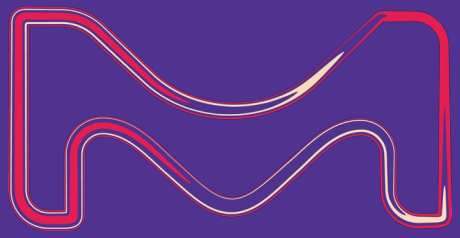
What compound to synthesize next?

Wie Machine Learning und KI das Wirkstoffdesign beeinflussen

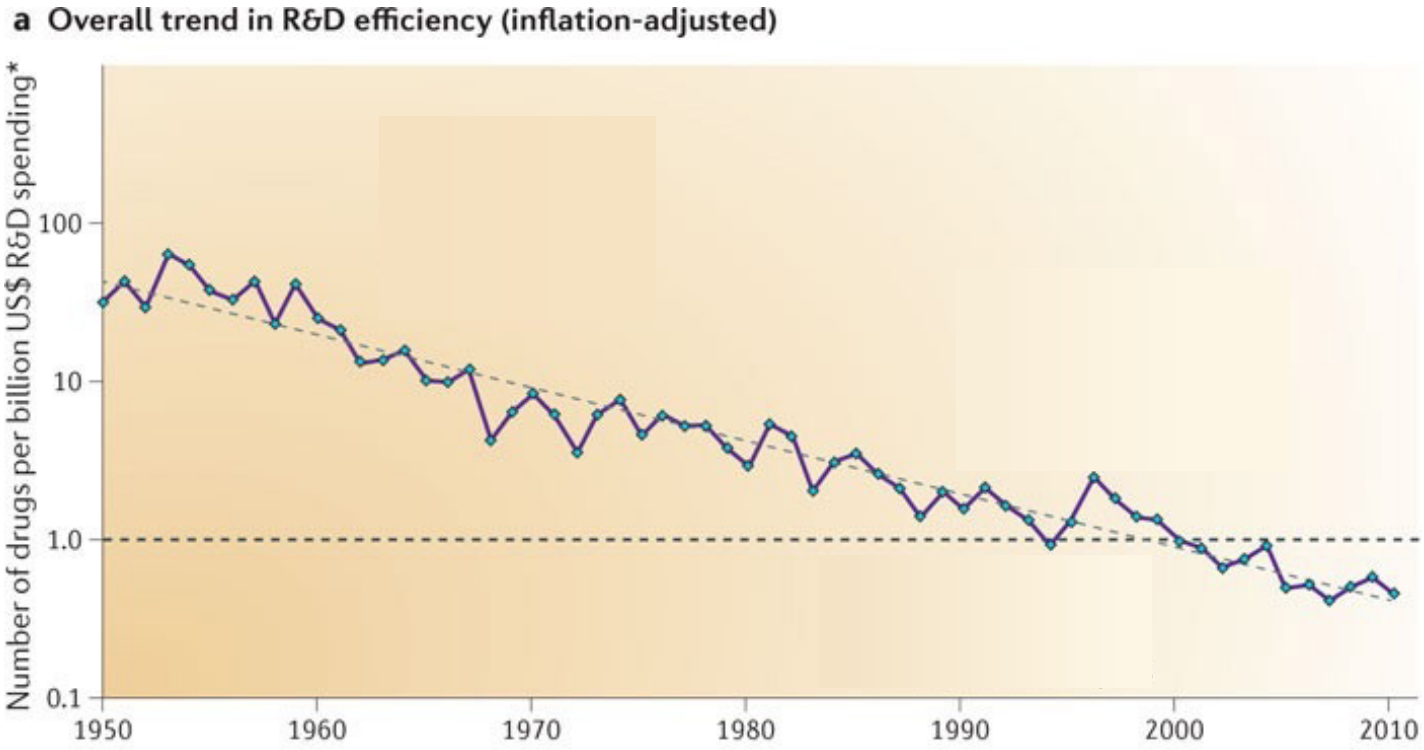
Daniel Kuhn
Computational Chemistry & Biologics
Merck Healthcare KGaA

