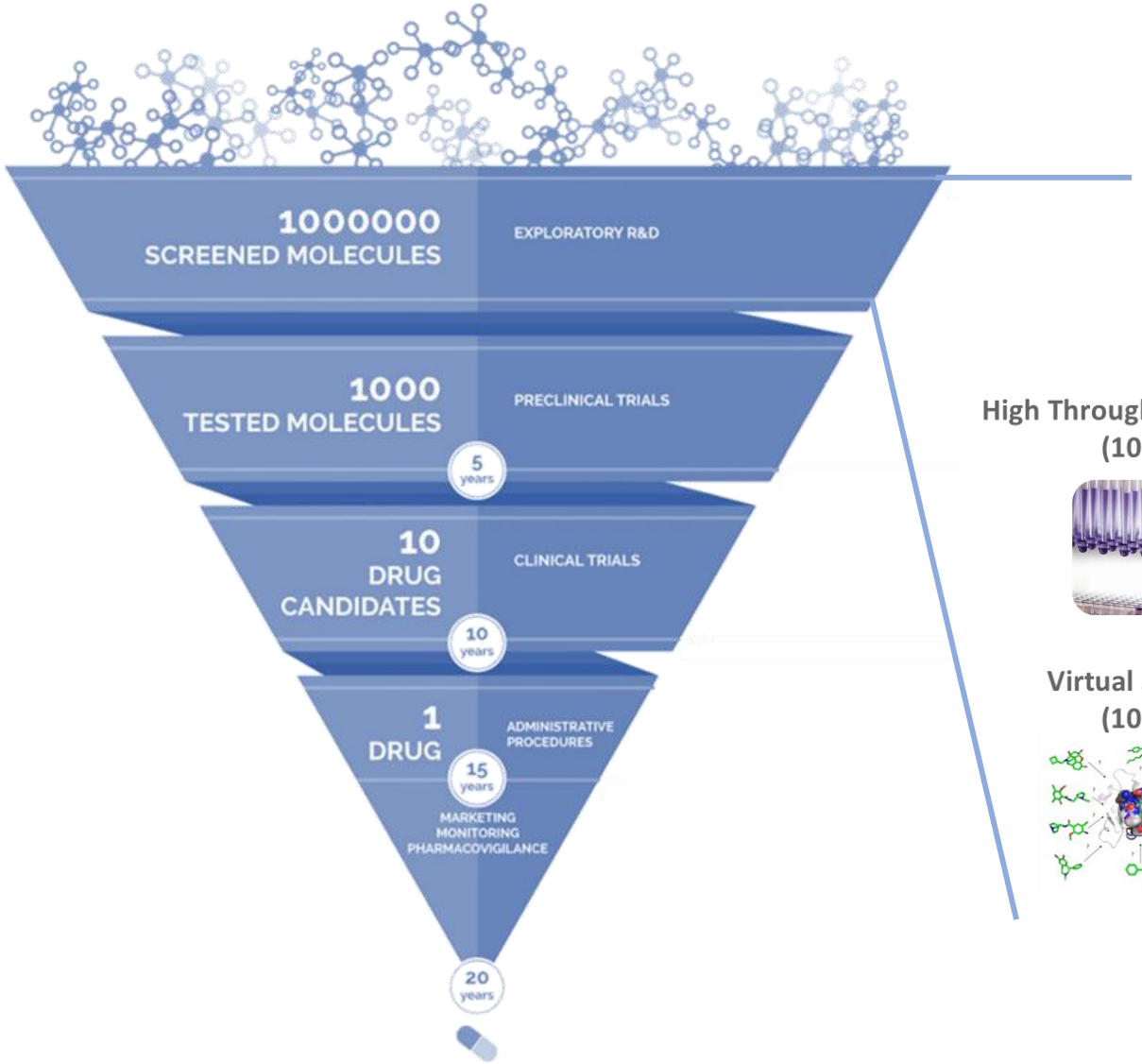
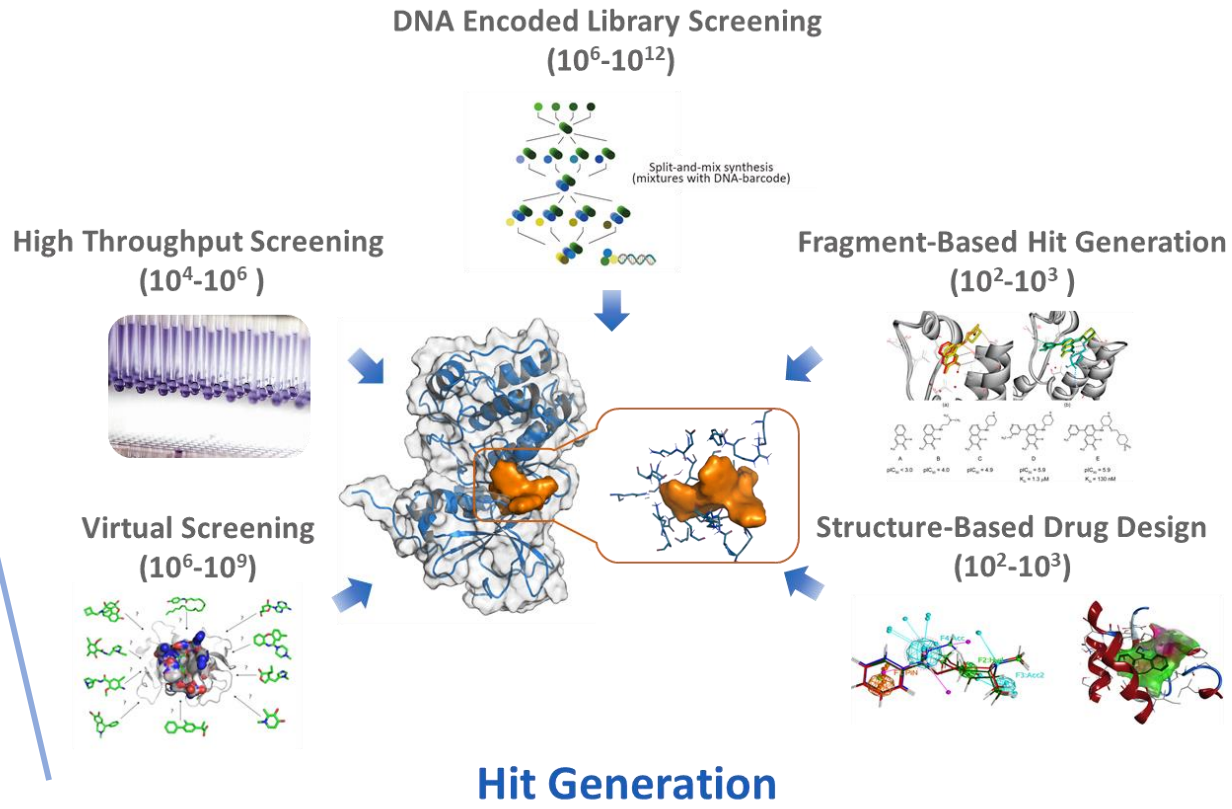
Big Data Management and Applications in WuXi AppTec DEL

