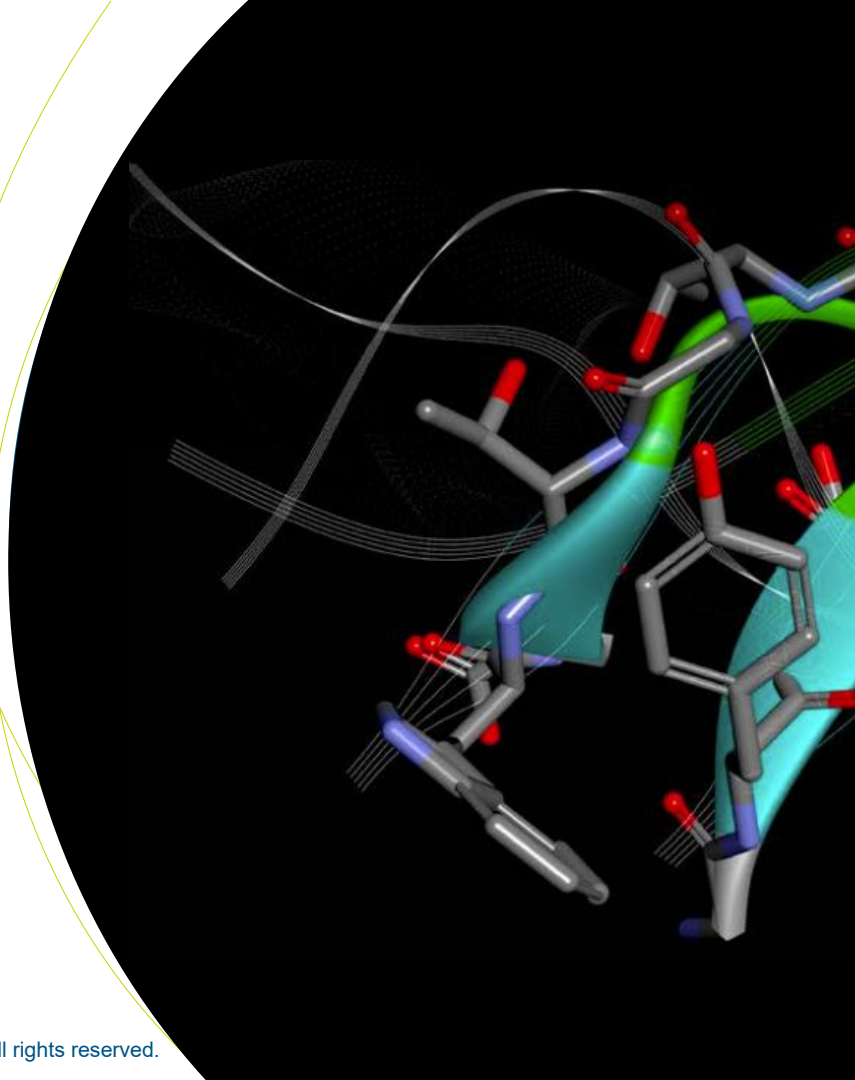
Accelerate Drug Design with AI and 3D Models



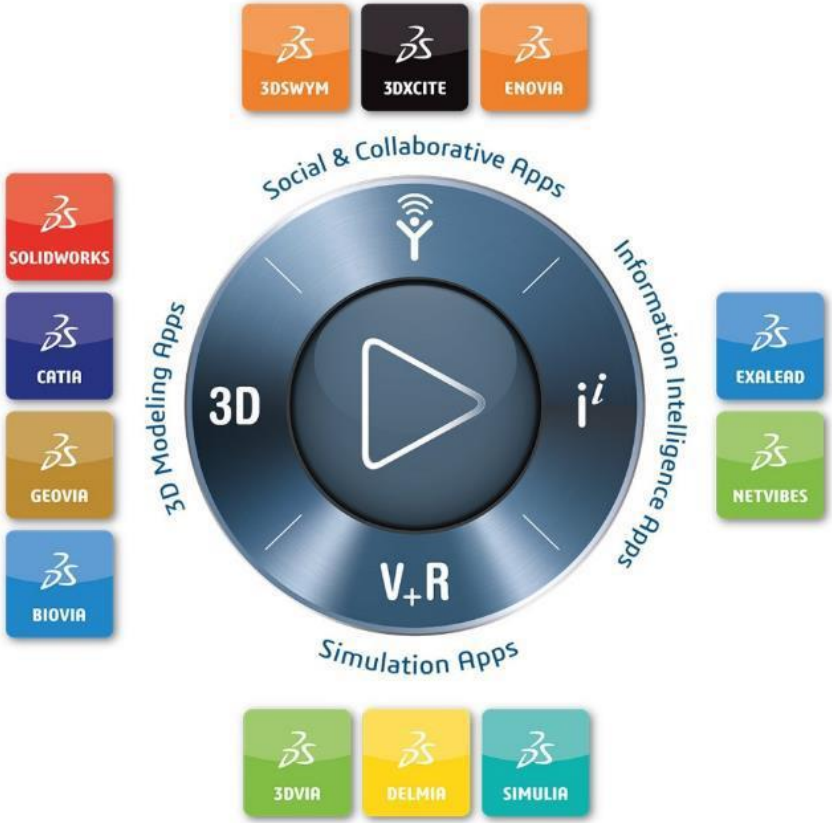
Cloud and On-Prem Developments in CY 2025

Novelyn Tsai
Engineer Specialist

3DEXPERIENCE®

Copyright and Disclaimer

- Copyright © 2025 GGA corp. All rights reserved.
  - This presentation and/or any related documents contains statements regarding our plans or expectations for future features, enhancements or functionalities of current or future products (collectively "Enhancements"). Our plans or expectations are subject to change at any time at our discretion. Accordingly, GGA Corp. is making no representation, undertaking no commitment or legal obligation to create, develop or license any product or Enhancements.
  - The presentation, documents or any related statements are not intended to, nor shall, create any legal obligation upon GGA Corp., and shall not be relied upon in purchasing any product. Any such obligation shall only result from a written agreement executed by both parties.
  - In addition, information disclosed in this presentation and related documents, whether oral or written, is confidential or proprietary information of GGA Corp.. It shall be used only for the purpose of furthering our business relationship, and shall not be disclosed to third parties.



3DEXPERIENCE Cloud Security & Privacy Overview

Dassault Systèmes Commitment

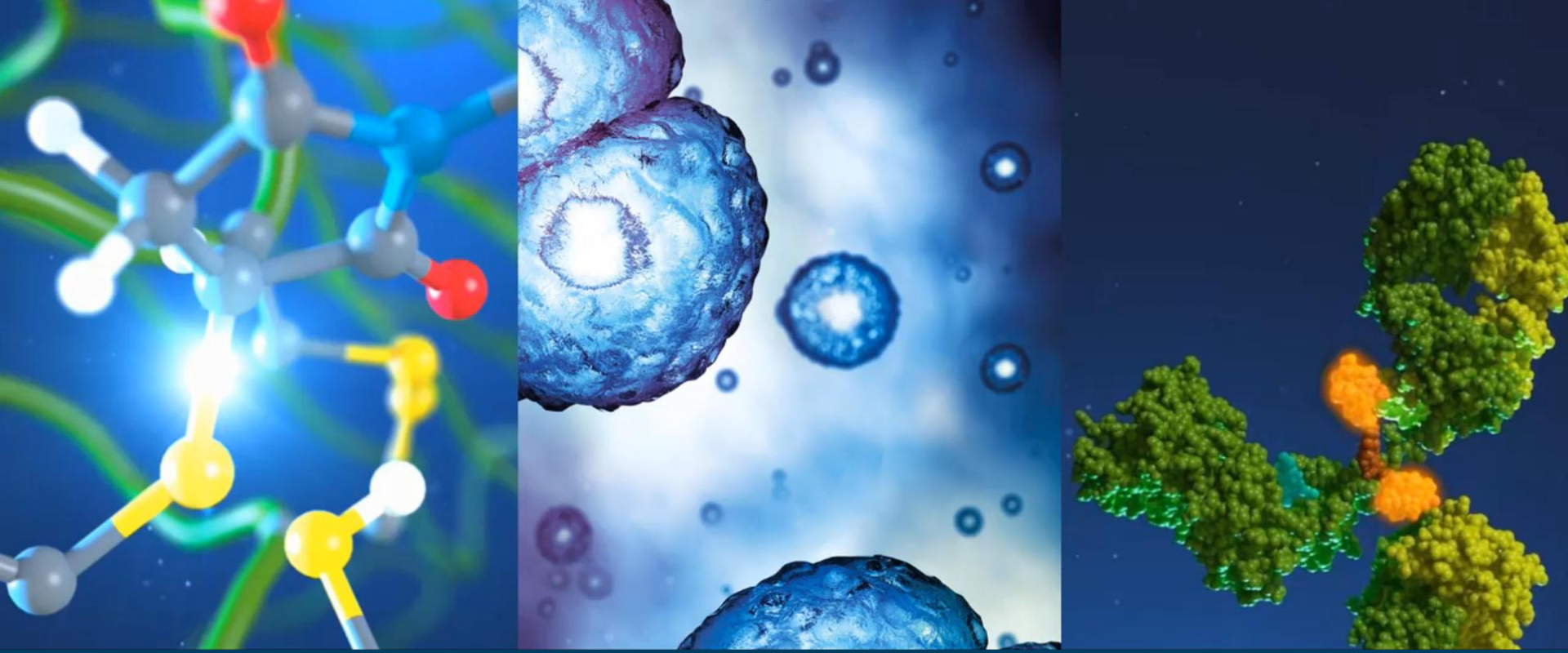
- Security and privacy are core to the **3DEXPERIENCE® platform**, supporting trust and compliance globally.
- ISO/IEC 27001:2017 & ISO/IEC 27701:2019 certified Information Security & Privacy Management System (ISPMS).
- Built on principles of **confidentiality, integrity, availability, and accountability**.

Key Highlights

- **Governance:** Centrally managed ISPMS with continuous audits and improvements.
- **Standards:** OWASP, NIST SP 800-53, ISO/IEC frameworks guide security practices.
- **Authentication:** 3D Passport with SSO, MFA, and strict access control.
- **Data Protection:** GDPR-compliant; roles as Controller & Processor clearly defined.
- **Operational Security:** Multi-layer approach (SaaS, PaaS, IaaS) with encryption, anti-DDoS, and vulnerability management.
- **Incident Response:** 24/7 monitoring via SOC, SIEM, and robust BCP/DRP for business continuity

研究發表著作

- Structure-Guided Discovery of PD-1/PD-L1 Interaction Inhibitors: Peptide Design, Screening, and Optimization via Computation-Aided Phage Display Engineering, Journal of Chemical Information and Modeling, 2024
- Biophysical mechanisms underlying tefluthrin-induced modulation of gating changes and resurgent current generation in the human Nav1.4 channel, Pesticide Biochemistry and Physiology, 2024
- Collagen-binding peptides for the enhanced imaging, lubrication and regeneration of osteoarthritic articular cartilage, Nature Biomedical Engineering, 2022
- Imaging the Cytokine Receptor CXCR4 in Atherosclerotic Plaques with [68Ga]-APD: A Novel Agent on Computer Simulation Approach, Journal of Clinical and Cellular Immunology, 2022
- In silico and in vitro studies of Taiwan Chingguan Yihau (NRICM101) on TNF-α/IL-1βinduced Human Lung Cells, BioMedicine, 2022
- Helicobacter pylori Targets in AGS Human Gastric Adenocarcinoma: In Situ Proteomic Profiling and Systematic Analysis, ANTICANCER RESEARCH, 2022



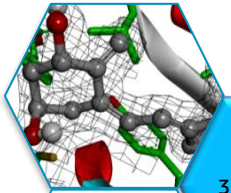
Discovery Studio

Small Molecule and Biologics Lead Identification & Optimization

DISCOVERY STUDIO = SIMULATIONS FOR DRUG DESIGN

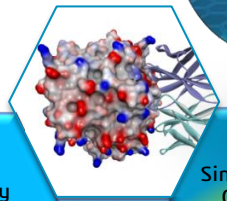
Bio-therapeutics Design

Small Molecules Design



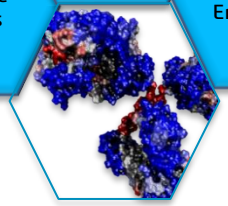
3D structure

Homology Modeling



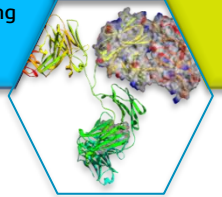
Protein-Protein Docking

Antibody Design



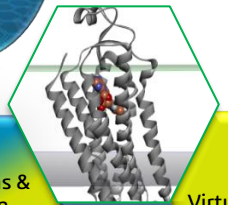
Sequence Analysis

Protein Engineering



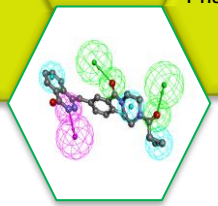
Protein Developability

Simulations & Quantum Mechanics



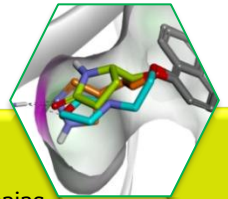
Free Energy Simulations

Scaffold Hopping



Fragment Replacement

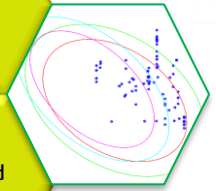
Virtual Screening



Library Enumeration & Analysis

Pharmacophore Profiling

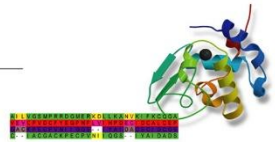
QSAR



ADME and Toxicity

Modeller

Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints



NAMD Scalable Molecular Dynamics

ILLINOIS

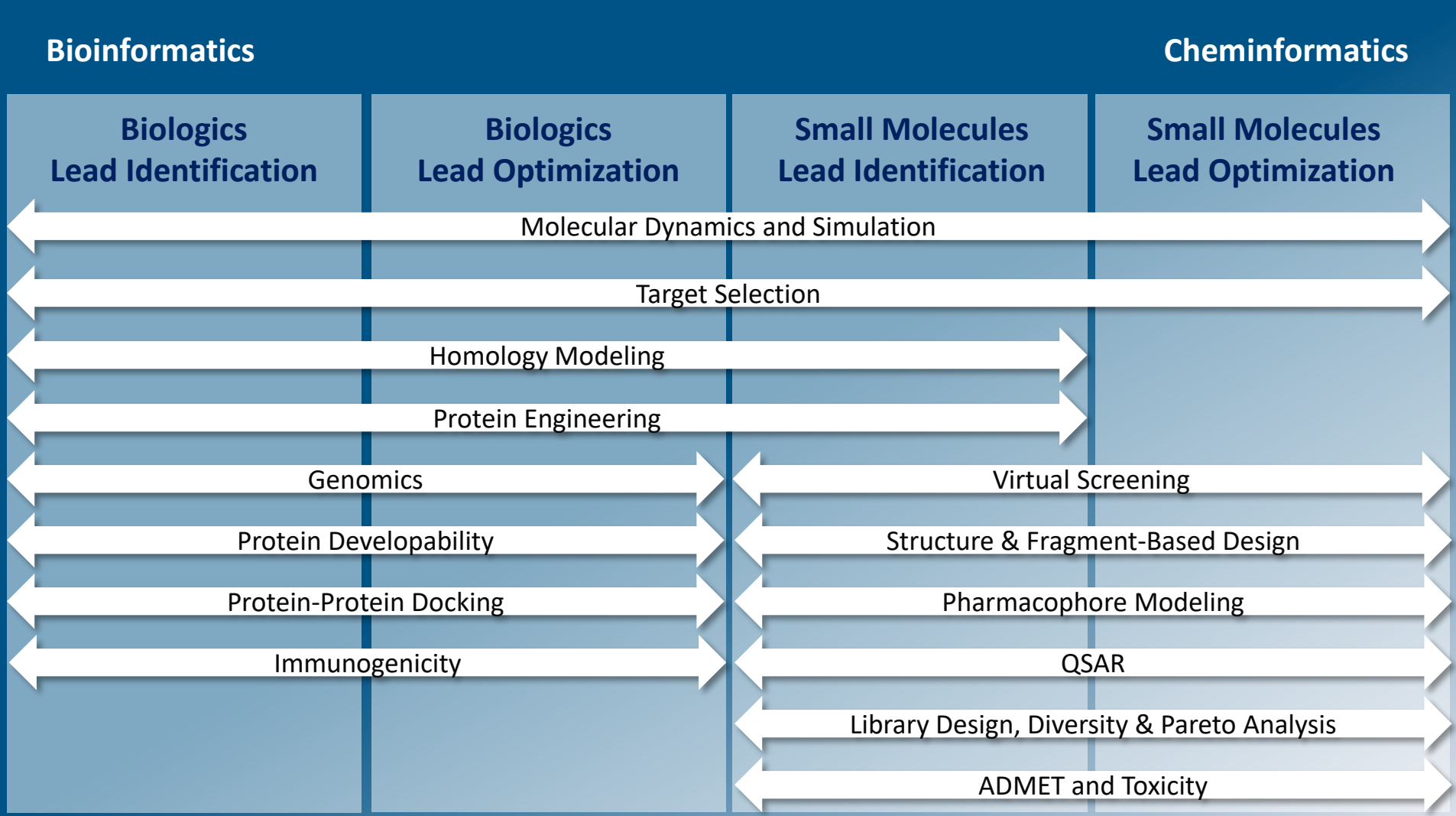
UCSF University of California San Francisco

Massachusetts Institute of Technology

HARVARD

BIOVIA

DASSAULT SYSTEMES


Bioinformatics
Cheminformatics
Biologics Lead Identification
Biologics Lead Optimization
Small Molecules Lead Identification
Small Molecules Lead Optimization
Molecular Dynamics and Simulation
Target Selection
Homology Modeling
Protein Engineering
Genomics
Protein Developability
Protein-Protein Docking
Immunogenicity
Virtual Screening
Structure & Fragment-Based Design
Pharmacophore Modeling
QSAR
Library Design, Diversity & Pareto Analysis
ADMET and Toxicity

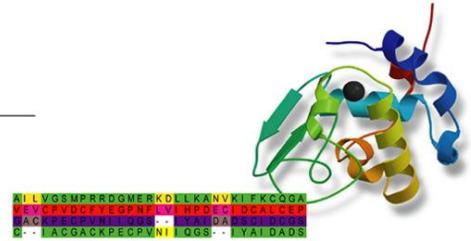
Best Validated Science – 30+ Years History

NAMD
Scalable Molecular Dynamics

- Force-field simulations: CHARMm
- Force-field simulations: NAMD
- Protein homology modeling: MODELLER
- Protein-protein docking: ZDOCK
- Protein aggregation & viscosity: AggMap, SCM
- Pharmacophore: Catalyst
- And Many more novel, internally developed, peer reviewed scientific algorithms

Modeller

Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints



UCSF
University of California San Francisco

Massachusetts Institute of Technology

accelrys®

UMASS MEDICAL SCHOOL

VE RI TAS HARVARD

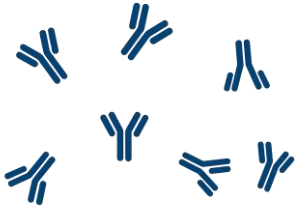
Agenda

- Scientific Platform Portfolio – R&D to Manufacturing
- **AI in Drug Discovery**
- Q & A