2021-04-14



MERCK

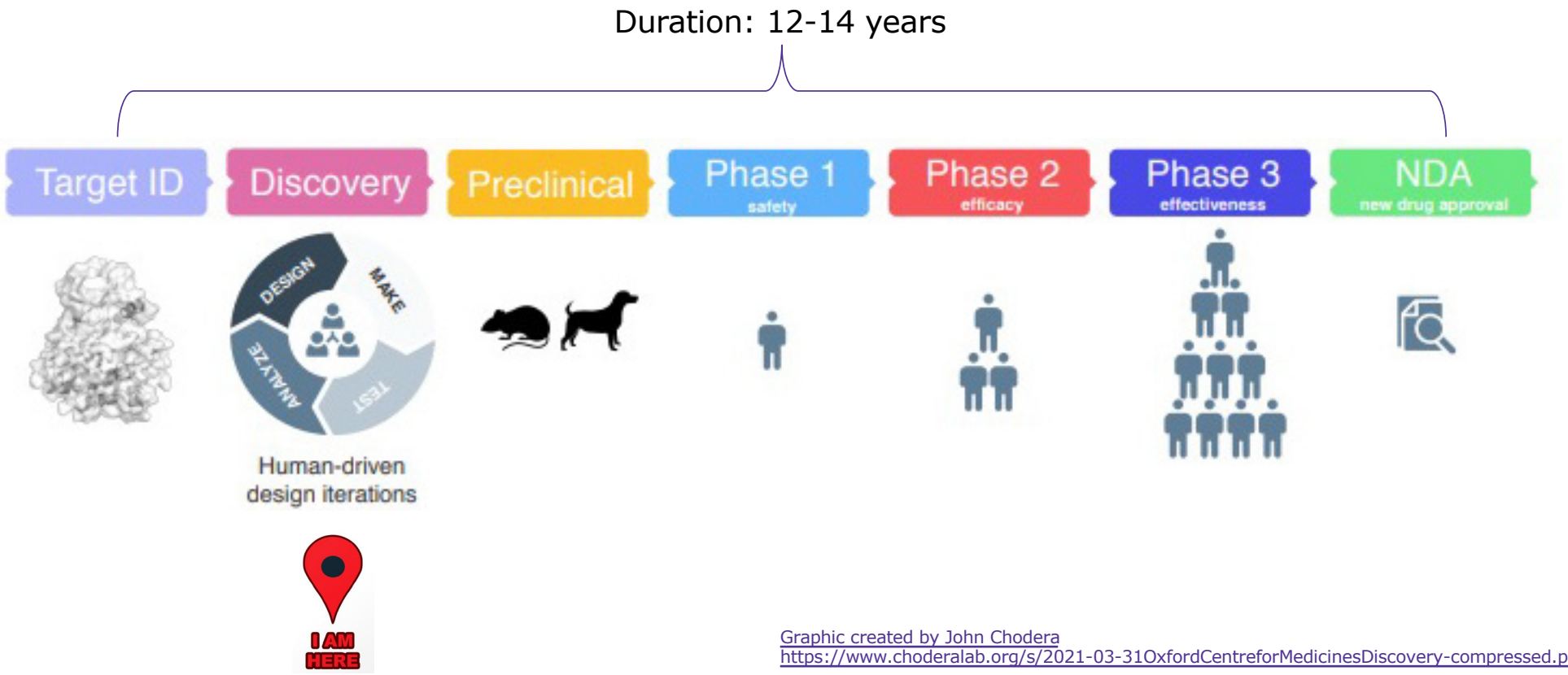
Efficiency in drug discovery is declining over years