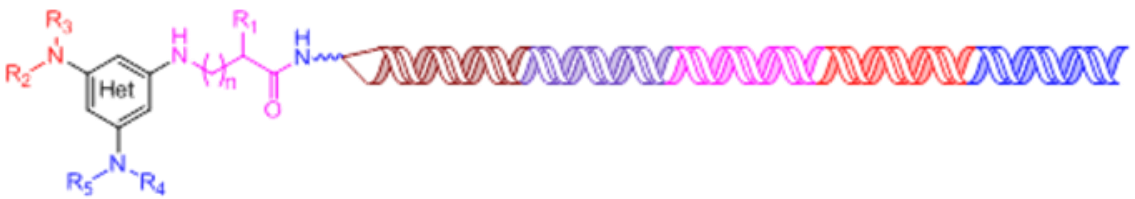
Jason Deng

Drug Discovery is an exploration of chemical spaces





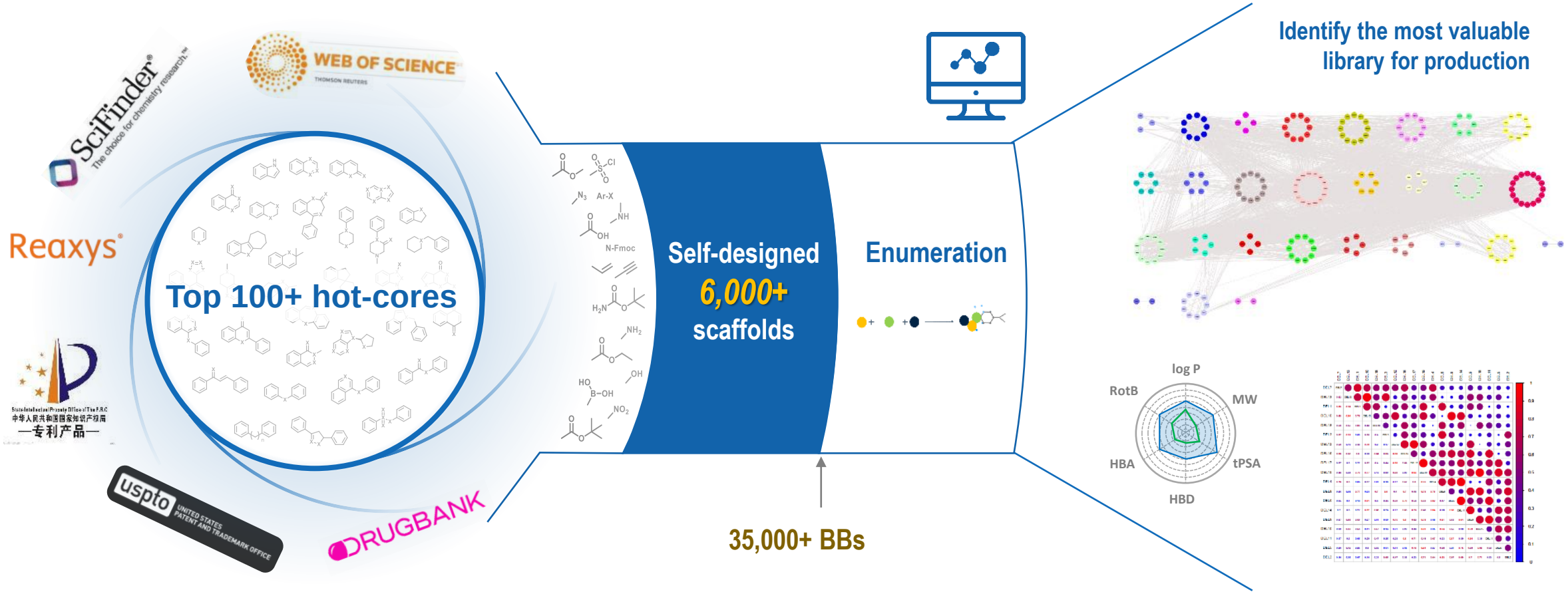
What is DNA-Encoded Library



- **A combinatorial chemistry library**
  - chemical compounds bar-coded by DNA
  - all compounds in a mixture
  - tag-based de-convolution

- **An affinity-based selection system**
  - minimum assay development
  - no cost-per-compound for screening
  - rich SAR reported in one assay

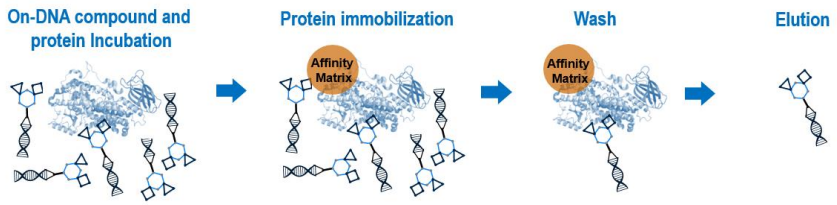
DNA-Encoded Library Design

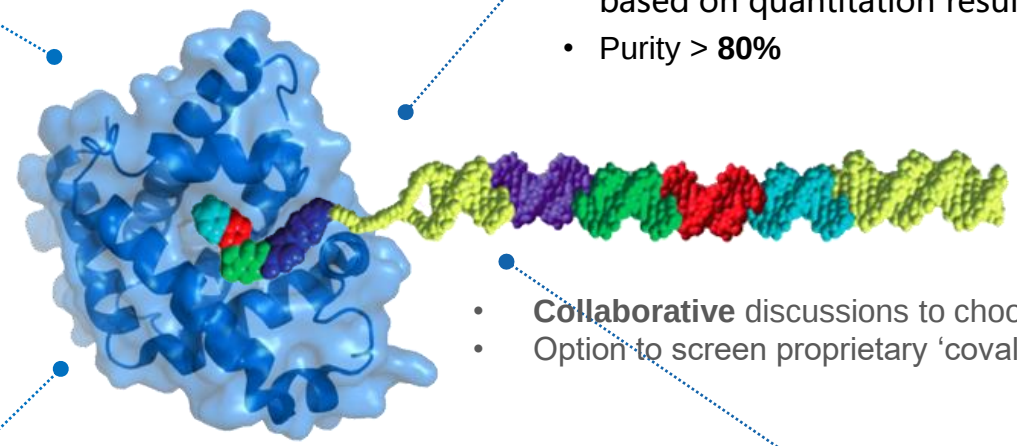


DEL Affinity Selection

1. Affinity Selection

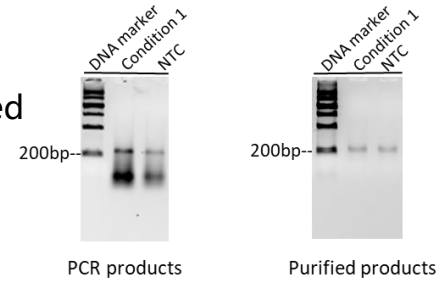
- Multiple parallel conditions
- No target control



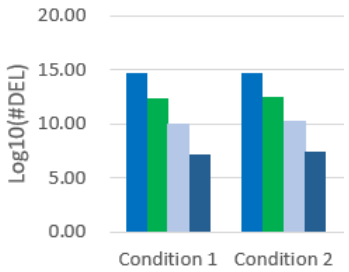


3. PCR

- Number of cycles pre-determined based on quantitation result
- Purity > **80%**

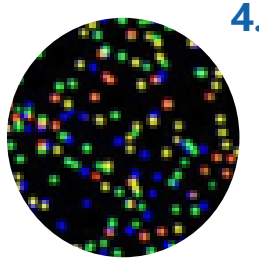


- **Collaborative** discussions to choose up to 8 selection conditions
- Option to screen proprietary 'covalent' DELs



