MET inhibitor CHEMBL3402750



c-MET inhibitor CHEMBL3402755



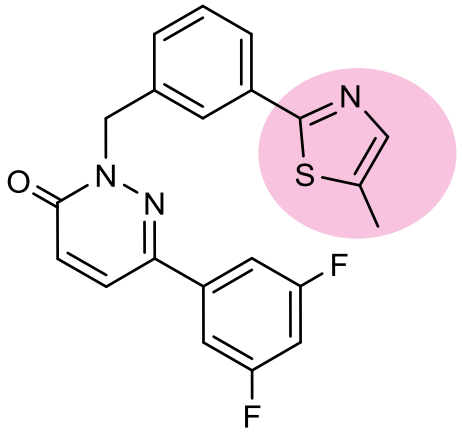
c-MET inhibitor CHEMBL3402751



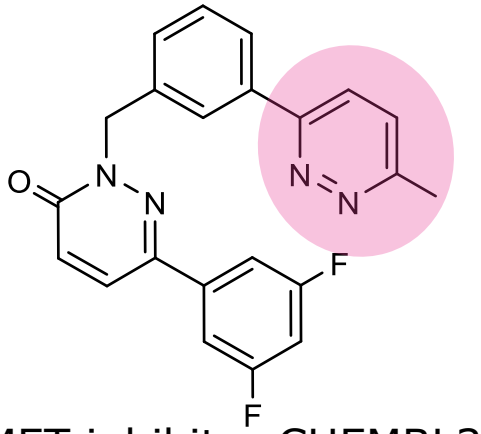
c-MET inhibitor CHEMBL3402754

Which compounds to synthesize next?

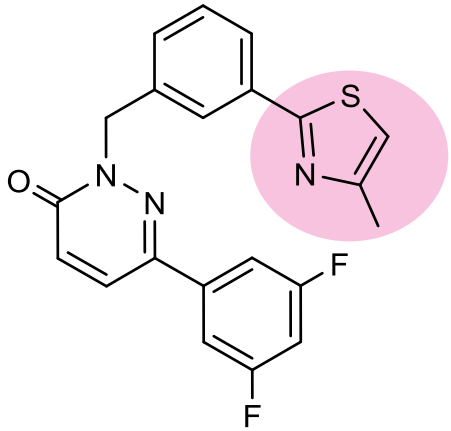
The holy grail of drug design: Prediction of binding constants



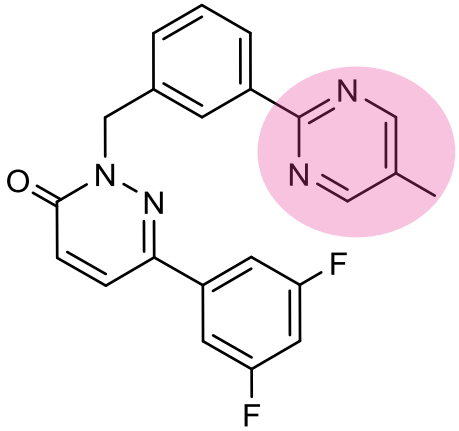
c-MET inhibitor CHEMBL3402750



c-MET inhibitor CHEMBL3402755



c-MET inhibitor CHEMBL3402751

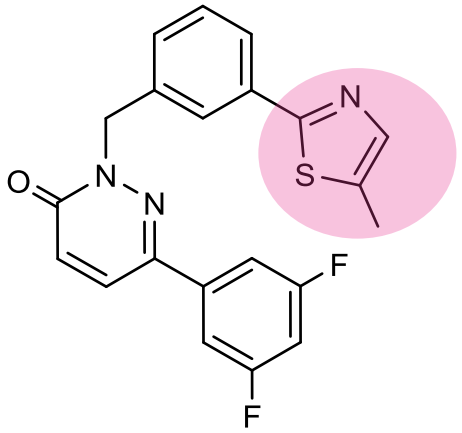


c-MET inhibitor CHEMBL3402754

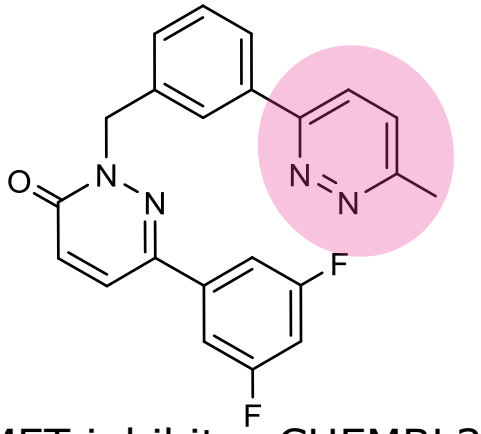
MERCK

Which compounds to synthesize next?

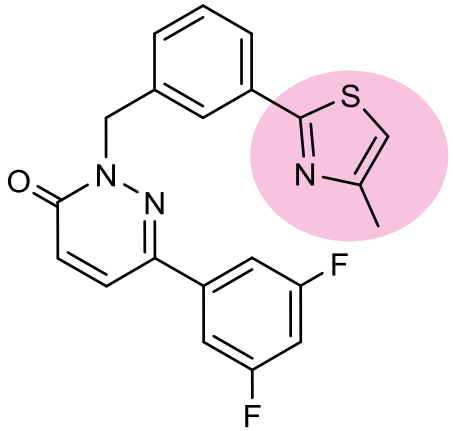
The holy grail of drug design: Prediction of binding constants



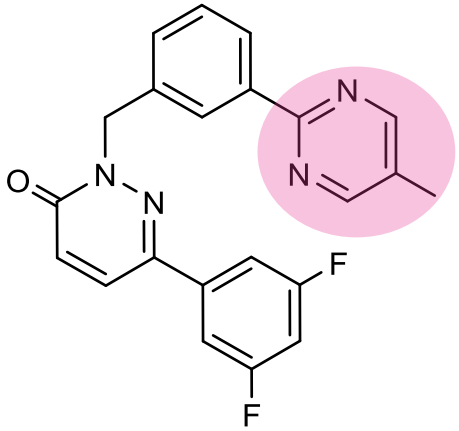
c-MET inhibitor CHEMBL3402750 (400 nM)



c-MET inhibitor CHEMBL3402755 (4200 nM)



c-MET inhibitor CHEMBL3402751 (2100 nM)

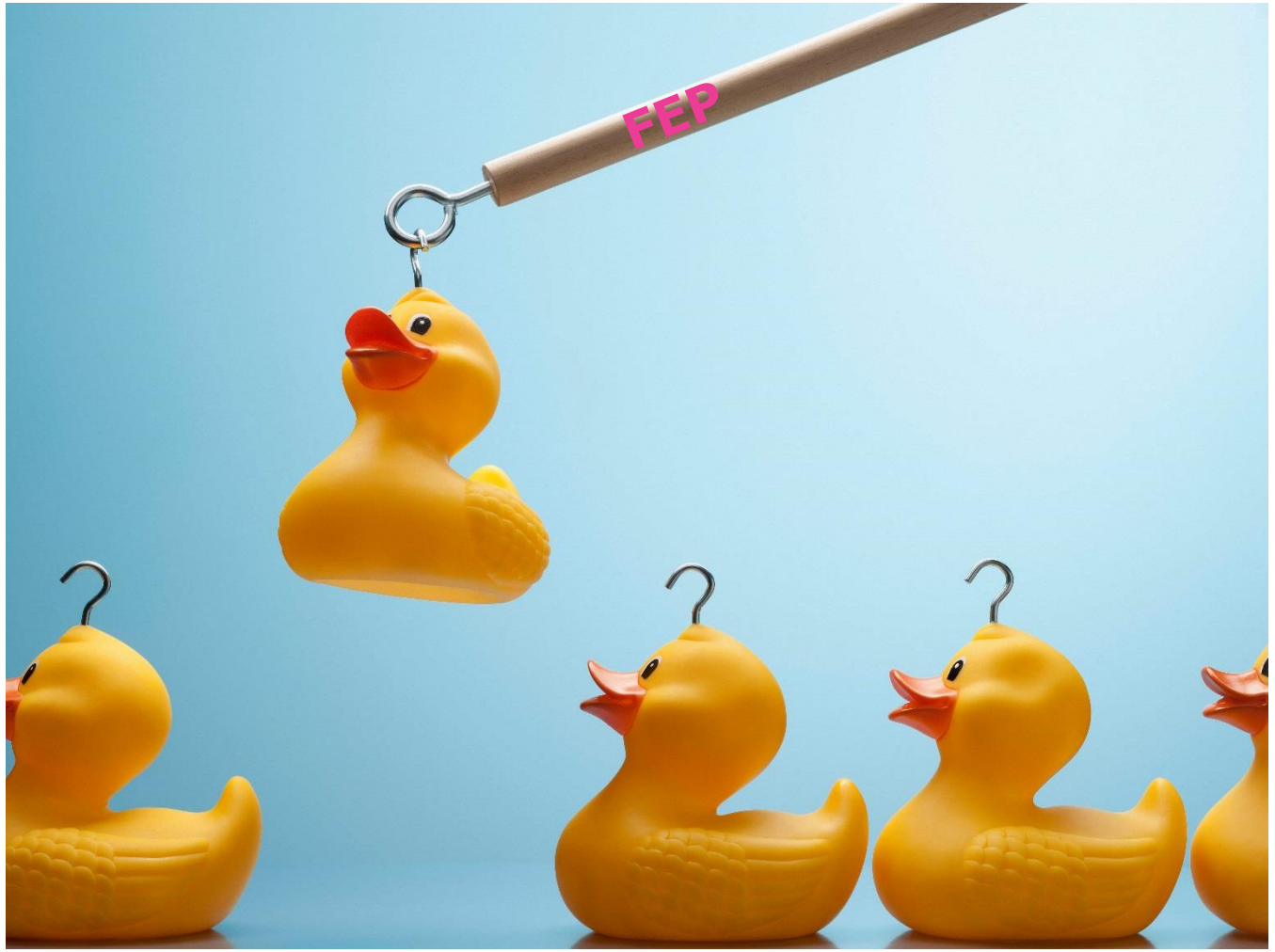


c-MET inhibitor CHEMBL3402754 (40 nM)

MERCK

Finding promising molecules

Opportunities for FEP in drug discovery projects



Finding promising molecules
Opportunities for FEP in drug discovery projects

Focus on potent molecules by prioritizing ideas for synthesis

Finding promising molecules
Opportunities for FEP in drug discovery projects

Focus on potent molecules by prioritizing ideas for synthesis

Increase structural diversity and broadly explore chemical space with large libraries

MERCK

Finding promising molecules

Opportunities for FEP in drug discovery projects

Focus on potent molecules by prioritizing ideas for synthesis

Increase structural diversity and broadly explore chemical space with large libraries

De-risk challenging synthesis e.g. for scaffold hops

Finding the right balance

Challenges for FEP in drug discovery projects

Finding the right balance
Challenges for FEP in drug discovery projects

Scientific:

Achieve sufficient prediction accuracy (<1.4 kcal/mol)

Finding the right balance
Challenges for FEP in drug discovery projects

Technical:

Manage large-scale computations on specialized hardware

Scientific:

Achieve sufficient prediction accuracy (<1.4 kcal/mol)

Finding the right balance
Challenges for FEP in drug discovery projects

Technical:

Manage large-scale computations on specialized hardware

Scientific:

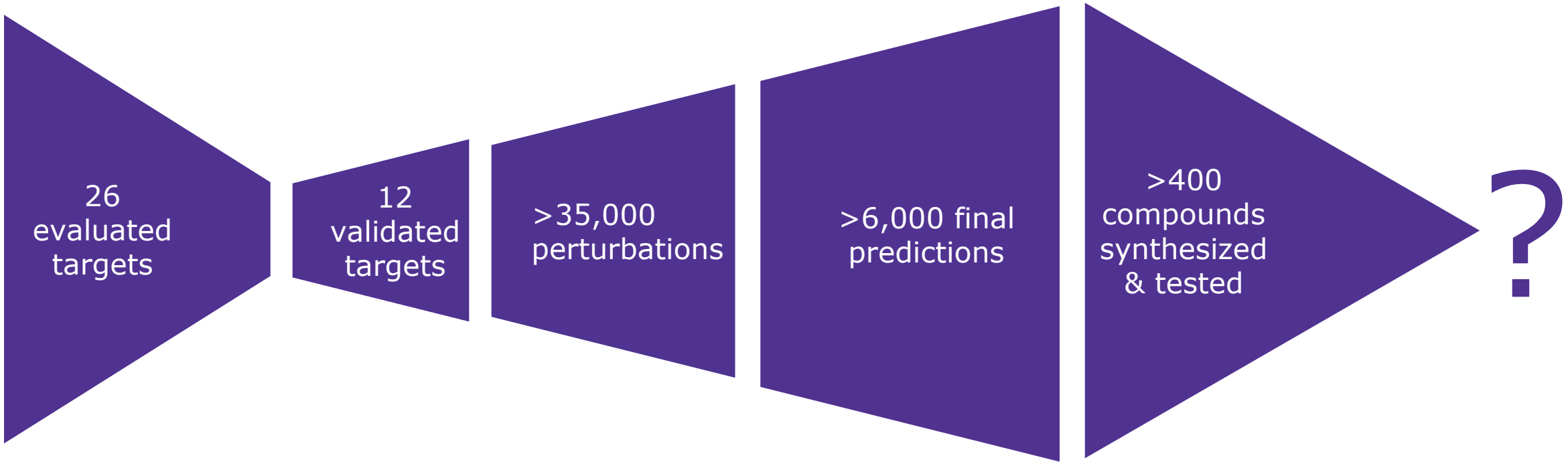
Achieve sufficient prediction accuracy (<1.4 kcal/mol)

