evotec

#RESEARCHNEVERSTOPS

Leveraging Evotec AI/ML-DMTA engine for project acceleration

Evotec Drug Discovery Chemistry

Disclaimer

This presentation (including any information which has been or may be supplied in writing or orally in connection herewith or in connection with any further inquiries) is being delivered on behalf of Evotec SE (the "Company", "we," "our" or "us"). This presentation is made pursuant to Section 5(d) and/or Rule 163B of the Securities Act of 1933, as amended, and is intended solely for investors that are qualified institutional buyers or certain institutional accredited investors solely for the purposes of familiarizing such investors with the Company. This presentation shall not constitute an offer to sell or the solicitation of an offer to buy Evotec securities, nor shall there be any sale of these securities in any state or jurisdiction in which such offer, solicitation or sale would be unlawful prior to registration or qualification under the securities laws of any such state or jurisdiction. No representations or warranties, express or implied, are made as to the accuracy or completeness of the statements, estimates, projections or assumptions contained in the presentation, and neither the Company nor any of its directors, officers, employees, affiliates, agents, advisors or representatives shall have any liability relating thereto.

Cautionary Note Regarding Forward-Looking Statements

This presentation contains forward-looking statements concerning our business, operations and financial performance and condition, as well as our plans, objectives and expectations for our business operations and financial performance and condition. Many of the forward-looking statements contained in this presentation can be identified by the use of forward-looking words such as "anticipate," "believe," "could," "estimate," "expect," "intend," "may," "might," "plan," "potential," "should," "target," "would" and other similar expressions that are predictions of or indicate future events and future trends, although not all forward-looking statements contain these identifying words. Forward-looking statements are based on our management's beliefs and assumptions and on information currently available to our management. Such statements are subject to risks and uncertainties, and actual results may differ materially from those expressed or implied in the forward-looking statements due to a variety of factors. The forward-looking statements contained in this presentation speak only as of the date of this presentation, and unless otherwise required by law, we do not undertake any obligation to update them in light of new information or future developments or to release publicly any revisions to these statements in order to reflect later events or circumstances or to reflect the occurrence of unanticipated events.

Drug Discovery Chemistry community

Rich pool of talent at all levels

Drug Discovery Chemistry Leadership Team

- Wealth of drug discovery experience and insight applied to all projects and collaborations

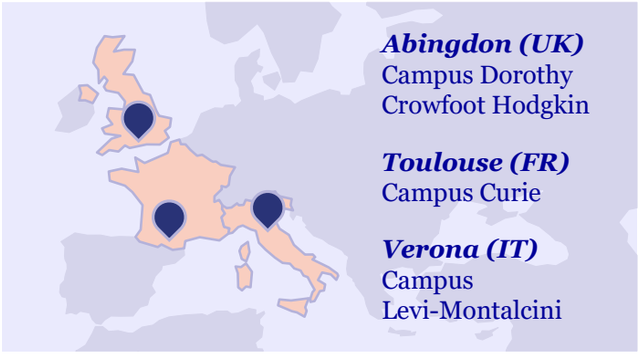
Project Leaders

- Experienced cohort of project leaders (typically 10-15 years' experience)
- Single point of contact with client
- Talented scientists with a proven track record of delivery (quality chemical assets, publications, patents, keynote presentations etc.)

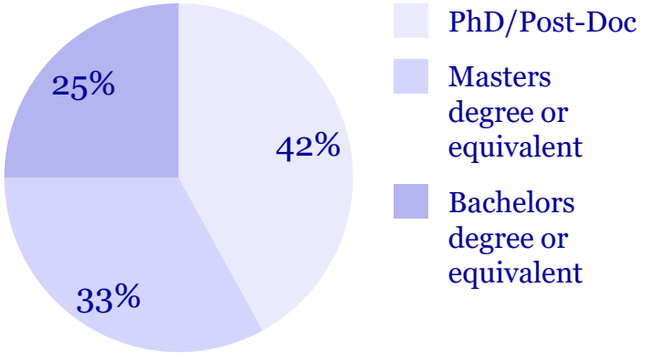
Chemistry Resource

- Global capacity of 500 chemists encompassing medicinal, analytical and computational chemistry
- High quality design and synthetic capabilities (~350 lab chemists; 40% female: 60% male)
- High proportion (>42% lab-based and ~50% overall) of PhDs
- Knowledge and experience in AI/ML generative and computational design and cheminformatics

Integrated Drug Discovery sites



Diploma level of lab chemists



Drug Discovery Chemistry at Evotec: the right organizational synergies

Our key ingredients for a successful Drug Discovery Chemistry collaboration

Technology differentiators

Industry-leading suite of AI/ML and advanced computational tools

- Development of cutting-edge in silico predictive, generative and analytical tools
  - Multi objective AI/ML-based design workflow
    - Emphasis on design precision and smooth integration with lab expertise
      - Predictive safety (Transcriptomics, Cell Painting, …)

Our seasoned drug hunters "The human catalyst"

Our people – passion, expertise, problem solving

- Strong disease understanding and target family expertise
  - Rapid response time & decision-making
    - MedChem extensive experience with strong track record in CD nominations and inventorship contribution

Process fundamentals

Flexibly accessible chemistry platform

- Rapid cycle times, e.g., DMTA
  - High quality delivery
    - Multiple modality, efficient synthetic execution platforms

DMTA: Design, Make, Test, Analyze CD: Candidate Drug

A track record of success: selected Discovery & Development KPIs

"Evotec Inside" our partners' pipeline contributions

>16

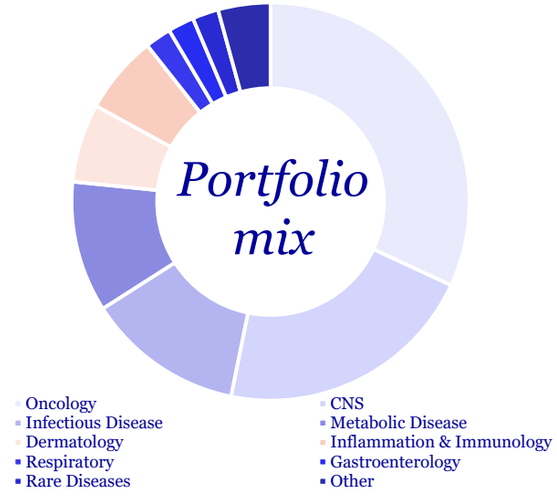
PDCs nominated since 2020

28

IND packages supported since 2020

>5

Manufactured clinical agents and commercial APIs since 2020



>100

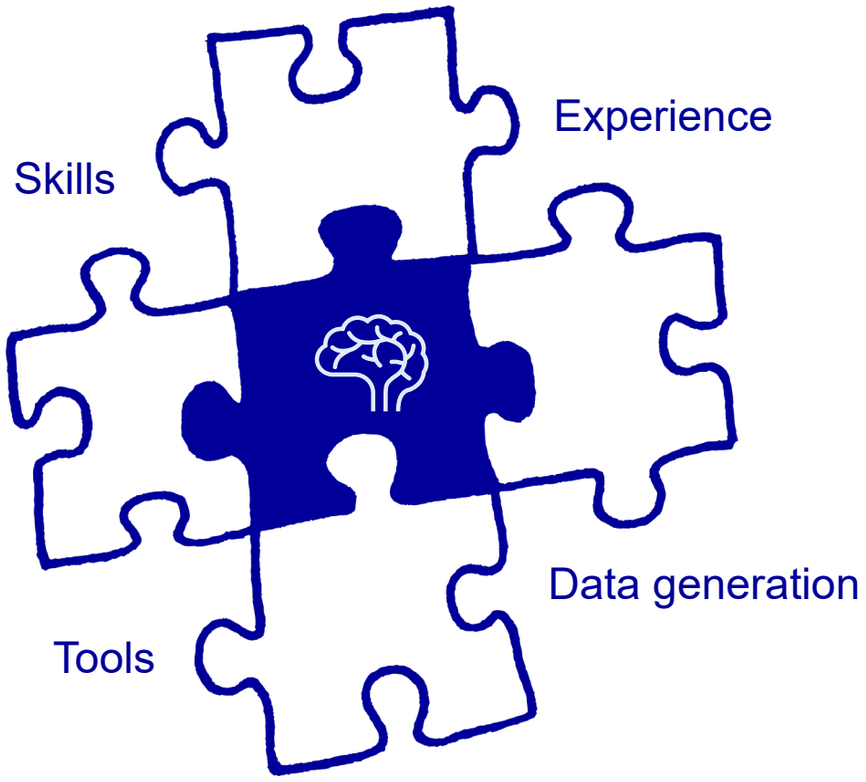
Patents with Evotec inventors published since 2020

>250

Research articles published since 2020

The secret sauce for Excellence in Molecular Design at Evotec

Combining skills, tools, experience, experimental data and integration/communication



Molecular Architects

CompChem & MedChem

CompChem expertise and access to state-of-the-art in silico tools as well as highly sophisticated AI/ML tools developed by in silico R&D

MedChem extensive experience and expertise with strong track record of CD identification

In silico R&D

CompChem & Software Development

Develop cutting-edge tools and methods

Laboratory data generation

An "all-under-one-roof" molecular design solution to accelerate the path to CDs

isR&D in a Nutshell

Integrating data science & AI

Structured

ODA

Data Mesh

IDD

Unstructured

Evotec ELNs

Scientific publications

Patent filings

Clinical trial notes

NLP

Knowledge Graph

GPT3 Web

Neo4J

Disease Models

Services

Bayesian Optimization

EvoFold

Generative Design

EvoML …

Compute

DC

Training

User Base

Web

Client

Data | Information | Knowledge | Impact

[1] Highly interdependent with DDS and IVB

Evotec's "Design EVOlution" philosophy

Our core design and make principles to enable access to tomorrow's medicines

Gap Analysis

How do compound properties differ from the target profile? How can this be improved?

Define the Pharmacophore

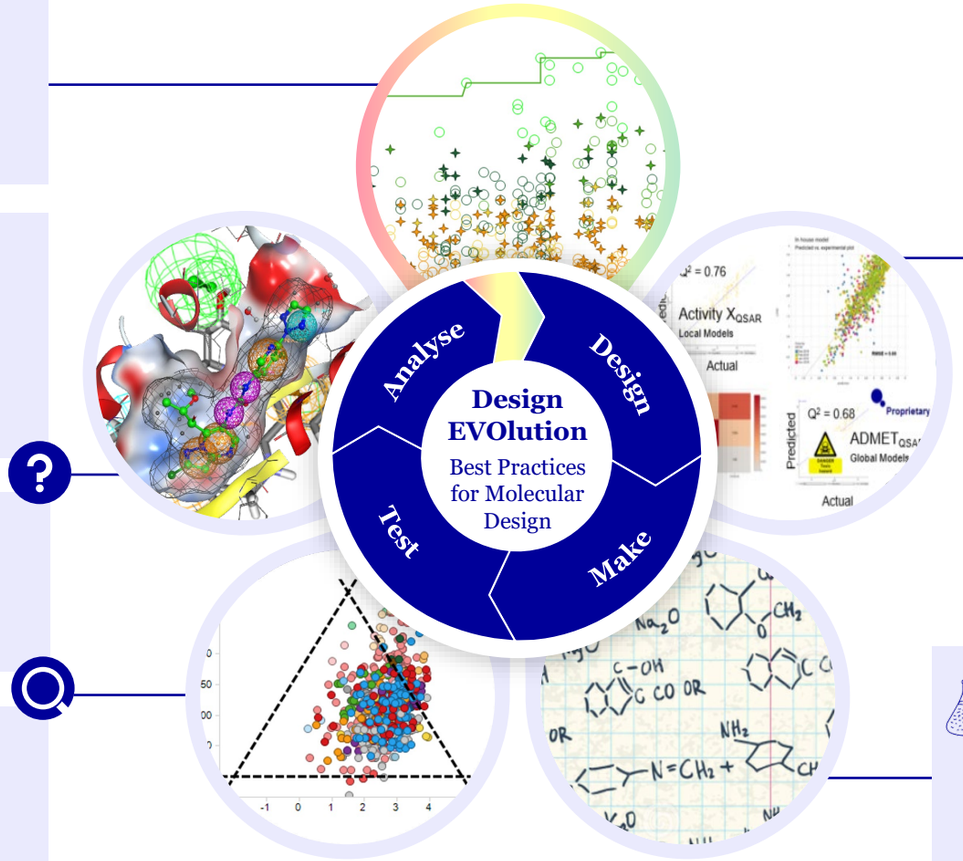
Which molecular features are driving compound potency?