Scannell et al., Nat Rev Drug Disc 2012

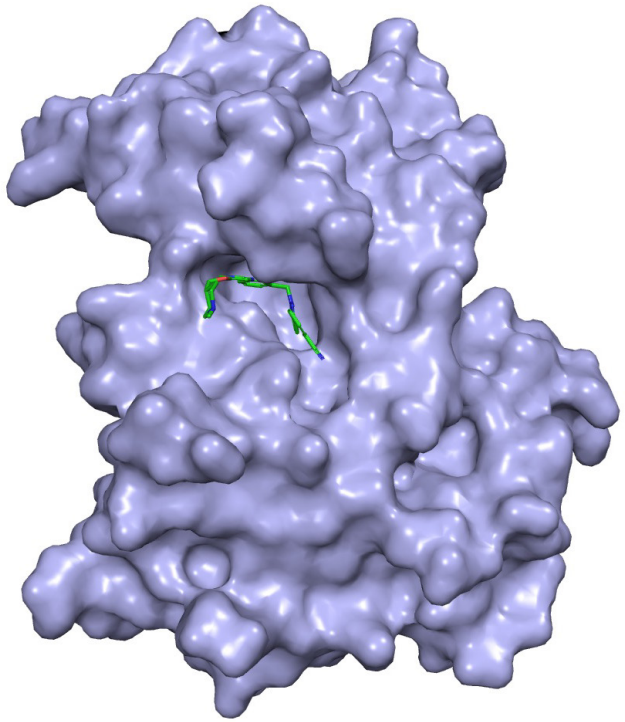
https://www.nature.com/articles/nrd3681

Drug discovery development is costly and timely

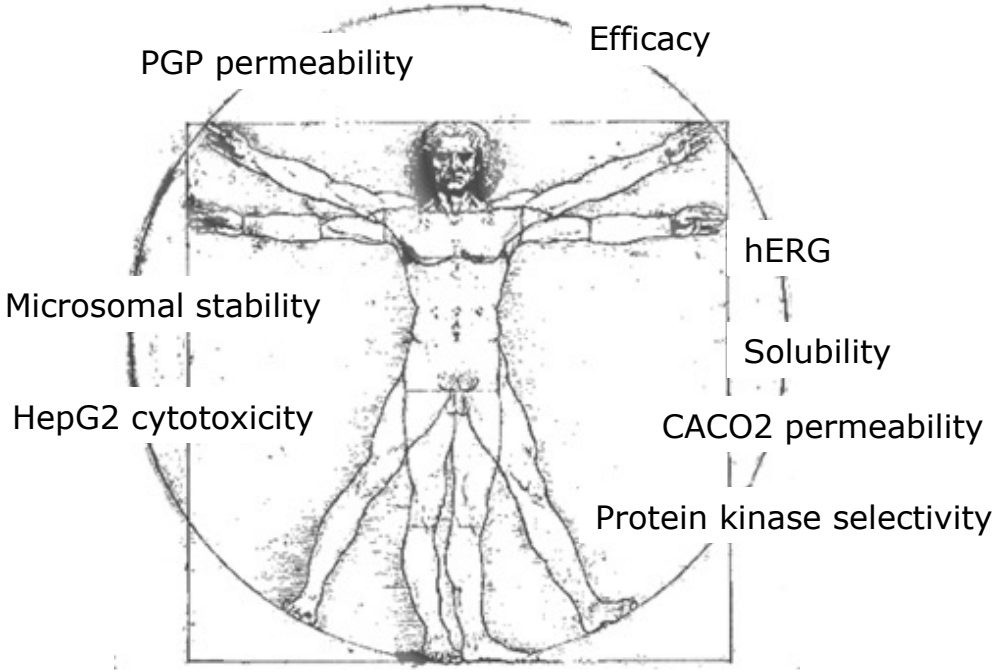


Graphic created by John Chodera
https://www.choderalab.org/s/2021-03-31OxfordCentreforMedicinesDiscovery-compressed.pdf

Drug development
A multi-parameter optimization problem



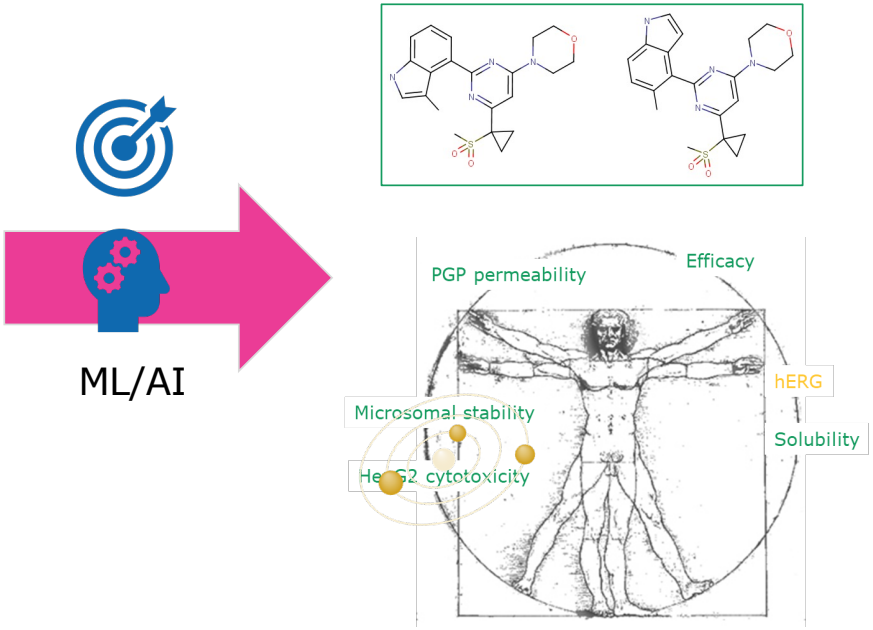
TEPMETKO (Tepotinib) bound to c-Met



Which compounds to make next?

Challenge: Chemical space is huge – which compounds to make next?

Chemical space is huge



Foote et al., *J. Med. Chem.* **2013**, *56*, 2125–2138

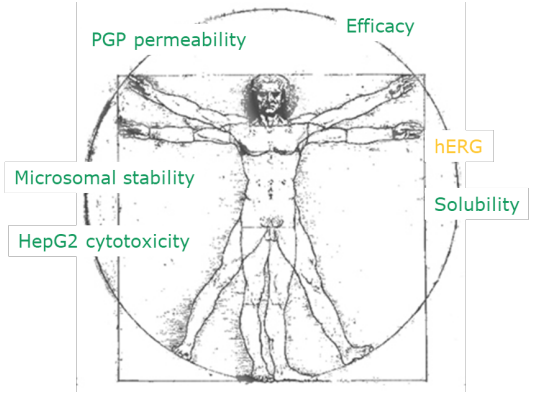
AI creatively inventing novel stuff

Neural network DALL·E is a 12-billion parameter version of GPT-3 language model trained to generate images from text descriptions