Operational:
Communication, timing and impact in projects

MERCK

Broad application across multiple targets and series

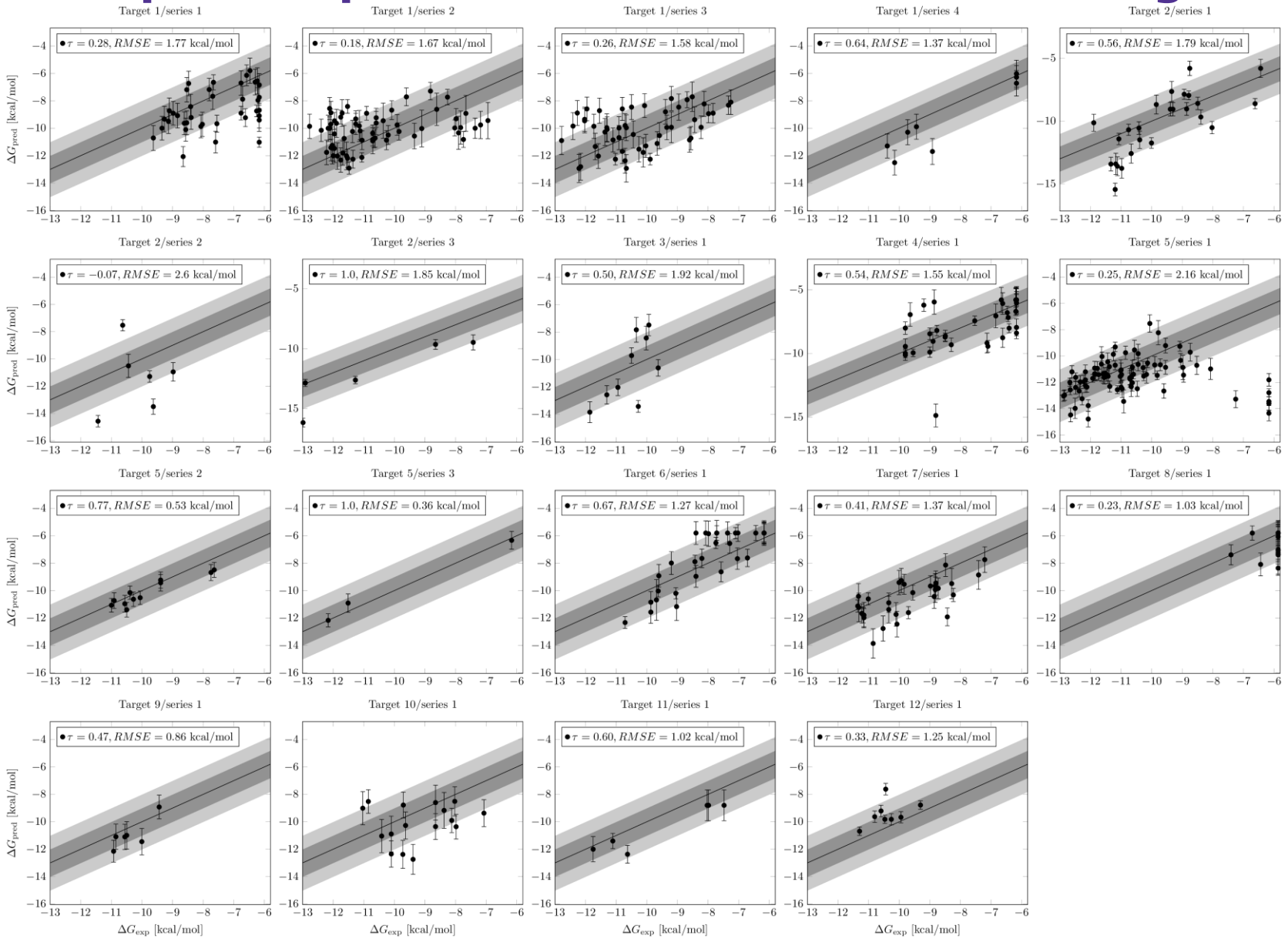
FEP+ in drug discovery at Merck KGaA, Darmstadt, Germany



- **How does FEP perform?**
- **What are the learnings?**
- **What is the impact on projects?**

Prospective data for 19 chemical series available

Prospective performance varies between targets and series



N = 463

Average Kendall tau = 0.37

Average RMSE = 1.64 kcal/mol

Cohen's d for R2

- Glide: 0.76
- Prime: 0.74

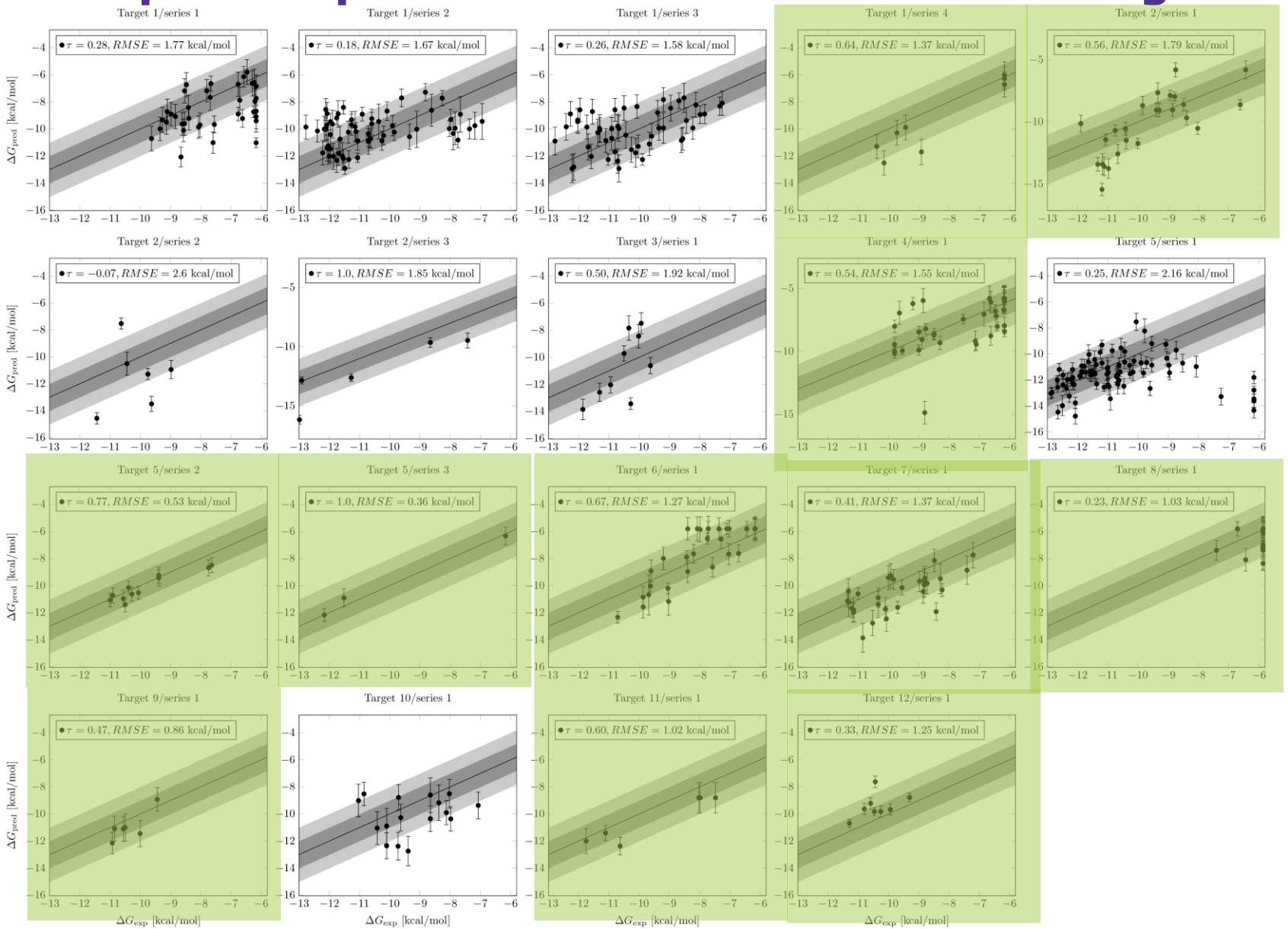
Challenges

- Predictions often not possible in certain parts of the molecule
- Transformation from short R-group to long, flexible chains
- Changes in net charge and charge distribution

MERCK

Prospective data for 19 chemical series available

Prospective performance varies between targets and series



N = 463

Average Kendall tau = 0.37

Average RMSE = 1.64 kcal/mol

Cohen's d for R2

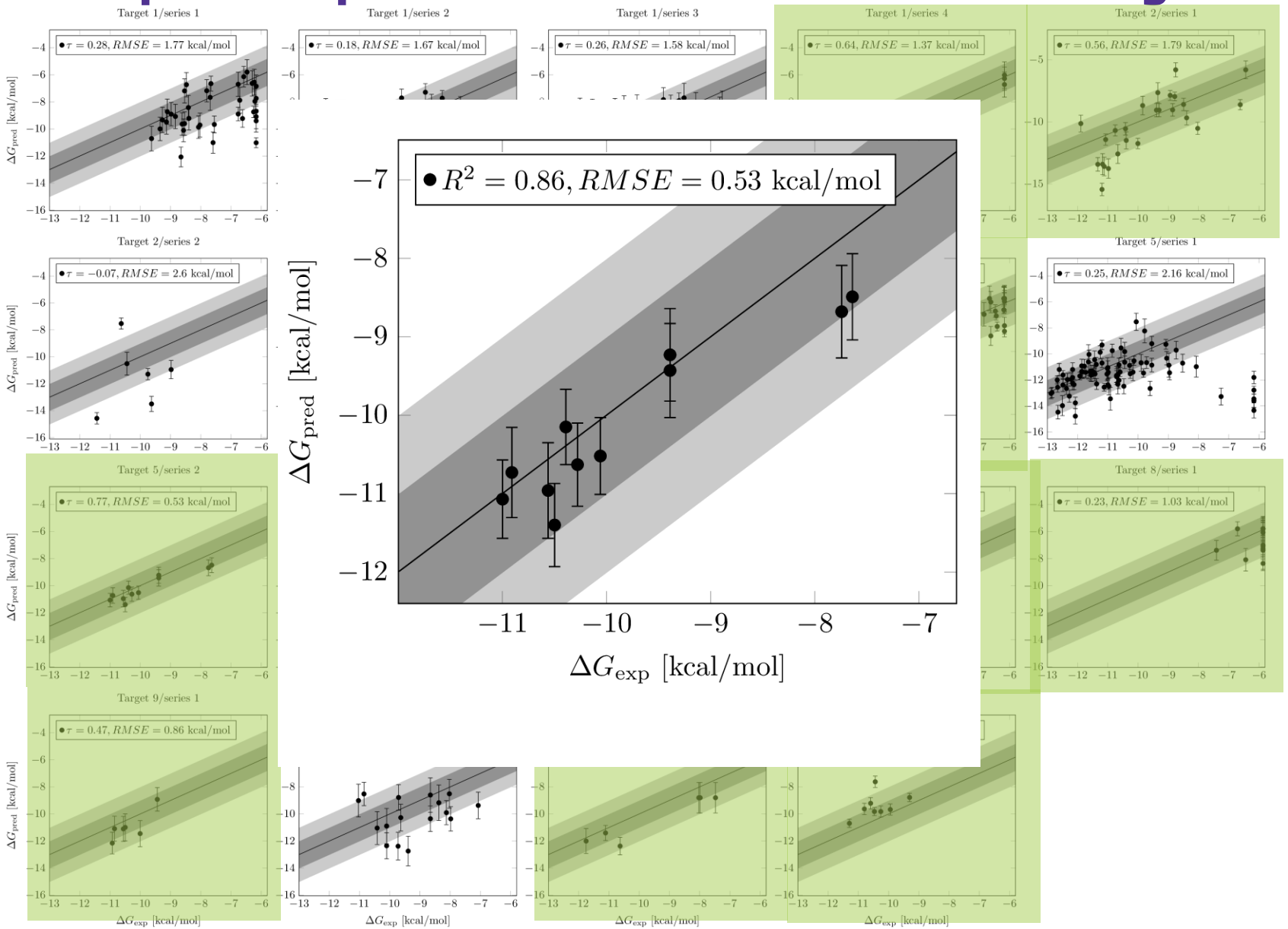
- Glide: 0.76
- Prime: 0.74

Challenges

- Predictions often not possible in certain parts of the molecule
- Transformation from short R-group to long, flexible chains
- Changes in net charge and charge distribution

Prospective data for 19 chemical series available

Prospective performance varies between targets and series



N = 463

Average Kendall tau = 0.37

Average RMSE = 1.64 kcal/mol

Cohen's d for R2

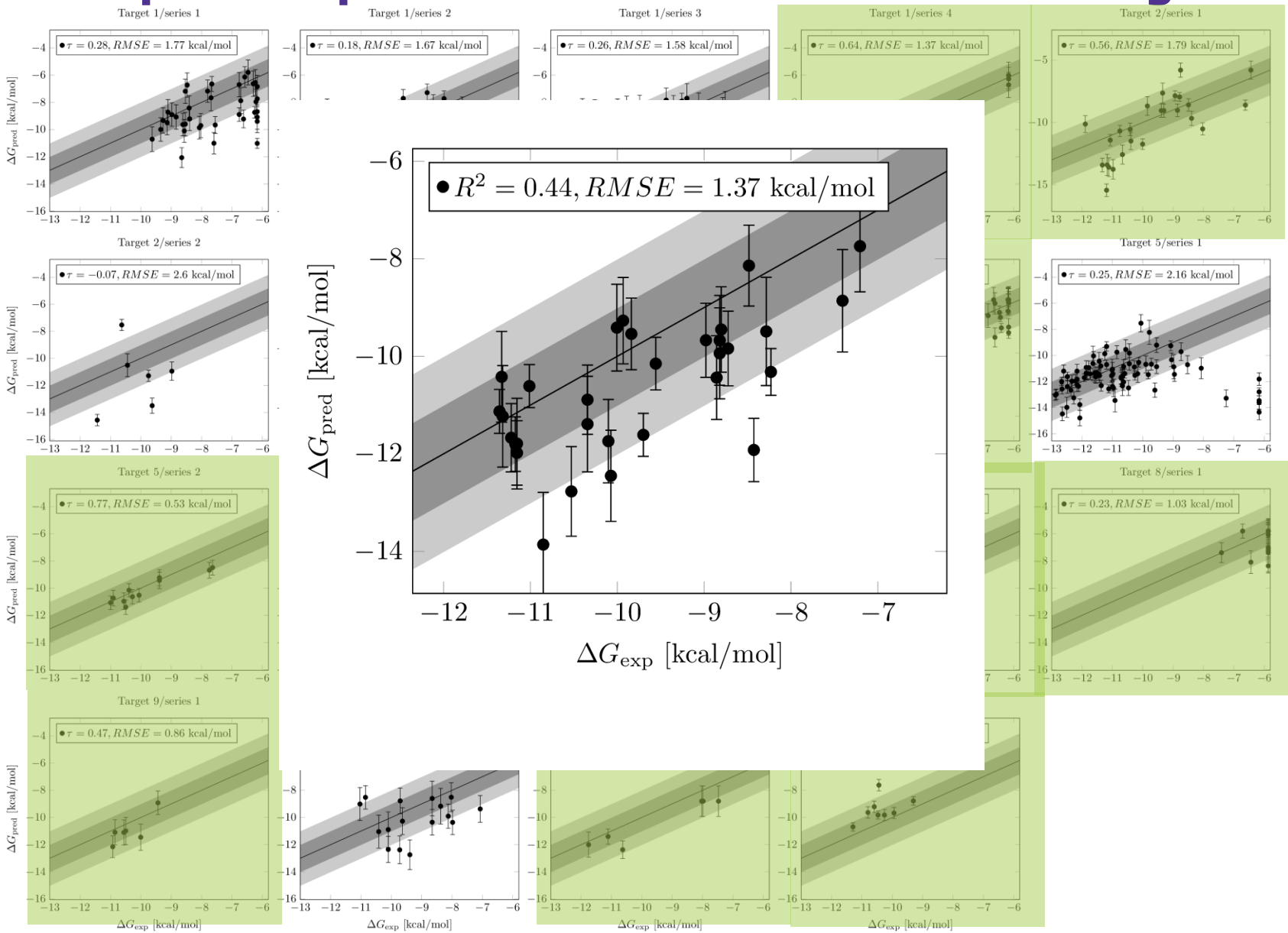
- Glide: 0.76
- Prime: 0.74

Challenges

- Predictions often not possible in certain parts of the molecule
- Transformation from short R-group to long, flexible chains
- Changes in net charge and charge distribution