Molecular conformation

Understand impact on biological & property profile

Focus on properties

Align with desirable property space.
Make the right compounds first time.



Design Efficiency

Rapid iterative design. Matched pairs. Maximum information from fewest compounds.

In Silico Methods

DMPK models. Local ML models. Prediction-driven, generative design and active learning (AI/ML/DL).

Make-Synthesis

Right first time. Good chemical route planning & execution. Minimize needs for re-synthesis. Apply best technologies.

Industry-leading suite of AI/ML and advanced computational tools

Experts in application of E.INVENT-AI design toolkit and Design EVOlution philosophy

AI Structure Generation
Molecular autoencoders, SLERP latent space exploration

Generative Design
Reinforcement learning based generative design

Predictive Models
Streamlined ML model training & application – global and local models for virtual selection

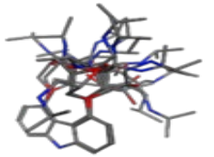
Transforms
Molecular optimization using coded expert medchem transformations

Bayesian Optimization
Design for model construction and optimization

3D Protein Structure prediction and exploitation
CryoEM, MD, FMO, MMGBSA, FEP determination & exploitation of compound-protein interactions

AI Assisted Design
AI-guided virtual library generation using known reaction data

Matched Molecular Pairs
Prediction of properties using statistical historical data



4D NMR conformation & QM conformation prediction
Exploitation of conformational design for activity and disposition

Generative AI, evaluative prediction tools, experimental techniques and drug-hunting expertise creates current state-of-the-art design

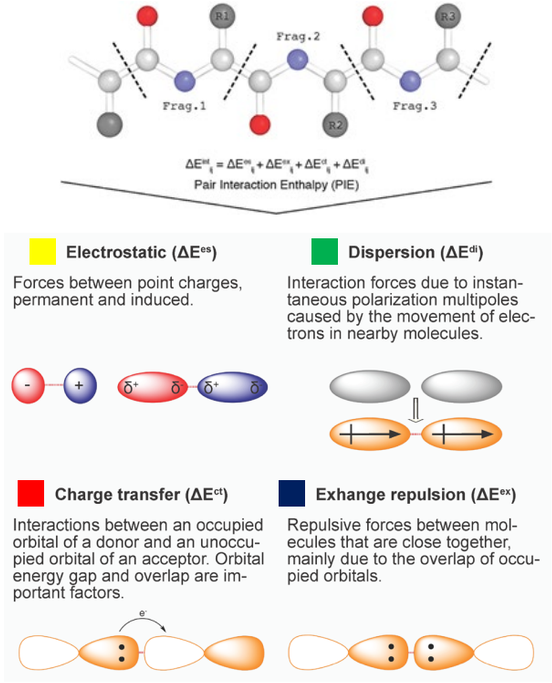
QM, AIML and High Performance Computing (HPC)

Cutting-edge technologies for excellence in molecular design

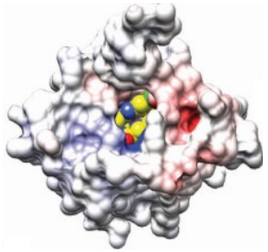
QM together with **AI** and **HPC** are the cornerstones and present/ future of modern drug design

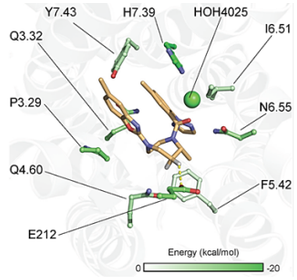
High Throughput Fragment Molecular Orbital (HT-FMO) is a unique selling point of Evotec[1]

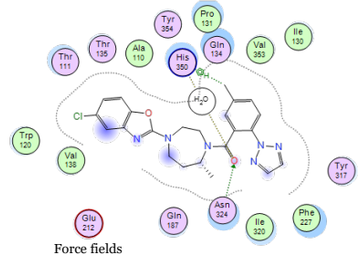
- "Visual inspection" and molecular mechanics do not fully explain complex protein-ligand interactions
- FMO is a QM application for drug design used at Evotec in almost every SBDD client project for the last 8 years[2]
- Protein-ligand or protein- protein interactions in kcal/mol and their chemical nature (electrostatic or hydrophobic)

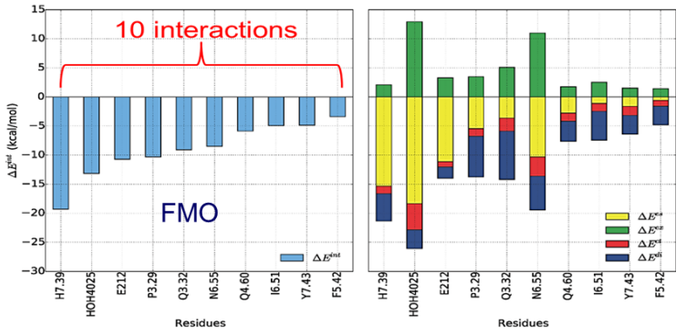


Suvorexant::hOX2R Ki = 0.35 nM)[3]





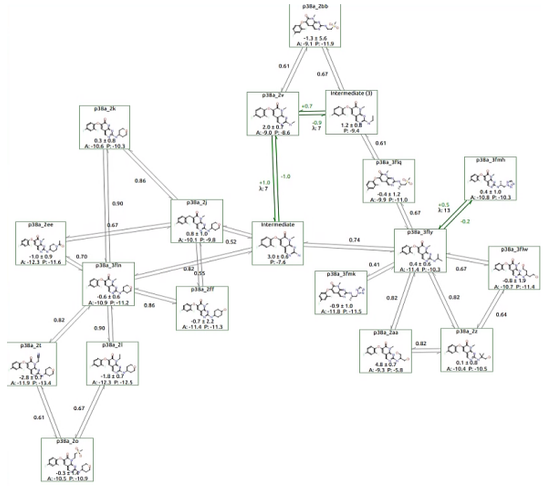


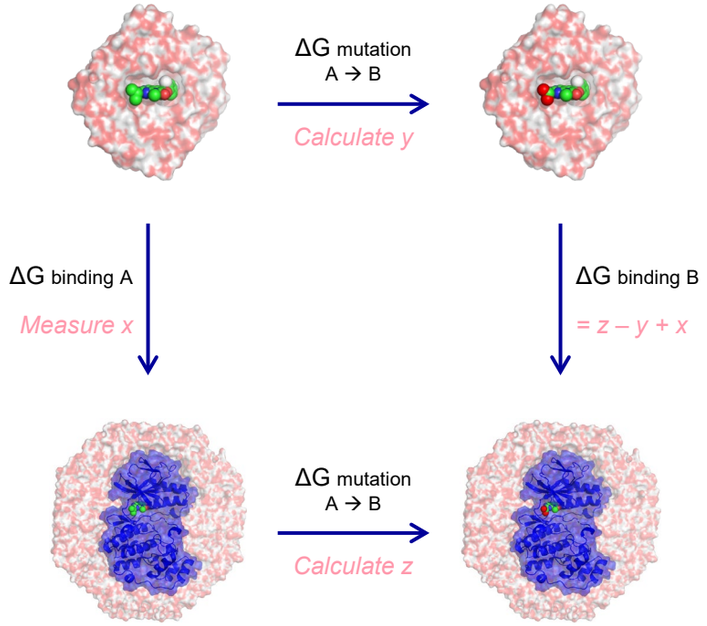


1 Kitaura K. *et coll.*, (1999). "Fragment molecular orbital method: an approximate computational method for large molecules". **Chem. Phys. Lett. 313 (3–4): 701–706**
2 Heifetz A, *et al.* (2016). "Accurate calculation of the absolute free energy of binding for drug molecules". **J. Chem. Inf. Model. Jan 25;56(1):159-72**
3 Yin J, *et al.* (2015). "Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant" **Nature 519: 247-250**

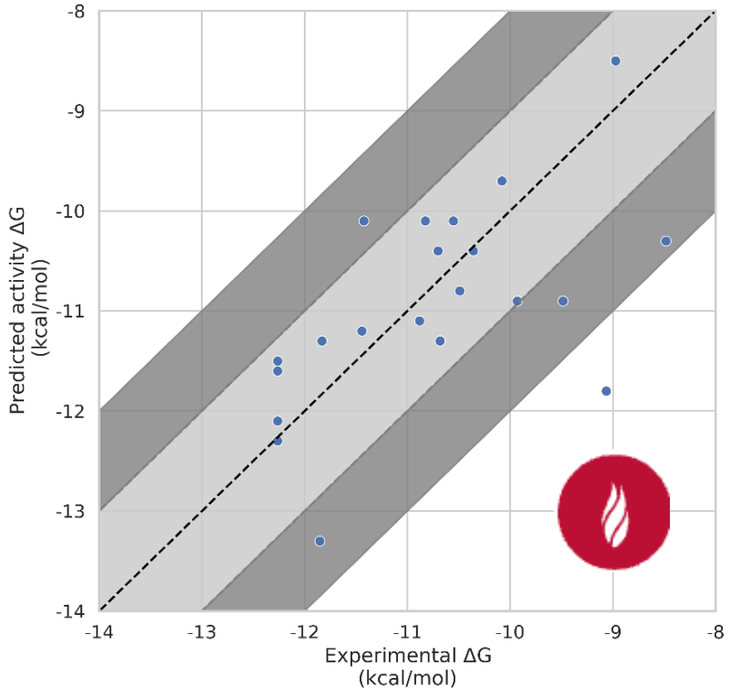
Cutting-edge technologies and High Computing Power for excellence in Molecular Design – FEP

- FEP uses molecular dynamics (MD) simulations to calculate the energy of mutating one ligand into another
- Many of these simulations are carried out at Evotec's Frankfurt Data Center to calculate the binding affinity of AI generated designs





Data generated for p38 kinase ligands using Flare (Cresset) FEP software



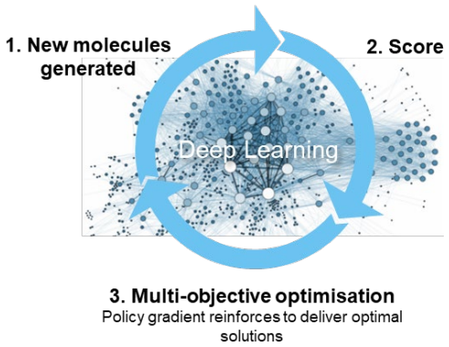
Accurate prediction of potency for final prioritization

Small Molecule Generative Design using AI/ML

Examples of Deep Learning, Transformers, Reinforcement Learning, Autoencoders, Bayesian Optimization

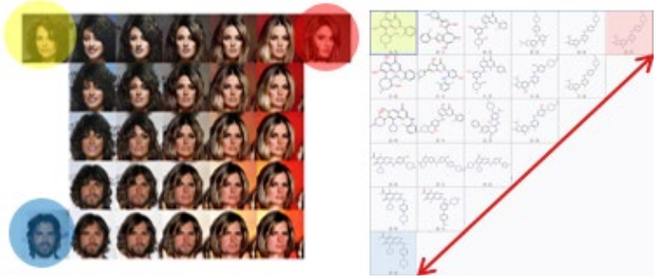
eINVENT

Reinforcement learning of a SMILES generator



SLERP

Spherical interpolation in the Hypercube



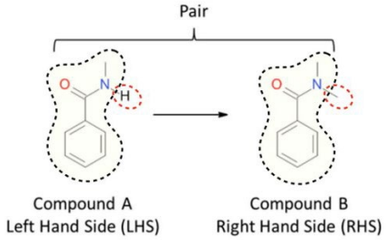
Pocket2Mol

3D exploration of pockets

Including fragment growing and linking

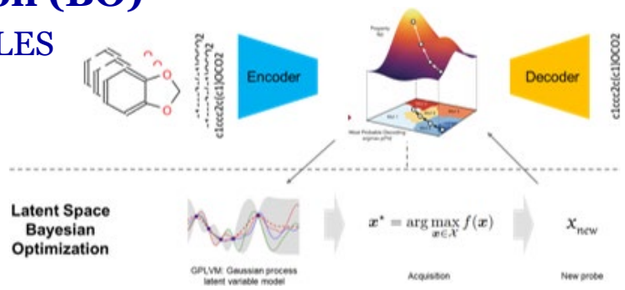
Asphodel

Ultra large library generator based on molecular pairs



Generative Bayesian Optimization (BO)

BO-based SMILES generation



in-house tools built on top of the mmpdb3.1 library predicting property changes and generating new molecular structures