2. qPCR Quantitation

- DEL quantity determination
- Recovery rate monitoring



4. NGS

- % of >= Q30 Bases (PF) > 70%

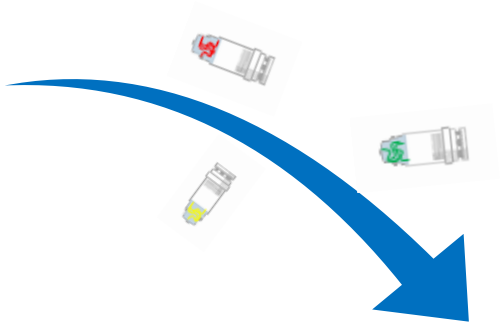
Additionally:

- **Collaborative** discussions to choose up to 8 selection conditions
- Option to screen proprietary 'covalent' DELs

Big Data Management in DEL

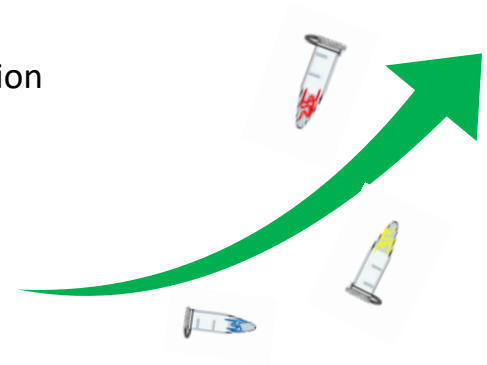
Single Library

10M ~ 3B



Single Selection Condition

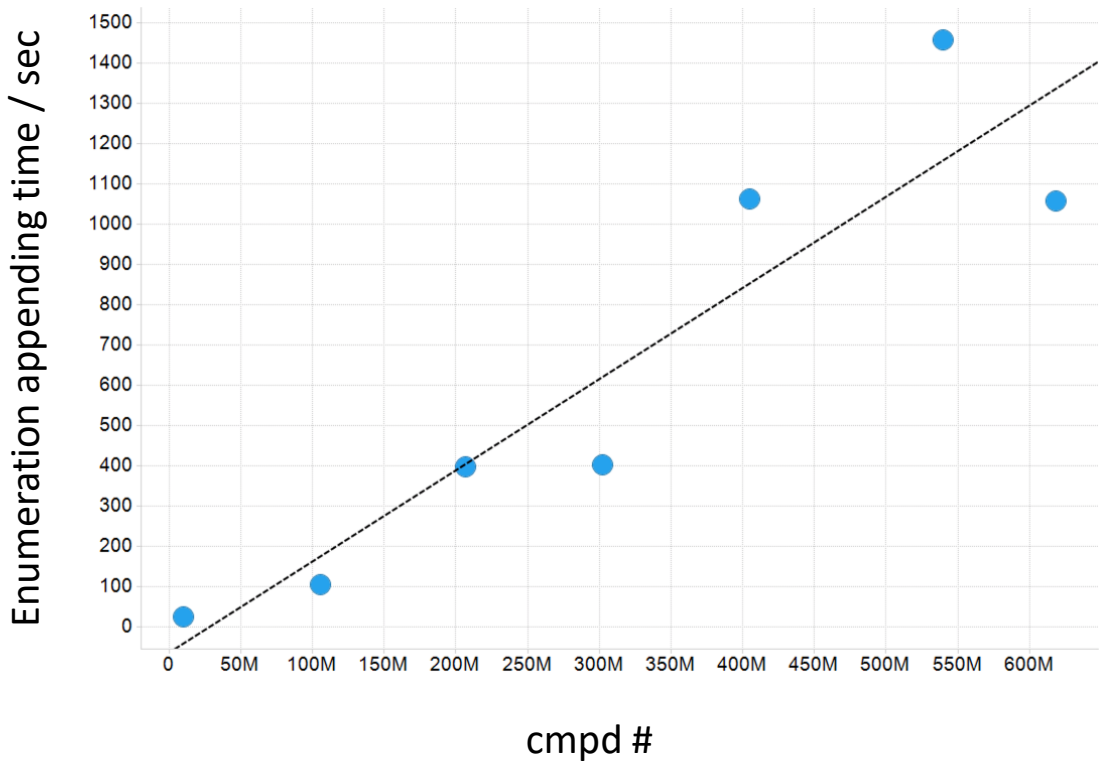
1M ~ 100M

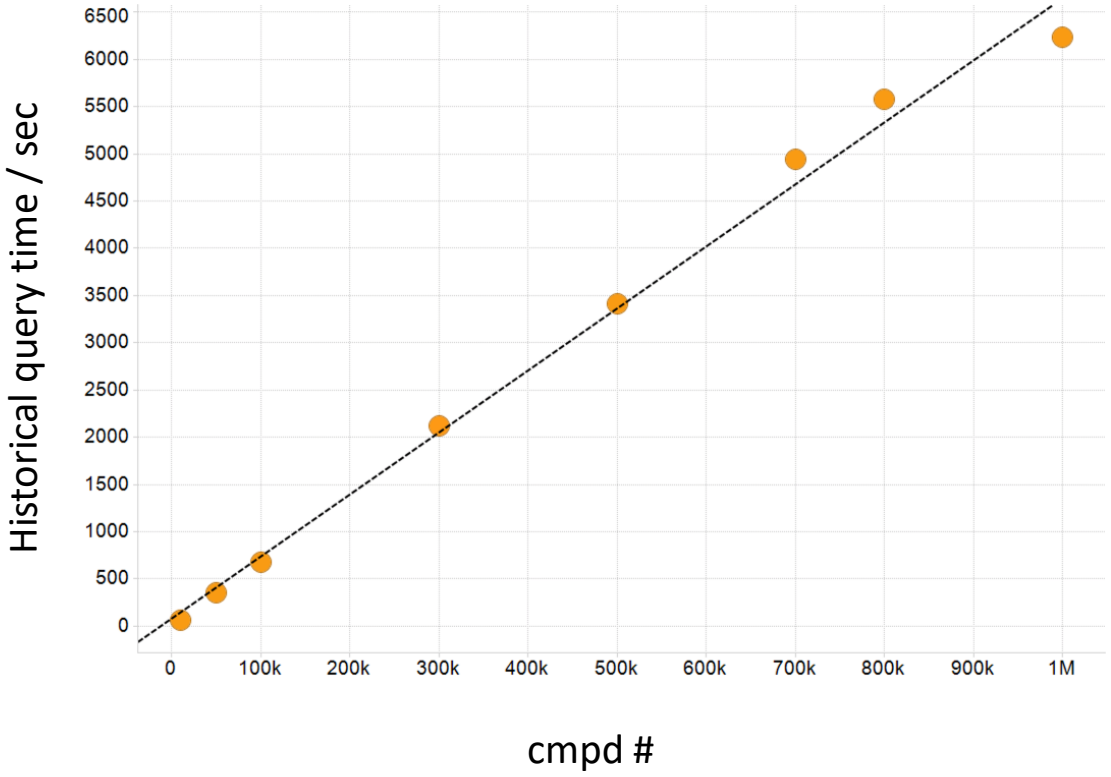


TB ~ PB sized

- Enumerated Structures
- Campaign Based Selection Data
- Manipulation within Campaign
- Historical Performance
- Free Style Query