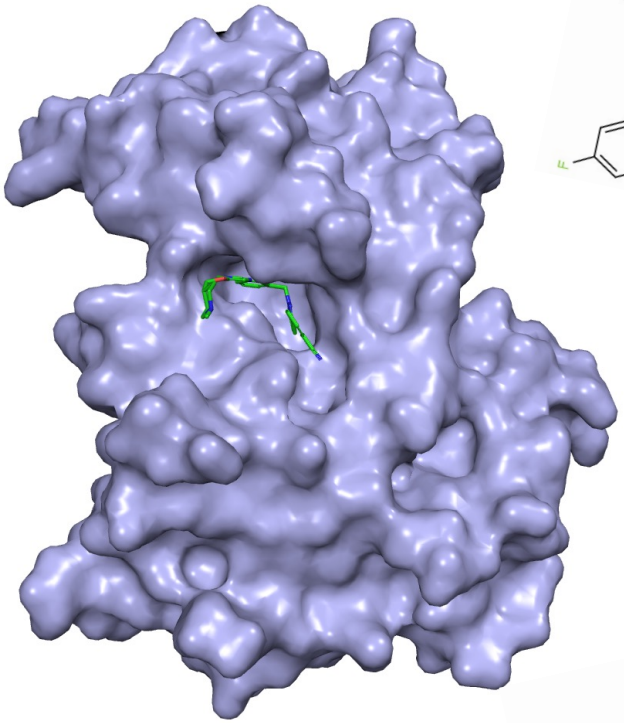
„Draw a picture of an armchair in the shape of an avocado “

http://openai.com - DALL·E: Creating images from text

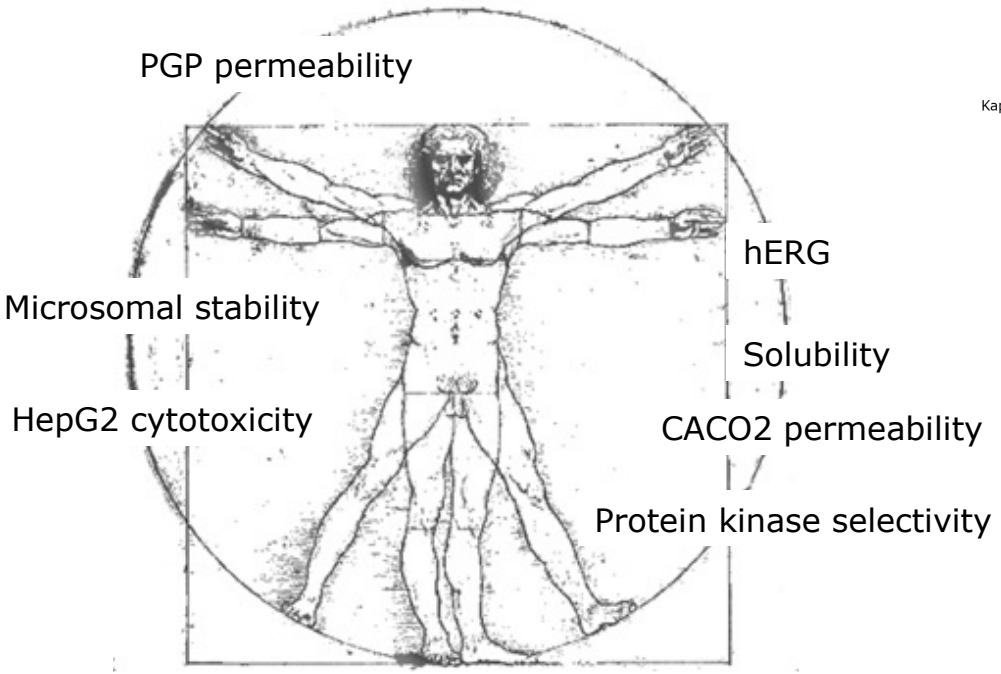
„Design a cellular potent c-Met inhibitor with good microsomal stability and high CACO2 permeability“

