



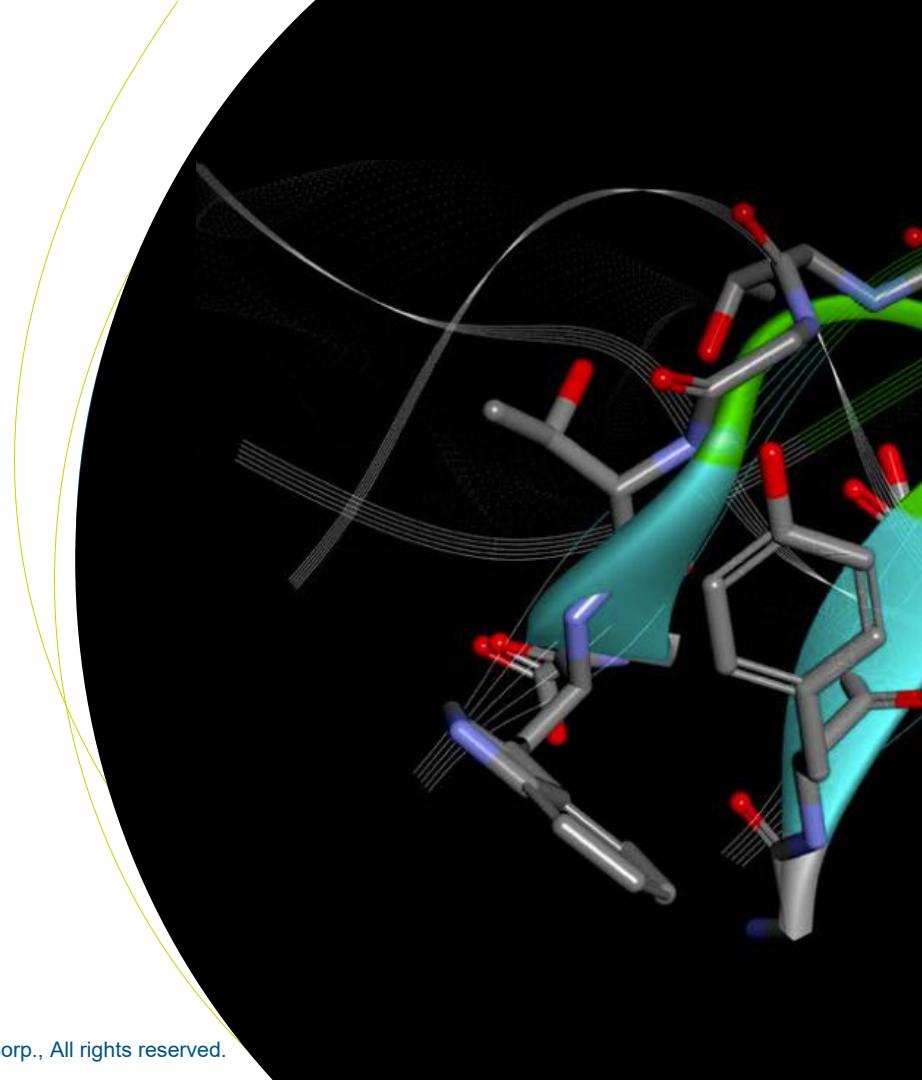
Accelerate Drug Design with AI and 3D Models

Cloud and On-Prem Developments in CY 2025



3DEXPERIENCE®

Novelyn Tsai
Engineer Specialist



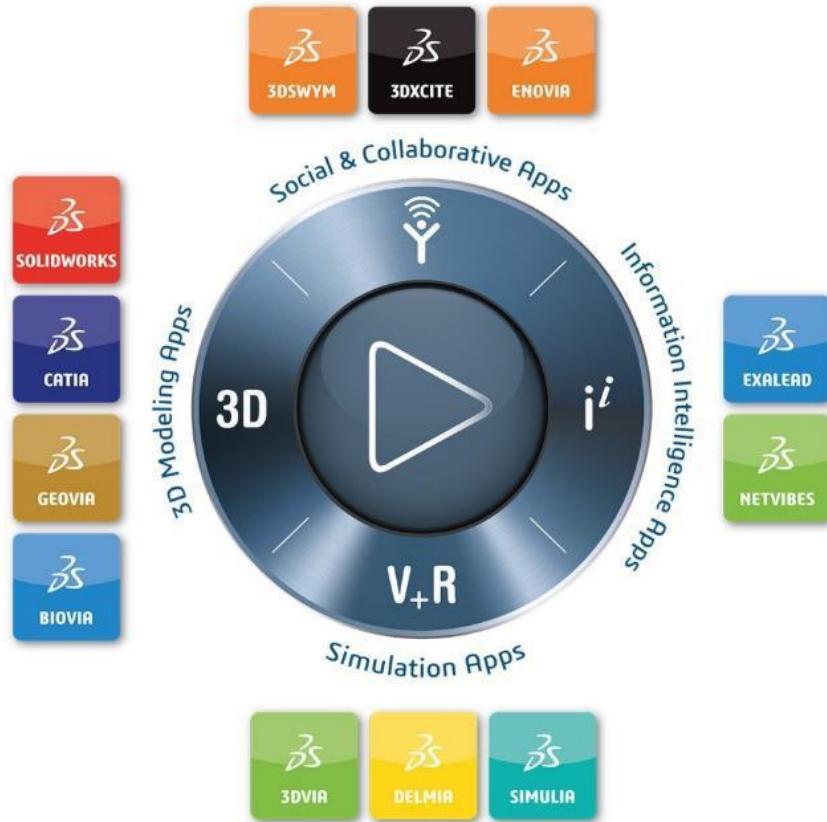
訊聯基因數位

Copyright©2025 GGA Corp., All rights reserved.

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

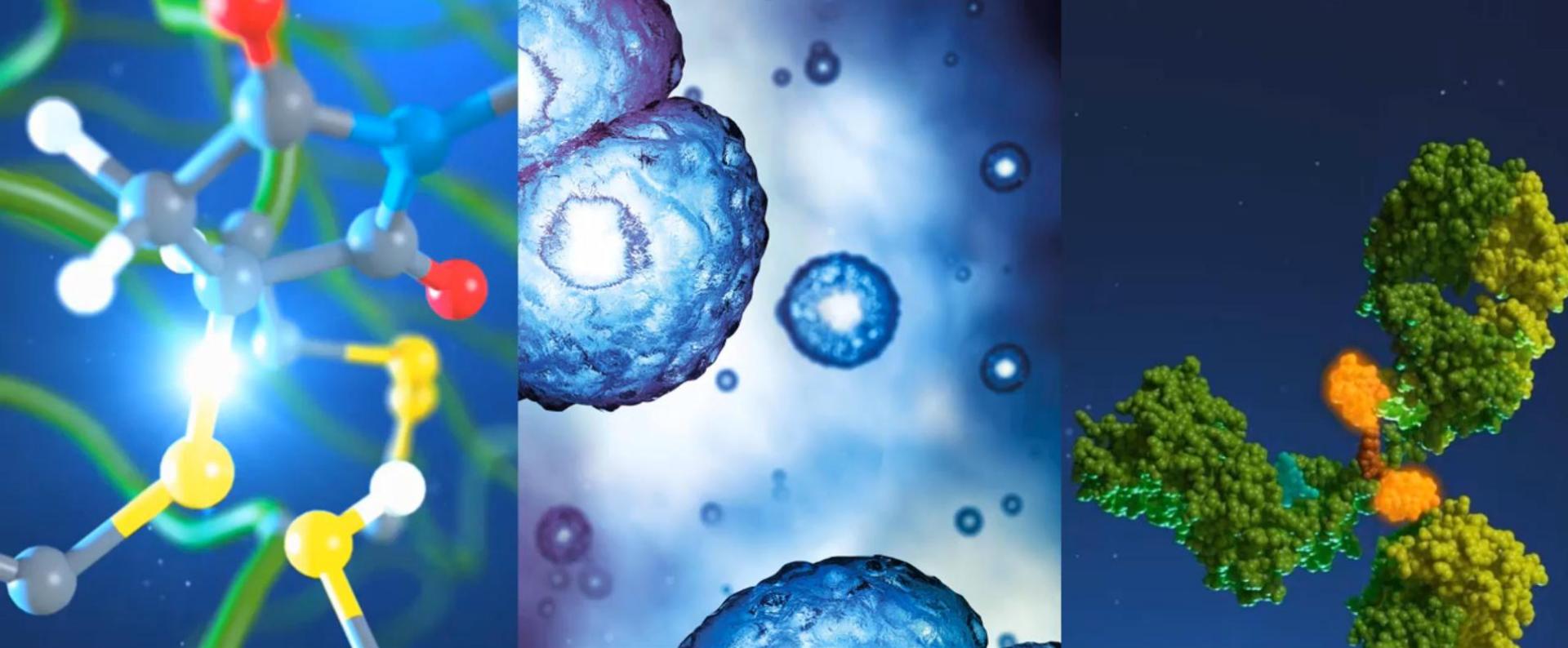
- Security and privacy are core to the **3DEXPERIENCE® platform**, supporting trust and compliance globally.
- ISO/IEC 27001:2017 & ISO/IEC 27701:2019 certified Information Security & Privacy Management System (ISPMS).
- Built on principles of **confidentiality, integrity, availability, and accountability**.

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

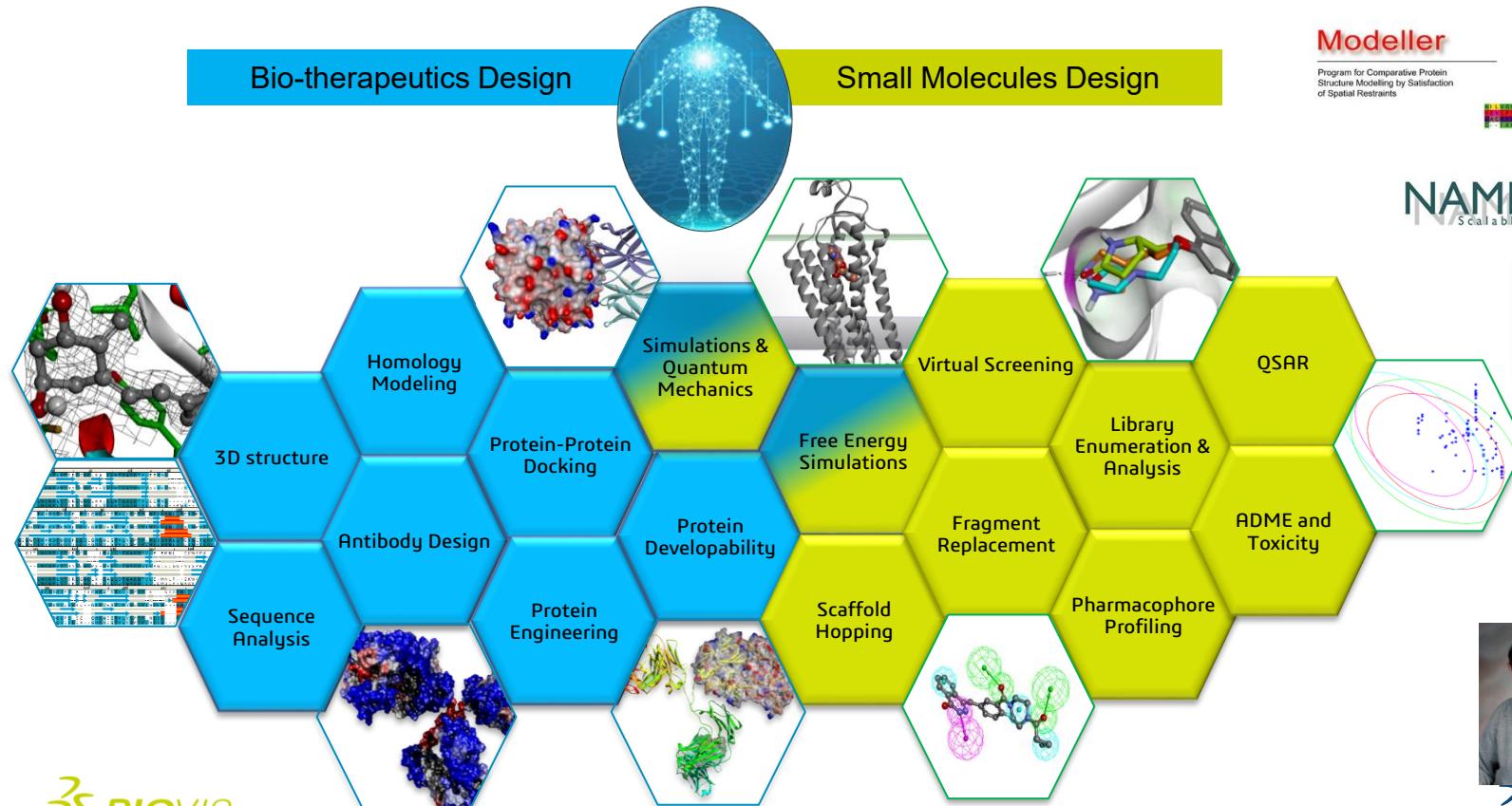
- Structure-Guided Discovery of PD-1/PD-L1 Interaction Inhibitors: Peptide Design, Screening, and Optimization via Computation-Aided Phage Display Engineering, *Journal of Chemical Information and Modeling*, 2024
- Biophysical mechanisms underlying tefluthrin-induced modulation of gating changes and 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-1 β induced 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



DS BIOVIA



訊聯基因數位 Copyright©2025 GGA Corp., All rights reserved.

Modeller

Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints



NAMD  **ILLINOIS**
Scalable Molecular Dynamics

UCSF

University of California
San Francisco

MIT
Massachusetts
Institute of
Technology



Bioinformatics

Biologics Lead Identification

Biologics Lead Optimization

Small Molecules Lead Identification

Cheminformatics

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

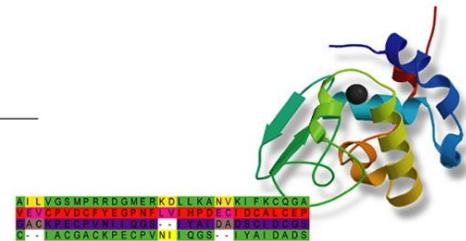
Best Validated Science – 30+ Years History

- 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

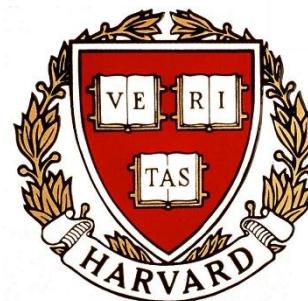


Modeller

Program for Comparative Protein
Structure Modelling by Satisfaction
of Spatial Restraints



University of California
San Francisco



Agenda

Scientific Platform Portfolio – R&D to Manufacturing

AI in Drug Discovery

Q & A



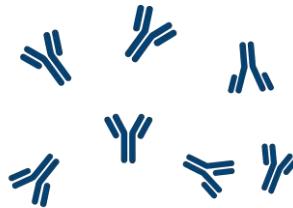
THE DISCOVERY BOTTLENECK



5+ Years

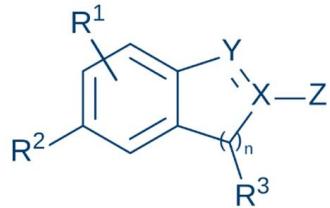
Average time spent in discovery

Can AI do better?