THE DISCOVERY BOTTLENECK

5+ Years

Average time spent in discovery



Thousands of Biologics

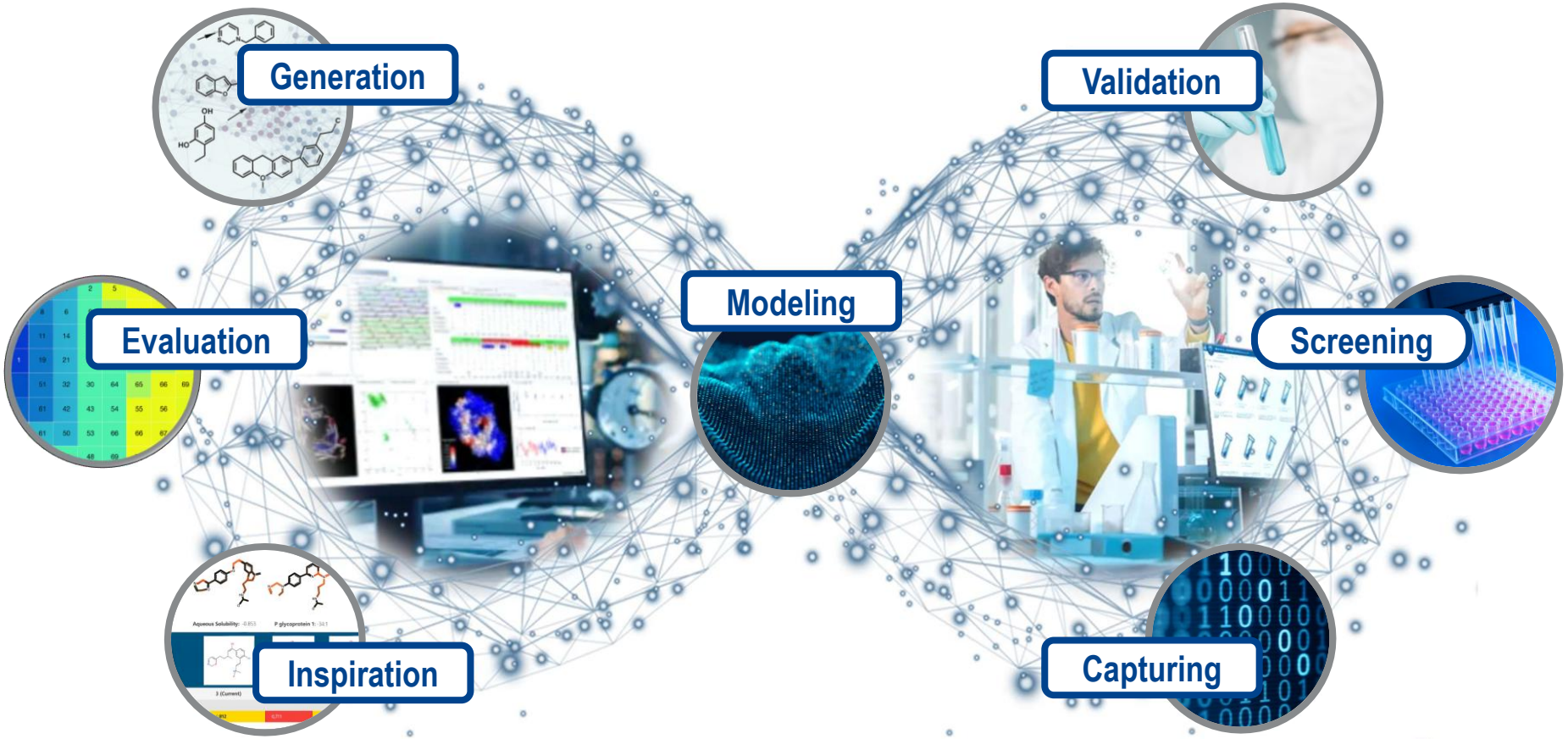
To find a viable candidate

4,000 Compounds

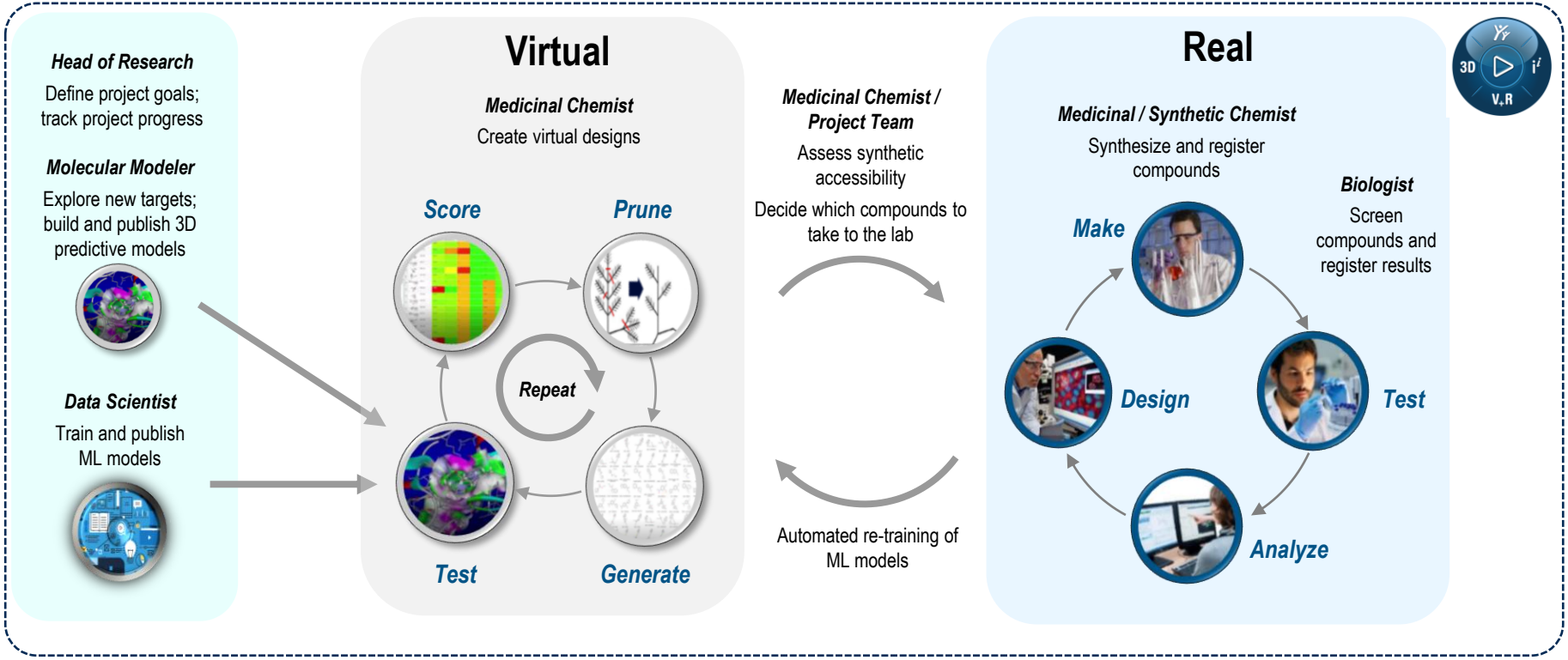
To find a viable candidate

Can AI do better?

ACTIVE AI-ASSISTED V+R CYCLE AIDS DRUG DISCOVERY

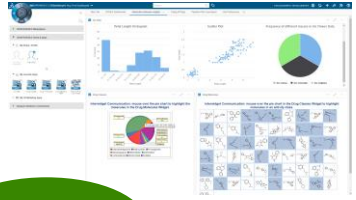


GENERATIVE THERAPEUTICS DESIGN PROCESS



COMPREHENSIVE SOLUTION FOR EARLY DISCOVERY with 3DE

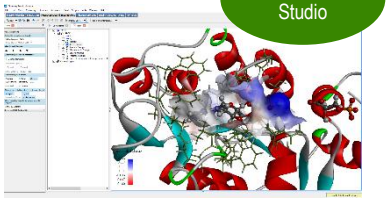
1 *Search Prior Art*



Experimental knowledge

2 *Build 3D Models*

Discovery Studio



3 *Build Machine Learning Models*

Model Report for FPR1_IC50 version 1

Type	Target	Algorithm	ROC	Accuracy	Precision	Recall
Classification	IC50	Bayesian - Chemistry	0.8234/NaN	0.8175/NaN	0.3609/NaN	0.6154/NaN

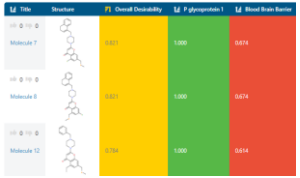
Training Set

Confusion Matrix

Actual\Pred.	Active	not_Active
Active	48	30
not_Active	85	467

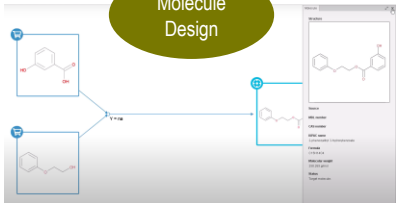
Machine Learning Workbench

4 *Generative Therapeutics Design*



5 *Make & Register Compounds*

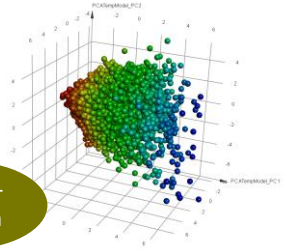
Molecule Design



6 *Test Compounds*

Scientific Notebook

7 *Analyze V+R Results (SAR)*



Insight for Research

8 *Retrain Machine Learning Models*

Model Report for FPR1_IC50 version 1

Type	Target	Algorithm	ROC	Accuracy	Precision	Recall
Classification	IC50	Bayesian - Chemistry	0.8234/NaN	0.8175/NaN	0.3609/NaN	0.6154/NaN

Training Set

Confusion Matrix

Actual\Pred.	Active	not_Active
Active	48	30
not_Active	85	467

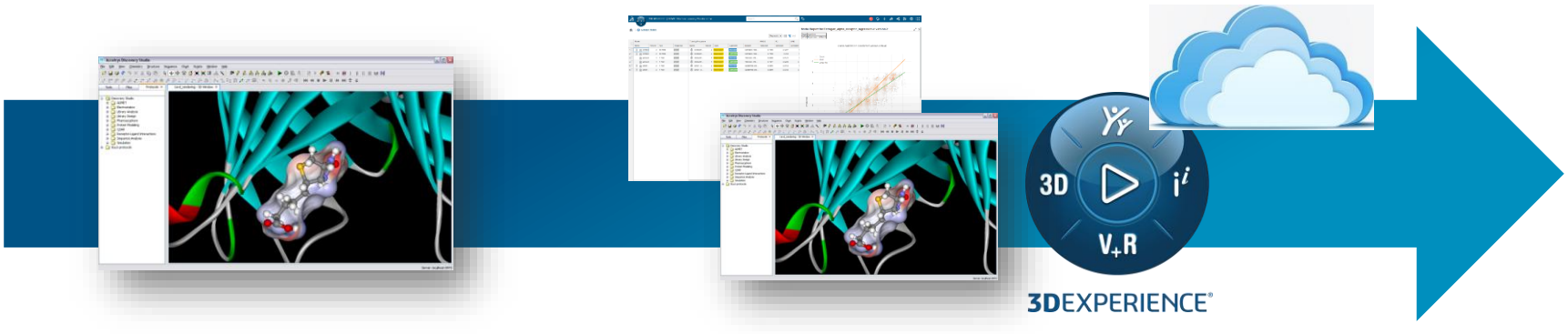
3D V+R

3DEXPERIENCE

The Public Cloud Biosphere and Materials

DISCOVERY STUDIO SIMULATION

Bringing 30+ years of molecular modeling and therapeutics design to 3DEXPERIENCE Cloud



Discovery Studio

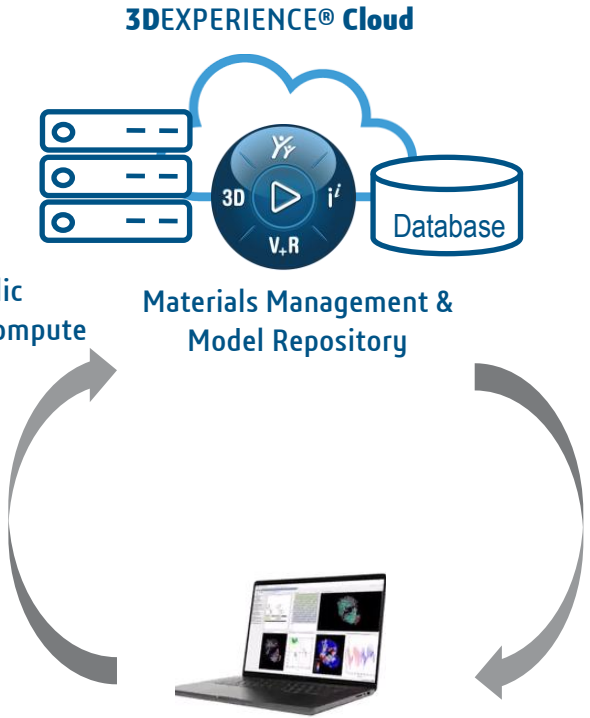
On-premise application
Expert users working with a rich client
Shareable user licenses

Discovery Studio Simulation

SaaS 3DEXPERIENCE Platform
Simple access to high-performance cloud computing
Regular updates and new scientific functionality
Industry process focus – beyond an application
Easy install of rich client, familiar to expert users
Named user with usage-based licenses

DISCOVERY STUDIO - CLOUD

Powered by 3DEXPERIENCE SaaS Cloud



DISCOVERY STUDIO SIMULATION

Benefits of 3DEXPERIENCE Cloud

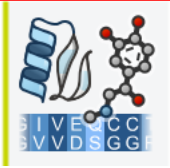
AVAILABILITY OF DISCOVERY STUDIO SIMULATION

Applications	Computational Chemist	Computational Structural Biologist	Medicinal Chemist
3DE Platform	O	O	O
Discovery Studio Simulation	O	O	X
Generative Therapeutics Design	O	X	O
Insight for Research	O	O	O
Job Management	O	O	O
Molecular Design	O	O	O
Scientific Notebook	O	O	O
Scientific Search Intelligence	O	O	O
Scientific Settings and Administration	O	O	O
Materials Management	O	O	O
Machine Learning Workbench	O	O	X
Reaction Planner	X	X	O

AI in Drug Discovery

Comprehensive Applications in Small and Biological Drug Design

3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

Use AI strategies to design protein binders and predict structures with AlphaFold.

Data visualization and Analysis

Insight for Research

Visualize your data and connect it to physics-based models.

AI in Small Molecule Design

Machine Learning Workbench

Train ML models with your experimental data.

Generative Therapeutics Design

Automate the virtual creation, testing and selection of drug-like compounds using ML models and scientific methods.

Reaction Planner

AI-based retrosynthesis prediction tool

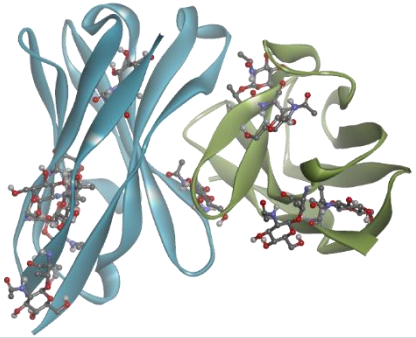
RELEASES

DRUG THERAPEUTICS DESIGN

What Therapeutic Do You Have?

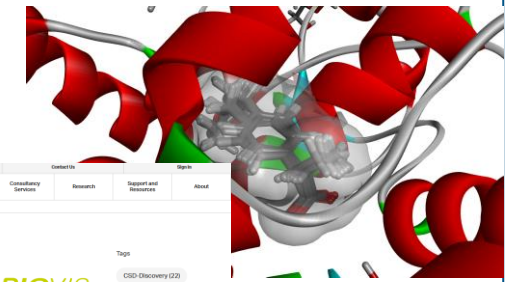
Biological Therapeutics (Biotherapeutics)

Recombinant proteins
Monoclonal antibodies, bi-specifics, antibody drug conjugates (ADCs)
Antibody-like modalities (eg. darpins)
scFv and nanobody-based CAR T-cells
Vaccine design

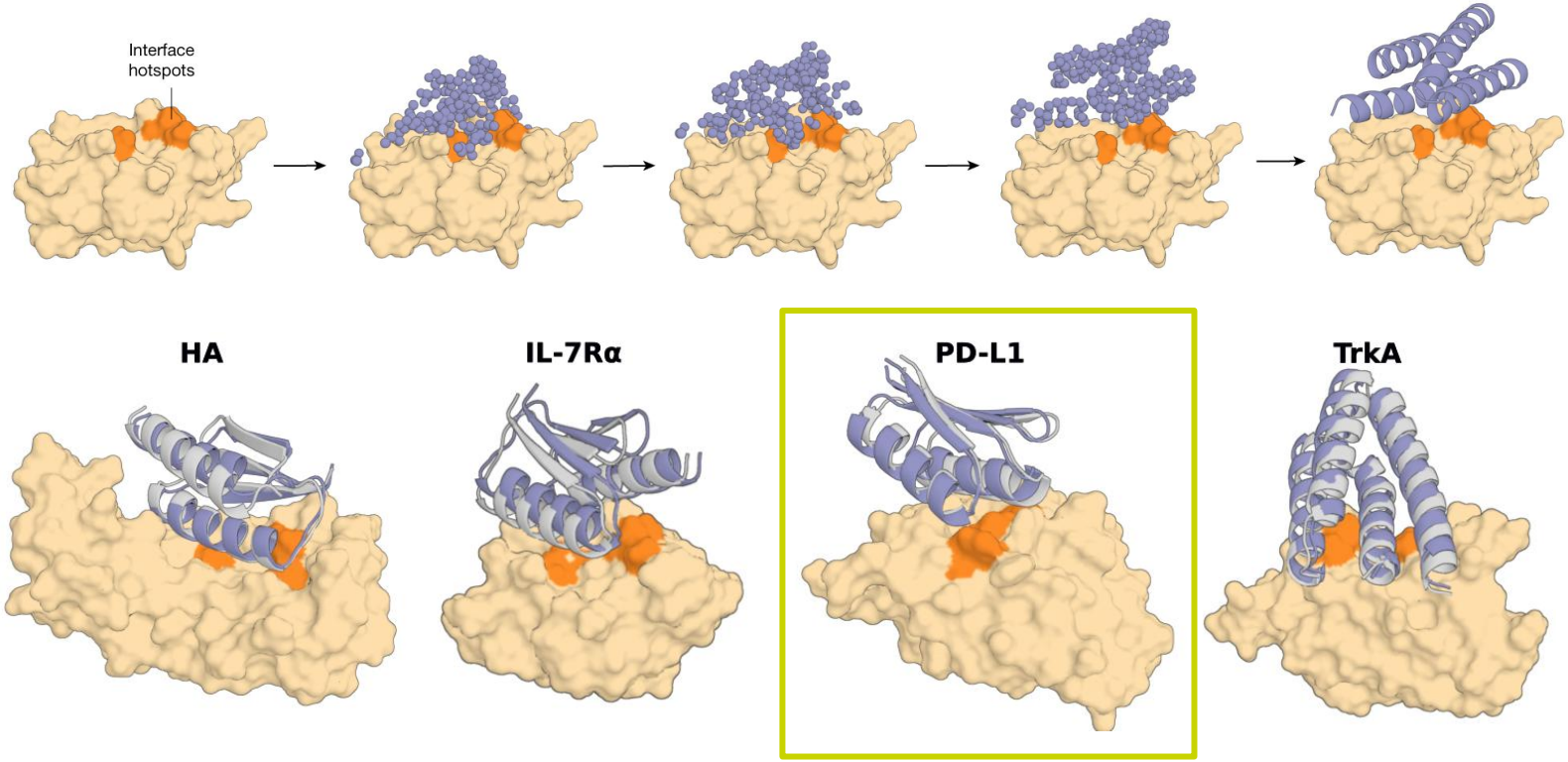


Small Molecule Therapeutics

Chemical compounds
Small molecule inhibitors
Small peptides



DE NOVO DESIGN OF PROTEIN-BINDING PROTEINS



Watson et al. De novo design of protein structure and function with Rfdiffusion. *Nature* (2023).

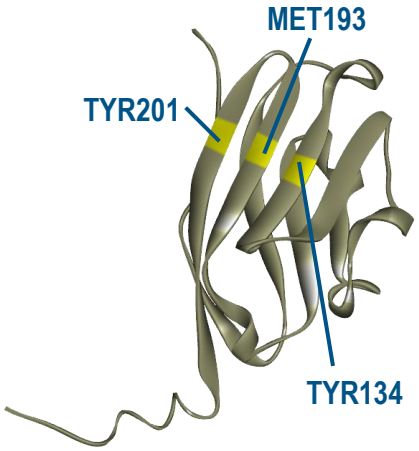
Generate protein scaffolds from a template structure using RFDiffusion.

