CelerisTx

Insights from a Biotech Startup: How Computational Chemists Shape the Design of Proximity-Inducing Compounds

Alzbeta Tuerkova, Ph.D.
Principal Computational Chemist

29th January 2024

AGENDA

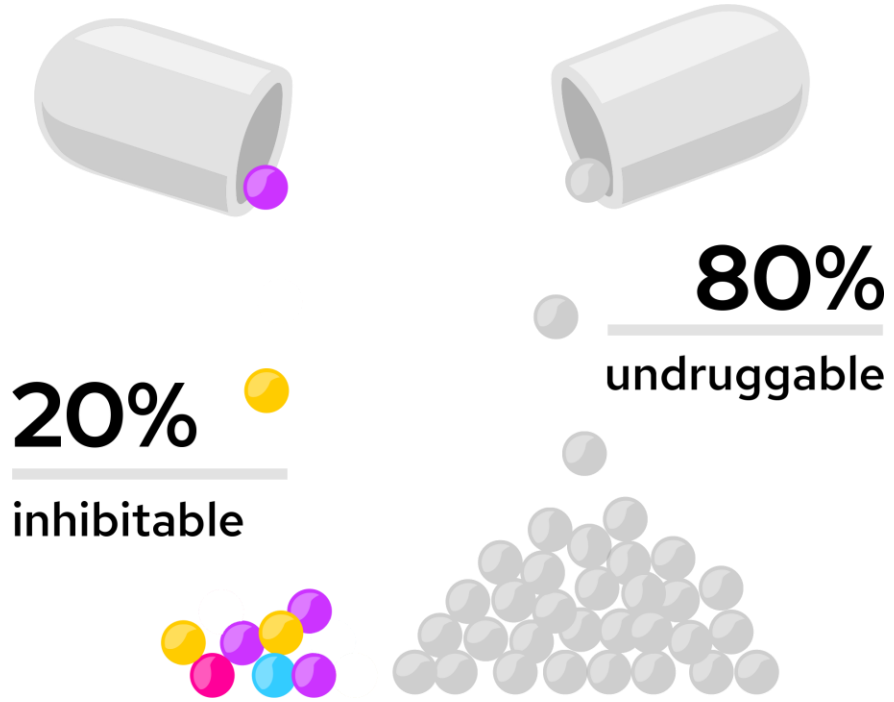
AGENDA

The Problem of Undruggable Targets

Most traditional small-molecule drugs deactivate functions of proteins

The typical mechanism is of action is to target the active site, resulting in inhibition

Only 10-20 %[1] of all pathogenic proteins in the human proteome possess such sites



[1] DANG, Chi V., et al. Drugging the 'undruggable' cancer targets. *Nature Reviews Cancer*, 2017, 17.8: 502-508.

From Protein Inhibition to Degradation

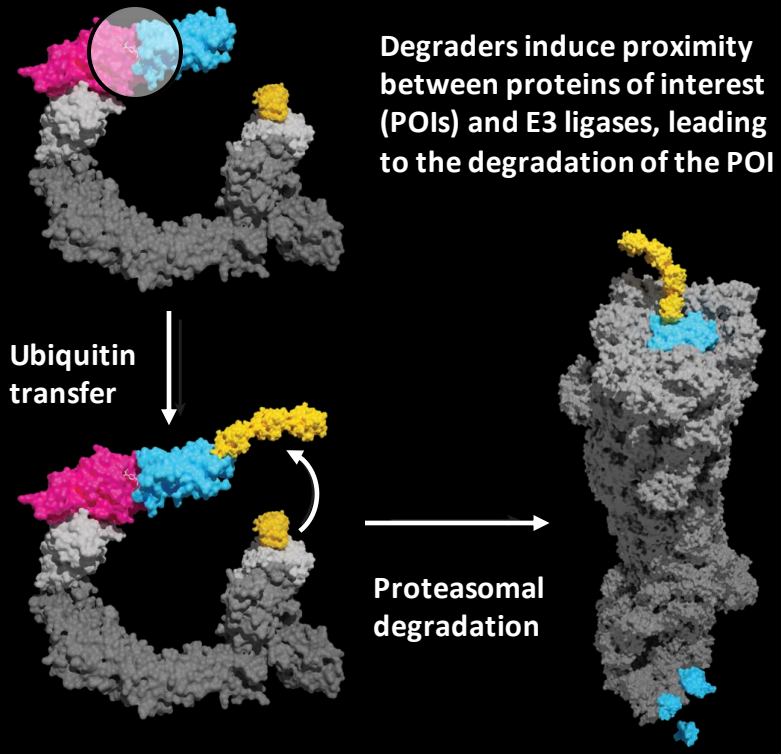
Targeted Protein Degradation

- Humans have different systems to remove dysfunctional/misfolded proteins.
- One dominant system is the intracellular ubiquitin-proteasome complex.
- Degraders are drugs that take advantage of the ubiquitin-proteasome pathway.
- We can hijack these systems to tag specific proteins.

What is a PIC™ Degrader?

Proximity-inducing Compound (PIC™) Degraders are the next generation of highly potent, rationally designed degraders that feature **E3 ligase selection** based on **protein-protein interfaces**.

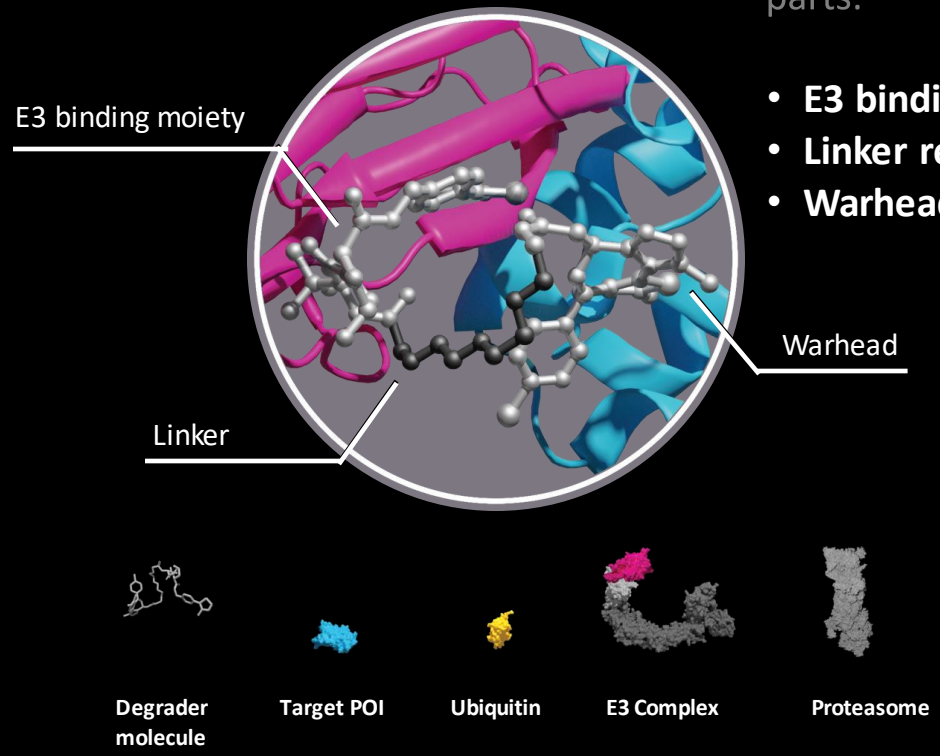
TARGETED PROTEIN DEGRADATION PATHWAY



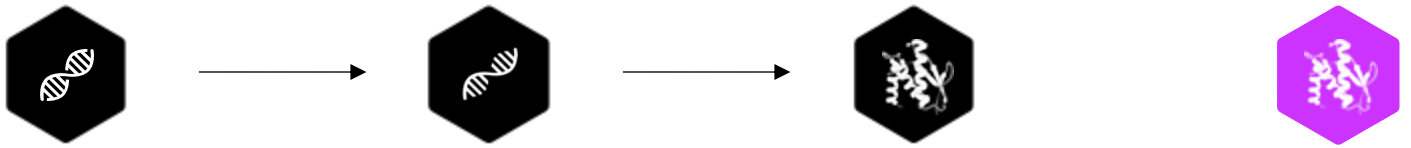
BIVALENT DEGRADER MOLECULES

A traditional bivalent degrader consists of three parts:

- **E3 binding moiety**
- **Linker region**
- **Warhead (POI ligand)**



Advantages of Targeted Protein Degradation



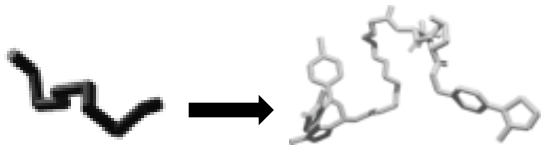
	CRISPR-CAS9	RNAi	Traditional small molecules	**Degraders**
Stability	X	X	✓	✓
Going Beyond Druggables	✓	✓	X	✓
Working Directly on Proteins	X	X	✓	✓
Clinical Validation	✓	✓	✓	✓
Oral Application	X	X	✓	✓
Low Dosage / Tox	X	X	X	✓

The 'Success' in Designing PICs is Multi-factorial

Degradation = Solubility x PPB* x Permeability x ([Complex] & Ubiquitination) x POI synthesis rate

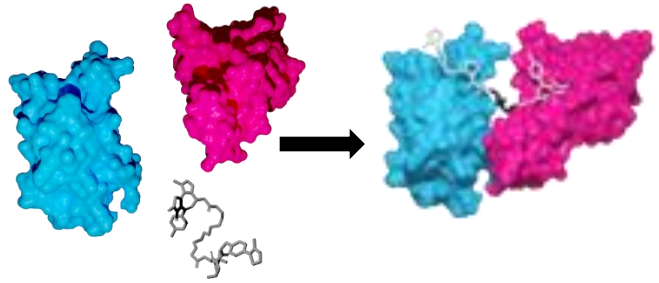
*) **P**lasma **P**rotein **B**inding

The 'Success' in Designing PICs is Multi-factorial



Linker generation

Ternary Complex prediction



Degradation = Solubility x PPB x Permeability x ([Complex] & Ubiquitination) x POI synthesis rate

Distance measurements

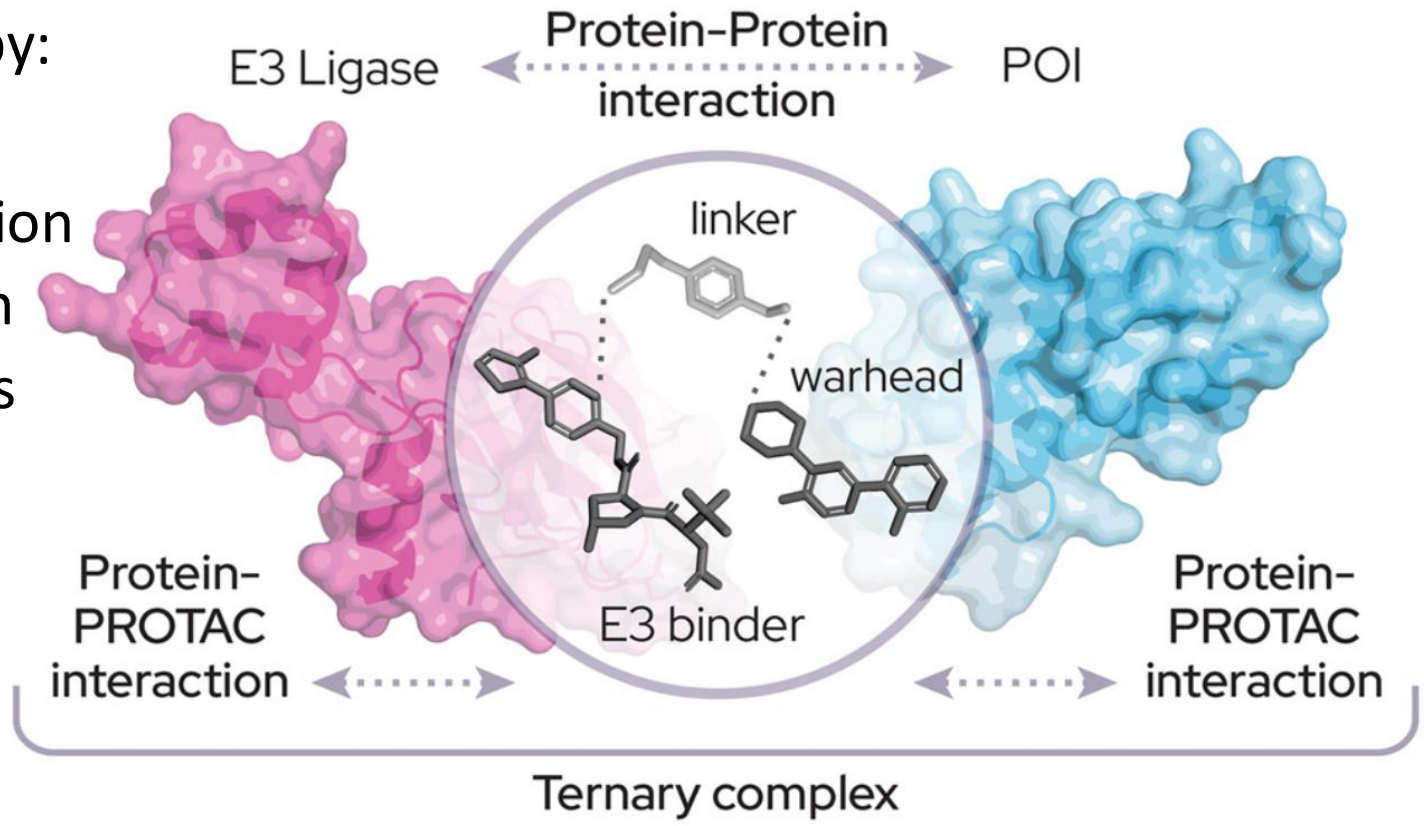
AGENDA

Ternary Complex Prediction: Challenges

1) Prediction of key interactions

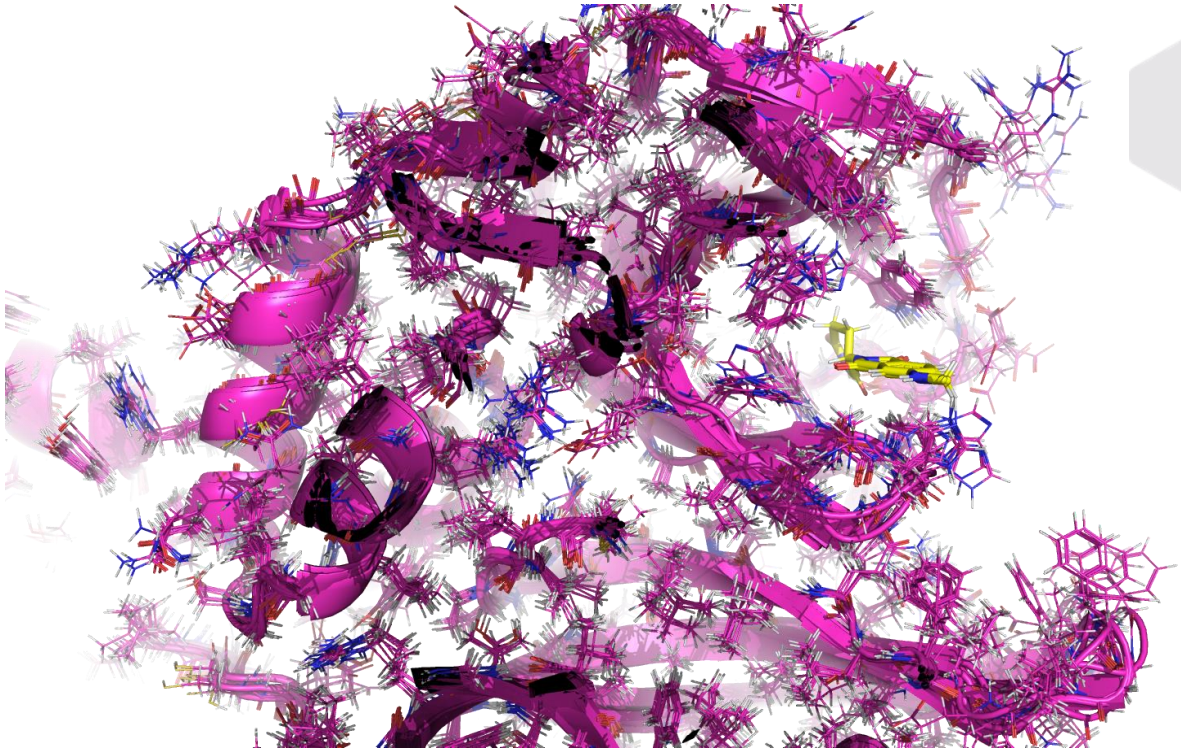
Ternary complexes are mediated by:

1. **Fragment-Protein** Interaction
2. **Protein-Protein** Interaction
3. **Linker** enforced constraints



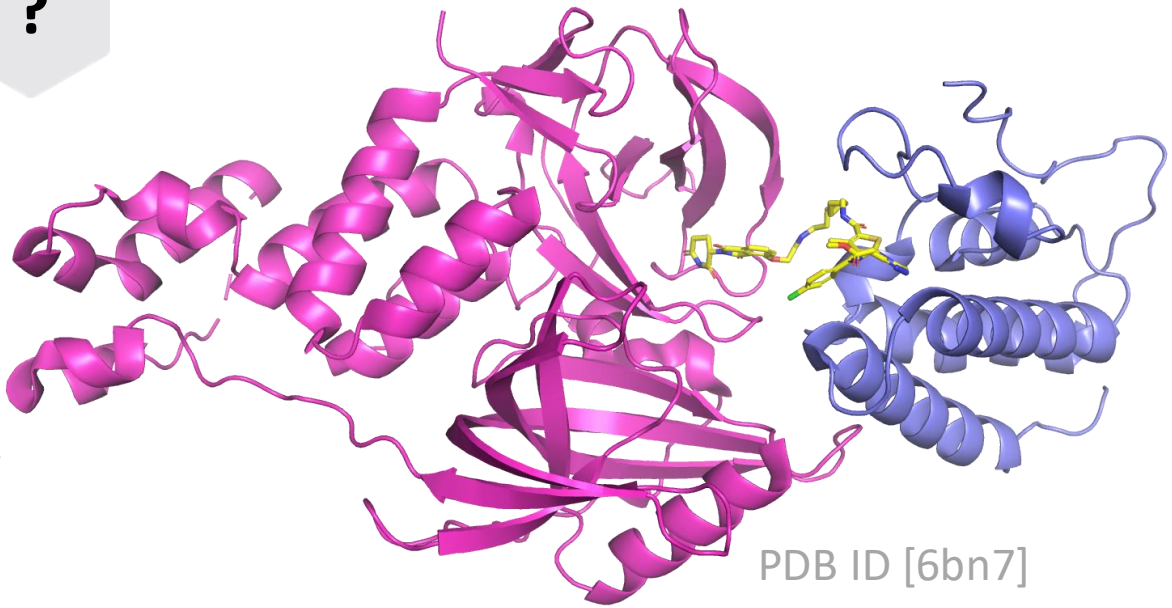
Ternary Complex Prediction: Challenges

2) Accounting for protein conformational flexibility



"unbound" structures (CRBN)

?

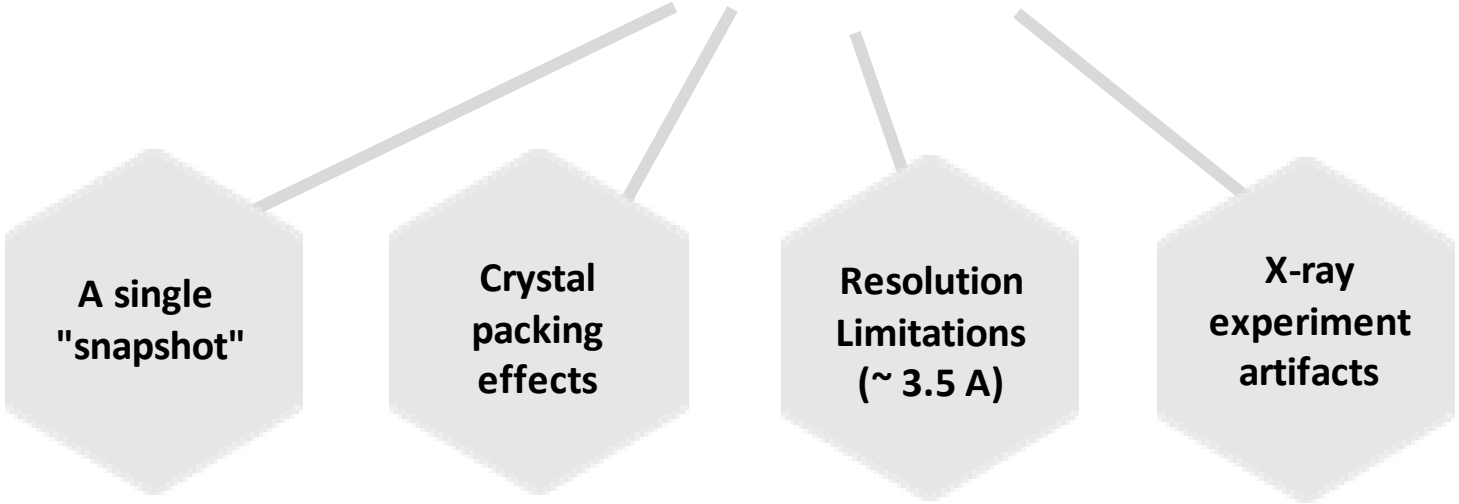


"bound" structure

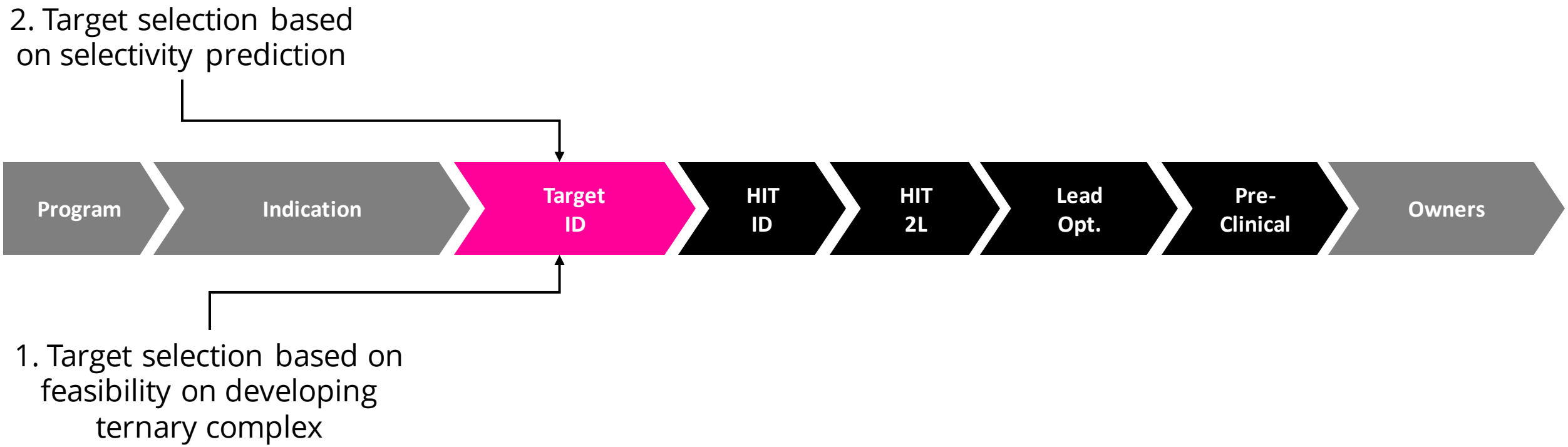
Ternary Complex Prediction: Challenges

3) Validation of the approach

- **Limited number** of experimentally determined ternary complex structures (n=22)
- All available structures originate from **X-ray measurements**