Thousands of Biologics

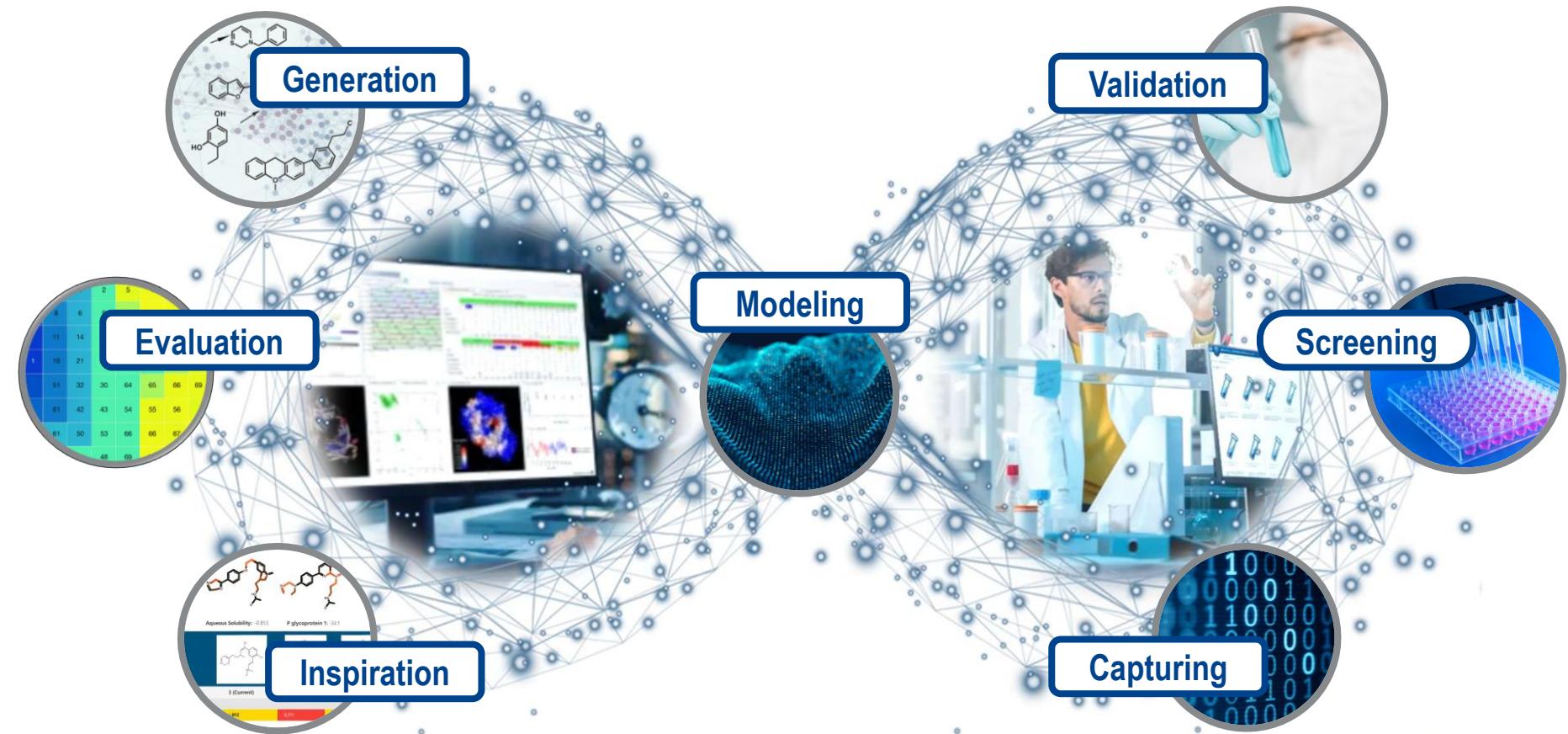
To find a viable candidate



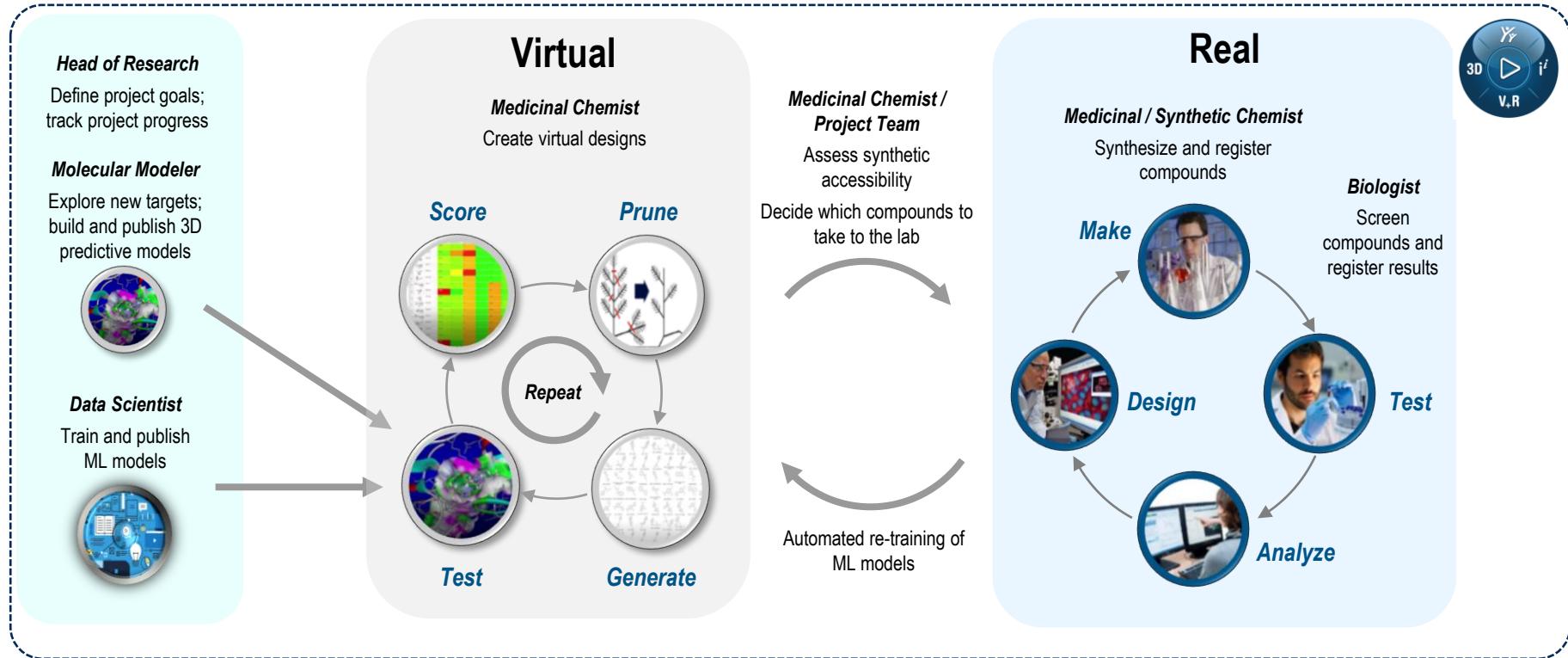
4,000 Compounds

To find a viable candidate

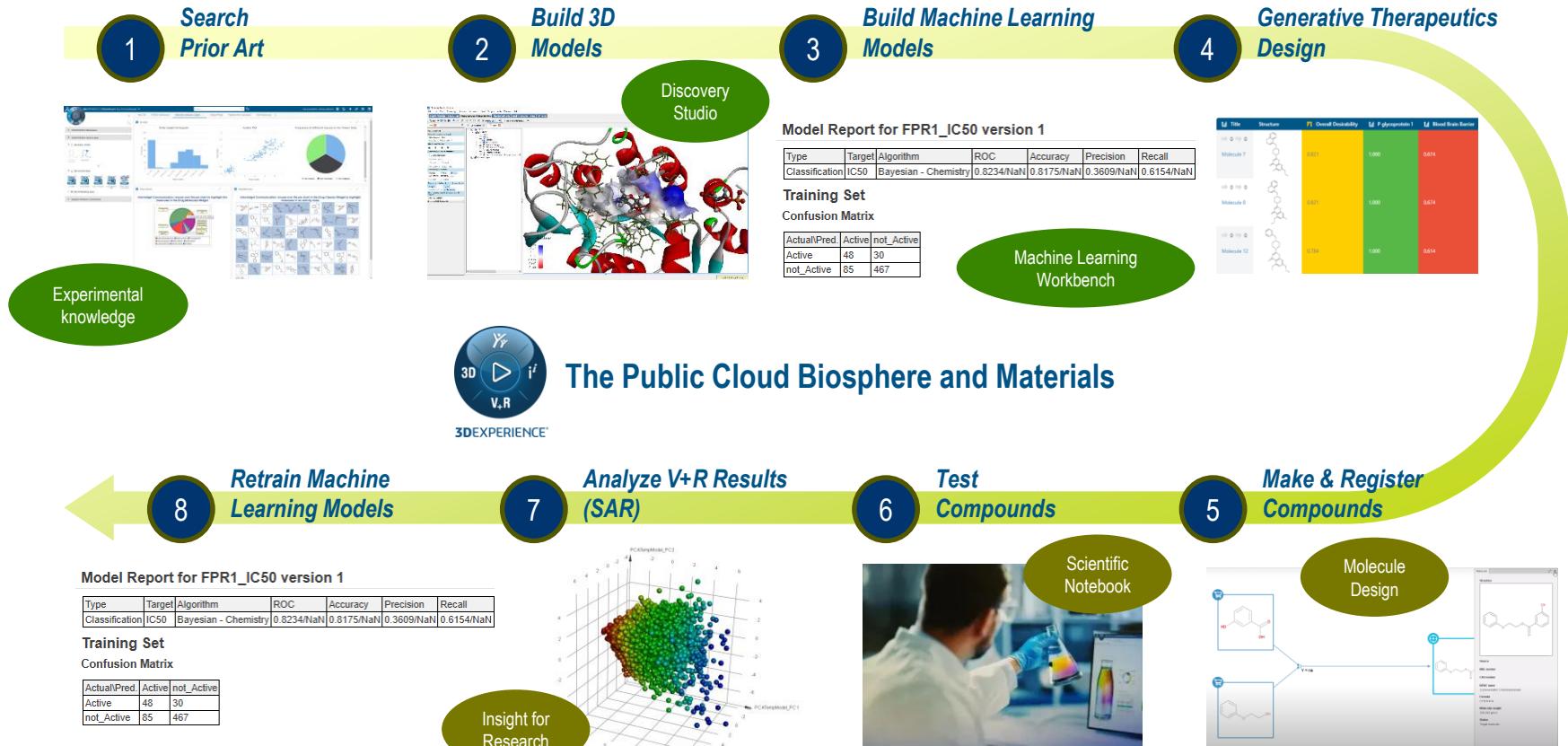
ACTIVE AI-ASSISTED V+R CYCLE AIDS DRUG DISCOVERY



GENERATIVE THERAPEUTICS DESIGN PROCESS

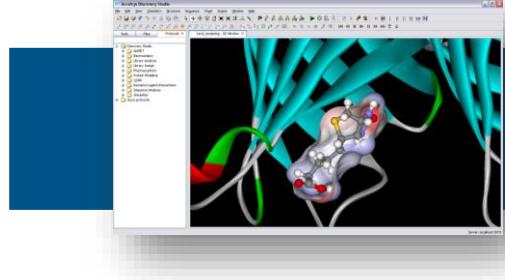


COMPREHENSIVE SOLUTION FOR EARLY DISCOVERY with 3DE



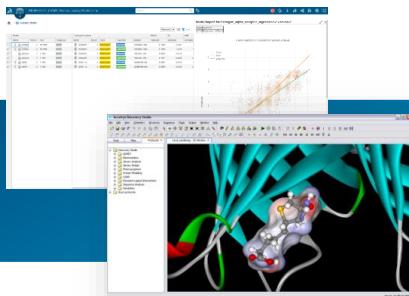
DISCOVERY STUDIO SIMULATION

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



3DEXPERIENCE®

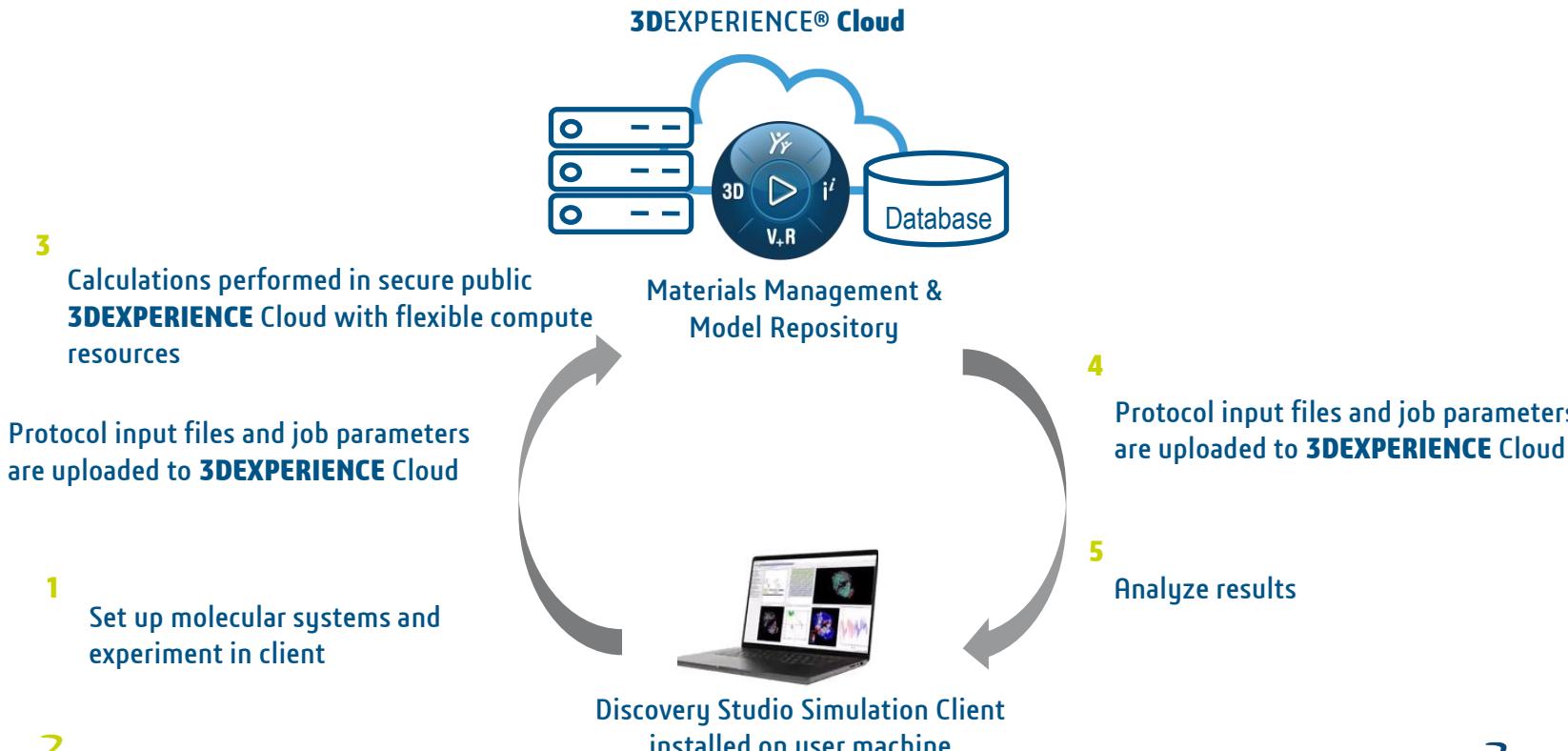
Discovery Studio Simulation

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	○	○	○
Discovery Studio Simulation	○	○	✗
Generative Therapeutics Design	○	✗	○
Insight for Research	○	○	○
Job Management	○	○	○
Molecular Design	○	○	○
Scientific Notebook	○	○	○
Scientific Search Intelligence	○	○	○
Scientific Settings and Administration	○	○	○
Materials Management	○	○	○
Machine Learning Workbench	○	○	✗
Reaction Planner	✗	✗	○



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.