Prospective data for 19 chemical series available

Prospective performance varies between targets and series



N = 463

Average Kendall tau = 0.37

Average RMSE = 1.64 kcal/mol

Cohen's d for R2

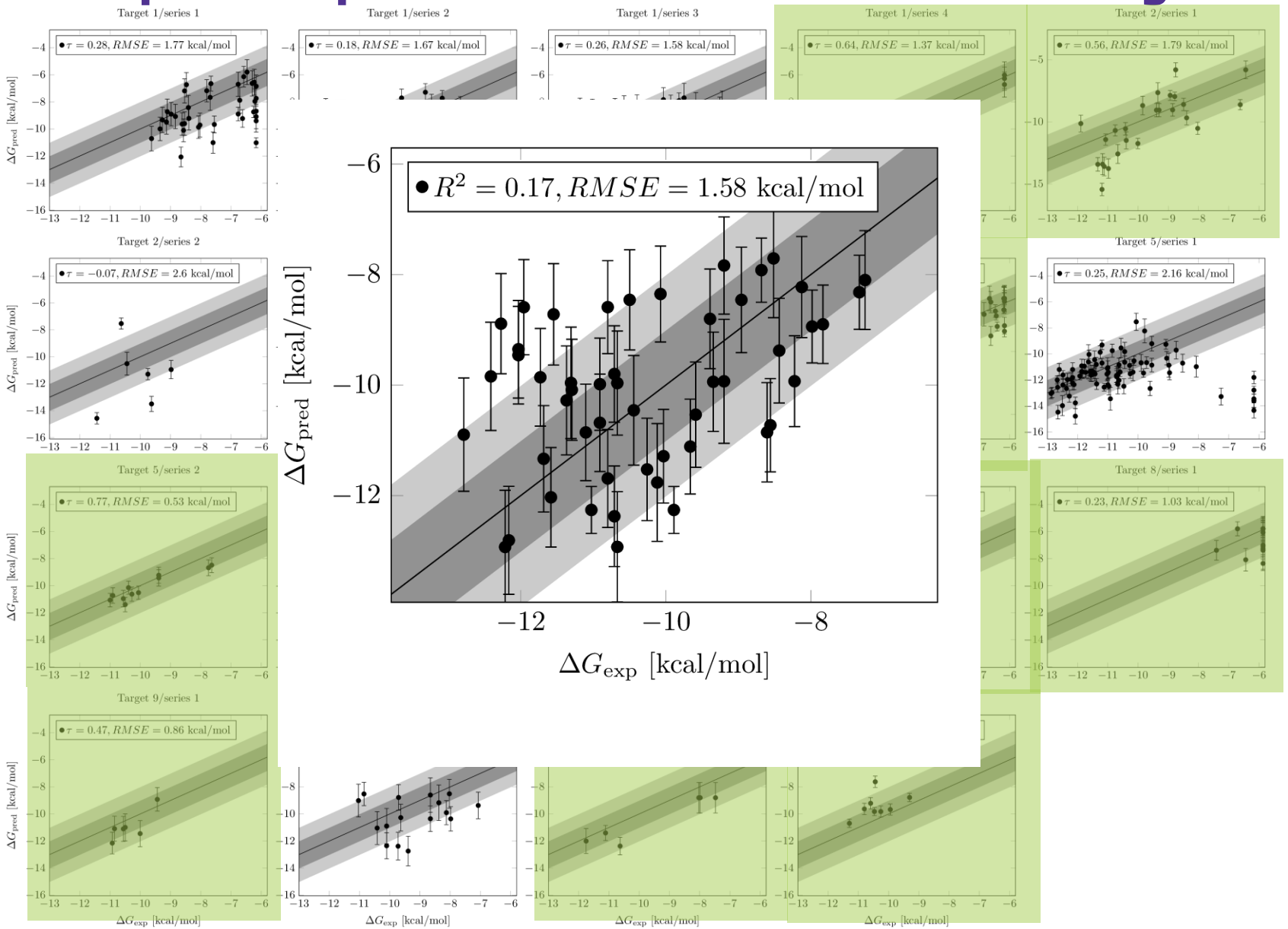
- Glide: 0.76
- Prime: 0.74

Challenges

- Predictions often not possible in certain parts of the molecule
- Transformation from short R-group to long, flexible chains
- Changes in net charge and charge distribution

Prospective data for 19 chemical series available

Prospective performance varies between targets and series



N = 463

Average Kendall tau = 0.37

Average RMSE = 1.64 kcal/mol

Cohen's d for R2

- Glide: 0.76
- Prime: 0.74

Challenges

- Predictions often not possible in certain parts of the molecule
- Transformation from short R-group to long, flexible chains
- Changes in net charge and charge distribution

264 ligands for eight pharmaceutically relevant targets
New benchmark for free energy calculations created

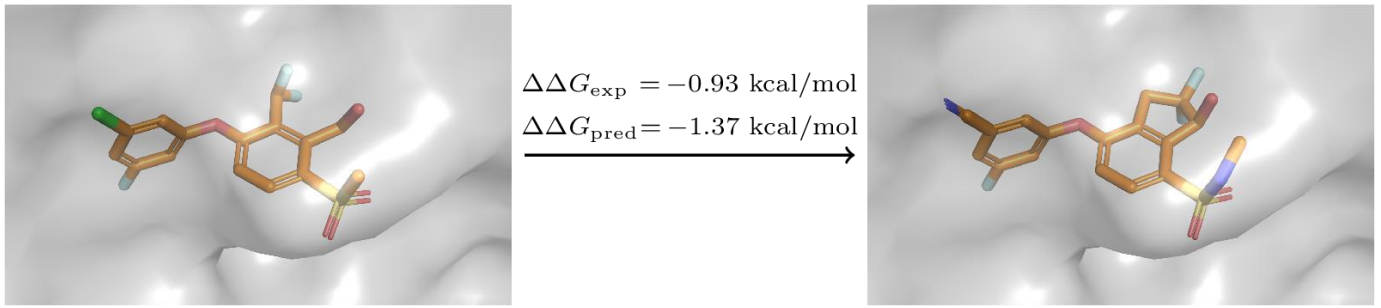
Recent ligand-target pairs collected from literature

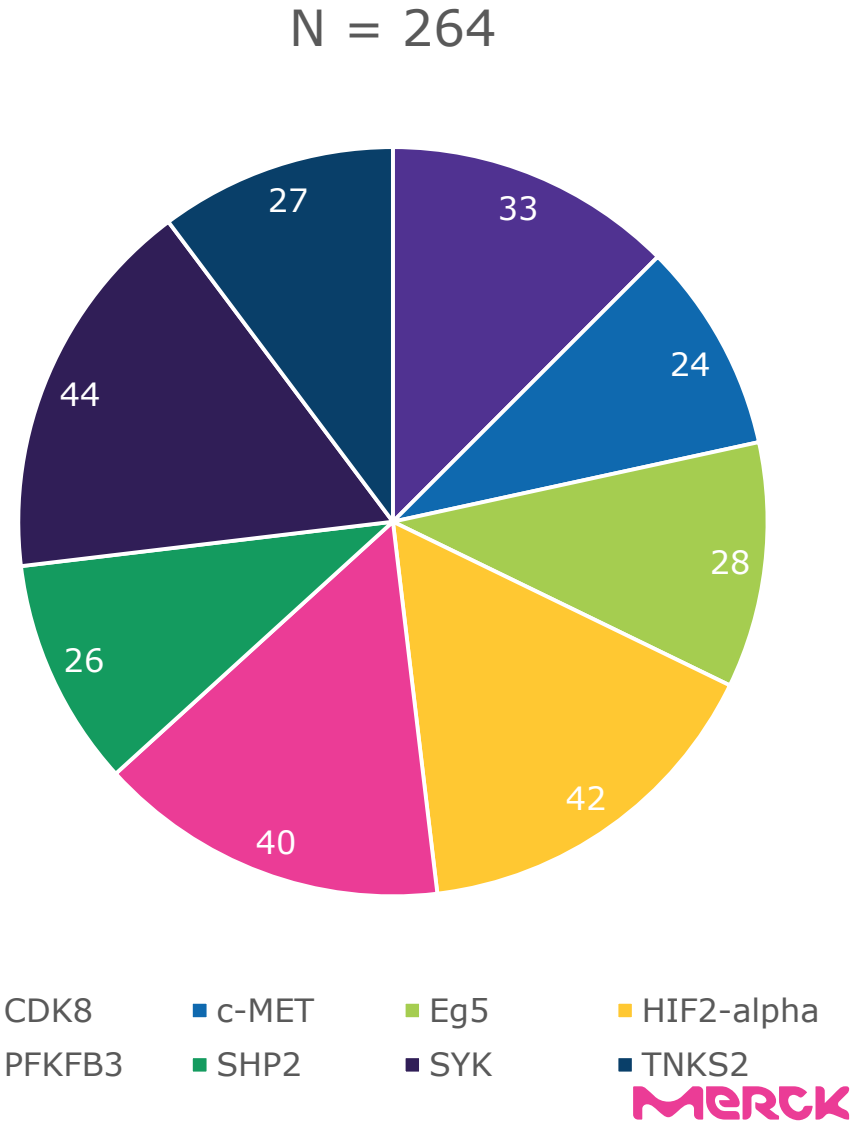
Diverse chemical transformations representative of modifications in (early) compound optimization

Challenging transformations to test methodological advances in FEP+: charge changes and ring openings

Available on github:

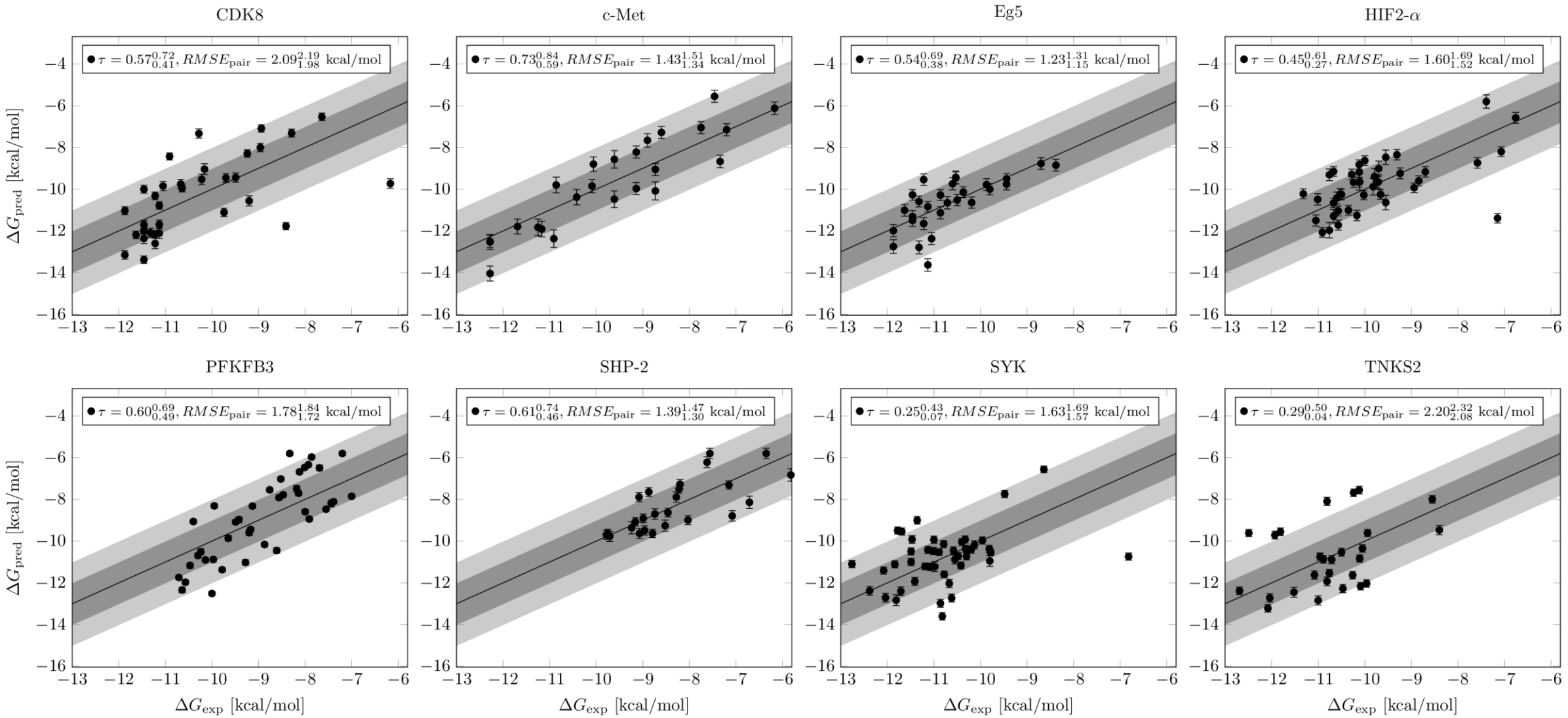
github.com/MCompChem/fep-benchmark



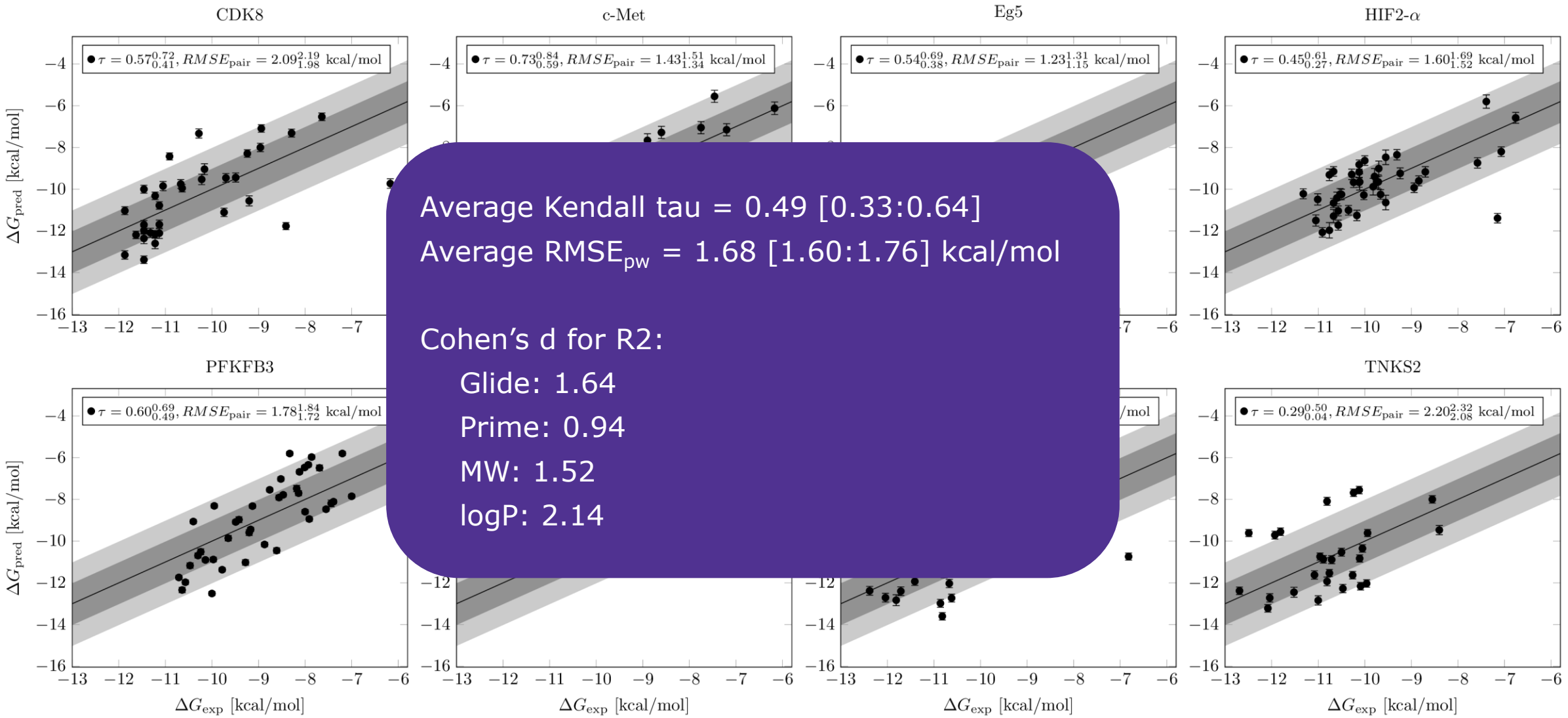


Large public benchmark with eight pharmaceutically relevant targets

Good performance on challenging data set



Large public benchmark with eight pharmaceutically relevant targets
Good performance on challenging data set



Average Kendall tau = 0.49 [0.33:0.64]

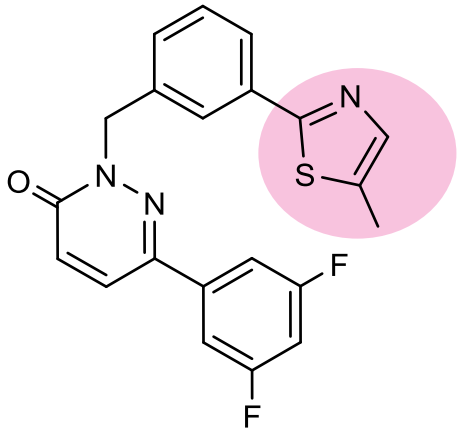
Average $RMSE_{pw}$ = 1.68 [1.60:1.76] kcal/mol

Cohen's d for R2:

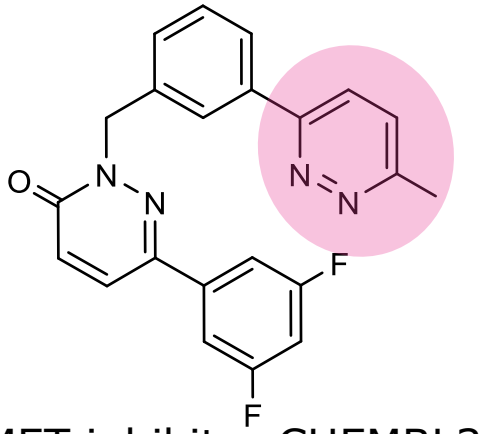
- Glide: 1.64
- Prime: 0.94
- MW: 1.52
- logP: 2.14

Which compounds to synthesize next?

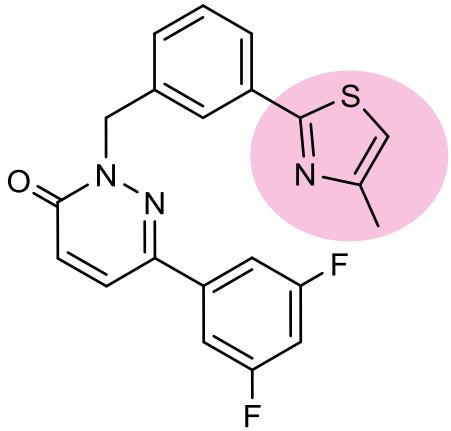
The holy grail of drug design: Prediction of binding constants



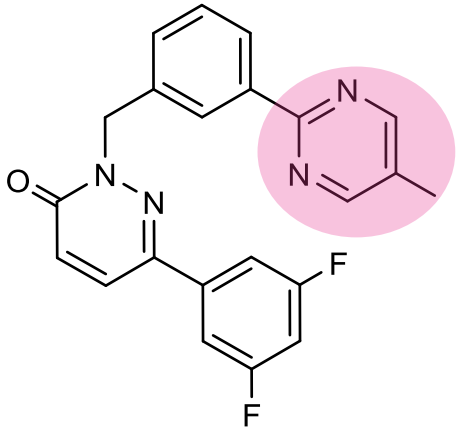
c-MET inhibitor CHEMBL3402750



c-MET inhibitor CHEMBL3402755



c-MET inhibitor CHEMBL3402751

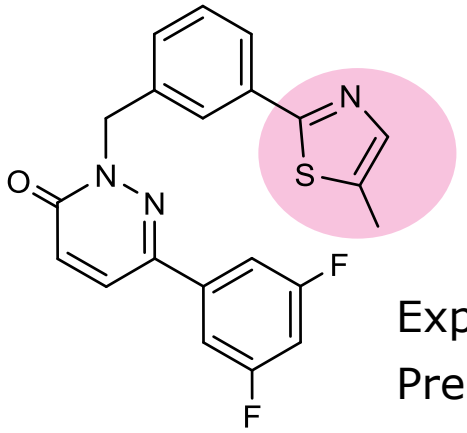


c-MET inhibitor CHEMBL3402754

MERCK

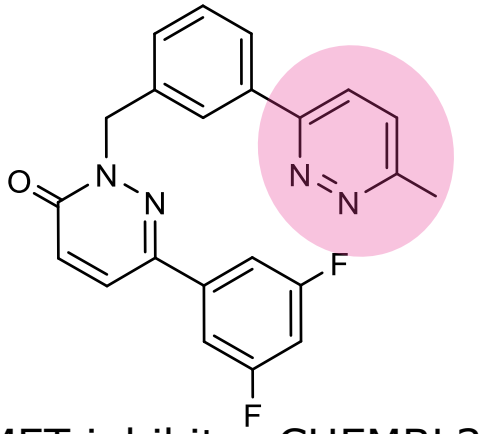
Which compounds to synthesize next?

The holy grail of drug design: Prediction of binding constants



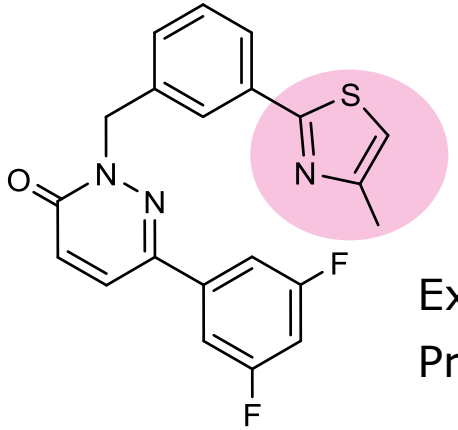
Exp. IC50 = 400 nM
Pred. IC50 = 232 nM

c-MET inhibitor CHEMBL3402750



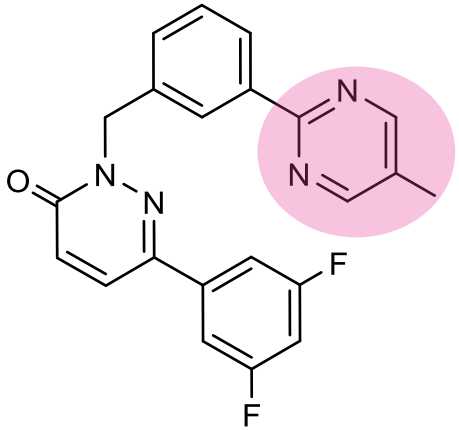
Exp. IC50 = 4200 nM
Pred. IC50 = 441 nM

c-MET inhibitor CHEMBL3402755



Exp. IC50 = 2100 nM
Pred. IC50 = 5020 nM

c-MET inhibitor CHEMBL3402751



Exp. IC50 = 40 nM
Pred. IC50 = 60 nM

c-MET inhibitor CHEMBL3402754

MERCK

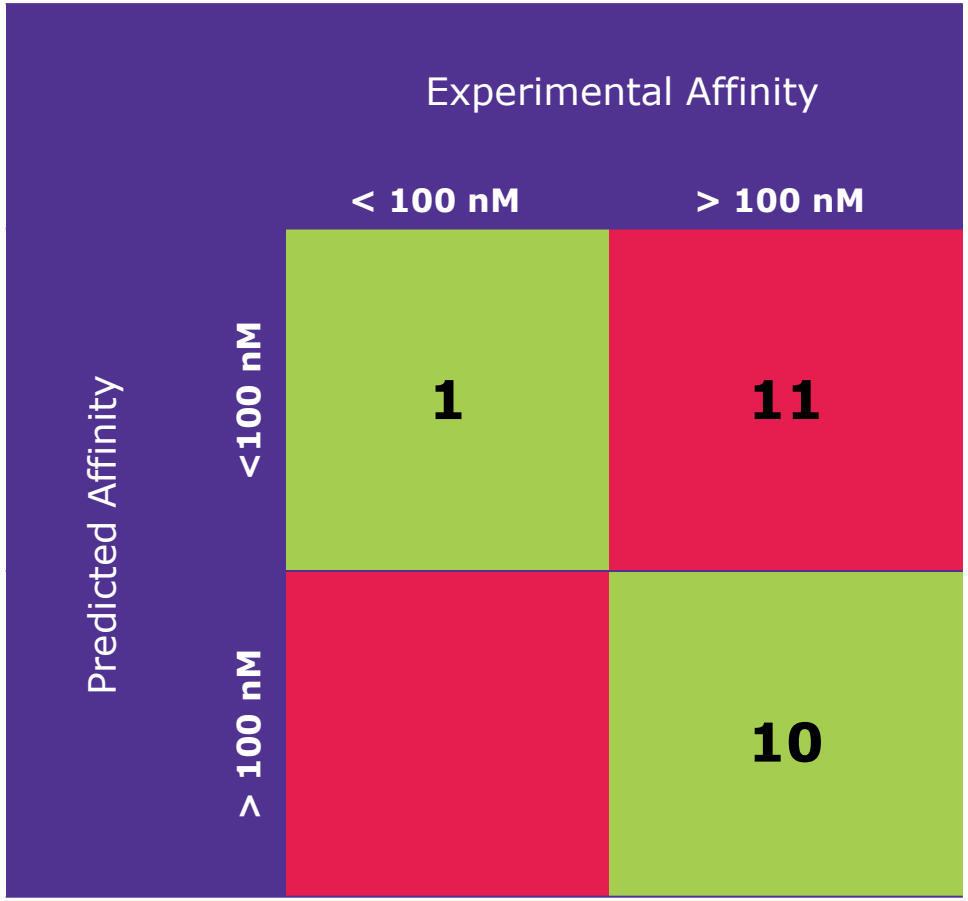
We can predict binding affinity with good accuracy…

…but how to best use it?

