CDD WEBINAR

Recorded Live

Drug Design Trends in the Age of AI

LIVE

September 14, 2023
8:00AM (PDT), 11:00AM (EDT), 16:00 (BST)

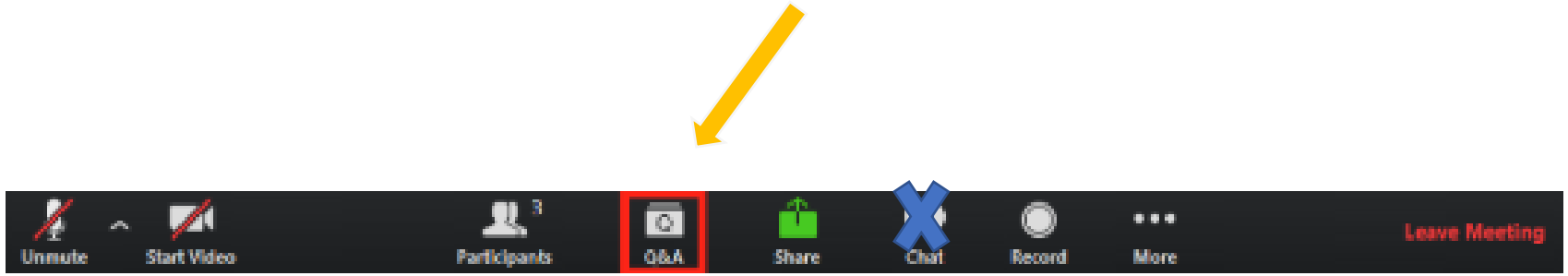
BARRY BUNIN, PHD
CEO, Collaborative Drug Discovery

DANIEL ERLANSON, PHD
SVP, Innovation and Discovery,
Frontier Medicines

ERIC MARTIN, PHD
Director, Computational Chemistry,
Novartis

CDD VAULT®
Complexity Simplified

Have a question to ask our panel?
Open the ZOOM Q&A and type in your question at anytime !



Saving your Questions to the end

Featuring these leading scientists...

Barry Bunin, PhD

CEO, Collaborative Drug Discovery

Barry Bunin founded CDD in 2004 to pioneer CDD Vault, a hosted research data management system. Prior to CDD, he was an Entrepreneur in Residence with Eli Lilly, as well as the founding CEO, President, & CSO of Libraria (now Eidogen-Sertanty). On the scientific side, Dr. Bunin is an expert in cheminformatics and combinatorial chemistry, and has written two books in those fields. He is also on the patent of Kyprolis, a selective proteasome inhibitor for treating multiple myeloma. Dr. Bunin received his B.A. from Columbia University and his Ph.D. from UC Berkeley, where he synthesized and tested the initial 1,4-benzodiazepine libraries with Professor Jonathan Ellman

Eric Martin, PhD

Director, Computational Chemistry at Novartis

Eric Martin has a Ph.D. in physical organic chemistry from Yale University. He has worked in computational drug design and herbicide design for 40 years at Dow, DowElanco, Chiron and Novartis. He is currently developing novel methodologies for two areas of drug discovery: 1) Developing “Profile-QSAR”, a massively multitask machine learning method that builds experimental-quality virtual screening models for over 9000 IC50 assays, and 2) “rational oral bioavailability design” during lead optimization, by applying global sensitivity analysis to physiologically-based pharmacokinetics simulations. Eric was awarded the lifetime title of Novartis Leading Scientist for the former.

Daniel Erlanson, PhD

SVP, Innovation and Discovery at Frontier Medicines

Daniel Erlanson is an expert in fragment-based and covalent drug discovery. Prior to Frontier, Dan co-founded Carmot Therapeutics, where he developed drug discovery technologies and led chemistry efforts that resulted in three clinical-stage molecules and partnerships with biotech companies including Amgen and Genentech. One of the molecules advanced by Amgen is LUMAKRAS™ (sotorasib), the first FDA and EMA approved inhibitor of KRASG12C, a previously undruggable target. Dr. Erlanson earned his Ph.D. in Chemistry from Harvard University in the laboratory of Gregory L. Verdine.

Interpreting the impact of AI large language models on chemistry

BY PHILIP BALL | 5 APRIL 2023

Chemistryworld.com 5 April 2023 Philip Ball

LLMs may outperform Alphafold, but currently struggle to identify simple chemical structures

Is AI on the brink of something massive? That's been the buzz over the past several months, thanks to the release of improved 'large language models' (LLMs) such as OpenAI's GPT-4, the successor to ChatGPT. Developed as tools for language processin these algorithms respond so fluently and naturally that some users become convinced th are conversing with a genuine intelligence. Some researchers have suggested that LLMs beyond traditional deep-learning AI methods by displaying emergent features of the human mind, such as a theory of mind that attributes other agents with autonomy and

UNIVERSITY of ROCHESTER

News Center | SCIENCE & TECHNOLOGY | SOCIETY & CULTURE | CAMPUS LIFE | UNIVERSITY NEWS | ARTS

search news center SEARCH

Large language models could be the catalyst for a new era of chemistry

May 23, 2023

CAS A division of the American Chemical Society

SOLUTIONS ▾ RESOURCES ▾ CAS DATA ▾ ABOUT ▾

Home > Resources > CAS Insights > Are large language models right for scientific research

Are large language models right for scientific research

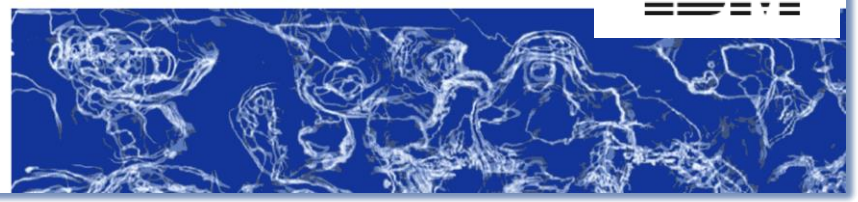
Philippe Ayala, Data Science Technical Manager

August 11, 2023

An AI foundation model that learns the grammar of molecules

Meet MoLFormer-XL, a pretrained AI model that infers the structure of molecules from simple representations, making it faster and easier to screen molecules for new applications or create them from scratch.