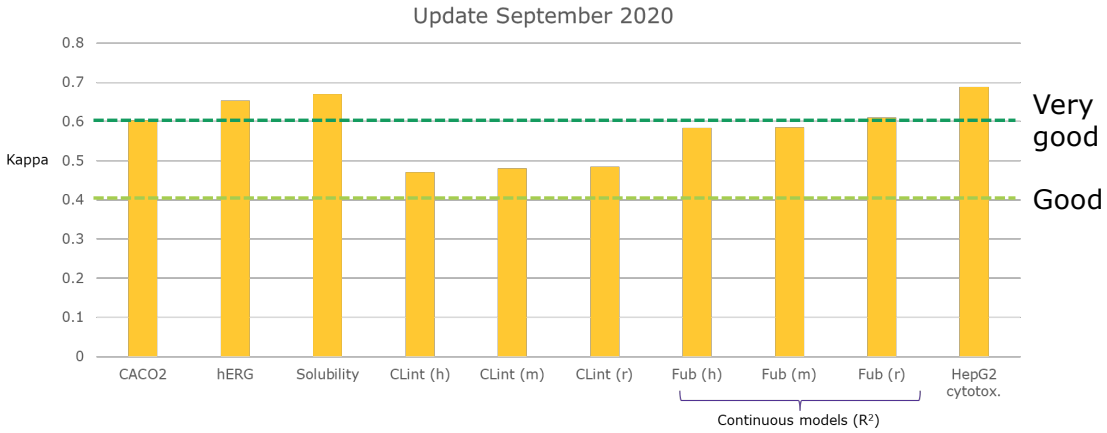


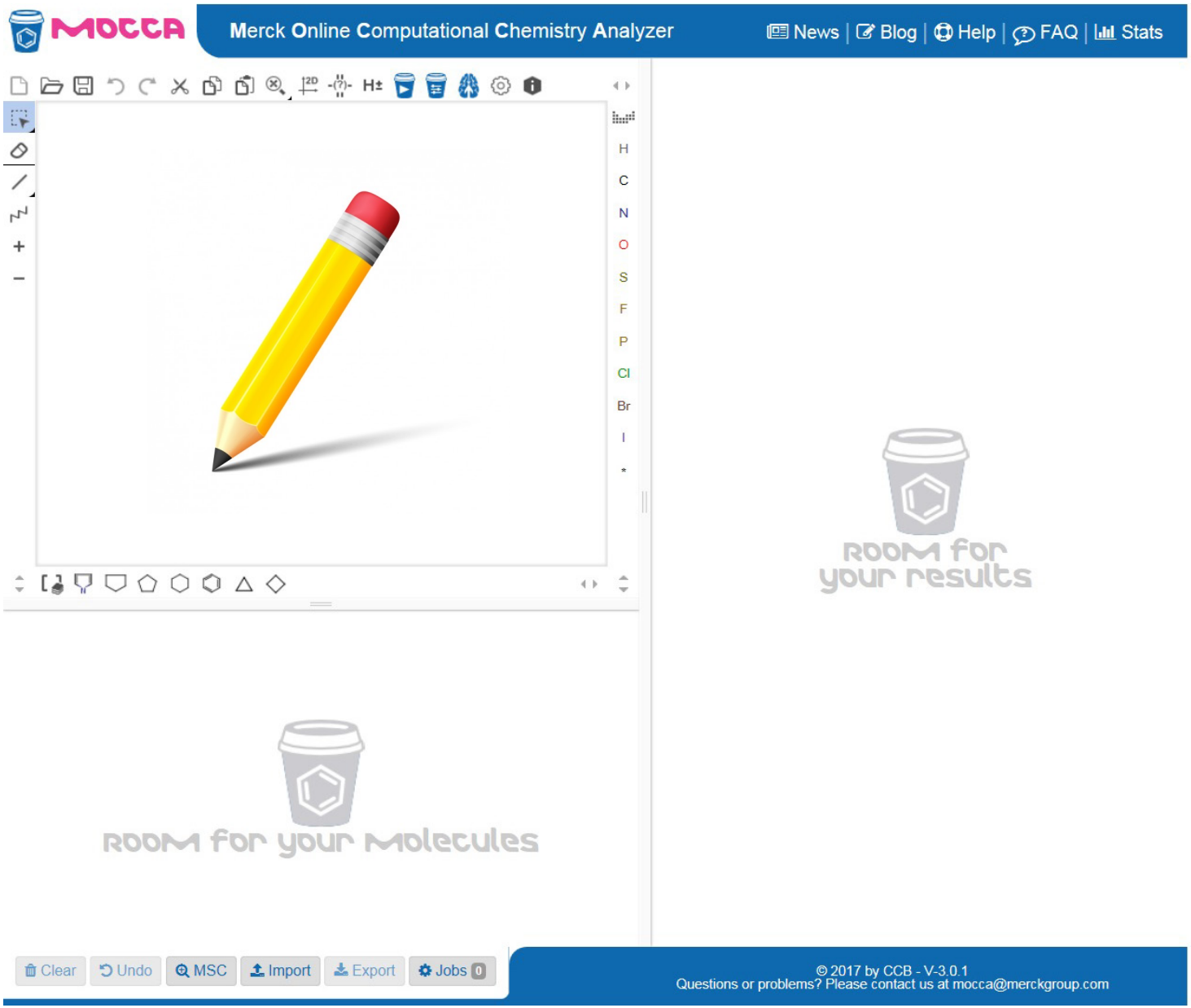
AI-NN

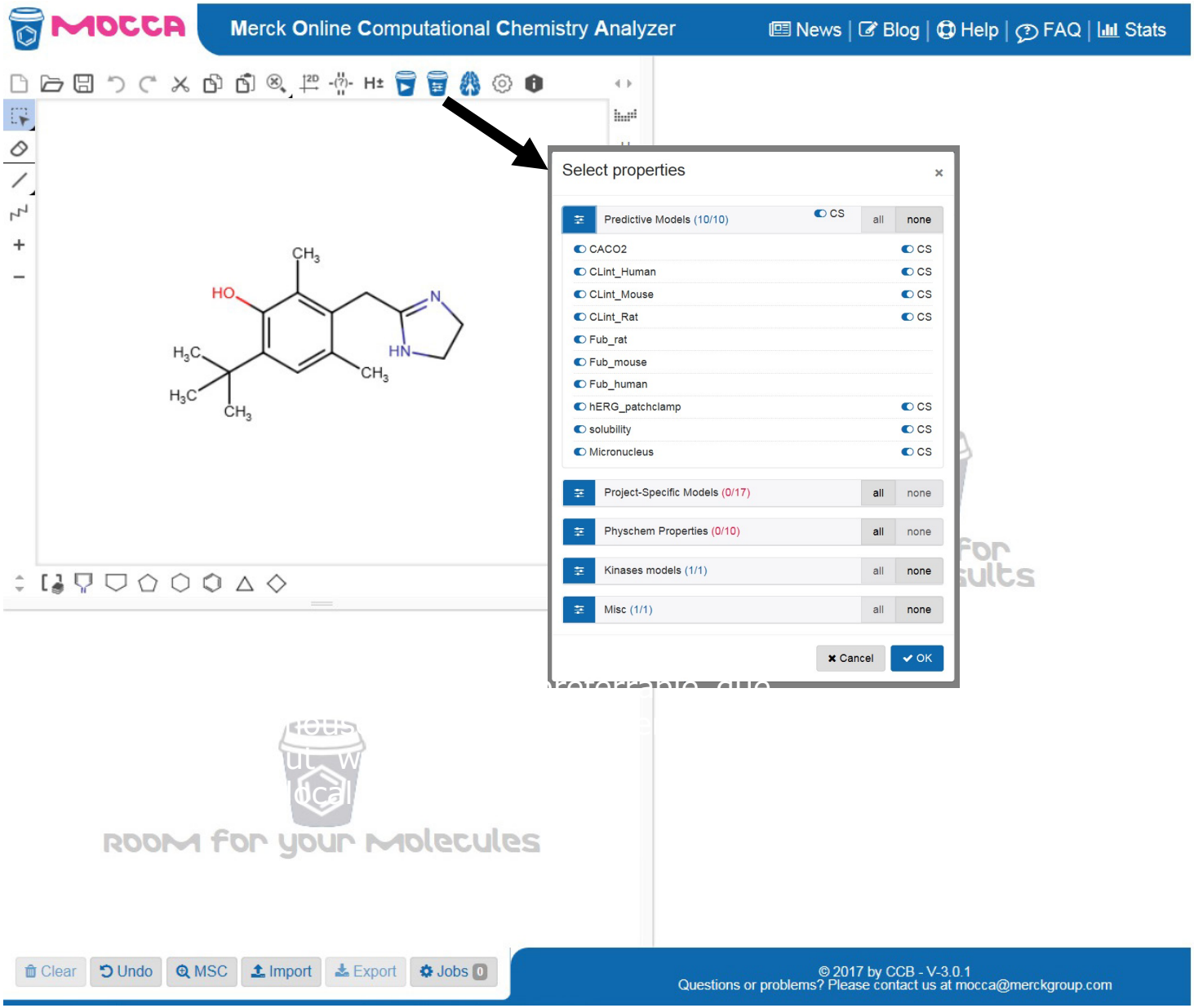


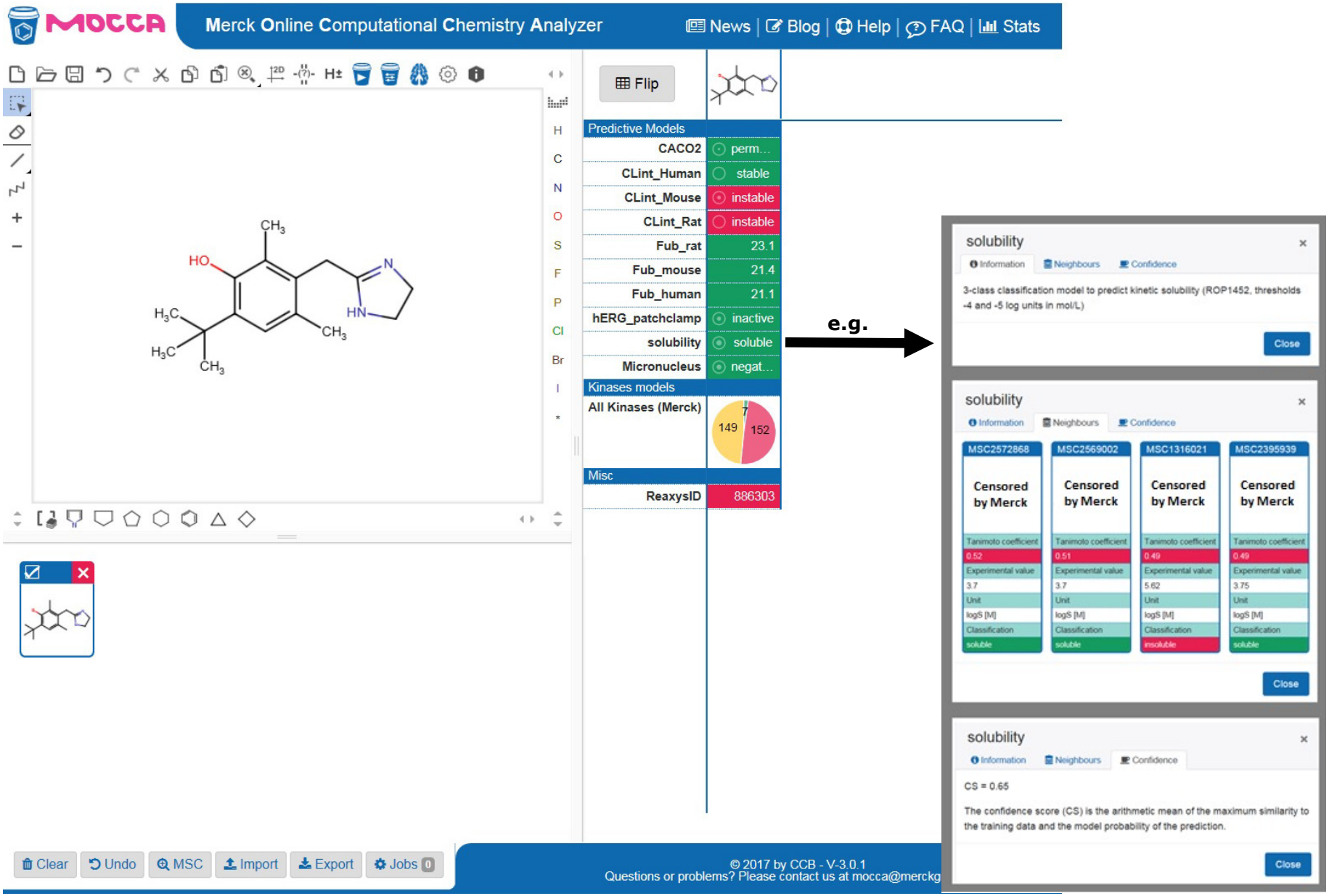
Predictive models in MOCCA: Endpoints & performance






MOCCA
Merck Online Computational Chemistry Analyzer
News | Blog | Help | FAQ | Stats
H
C
N
O
S
F
P
Cl
Br
I
*
ROOM FOR YOUR RESULTS
ROOM FOR YOUR MOLECULES
Clear
Undo
MSC
Import
Export
Jobs 0
© 2017 by CCB - V-3.0.1
Questions or problems? Please contact us at mocca@merckgroup.com


MOCCA
Merck Online Computational Chemistry Analyzer
News | Blog | Help | FAQ | Stats
Select properties
Predictive Models (10/10)
CS
all
none
CACO2
CLint_Human
CLint_Mouse