Use cases for FEP in projects

Prioritize synthesis proposals & focus on promising molecules

22 compounds synthesized and tested



Scenario:

Continuous evaluation of synthesis proposals and prioritization with FEP

Benefit:

Save synthesis resources spent on inactive molecules

Problems:

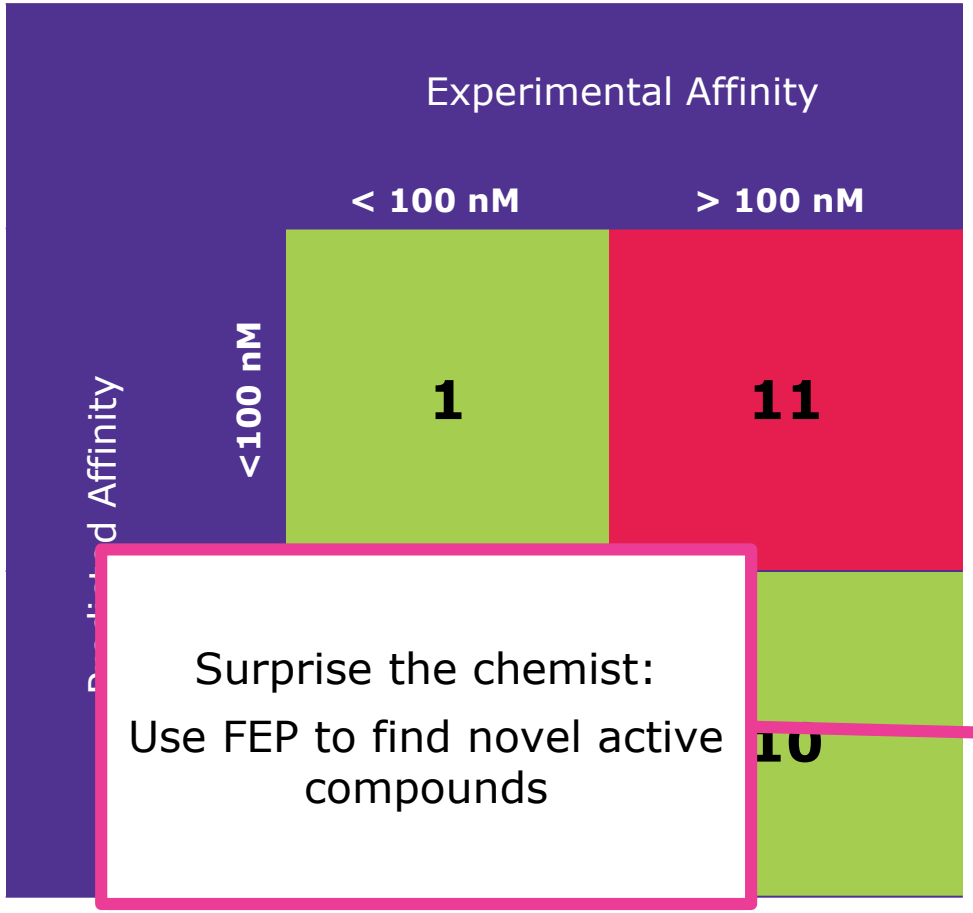
- Number of compounds proposed often not much larger than compounds synthesized
- Chemists don’t want to wait for FEP predictions
- After 6 months, nobody will remember the bad molecules that were not made

→ Project impact limited

Use cases for FEP in projects

Prioritize synthesis proposals & focus on promising molecules

22 compounds synthesized and tested



Surprise the chemist:

Use FEP to find novel active compounds

Scenario:

Continuous evaluation of synthesis proposals and prioritization with FEP

Benefit:

Save synthesis resources spent on inactive molecules

Problems:

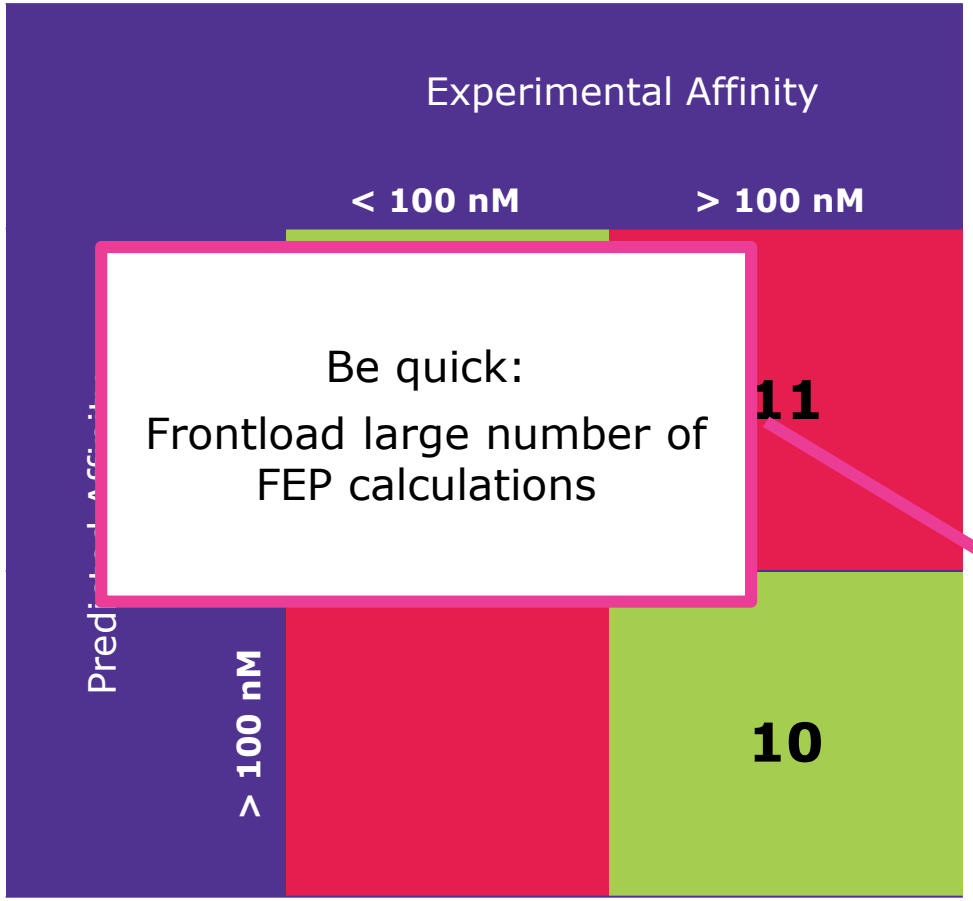
- Number of compounds proposed often not much larger than compounds synthesized
- Chemists don't want to wait for FEP predictions
- After 6 months, nobody will remember the bad molecules that were not made

→ Project impact limited

Use cases for FEP in projects

Prioritize synthesis proposals & focus on promising molecules

22 compounds synthesized and tested



Scenario:

Continuous evaluation of synthesis proposals and prioritization with FEP

Benefit:

Save synthesis resources spent on inactive molecules

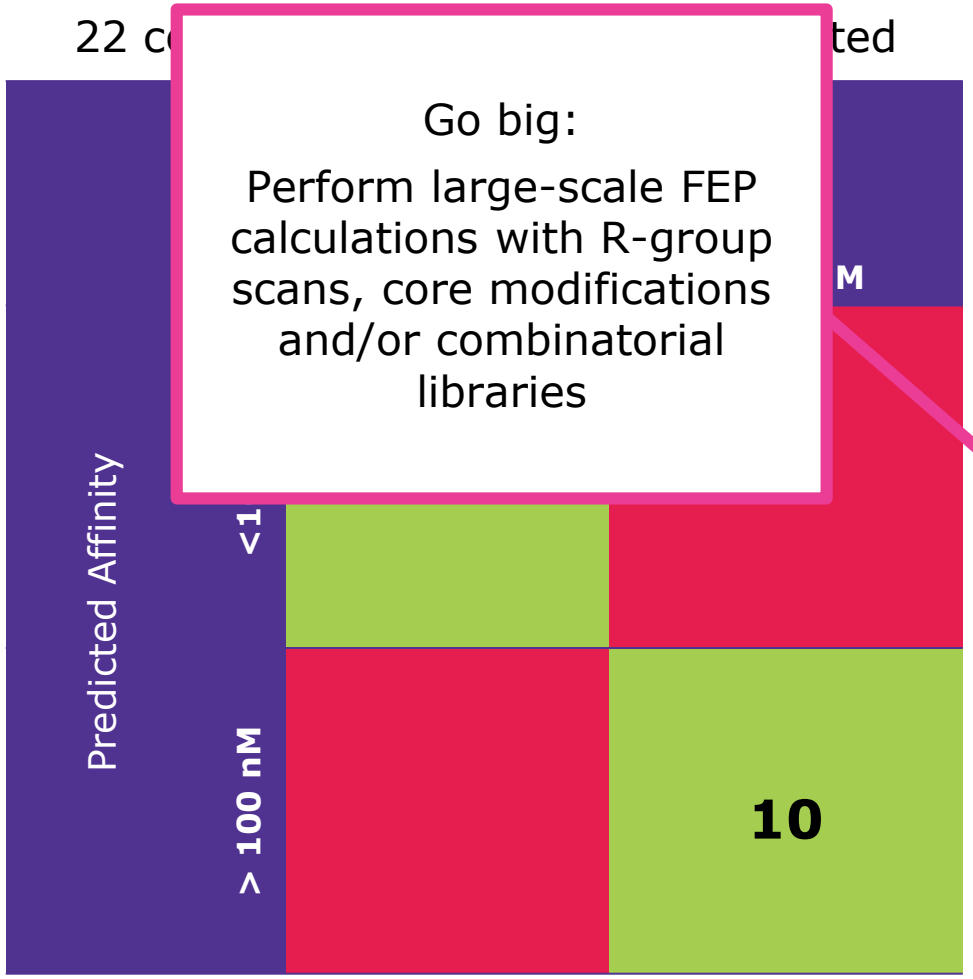
Problems:

- Number of compounds proposed often not much larger than compounds synthesized
- Chemists don’t want to wait for FEP predictions
- After 6 months, nobody will remember the bad molecules that were not made

→ Project impact limited

Use cases for FEP in projects

Prioritize synthesis proposals & focus on promising molecules



Go big:

Perform large-scale FEP calculations with R-group scans, core modifications and/or combinatorial libraries

Scenario:

Continuous evaluation of synthesis proposals and prioritization with FEP

Benefit:

Save synthesis resources spent on inactive molecules

Problems:

- Number of compounds proposed often not much larger than compounds synthesized
- Chemists don’t want to wait for FEP predictions
- After 6 months, nobody will remember the bad molecules that were not made

→ Project impact limited

Chemical space exploration with FEP
Prioritizing promising compounds from a large library

Jakub Gunera

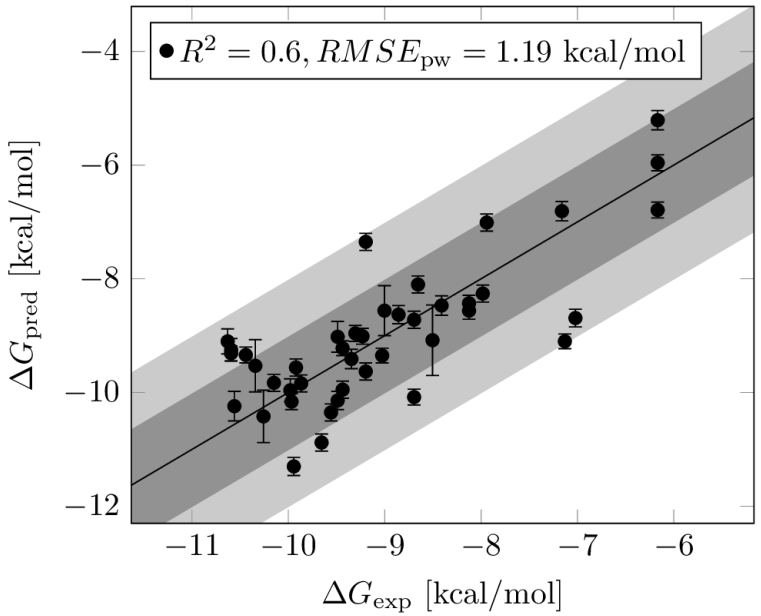
IC50 = 92 nM

Chemical space exploration with FEP
Prioritizing promising compounds from a large library

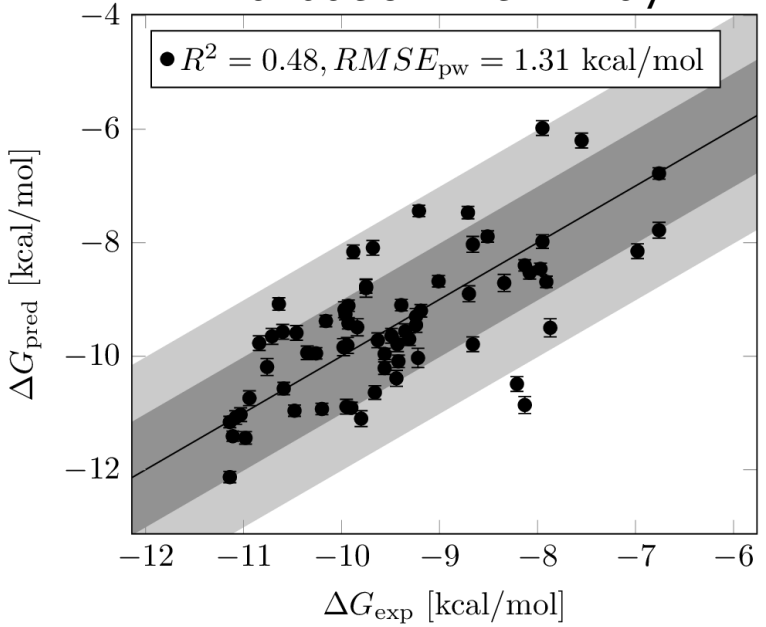
Jakub Gunera

IC50 = 92 nM

Validation with docking pose



Validation with X-ray



Chemical space exploration with FEP
Prioritizing promising compounds from a large library

Jakub Gunera

IC50 = 92 nM

variation of R-groups

Library constructed from commercially available building blocks

15,600 cpds

MERCK

Chemical space exploration with FEP
Prioritizing promising compounds from a large library

Jakub Gunera

IC50 = 92 nM

variation of R-groups

Library constructed from commercially available building blocks

15,600 cpds

Filtering by properties and substructures

307 cpds

MERCK

Chemical space exploration with FEP
Prioritizing promising compounds from a large library