IBM



nature computational science

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > nature computational science > research highlights > article

Research Highlight | Published: 23 January 2023

Cheminformatics

Large language model for molecular chemistry

Jie Pan

Nature Computational Science **3**, 5 (2023) | Cite this article

922 Accesses | **1** Citations | **10** Altmetric | Metrics

Machine learning (ML) has disruptively changed the way scientists predict molecular structure and properties that are relevant to chemical and materials design. Graph neural

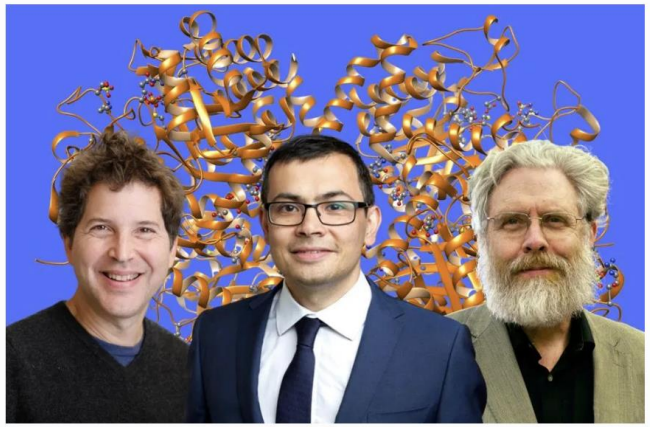
Forbes

The Next Frontier For Large Language Models Is Biology

Rob Toews Contributor

I write about the big picture of artificial intelligence.

Follow



David Baker (University of Washington), Demis Hassabis (DeepMind) and George Church (Harvard) have ... [+] PHOTO SOURCE: U OF W, ROYAL SOCIETY, HARVARD

Large language models like GPT-4 have taken the world by storm thanks to their astonishing command of natural language. Yet the

Eric's long ML career in brief

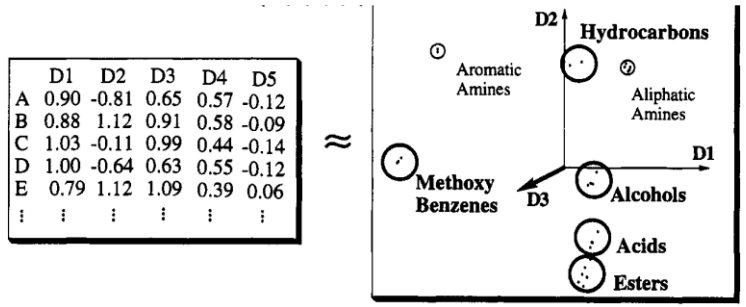
K college:
- 1978 with Sam Yalkowsky, delS.fus term in General Solubility Equation (mpt & logP)

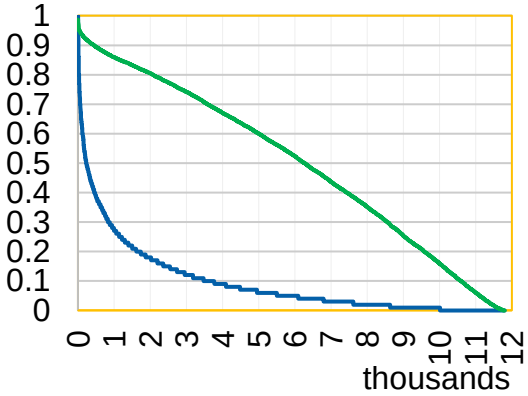
Yale:
- 1978-83 built automated reaction calorimeter to study Taft Es steric parameter a term in Hansch QSAR. (machine language)

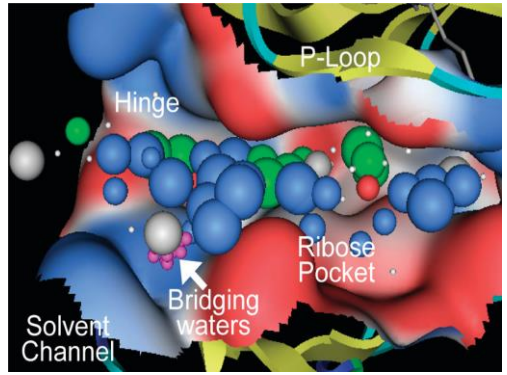
Dow/DowElanco:
- 1984 1st DSC-XRD-MS enabled Seldane 1st drug withdrawn for hERG
- 1985 Phloem transport simulations Ro6
- 1986 AutoSar 1st automated Hansch QSAR
- 1986 1st automated direct logP (w/ CPC)
- 1987 GASP 1st global leaching into ground water

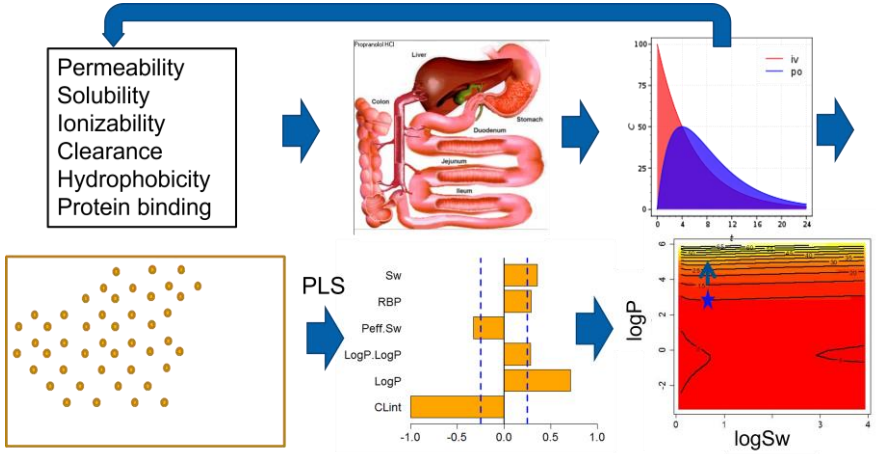
Chiron:
- 1992 1st chemical spaces & library design
- 2005 pQSAR 1st massively-multitask bioactivity models