Overall Performance from Managed Data

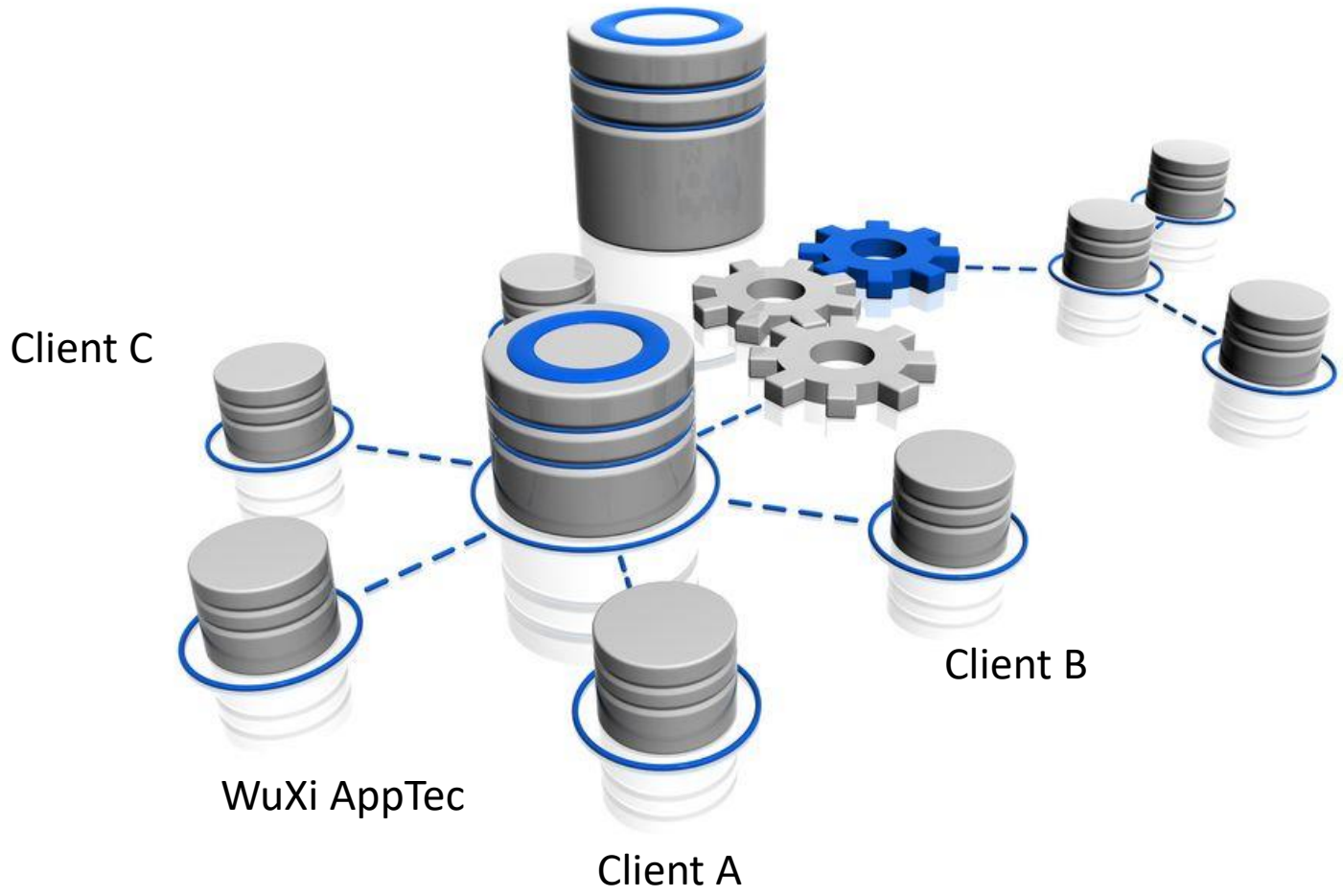




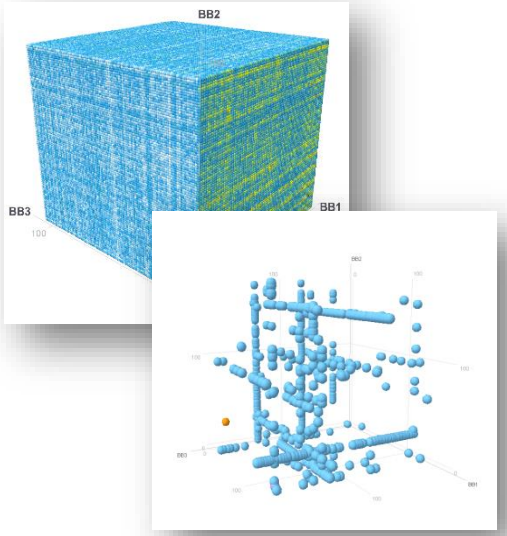
Case 1: Export compounds with enumerated structures directly

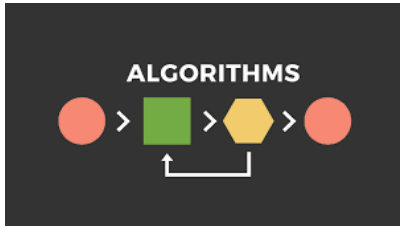
Case 2: Historical performance query test

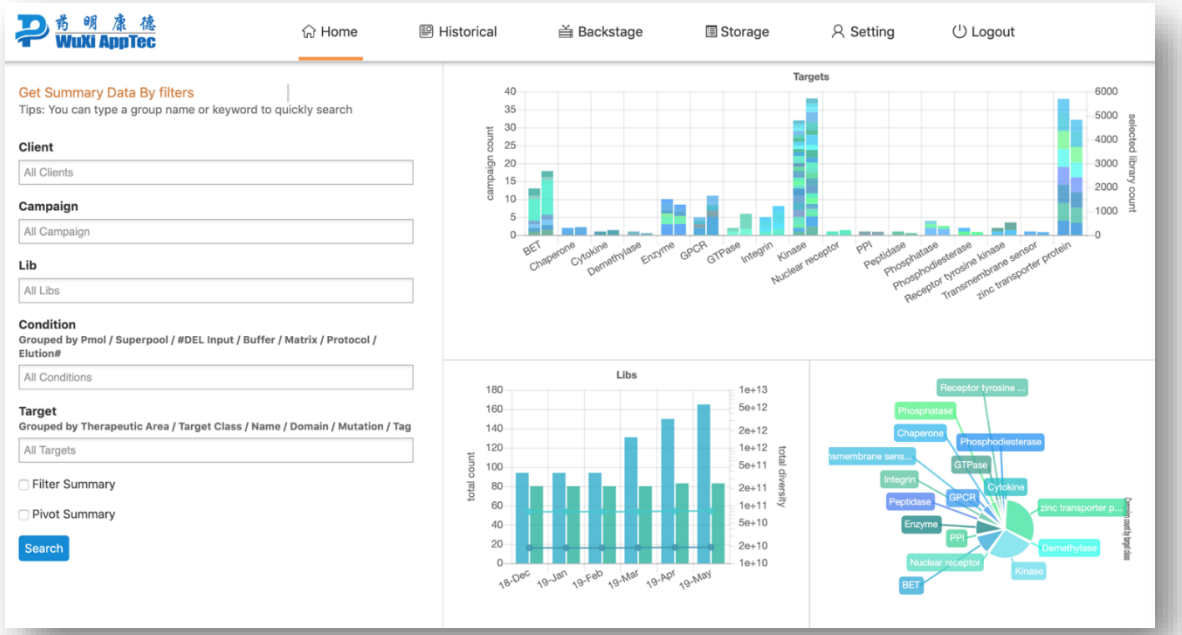
Isolated Client Data

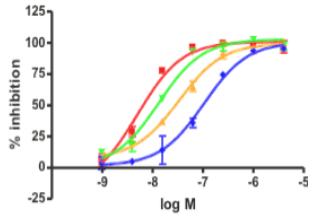


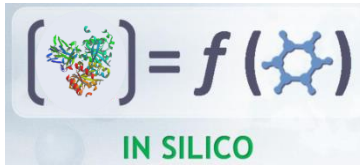
DEL Lake: enable fishing in the lake of DEL cmpds











DELopen

DELight

...