Jakub Gunera

IC50 = 92 nM

variation of R-groups

Library constructed from commercially available building blocks

15,600 cpds

→

Filtering by properties and substructures

307 cpds

→

FEP

20 cpds selected

Synthesis on-going

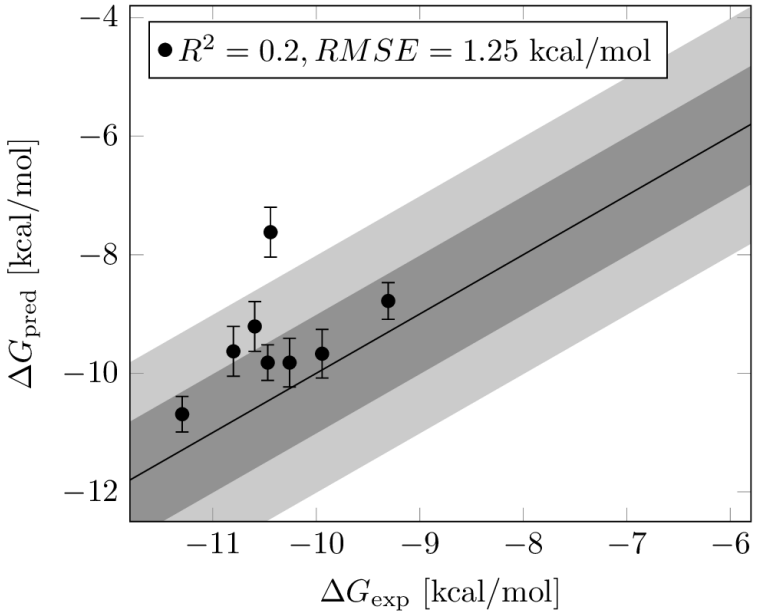
MERCK

Chemical space exploration with FEP
Prioritizing promising compounds from a large library