NVS:
- 2006 AutoShim: 1st? ML scoring functions for docking
- 2016 1st Adapting PBPK to LO. Machine learning on synthetic data from 10,000 PBPK simulations for %F & AUC (for MPO)





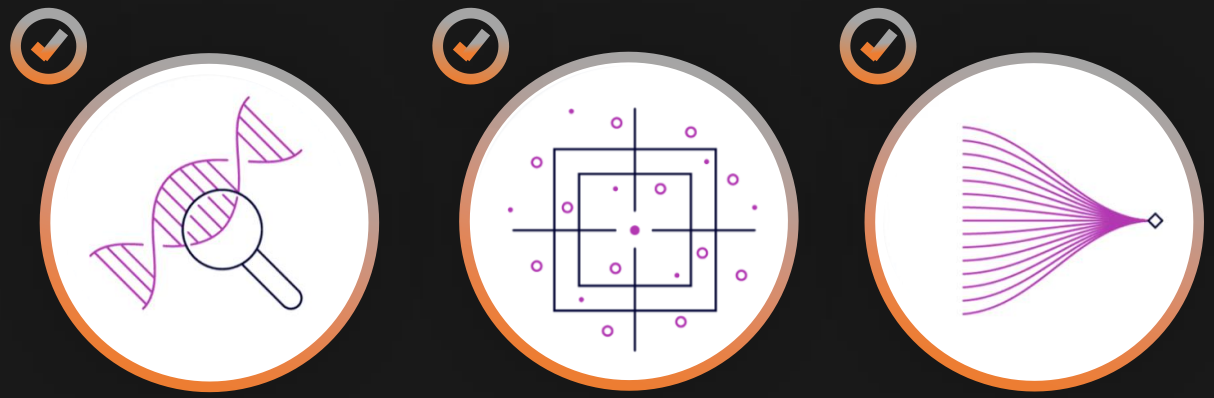




The Frontier™ Platform

FRONTIER MEDICINES

Frontier: A leader in chemoproteomic covalent drug discovery



Mapping the proteome and prioritize targets

Targeting any protein of interest

Covalent screening

Identifying high-quality covalent hits

Accelerated covalent drug optimization

Efficiently developing covalent fragment drug candidates

Daniel Erlanson, PhD

Senior Vice President of Innovation and Discovery, Frontier Medicines

- PhD in Chemistry, Harvard University, lab of Greg Verdine
- Postdoc, Jim Wells, Genentech
- Inventor on more than 15 issued US patents and author of more than 50 publications
- Co-editor of *Fragment-based Approaches in Drug Discovery* and *Fragment-based Drug Discovery: Lessons and Outlook*
- Co-Founder, Carmot Therapeutics
  3 clinical-stage molecules, including the first FDA-approved treatment targeting $KRAS^{G12C}$, LUMAKRAS™ (advanced by Amgen)
- Editor, *Practical Fragments*: *practicalfragments.blogspot.com*

Forty years of fragments

1981 1986 1991 1996 2001 2006 2011 2016 2021 2026

William Jencks

Additivity of Binding Energies

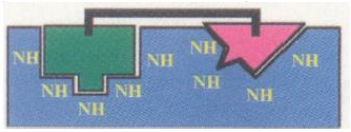
PNAS 1981

$$\Delta G^0_{AB} = \Delta G^i_A + \Delta G^i_B + \Delta G^s.$$

Abbott, SAR by NMR

Shuker, Hajduk, Meadows, Fesik,

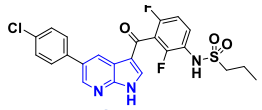
Science 1996



Multiple companies form, multiple drugs enter clinical development

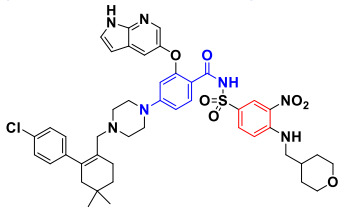
1st FBLD-derived drug approved

vemurafenib (Plexxikon/Roche)



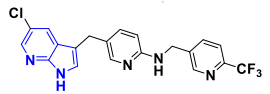
2nd FBLD approval

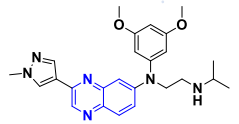
venetoclax (AbbVie/Genentech)



3rd, 4th approvals

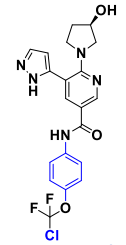
pexidartinib (Daiichi Sankyo)
erdafitinib (Astex/J&J)

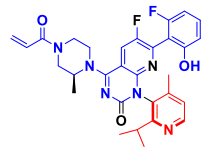




5th, 6th approvals

sotorasib (Amgen)
asciminib (Novartis)





>50 fragment-based drugs have entered the clinic

Approved

- **Asciminib Novartis BCR-ABL1**
- **Erdafitinib J&J/Astex FGFR1-4**
- **Pexidartinib Plexxikon CSF1R, KIT**
- **Sotorasib Amgen KRAS^{G12C}**
- **Vemurafenib Plexxikon/Roche B-RAF^{V600E}**
- **Venetoclax AbbVie/Genentech BCL-2**

Phase 3 (3/5 active)

- **Capivasertib (AZD5363) Astex/AstraZeneca/CR-UK AKT**
- Lanabecestat Astex/AtraZeneca/Lilly BACE1
- **Navitoclax (ABT 263) AbbVie BCL-2/BCL$_{xL}$**
- **Pelabresib (CP-0610) Constellation BET**
- Verubecestat Merck BACE1

Bold: Still active*

Phase 2 (11/22 active)