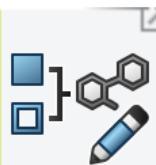
AI in Small Molecule Design



Machine Learning Workbench



Generative Therapeutics Design



Reaction Planner

Train ML models with your experimental data.

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

AI-based retrosynthesis prediction tool

RELEASES



DRUG THERAPEUTICS DESIGN

What Therapeutic Do You Have?



Biological Therapeutics (Biotherapeutics)

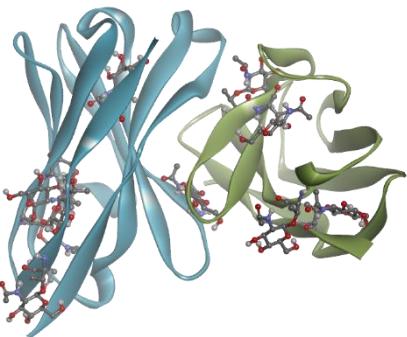
Recombinant proteins

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



CCDC

Deposit Structures Access Structures Contact Us

Home Solutions Community Discover Consultancy Services Research Support and Resources About

Mark to Discover

Written by
CCDC Team

Posted on
February 17, 2022

CCDC  BIOVIA

Virtual drug design platform
set to accelerate drug
discovery

Tags

CSD-Discovery (22)

Docking (115)

Drug Discovery (98)

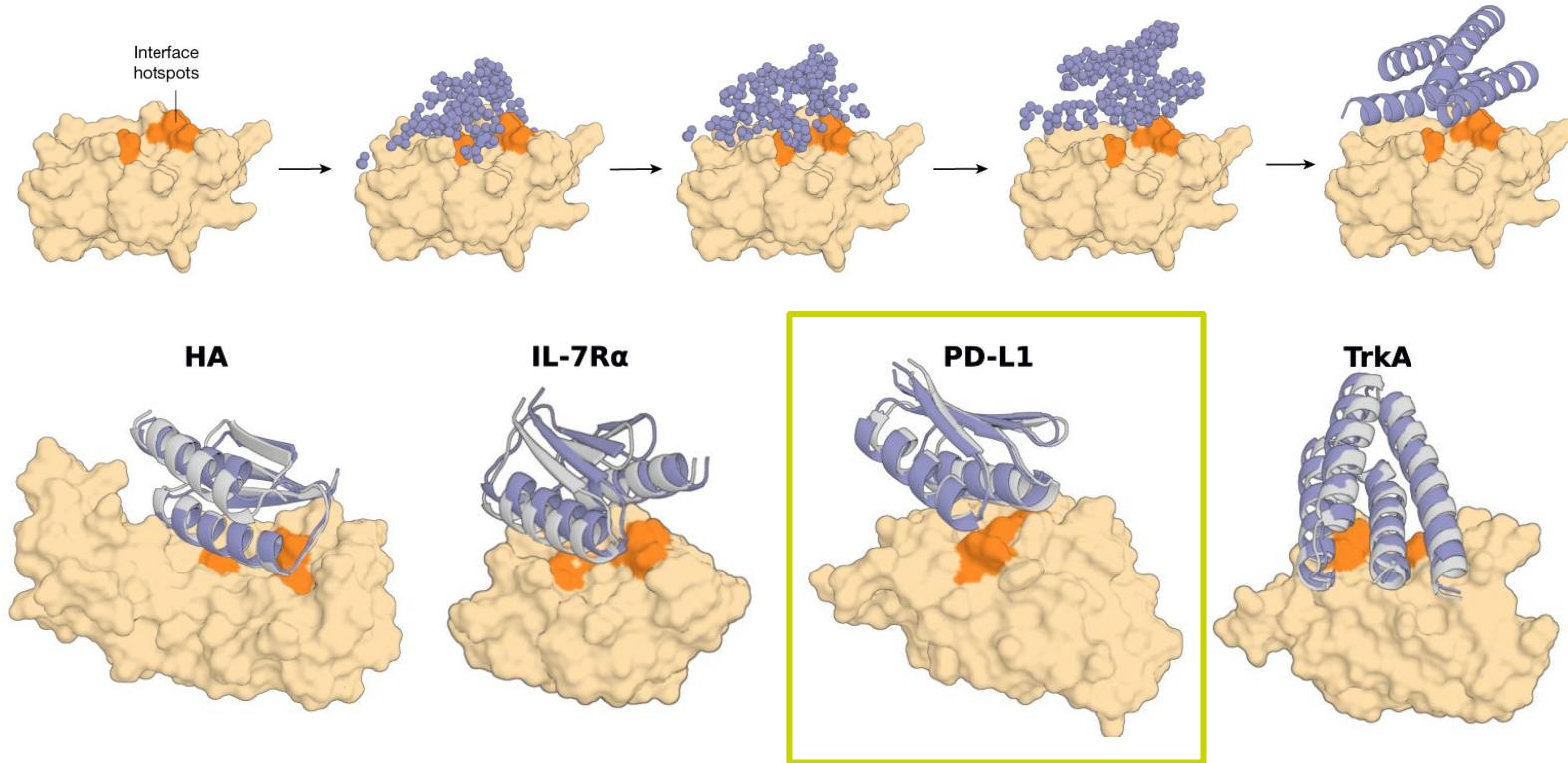
GOLD (29)

Pharmaceutical Discovery (38)

Pharmaceuticals (51)

Virtual Screening (5)

DE NOVO DESIGN OF PROTEIN-BINDING PROTEINS



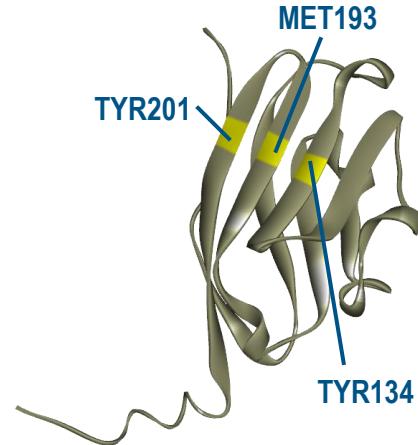
Watson et al. De novo design of protein structure and function with Rfdiffusion. *Nature* (2023).



Generate protein scaffolds from a template structure using RFDiffusion.

PDB: 5O45

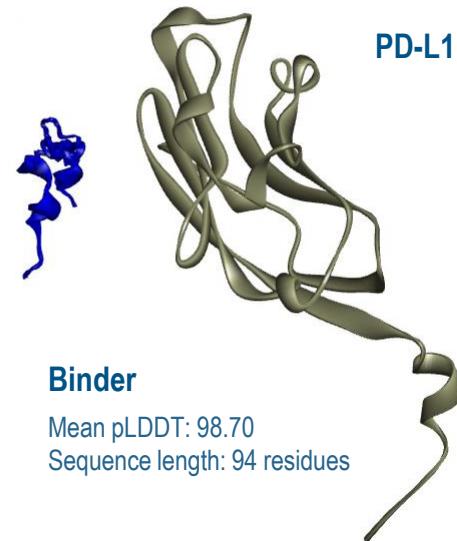
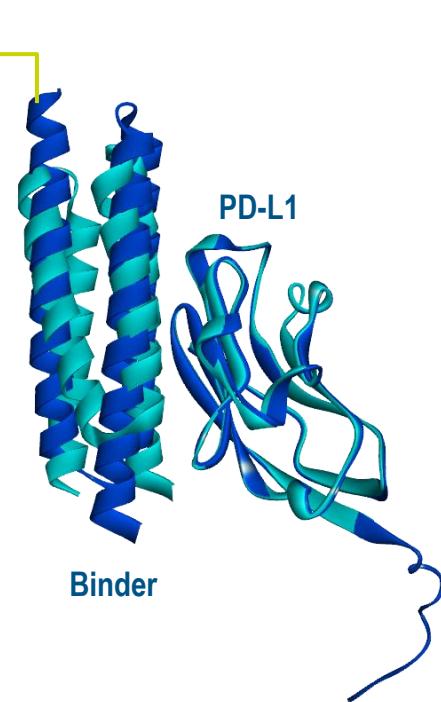
Select hotspot residues for binder design



© Dassault Systèmes | Confidential Information | 12/24/2025 | ref.: 3DS_Document_L2025

Parameter Name	Parameter Value
Template Scaffold	PD-L1_RFDDiff: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

RFDDiff PDL1_binder_design *Nature* volume 620, pages 1089–1100 (2023)
RFDDiff PDL1_binder_design (Discovery Studio)



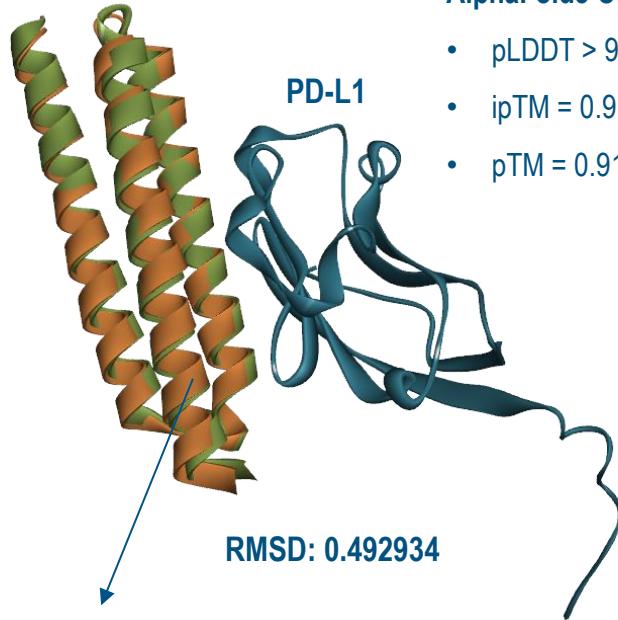
→ A:17-145 60-100 Generate a potential binder containing 60-100 residues
→ Targeting hotspot residues 17-145 from chain A in the template.

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	AVEEARRRLEELLERARRELERLLARVAAAAPAELDAALTDEFLLALSRLRRETRRLIER DPEIAEEMRRRTEELIRRALAEFDAVRAKLA
PD-L1_RFDiff_77	77	0.1	111	0.3963	0.3963	0.2511	AAAAARAELDALLAADATAAAALLAAVAAAAPAELTALRDAFMLALSDLARATAALV ARDPALAAEFRARTEARIRRWLTEFDAVRAK
PD-L1_RFDiff_35	35	0.1	111	0.3939	0.3939	0.2422	MEEEARRRLELLREFEERLRRLLARVAAAADPAELDALRDEFLLAISELSRRTKELIEE YPELAEEFRRRLEEAIRRAQREFDAVRAH
PD-L1_RFDiff_27	27	0.1	111	0.3931	0.3931	0.2466	AAAEARARLDALLAEAEATARRLLERVAAAAPEELDALRDEALLAVSRLRRETAALIA EHPDLAAEMRERTEAAILRWLREFDAVRAKRAAE
PD-L1_RFDiff_29	29	0.1	111	0.3926	0.3926	0.2287	SMEARKKLEELLEEAAERRLRELLARVAAAAPAELDAALTDEFLLALSELSRRTKELIEK YPEIAEEARKKTEELIRRHVEFTA
PD-L1_RFDiff_8	8	0.1	111	0.3923	0.3923	0.2466	ATAEADARLDALLAADATAAAALLARVAAAAPAELSDLRDEAMLAISRLRRETAALIA ADPARAEEFRKRTEERTLEHLRALDAFRAH
PD-L1_RFDiff_25	25	0.1	111	0.3919	0.3919	0.2511	SAAAAEARLDALLEADAAAAALLARVAAAEPAEGLRDEFLLAVSDLARATAALIA EDSELAAEFRERTEARIDQWFRDFDAVRAH
PD-L1_RFDiff_49	49	0.1	111	0.3906	0.3906	0.2466	MEEEARRRLELLRRAEEHRELLERVEAELEELPELKNEFMLALSRLRRETAALIAE YPELAEEMRRRTRELIKMTREFIEAVRKREEAK

Sequence length: 94 residues

Predict protein structure for PD-L1 Binder



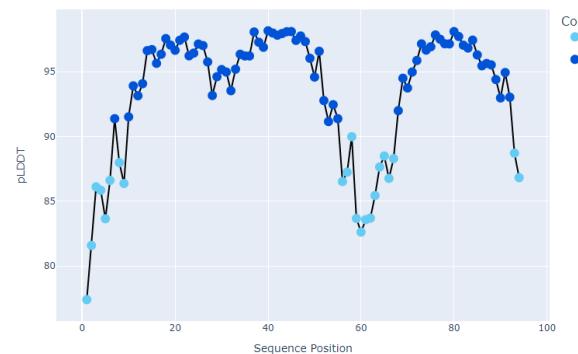
AlphaFold3 Confidence:

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

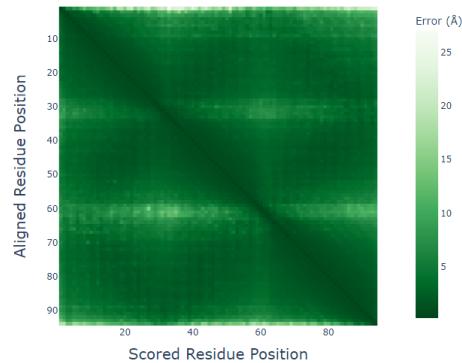
RFDiff Structure Prediction Confidence:

- Mean pLDDT = 93.44

Model Confidence: PD_L1_RFDiff_73 chain A



Predicted Aligned Error: PD_L1_RFDiff_73



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)

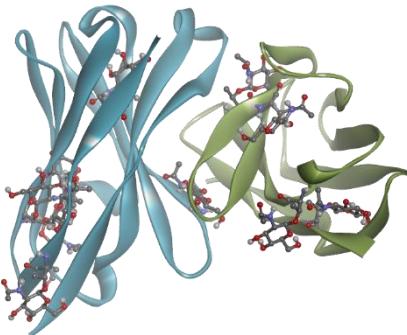
Recombinant proteins

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



CCDC

Deposit Structures Access Structures Contact Us

Home Solutions Community Discover Consultancy Services Research Support and Resources About

Mark to Discover

Written by

CCDC Team

Posted on

February 17, 2022

CCDC  BIOVIA

Virtual drug design platform
set to accelerate drug
discovery

Tags

CSD-Discovery (22)

Docking (115)

Drug Discovery (98)

GOLD (29)

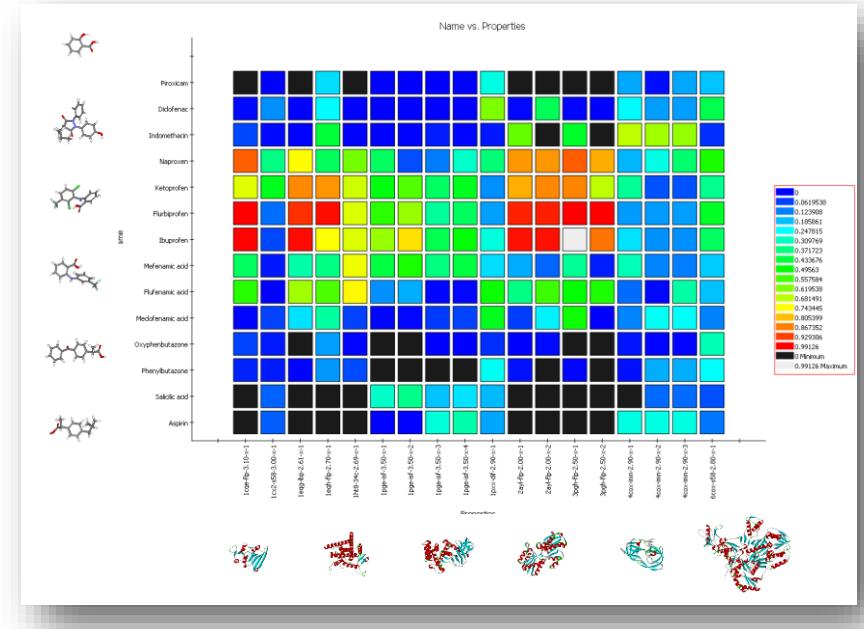
Pharmaceutical Discovery (38)

Pharmaceuticals (51)

Virtual Screening (5)

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

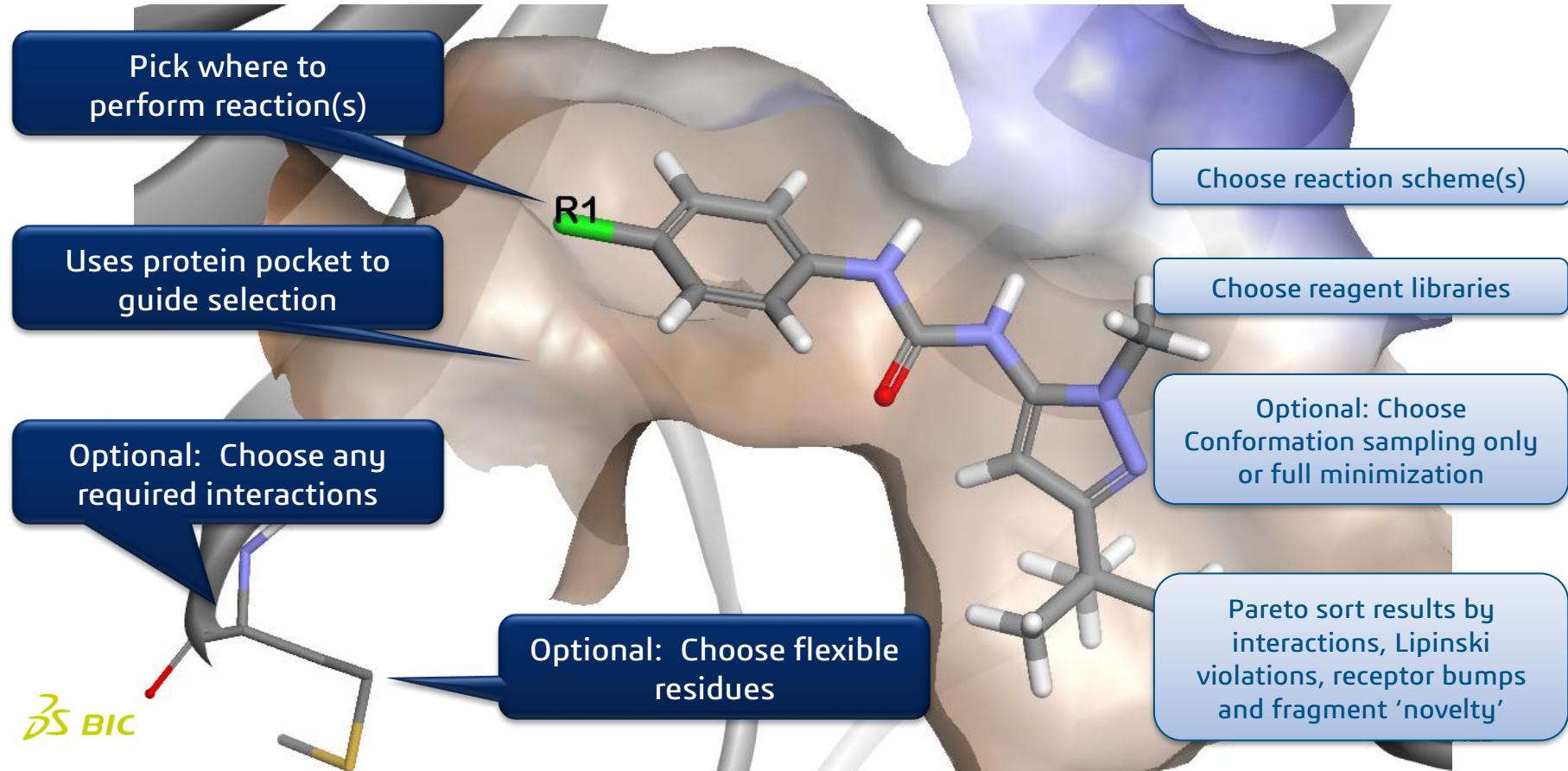
* Kellenberger *et al*, *J Chem Info Model*, **2006**, *46*, 717-727

What if the hit has multiple targets?

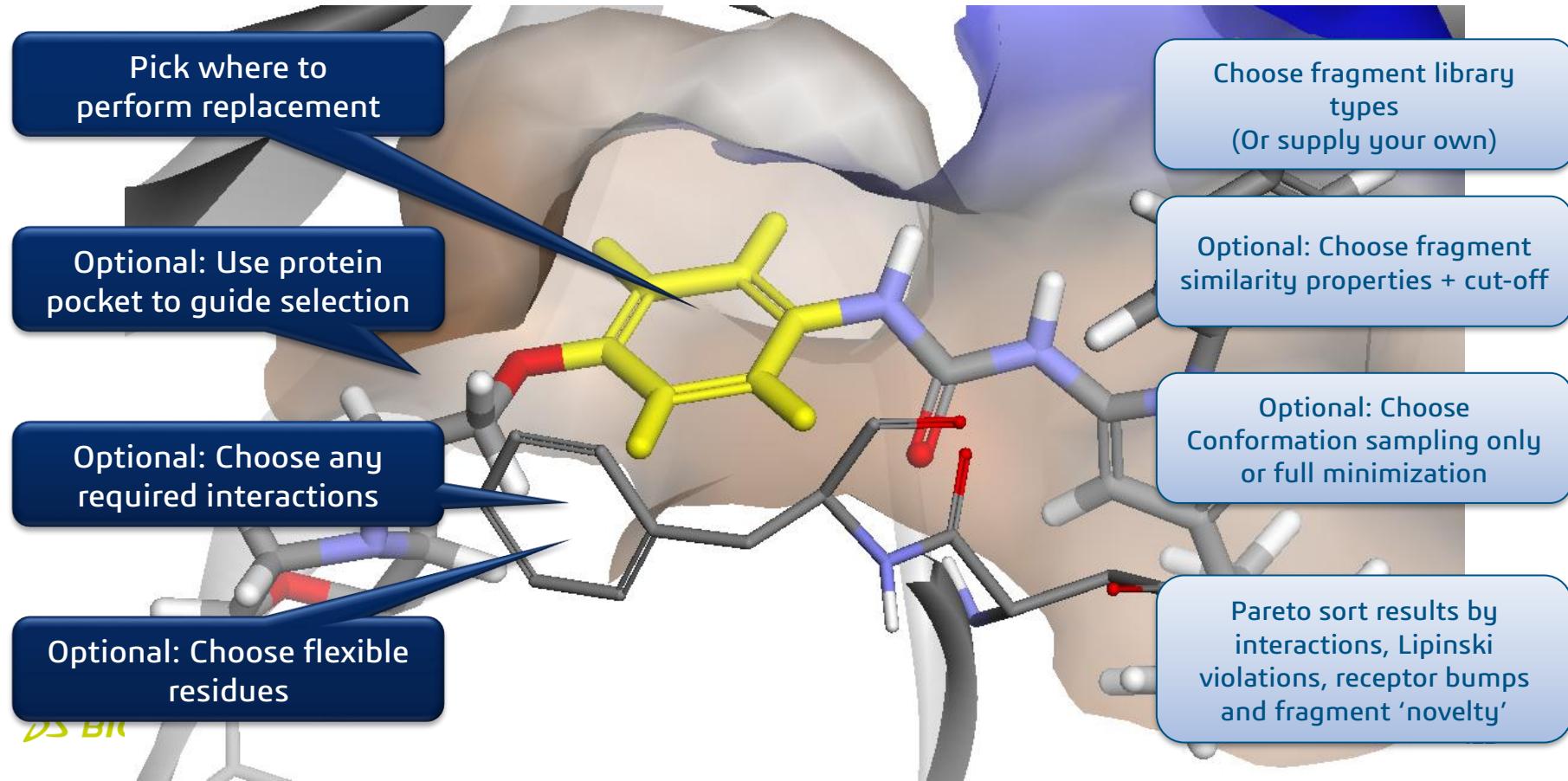
SBD: Fragment-Based Design Methods

- 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

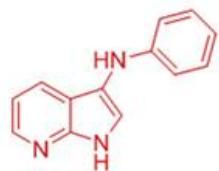


REPLACE: Fragment Based *In-Situ* Substitution



Fragment-based design of the BRAF inhibitor vemurafenib.

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

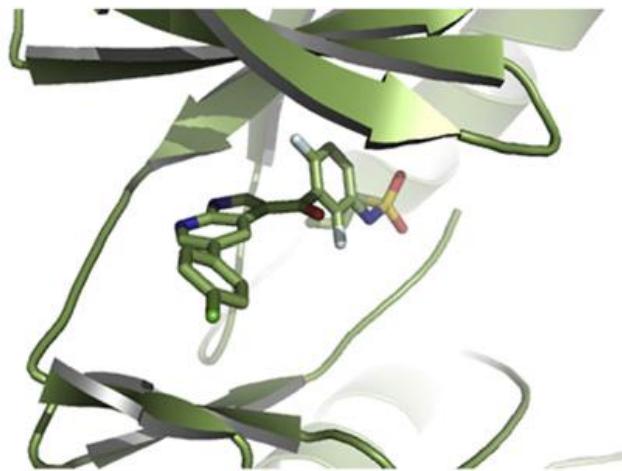


Unselective, weakly potent fragment hit ($IC_{50} > 100\mu M$)
PIM-1 $IC_{50} \sim 100\mu M$

Fragment growing



BRAF (V600E) $IC_{50} = 31nM$
High degree of selectivity against other kinases
PIM-1 $IC_{50} > 100\mu M$

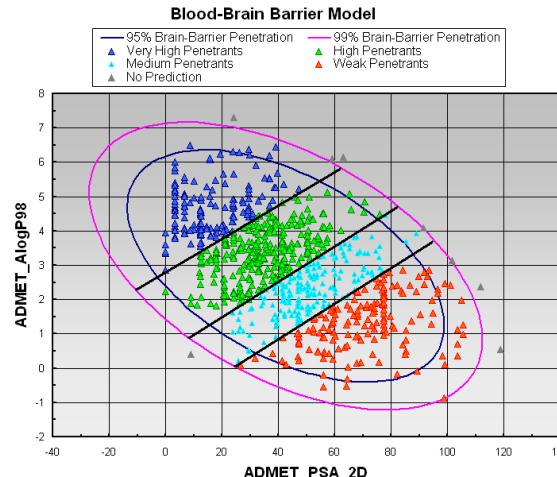
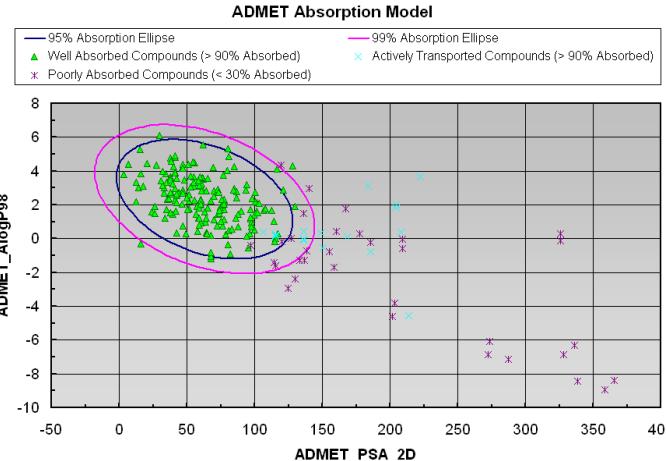


Swen Hoelder, Paul A. Clarke, Paul Workman, 2012

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



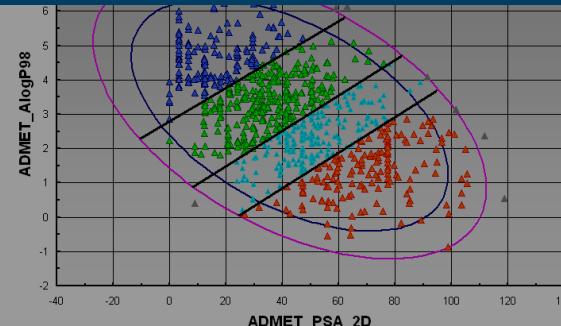
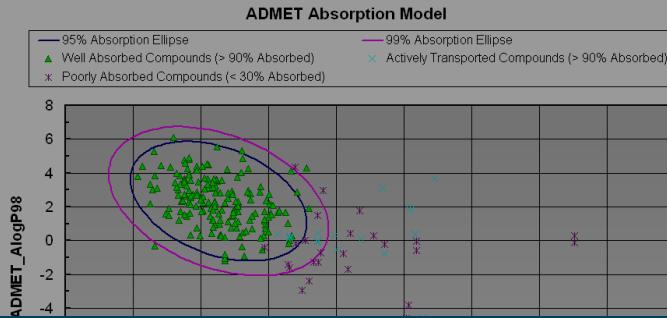
On Premise

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



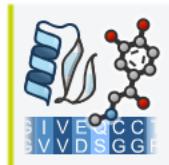
Can we use AI for optimization?