Application of TCP in PIC design



Application of TCP in PIC design

3. Calculate minimum linker length

4. Decide which compounds to synthesize based on ternary complex formation ability

a) Different warhead
    a) Different binding pockets
    b) Different E3 ligase
b) Linker enumeration

Application of TCP in PIC design

5. Decide which compounds to synthesize based on selectivity
   a) Different warhead
      a) Different binding pockets
   b) Linker enumeration

7. Accurate structural prediction

6. Decide which compounds to synthesize based on ternary complex formation ability
   a) Linker modifications
   b) Warhead modifications

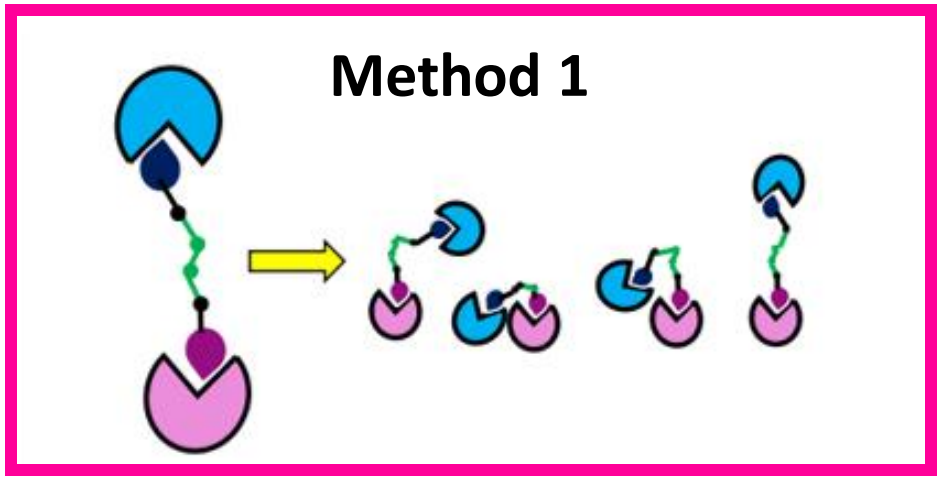
Application of TCP in PIC design

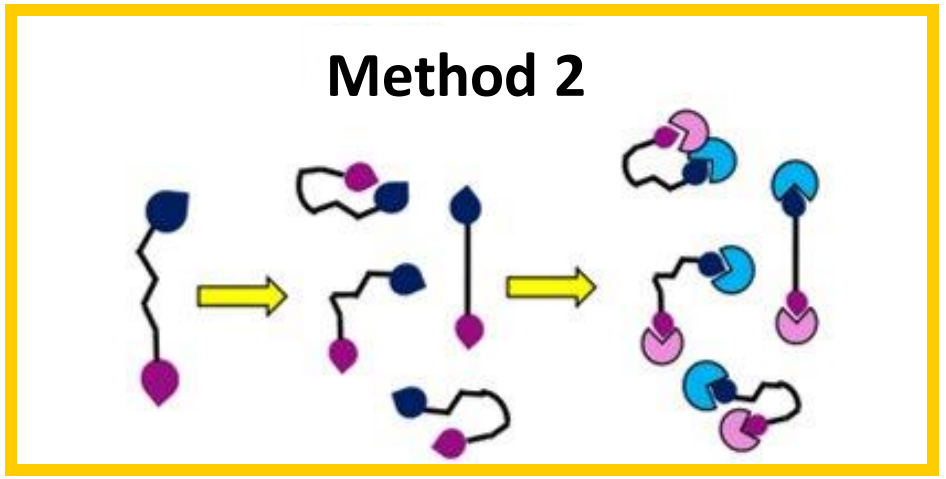
5. Decide which compounds to synthesize based on selectivity
   a) Different warhead
      a) Different binding pockets
   b) Linker enumeration

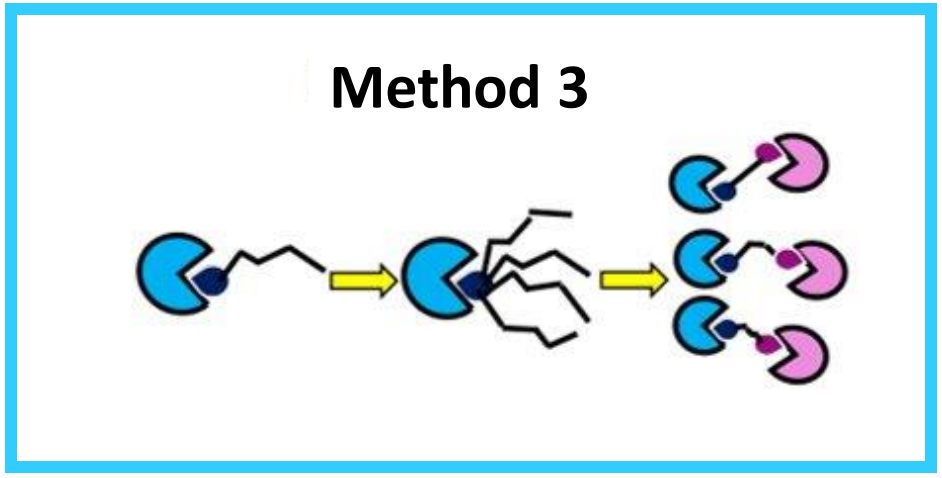
7. Accurate structural prediction

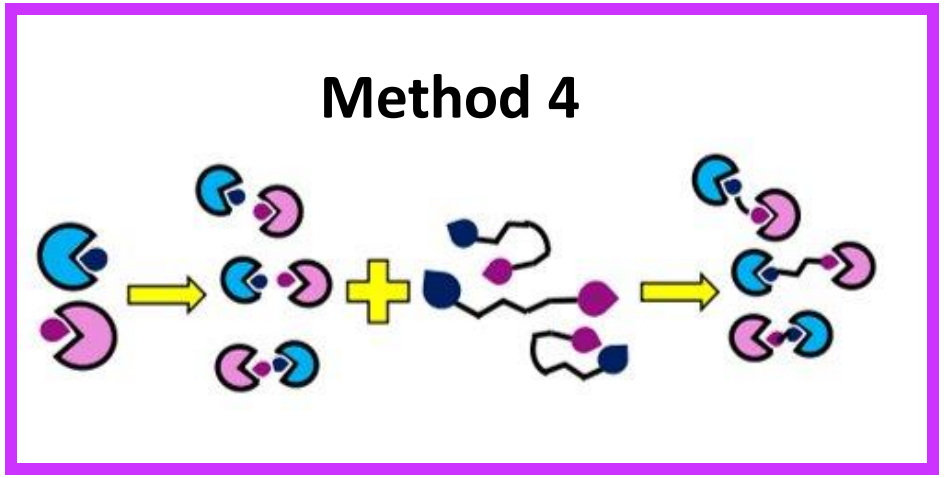
6. Decide which compounds to synthesize based on ternary complex formation ability
   a) Linker modifications
   b) Warhead modifications

Four Different Approaches to TCP









DRUMMOND, Michael L.; WILLIAMS, Christopher I. In silico modeling of PROTAC-mediated ternary complexes: validation and application. JCIM, 2019, 59.4: 1634-1644.