- **ASTX029 Astex ERK1,2**
- **ASTX660 Astex XIAP, cIAP1**
- AT7519 Astex CDK1,2,4,5,9
- AT9283 Astex Aurora, Janus Kinase 2
- AUY-922 Vernalis/Novartis HSP90
- AZD5991 AtraZeneca MCL1
- DG-051 deCODE LTA4H
- **eFT508 eFFECTOR MNK1/2**
- Indeglitazar Plexxikon PPAR agonist
- LY2886721 Lilly BACE1
- LY3202626 Lilly BACE1
- **LY3372689 Lilly OGA**
- LY517717 Lilly/Protherics FXa
- **LYS006 Novartis LTA4H**
- **MK-8189 Merck PDE10A**
- **MAK683 Novartis PRC2 EED**
- Onalespib Astex HSP90
- **PF-06650833 Pfizer IRAK4**
- PF-06835919 Pfizer KHK
- **PLX51107 Plexxikon BET**
- **S64315 Vernalis/Servier/Novartis MCL1**
- **VK-2019 Culinan Oncology/Wistar EBNA1**

Phase 1 (5/26 active)

- AG-270 Agios/Servier MAT2A
- **ABBV-744 Abbott BD2-selective BET**
- ABT-518 Abbott MMP-2 & 9
- ABT-737 Abbott BCL-2/BCL-xL
- AT13148 Astex AKT, p70S6K
- AZD3839 AstraZeneca BACE1
- AZD5099 AstraZeneca Bacterial Topo II
- **BI 1823911 BI KRAS^{G12C}**
- BI 691751 BI LTA4H
- **CFTX-1554 Confo AT$_2$ receptor**
- ETC-206 D3 MNK1/2
- GDC-0994 Genentech/Array ERK2
- **HTL0014242 Sosei Heptares mGlu5 NAM**
- HTL0018318 Sosei Heptares M1R p agonist
- HTL9936 Sosei Heptares M1R part. agonist
- IC-776 Lilly/ICOS LFA-1
- LP-261 Locus Tubulin
- LY2811376 Lilly BACE1
- Mivebresib AbbVie BRD2-4
- **MRTX1719 Mirati MTAP**
- Navoximod NewLink/ Genentech IDO1
- PLX5568 Plexxikon RAF
- SGX-393 SGX BCR-ABL
- SGX-523 SGX MET
- SNS-314 Sunesis Aurora
- TAK-020 Takeda BTK

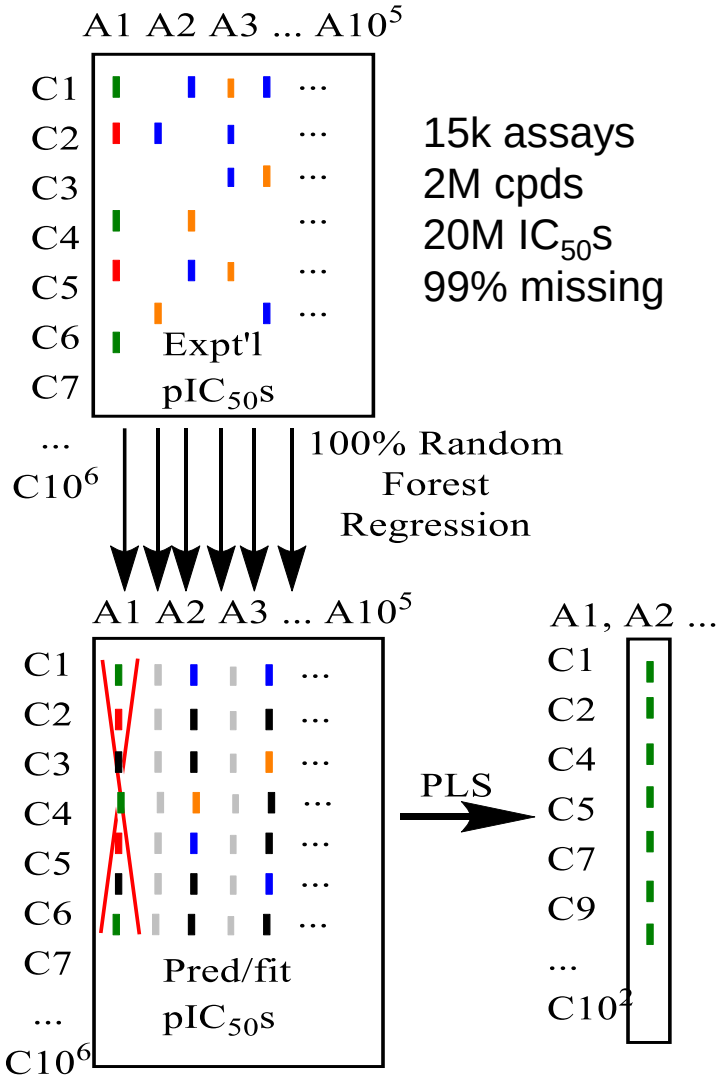
*Updated from Erlanson et al. *Nat. Rev. Drug Disc.* 2016
https://practicalfragments.blogspot.com/2022/11/fragments-in-clinic-2022-edition.html