RFDiff PDL1_binder_design *Nature* volume 620, pages1089–1100 (2023)

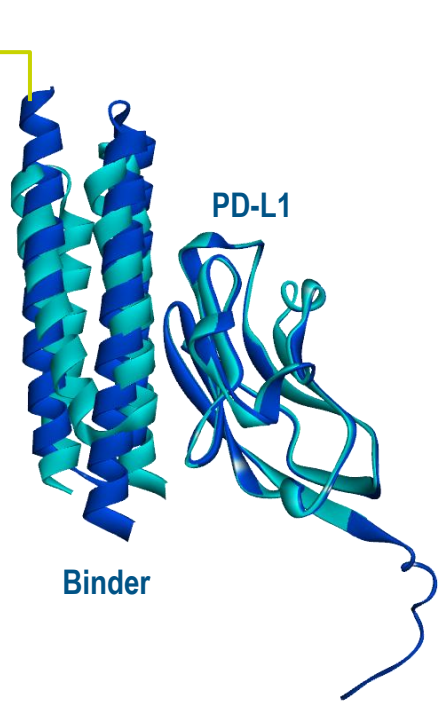
RFDiff PDL1_binder_design (Discovery Studio)

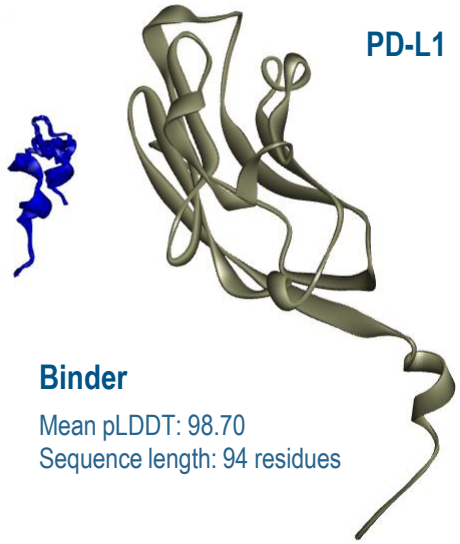
PDB: 5O45

Select hotspot residues for binder design



Parameter Name	Parameter Value
Template Scaffold	PD-L1_RFDiff:5O45_0
Use Hotspots	True
Hotspot Residues	5O45_0:Hotspot Residues
Number of Designs	10
Diffusion Time Steps	50
Design Configuration	## This file is for the "Generate Protein Scaffold"
Model Weights	Complex Base





A:17-145
60-100

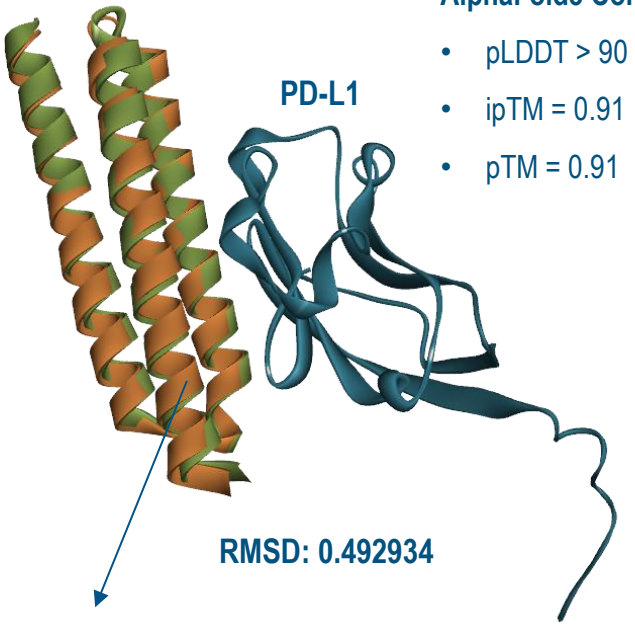
Generate a potential binder containing 60-100 residues
→ Targeting hotspot residues 17-145 from chain A in the template.

Generate protein sequences for PD-L1 Binder

structure_name	id	T	seed	overall_confidence	ligand_confidence	seq_rec	sequence
PD-L1_RFDiff_73	73	0.1	111	0.3968	0.3968	0.2735	AVEEARRRLEELLERARRELERLLARVAAAAPAELDALTDEFLLALSRLRRETRRLIER DPEIAEEMRRRTEELIRRALAEFDAAVRAKLAAAA
PD-L1_RFDiff_77	77	0.1	111	0.3963	0.3963	0.2511	AAAAARAELDALLAAADATAAALLAAVAAAAPAELTALRDAFMLALSDLARATAALV ARDPALAAEFRARTEARIRRWLTEFDAAVRAKAAAAA
PD-L1_RFDiff_35	35	0.1	111	0.3939	0.3939	0.2422	MEEEARRRLEELLREFEERLRRLLARVAAADPAELDALRDEFLLAISELSRRTKELIEE YPELAEEFRRRLEEAIRRAQREFDAAVRAHLAAAA
PD-L1_RFDiff_27	27	0.1	111	0.3931	0.3931	0.2466	AAAEARARLDALLAEAEATARRLLERVAAAAPEELDALRDEALLAVSRLRRETAALIA EHPDLAAEMRERTEAAILRWLREFDAAVRAKLRAAE
PD-L1_RFDiff_29	29	0.1	111	0.3926	0.3926	0.2287	SMEEARKKLEELLEEAERRLRELLARVAAAAPAELDALTDEFLLALSELSRRTKELIEK YPEIAEEARKKTEELIRRHVEEFTAAVRAKRAAAA
PD-L1_RFDiff_8	8	0.1	111	0.3923	0.3923	0.2466	ATAEADARLDALLAAADAAAAALLARVAAAAPAELSDLRDEAMLAISRLRRETAALIA ADPARAEEFRKRTEERTLEHLRALDAAFRAHAAAAA
PD-L1_RFDiff_25	25	0.1	111	0.3919	0.3919	0.2511	SAAAAEARLDALLEAADAAAAALLARVAAAEPAELGALRDEFLLAVSDLARATAALIA EDSELAAEFRERTEARIDQWFRDFDAAVRAHAAAAA
PD-L1_RFDiff_49	49	0.1	111	0.3906	0.3906	0.2466	MEEEARRRLEELLRRAEEEHRELLERVEEAELEELPELKNEFMLALSRLRRETAALIAE YPELAEEMRRRTRELILKMTREFIEAVRRKREEAK

Sequence length: 94 residues

Predict protein structure for PD-L1 Binder

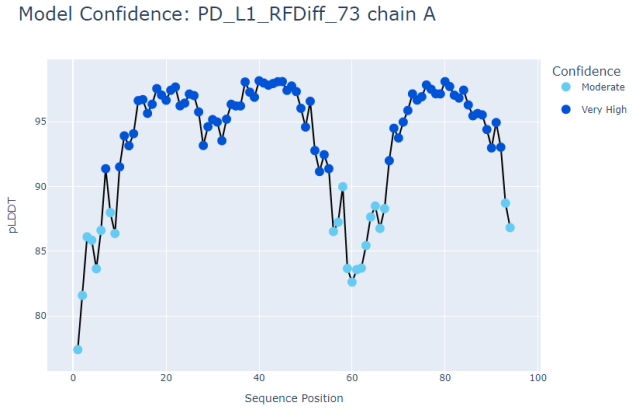


AlphaFold3 Confidence:

- pLDDT > 90 (Very High)
- ipTM = 0.91
- pTM = 0.91

RFDiff Structure Prediction Confidence:

- Mean pLDDT = 93.44



PDL1_binder_design (Alphafold3)
PDL1_binder_design (RFDiff)

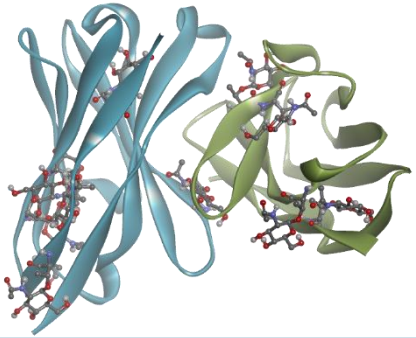
Lower RMSD indicates closer structural alignment between designed binder from RFDiffusion and Alphafold3

DRUG THERAPEUTICS DESIGN

What Therapeutic Do You Have?

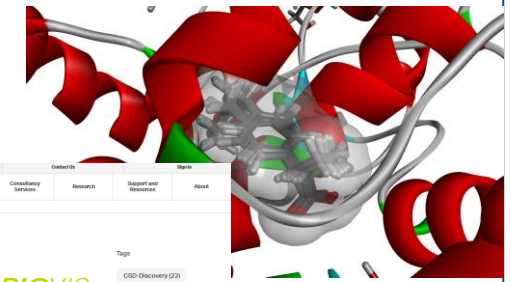
Biological Therapeutics (Biotherapeutics)

Recombinant proteins
Monoclonal antibodies, bi-specifics, antibody drug conjugates (ADCs)
Antibody-like modalities (eg. darpins)
scFv and nanobody-based CAR T-cells
Vaccine design



Small Molecule Therapeutics

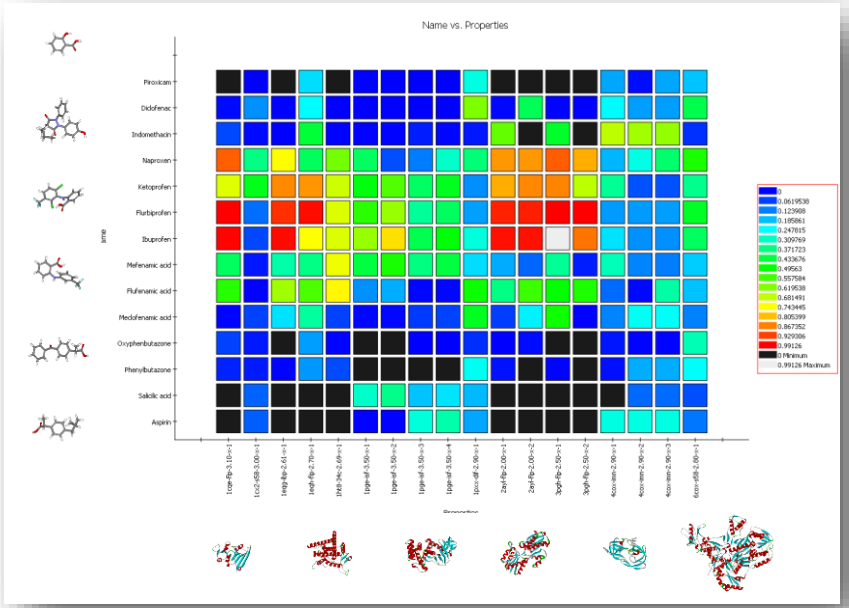
Chemical compounds
Small molecule inhibitors
Small peptides



On Premise

PharmaDB Profiling

- Rapidly screen libraries of ligands against multiple pharmacophores
  - Predicting protein-drug off-targets (side effects)
  - Repositioning/repurposing existing drugs
  - In silico target fishing



- PharmaDB
  - Validated in collaboration with Prof. Rognan at University of Strasbourg*
  - Derived from the scPDB (http://bioinfo-pharma.u-strasbg.fr/scPDB)
  - More than 40,000 validated models
  - Classified using Kyoto Encyclopedia of Genes and Genomes (KEGG)-BRITE

* Kellenberger *et al*, *J Chem Info Model*, **2006**, 46, 717-727

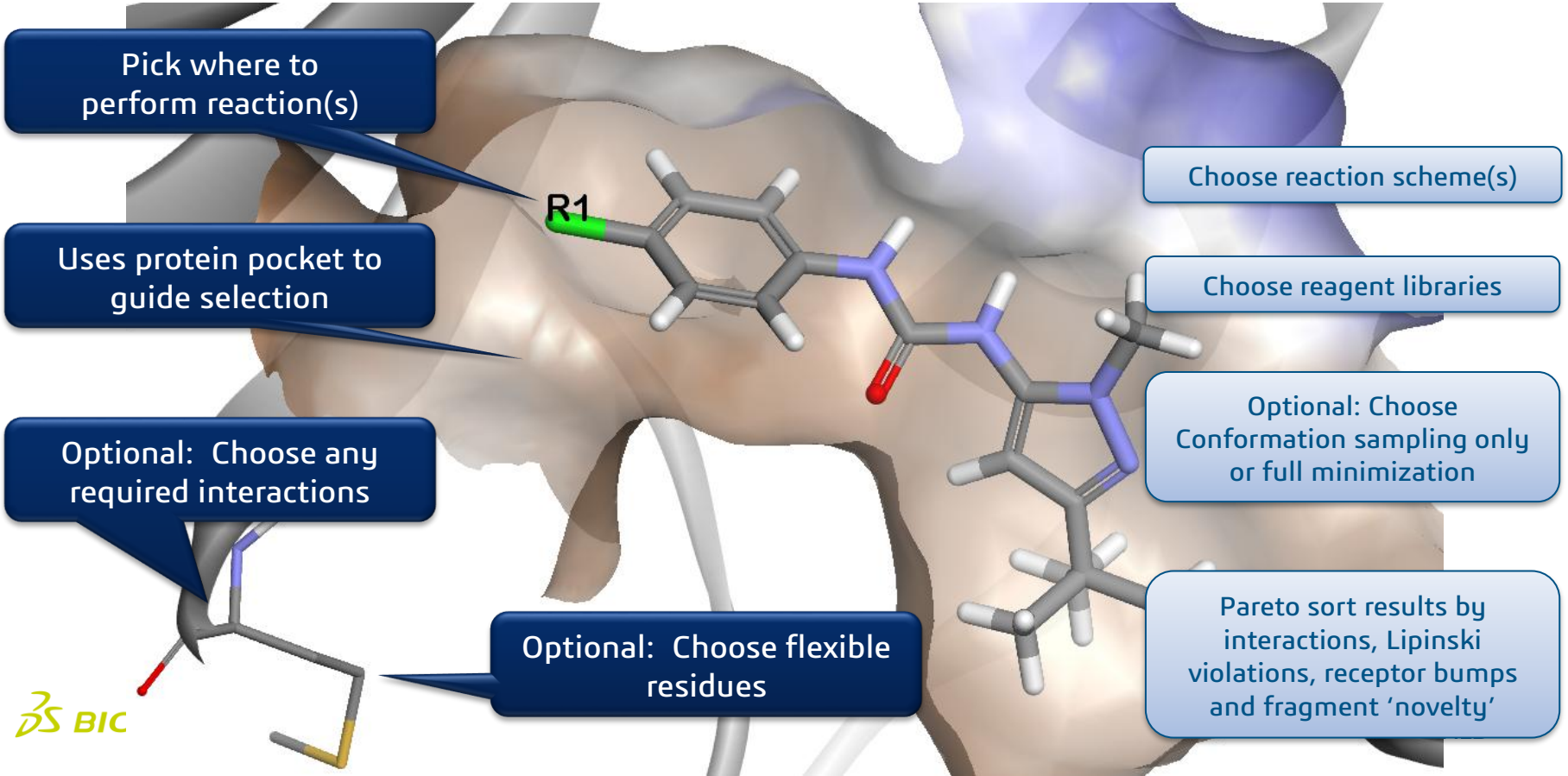
What if the hit has multiple targets?

SBD: Fragment-Based Design Methods

On Premise

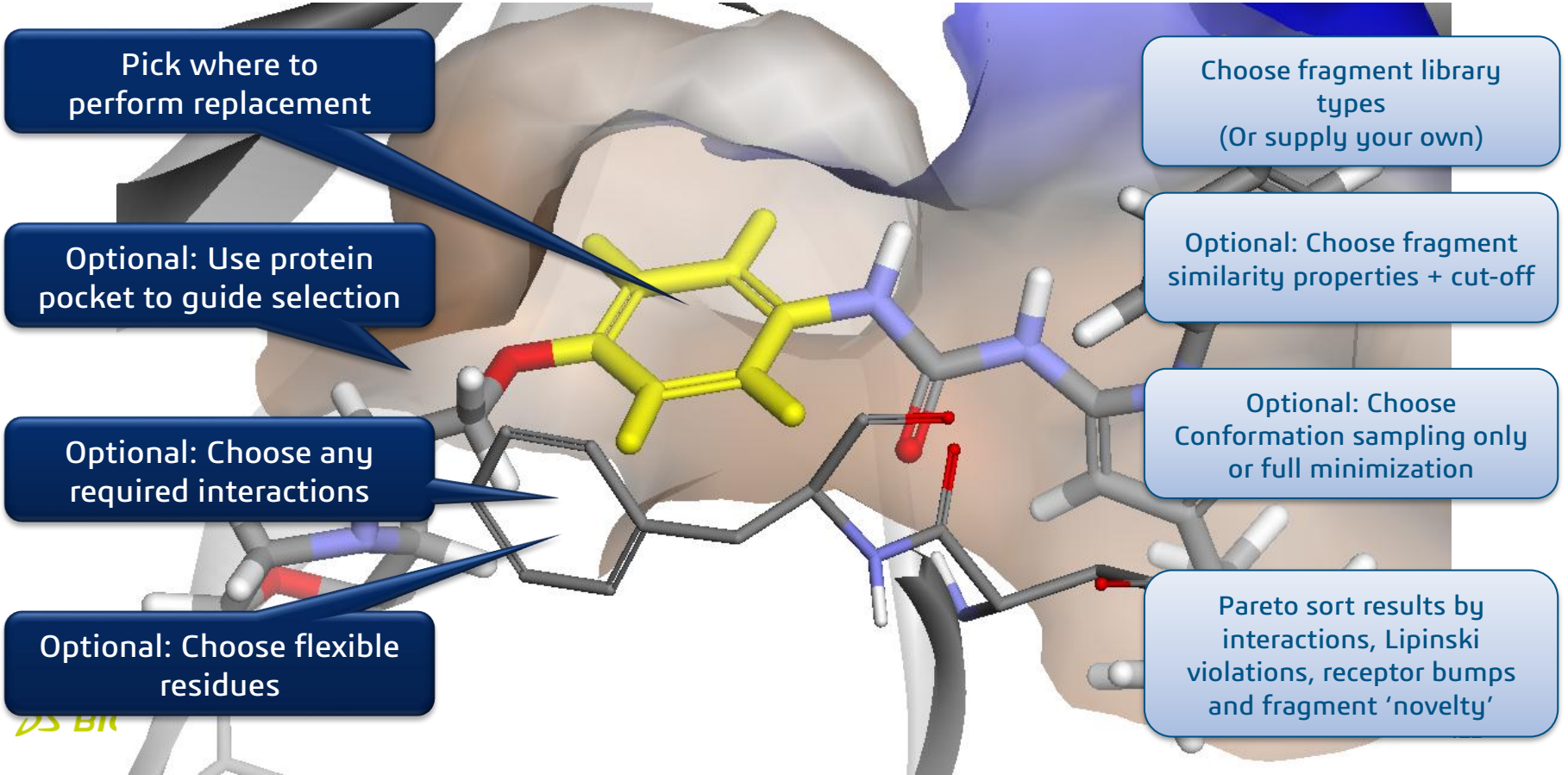
- GROW
  - Reaction-based in situ ligand enumeration
  - E.g., Amide synthesis, Esterification, Hiyama, Kuyama, Negishi, Stille, Suzuki, Williamson Ether
  - Pre-filtered sets of reagents selected from ACD
- REPLACE
  - Fragment based in situ isostere replacement
  - E.g., scaffold-hopping, R-group replacement
  - Pre-filtered set of 1.5M fragments generated from SCD

GROW: Reaction-based *in-situ* Ligand Optimization



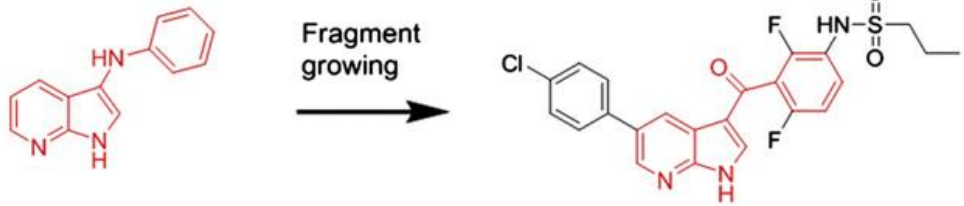
On Premise

REPLACE: Fragment Based *In-Situ* Substitution



Fragment-based design of the BRAF inhibitor vemurafenib.

First fragment-based drug (Zelboraf) approved in 2011!