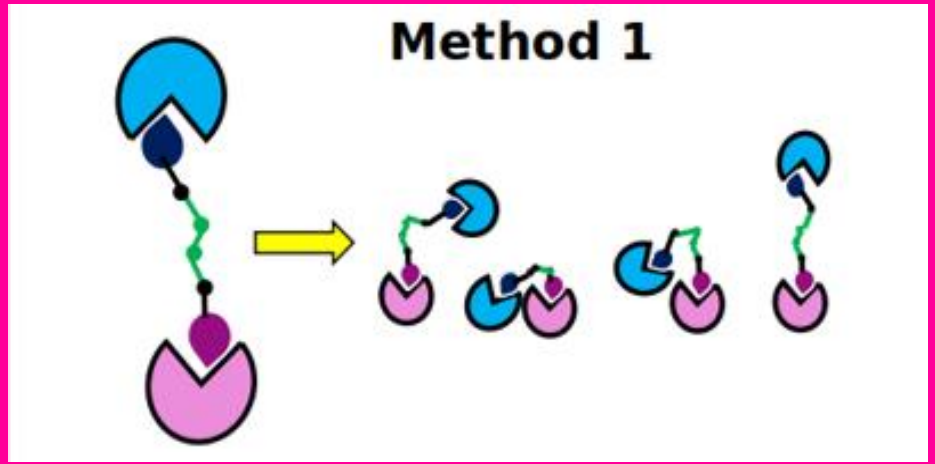
Method 1

Attachment of each fragment (warhead/E3 binder) to the initial conformation of linker.

Positioning of **proteins** at around their respective fragments.

Protein-ligand and protein-protein clashes due to the separate starting environments (protein-fragment groups vs. linker conformation).

The PROTAC conformation is automatically adjusted to adopt an **extended conformation.**



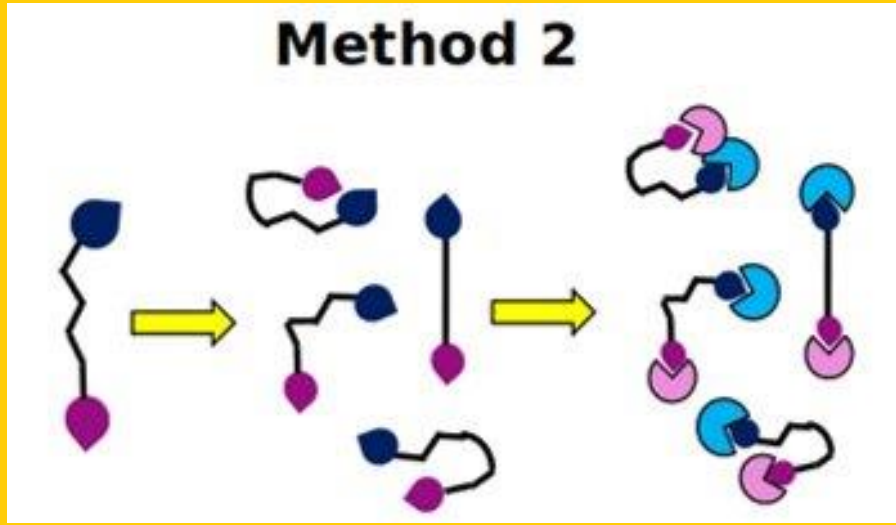
The entire ternary complex is sampled at once.

Method 2

Two protein-fragment complexes and **a full PROTAC** are required as inputs.

Different PROTAC conformers are pre-generated/sampled on the fly (sampling may affect fragment conformer).

Because the PROTAC sampling occurs in the absence of any proteins, there is often significant overlap after the proteins are reintroduced.



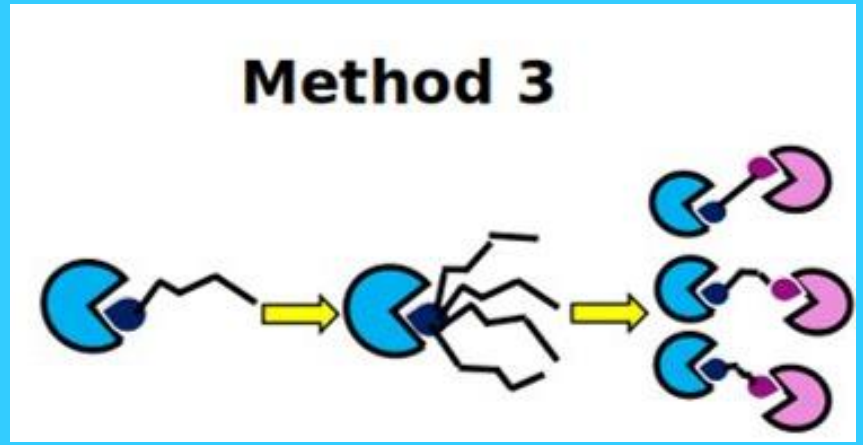
PROTAC conformations are sampled independently, followed by post hoc addition of rigid body proteins.

Method 3

Only one of the proteins is included in the conformers sampling phase (the smaller of the two).

The fragment belonging to the binding moiety of the protein included in the sampling is tethered during the conformational sampling.

The second fragment is kept rigid to prevent deformations.

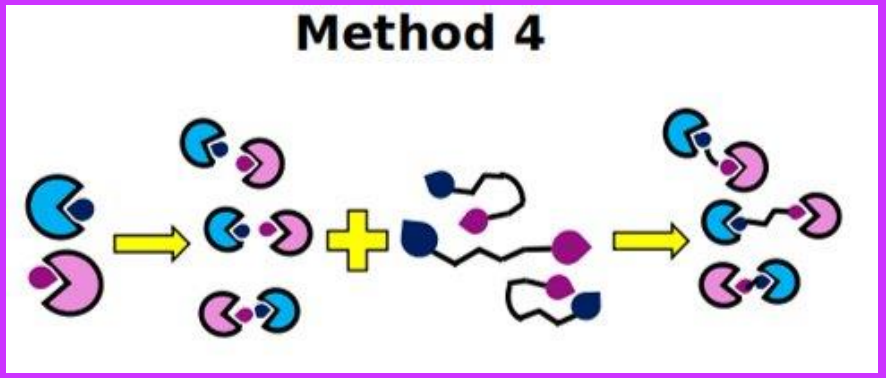


The PROTAC is sampled in the context of one of the proteins, with the second added afterwards.

Method 4

Two protein-fragment complexes and different **PROTAC conformations** are required as inputs.

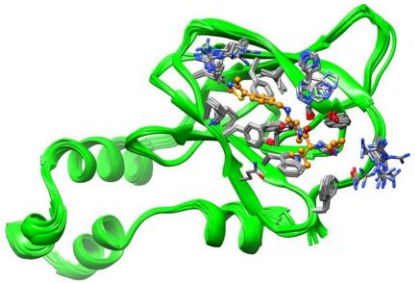
- Phase 1: Protein–protein docking
- Postfiltering based on patch-based descriptors.
- Phase 2: PROTAC conformational ensemble.

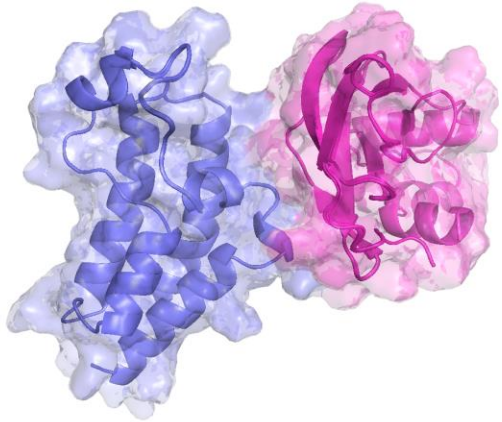


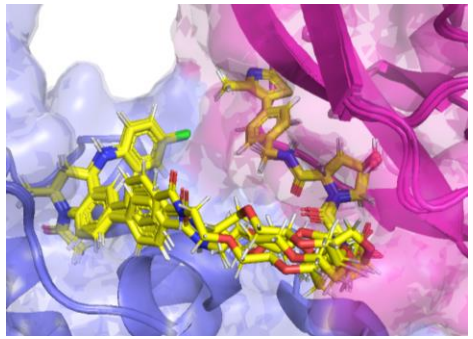
PROTAC conformations are sampled independently of the proteins, but possible POI/E3 ligase are sampled via protein-protein docking.