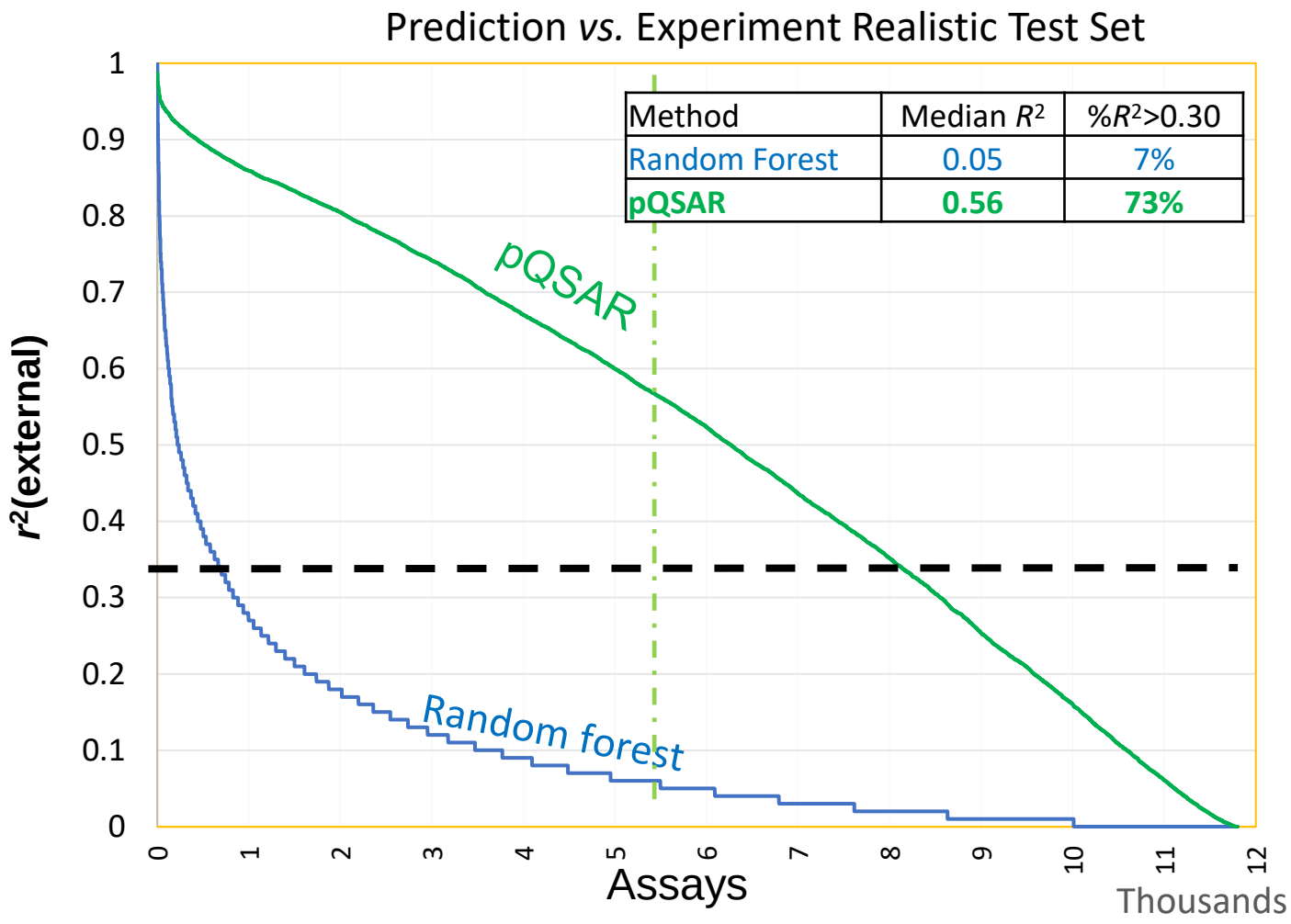
Poll question: what are your favorite methods for finding hits (Hit ID)?

1. HTS - Phenotypic Screen
2. HTS – Target Based
3. FBDD
4. Structure based virtual screening (Docking)
5. ML/Virtual Screening (QSAR)
6. DEL (DNA encoded libraries)
7. Me too-ing (patent/literature busting) “best way to find a new drug is to start w/ an existing one”

Massively multitask pQSAR: Successful models for ¾ (9400) of Novartis assays

*C.f. MTS 4-point pIC_{50} vs. lead op 8- to 12-point pIC_{50} average r^2=**0.54***





Martin *et al.*, *J. Chem. Inf. Model* (2019)

What if you could send 50-500,000+ cpds for 4-pt IC_{50} screens on 5,000+ assays for <$1000, with next-day results?

Most important ML applications are profile predictions for existing compounds

- Monthly, ~6.5M compounds on ~15,000 assays => ~100B IC_{50} predictions.
  - Includes registered synthesis candidates
  - Instant virtual screens (biochemical and phenotypic)
  - **Off-target predictions for your hits or lead series**
  - **Polypharmacology discovery**
  - **Mechanism-of-action discovery from phenotypic hits**
  - Artifact ID and virtual counter-screens
  - Detailed triaging for advancement (pred. IC_{50}, LE, lipE, tox, selectivity, etc.)
  - pQSAR biological profile fingerprints found best for scaffold hopping
  - "Give me all compounds sub-uM in one of these 6 biochemical assays, selective against these anti-targets, and inhibit proliferation in 1+ of these 5 cancer cell lines driven by this pathway."
- Also, custom models and predictions on *ad hoc* virtual compounds

GitHub - Novartis/pQSAR

Discussion: When better to use Fragments vs Whole Molecule SAR?

…and what's the same and different for QSAR vs SAR with each approach?

Experimental and AI Drug Discovery (FBDD + Whole Molecules)

When and where is one approach will be better or worse than the other...

- ✓ **Question: Is the more elegant approach specific to certain types of chemotypes? Proteins? Assays? and/or Therapeutic Areas? Why?**

Fragments are possibly the best way to find additional binding sites, but they are often so weak that you can't tell if the binding sites are functional...

- ✓ **Question: What's the best way(s) to find out if your binding sites <u>are</u> functional?**

Experimental and AI Drug Discovery (FBDD + Whole Molecules)

The advantage of fragment-based drug design is that it covers a huge virtual chemistry space, but now there are huge virtual libraries from vendors (like Enamine)…

- ✓ **Question: Does (and if so how does) VLs impact the calculation of doing fragment and/or whole molecule screening?**

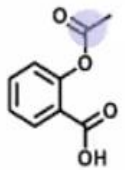
Given what's available virtual vs off the shelf…one can have virtual libraries of fragments and/or whole molecules (and w/ reactions one can go from A to B)…

- ✓ **Question**: **How does the availability of "on-demand" libraries built from fragments impact the fragment-library design/screening vs whole molecule library design & screening tradeoffs? (in house vs off the shelf / outsourced)**

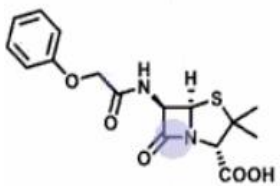
Other Topics: How do the superior methods depend on the optimization strategy (allosteric vs active site, potency vs novelty, big pharma vs small co., covalent vs non-covalent, etc.)…