4

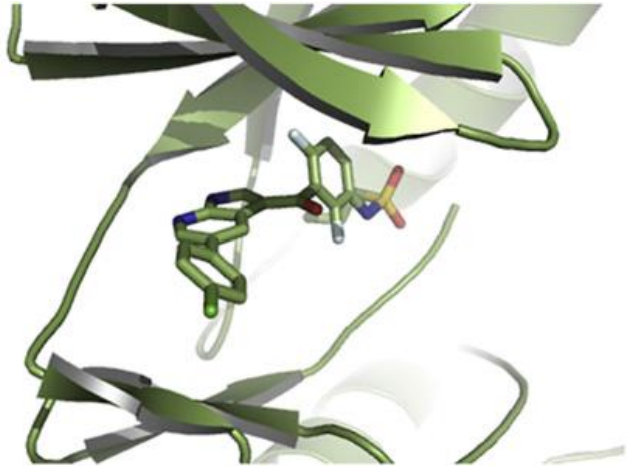
Unselective, weakly potent fragment hit (IC_{50} > 100μM)

PIM-1 IC_{50} ~ 100μM

5 (Vemurafenib)

BRAF (V600E) IC_{50} = 31nM
High degree of selectivity against other kinases

PIM-1 IC_{50} > 100μM



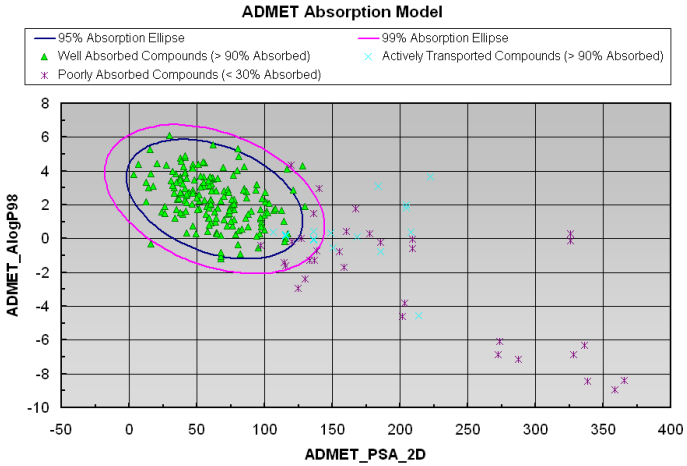
Swen Hoelder, Paul A. Clarke, Paul Workman, 2012

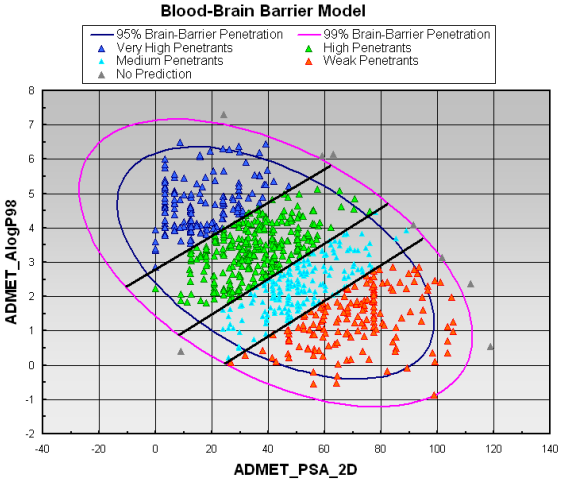
On Premise

ADME Descriptors

Used the **QSAR models** to estimate a range of ADMET related properties for small molecules. The following properties, and classes of properties, can be computed:

- Aqueous solubility
- Blood brain barrier penetration (BBB)
- Cytochrome P450 (CYP450) 2D6 inhibition
- Hepatotoxicity
- Human intestinal absorption (HIA)
- Plasma protein binding

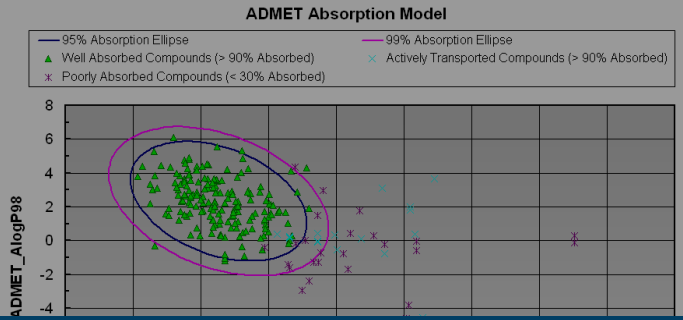




On Premise

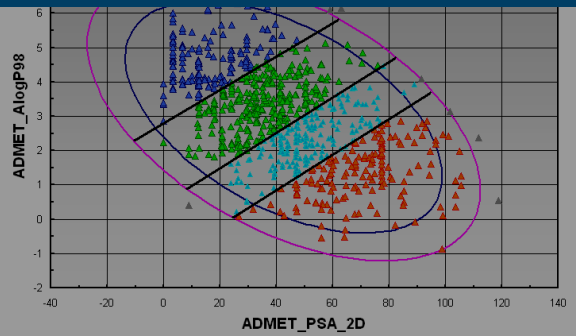
ADME Descriptors

Used the **QSAR models** to estimate a range of ADMET related properties for small molecules. The following properties, and classes of



The limitation of ADME descriptor tools in Discovery Studio on-premise is that they **cannot optimize for multiple TPP profiles simultaneously** and provide **no guidance on how to modify compounds to meet those profiles.**

- Hepatotoxicity
- Human intestinal absorption (HIA)
- Plasma protein binding



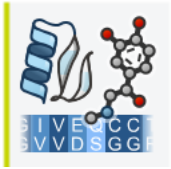
Can we use AI for optimization?

Generative Therapeutic Design

AI in Drug Discovery

Comprehensive Applications in Small and Biological Drug Design

3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

Use AI strategies to design protein binders and predict structures with AlphaFold.

Data visualization and Analysis

Insight for Research

Visualize your data and connect it to physics-based models.

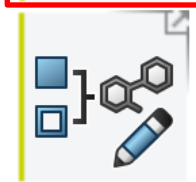
AI in Small Molecule Design

Machine Learning Workbench

Train ML models with your experimental data.

Generative Therapeutics Design

Automate the virtual creation, testing and selection of drug-like compounds using ML models and scientific methods.

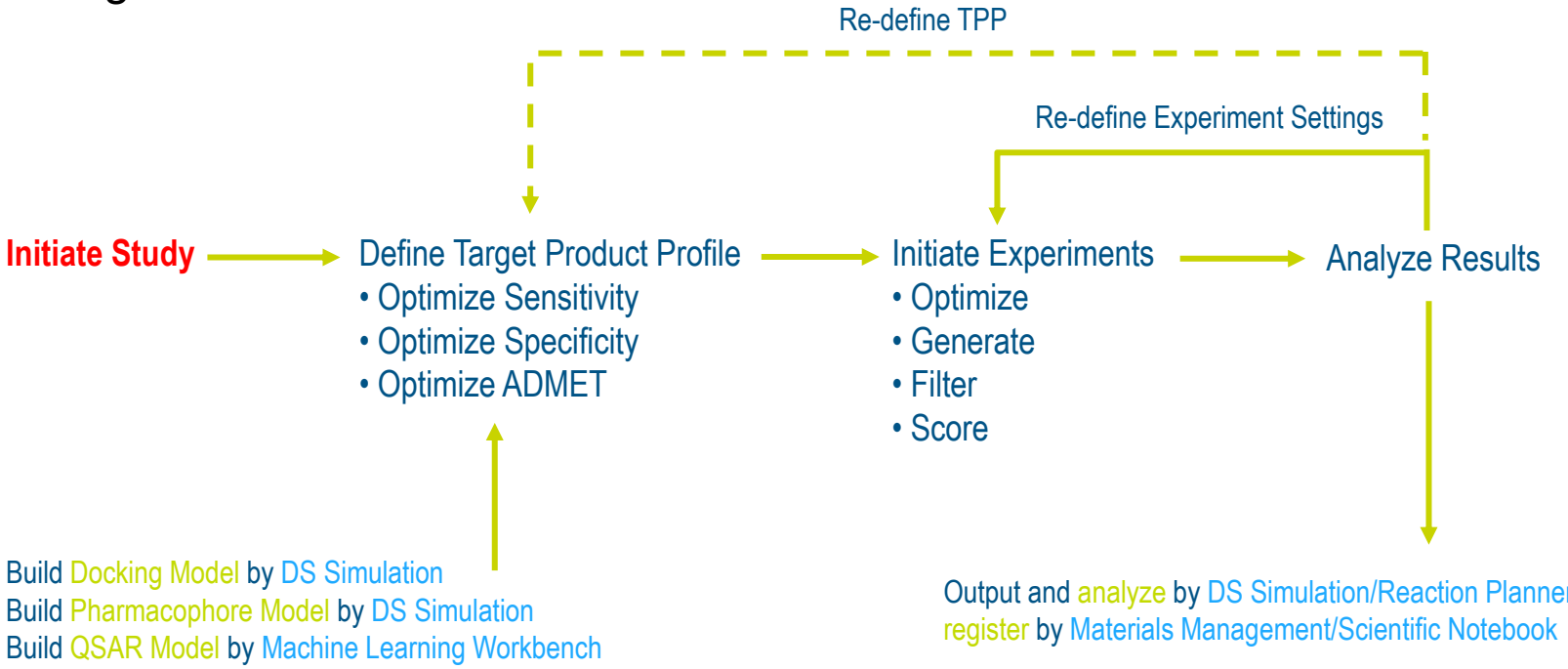


Reaction Planner

AI-based retrosynthesis prediction tool

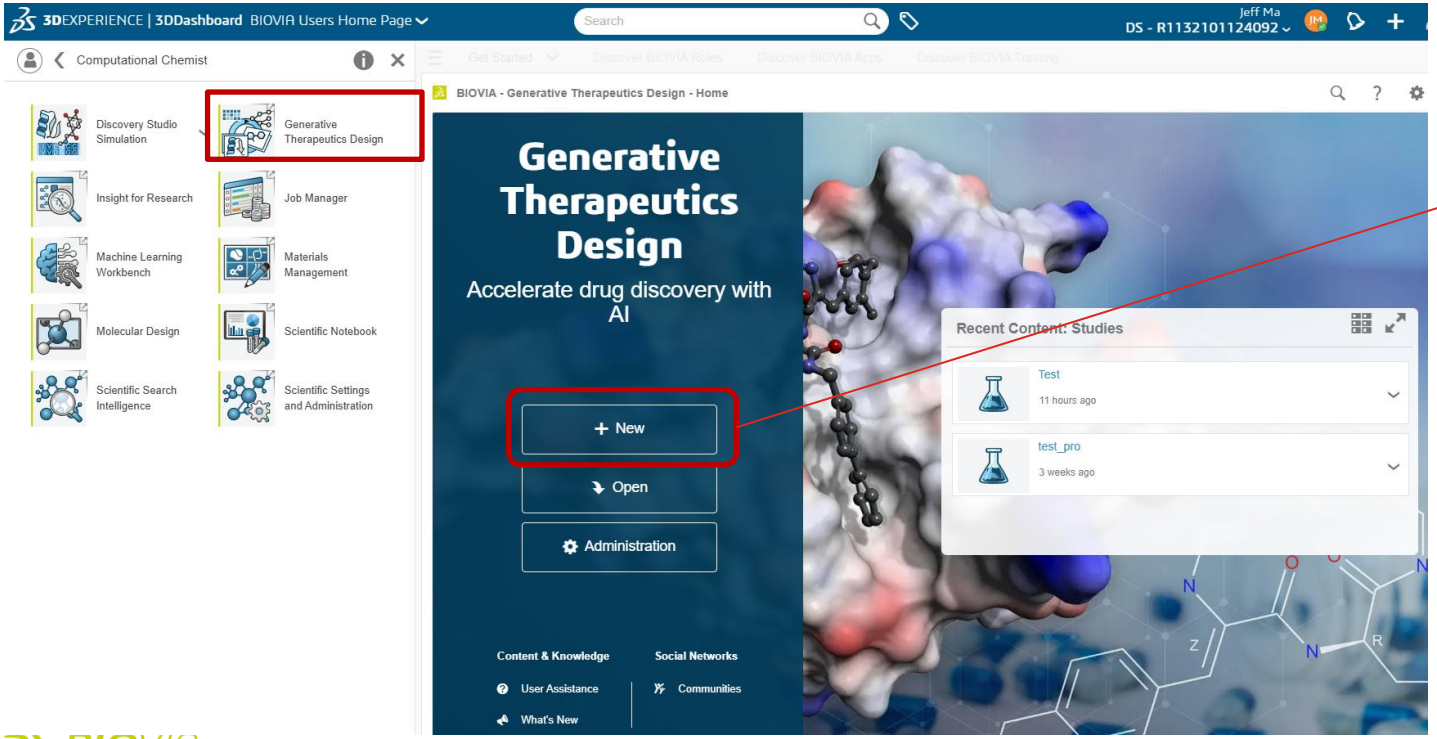
Generative Therapeutics Design (GTD)

Design Workflow



Generative Therapeutics Design (GTD)

GTD combines data science, machine learning (ML), cheminformatics and structure-based modeling to explore chemical space and automate the virtual creation, testing and selection of novel compounds.



Project

↓

Studies

↓

Experiments

BIOVIA

DASSAULT SYSTEMES

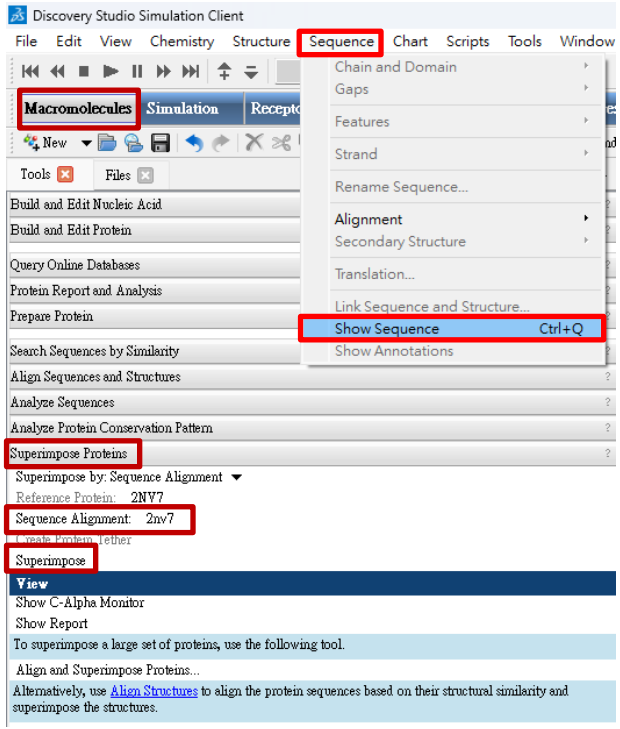
Training Materials

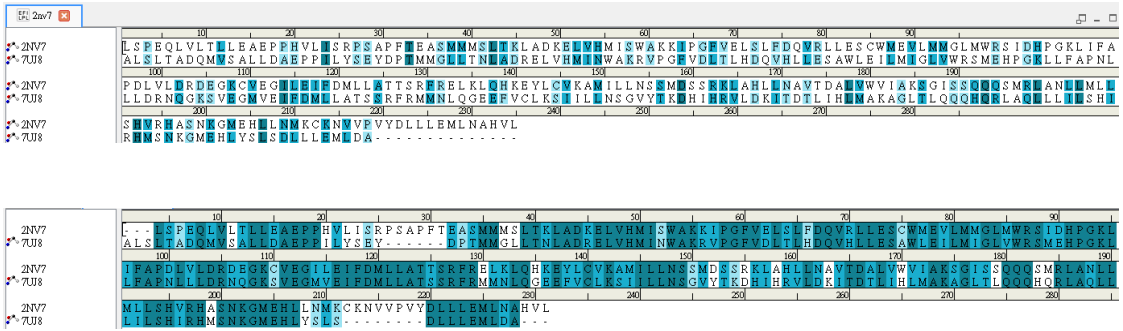
Selective Estrogen Receptor Modulator

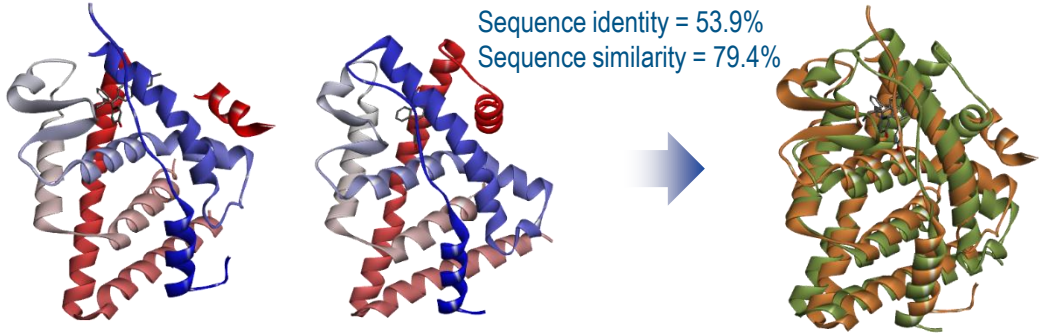
- Two different subtypes of ER have been identified, ERα and Erβ
- ERα and ERβ share approximately 97% of the amino-acid sequence identity in the DNA-binding domain and about 56% in the ligand-binding domain
- The main difference of the ligand-binding domains is determined by Leu 384 and Met-421 in ERα, which are replaced by Met-336 and Ile-373, respectively, in ERβ.

1. Check For Protein Sequence Similarity

Conserved vs Non-Conserved Residues

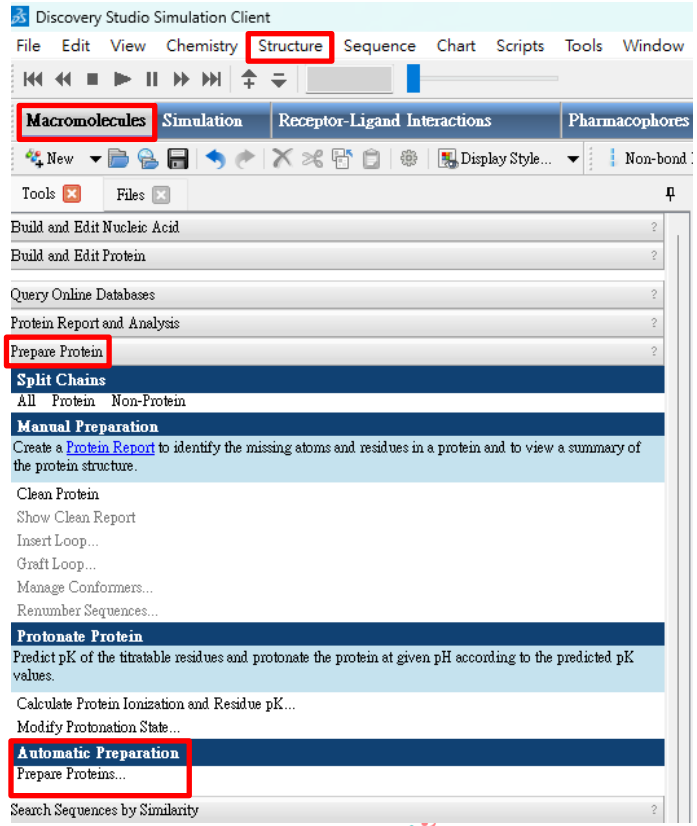


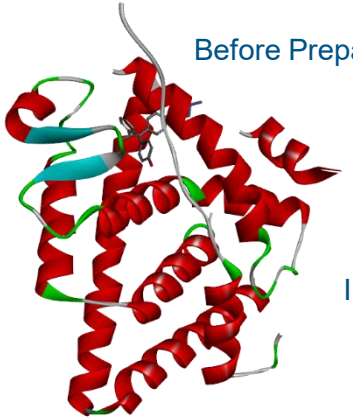




ERalpha: 7UJ8 **ERbeta: 2NV7**

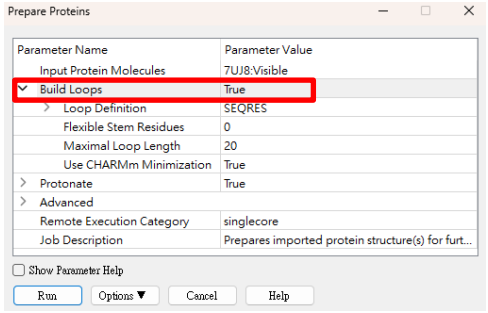
2. Protein Structure Preparation

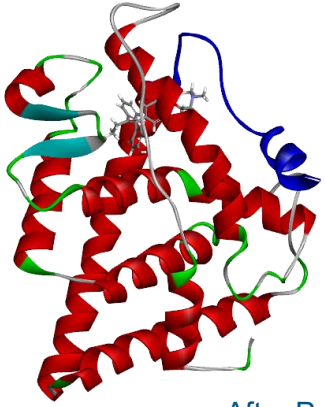




Before Preparation

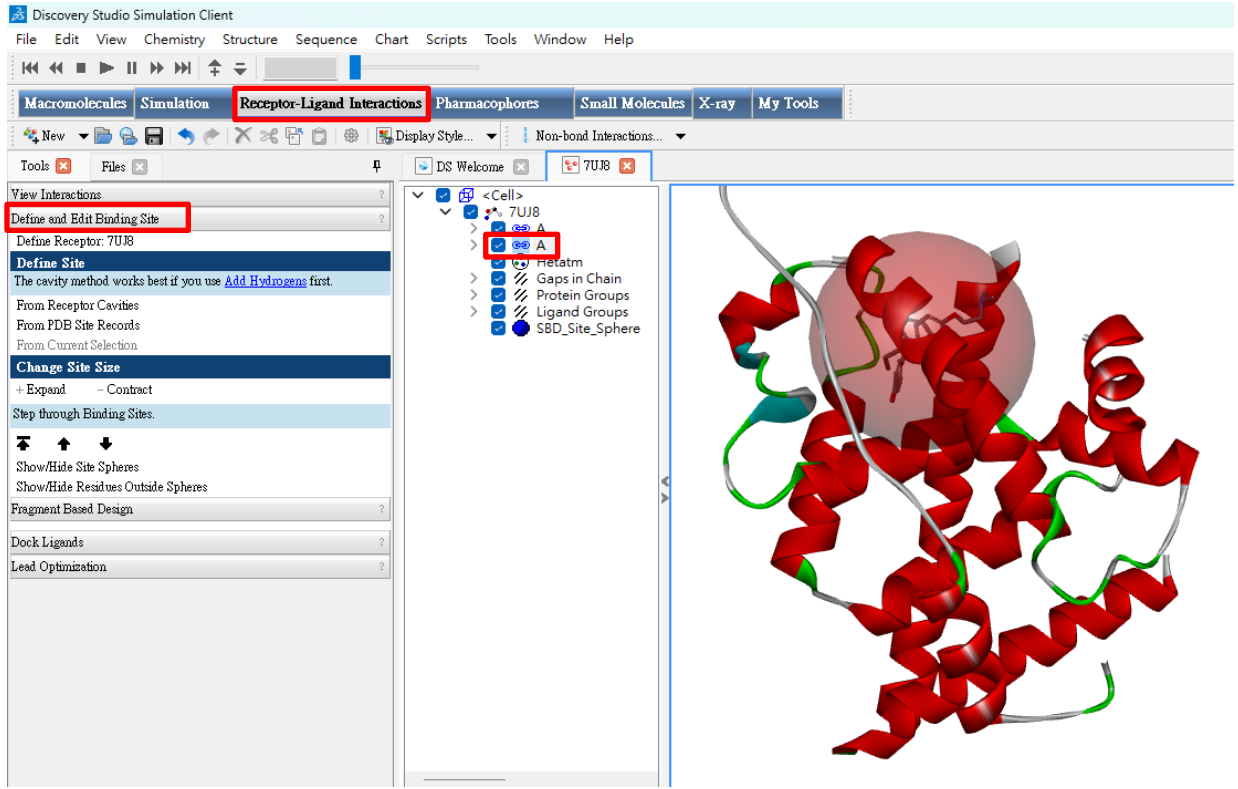
Insert missing loops.