WuXi AppTec DEL Screening Business Model

Pre-Selection Tests → Affinity Selection → NGS → Data Analysis

WuXi AppTec DEL Selection Package

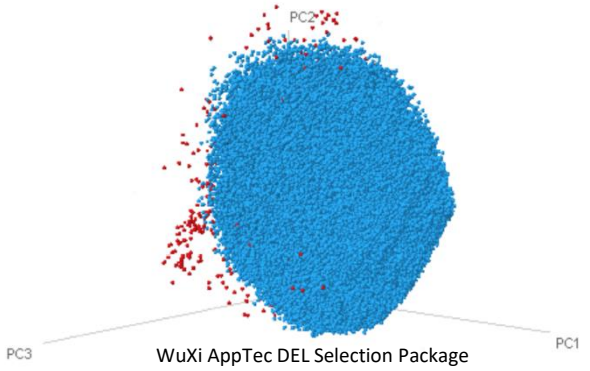
DELight

DELopen

Library NO.

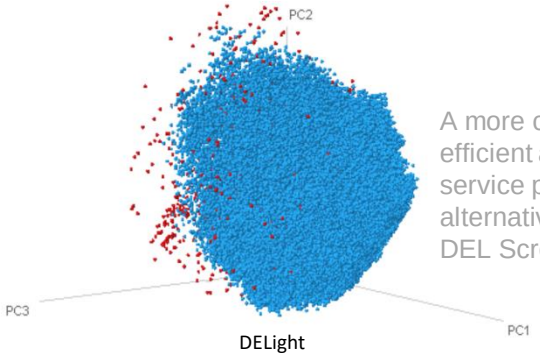
WuXi AppTec DEL Selection Package

- ***50+ billion*** compounds
- For all users at competitive market price



DELight

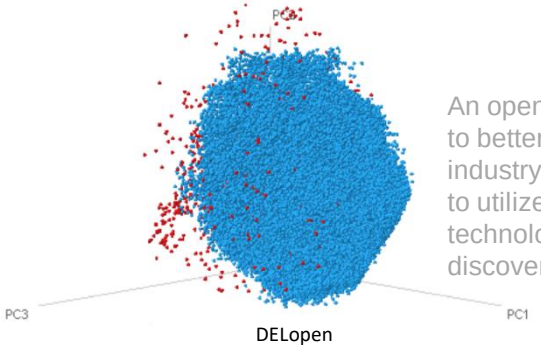
- ***14.1 billion*** compounds
- For all users at economic price



A more convenient, efficient and economical service package as alternative to traditional DEL Screening service

DELopen

- ***4.4 billion*** compounds
- ***Free*** for academic user



An open source platform to better connect industry and academia to utilize DEL technology in drug discovery

WuXi AppTec Virtual Library

药明康德
WuXi AppTec

WuXi AppTec Virtual Library

General Introduction

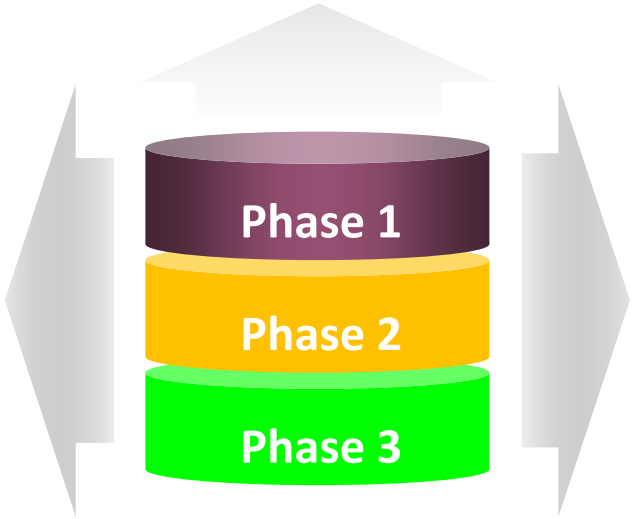
1

Final compounds enumerated

- ~50 M for phase 1
- ~1.7 B for phase 2
- ~0.6 B for phase 3
- SMILES for virtual library

- **>300 GB for csv file**
- **~20 GB compressed**

2 models



2

GalaXi Space

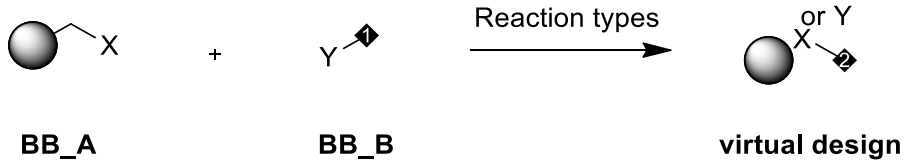
BioSolveIT

- Collaboration with BioSolveIT
- Fragment based
- ~50 M for phase 1
- ~1.7 B for phase 2
- ~0.6 B for phase 3

- **File size ~60 MB**
- **Similarity searching in 1 min**

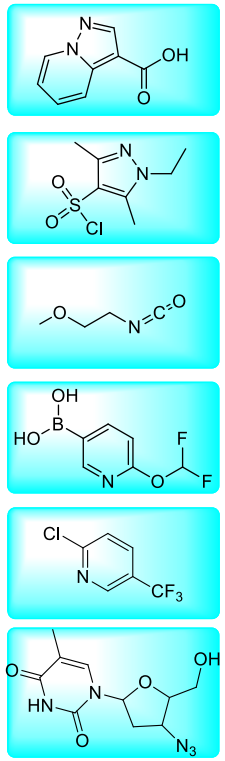
Phase 1 Building Block Docking

Wide Coverage, Diversity, Novelty

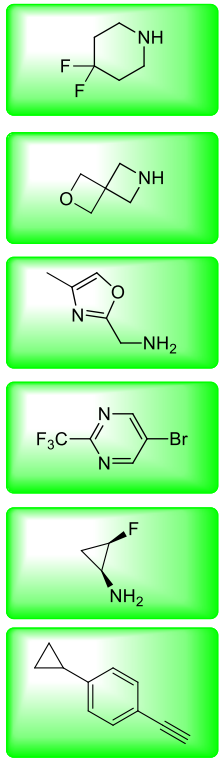


- 50M compounds
- 47,879 commercial BBs, including novel WuXi BBs
- 38 reaction types, 128 validated reaction protocols
- 75%+ success rate, delivered in 4-6 weeks