Jakub Gunera

IC50 = 92 nM

Prospective



variation of R-groups

Library constructed from commercially available building blocks

15,600 cpds

Filtering by properties and substructures

307 cpds

FEP

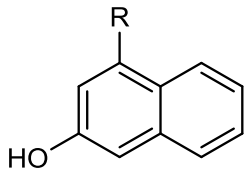
20 cpds selected

Synthesis on-going

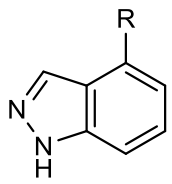
Library scanning with covalent FEP

Replacement for unwanted R-group needed

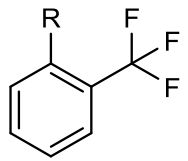
Known SAR



IC50 = 15 nM

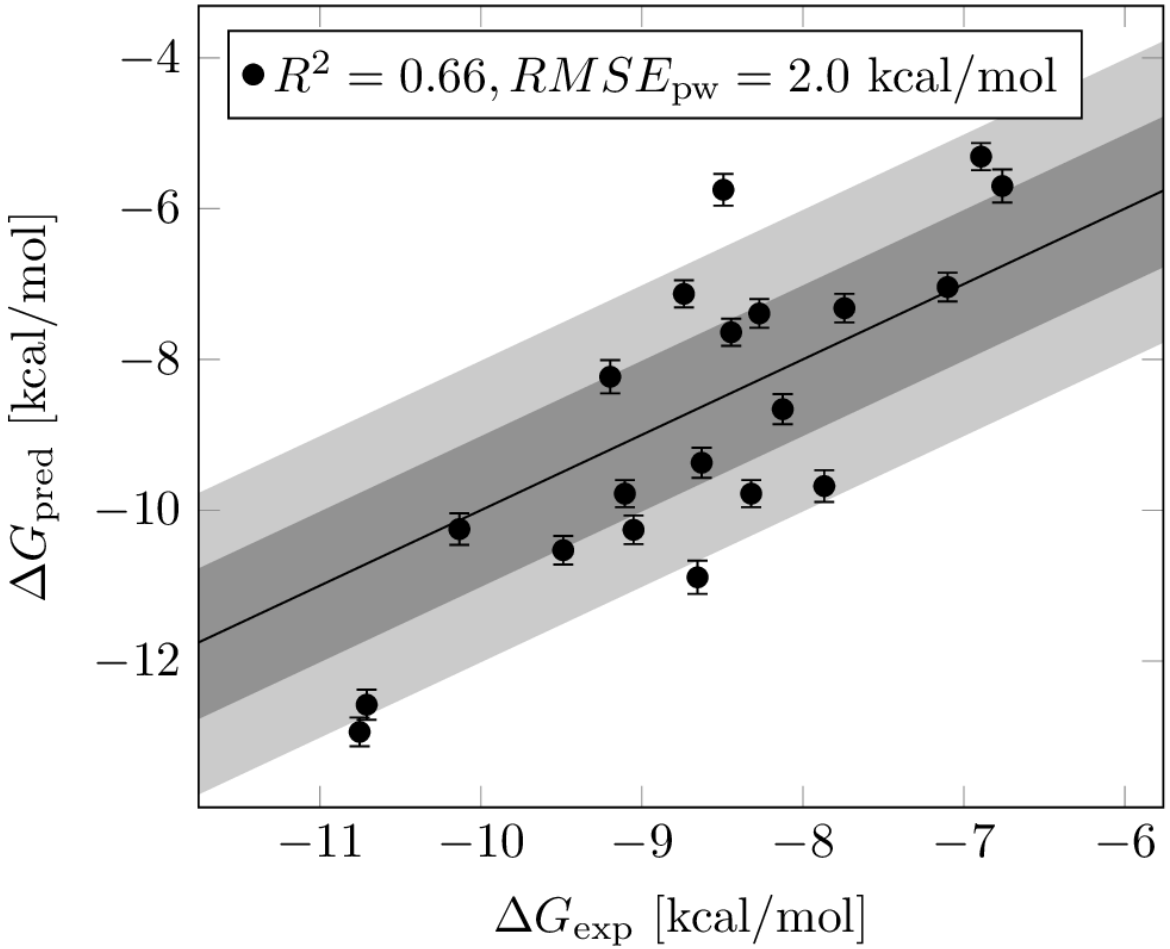


IC50 = 230 nM

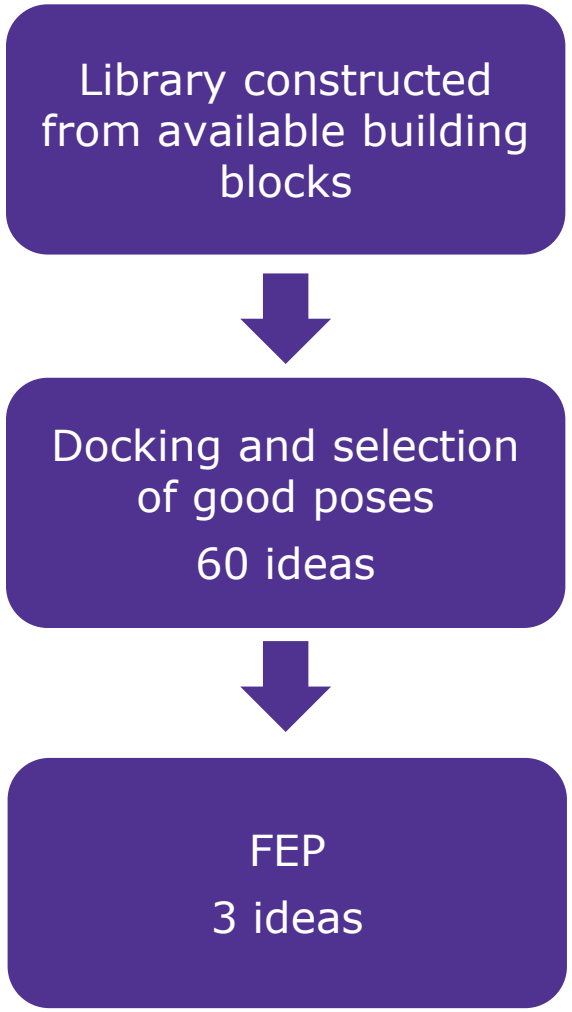


IC50 = 2000 nM

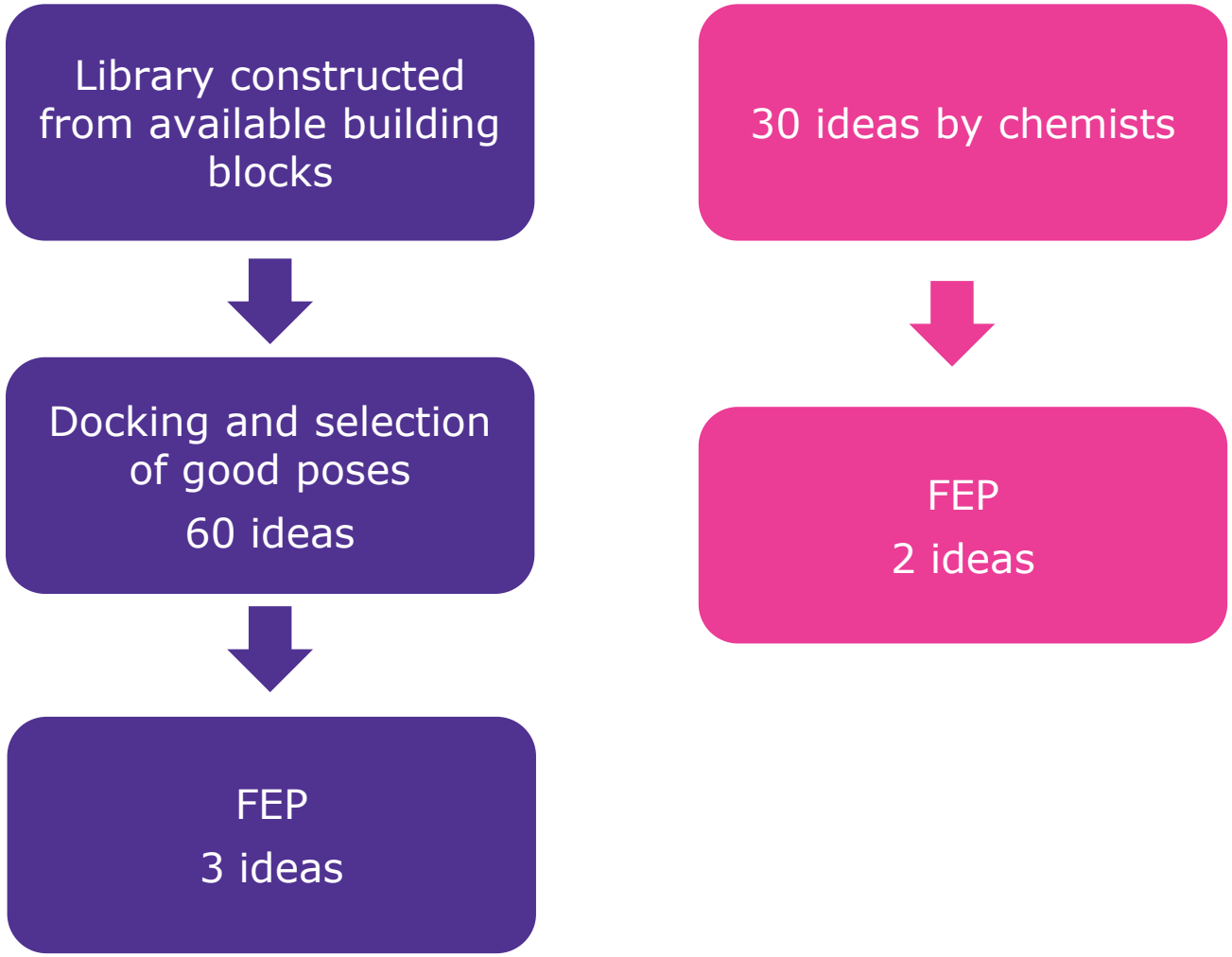
FEP validation study



Library scanning with covalent FEP
Successful replacement of unwanted R-group



Library scanning with covalent FEP
Successful replacement of unwanted R-group



Library scanning with covalent FEP
Successful replacement of unwanted R-group

Library constructed from available building blocks

↓

Docking and selection of good poses

60 ideas

↓

FEP

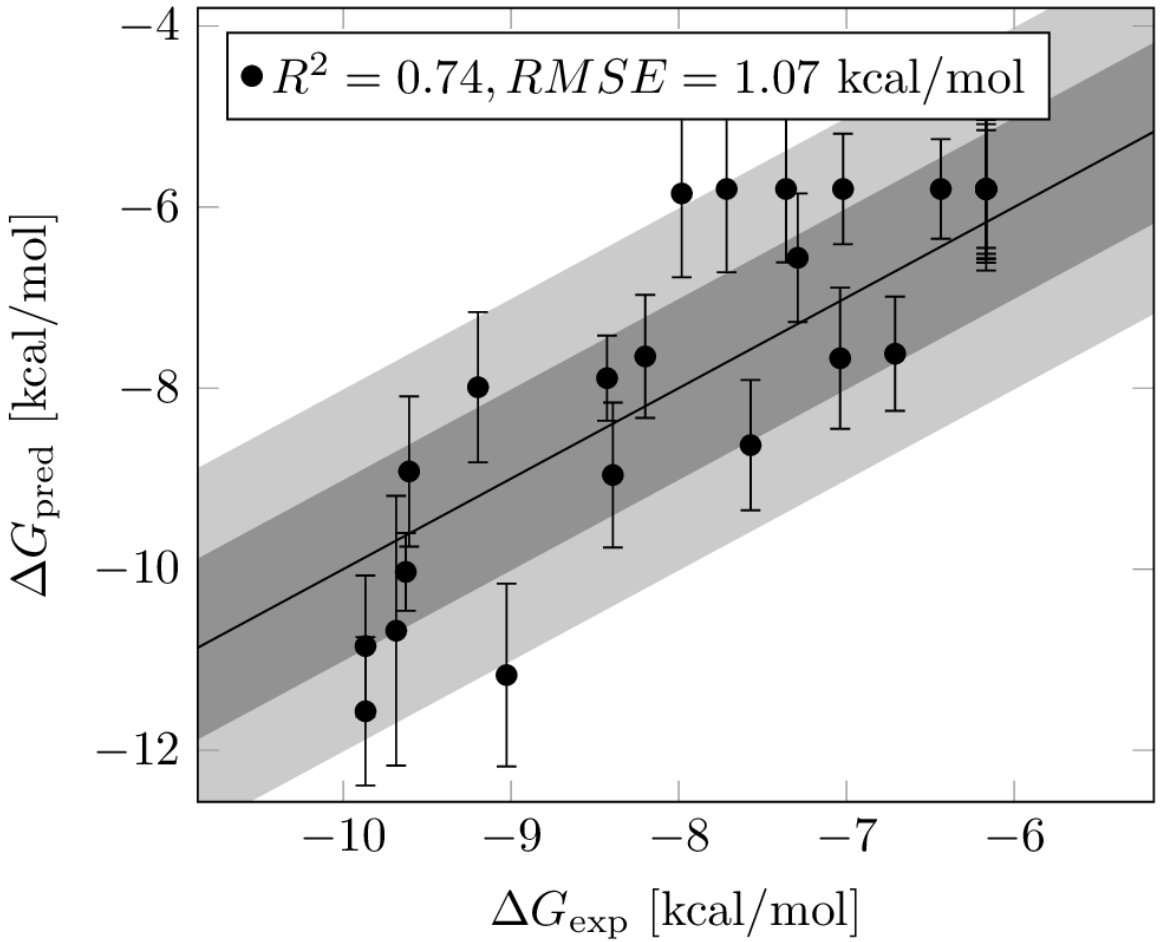
3 ideas

30 ideas by chemists

↓

FEP

2 ideas



Library scanning with covalent FEP
Successful replacement of unwanted R-group

Library constructed from available building blocks

↓

Docking and selection of good poses

60 ideas

↓

FEP

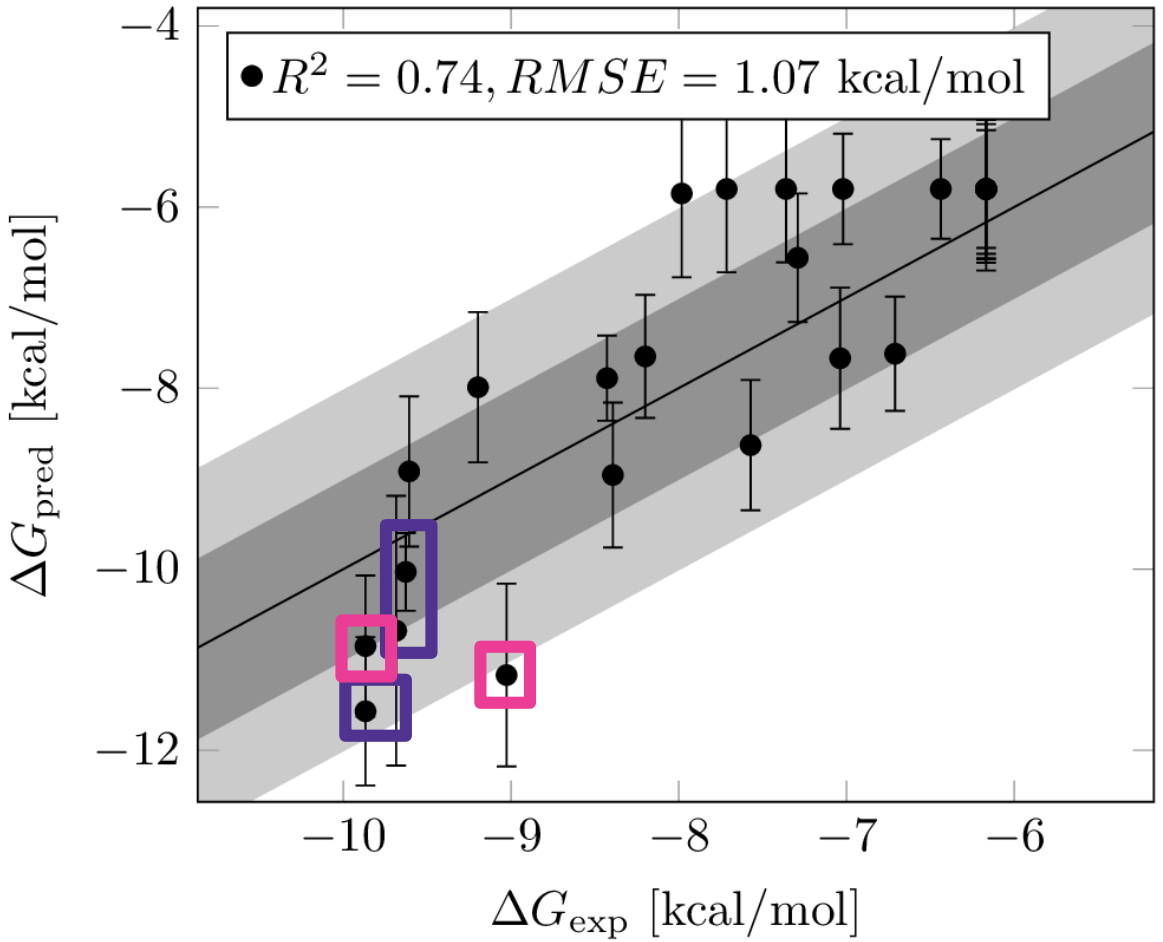
3 ideas

30 ideas by chemists

↓

FEP

2 ideas



Library scanning with covalent FEP
Successful replacement of unwanted R-group

Library constructed from available building blocks

↓

Docking and selection of good poses

60 ideas

↓

FEP

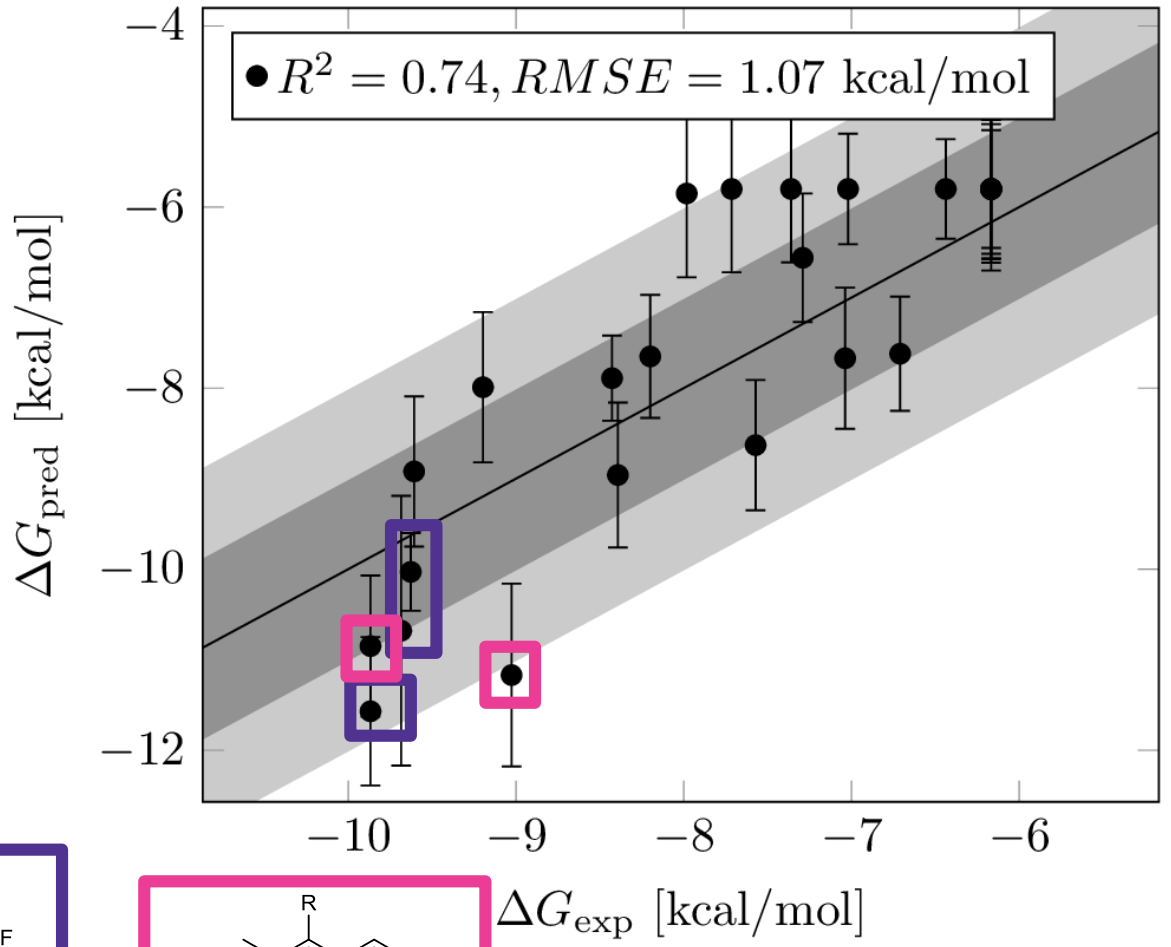
3 ideas

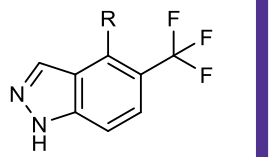
30 ideas by chemists

↓

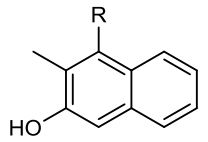
FEP

2 ideas





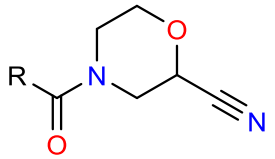
IC50 = 58 nM



IC50 = 79 nM

Discovery of new chemical starting points with FEP

From fragment to hit: Proof-of-principle for in silico optimization

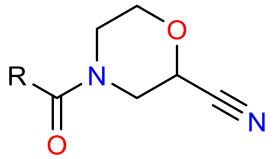


SPR KDss = 300 µM

LE = 0.25

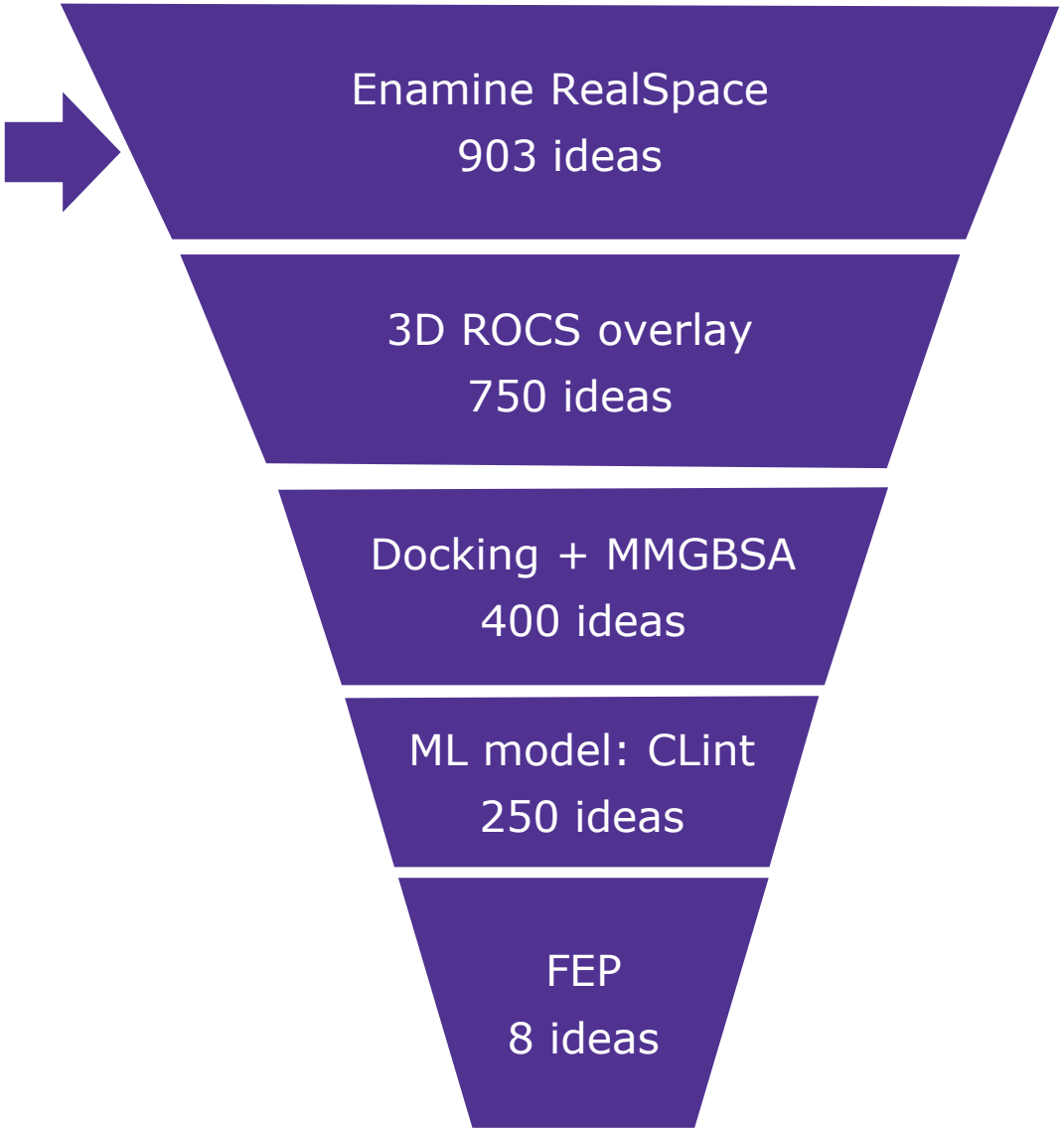
Discovery of new chemical starting points with FEP