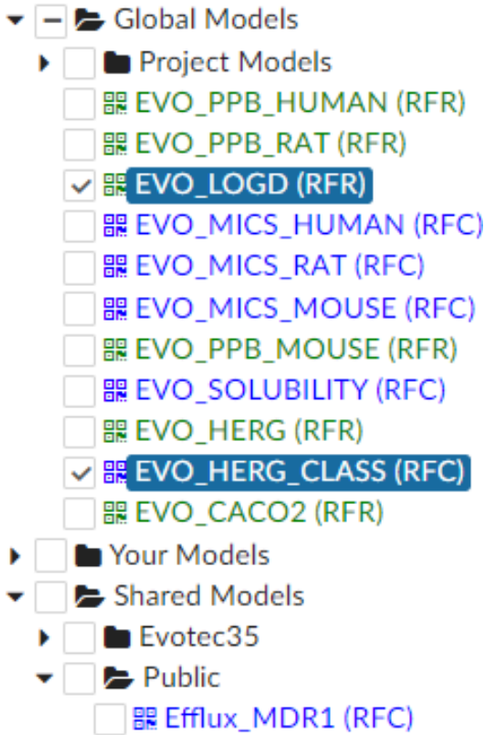
A variety of AI/ML algorithms accessible for small molecule generative design

J. Chem. Inf. Model. 2020, 60, 12, 5918–5922
Chem. Sci., 2019,10, 1692-1701
arXiv:2205.07249 [cs.LG]

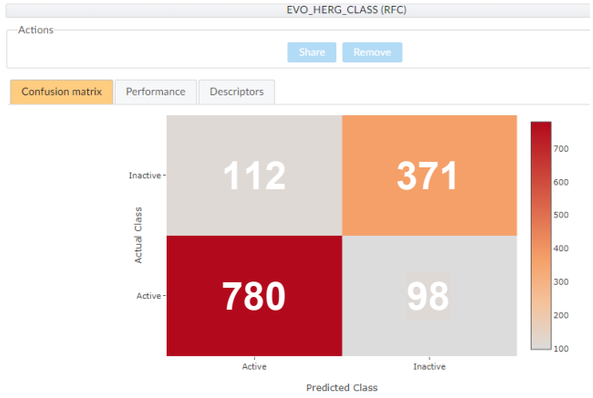
DMPK, safety & developability ML models

Mitigating compound attrition with smart prediction

Global Models



hERG



Model Performance (5-fold stratified cross validation)

Overall

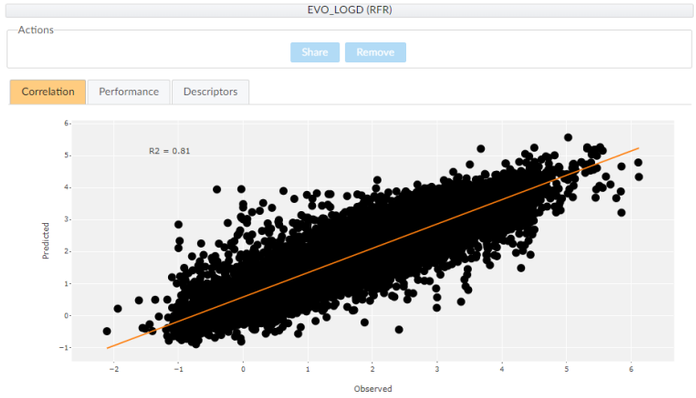
Accuracy: 0.85

Cohen's kappa: 0.66

Per class

	Active	Inactive
Recall	0.89	0.77
Precision	0.87	0.79
F1 Score	0.88	0.78

LogD



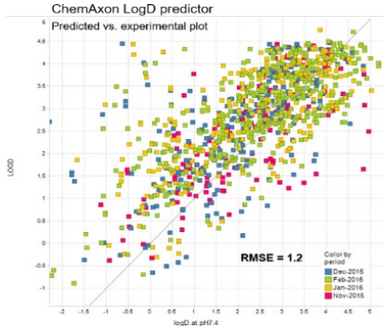
Model Performance (5-fold stratified cross validation)

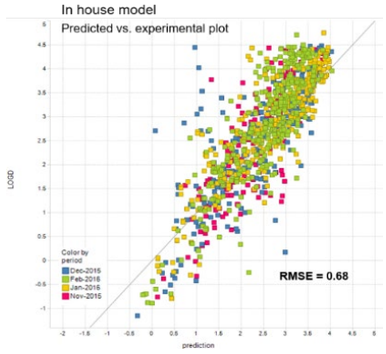
Overall

Q^2: 0.81

MAE: 0.37

RMSE: 0.26





Predictive in-silico DMPK Machine Learning models that allow projects to progress faster through more effective compound design

Track record of developing predictive models with demonstrated impact

Support to generative design and design selection steps

Public / Partner / Evotec data

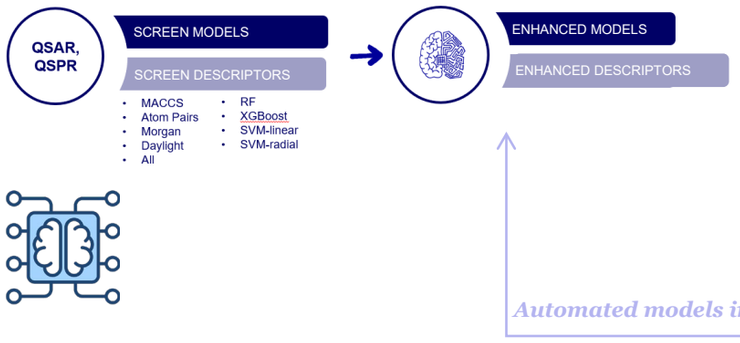
Target	Total compounds (Annotations)	Total datapoints
Target 1	n	n
Target 2[1]	n	n
Target 3[1]	n	n
Target 4[1]	n	n

Data preparation

Project data curation & standardization
Molecules preparation

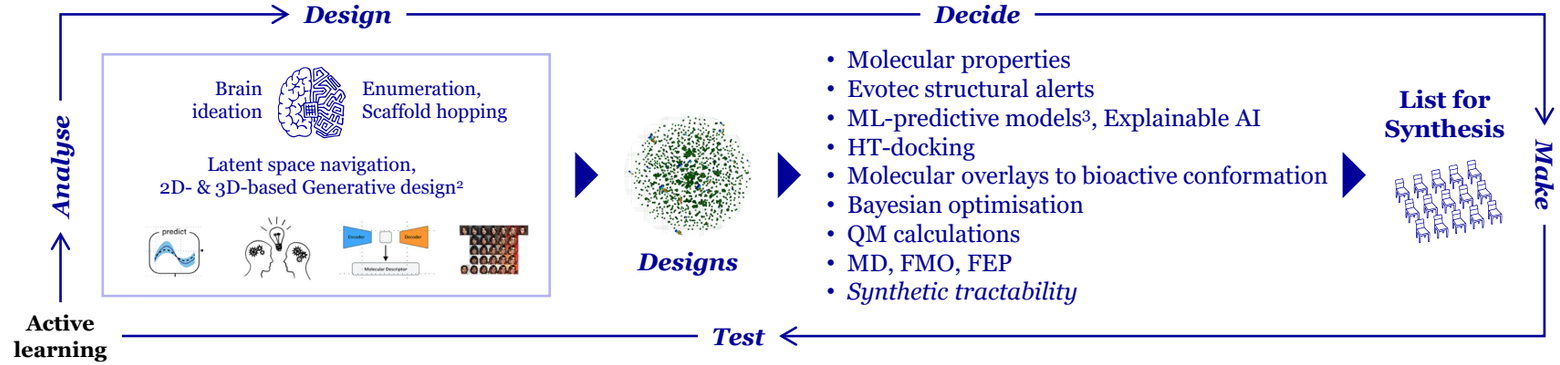
Development of predictive models

Local model development



Pipelined design (illustration)

Workflow from human/machine ideation up to compound selection for synthesis

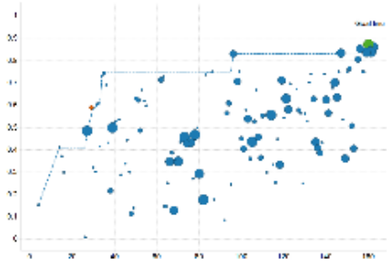


1- Polypharmacology or off targets
2- eINVENT, SLERP, Generative Bayesian, Pocket2Mol, etc.
3- Activity, selectivity, ADME, safety, etc. data

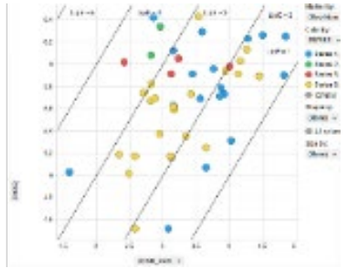
Informatics platforms to optimize DMTA speed and quality

Tools and processes for data integration and visualisation to aid design & synthesis decision making

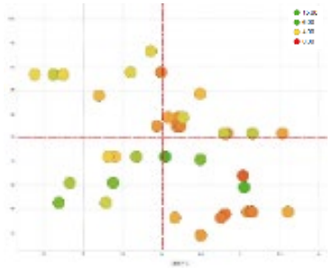
MPO[1] Telemetry



LiPE

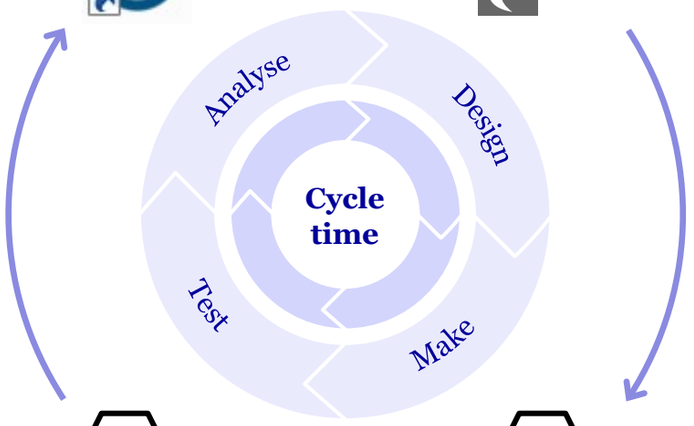


AEI[2]



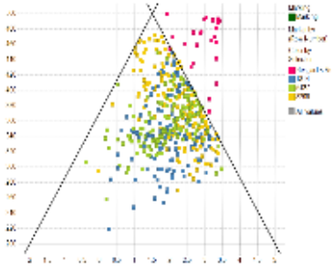
Graphical analysis tools

Central data repository of target ideas

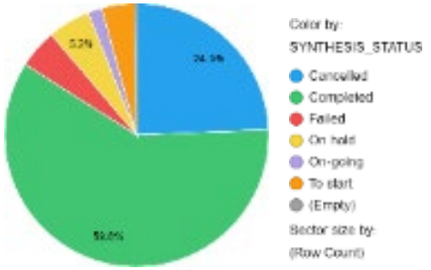


Project data warehouse rehouse

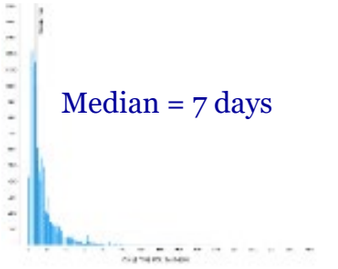
Central data repository of synthesis targets



Golden Triangle Plot



Synthesis Outcome



Synthesis Cycle time

1 Multi-parameter optimisation
2 ADME Efficiency Index (https://doi.org/10.1080/17460441.2019.1660642)

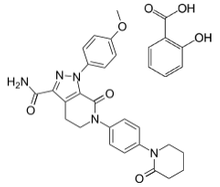
We solve complex problems in small molecule discovery

Multi-modalities and synthesis efficiency

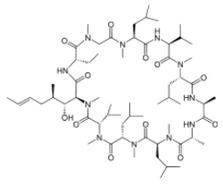
Synthetic chemistry must not be the limiting step for the preparation of therapeutically relevant molecules