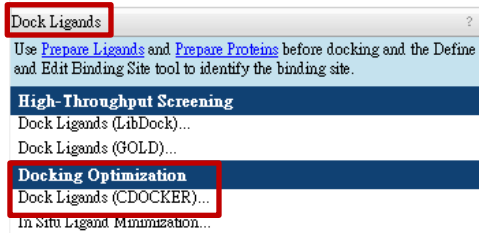




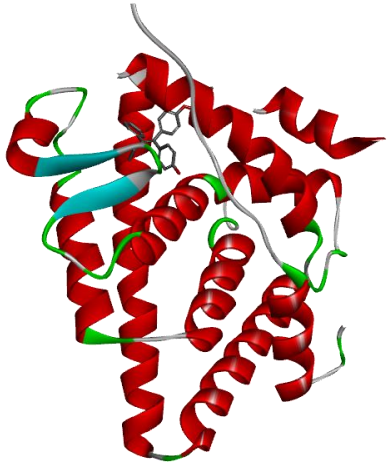
After Preparation

3. Protein-Ligand Molecular Docking

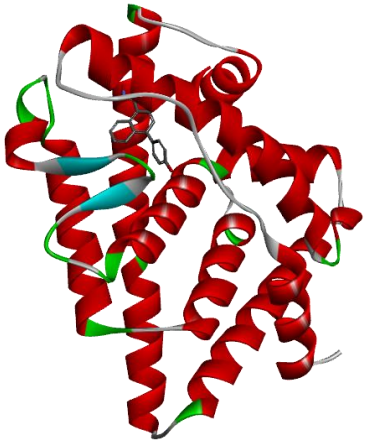




4. Publish 3D Model to GTD (software → web)



ERalpha: 7UJ8

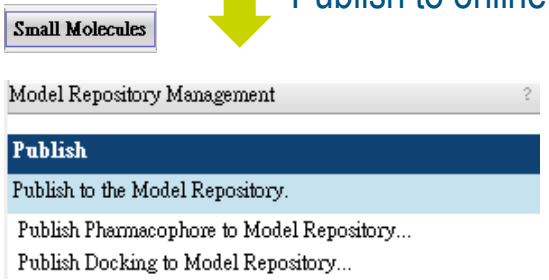


ERbeta: 2NV7

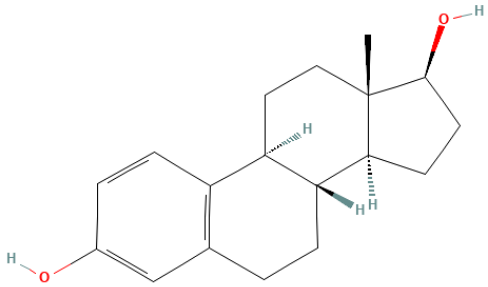
Build Docking (Dock Ligand (CDOCKER)) and Pharmacophore (Interaction Pharmacophore Generate) model for GTD optimization

- Improve sensitivity ⟶ Target optimization
- Improve specificity ⟶Anti-target optimization

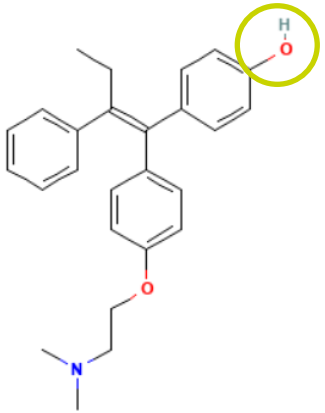
Publish to online Repository



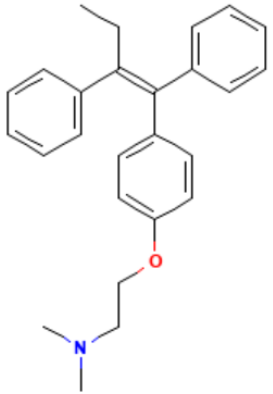
ER inhibitors



17beta-Estradiol



4-Hydroxytamoxifen

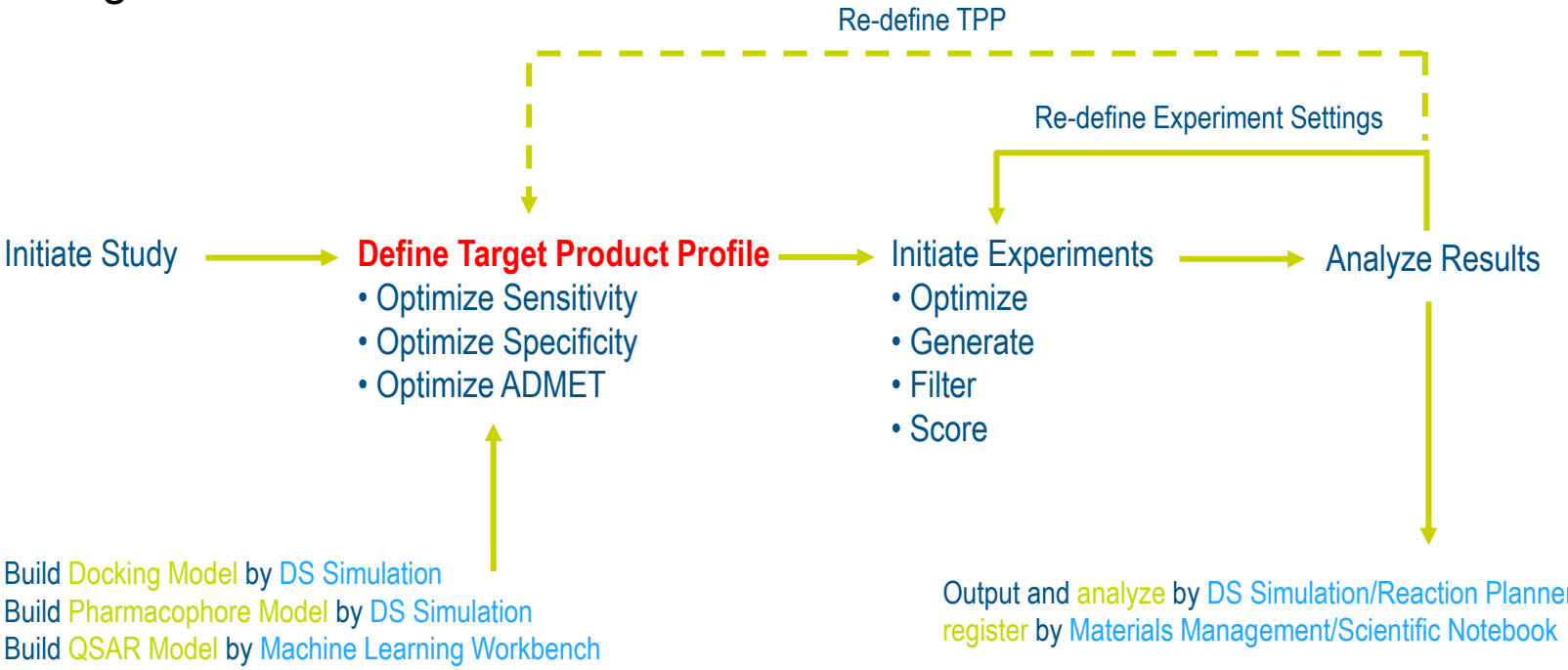


Tamoxifen

- Tamoxifen has become the treatment of choice for women diagnosed with all stages of hormone responsive breast cancer.
- 4-hydroxytamoxifen has more than 100 times higher relative binding affinity than tamoxifen. If its OH group is eliminated or its position is changed the binding affinity is reduced.

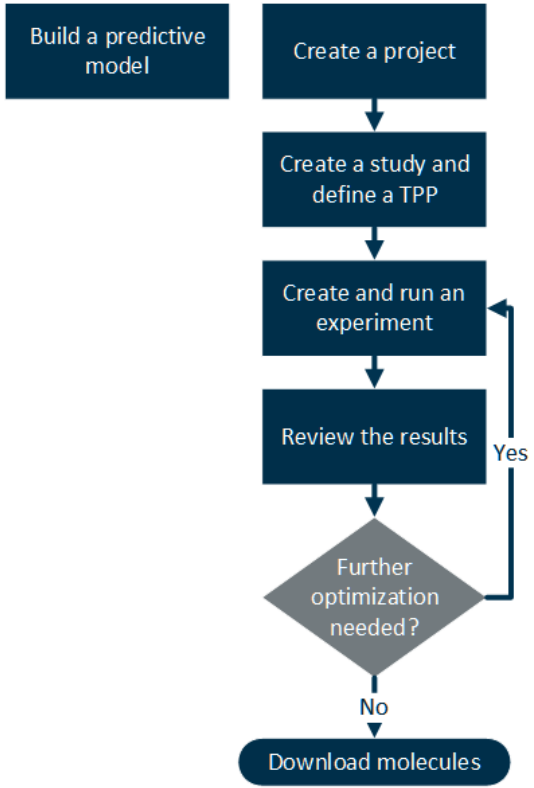
Generative Therapeutics Design (GTD)

Design Workflow



Generative Therapeutics Design (GTD) Workflow

GTD offers four categories of model: Target, Anti-Target, ADME, and Toxicity.
GTD allows you to build your own models based on SAR data. (Machine Learning Workbench)



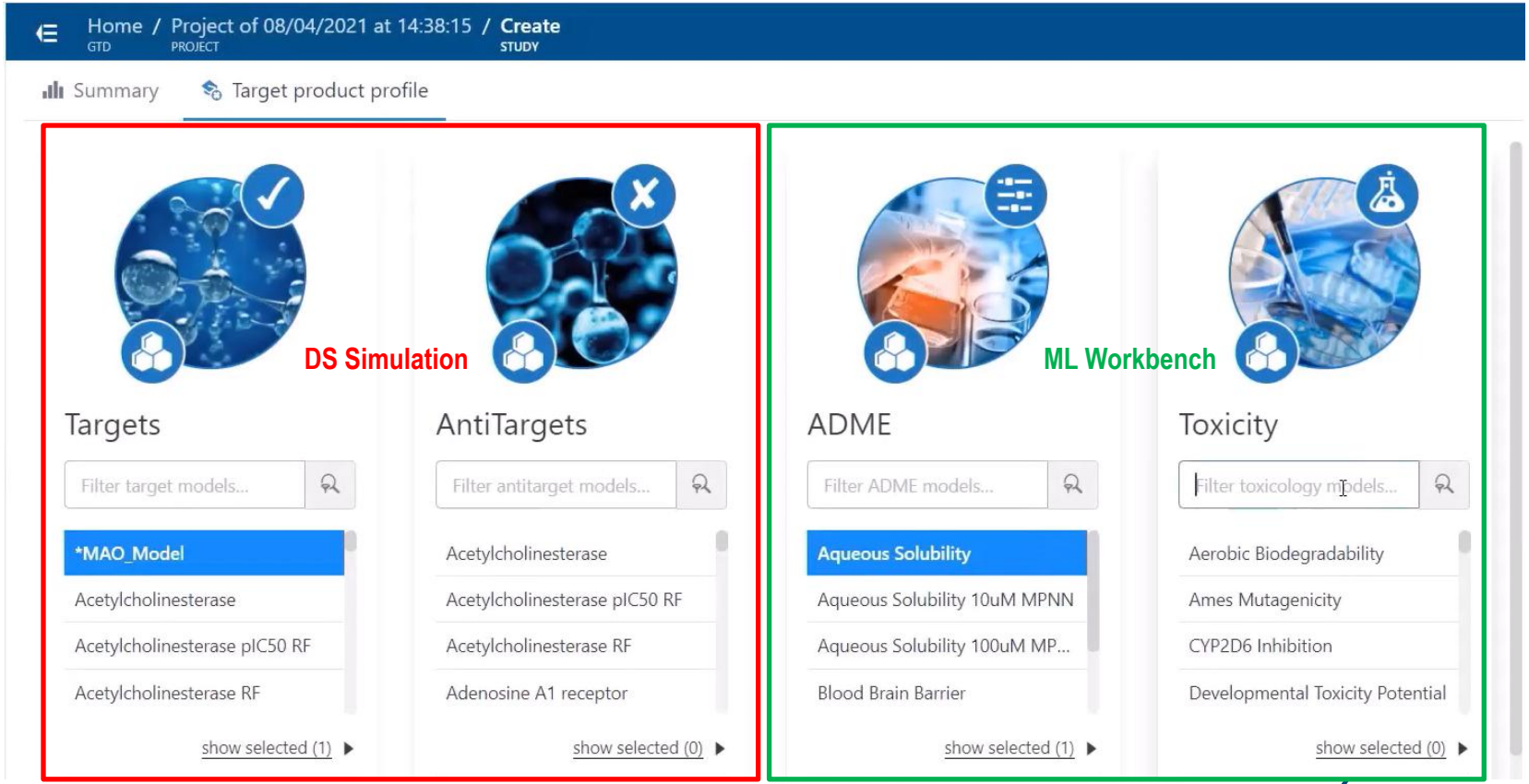
Top-level organizational element.
Each project contains one or more studies.

A study defines the Target Product Profile (TPP) for a group of experiments.

The Target Product Profile (TPP) is a collection of predictive models for:

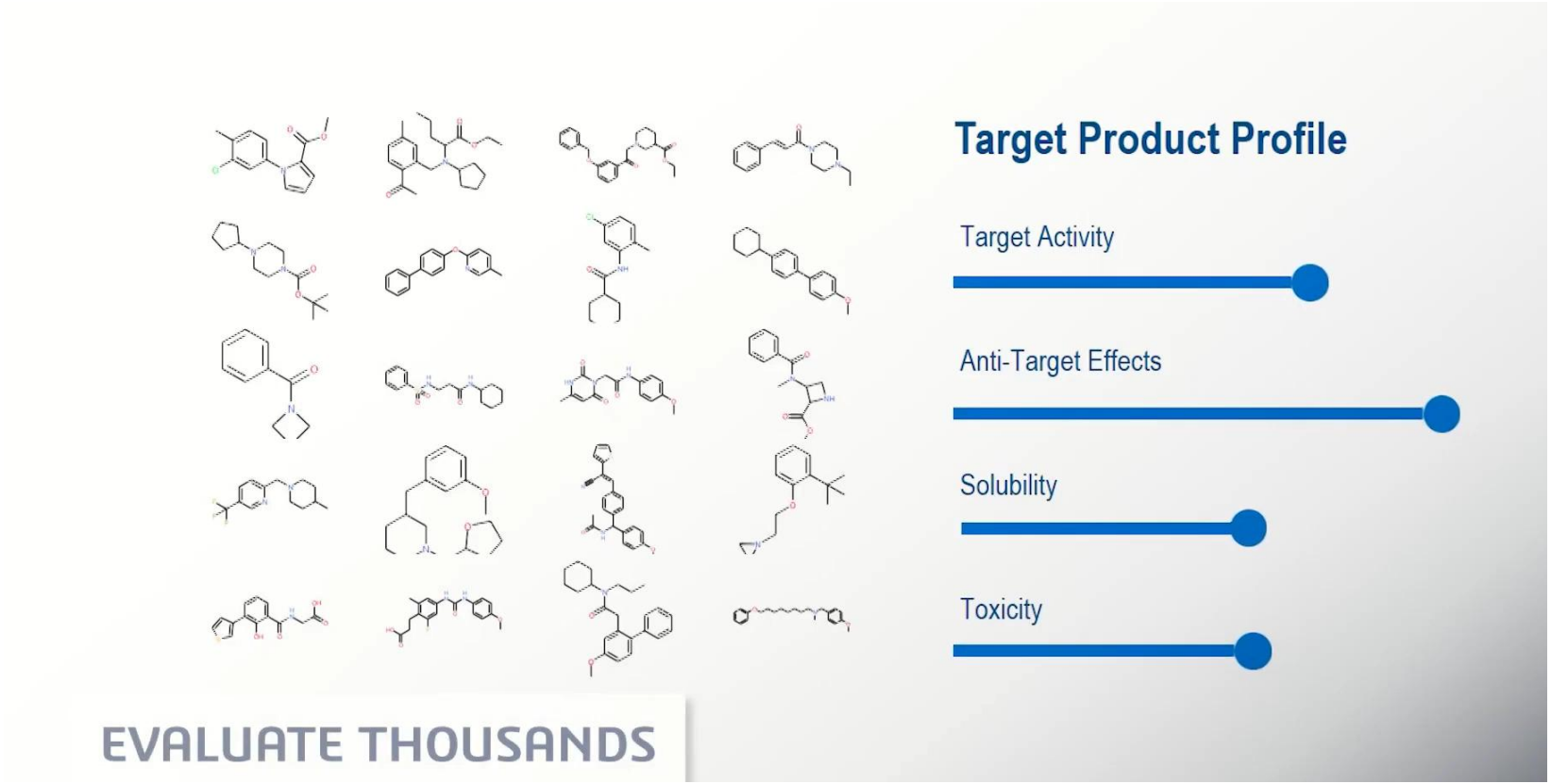
- **Targets**
- **Anti-targets**
- **ADME**
- **Toxicity.**

Define Target Product Profile (TPP) in study



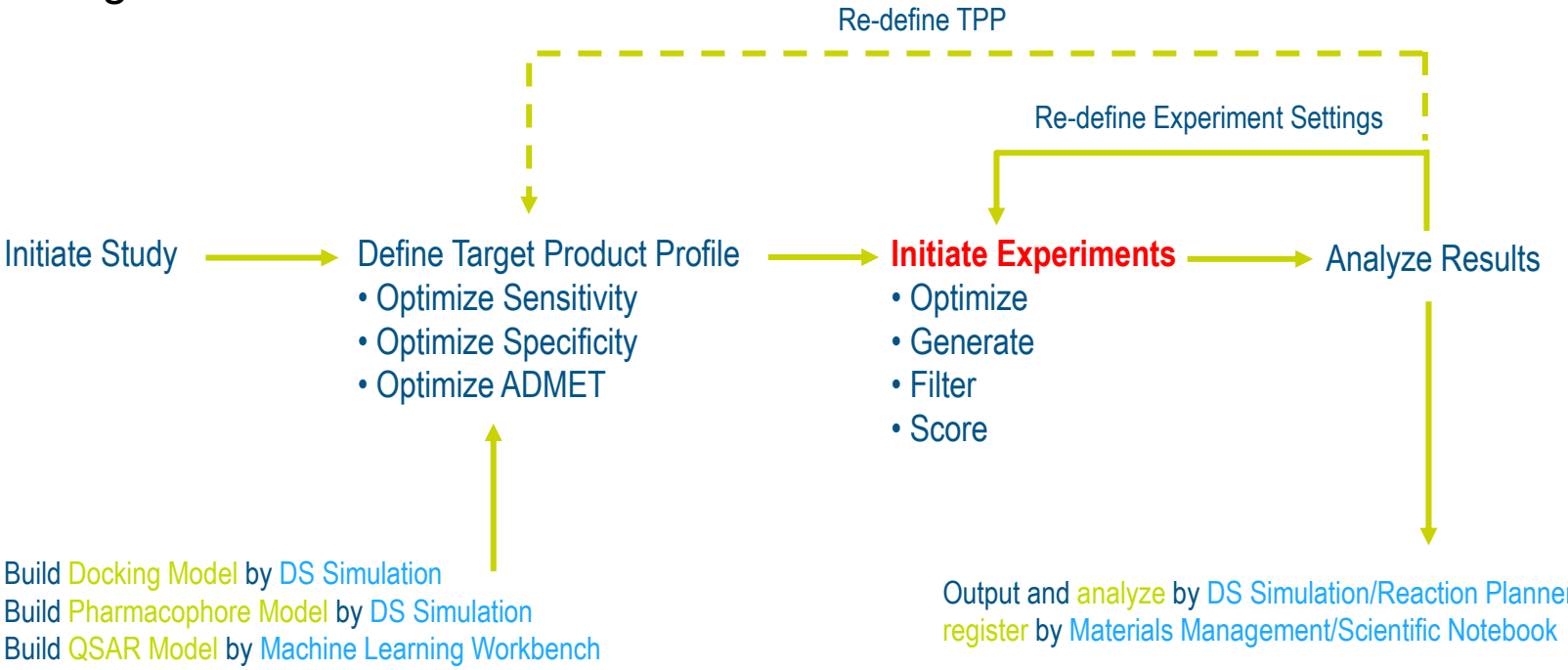
https://www.youtube.com/watch?v=3zP01xdcfO0

Tailor-made compounds



Generative Therapeutics Design (GTD)

Design Workflow



Perform new experiment: Iteration of 4 steps

Optimize:
Perform the multi-objective optimization to achieve the Target Product Profile (TPP).
Find compounds that are predicted to be the best based on model scores.
BIOVIA retrosynthesis score: Commercially available or Synthesis route exists.