From fragment to hit: Proof-of-principle for in silico optimization



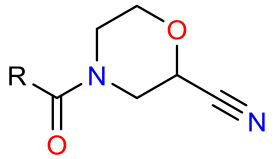
SPR KDss = 300 µM

LE = 0.25



Discovery of new chemical starting points with FEP

From fragment to hit: Proof-of-principle for in silico optimization



SPR KDss = 300 µM

LE = 0.25

Enamine RealSpace
903 ideas

3D ROCS overlay
750 ideas

Docking + MMGBSA
400 ideas

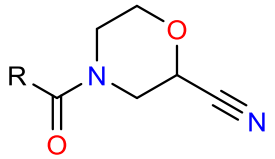
ML model: CLint
250 ideas

FEP
8 ideas

Synthesis at Enamine
- 4 weeks
- < 100 EUR per compound

Discovery of new chemical starting points with FEP

From fragment to hit: Proof-of-principle for in silico optimization



SPR KDss = 300 µM

LE = 0.25

Enamine RealSpace
903 ideas

3D ROCS overlay
750 ideas

Docking + MMGBSA
400 ideas

ML model: CLint
250 ideas

FEP
8 ideas

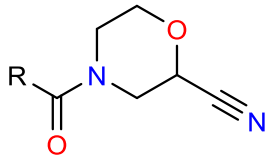
5 out of 8 molecules have IC50 < 100 µM

Synthesis at Enamine

- 4 weeks
- < 100 EUR per compound

Discovery of new chemical starting points with FEP

From fragment to hit: Proof-of-principle for in silico optimization



SPR KDss = 300 µM

LE = 0.25

Enamine RealSpace
903 ideas

3D ROCS overlay
750 ideas

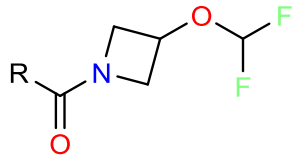
Docking + MMGBSA
400 ideas

ML model: CLint
250 ideas

FEP
8 ideas

5 out of 8 molecules have IC50 < 100 µM

Synthesis at Enamine
- 4 weeks
- < 100 EUR per compound

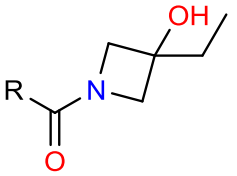


Top 1 in FEP

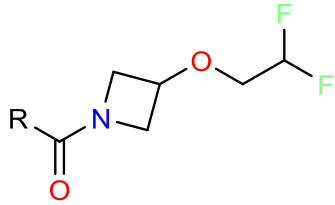
IC50 = 1.2 µM

ITC KD = 1 µM

LE = 0.41



IC50 = 24 µM



IC50 = 47 µM

Summary

- Large-scale prospective benchmarking demonstrated that accuracy of 1.6 kcal/mol can be obtained for diverse and challenging targets in an industry setting
- Large new public benchmark for free energy calculations created
- Accuracy on benchmark is in line with prospective results from in-house projects
- Recommended use case: large library scanning with FEP
- Successful in-silico optimization of fragment to hit

FEP has become a mainstay in computational chemistry support at Merck KGaA, Darmstadt, Germany

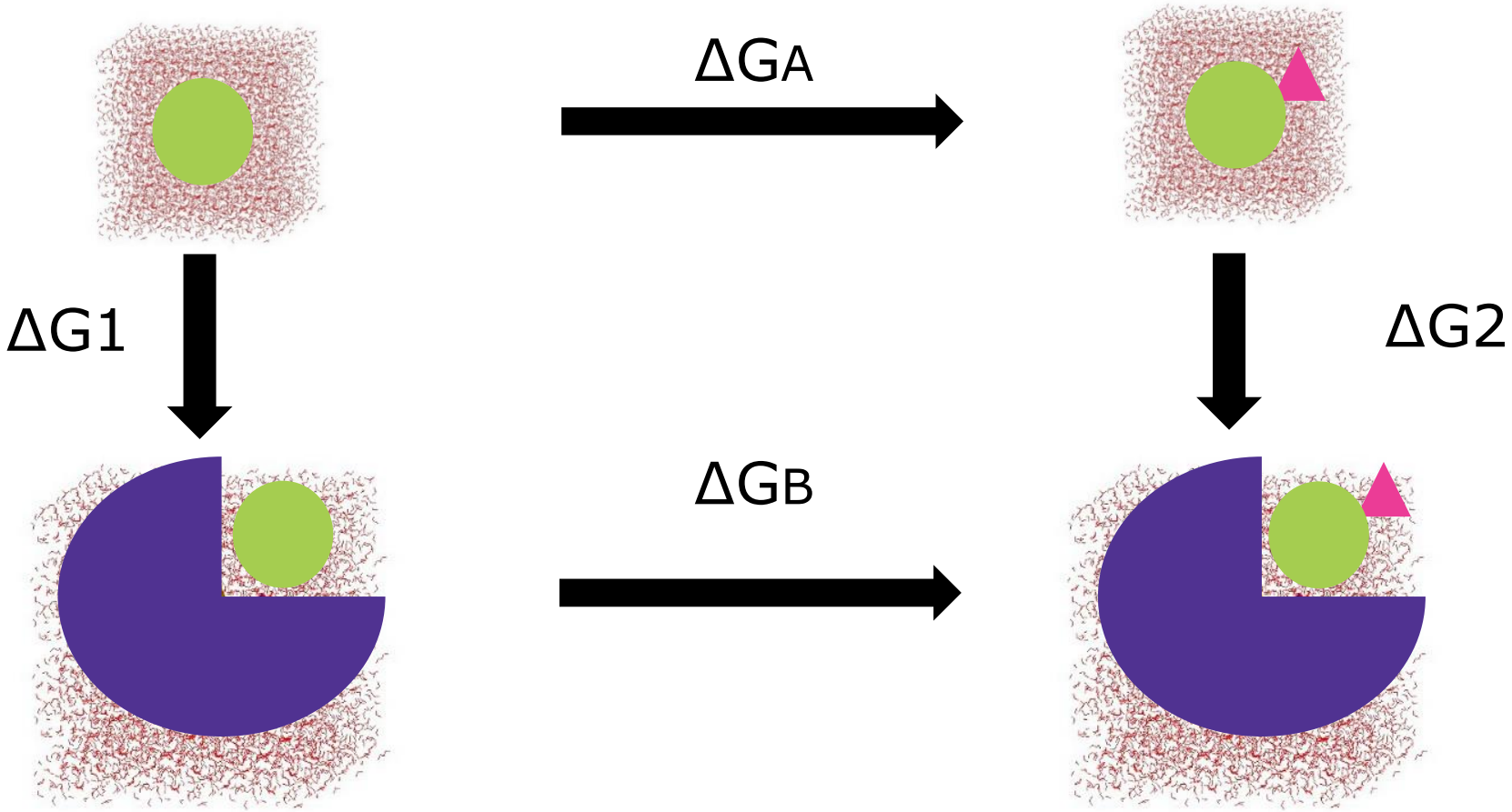
FEP benchmark available on github:

github.com/MCompChem/fep-benchmark

MERCK

Free energy perturbation (FEP)

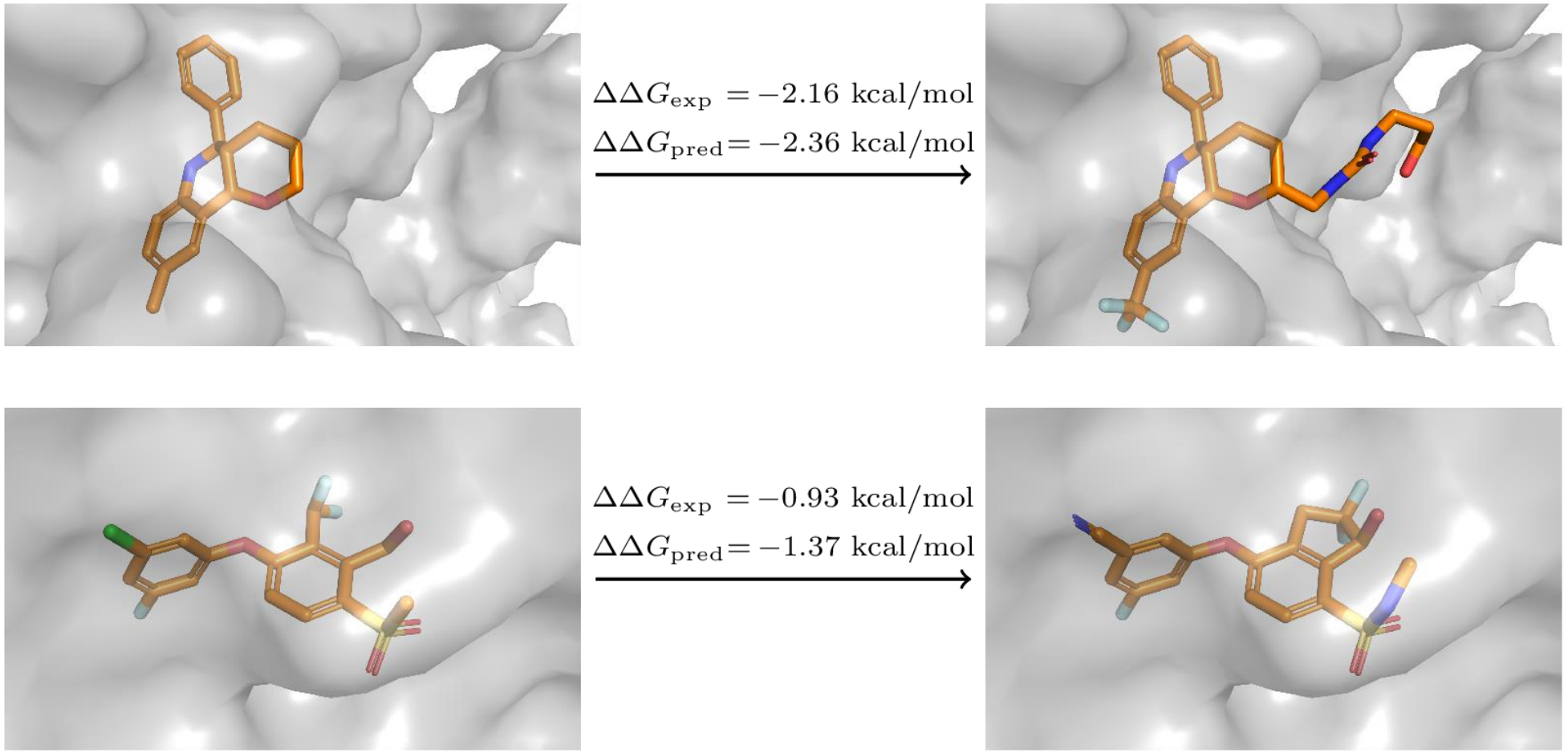
A physics-based method for computing binding affinity differences with molecular dynamics simulations



$$\Delta\Delta G = \Delta G1 - \Delta G2 = \Delta G_B - \Delta G_A$$

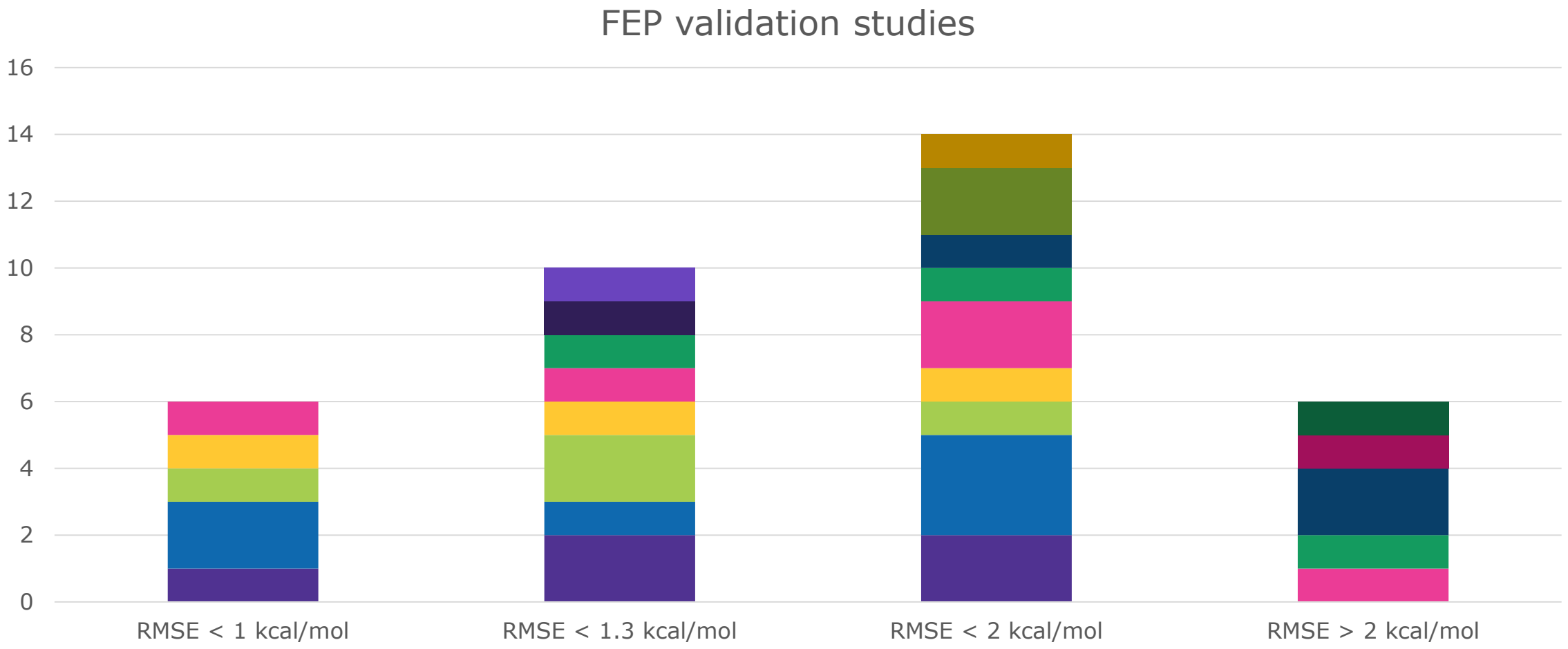
264 ligands for eight pharmaceutically relevant targets

New benchmark for free energy calculations created



Desired accuracy below 1 kcal/mol

Validation results vary across targets and (sub-)series



Pushing the technology to the edge
Observed limitations in FEP calculations (qualitative)

- Predictions often not possible in certain parts of the molecule (indicated by validation study)
- Difficulty in predicting solvent accessible R-groups (often overestimated)
- Transformation from short R-group to long, flexible chains
- Transformation from aromatic to aliphatic ring → improvement in OPLS3e
- Changes in net charge and charge distribution
- Substituted aliphatic rings → improvement with new torsion fitting feature in release 19-3