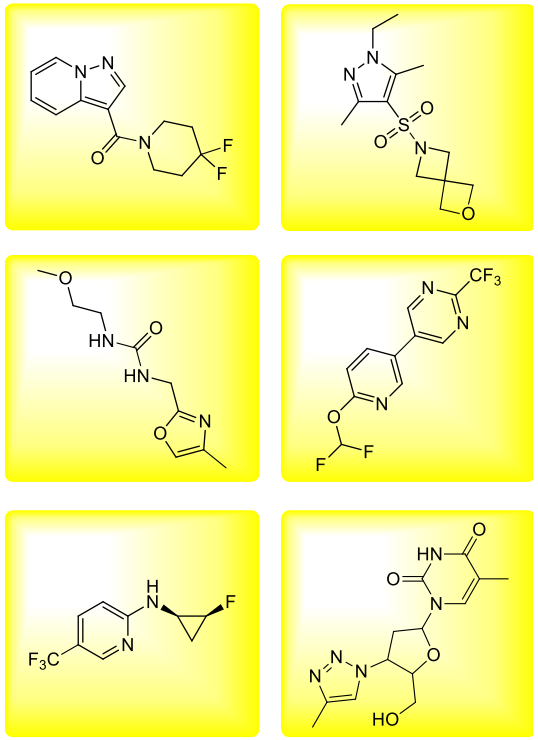
Examples:



Examples:



Examples:

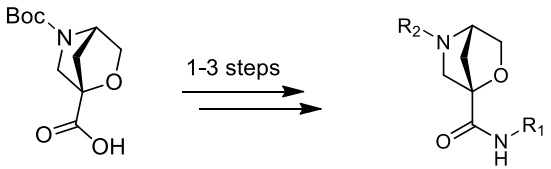


- Acylation
- Aldol condensation
- Amide Formation
- Amine-ester exchange
- Appel Reaction
- Buchwald reaction
- Carbamate formation
- Chan-Lam Coupling
- Click Chemistry
- Condensation
- Conjugate addition
- De-methylation
- De-protection
- Ester formation
- Halogenation reactio
- Heck reaction
- Hydrolysis
- Mitsunobu reaction
- N-Displacement
- Negishi Coupling
- O-Displacement
- Oxidation reaction
- Protection
- Reduction
- Reduction Amination
- Ring Formation
- S-Alkylation
- Saponification
- Sonogashira reaction
- Stille Coupling
- Strecker Synthesis
- Sulphonamide Formation
- Suzuki Coupling
- Ugi multicomponent reaction
- Ullmann reaction
- Urea Formation
- Weinreb ketone synthesis
- Wittig reaction

Phase 2 WuXi Template Based

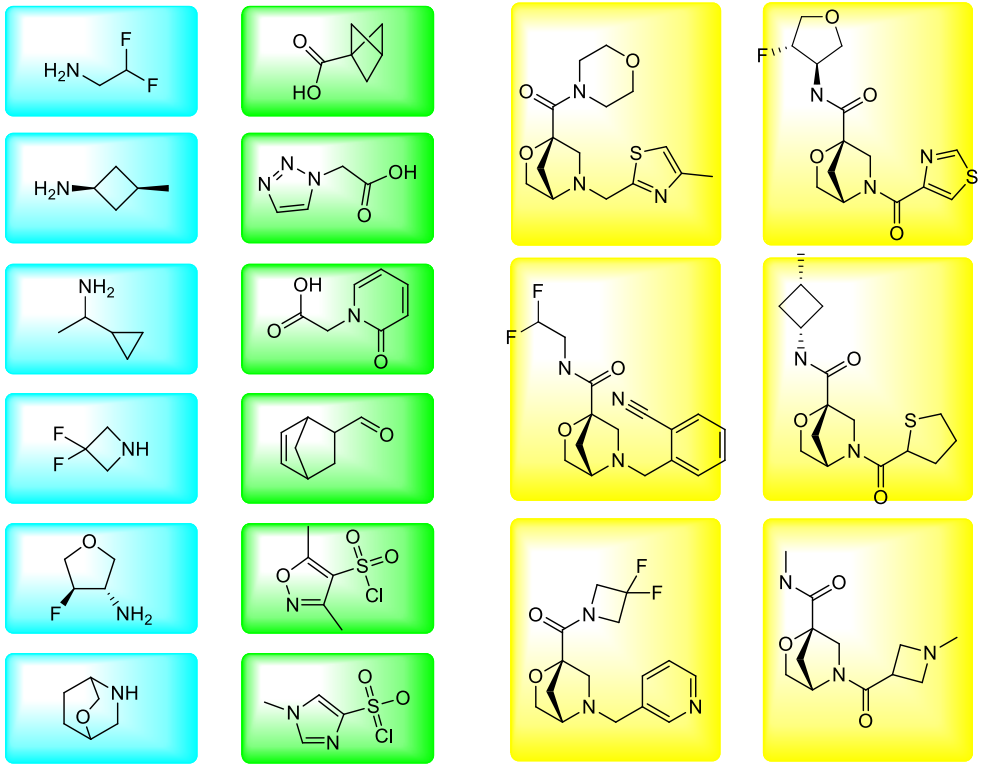
Novelty & Diversity

➢ **WX120709 as a example:**



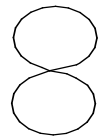
WX120709 **WX120709_Xi*****

Examples: **Examples:** **Examples:**

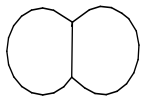


- **~1.7 B virtual compounds designed to date**
- **1,000+ novel templates for future virtual library enumeration.**

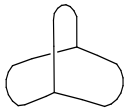
Diverse chemo-types



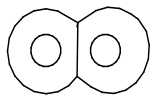
Spiro aliphatic ring SP



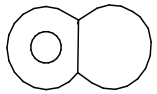
Fused bicyclic aliphatic ring BC



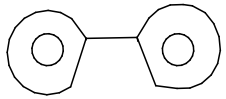
Bridged aliphatic ring BR



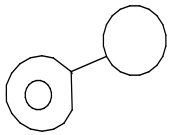
Fused aromatic ring FAR



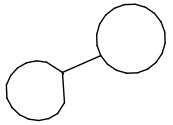
Fused aromatic -aliphatic ring FAA



Bicyclic aromatic ring BAR

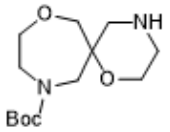


Bicyclic aromatic -aliphatic ring BAA

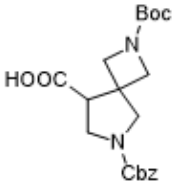


Bicyclic aliphatic ring BAL

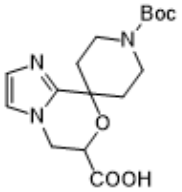
Examples of Templates



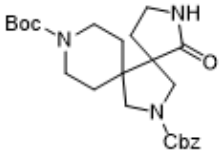
WX100112



WX100287



WX105122



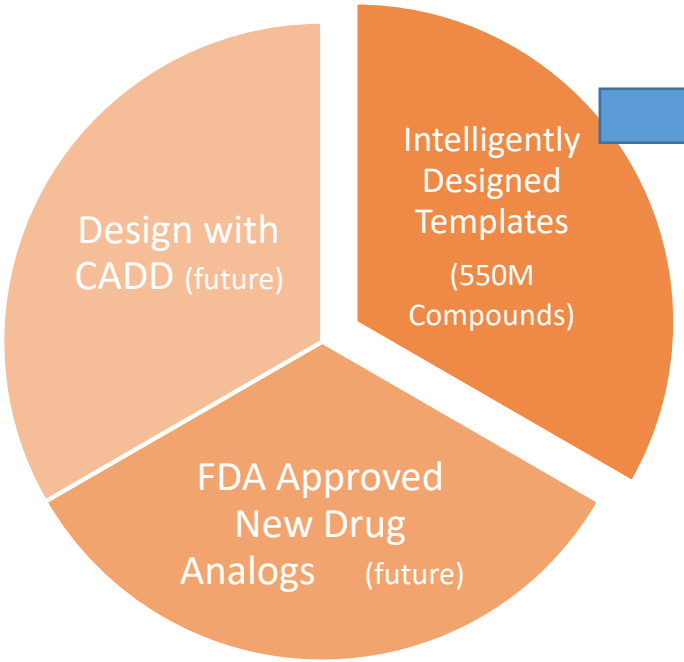
WX105147

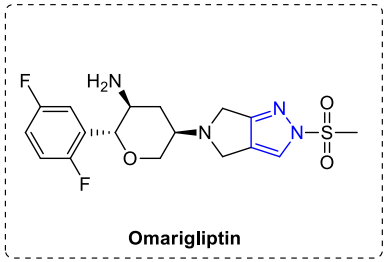
Phase 3 Drug-Like

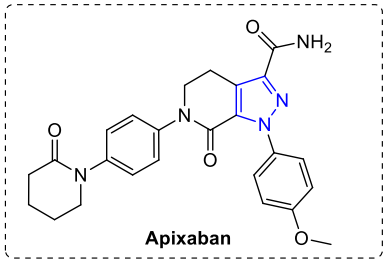
Segment 1 (Intelligently Designed) of 3 Just Released