Generate:
Generate new molecules from each input by applying different changes, like swapping parts, modifying rings, adjusting atom types, or making other alterations such as splitting, trimming, or rearranging the structure.)

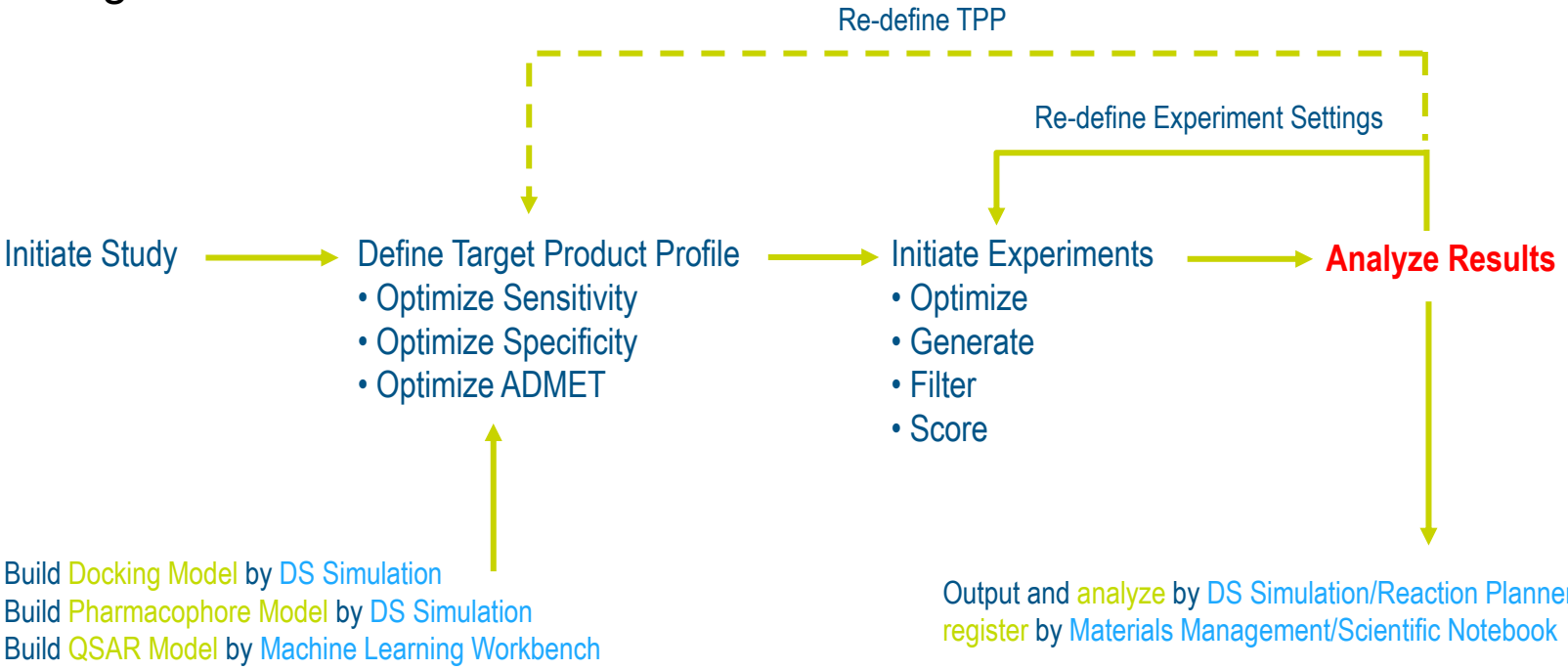
Filter
Filters molecules based on the properties.
Physical Properties: Molecular Weight, AlogP, Polar Surface Area, and Fraction of SP3
Property Counts Filter: Number of rings, rotatable bonds, and others.
Druglike Models Filter: Max Lipinski Violations

Score
Based on 4 models in TPP: Cross-Validation Prediction Distribution, Positive/Negative Category Distribution.

Tailor-made compounds

Generative Therapeutics Design (GTD)

Design Workflow



Analysis GTD Results In All Iterations

Title	Optimization Properties											
	Overall Desirability		Estrogen Receptor Alpha		Aqueous Solubility		Synthetic Accessibility Score		Hepatotoxicity		ERbeta_2NV7_Nove	
Iteration 10	0.988	0.910 ± 0.034 [0.855 ; 0.988]	1.000	0.818 ± 0.116 [0.572 ; 1.000]	1.000	0.913 ± 0.098 [0.564 ; 1.000]	1.000	0.965 ± 0.059 [0.727 ; 1.000]	1.000	0.971 ± 0.053 [0.749 ; 1.000]	0.526	0.437 ± 0.081 [0.000 ; 0.526]
Iteration 9	0.988	0.900 ± 0.035 [0.841 ; 0.988]	1.000	0.809 ± 0.123 [0.518 ; 1.000]	1.000	0.877 ± 0.120 [0.524 ; 1.000]	1.000	0.967 ± 0.054 [0.783 ; 1.000]	1.000	0.976 ± 0.049 [0.730 ; 1.000]	0.572	0.444 ± 0.070 [0.000 ; 0.572]
Iteration 8	0.988	0.887 ± 0.032 [0.840 ; 0.988]	1.000	0.778 ± 0.124 [0.508 ; 1.000]	1.000	0.879 ± 0.119 [0.563 ; 1.000]	1.000	0.961 ± 0.058 [0.759 ; 1.000]	1.000	0.966 ± 0.062 [0.716 ; 1.000]	0.545	0.441 ± 0.065 [0.000 ; 0.545]
Iteration 7	0.979	0.864 ± 0.036 [0.810 ; 0.979]	1.000	0.758 ± 0.133 [0.483 ; 1.000]	1.000	0.834 ± 0.127 [0.517 ; 1.000]	1.000	0.960 ± 0.066 [0.707 ; 1.000]	1.000	0.947 ± 0.070 [0.723 ; 1.000]	0.526	0.433 ± 0.064 [0.000 ; 0.526]
Iteration 6	0.953	0.819 ± 0.044 [0.759 ; 0.953]	1.000	0.702 ± 0.142 [0.452 ; 1.000]	1.000	0.799 ± 0.141 [0.418 ; 1.000]	1.000	0.943 ± 0.095 [0.569 ; 1.000]	1.000	0.891 ± 0.119 [0.443 ; 1.000]	0.560	0.445 ± 0.057 [0.219 ; 0.560]
Iteration 5	0.893	0.735 ± 0.051 [0.661 ; 0.893]	1.000	0.615 ± 0.185 [0.224 ; 1.000]	1.000	0.741 ± 0.163 [0.333 ; 1.000]	1.000	0.910 ± 0.125 [0.461 ; 1.000]	1.000	0.778 ± 0.158 [0.349 ; 1.000]	0.554	0.432 ± 0.072 [0.000 ; 0.554]
Iteration 4	0.818	0.558 ± 0.085 [0.439 ; 0.818]	1.000	0.510 ± 0.161 [0.177 ; 1.000]	1.000	0.573 ± 0.184 [0.118 ; 1.000]	1.000	0.751 ± 0.197 [0.146 ; 1.000]	1.000	0.529 ± 0.189 [0.223 ; 1.000]	0.522	0.422 ± 0.046 [0.225 ; 0.522]
Iteration 3	0.708	0.275 ± 0.154 [0.055 ; 0.708]	0.919	0.410 ± 0.200 [0.010 ; 0.919]	0.960	0.405 ± 0.250 [0.010 ; 0.960]	1.000	0.393 ± 0.351 [0.010 ; 1.000]	1.000	0.327 ± 0.228 [0.010 ; 1.000]	0.524	0.413 ± 0.083 [0.000 ; 0.524]
Iteration 2	0.342	0.146 ± 0.080 [0.024 ; 0.342]	0.848	0.423 ± 0.239 [0.010 ; 0.848]	0.743	0.230 ± 0.258 [0.010 ; 0.743]	1.000	0.404 ± 0.369 [0.010 ; 1.000]	0.391	0.126 ± 0.125 [0.010 ; 0.391]	0.483	0.367 ± 0.107 [0.000 ; 0.483]
Iteration 1	0.088	0.062 ± 0.026 [0.025 ; 0.088]	0.981	0.729 ± 0.181 [0.453 ; 0.981]	0.010	0.010 ± 0.000 [0.010 ; 0.010]	0.724	0.358 ± 0.300 [0.010 ; 0.724]	0.010	0.010 ± 0.000 [0.010 ; 0.010]	0.421	0.130 ± 0.159 [0.000 ; 0.421]
Input	0.088	0.088 ± 0.000 [0.088 ; 0.088]	1.000	1.000 ± 0.000 [1.000 ; 1.000]	0.010	0.010 ± 0.000 [0.010 ; 0.010]	0.623	0.623 ± 0.000 [0.623 ; 0.623]	0.010	0.010 ± 0.000 [0.010 ; 0.010]	0.207	0.207 ± 0.000 [0.207 ; 0.207]

Analysis | Compare iterations | Save | Run

Plot properties and Download molecules

Summary | Molecules 1 | Parameters | Convergence | Iterations 10 | Comments 1 | Analysis

Iterations

10 | 9 | 8 | 7 | 6 | 5

4 | 3 | 2 | 1 | 0

Properties

Y: Overall Desirability

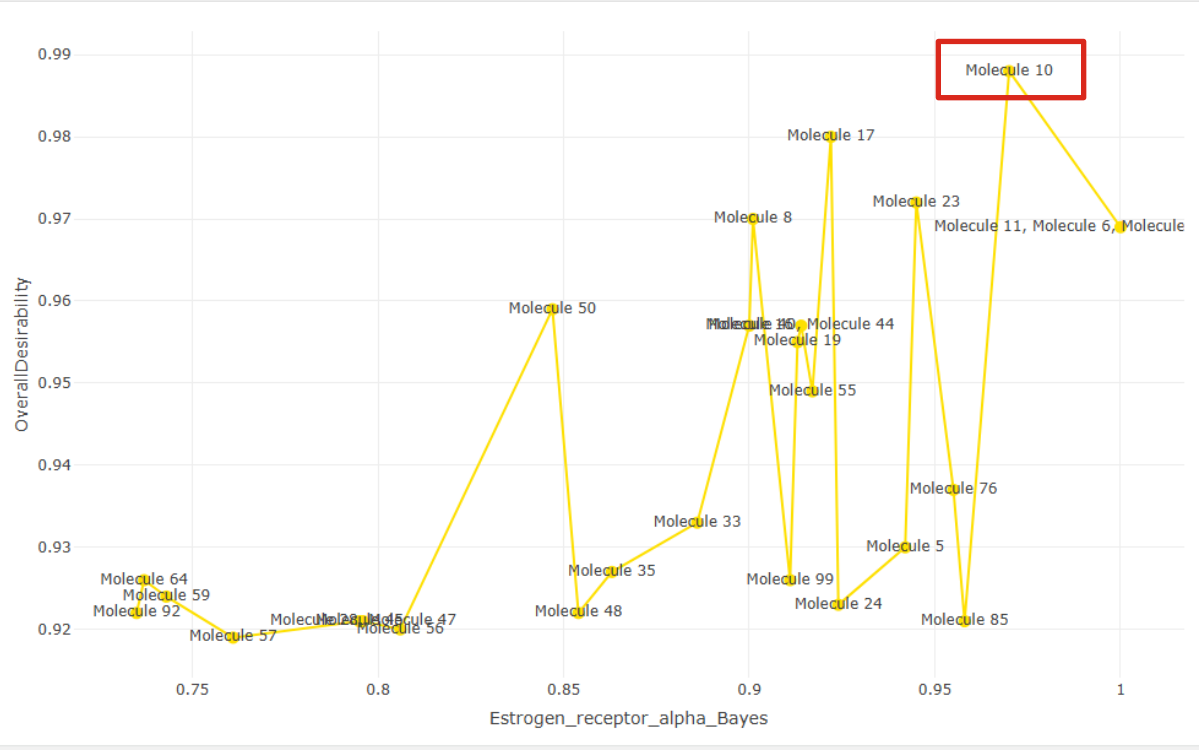
X: Estrogen receptor alpha

Aqueous Solubility

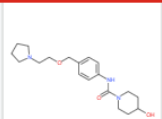
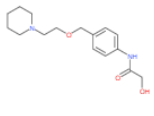
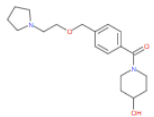
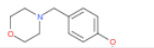
Synthetic Accessibility Score

Hepatotoxicity

ERbeta_2NV7_Nove



Plot properties and Download molecules

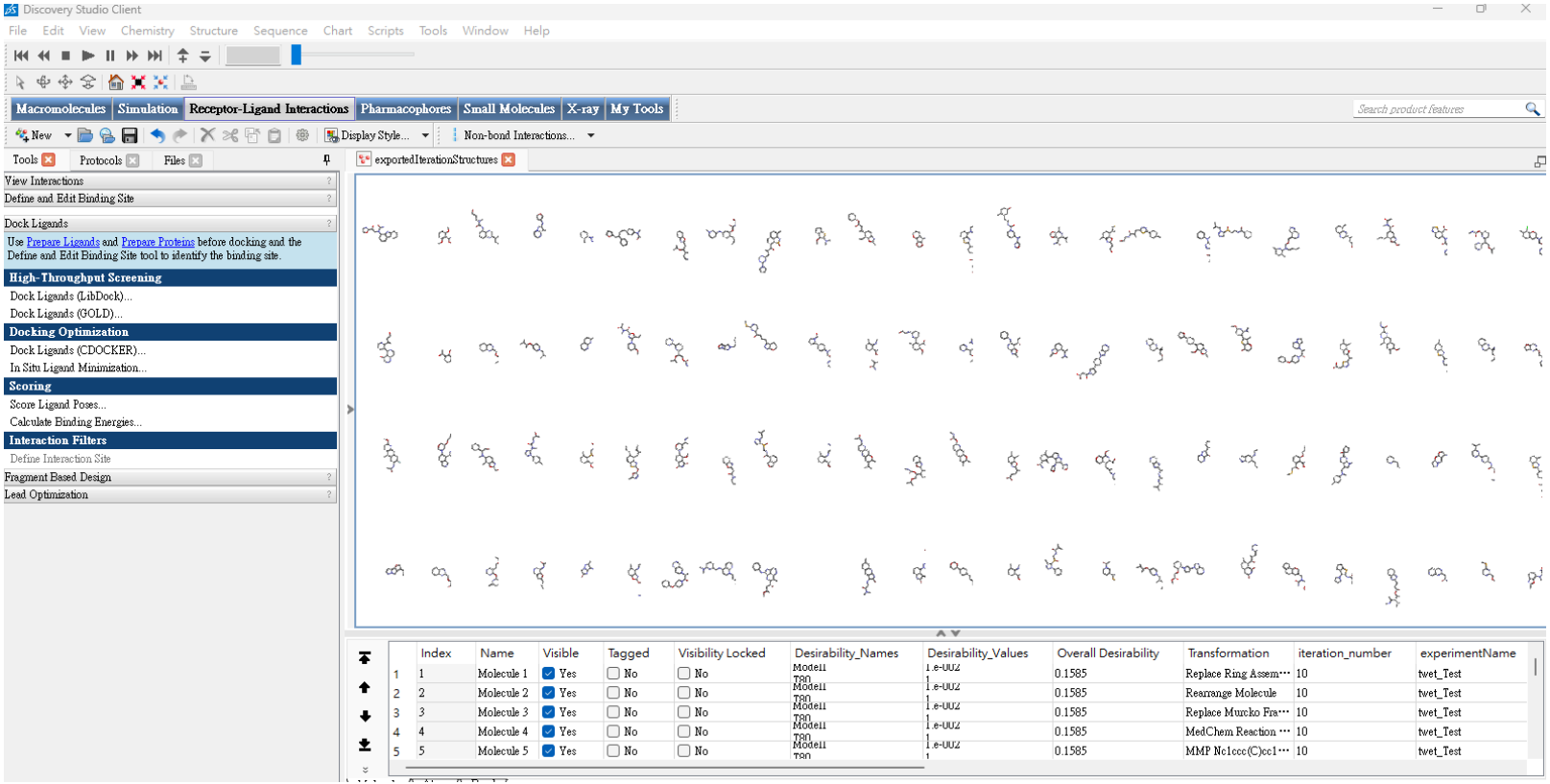
Title	Optimization Properties							Generator And Filte	
	Structure	Overall Desirability	Estrogen Receptor Alpha	Aqueous Solubility	Synthetic Accessibility Score	Hepatotoxicity	ERbeta_2NV7_Nove	ALogP	M
0 0 Molecule 10		0.988	0.970	0.982	1.000	1.000	0.450	0.988	65.040
0 0 Molecule 17		0.980	0.922	1.000	1.000	1.000	0.455	1.138	61.800
0 0 Molecule 23		0.972	0.945	0.944	1.000	1.000	0.482	1.080	53.010
0 0									

Showing 1 to 30 of 100 rows. (Show 30 more) (Show all)

‹ 1 2 3 4 ›

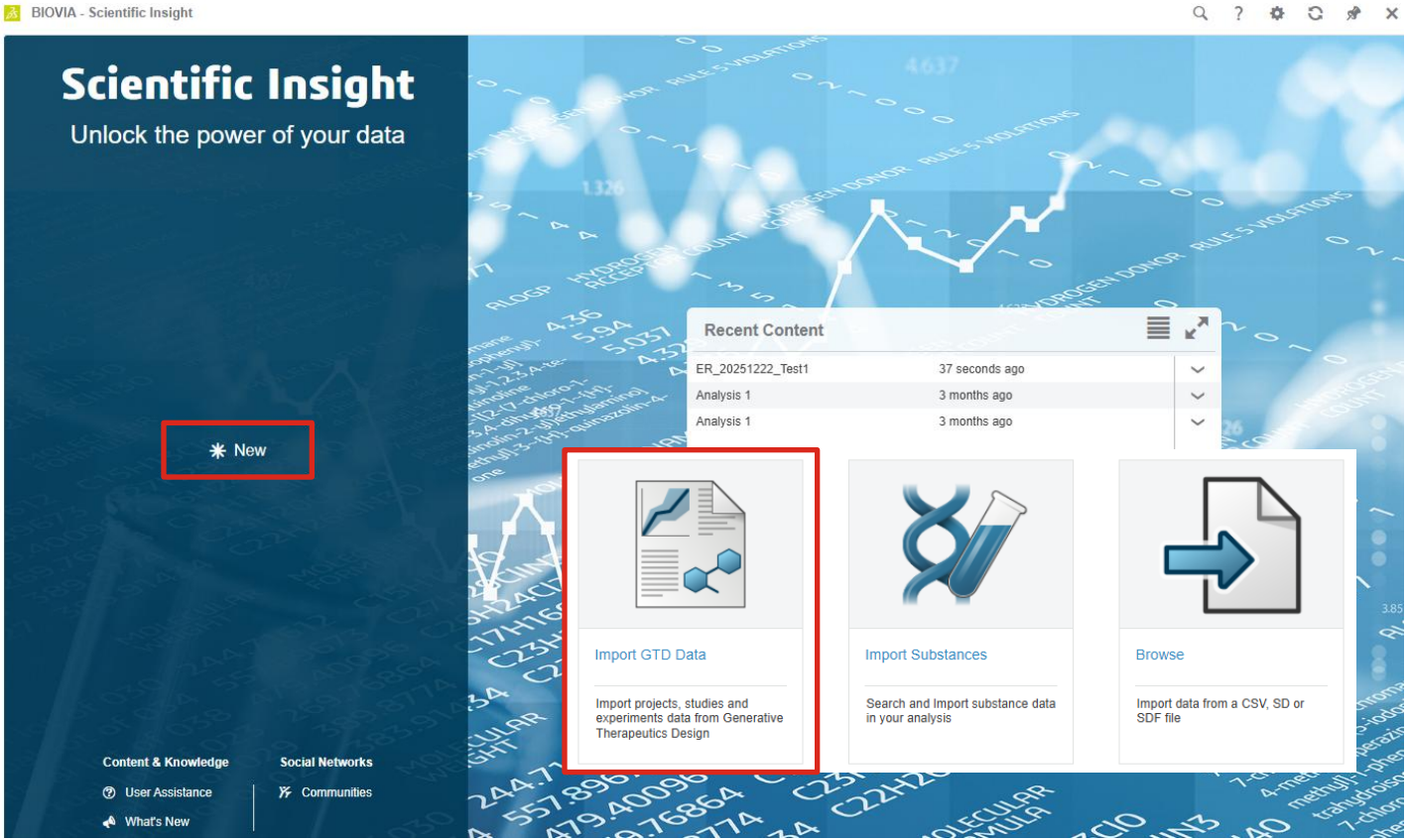
Analysis | Molecules review | Download molecules | Create an experiment