CLint_Rat
Fub_rat
Fub_mouse
Fub_human
hERG_patchclamp
solubility
Micronucleus
Project-Specific Models (0/17)
Physchem Properties (0/10)
Kinases models (1/1)
Misc (1/1)
Cancel
OK
Room for your Molecules
Clear
Undo
MSC
Import
Export
Jobs 0
© 2017 by CCB - V-3.0.1
Questions or problems? Please contact us at mocca@merckgroup.com

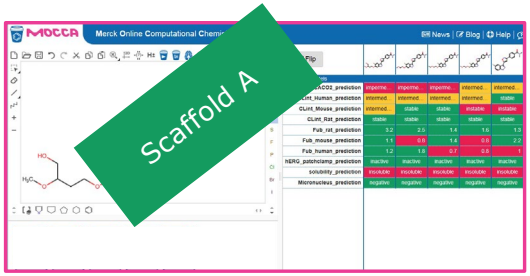

Merck Online Computational Chemistry Analyzer
News | Blog | Help | FAQ | Stats
Flip
Predictive Models
CACO2 perm...
CLint_Human stable
CLint_Mouse instable
CLint_Rat instable
Fub_rat 23.1
Fub_mouse 21.4
Fub_human 21.1
hERG_patchclamp inactive
solubility soluble
Micronucleus negat...
Kinases models
All Kinases (Merck)
149
152
Misc
ReaxysID 886303
e.g.
solubility
Information
Neighbours
Confidence
3-class classification model to predict kinetic solubility (ROP1452, thresholds -4 and -5 log units in mol/L)
Close
MSC2572868
MSC2569002
MSC1316021
MSC2395939
Censored by Merck
Tanimoto coefficient
0.52
0.51
0.49
0.49
Experimental value
3.7
3.7
5.62
3.75
Unit
logS [M]
Classification
soluble
soluble
insoluble
soluble
CS = 0.65
The confidence score (CS) is the arithmetic mean of the maximum similarity to the training data and the model probability of the prediction.
Clear
Undo
MSC
Import
Export
Jobs 0
© 2017 by CCB - V-3.0.1
Questions or problems? Please contact us at mocca@merckg