Strong and Diverse Expertise

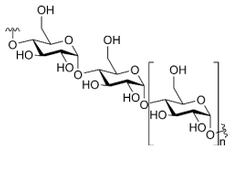
Small molecules / NP



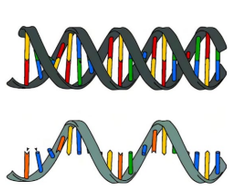
Peptides



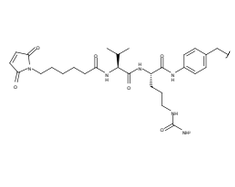
Carbohydrates



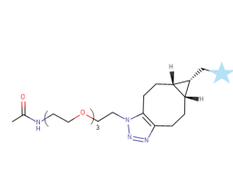
Oligonucleotides



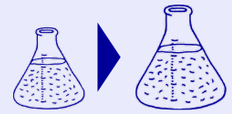
Linker Chemistry



Bio-orthogonal Chemistry



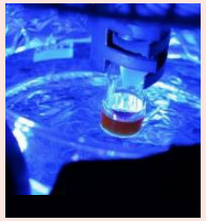
Scale-Up is … something more than just increase the reaction size



State of the art Technologies/ Strategies

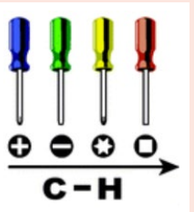
Peptide Synthesis

Photo-chemistry



Flow Chemistry

Late stage Func-tionalization

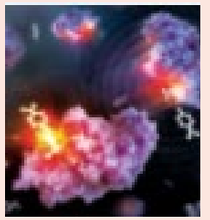


Parallel Synthesis

Oligonucleotides Synthesis

Electro-chemistry

Enzymatic catalysis



Capsules strategy

In addition: *Green chemistry, special Gas handling (H_2, CO, HCl, NH_3, etc.) capabilities, Scavengers & Resins, Mechano-chemistry, Glovebox facilities OEB5 and more …*

NP: Natural Products

Evotec's "every compound is a CanDidate" testing paradigm

Focus on parallel approach to de-risking and control of cycle times

Evotec's parallel de-risking approach converts uncertainty-to-knowledge to aid direct navigation to high-value outcomes: e.g., delivery of a pre-clinical development candidate (CD) suitable for IND-enabling studies

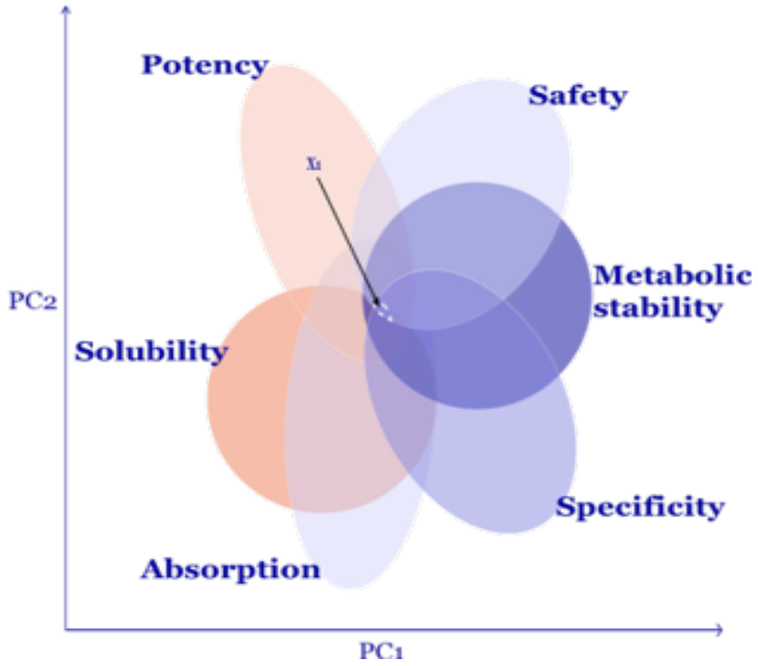
Critical path activities

Qualitative TPP		Risk assessment Chemistry		Risk assessment Biology	
Primary Assay / Hit Finding			Target Validation		
Design – Synthesis					
Binding & Function	Species Homologues	Stability	Phys Chem *in vitro* ADME	Permeability	Preliminary Safety
◆					
Tx cell biology and cytotox	Rodent PK	Early Hu. PK prediction	CYP Induction and DDI	Secondary Pharmacology[1]	Tox_CV and DILI
◆ *Resynthesis*					
Acute PK / PD	Dose Range	Dog PK	Human Prediction	Solid State	AMES/ Other tox
◆ *Resynthesis?*					
Disease Model Efficacy	Rat Safety	Dose Linearity/ formulation	Risk assessment – Process Chem	Other studies?	
IND-enabling studies					

◆ Active decision Project dependent

Cycle Times

Synthesis 7 d	
5-10 d	
5-10 d	
7-14 d	
<28 d	



Delineating multi-tiered readouts to maximise data-led predictions within each DMTA cycle

1 Outsourced off-target panel screens
Cascade can be supported by *in silico* global and/or local models

Powerful combination of AIML and efficient DMTA

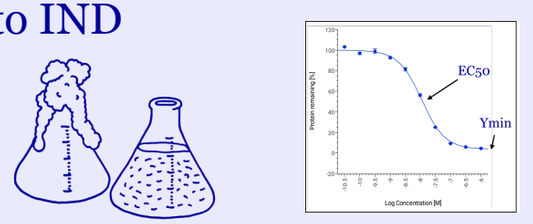
Reducing timelines to candidate nomination

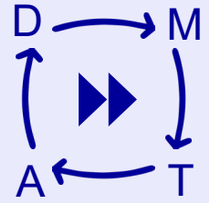
Advanced data curation and data-driven quality design

- Careful selection, cleaning and organisation of data for predictive modelling
- Data analysis and interpretation for project enablement and hypothesis generation
- Generative AI/ML & advanced computational design combined with drug hunting expertise

High-Speed Synthesis and efficient DMTA

- High speed synthesis (median TAT of 7 days) supported by access to state-of-the-art synthetic technologies
- Rapid DMTA cycles are enabled by full integration of Molecular Architects, Chemistry, DMPK and Biology
- Therapeutic area and development expertise, enables accelerated progression from LO to PDC and to IND





From target to advanced leads in less than 18 months

Case study 1: integrated small molecule drug discovery

Starting point

Data mining
Database built

- Competitive landscape assessment
- Pharmacophore elucidation (FMO)
- ML model on target activity
- Docking models & MD simulation

→ Transition to Hit identification

Hit identification – 4 months

Drug hunter knowledge-based design

AI/ML generative design

21K ideas

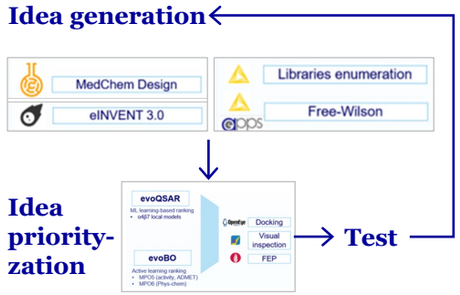
- Activity prediction
- Physchem properties
- Docking and IP assessment
- ~ 45 ideas

*2 **novel** and **selective** hit series identified*
+ 1 novel non-selective macrocycle

→ Transition to H2L

Hit to Lead – 3 months

- ✓ Refined Machine Learning models
- ✓ Efficient DMTA & decision-making
- ✓ Active learning models



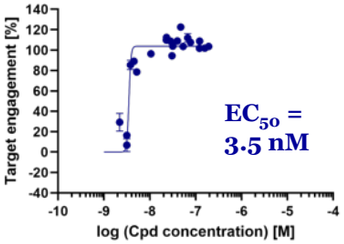
Lead showing similar potency to ref. clinical mAb in PK-PD model

→ 3 series transitioned to LO

Lead Optimization –8 months

- ✓ Continue leveraging the AI/ML platform
- ✓ Exceptional Evotec capabilities and expertise to improve drug properties

100% Receptor Occupancy for at least 8h with EC50 = 3.5 nM



→ ***Advanced leads in < 18 months total***

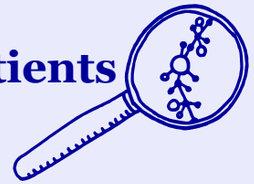
→ ***Adv. dose to man prediction < 90 mg BID***

Optimal combination of drug-hunting knowledge and cutting-edge AI/ML tools led to accelerated identification of Advanced Leads

MD: Molecular Dynamics; H2L: Hit to Lead; LO: Lead Optimization; IP: Intellectual Property; mAb: Monoclonal Antibody: DMTA: Design, Test, Make, Analyse; PK: Pharmacokinetics; PD: Pharmacodynamics

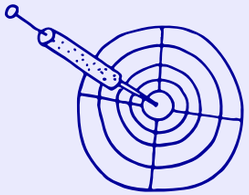
Exploring opportunities to work together to accelerate your ambition

Your organization aims to accelerate drug discovery in order to provide patients with drugs which matter?



Evotec leverages data and *in silico* physics & AIML-based tools to accelerate projects in a cost-effective manner

A unique opportunity for your organization and Evotec to partner towards your projects' acceleration