Use GTD Result as a Drug Library in DS

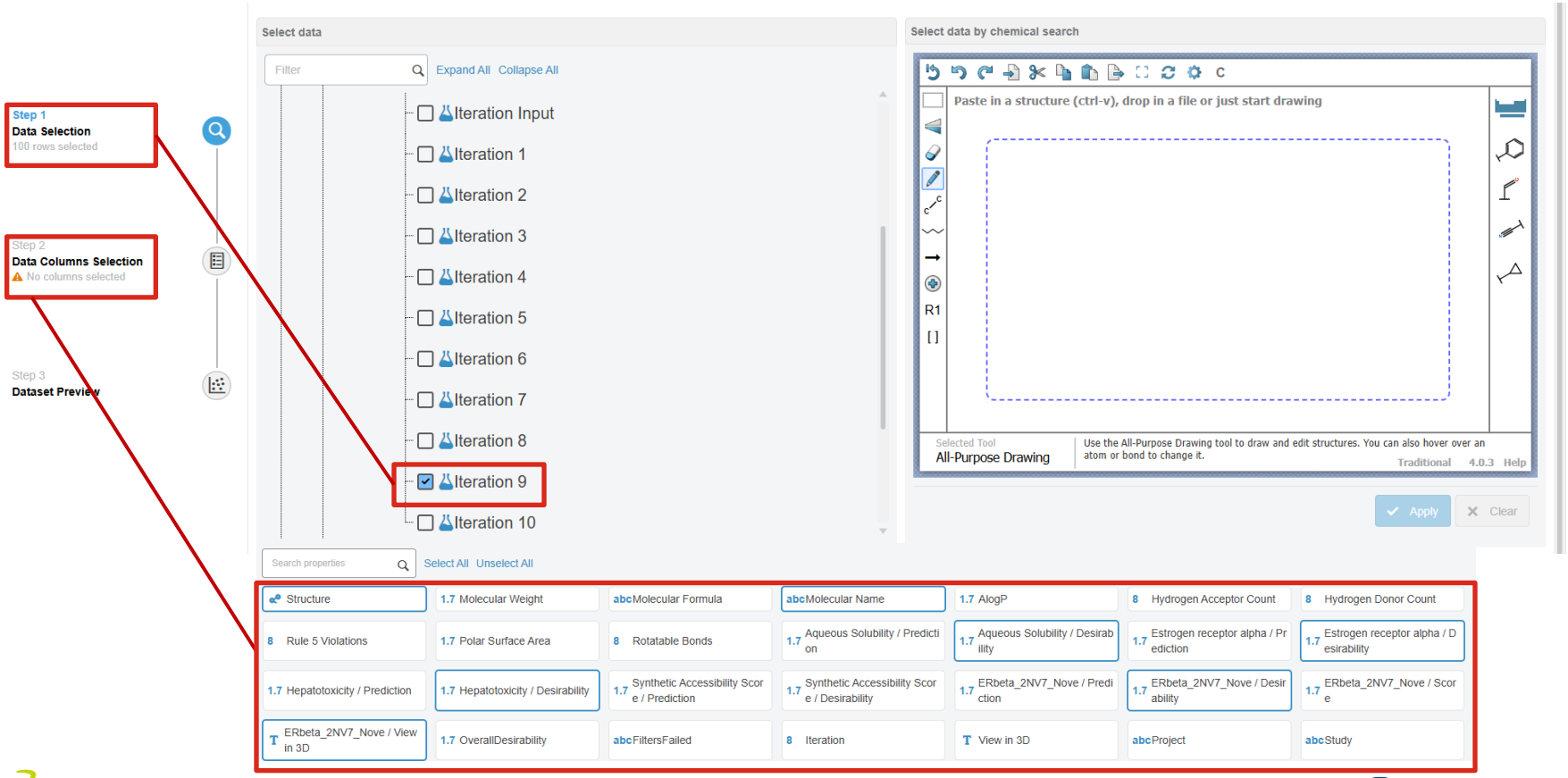


Scientific Insight - Visualize and interact with data

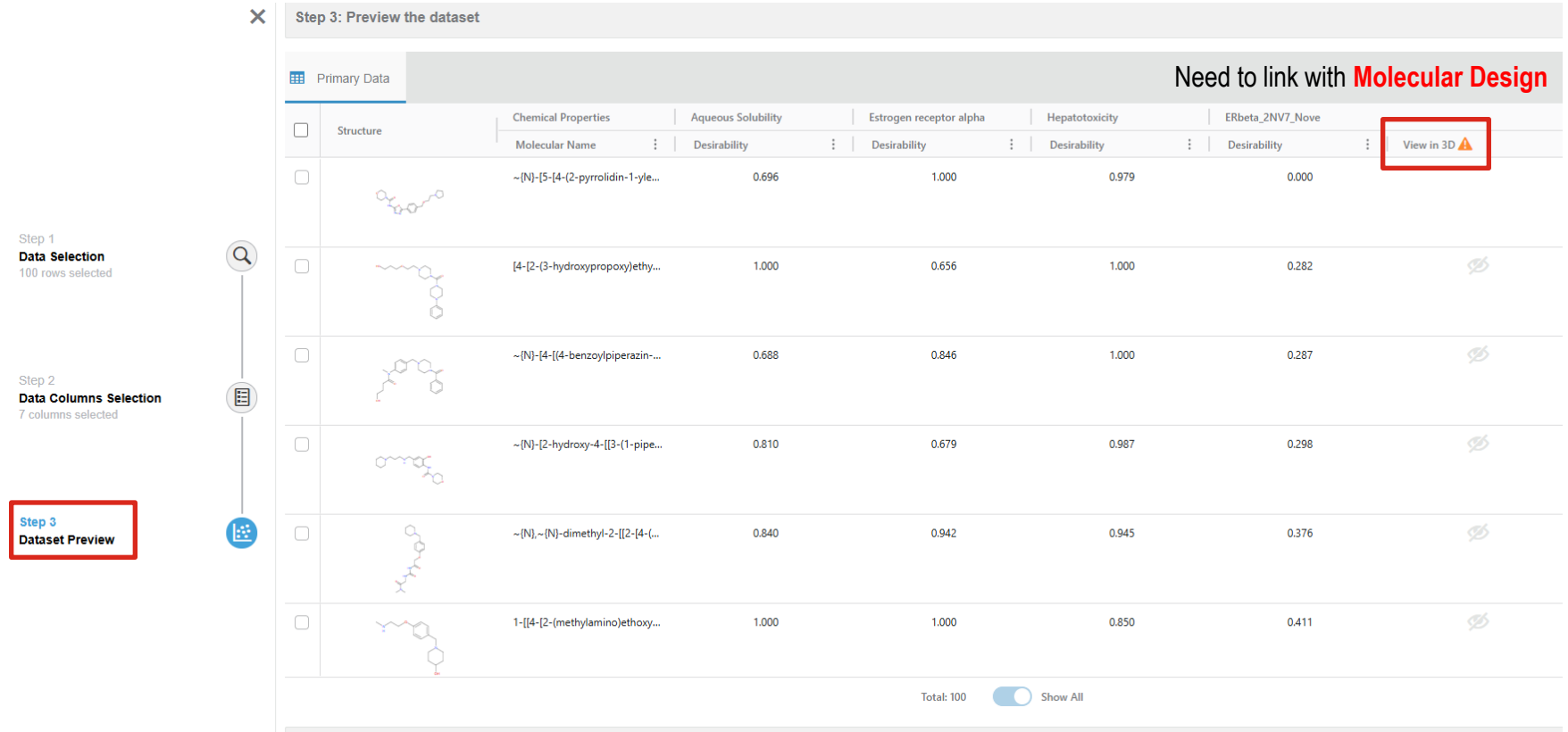
Scientific Insight



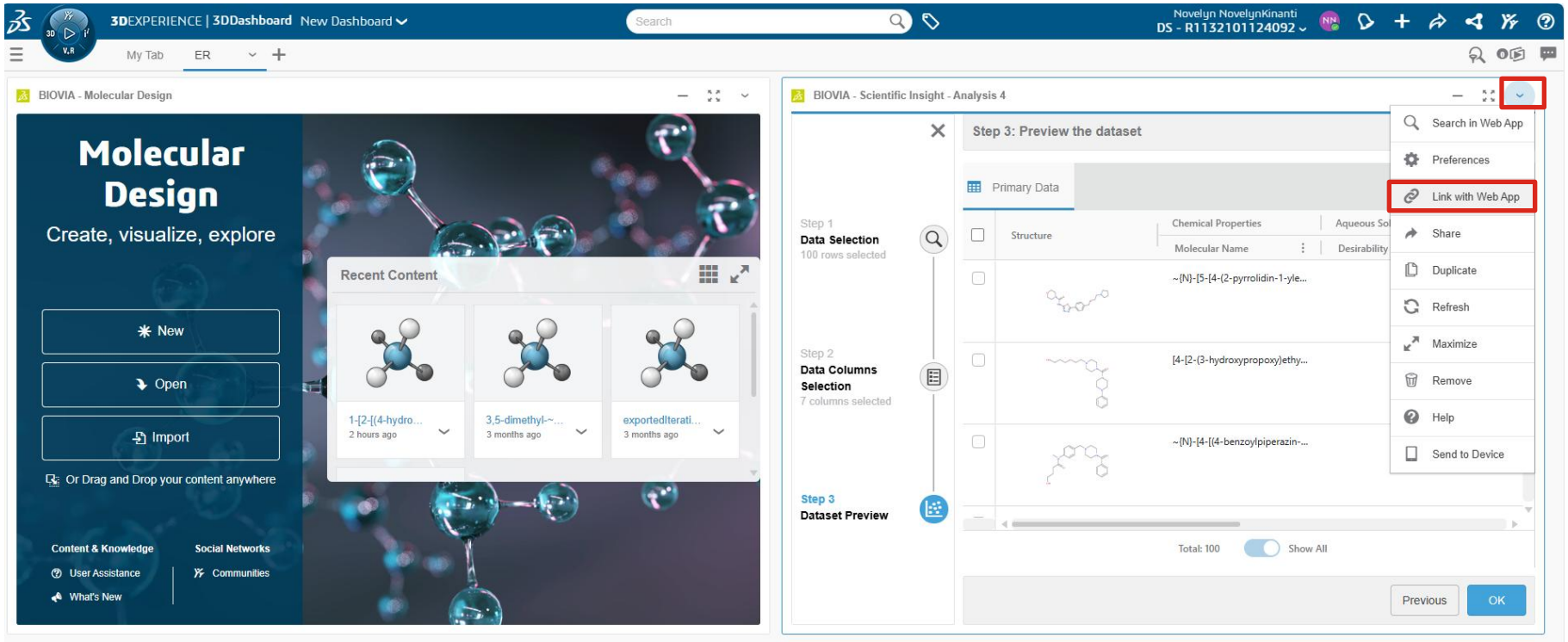
Scientific Insight - Visualize and interact with data



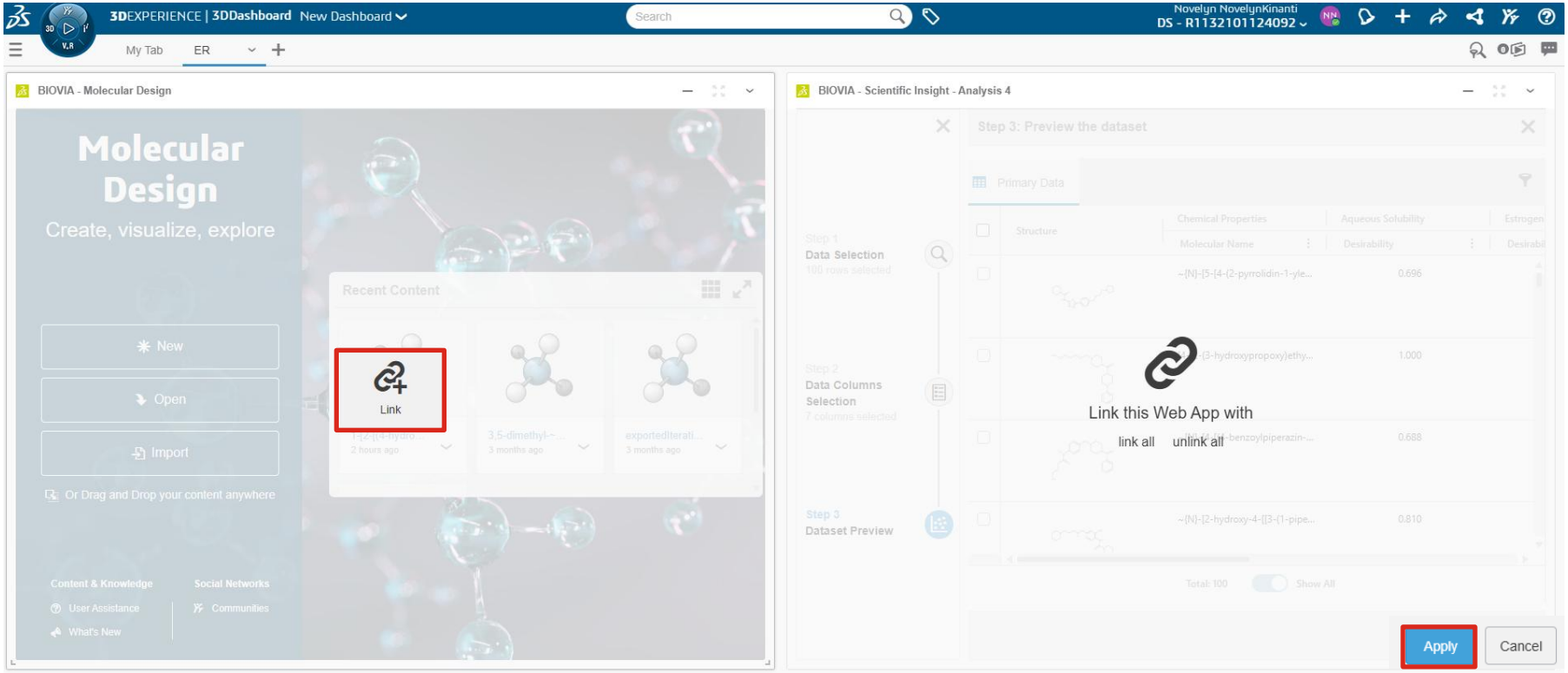
Scientific Insight - Visualize and interact with data



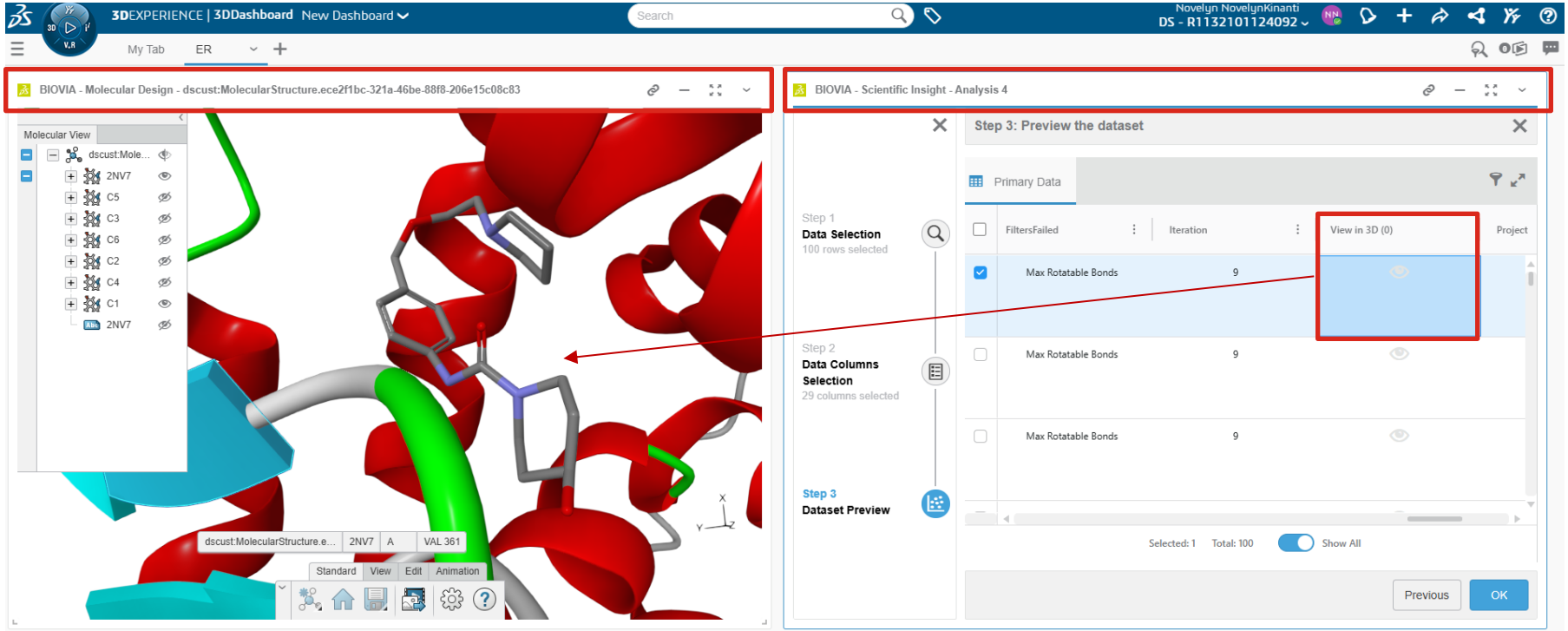
Link Scientific Insight with Molecular Design



Link Scientific Insight with Molecular Design



Link Scientific Insight with Molecular Design

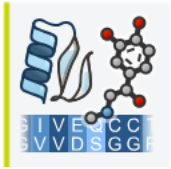


Machine Learning Workbench

AI in Drug Discovery

Comprehensive Applications in Small and Biological Drug Design

3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

Use AI strategies to design protein binders and predict structures with AlphaFold.

Data visualization and Analysis

Insight for Research

Visualize your data and connect it to physics-based models.

AI in Small Molecule Design

Machine Learning Workbench

Train ML models with your experimental data.

Generative Therapeutics Design

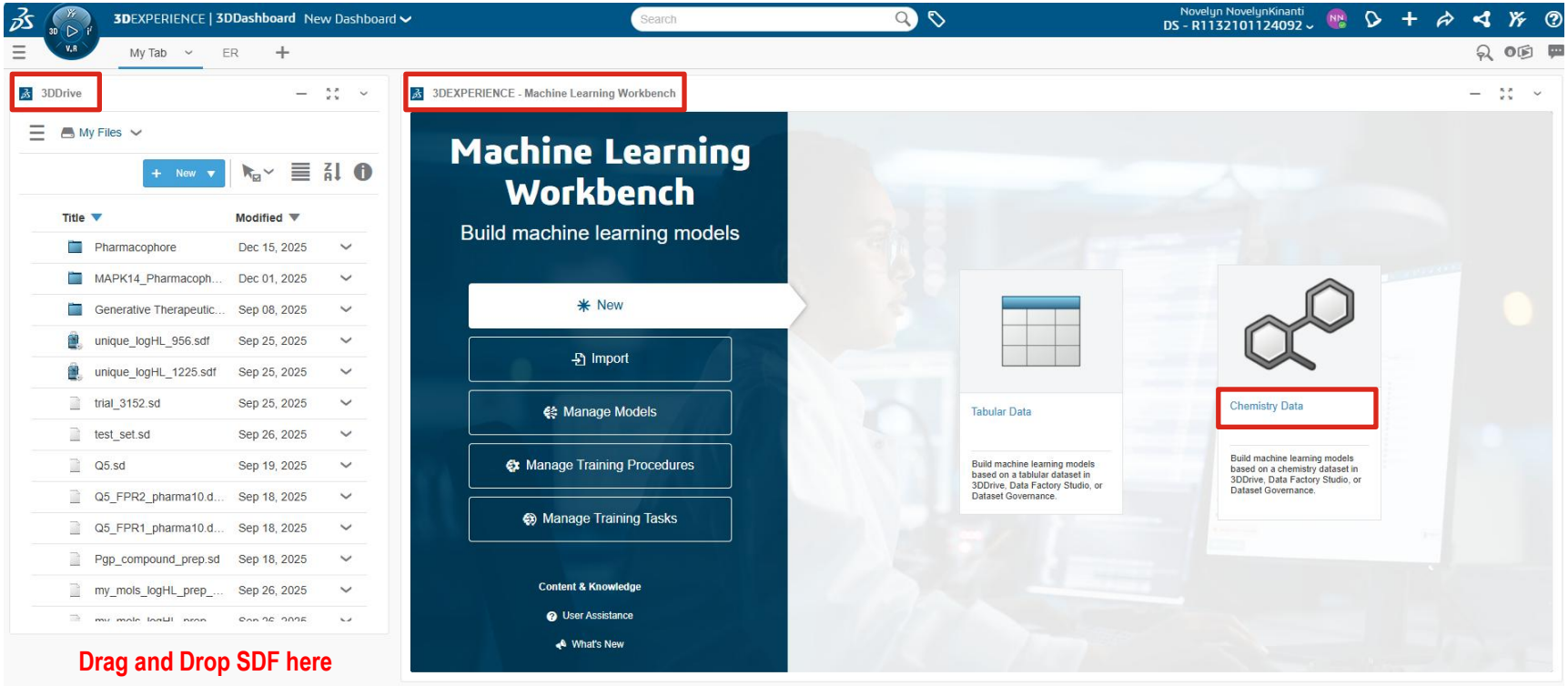
Automate the virtual creation, testing and selection of drug-like compounds using ML models and scientific methods.