Recent surge in approved targeted covalent drugs

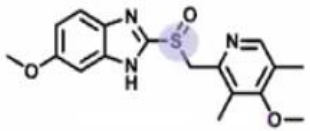
aspirin
1890s



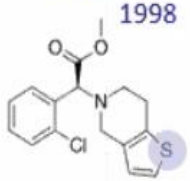
penicillin
1920s



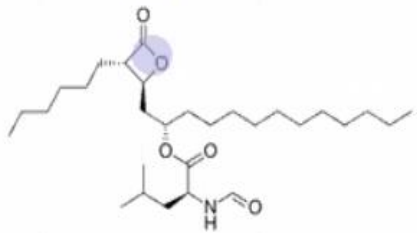
omeprazole
1988



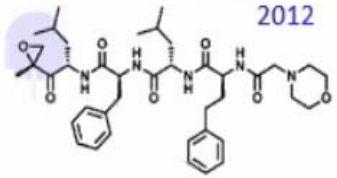
clopidogrel
1998



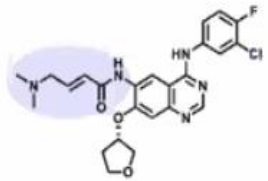
orlistat
1999



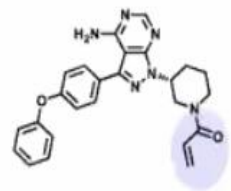
carfilzomib
2012



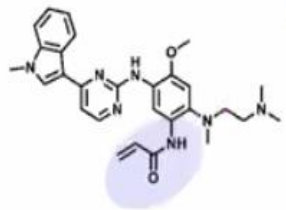
afatinib
2013



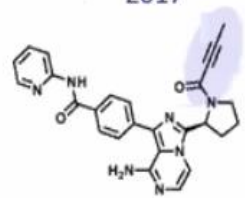
ibrutinib
2013



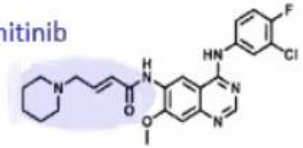
osimertinib
2015



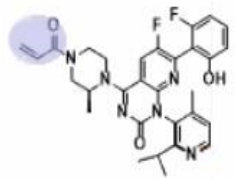
acalabrutinib
2017



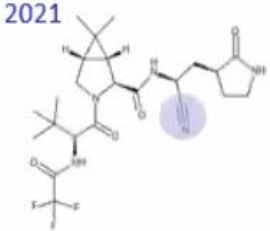
dacomitinib
2018



sotorasib
2021



nirmatrelvir
2021

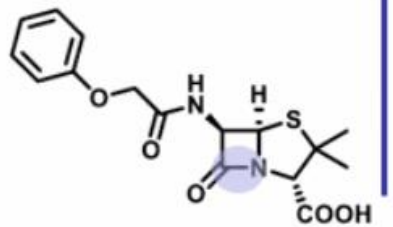


Origins of covalent drugs

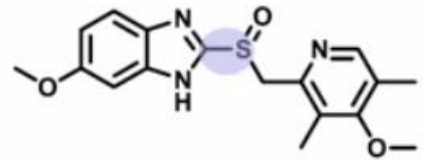
Historic

phenotypic screening

penicillin
1920s



omeprazole
1988



Modern

rationally designed

Kinase "hinge-binding" element

High affinity binder
IC_{50} = 8.2 nM

Kinase "hinge-binding" element

Ibrutinib

Approved 2013

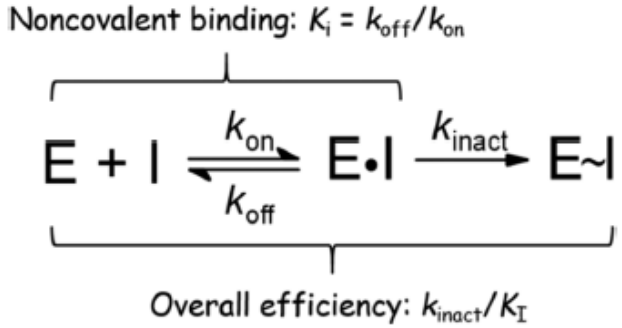
$10 billion sales worldwide in 2021

Cysteine-reactive acrylamide

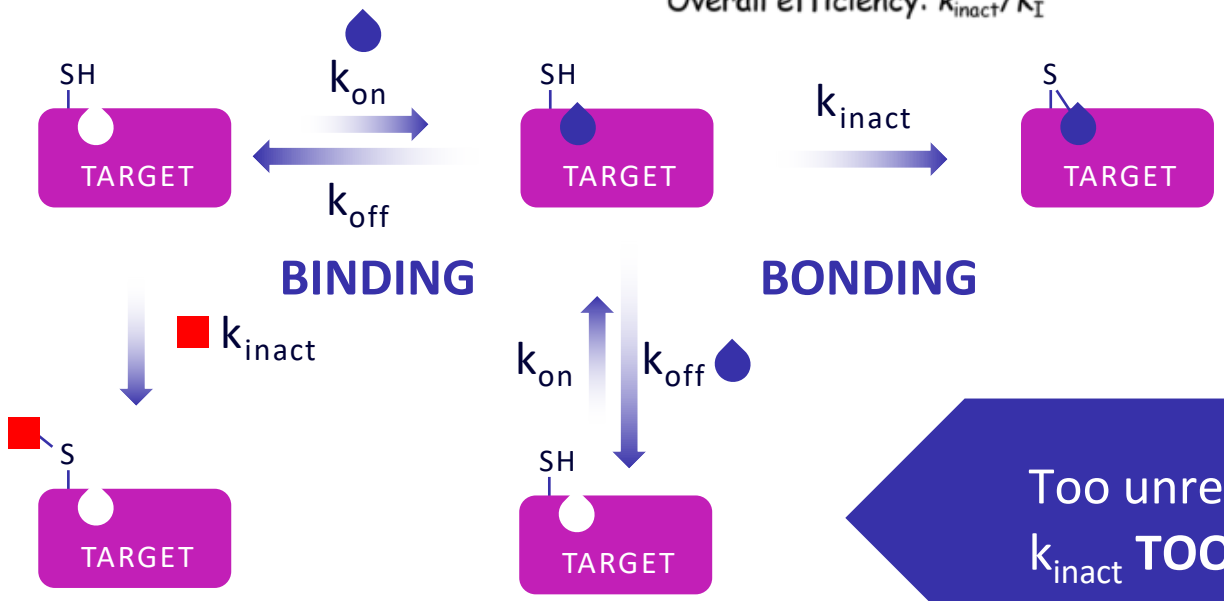
Pan et al. (2007) *ChemMedChem*. 2(1):58 10.1002/cmdc.200600221