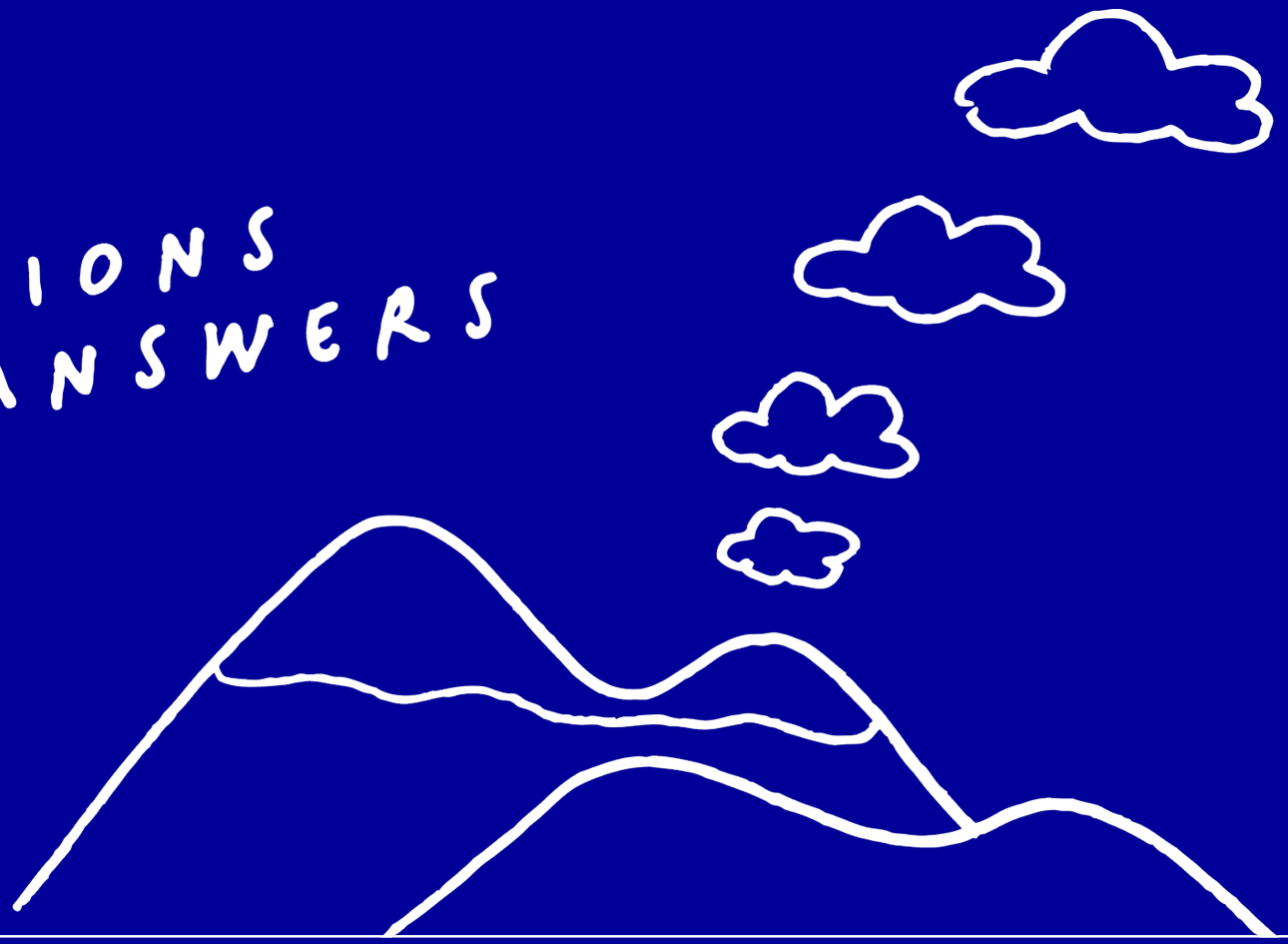


A powerful partnership to deliver medicines that matter to patients

evotec

#RESEARCHNEVERSTOPS

QUESTIONS AND ANSWERS



#RESEARCHNEVERSTOPS

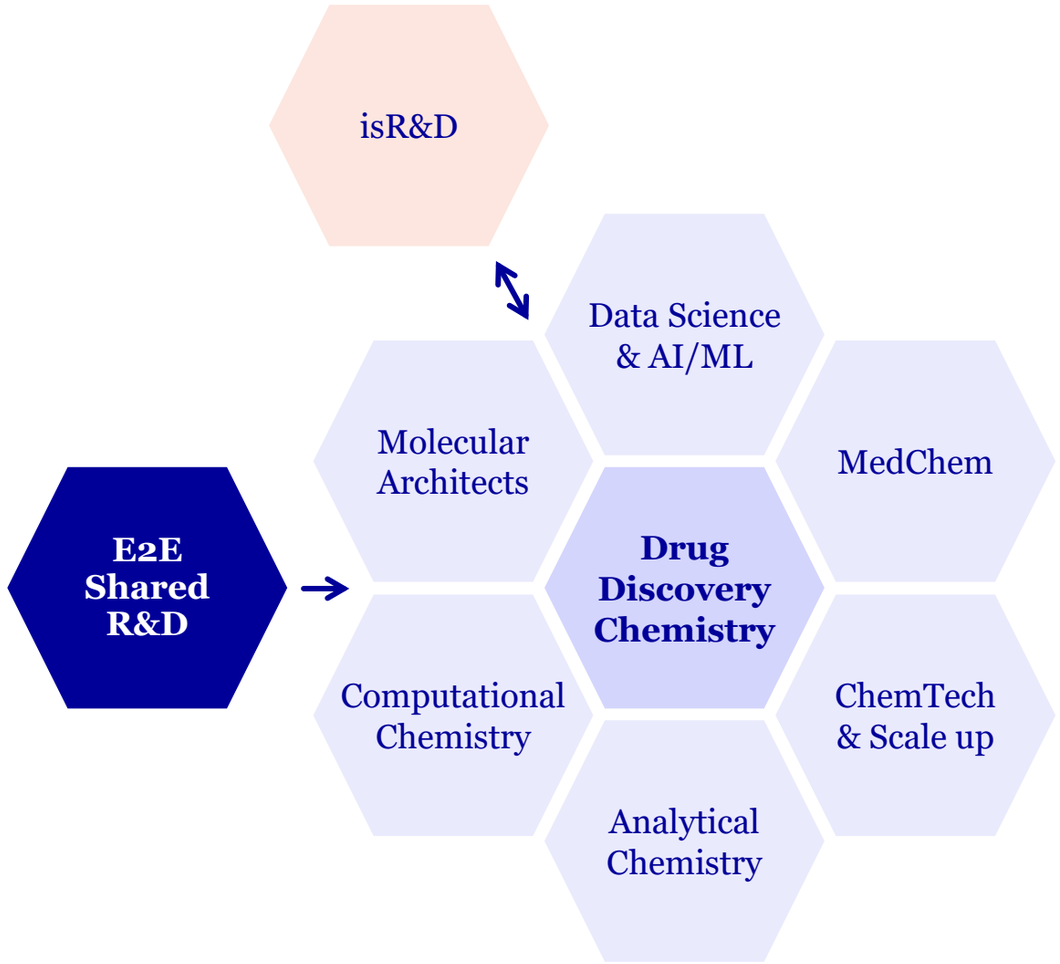
Your contact:

Christophe Boldron
Vice President, Head of Molecular Architects, Toulouse

Christophe.Boldron@evotec.com

Drug Discovery Chemistry building blocks

A key component of Evotec's E2E Shared R & D continuum



Key features

- Computationally-enabled **Molecular Architects**; prediction-driven drug hunters deployed to navigate the expanding data landscape, utilising & democratising *in silico* (AI/ML) tools for rapid, enhanced design decision making
- Goal-focused **Medicinal Chemists**; emphasis on design precision and synthesis efficiency (as part of rapid DMTA[1] approach)
- Investment in **chemistry technologies** and innovations to accelerate new discoveries and enable difficult-to-access NME[2] synthesis
- Dedicated **Discovery scale-up** teams providing "just-in-time" key intermediates and pilot compound batches and facilitating seamless exchange with API Chemistry
- State-of-the-art **Analytical Chemistry**; supporting compound analysis, purification, characterization and specialist drug discovery applications

- Development of cutting-edge *in silico* predictive, generative and analytical tools in **isR&D[3]**

1 DMTA = Design-Make-Test-Analyze
2 NME = New Molecular Entity
3 *in silico* Research and Development unit

Contemporary Requirements and Toolboxes for Drug Hunters

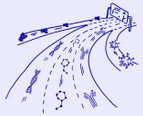
Contemporary requirements for drug discovery

- Clear clinical line of sight
- Target validation supporting the therapeutic hypothesis
- Translational strategy
- Data-rich processes – Encode all the science relevant to the project

Data mining

The Modality toolbox

- ***Small molecules*** *also including NPs and NP-based molecules, macrocycles, covalent, etc.*
- **Gene therapy** – strong focus on delivery technologies
- **RNA molecules** – ASOs or siRNAs
- **Peptides** – including chameleonic effect
- **Antibodies** – already major actor on the drug marketplace
- **Bioconjugates** including **ADCs** – fantastic technological breakthroughs



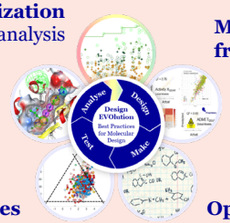
The Disease-modifying toolbox

- **Targeting proteins** – modulation of the function (ortho/allosteric, PPI)
- **Targeting complex 3D-structures of RNA** – modulate the expression of the POI
- **Degrading proteins**
- Molecular Glues or Protac
- Stabilizing non-folded protein conformations
- **Etc.**

The Augmented Design Toolbox

Multiparametric optimization
Gap analysis

Cutting-edge *in silico* tools
Continuously develop new *in silico* science



Maximum information from fewest compounds

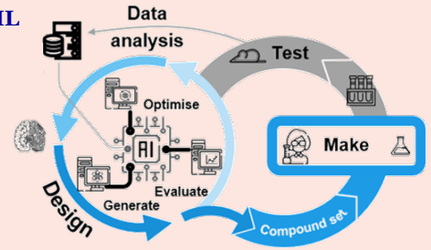
Active learning

Molecular properties

Optimal DMTA processes

Efficient navigation of the chemical space

AI/ML

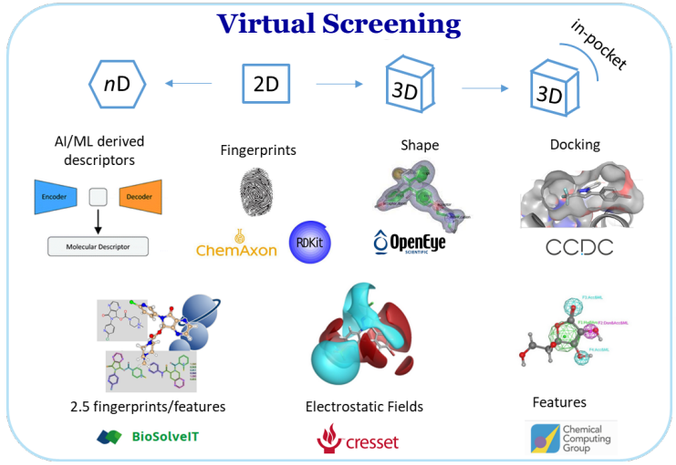


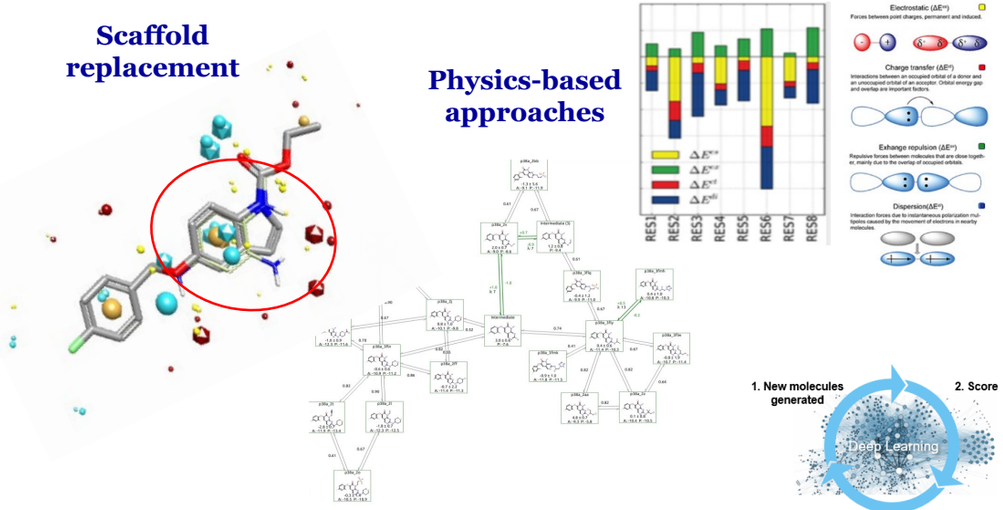
Evotec developed solutions to support all type of modalities and disease-modifying approaches

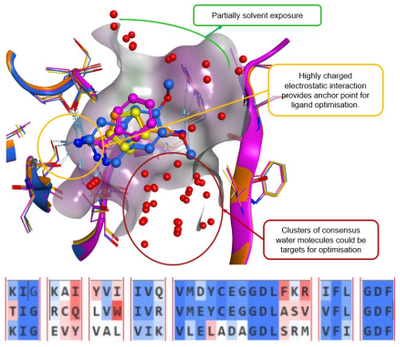
NP: Natural Products ; ASO: Antisense Oligonucleotide ; POI: Protein Of Interest

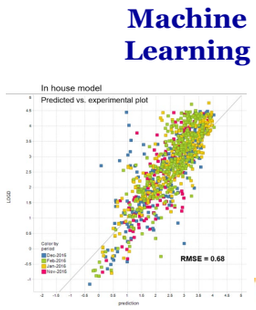
Advanced 2D- and 3D- Molecular Design at Evotec

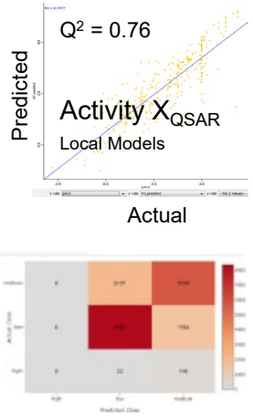
Directed by computationally-enabled Molecular Architects

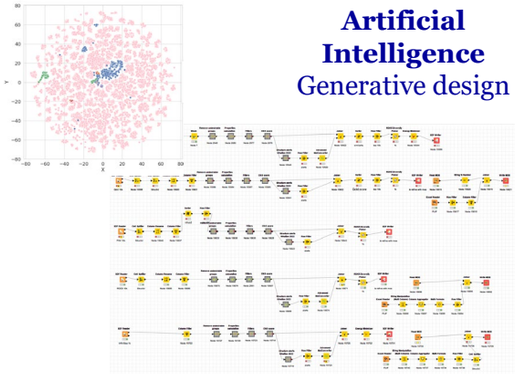


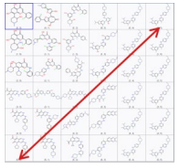












Active Learning

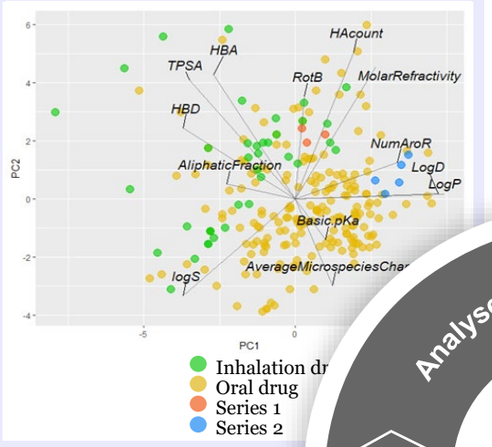
- Over 45 FTEs across four sites dedicated to delivering high quality molecular drug hunting
- Comprehensive portfolio of commercial and proprietary software tools
- Appropriate deployment of *in silico* tools across the drug discovery continuum
- Vast experience with a multitude of partners over a range of target classes and therapeutic areas
- Strong partnerships across disciplines including medicinal chemistry, DMPK and structural biology