CelerisTx: TCP Pipeline Introduction

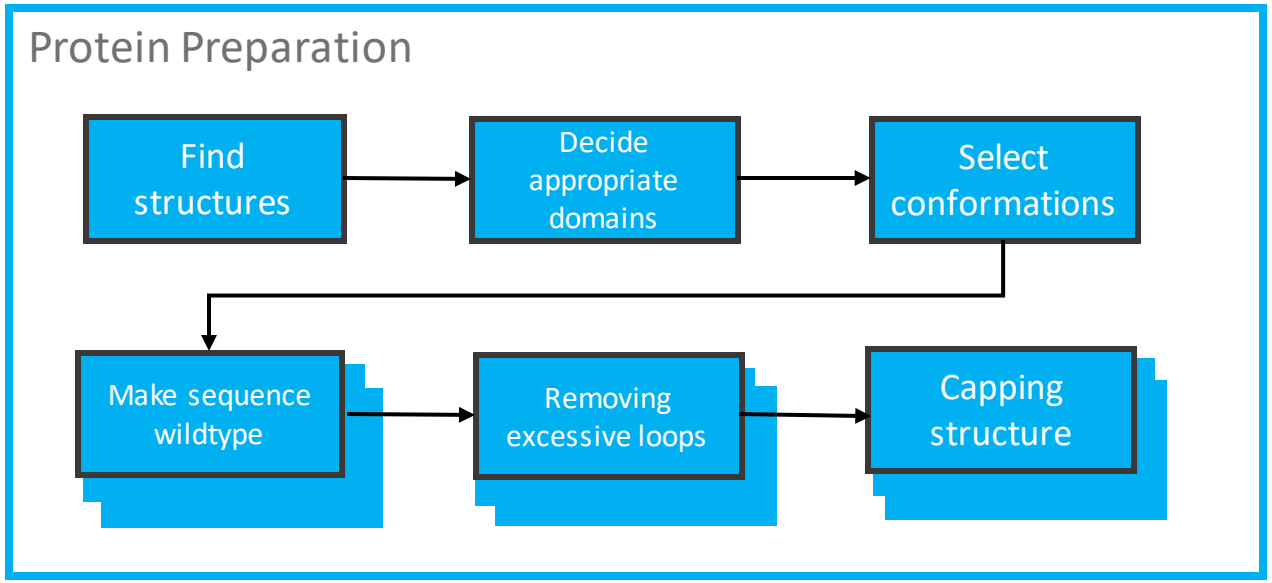
1. Establishing protein-fragment interactions
2. Protein-Protein Interactions Prediction
3. Linker insertion, TC optimization, PIC ranking



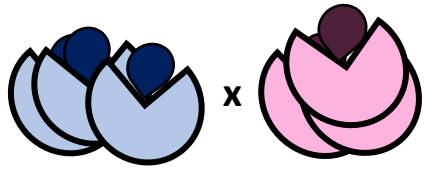




Conformational Ensemble Generation



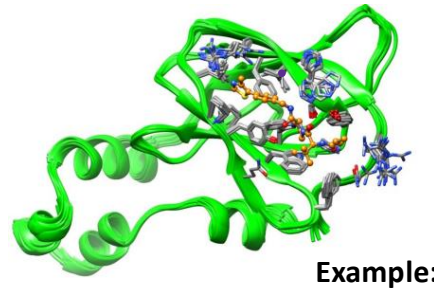
 # PPI runs =



 # POI confs. **# E3 confs.**

Typically 10-20 distinct structures per a given protein

- X-Ray
- NMR
- Cryo-EM
- *Ab initio:*
  - Molecular dynamics
  - Normal mode analysis
  - Homology modelling



Example: VHL ligase

Binary Inputs Preparation

Data acquisition

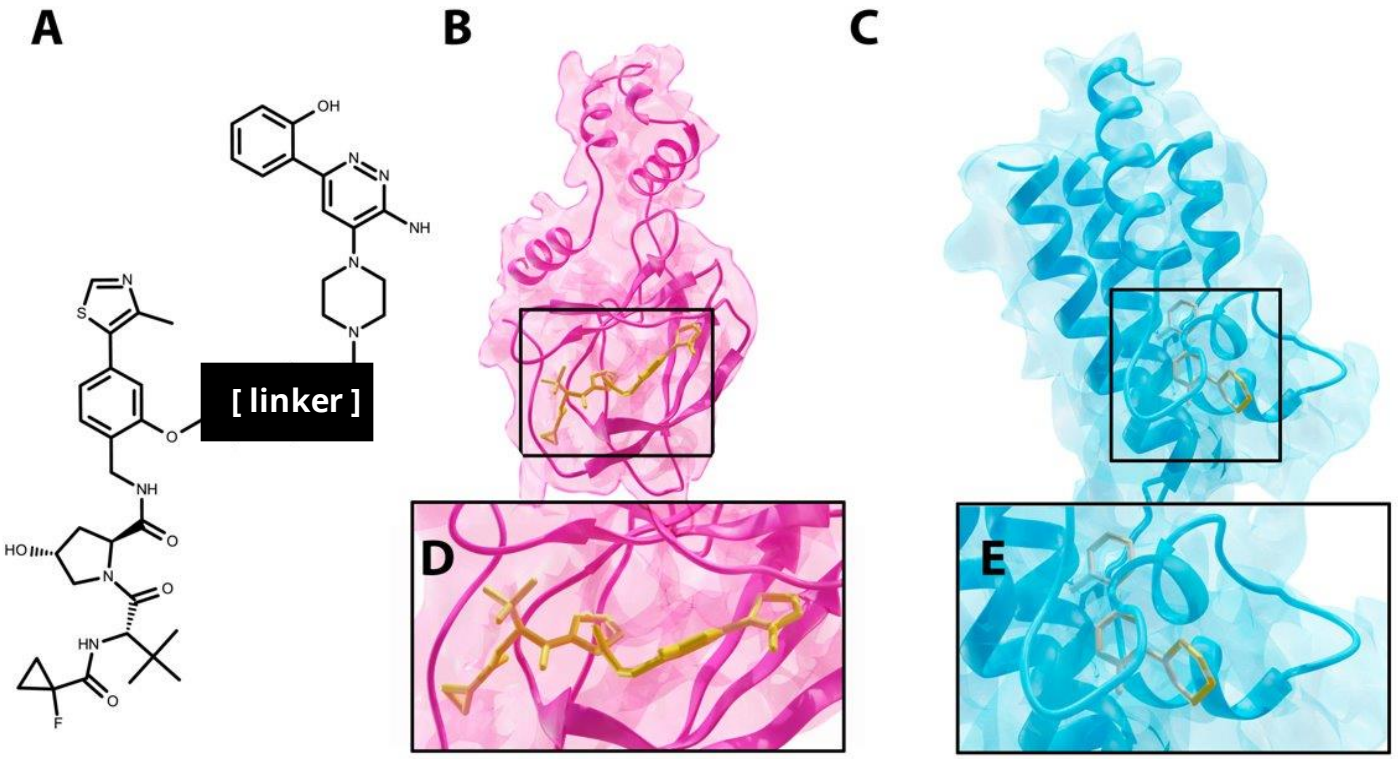
- **Retrieval of** protein structures
- Automated download of ligand **bioactivity data**

Protein/ ligand structure preparation

- Optimizing structure with co-resolved warhead
- Rigid body docking
- (Ensemble docking)

Generating binary complexes

- **pKa predictions**
- Adding explicit hydrogens, fixing missing chains
- **Energy minimization**