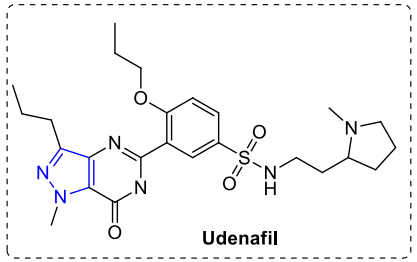
Drug-like Cores + R1 —2 steps→ **Key Intermediate** —R2 / reactions type→ **Virtual Library**

Drug-like Cores Three Segment Approach

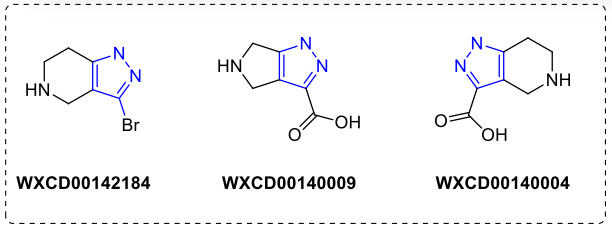


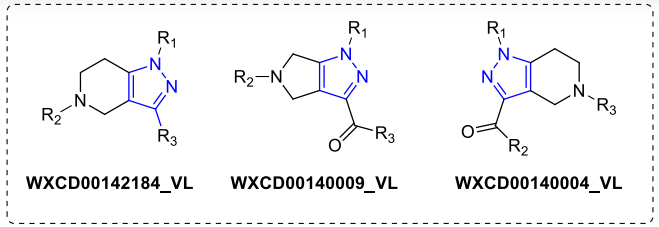






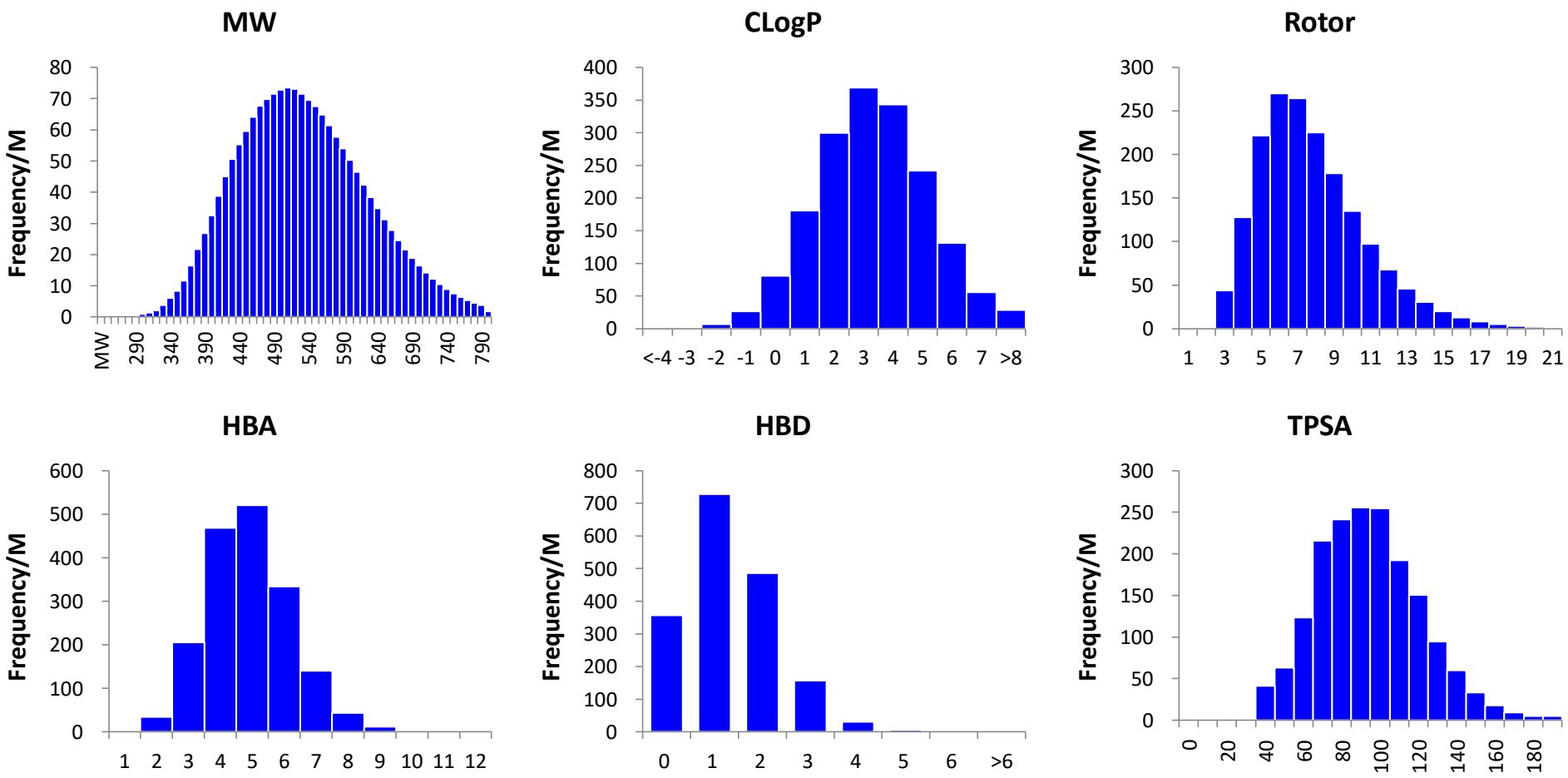
Example: Pyrazols play an important role in drugs or drug candidates, as shown above. We selected pyrazols and intelligently designed using popular catalog compounds and reaction schemes as cores for drug-like virtual compounds.