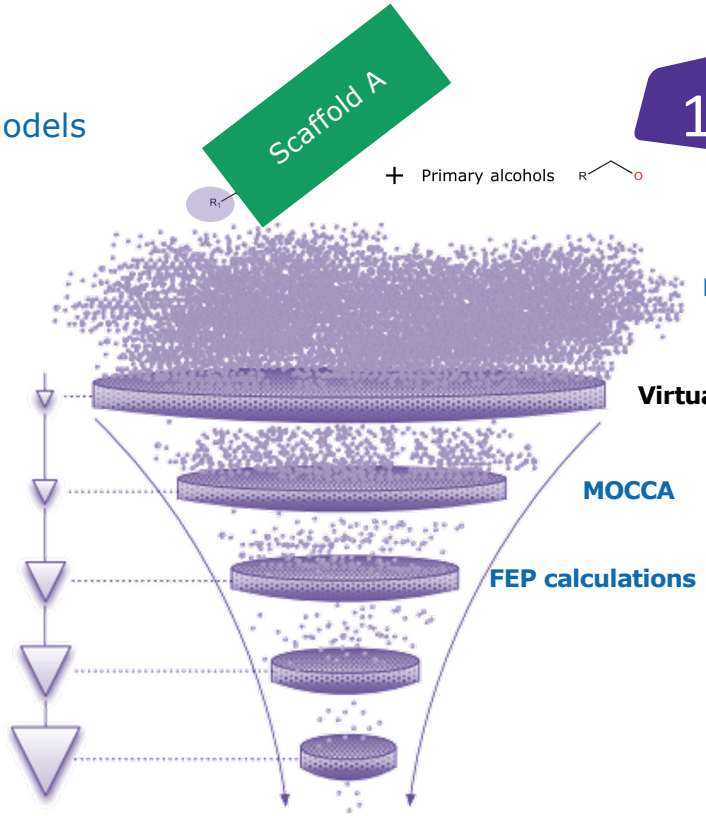
Combination is key for impact in compound optimization

2

MOCCA:

Application of ML/DNN predictive models





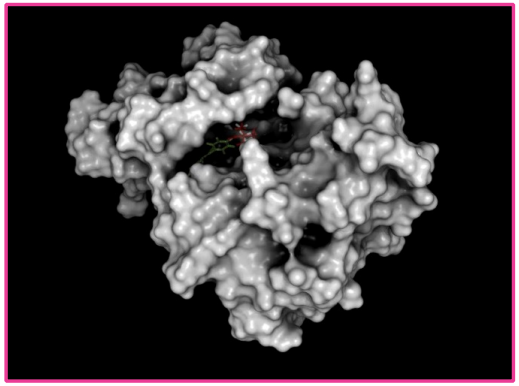
1

MASSIV: Enumeration of synthetically accessible chemical space

Generative design

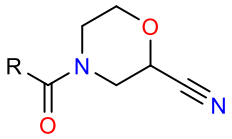
3

FEP: Binding constant prediction



Discovery of new chemical starting points with FEP+ML

Use case 3: From fragment to hit



SPR KDss = 300 µM

LE = 0.25

Enamine RealSpace
903 ideas

3D ROCS overlay
750 ideas

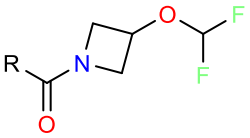
Docking + MMGBSA
400 ideas

ML model: CLint
250 ideas

FEP
8 ideas

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

5 out of 8 molecules have IC50 < 100 µM

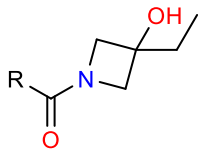


Top 1 in FEP

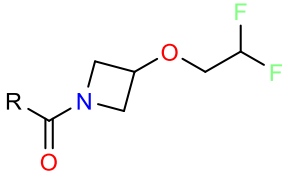
IC50 = 1.2 µM

ITC KD = 1 µM

LE = 0.41



IC50 = 24 µM



IC50 = 47 µM

Christina Schindler

The next ten years

Predictive modelling & AI driving drug design

Merck celebrated it's 352th birthday on August 26th 2020

MERCK

Acknowledgements

Christina Schindler
Lukas Friedrich
Friedrich Rippmann
Vanita Sood
Mireille Krier
Tim Knehans
Cornelius Kohl
Gianna Pohl
Michael Krug
Jakub Gunera
Theresa Johnson
Hans-Peter Buchstaller

Johannes Schimunek
Kristina Preuer
Günter Klambauer
Sepp Hochreiter

Quentin Perron
Yann Gaston-Mathe

Paul Czodrowski
Marcel Baltruschat

IT-Consultants

Jan Fiedler
Christian Röder
Samo Turk
Andrew Dalke

Franca Klingler
Christian Lemmen