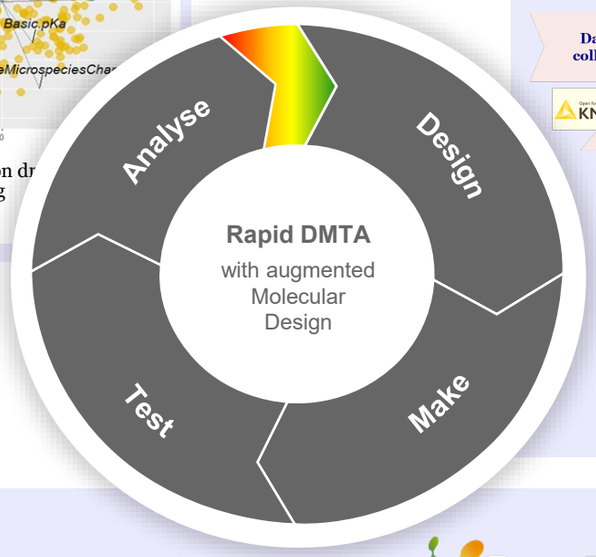
AI/ML capabilities

ADME(T) models for augmented DMTA cycles

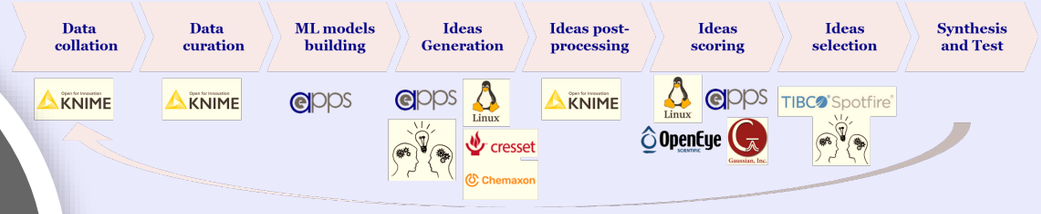
Therapeutic area map (e.g CNS-map and Inhalation-map) to assess the physicochemical property space:

- Projection of project compounds on PCA map to understand compounds potential and type of PhysChem modulation needed



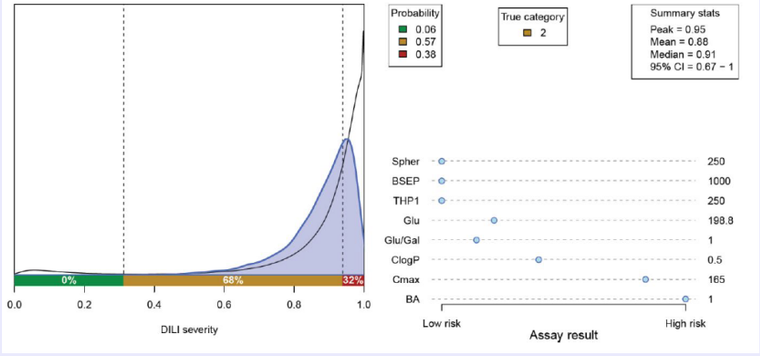


- **evoQSAR** and **evoDeep**: *in house* developed tools to build AI/ML models
  - Models can be **easily integrated** into project-specific workflows to support the design of next round of compounds



Global models using in vitro ADME data, available automatically through ideas&target DB and eapps, in house developed suite of molecular modeling tools and command line/Knime
Local models built using project-specific data

DILI risk assessment: predict risk of DILI using Bayesian Machine Learning[1]

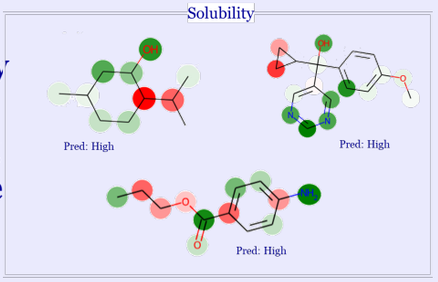


Metabolic soft-spots prediction using multiple approaches to help the design of new compounds

StarDrop

Chemical Computing Group

Explainable AI: xAI methods can potentially help rationalize deep (and machine) learning models and support the design of better molecules



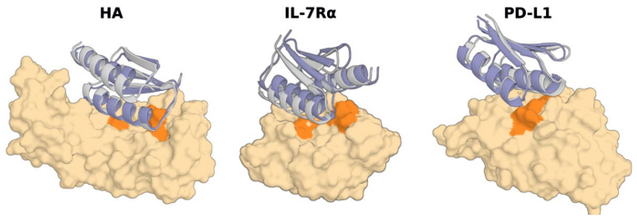
1) Predicting Drug-Induced Liver Injury with Bayesian Machine Learning, Williams D. P., Lazic S. E., Foster A. J. , Semenova E. and Morgan P., Chem. Res. Toxicol. 2020, 33, 1, 239–248

How to start a PPI project when no ligands are available?

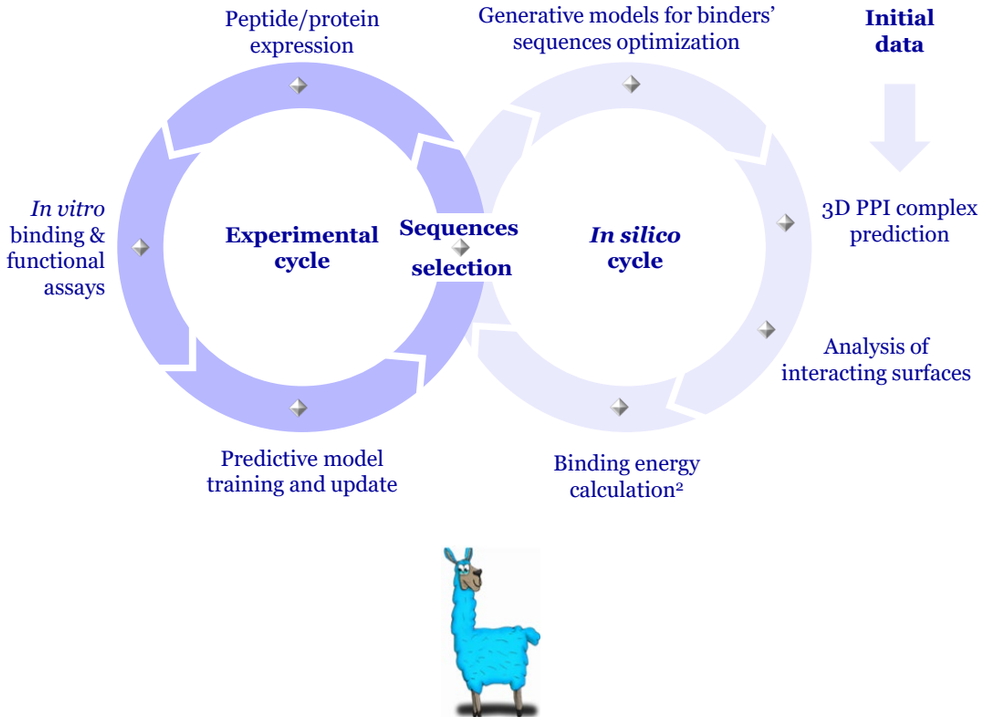
EvoLlama: design of peptides or proteins binding partners

Use of the natural PPI complex as starting point to design allosteric binders

- **3D structure prediction** of the natural protein-protein complexes using PDB structures and/or AlphaFold[1] and FoldDock[1]
- **Design peptide or protein binders** with EvoLlama platform
- Combination of deep learning, machine learning and physics based computational methods



Alliance of *in silico* and *in vitro* methods



- Robust data generation to train **ML models**
- **DL-based generative and optimization models** (EvoFold, MCTS, Bayesian Optimization)
- **Established rational design methods** (FMO[3], PLIF)
- No constraints on the peptide length
- Currently limited to natural a.a and linear sequences

Joint work & decision by highly skilled structural biologists, medicinal chemists and ML experts

1 https://www.nature.com/articles/s41467-022-28865-w
2 https://www.biorxiv.org/content/10.1101/2021.10.04.463034v2
3 Monteleone, Stefania *et al.*, Hotspot Identification and Drug Design of Protein-Protein Interaction Modulators Using the Fragment Molecular Obital Method « Journal of Chemical Information and Modeling (2022)62, 16, 3784-3799

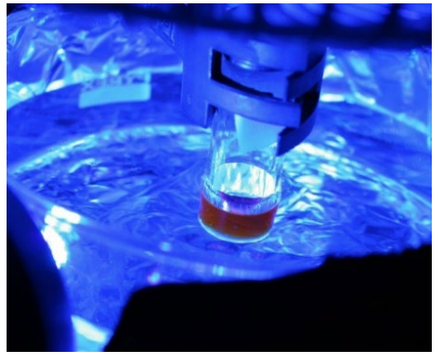
Expertise in organic synthesis

Applying the right technology to deliver the molecules that matter

Quick and easy access to the correct technology to deliver the desired compound is essential for success.
Evotec has both the expertise and technology platforms to access the most challenging of targets in a rapid fashion:

Photochemistry

- Access to new and shorter chemical routes
- In-house built equipment
- In-house specialists
- Applied to multiple projects

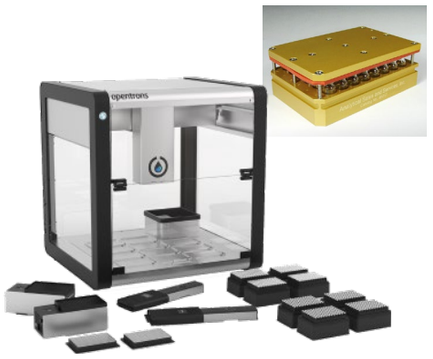


Flow Chemistry

- Links with academic groups
- Vapourtec, H-Cube Pro & In-house built equipment
- In-house specialists
- Applied to multiple projects

Parallel / Capsules / HTE Chemistry

- Wide range of equipment available
- In-house specialists
- Several operational areas providing support
- Applied to multiple projects



Electrochemistry

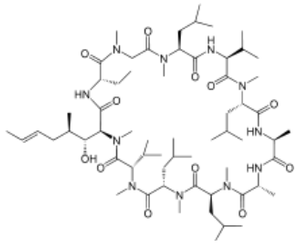
- Access to new and shorter chemical routes
- IKA Electrasyn 2.0 pro providing reproducibility & standardisation
- In-house specialists

Analytical Chemistry expertise, experience and technologies

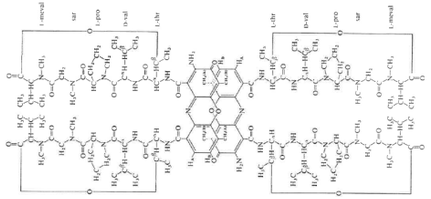
To support the End-to-End shared R&D continuum

Strong and diverse expertise combined with state-of-the-art technologies

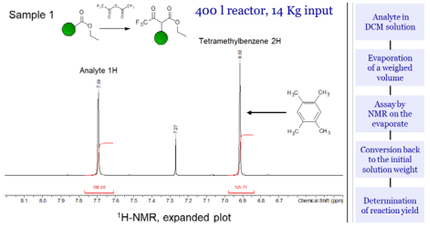
Peptides/oligos QC & purification



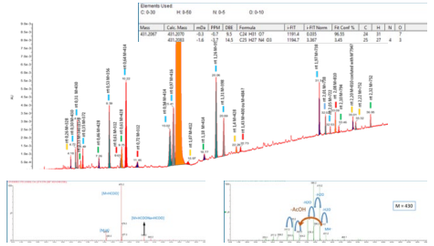
Drug self association study



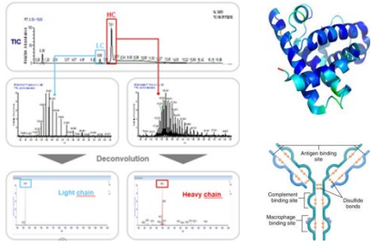
qNMR and in process control



Impurity identification



Intact mass screening

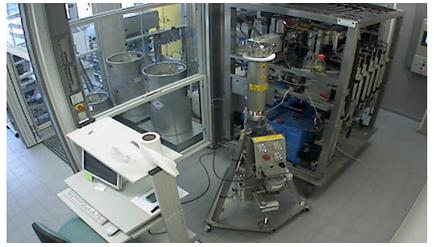


NMR up to 600MHZ

Oligos QC & purification platform

HRMS-MS

Pilot Lab: Large scale purification



Prep SFC capabilities

Capacity: 45 analysts with analyst/bench chemist ratio 1/10.
A mix of Open-Access facilities and dedicated Analytical expertise to maximize efficiency.