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



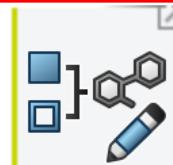
Machine Learning Workbench

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.

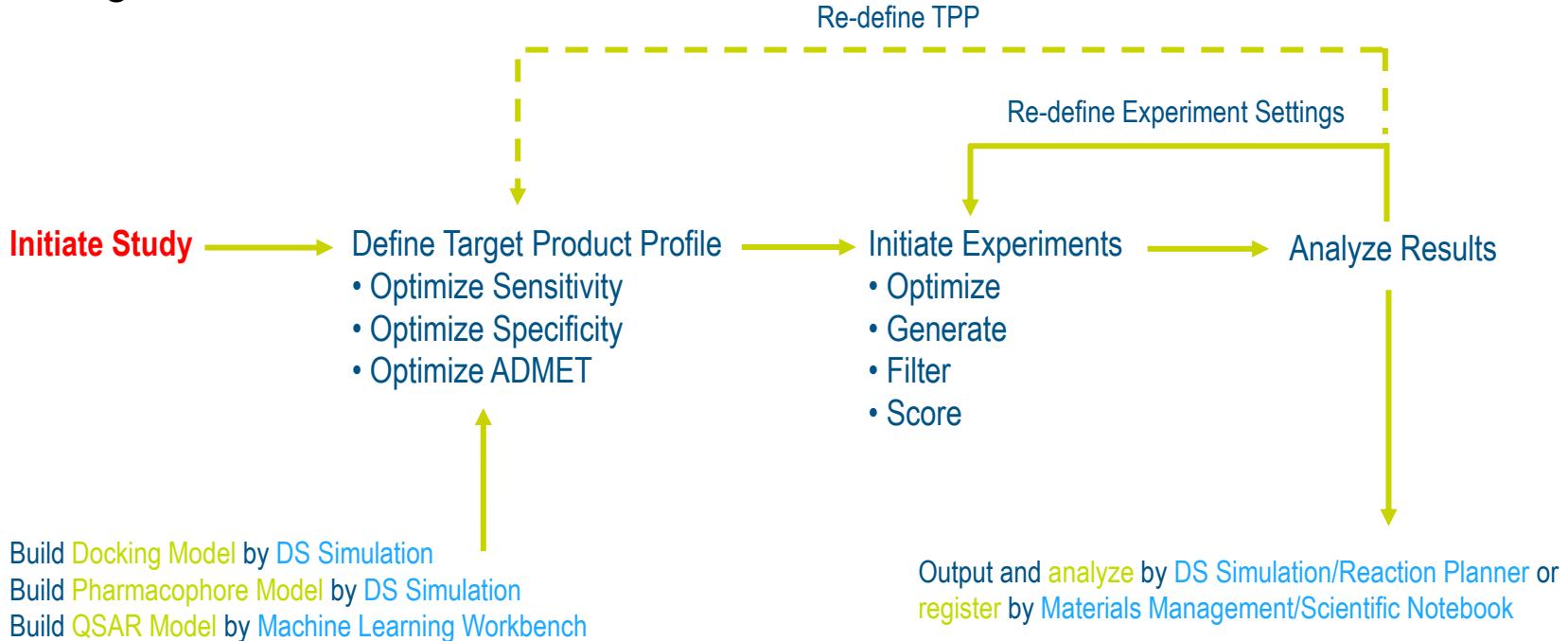


Reaction Planner

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.

DS 3DEXPERIENCE | 3DDashboard BIOVIA Users Home Page

Jeff Ma DS - R1132101124092

Computational Chemist

Discovery Studio Simulation

Generative Therapeutics Design

Insight for Research

Job Manager

Machine Learning Workbench

Materials Management

Molecular Design

Scientific Notebook

Scientific Search Intelligence

Scientific Settings and Administration

Get Started

Discover BIOVIA Roles

Discover BIOVIA Apps

Discover BIOVIA Training

BIOVIA - Generative Therapeutics Design - Home

Generative Therapeutics Design

Accelerate drug discovery with AI

+ New

Open

Administration

Recent Content: Studies

Test 11 hours ago

test_pro 3 weeks ago

Content & Knowledge

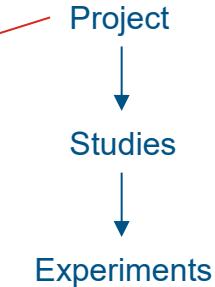
Social Networks

User Assistance

Communities

What's New

DS BIOVIA



Training Materials

Selective Estrogen Receptor Modulator

- Two different subtypes of ER have been identified, ER α and ER β
- ER α and ER β 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 ER α , which are replaced by Met-336 and Ile-373, respectively, in ER β .

1. Check For Protein Sequence Similarity

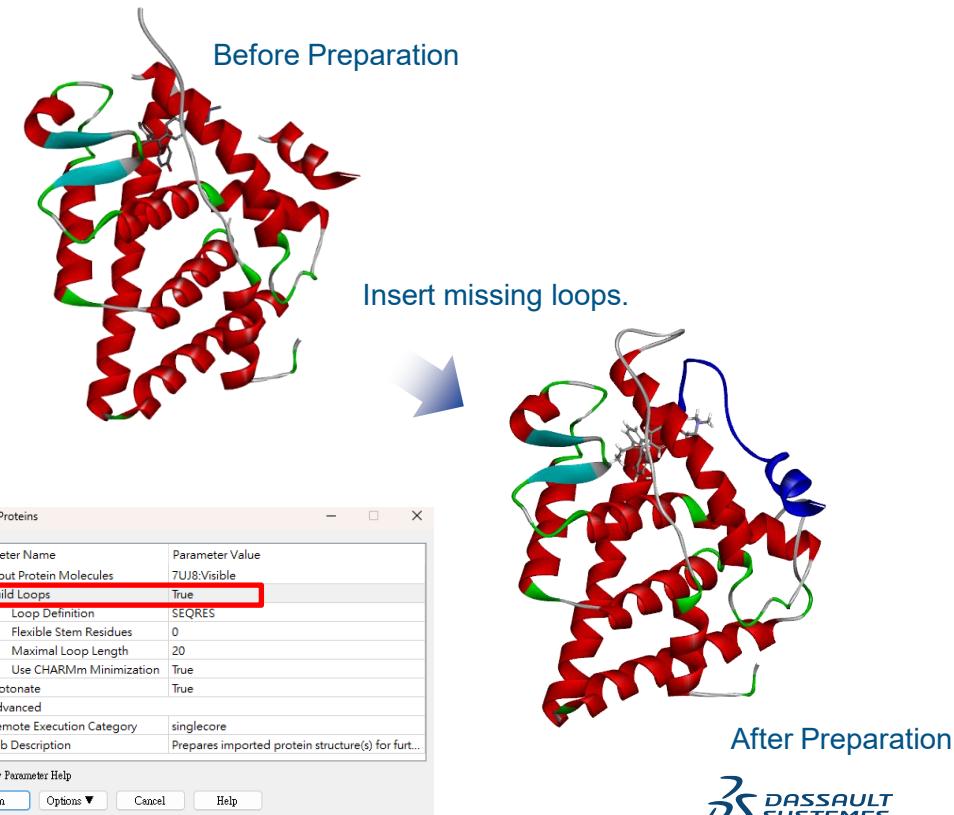
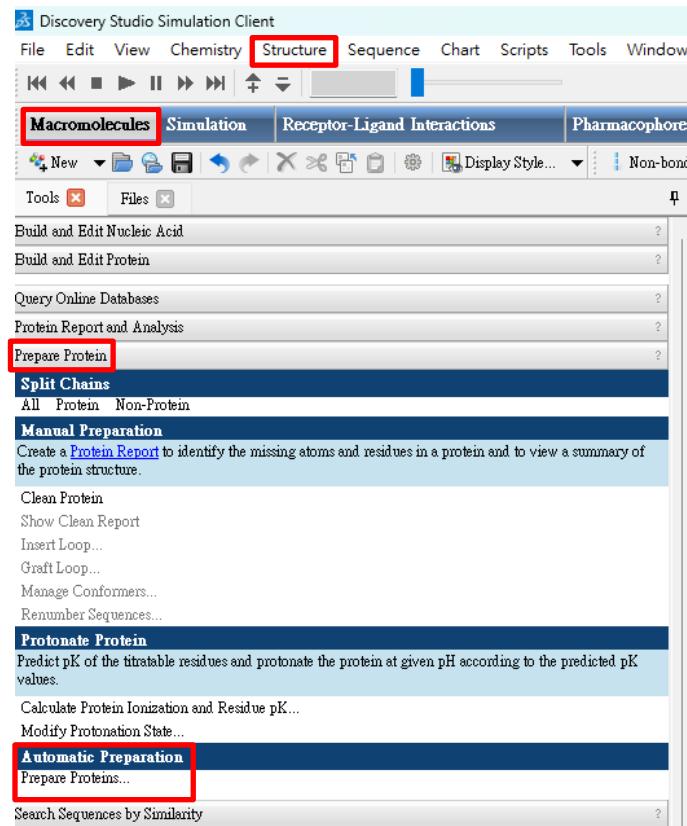
Conserved vs Non-Conserved Residues

The figure shows the Discovery Studio Simulation Client interface. The 'Sequence' menu is highlighted with a red box. The 'Show Sequence' option is also highlighted with a red box. The 'Sequence' menu contains the following items: Chain and Domain, Gaps, Features, Strand, Rename Sequence..., Alignment, Secondary Structure, Translation..., Link Sequence and Structure..., Show Sequence (Ctrl+Q), and Show Annotations. The 'Sequence' menu is open, showing the 'Show Sequence' option. The 'Macromolecules' tab is also highlighted with a red box. The main workspace displays two protein structures (2NV7 and 7U8) in ribbon format, with one structure colored red and the other blue. A sequence alignment window is open, showing the amino acid sequences of 2NV7 and 7U8. The alignment shows a high degree of sequence identity and similarity. Text in the workspace indicates: 'Sequence identity = 53.9%' and 'Sequence similarity = 79.4%'. A large blue arrow points from the aligned sequences towards the superimposed protein structures.

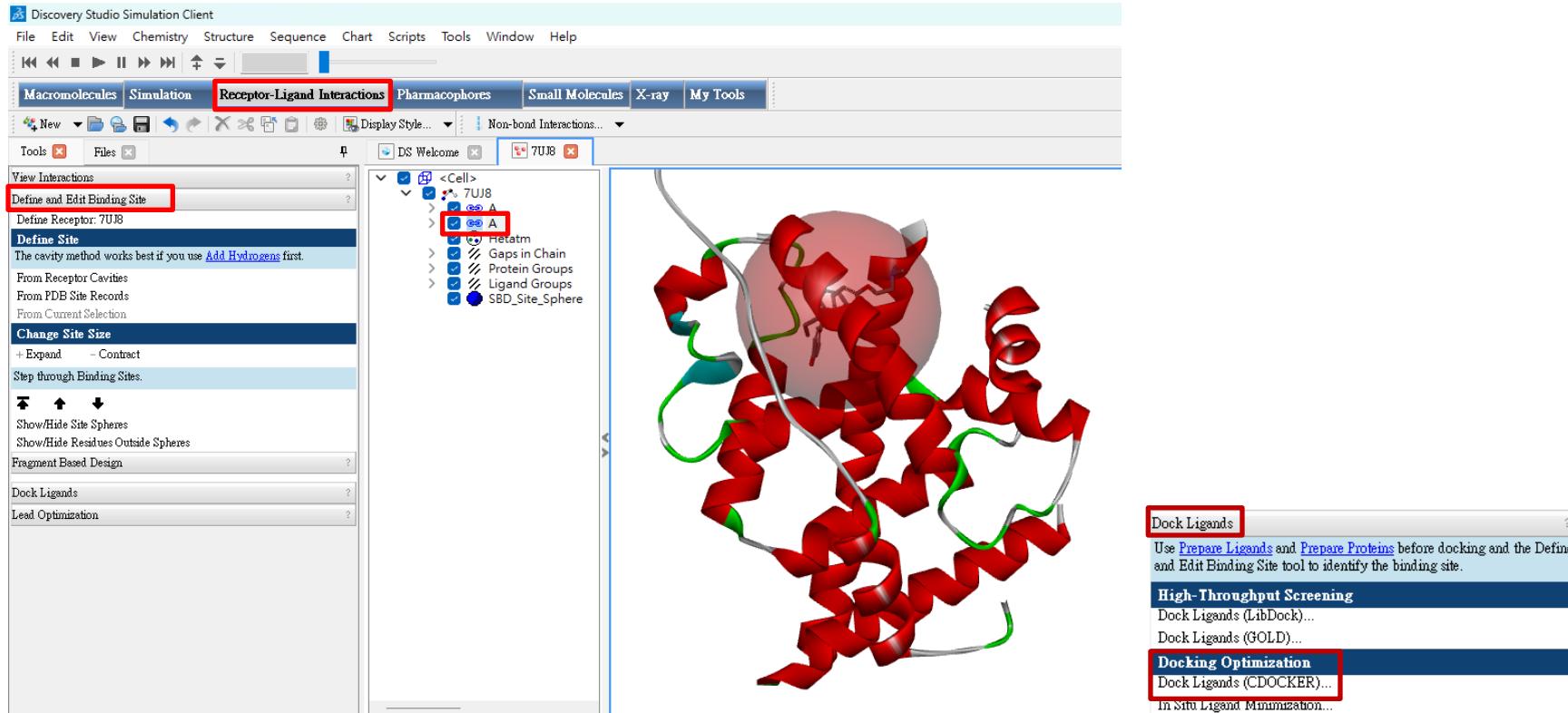
ERalpha: 7UJ8

ERbeta: 2NV7

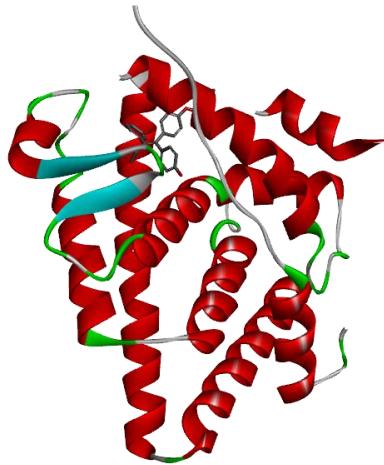
2. Protein Structure Preparation



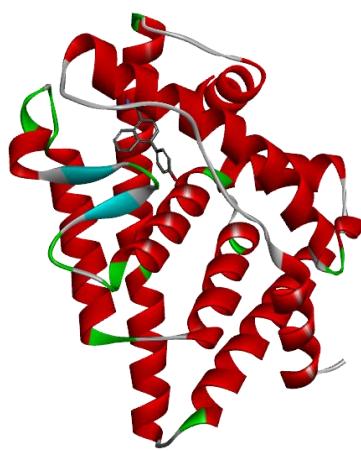
3. Protein-Ligand Molecular Docking



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

Small Molecules

Model Repository Management

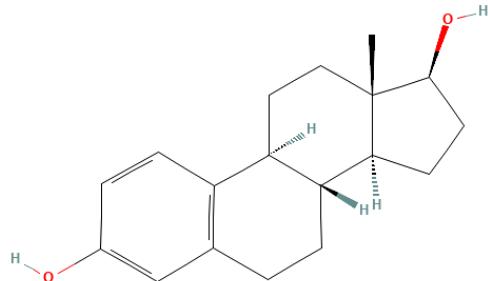
Published

Published to the Model Repository.