Reaction Planner

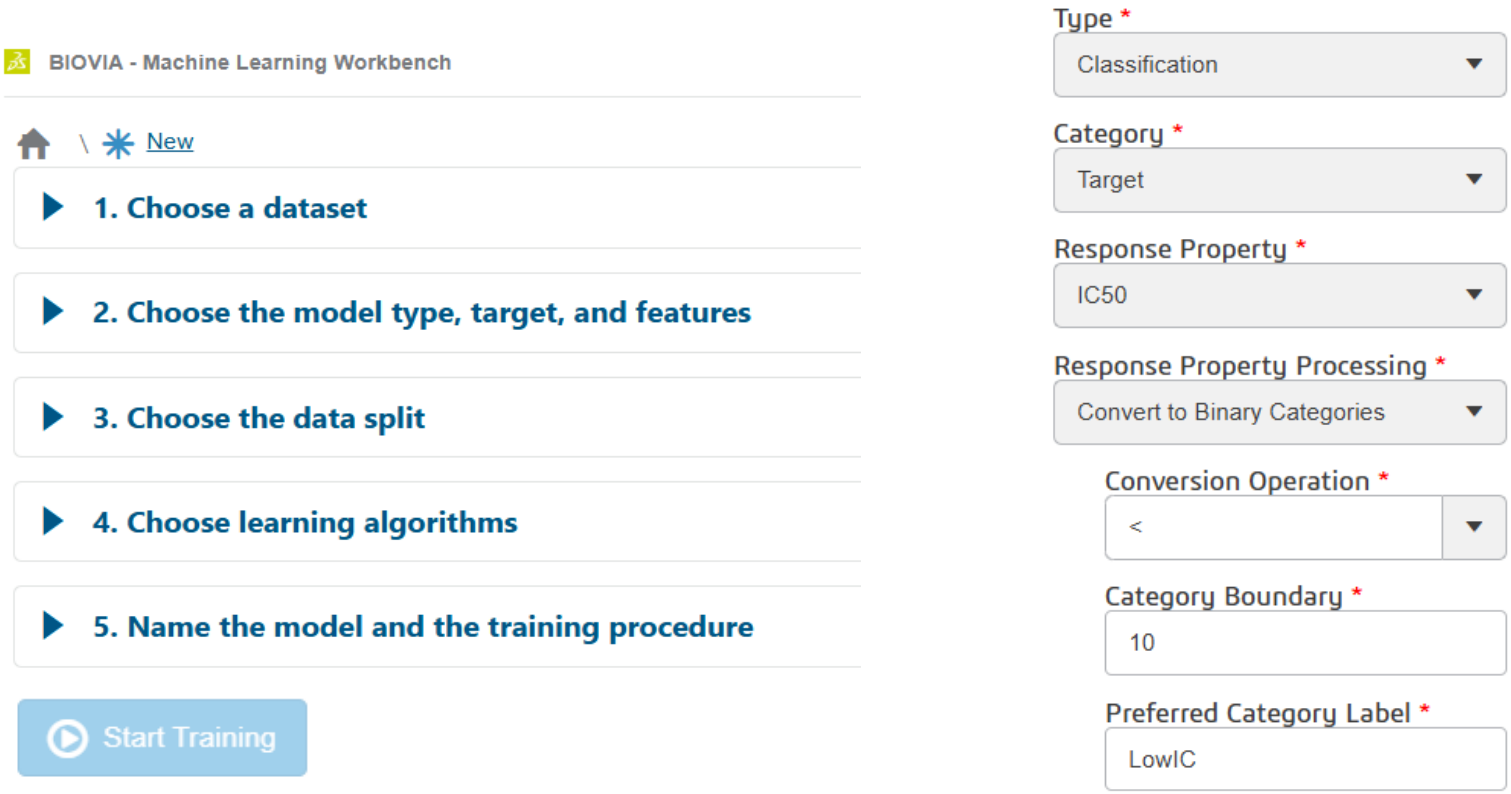
AI-based retrosynthesis prediction tool

Import Datasets

In Structure Data Format (.sd / .sdf)



Machine Learning Workbench - User-defined ML model



Half-Life Pharmacokinetics: AI Model Prediction

MACHINE LEARNING AND DEEP LEARNING | April 2, 2024

Predicting Elimination of Small-Molecule Drug Half-Life in Pharmacokinetics Using Ensemble and Consensus Machine Learning Methods

Jianing Fan, Shaohua Shi, Hong Xiang, Li Fu, Yanjing Duan, Dongsheng Cao*, and Hongwei Lu*

Access Through Your Institution | Other Access Options | Supporting Information (6)

	t12	smiles	group
1	0.160000006	Clc1cc(/C=C/C(=O)N	train
2	12.99998945	O=C([O-])[C@@H](O	train
3	0.166701689	S(=O)([O-])(=Nc1cc(	train
4	0.200000002	Clc1c(-c2ccc(NC(=O)	train
5	5.999996541	O=C(OC[C@H]([N+](	train
6	4.017001556	FC(F)(F)Oc1cc(F)c(-c	train
7	0.383301042	Clc1c(N(C(=O)NCC(=	train
8	0.133299954	O=C(NCCCc1ccccc1	train
9	2.582997584	S(=O)([O-])(=Nc1cc(	train
10	0.766708009	O=C(NCCCc1ccccc1	train
11	0.18329896	O=C(Nc1ncc(C(C)(C)	train
12	3.300000305	O(C(C)C)c1c2c(nc(-c	train
13	5.899999842	O=C1N(C)[C@H](c2c	train
14	0.800000024	O=C1Oc2c(cccc2)C=(	train
15	0.633300747	O=C(N[C@H](C(=O)	train
16	0.696706721	O=C(N1CCC2(C(C(=	train
17	0.833297349	S(=O)([O-])(=Nc1cc2	train
18	0.266698148	S(=O)(=O)(NC(=O)[C	train
19	16.00000064	O=[N+](OC(CO[N+](	train
20	3.84699835	Fc1c(Oc2ncnc3c2cc(C	train
21	0.133299957	O=C(NC(C)C12CC3C	train
22	1.298999548	O=C(N/N=C(\C)/c1nc	train
23	6.200000435	[N+H3]C1C(c2ccccc2	train
24	11.00000797	S(CC[C@H](NC(=O)[	train
25	0.599998273	P(=O)(OP(=O)(OC[C	train
26	0.800000024	[S+](CCOC(=O)N)=C	train
27	7.800007136	FC(F)(F)c1cc(CN(C(=	train
28	2.160001238	Clc1c(CO)n(Cc2ccc(-c	train
29	6.999999677	O=C1N(C)C(=O)NC1	train

SMILES to SDF

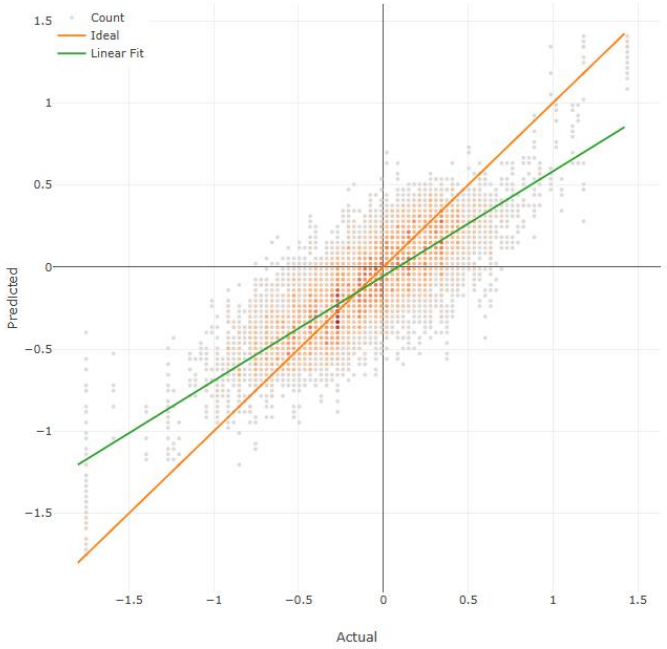
XGBoost / LightGBM Regression Model
Based on Features and Fingerprints

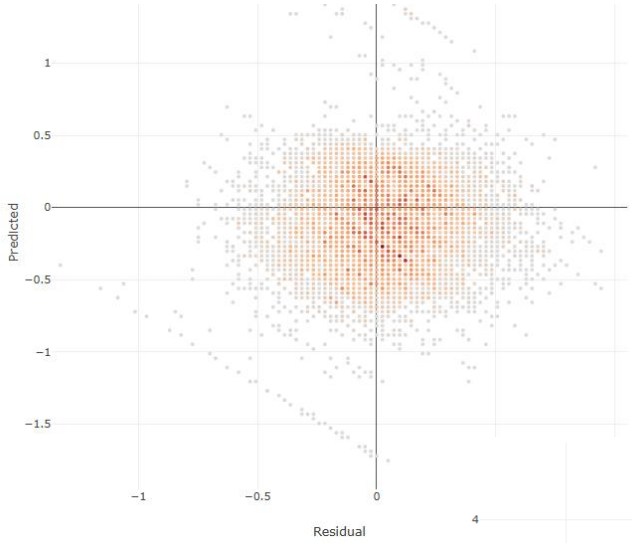
3512 ->2557 Molecules

Model Report for logHL_mol2557 version 1

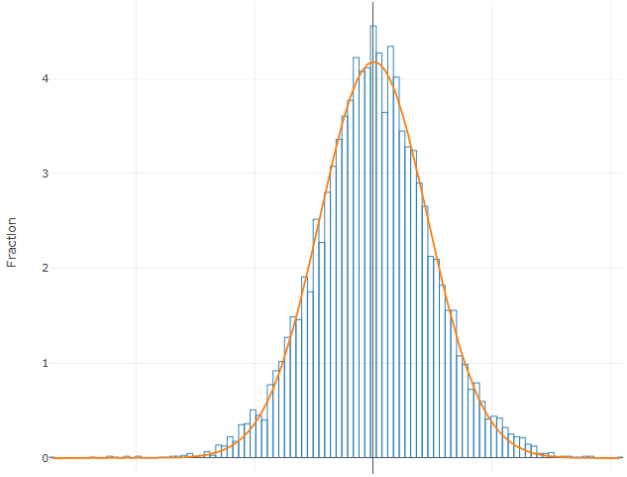
Type	Target	Algorithm	RMSE	R2	MAE
Regression	logHL	LightGBM	0.2414/0.2117	0.6674/0.7167	0.1887/0.1651

Cross-Validation Predicted Versus Actual





Cross-Validation Residual Distribution

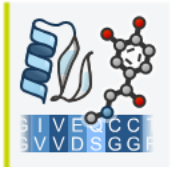


Light GBM Algorithm
R2=0.6906/0.7167

AI in Drug Discovery

Comprehensive Applications in Small and Biological Drug Design

3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

Use AI strategies to design protein binders and predict structures with AlphaFold.

Data visualization and Analysis

Insight for Research

Visualize your data and connect it to physics-based models.

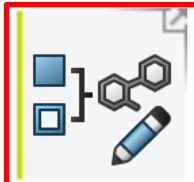
AI in Small Molecule Design

Machine Learning Workbench

Train ML models with your experimental data.

Generative Therapeutics Design

Automate the virtual creation, testing and selection of drug-like compounds using ML models and scientific methods.



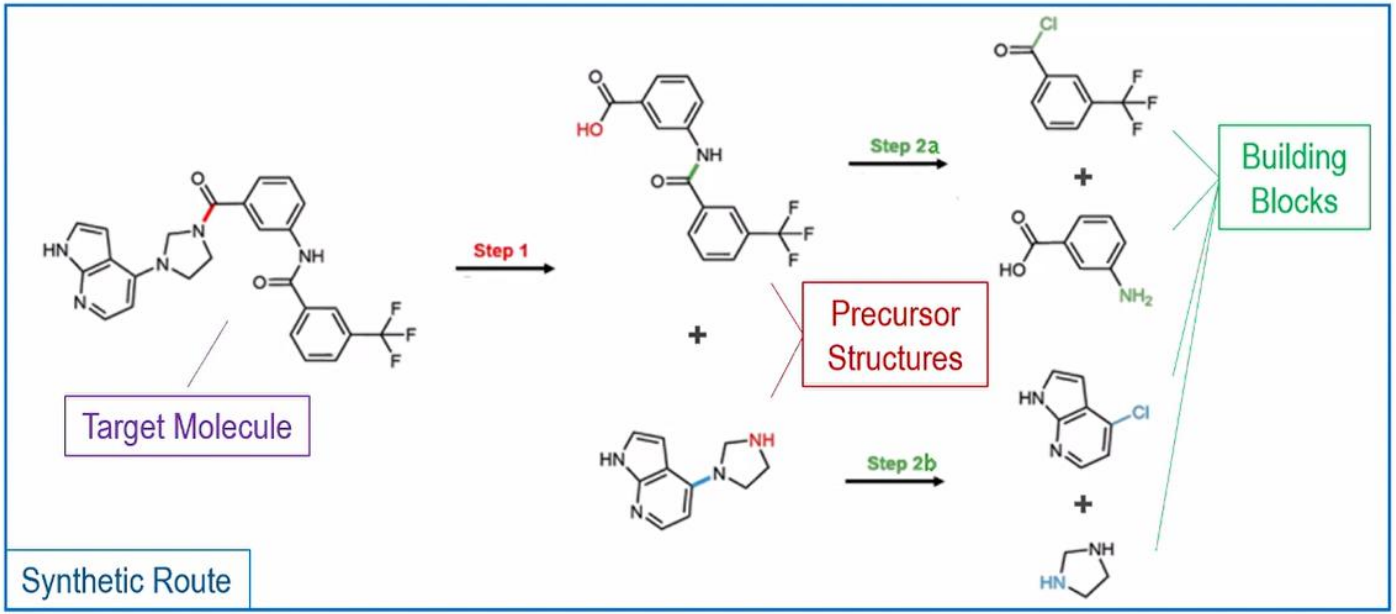
Reaction Planner

AI-based retrosynthesis prediction tool

DASSAULT SYSTEMES

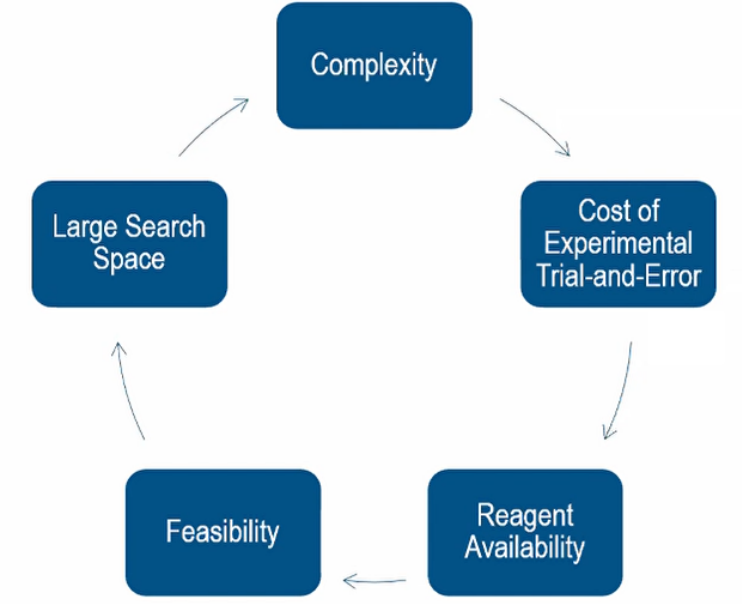
Reaction Planner

AI-based retrosynthesis prediction tool



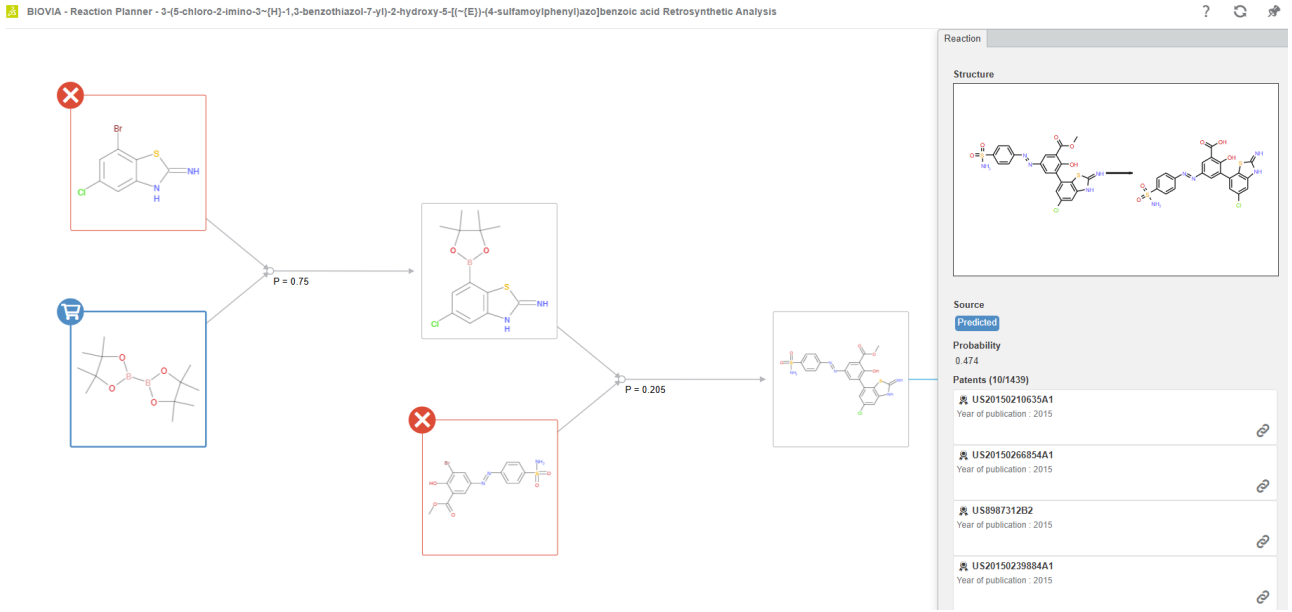
Retrosynthetic Analysis

Challenges



Reaction Planner

Retrosynthesis to design synthetic plan for small molecules



- Helps the Medicinal Chemists to identify synthetics routes to make targets
- Bridge the Virtual Design to organic molecule to the planning of Real Experiment

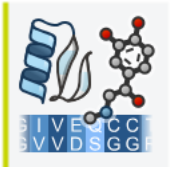
Main Capability

- Import std file, drop substance, sketch compound
- Fine-tune search parameters
- Preview the route, sort on scores and calculated properties, filter for reagent availability
- Investigate molecule and reaction details, review experimental conditions for known reactions, navigate to referenced US patents

Summary

AI in Drug Discovery

3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

Use AI strategies to design protein binders and predict structures with AlphaFold.

Data visualization and Analysis

Insight for Research

Visualize your data and connect it to physics-based models.

AI in Small Molecule Design

Machine Learning Workbench

Train ML models with your experimental data.

Generative Therapeutics Design

Automate the virtual creation, testing and selection of drug-like compounds using ML models and scientific methods.

Reaction Planner

AI-based retrosynthesis prediction tool

Agenda

- Scientific Platform Portfolio – R&D to Manufacturing
- AI in Drug Discovery
- **Q & A**

For more information please contact…

訊聯基因數位

台北市114內湖區新湖一路36巷28號

中華民國 台灣

Ph: (02) 2795 1777 x 3014 Fax: (02) 2793 8009

msc-support@gga.asia

www.gga.asia