Discovery Chemistry Scale-up

A key part of the Discovery & Development Chemistry continuum

Hit ID & Lead Gen. | **Lead Optimisation** | **IND enabling** | **Clinical** | **Market**

“Scale-up” is something more than just increase the reaction size

- Expertise in large-scale synthesis of intermediates and chemical developability, from g to kg scale
- Assessment of synthetic route, safety, scalability and new route development
- Full range of synthetic technologies and methodologies for reliable transfer to Development Chemistry (API)

The Discovery Chemistry Scale-up group plays a strategic role during Lead Optimisation, working in close collaboration with API Development Chemistry to ensure a smooth transition into the IND-enabling phase

Discovery Chemistry scale-up groups are available at all sites to ensure a harmonised, efficient chemistry de-risking approach

Binding site finding

Identification of shallow pocket binders for undruggable targets

	Search for	Primary objective	Hit identification strategies (in silico, **experimental**)	Confirm Binding site & enable SBDD	Disease-modifying approach(es)
Hetero-bifunctional approach Identify binders of the POI	**Novel chemical series**	• **Serendipity**	**Biophysical testing (ASMS, SPR, NMR, X-Ray, DEL, …)**		
	Novel chemical series *PDBs, Alphafold, Ortholog, Homolog proteins*	Cavities/sites analysis (see next slides)	Ligand- and/or Structure-Based Virtual Screening	• **HDX** • **Protein observed NMR** • **Structural Biology**	• **Protein homeostasis** (if validated binders) • **PPI inhibition** (if validated disruptors of relevant PPIs)
	Ligands from literature related to POI or close homologs	• Chemical starting points • Support cavity(ies) validation • Chemogenomic approach	Knowledge-based approach		
	In cell reactive cysteines (from our internal databases)	Cavity(ies) analysis	• Covalent Virtual screening[1] • **Screen 5K Evotec covalent library (Intact MS and/or in cell Cysteine Protein Profiling)**		
PPI inhibition	**Novel chemical series**	• **Serendipity**	• **PPI inhibition assay (e.g. HTRF)**		
	Homo-dimer analysis	• Identify PPIs relevant to the targeted MOA/indication • FMO analysis: Identification of key PPI residues	• Virtual screening • Design peptides/peptidomimetic from interactor's epitope • De novo peptide design (EvoLlama)		
	Interactome				

If primary objective is not met

- Small molecules targeting RNA
- ASO, siRNA

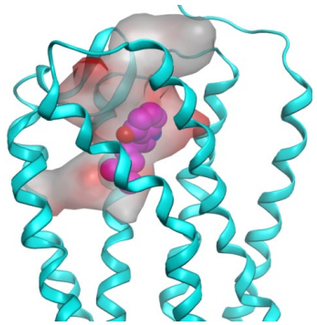
1- A toolkit for covalent docking with GOLD: from automated ligand preparation with KNIME to bound protein–ligand complexes https://doi.org/10.1093/bioadv/vbac090

Binding sites identification at Evotec

State of the Art methods and FMO

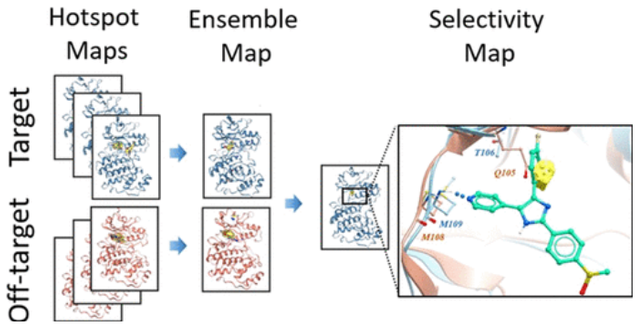
State of the Art methods

- Identification of potential active sites in targets has different applications in drug discovery:
  - **target ligandability**
  - elucidation of **protein function** and site-directed **mutagenesis** experiments
  - **virtual screening**, fragment-based **drug discovery** and **selectivity** analysis
- **Site Finder** (geometric method, CCG) and **P2rank** (ML, solvent accessible surface) are well-established tools that calculate potential active sites
- **Fragment Hotspots Mapping** (CCDC)[1] identify key interactions in ensembles of structures of the same protein (donor, acceptor, hydrophobic).



Site Finder

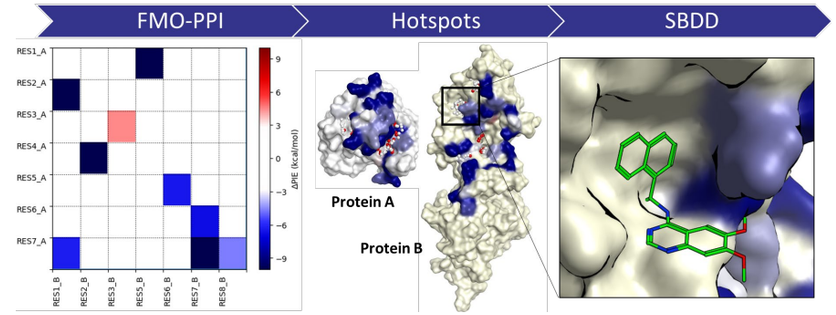
detected pocket in serotonin receptor. Agreement with the site described for psychedelic analogs.

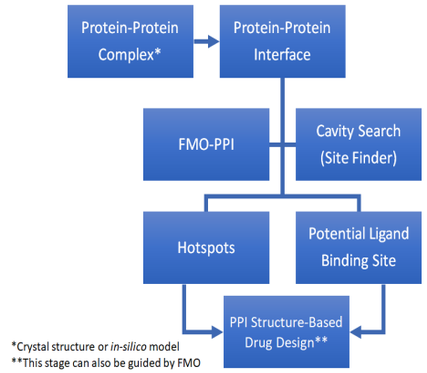


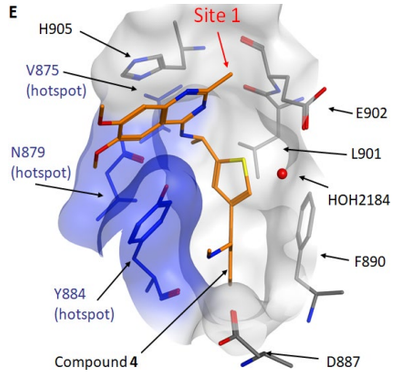
Fragment Hotspots Mapping

FMO guided SBDD for Protein-Protein Interactions

- Although **PPI interfaces** are large, a small molecule / peptidomimetic modulator / small peptide only needs to exploit **key 'hotspots' to modulate the interaction**
- We have developed a workflow deploying the FMO methodology to understand and capitalize on **PPI interaction hotspots to guide SBDD**[2]







1 M. Smilova *et al.* Fragment Hotspot Mapping to Identify Selectivity-Determining Regions between Related Proteins. J. Chem. Inf. Model. **2022**, 62, 16, 3784–3799
2 A. Heifetz *et al.* Hotspot Identification and Drug Design of Protein–Protein Interaction Modulators Using the Fragment Molecular Orbital Method. J. Chem. Inf. Model. **2022**, 62, 16, 3784–3799

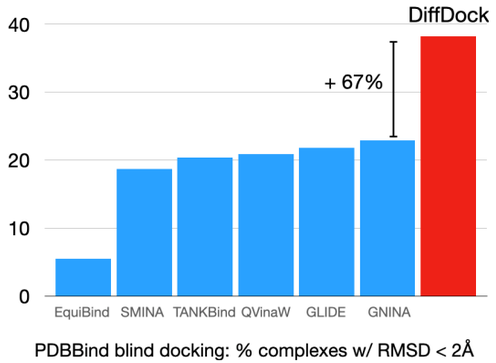
Binding sites identification at Evotec

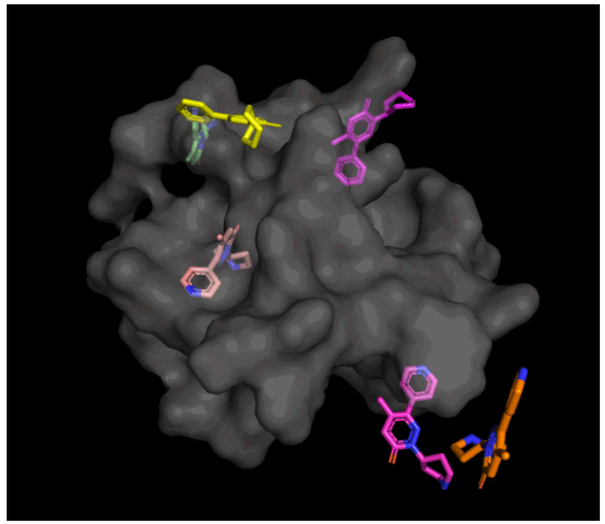
Diffdock

Diffdock and PocketMiner

Diffdock[1] (generative model based on diffusion)

- State-of-the-art on blind docking results[1]
- Diffdock allows multiple positions and poses to be generated
- Trained on a large number of protein-ligand pairs
- The model is able to identify multiple binding sites
- It uses two models:
  - generates multiple conformations in multiple sites
  - ranks the poses according to a confidence model

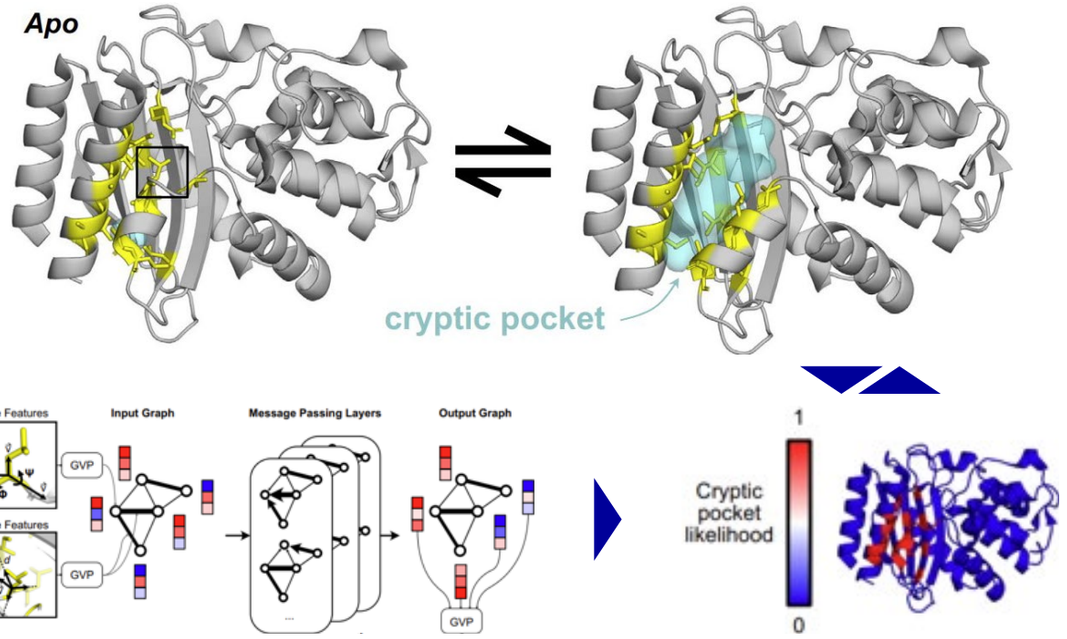




- 38% Top-1 prediction with RMSD<2°A on PDBBind
- Significantly outperforms traditional docking (23%) and deep learning (20%) methods

PocketMiner[2]

- Cryptic pocket formation likelihood: a graph neural network predicts where pockets are likely to open in molecular dynamics simulations
- Evaluates if each residue can rearrange as part of its thermal fluctuations
  - Does not necessitate existing ligands
  - Trained on examples of pocket opening events