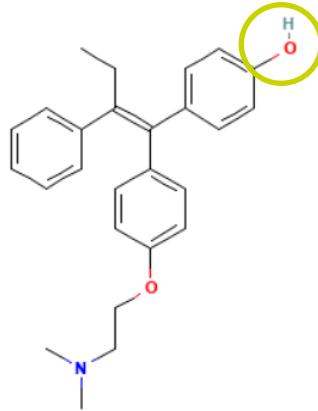
Published Pharmacophore to Model Repository...

Published Docking to Model Repository...

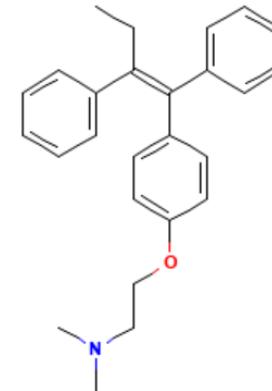
ER inhibitors



17 β -Estradiol



4-Hydroxytamoxifen

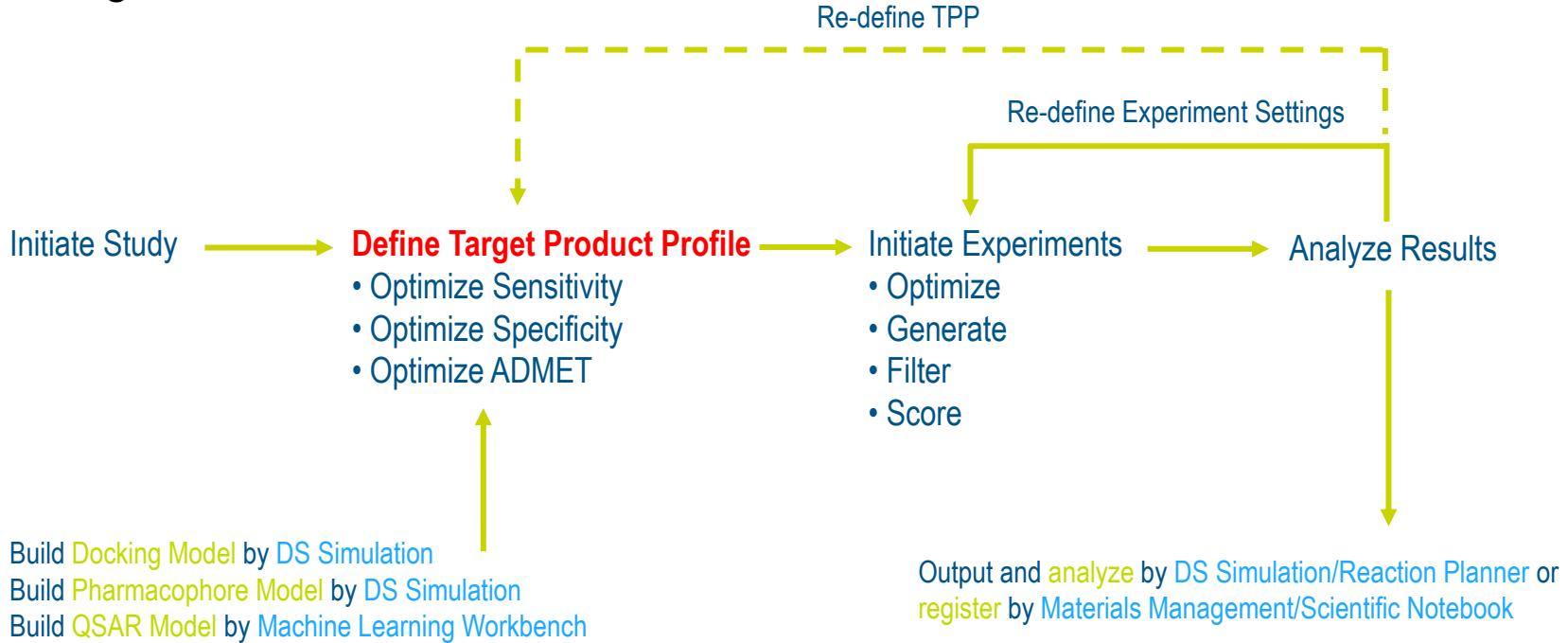


Tamoxifen

- 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.

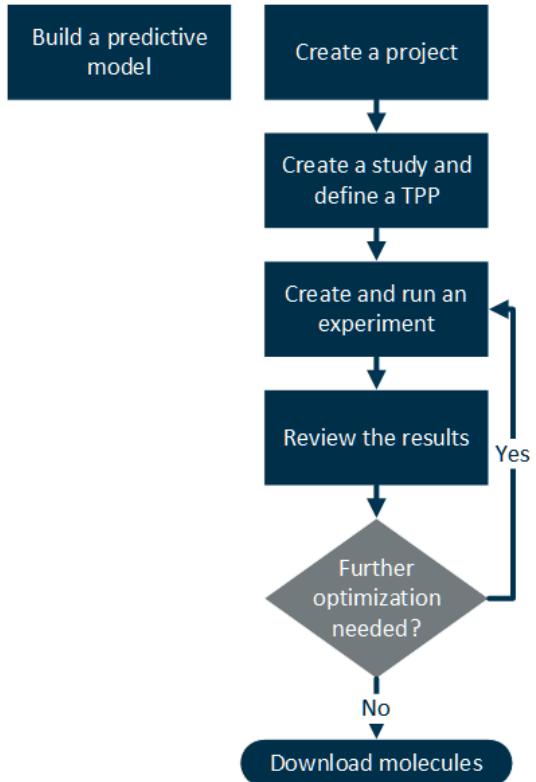
Generative Therapeutics Design (GTD)

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

Home / Project of 08/04/2021 at 14:38:15 / Create STUDY

GTD PROJECT

Summary Target product profile

**DS Simulation**

**AntiTargets**

**ADME**

**ML Workbench**

**Toxicity**

Targets

AntiTargets

ADME

ML Workbench

Toxicity

Filter target models... 

Filter antitarget models... 

Filter ADME models... 

Filter toxicology models... 

***MAO_Model**

Acetylcholinesterase

Acetylcholinesterase pIC50 RF

Acetylcholinesterase RF

Adenosine A1 receptor

Acetylcholinesterase

Acetylcholinesterase pIC50 RF

Acetylcholinesterase RF

Adenosine A1 receptor

Aqueous Solubility

Aqueous Solubility 10uM MPNN

Aqueous Solubility 100uM MP...

Blood Brain Barrier

Aerobic Biodegradability

Ames Mutagenicity

CYP2D6 Inhibition

Developmental Toxicity Potential

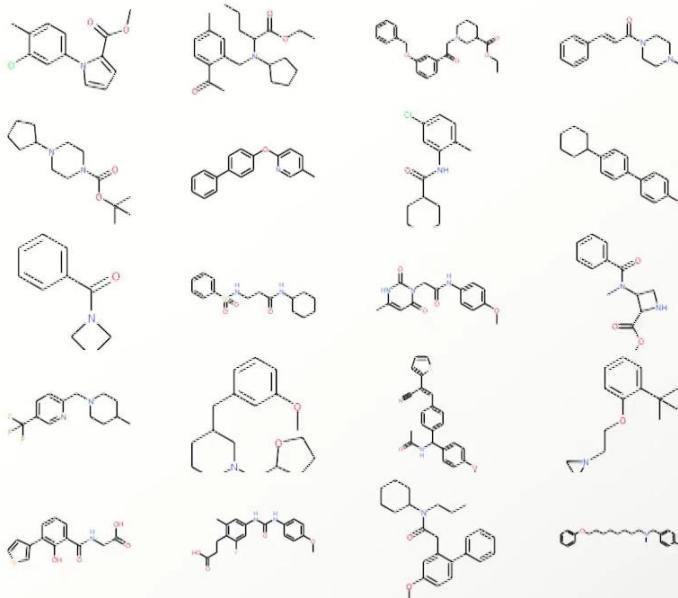
show selected (1) 

show selected (0) 

show selected (1) 

show selected (0) 

Tailor-made compounds



EVALUATE THOUSANDS

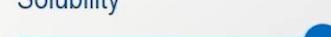
Target Product Profile

Target Activity

Anti-Target Effects

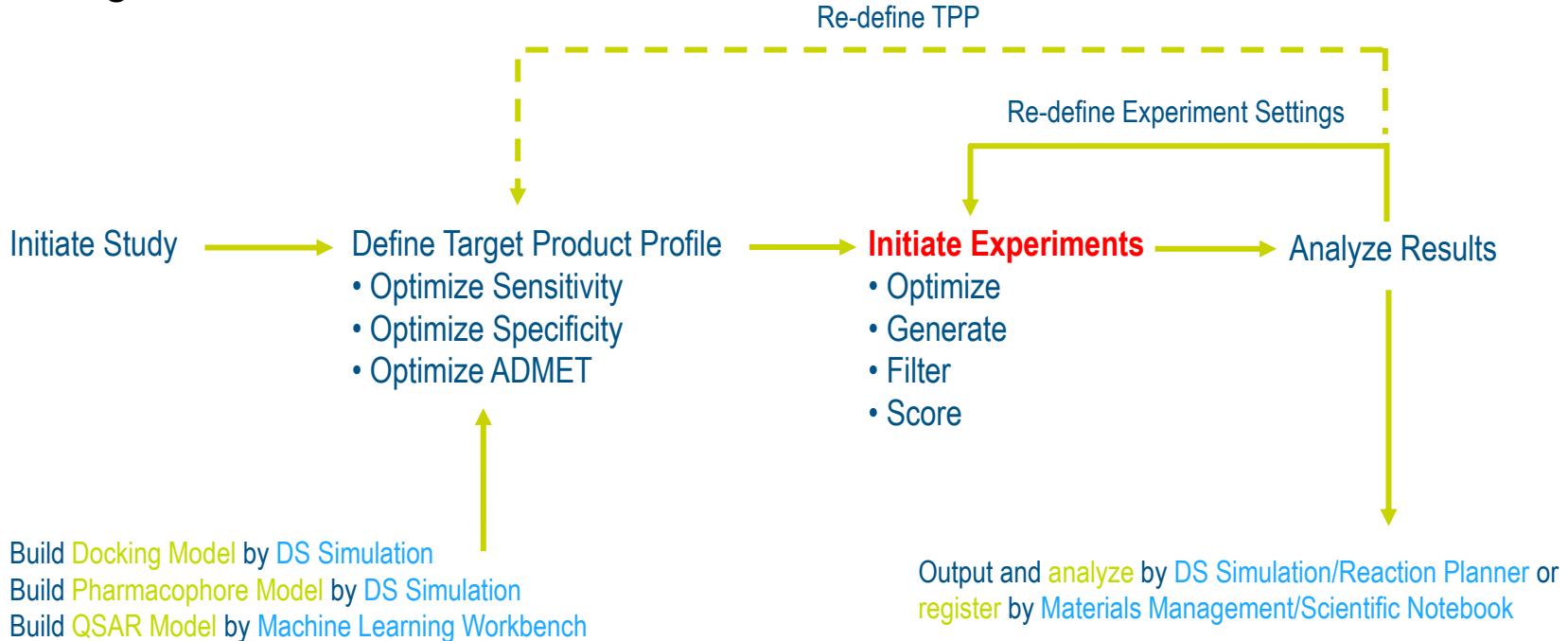
Solubility

Toxicity



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

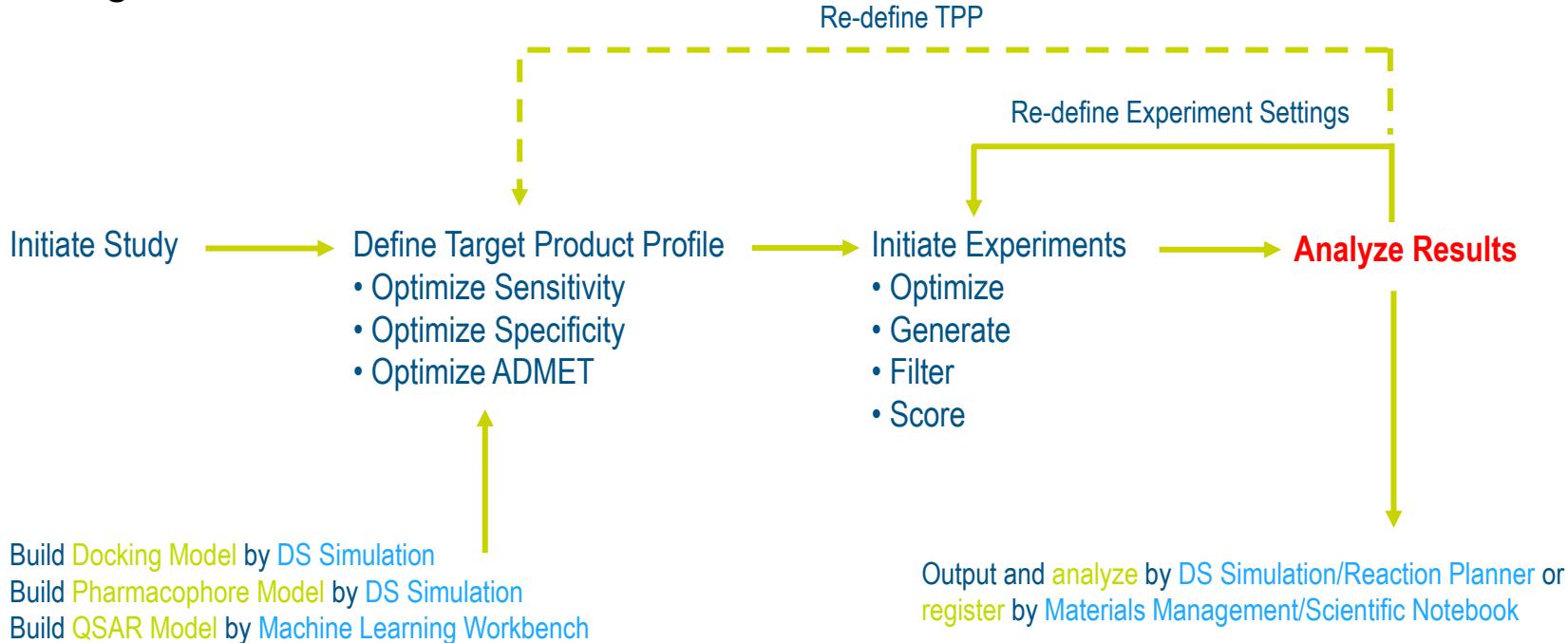
Based on 4 models in TPP: Cross-Validation Prediction Distribution, Positive/Negative Category Distribution.