The Goldilocks challenge: balancing fragment binding vs. reactivity

The Frontier™ Platform library is optimized for finding "Goldilocks" ligands



IDEAL:
Thermodynamic control



Reactivity-driven;
k_{inact} **TOO FAST**

Too unreactive;
k_{inact} **TOO SLOW**

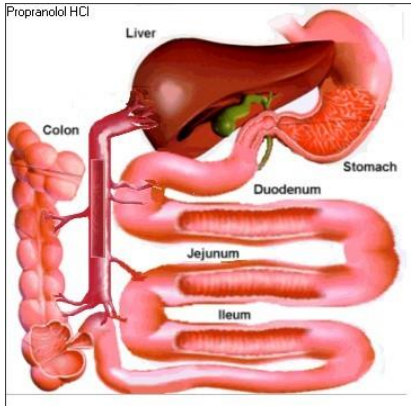
Adapting physiologically-based PK to whole medchem series

ML + PBPK => key properties and fast model for MPO

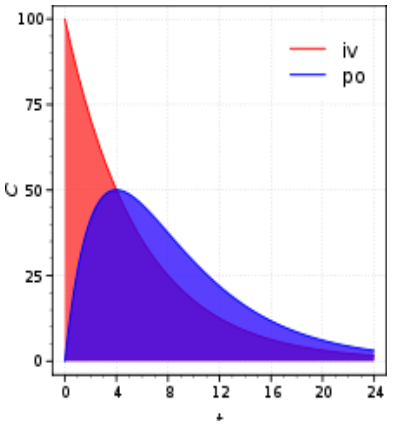
In Silico Properties

Permeability
Solubility
Ionizability
Clearance
Hydrophobicity
Protein binding

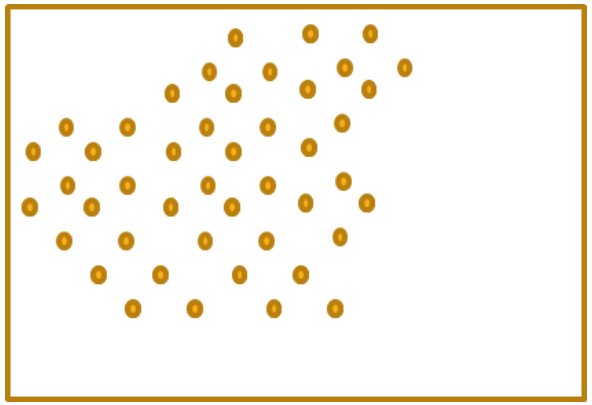
15+ rat %F's



Predicted %F

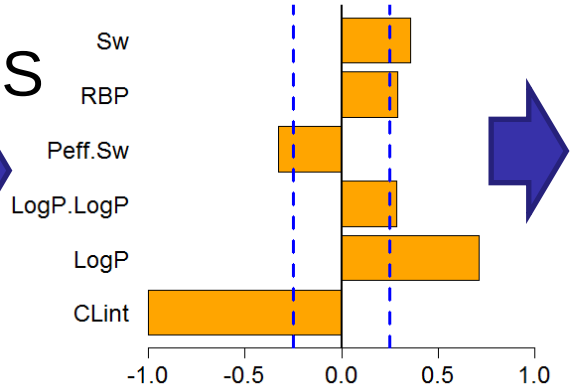


10,000 simulation

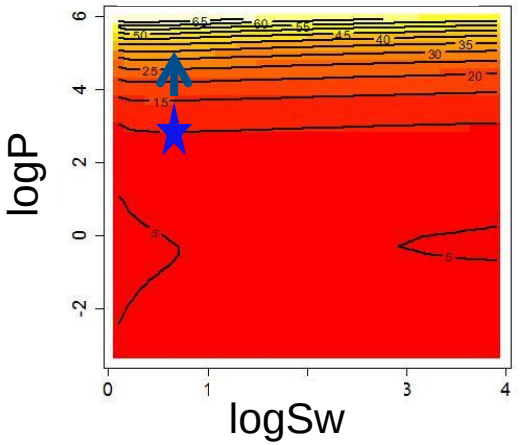


PLS

Global Sensitivities

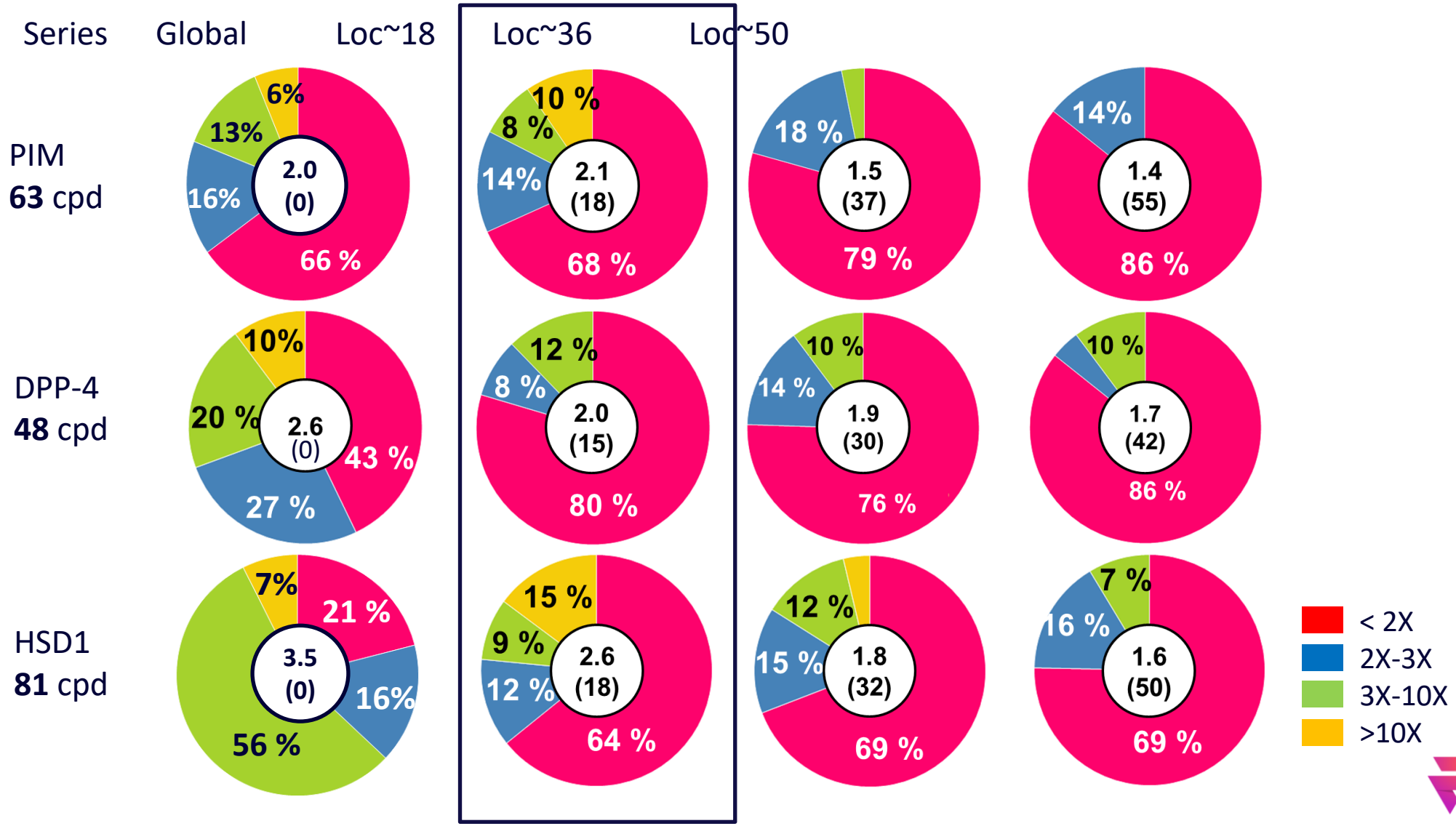


%F Landscape



OK model trained on 18 (3x6) Rat %F studies

Chronological test, avg. fold error in PBPK %F (no. training compounds)



Questions?

CDD VAULT®
COLLABORATIVE DRUG DISCOVERY

Upcoming Events