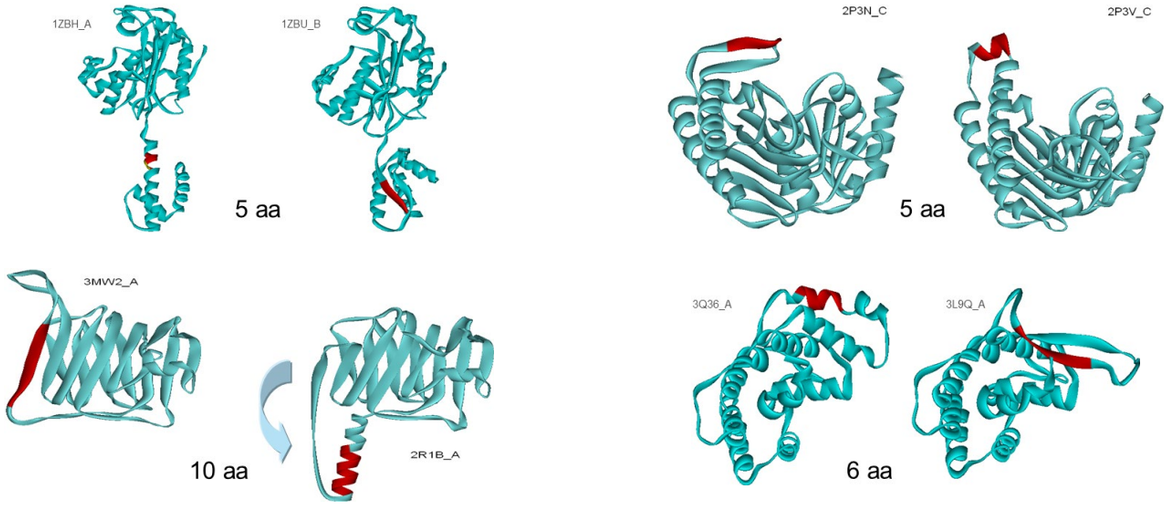
1 Corso *et al.* DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. Published as a *conference paper at ICLR* **(2023)**
2 Bowman *et al.* Predicting locations of cryptic pockets from single protein structures using the PocketMiner graph neural network, *Nature Communications*, **(2023)** 14, 1177

Binding sites identification at Evotec

α/β domains switchability for *in-silico* discovery of allosteric sites and modulators

- **Switchability**
  - Some protein segments can SWITCH their Secondary Structure between HELIX and STRAND → Ambivalent / Switchable segments
- **Allosteric modulation**
  - The switch, or the selection of a structure, may be favored by the binding of a modulator
- **Advantages**
  - Allosteric modulation → Higher specificity
  - New compounds classes → Patents

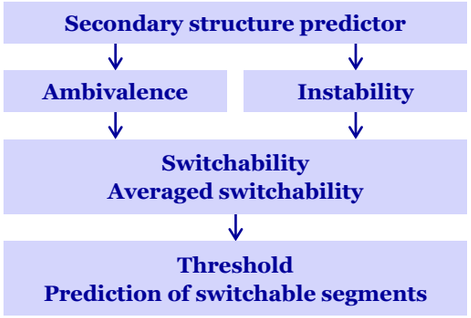
Examples of switched structures[1] in the PDB



Algorithm for predicting α/β switchs

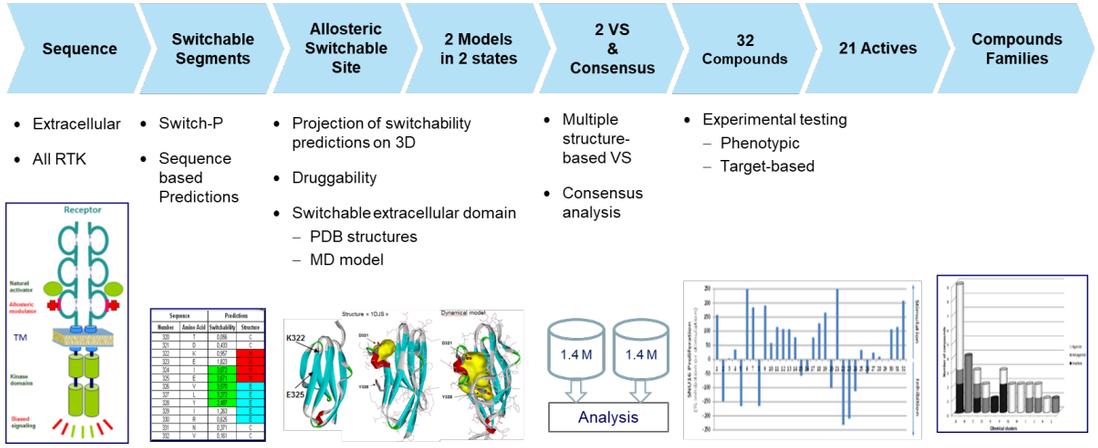
SWITCH-P: sequence-based switchability predictor

1) Secondary structure predictor
2) Ambivalence: Static property (wild-type)
3) Instability: Dynamic property (mutations)
4) Switchability & averaged switchability
5) Threshold & prediction of switchable segments



Example

Virtual screening on an a-helix to β-strand switchable region of the FGFR2 extracellular domain revealed positive and negative modulators[2]

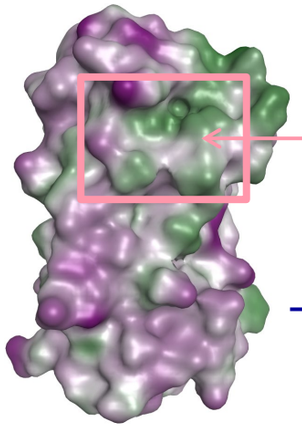


1 Unmodified sequences
2 Virtual screening on an a-helix to b-strand switchable region of the FGFR2 extracellular domain revealed positive and negative modulators. C. Diaz et coll, Proteins 2014; 82:2982–2997

Molecular Dynamics (MD) for cryptic site search

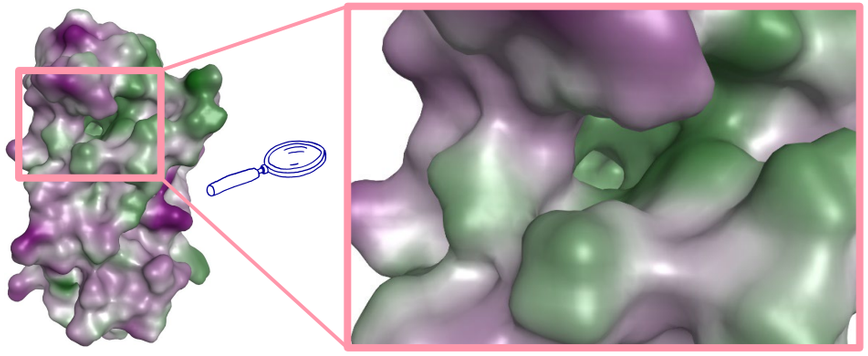
A test case: NPC2 protein: water, mixed and organic solvents MD

PDB entry 1NEP, **Apo**
starting structure (for MD)

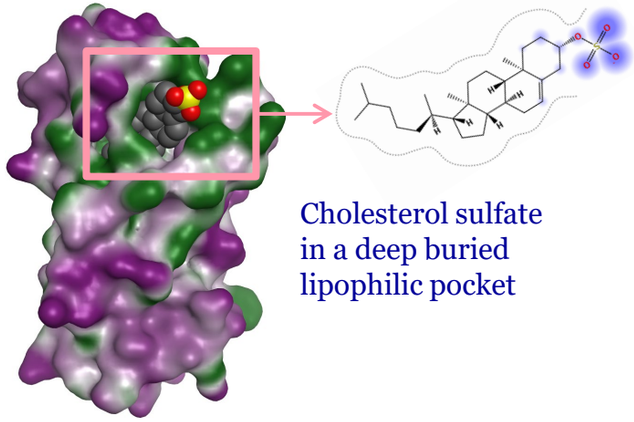


The cryptic site is not visible nor detectable by software

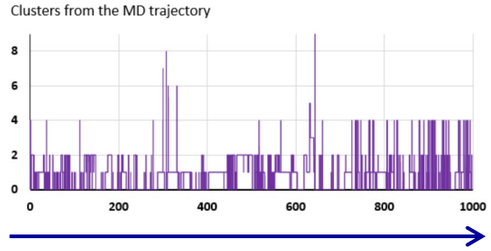
After~600 ns of MD of the Apo PDB, a deep pocket is visible and well discovered by site detection software tools



Holo PDB entry 2HKA
(for comparison)

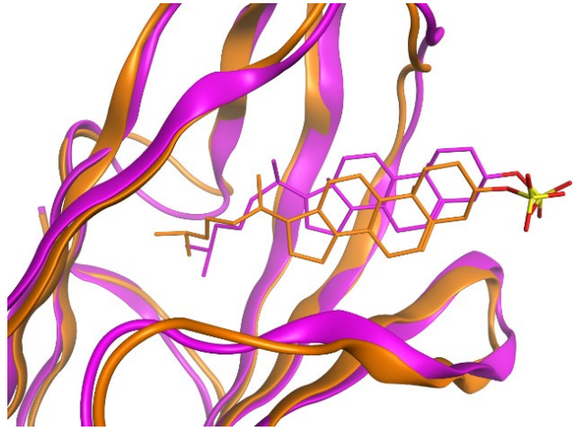


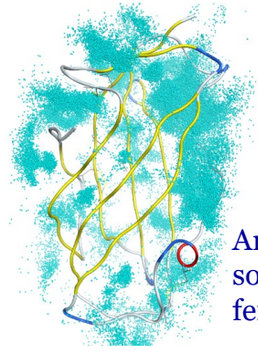
Cholesterol sulfate in a deep buried lipophilic pocket



Clusterings, Representatives, Docking, Scoring, Pose analysis

Best docking pose from MD snapshot aligned with HOLO pdb 2HKA





An organic solvent preference map

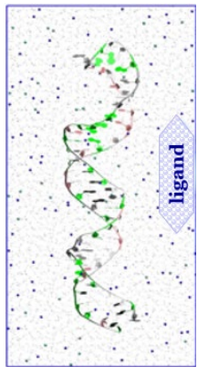
Evotec proprietary small molecule-targeting RNA design platform

From 3D Modeling to Virtual Screening

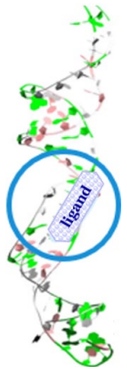
3D structure determination

Atomic solution model generated at Evotec

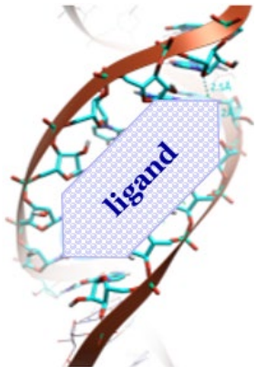
Pocket Search



Pocket identification

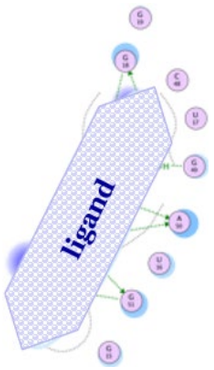


Potential Binding mode



Binding Mode-1 (~600ns)

RNA-Ligand interaction map



- Validated binders fit well
- Non-binders show poor fit
- Model able to discriminate

MD simulations

- X'-UTR RNA 3D Structure + Ligand (Random placement)
- FF parameters
- Charge Neutral [buffer conditions]
- Periodic Boundary Conditions + Particle Mesh Ewald
- Microsecond time scale simulations
- Clustering and pocket prioritization

Molecular Docking

- Docking in consensus pocket found via MD
- Full parametrization of compounds
- Pre-alignment and docking
- Docking pose rescoring

Ongoing assessment and implementation of published tools/methods

Identifying small-molecules binding sites in RNA conformational ensembles with SHAMAN

Panei et al;

Integrated Drug Discovery, Molecular Design Sciences, Sanofi, Vitry-sur-Seine, France
Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Structural Bioinformatics Unit, Paris, France

Predicting Small Molecule Ligand – RNA Binding Pocket Binding Modes Using Metadynamics

Zhixue Bai[1] and Alan Chen[1,*]

[1]Department of Chemistry and the RNA Institute, University at Albany, State University of New York, Albany, NY, 12222, USA

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

Mechanism of Ligand Binding to Theophylline RNA Aptamer

Sana Akhter, Zhichao Tang, Jinan Wang, Mercy Haboro, Erik D Holmstrom, Jingxin Wang,* and Yinglong Miao*

Cite This: https://doi.org/10.1021/acs.jcim.3c01454 Read Online

OXFORD

Briefings in Bioinformatics, 2023, **24(4)**, 1–11
https://doi.org/10.1093/bib/bbad187
Advance access publication date 18 May 2023
Problem Solving Protocol

Structural interaction fingerprints and machine learning for predicting and explaining binding of small molecule ligands to RNA