Tailor-made compounds

Generative Therapeutics Design (GTD)

Design Workflow



Analysis GTD Results In All Iterations

Title	Optimization Properties						Hepatotoxicity	ERbeta_2NV7_Nove
	Overall Desirability	Estrogen Receptor Alpha	Aqueous Solubility	Synthetic Accessibility Score	Hepatotoxicity			
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_Negative

Overall Desirability

Estrogen_receptor_alpha_Bayes

Molecule 10

Molecule 17

Molecule 8

Molecule 23

Molecule 11, Molecule 6, Molecule 12

Molecule 50

Molecule 15, Molecule 44

Molecule 19

Molecule 55

Molecule 33

Molecule 99

Molecule 24

Molecule 76

Molecule 5

Molecule 85

Molecule 64

Molecule 59

Molecule 92

Molecule 57

Molecule 28

Molecule 47

Molecule 35

Molecule 48

Molecule 99

Molecule 24

Molecule 76

Molecule 5

Molecule 85

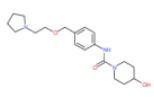
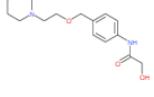
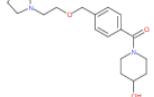
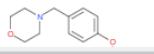
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0.92

0.75 0.8 0.85 0.9 0.95 1

Overall Desirability

Estrogen_receptor_alpha_Bayes

Plot properties and Download molecules

Title	Optimization Properties								Generator And Filter	
	Structure	Overall Desirability	Estrogen Receptor Alpha	Aqueous Solubility	Synthetic Accessibility Score	Hepatotoxicity	ERbeta_2NV7_Nova	AlogP	M	
Molecule 10		0.988	0.970	0.982	1.000	1.000	0.450	0.988	65.040	
Molecule 17		0.980	0.922	1.000	1.000	1.000	0.455	1.138	61.800	
Molecule 23		0.972	0.945	0.944	1.000	1.000	0.482	1.080	53.010	
										

Showing 1 to 30 of 100 rows. [\(Show 30 more\)](#) [\(Show all\)](#)

1 2 3 4 >

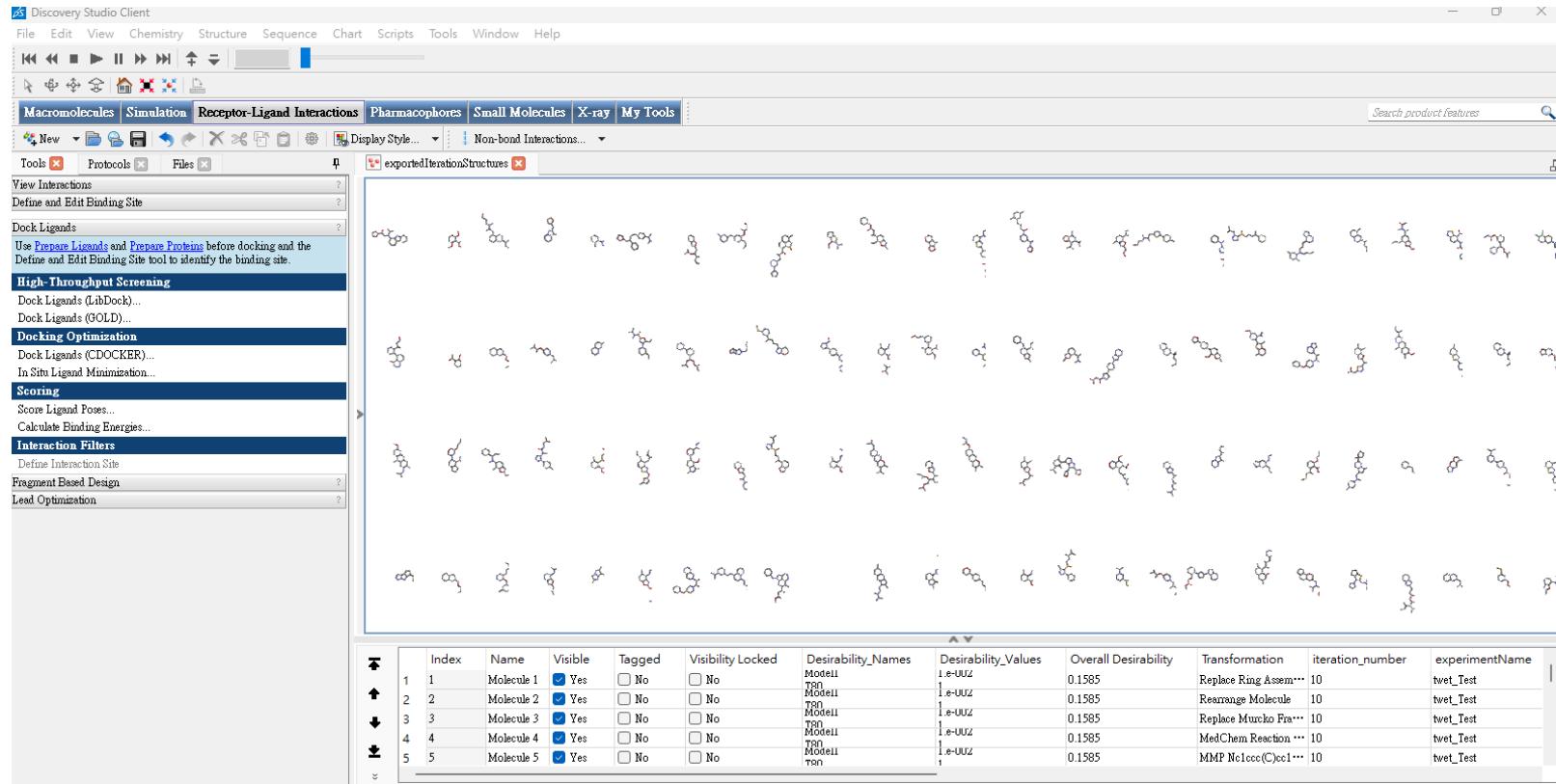
Analysis

Molecules review

Download molecules

Create an experiment

Use GTD Result as a Drug Library in DS



Scientific Insight - Visualize and interact with data

BIOVIA - Scientific Insight

Scientific Insight

Unlock the power of your data

* New

Recent Content

- ER_20251222_Test1 37 seconds ago
- Analysis 1 3 months ago
- Analysis 1 3 months ago

Import GTD Data

Import projects, studies and experiments data from Generative Therapeutics Design

Import Substances

Search and Import substance data in your analysis

Browse

Import data from a CSV, SD or SDF file

Content & Knowledge

Social Networks

User Assistance

Communities

What's New

Scientific Insight - Visualize and interact with data

The screenshot illustrates the workflow for data selection and visualization in the Scientific Insight application.

Step 1: Data Selection (100 rows selected) is highlighted in the top left.

Step 2: Data Columns Selection (No columns selected) is highlighted in the middle left.

Step 3: Dataset Preview is highlighted in the bottom left.

Chemical Search: On the right, a search interface allows pasting a structure, drawing one, or using a file. It includes a toolbar for drawing tools and a panel for drawing R1 structures.

Properties Preview: At the bottom, a grid of properties is shown, each with a brief description and a value.

Property	Description	Value				
Structure	1.7 Molecular Weight	abcMolecular Formula	abcMolecular Name	1.7 AlogP	8 Hydrogen Acceptor Count	8 Hydrogen Donor Count
8 Rule 5 Violations	1.7 Polar Surface Area	8 Rotatable Bonds	1.7 Aqueous Solubility / Prediction	1.7 Aqueous Solubility / Desirability	1.7 Estrogen receptor alpha / Prediction	1.7 Estrogen receptor alpha / Desirability
1.7 Hepatotoxicity / Prediction	1.7 Hepatotoxicity / Desirability	1.7 Synthetic Accessibility Score / Prediction	1.7 Synthetic Accessibility Score / Desirability	1.7 ERbeta_2NV7_Nove / Prediction	1.7 ERbeta_2NV7_Nove / Desirability	1.7 ERbeta_2NV7_Nove / Score
T ERbeta_2NV7_Nove / View in 3D	1.7 OverallDesirability	abcFiltersFailed	8 Iteration	T View in 3D	abcProject	abcStudy

Scientific Insight - Visualize and interact with data

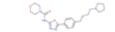
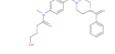
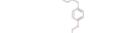
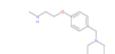
Step 1
Data Selection
100 rows selected

Step 2
Data Columns Selection
7 columns selected

Step 3
Dataset Preview

Step 3: Preview the dataset

Primary Data

Structure	Chemical Properties		Aqueous Solubility		Estrogen receptor alpha		Hepatotoxicity		ERbeta_2NV7_Nove	
	Molecular Name	Desirability	Desirability	Desirability	Desirability	Desirability	Desirability	Desirability		
	~{N}-[5-[4-(2-pyrrolidin-1-yle...	0.696	1.000	0.979	0.000	View in 3D				
	[4-[2-(3-hydroxypropoxy)ethy...	1.000	0.656	1.000	0.282	View in 3D				
	~{N}-[4-[4-benzoylpiperazin-...	0.688	0.846	1.000	0.287	View in 3D				
	~{N}-[2-hydroxy-4-[(3-(1-pipe...	0.810	0.679	0.987	0.298	View in 3D				
	~{N},~{N}-dimethyl-2-[(2-[4-(...	0.840	0.942	0.945	0.376	View in 3D				
	1-[[4-[2-(methylamino)ethoxy...	1.000	1.000	0.850	0.411	View in 3D				

Need to link with **Molecular Design**

Total: 100

Link Scientific Insight with Molecular Design

The image shows two applications side-by-side in the 3DEXPERIENCE platform:

- BIOVIA - Molecular Design:** A dashboard for molecular design. It features a main title "Molecular Design" with the subtext "Create, visualize, explore". Below this are three buttons: "New", "Open", and "Import". A "Recent Content" section displays three molecular structures with their names and creation dates: "1-[2-[(4-hydro... 2 hours ago", "3,5-dimethyl-~ 3 months ago", and "exportedliterati... 3 months ago". At the bottom, there are links for "Content & Knowledge", "User Assistance", and "What's New".
- BIOVIA - Scientific Insight - Analysis 4:** A data analysis application. It consists of three main steps:
 - Step 1: Data Selection** (100 rows selected): Shows a list of 100 rows with columns for "Structure", "Chemical Properties", "Molecular Name", and "Desirability".
 - Step 2: Data Columns Selection** (7 columns selected): Shows a list of 7 selected columns with their corresponding icons.
 - Step 3: Dataset Preview**: Shows a preview of the dataset with three chemical structures and their names: "1-[2-[(4-hydroxypropoxy)ethoxy]benzene", "[4-[(3-hydroxypropoxy)ethoxy]benzene", and "1-[2-[(4-benzoylpiperazin-1-yl)ethoxy]benzene".

On the right side of the interface, there is a vertical toolbar with the following options:

- Search in Web App
- Preferences
- Link with Web App** (highlighted with a red box)
- Share
- Duplicate
- Refresh
- Maximize
- Remove
- Help
- Send to Device

Link Scientific Insight with Molecular Design

The screenshot illustrates the integration of two 3DEXPERIENCE web applications: **BIOVIA - Molecular Design** and **BIOVIA - Scientific Insight - Analysis 4**.

BIOVIA - Molecular Design: This application is the primary interface. It features a header with tabs: **3D EXPERIENCE | 3DDashboard**, **New Dashboard**, **Search**, and **Novelyn NovelynKinanti DS - R1132101124092**. The main content area is titled **Molecular Design** with the sub-instruction **Create, visualize, explore**. It includes a sidebar with buttons for **New**, **Open**, and **Import**, and a note to **Or Drag and Drop your content anywhere**. A **Recent Content** section is displayed, with the **Link** button highlighted by a red box. The background features a molecular structure visualization.

BIOVIA - Scientific Insight - Analysis 4: This application is shown in a floating window. It is divided into three main steps: **Step 1: Data Selection** (100 rows selected), **Step 2: Data Columns Selection** (7 columns selected), and **Step 3: Dataset Preview**. The preview table includes columns for **Chemical Properties**, **Aqueous Solubility**, and **Desirability**. The first row of data is shown:

Chemical Properties	Aqueous Solubility	Desirability
<chem>[N+]1C=CC=C1</chem>	0.696	

Below the preview table, there is a section titled **Link this Web App with** with buttons for **link all** and **unlink all**. A red box highlights the **Apply** button in the bottom right corner of the window.

Link Scientific Insight with Molecular Design

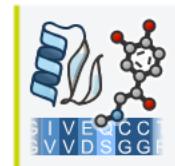
The screenshot illustrates the integration of molecular design and scientific insight within the 3DEXPERIENCE platform. On the left, the 'Molecular View' tab is active, displaying a 3D molecular structure (a protein-ligand complex) with atoms represented by spheres and bonds by lines. The interface includes a tree view of atoms (2NV7, C5, C3, C6, C2, C4, C1) and a toolbar with options like Standard, View, Edit, and Animation. On the right, the 'Scientific Insight' tab is active, showing a data analysis workflow. Step 1: Data Selection shows 100 rows selected. Step 2: Data Columns Selection shows 29 columns selected. Step 3: Dataset Preview shows a table with three rows of data, each with a 'View in 3D' button. The 'View in 3D (0)' button in the first row is highlighted with a red box. The top navigation bar shows the user 'Novelyn NovelunKinanti' and the dashboard ID 'DS - R1132101124092'.