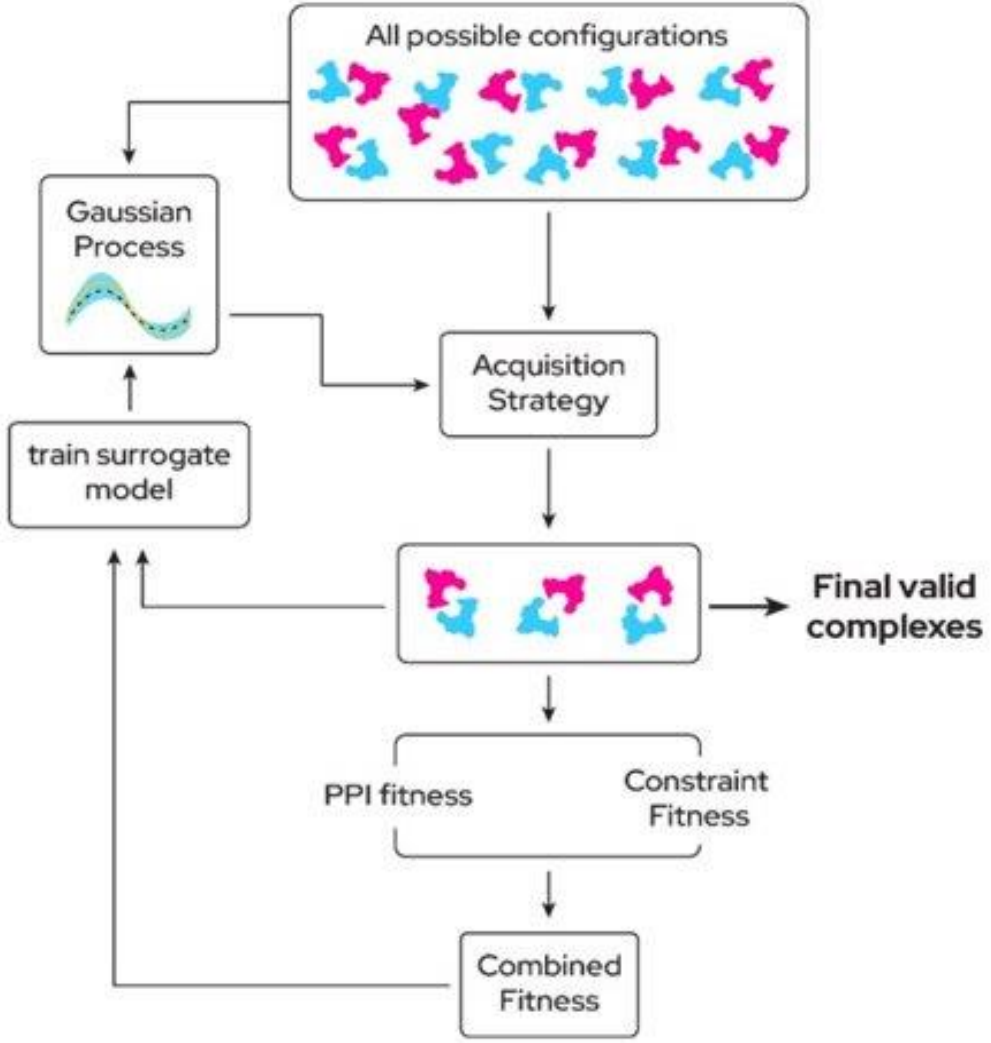
Bayesian-optimization PPI Prediction

Generating different relative **orientations** and **translations** helps **sample alternate ternary complex poses**.

Our goal is to optimize a **score** that describes the **quality of a ternary complex pose**.

Protein-protein interaction score, **linker constraint** score, **PIC stability score**.

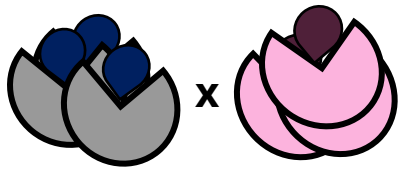
Clustering, reranking, and filtering final poses.



Physics-based PPI Prediction

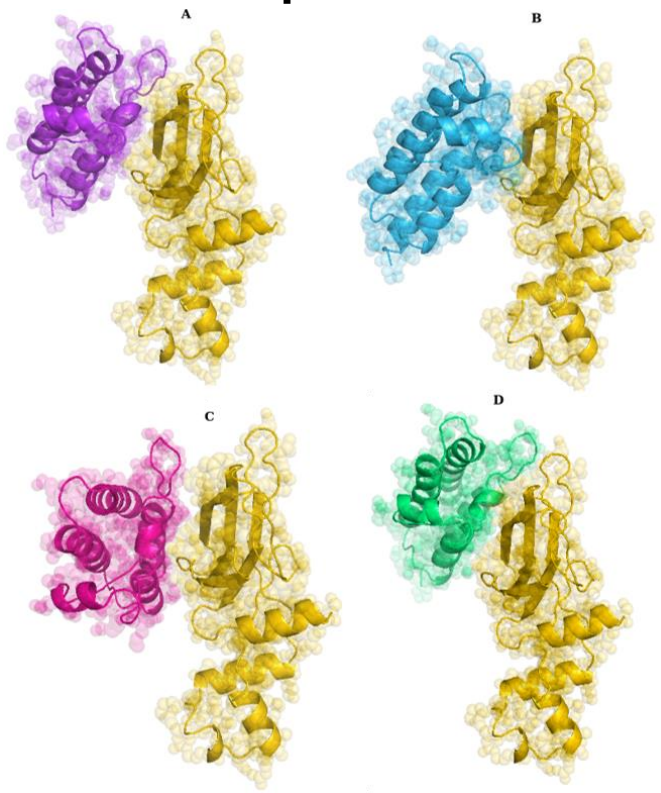
Conformational ensemble generation

 # PPI runs =

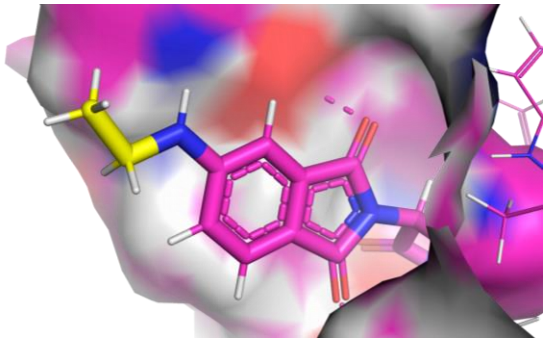


 # POI confs. x # E3 confs.

Exploring PPI space



Inclusion of a "spacer" to mimic presence of a linker

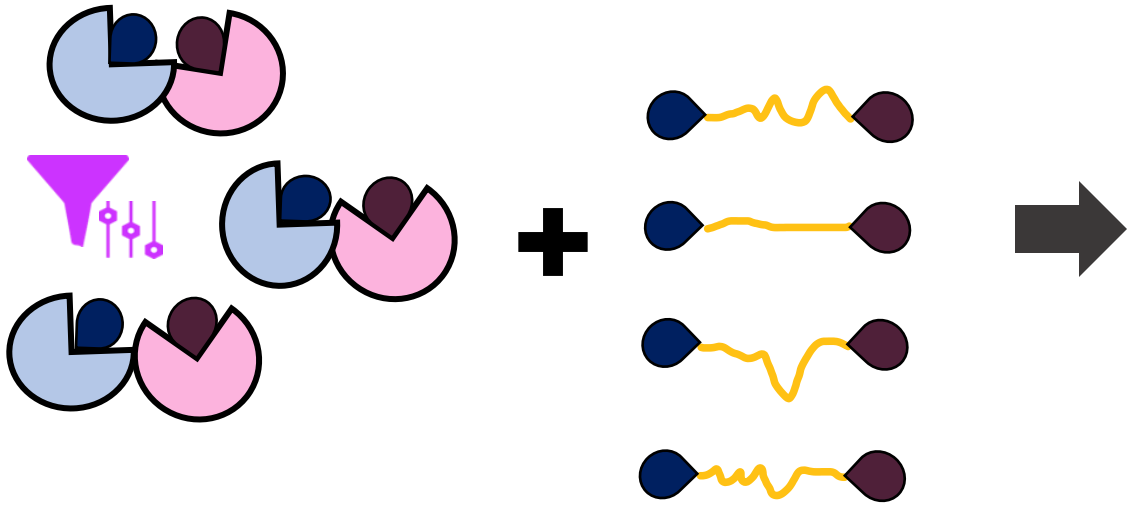


Pose filtering

1) Exit vector distance
2) PPI score
3) L-RMSD

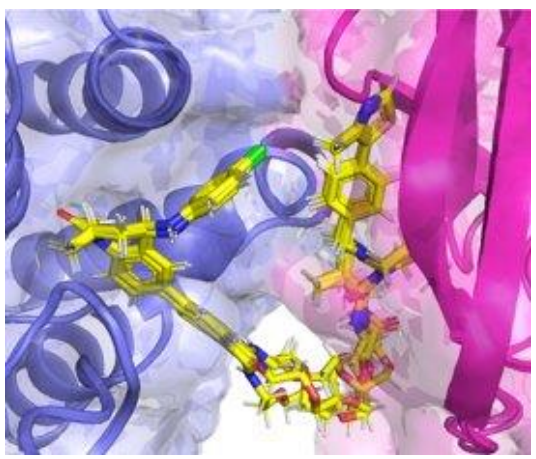
Linker Insertion and Structure Refinement