CDD Seminar Munich - Smart Labs: Enabling Data-Driven Discovery

- September 19th, 2023, Planegg/Martinsried, Germany
- Register: https://info.collaborativedrug.com/munich-seminar-2023

Conference: Discovery on Target 2023 (DOT)

- Sep 25 - Sep 28, 2023, Boston, MA: Sheraton Boston Hotel; Booth 209

Webinar: Manage Protocol/Assay Definitions in CDD Vault

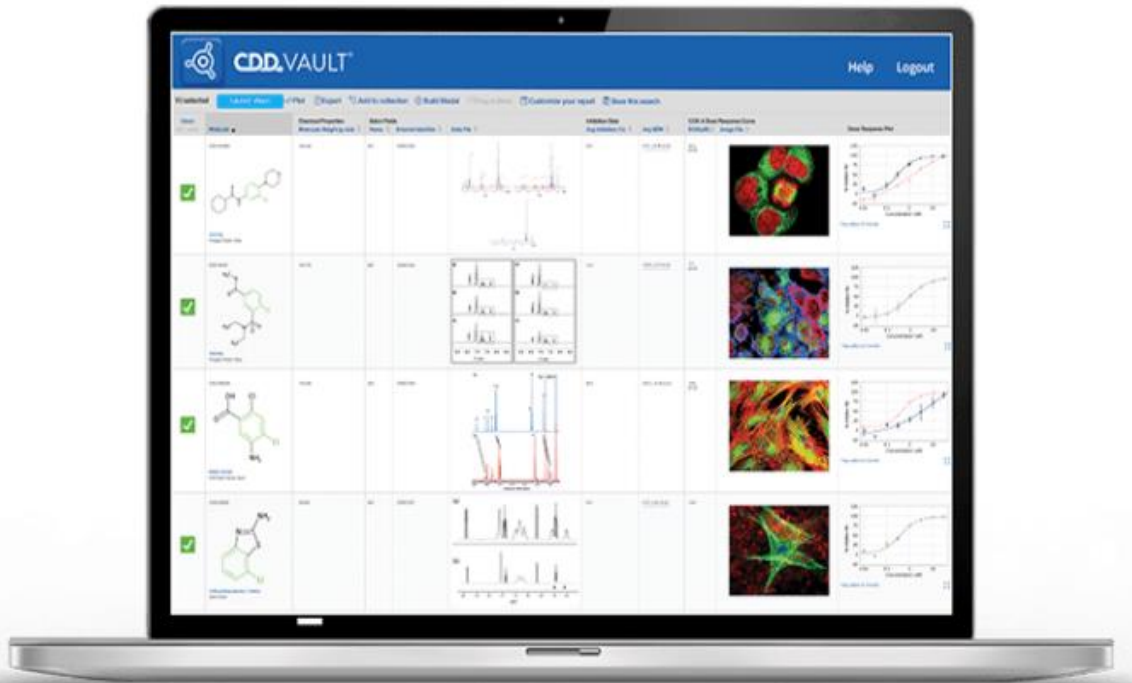
- Wednesday, September 27, 2023 9am PDT/11pm CDT/12pm EDT
- Register: https://info.collaborativedrug.com/manage-protocol-assay-definitions

CDD Vault User Group Meeting - Cambridge, MA

- Wednesday, Oct 4, 2023. 9 am - 6 pm, Cambridge MA
- Register: https://info.collaborativedrug.com/cdd-vault-user-group-meeting-cambridge-innovation-center

Activity & Registration
Store and organize your data

Inventory
Keep Track of Compounds



ELN
Document all your research

Visualization
Plot datasets and mine them

Learn More:
https://www.collaborativedrug.com

Personalized Demo & Free Trial:
info@collaborativedrug.com