Machine Learning Workbench

AI in Drug Discovery

Comprehensive Applications in Small and Biological Drug Design

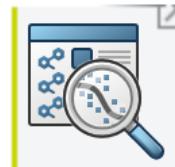
3D Modeling, Simulation and AI Prediction



Discovery Studio Simulation

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

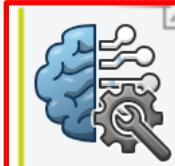
Data visualization and Analysis



Insight for Research

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

AI in Small Molecule Design



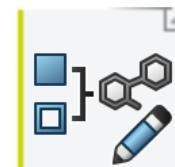
Machine Learning Workbench

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)

The screenshot shows the 3DEXPERIENCE Machine Learning Workbench interface. At the top, there is a navigation bar with the 3DEXPERIENCE logo, a search bar, and user information (Novelyn NovelynKinanti, DS - R1132101124092). Below the navigation bar, there are two main sections:

- 3DDrive:** A file browser showing a list of datasets in the 'My Files' folder. The list includes:
 - Pharmacophore (Dec 15, 2025)
 - MAPK14_Pharmacoph... (Dec 01, 2025)
 - Generative Therapeutic... (Sep 08, 2025)
 - unique_logHL_956.sdf (Sep 25, 2025)
 - unique_logHL_1225.sdf (Sep 25, 2025)
 - trial_3152.sd (Sep 25, 2025)
 - test_set.sd (Sep 26, 2025)
 - Q5.sd (Sep 19, 2025)
 - Q5_FPR2_pharma10.d... (Sep 18, 2025)
 - Q5_FPR1_pharma10.d... (Sep 18, 2025)
 - Pgp_compound_prep.sd (Sep 18, 2025)
 - my_mols_logHL_prep... (Sep 26, 2025)
- Machine Learning Workbench:** The main workspace for building machine learning models. It features a dark header with the title 'Machine Learning Workbench' and a sub-header 'Build machine learning models'. Below the header are five buttons:
 - * New
 - Import (highlighted with a red box)
 - Manage Models
 - Manage Training Procedures
 - Manage Training Tasks

On the right side of the workspace, there are two cards:

- Tabular Data:** Shows a grid icon and a description: 'Build machine learning models based on a tabular dataset in 3DDrive, Data Factory Studio, or Dataset Governance.'
- Chemistry Data:** Shows a chemical structure icon (a benzene ring fused with a cyclohexene ring) and a description: 'Build machine learning models based on a chemistry dataset in 3DDrive, Data Factory Studio, or Dataset Governance.'

At the bottom left of the workspace, there is a red box with the text 'Drag and Drop SDF here'.

Machine Learning Workbench - User-defined ML model

 BIOVIA - Machine Learning Workbench



► 1. Choose a dataset

► 2. Choose the model type, target, and features

► 3. Choose the data split

► 4. Choose learning algorithms

► 5. Name the model and the training procedure



Type *

Classification

Category *

Target

Response Property

IC50

Response Property Processing

Convert to Binary Categories

Conversion Operation

Category Boundary

10

Preferred Category Label *

LowIC

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](C)train
3 0.166701689 S(=O)([O-])=Nc1cc(C)train
4 0.20000002 Clc1cc-c2cc(NC(=O)O)train
5 5.999996541 O=C(OCC[C@H]J)N[=N+]train
6 4.017001556 FCF(F)F(F)Oc1cc(F)cc1train
7 0.383301042 Clc1c(NC(=O)NCC)cc1train
8 0.133299954 O=C(NCCCCc1cccc1)train
9 2.582997584 S(=O)([O-])=Nc1ccccc1train
10 0.766708009 O=C(NCCCCc1cccc1)train
11 0.18329896 O=C(Nc1nc(C(=O)C)C)train
12 3.300000305 O(CCC)C1c2cc(cnc1)train
13 5.899999842 O=C1NC(C(=O)H)cc2train
14 0.80000024 O=C1Oc2ccccc2C(=O)train
15 0.633300747 O=C(N(C(=O)H)C(=O)C)train
16 0.696706721 O=C(NCC2CC(C)C)train
17 0.833297349 S(=O)([O-])=Nc1cc2train
18 0.266698148 S(=O)(=O)(NC(=O)=O)Ctrain
19 16.00000064 O=[N-]JOC(CO)N+J)train
20 3.84699838 Fc1c(Oc2ncnc3cc2C)train
21 0.133299957 O=C(NCC(C)C2CC3C)train
22 1.298999548 O=C(NN=C(O)C)cc1train
23 6.200000435 [N-]H3C1C(C2CCCC2)train
24 11.00000797 S(CCC)C(=O)N(C)C)train
25 0.599998273 Pt(=O)(OP(=O)(OC)C)train
26 0.800000024 S(=O)(COCC(=O)N)Ctrain
27 7.800007136 FCF(F)F(F)Oc1cc(CN(C)C)train
28 2.160001238 Clc1c(CO)Cc2ccccc2train
29 6.99999677 O=C(N(C)C(=O)NCl)train

SMILES to SDF

XGBoost / LightGBM Regression Model
Based on Features and Fingerprints

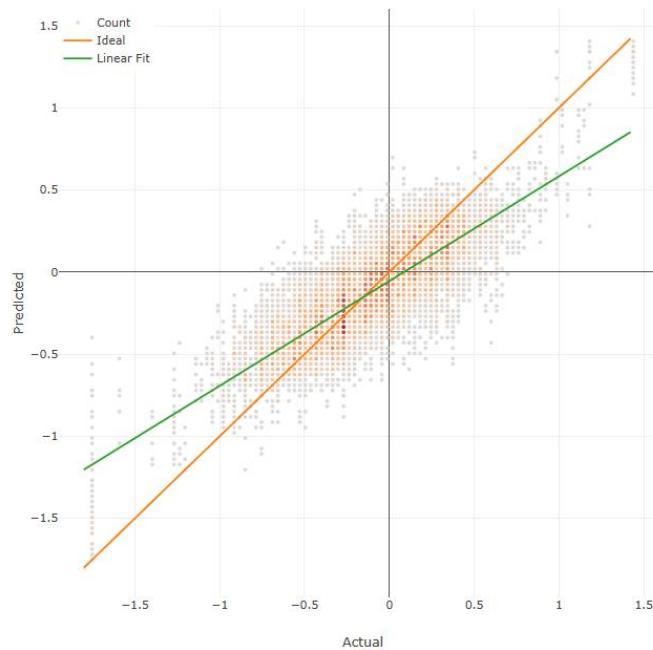
3512 ->2557 Molecules



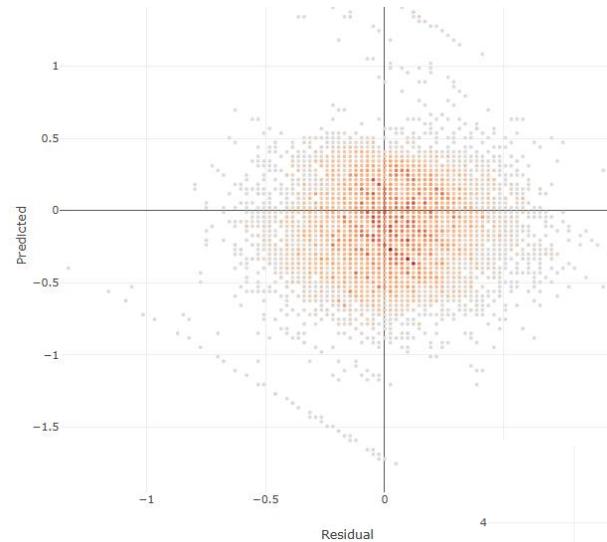
Model Report for logHL_mol2557 version 1

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

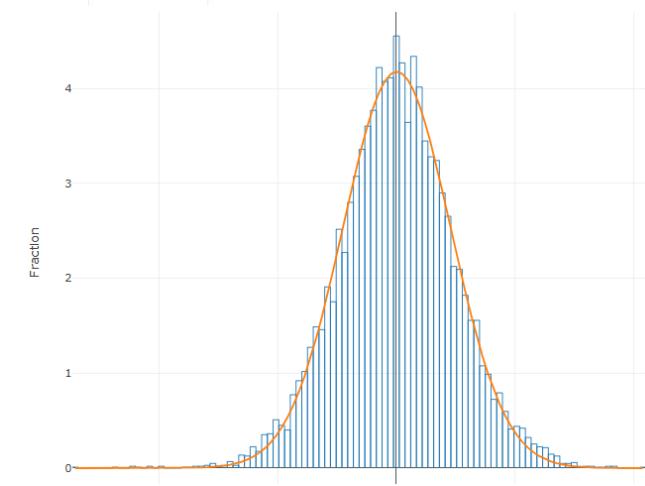
Cross-Validation Predicted Versus Actual



Cross-Validation Residual Distribution



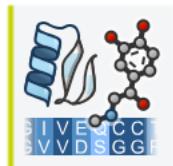
Light GBM Algorithm
R2=0.6906/0.7167



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



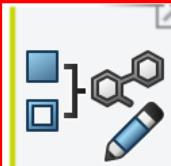
Machine Learning Workbench



Generative Therapeutics Design

Train ML models with your experimental data.

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

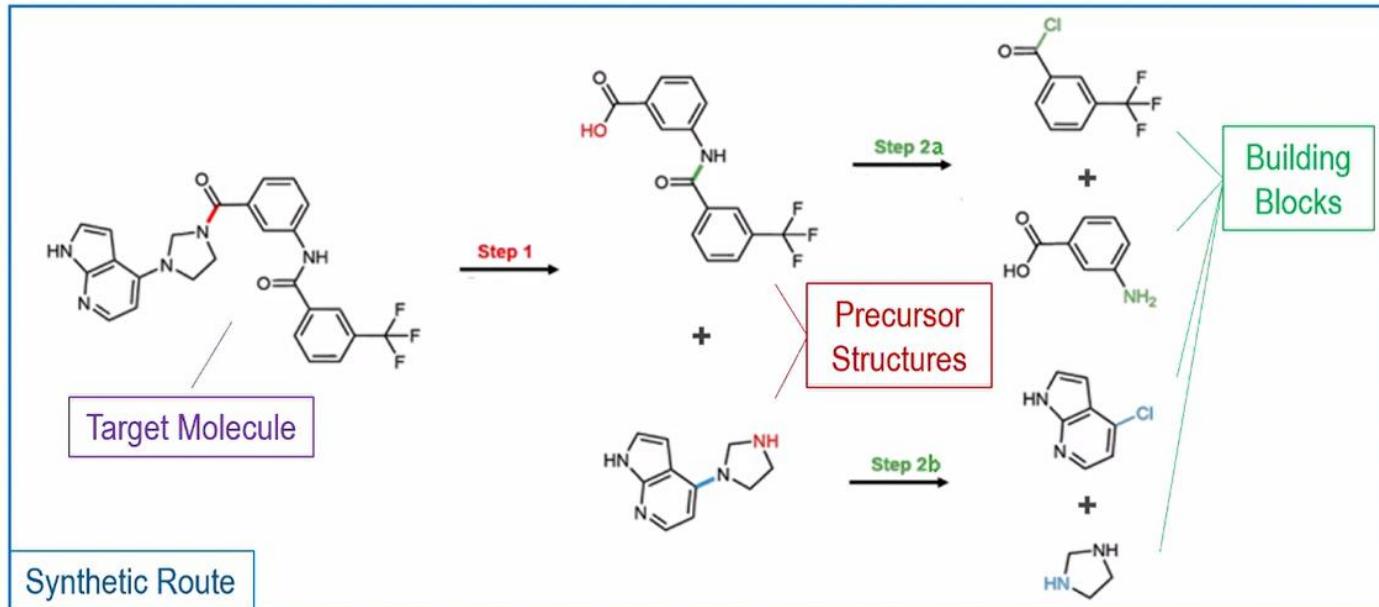


Reaction Planner

AI-based retrosynthesis prediction tool

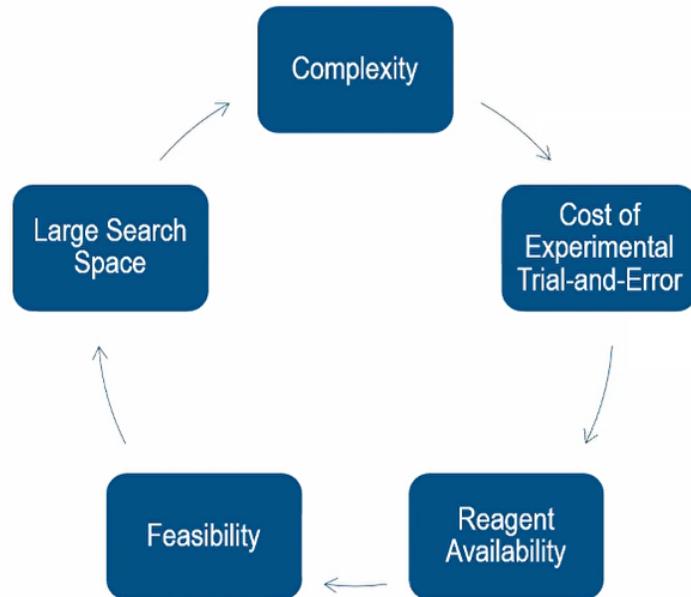
Reaction Planner

AI-based retrosynthesis prediction tool



Retrosynthetic Analysis

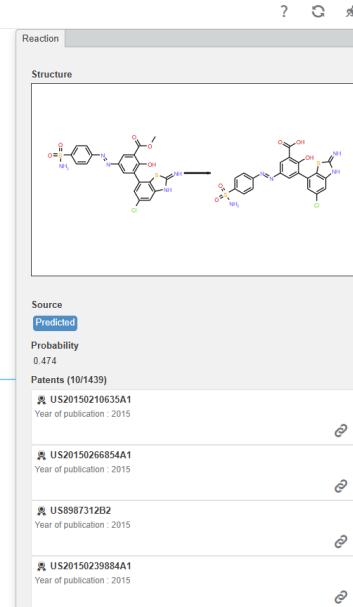
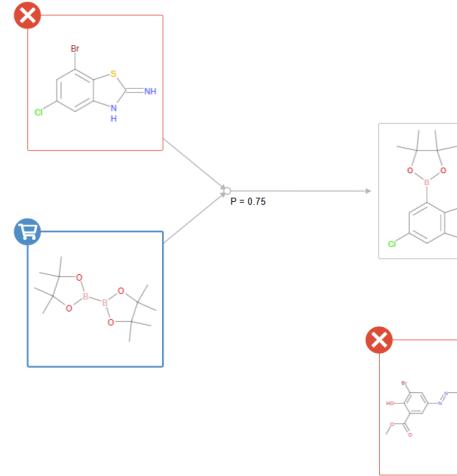
Challenges



Reaction Planner

Retrosynthesis to design synthetic plan for small molecules

BIOVIA - Reaction Planner - 3-(5-chloro-2-imino-3-[H]-1,3-benzothiazol-7-yl)-2-hydroxy-5-[{~(E)}-4-sulfamoylphenyl]azo]benzoic acid Retrosynthetic Analysis



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

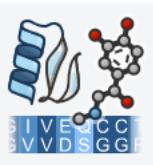
Main Capability

- 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.

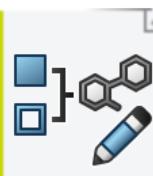
AI in Small Molecule Design



Machine Learning Workbench



Generative Therapeutics Design



Reaction Planner

Train ML models with your experimental data.

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

AI-based retrosynthesis prediction tool

Agenda

Scientific Platform Portfolio – R&D to Manufacturing

AI in Drug Discovery

Q & A



For more information please contact...



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