Virtual Compound Library

Physical Profiles

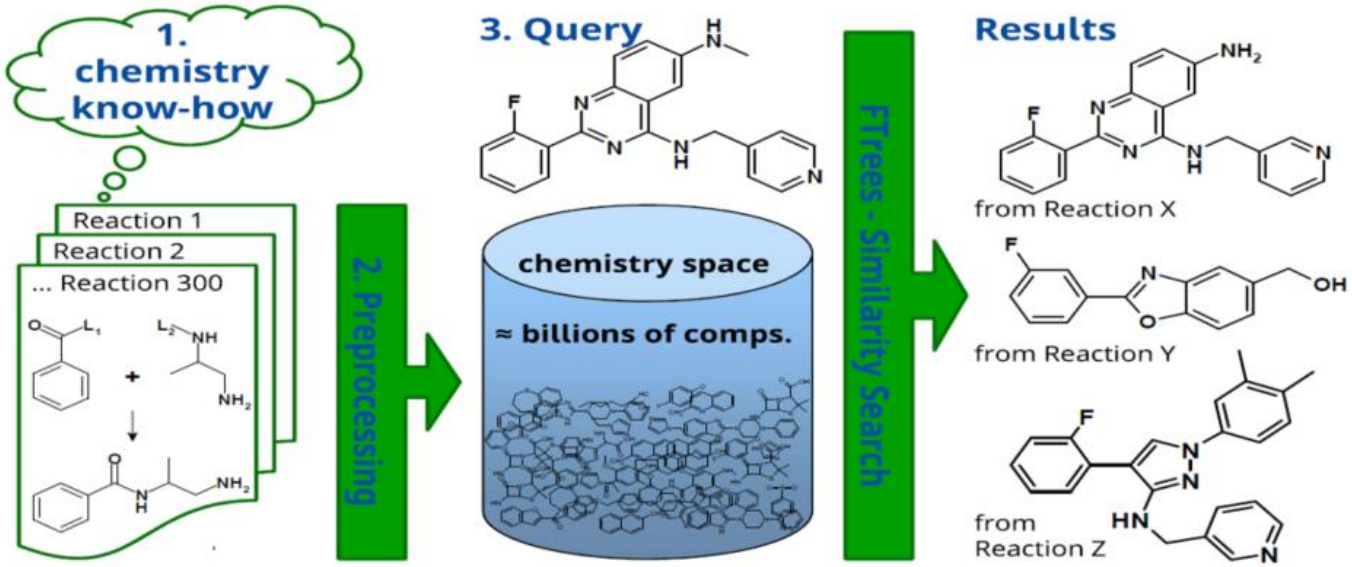


GalaXi™ Virtual Space

Navigate 2.3 x 10^9 in Minutes

www.biosolveit.de/infiniSee/#chemical_spaces

- Collaboration with BioSolveIT
- Similarity searches using BioSolveIT infiniSee software
- Quality processing ensured through CoLibri software
- Based on WuXi AppTec in-stock building blocks and proven reaction schemes

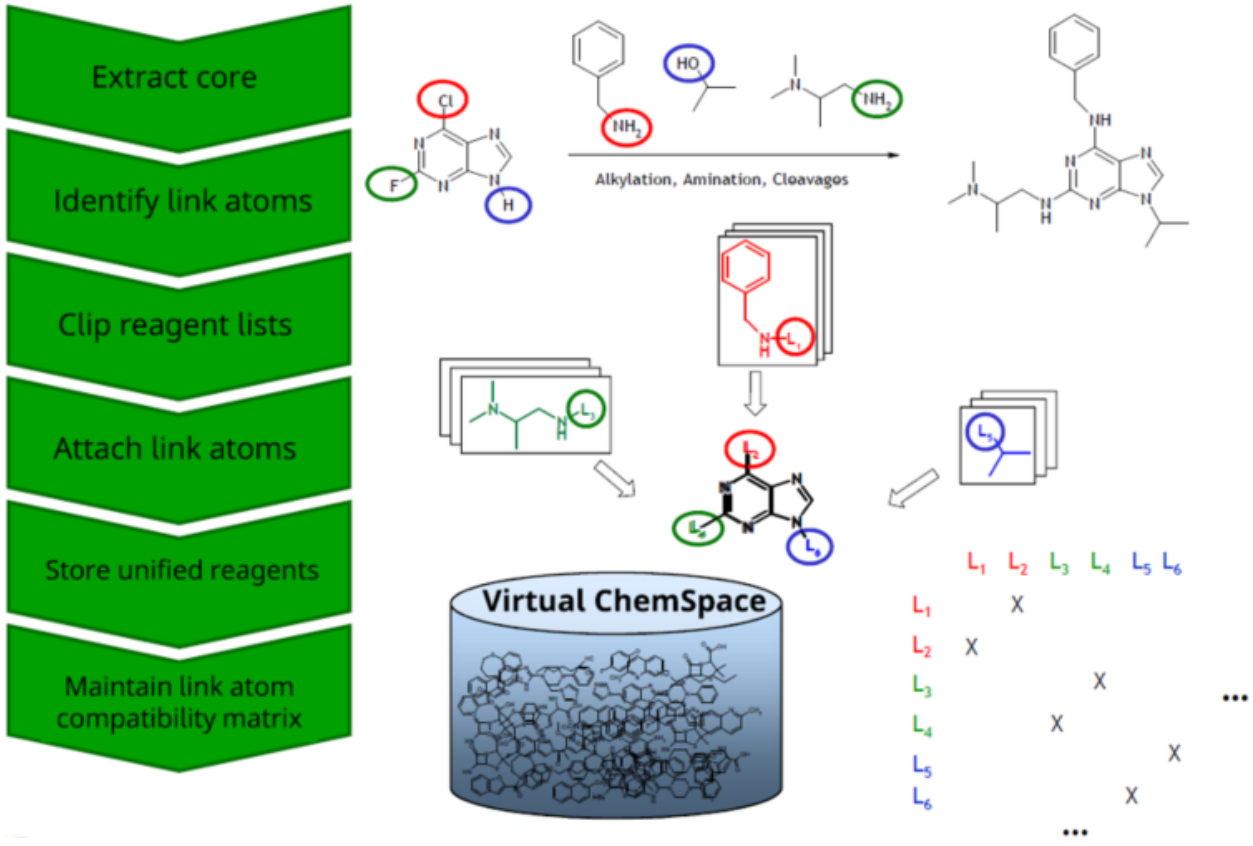


GalaXi™ Virtual Space

Navigate 2.3 x 10^9 in Minutes

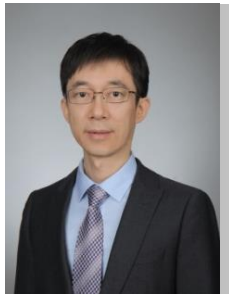
www.biosolveit.de/infiniSee/#chemical_spaces

CoLibri® – Capturing Chemical Reactions



The Library Team

Yan Xu
Executive Director; Head of Library Business
Joined WuXi in 2001
Ph.D. from Nanjing University
Over 16 years library synthesis experience



Yu Gao, Director
Joined WuXi in 2003
Ph.D. from Nankai University
Focusing on library synthesis for 14 years

Shanghai Site

Likai Xia, Associate Director
Joined WuXi in 2016
Ph.D. from Yeungnam University South Korea
Design and synthesis libraries with 10 years' experience

Shanghai Site

Lingbo Yu, Director
Joined WuXi in 2005
Focusing Library synthesis with new scaffolds over 11 years
Visiting scientist" from University of Dundee

Tianjin Site

- 17 years of experience in library design and synthesis; Delivered more than 3,000,000 compounds
- Dedicated purification team to accelerate library production
  - Typical timeline for focused libraries: ~ 4 to 8 weeks
  - As short as 8 calendar days to finish 100-member library
- Availability of diverse reagents to facilitate design and expedite production
  - 25 K+ reagents plus 8 K+ WuXi in-house building blocks in stock

Thank You

Contact

John Shirley, Sr. Director
Marketing & Sales, Research Services Division & LabNetwork

John.Shirley@WuxiAppTec.com
1-207-332-7988

Boston Site