Input: Post-filtered poses + 3D PIC structures

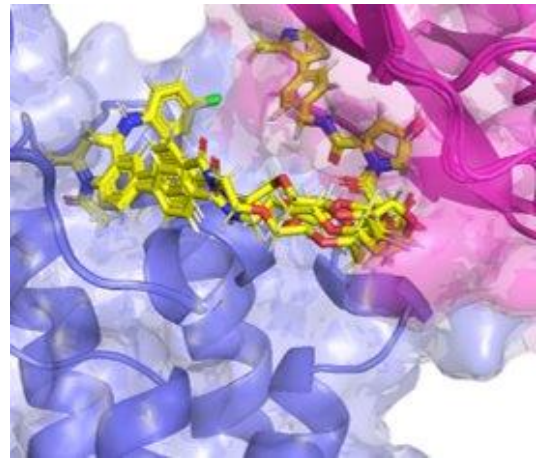


Modeling in a linker

(* fragments are constrained)



Top view



Side view

Linker Insertion & Structure Refinement

Molecular Mechanics with Generalized Born and Surface Area Solvation

Calculate partial charges for degrader molecules

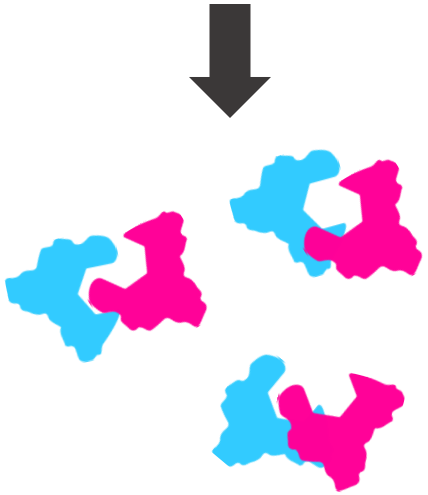
Optimize ternary complexes

Energy minimization

Simulated annealing T=300-0K

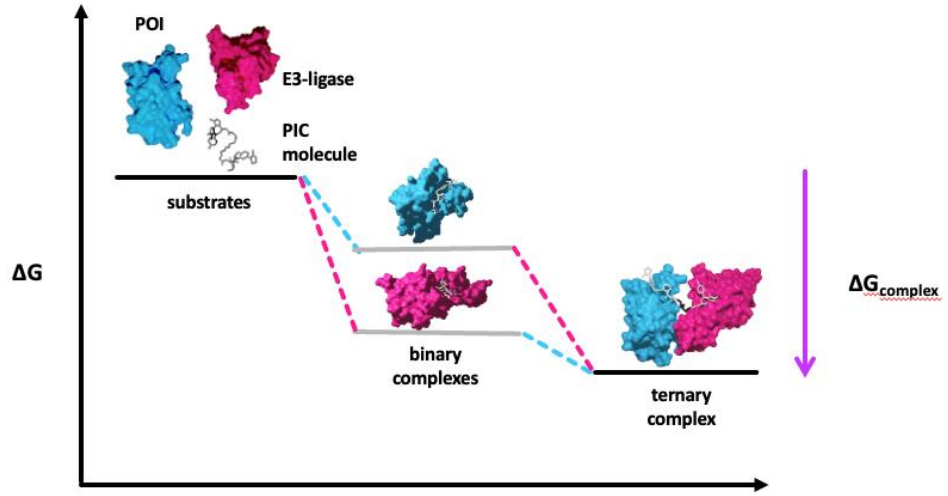
MMGBSA calculation

Approximate binding energy



Rank different Degraders

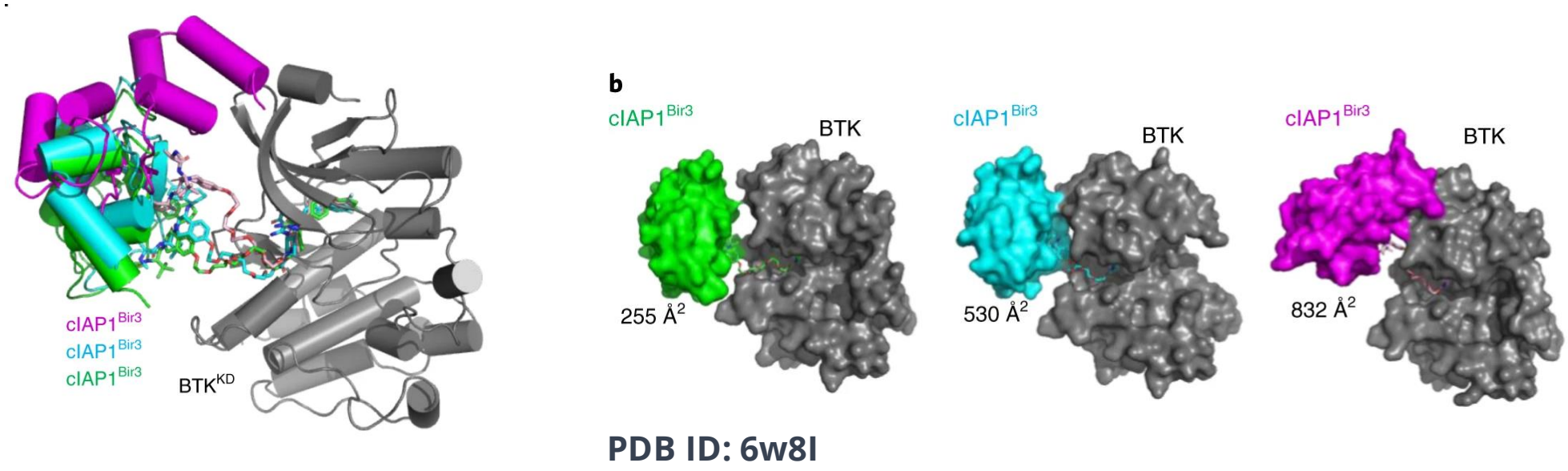
(Boltzmann average)



Validation

Moving away from structure to affinity-based comparisons

- **22 experimentally determined structures** available in PDB
- All available structures originate from **X-ray measurements** (a single "snapshot")



PDB ID: 6w8l

SCHIEMER, James, et al. Snapshots and ensembles of BTK and cIAP1 protein degrader ternary complexes. *Nature Chemical Biology*, 2021, 17. Jg., Nr. 2, S. 152-160.

Validation

Use case	Validation
1. Target selection based on feasibility on developing ternary complex	Known degraded targets vs random targets as negatives
2. Target selection based on selectivity prediction	Known example pairs of proteins where selectivity has been achieved vs random pairs
3. Calculate minimum linker length	Datasets with same warhead + E3 binder, different linker size
4. & 6. Decide which compounds to synthesize based on ternary complex formation ability	Compound series needed
5. Decide which compounds to synthesize based on selectivity	Compound pair data with the same E3 ligase but different POI tested
7. Accurate structural prediction	Experimentally-determined structures

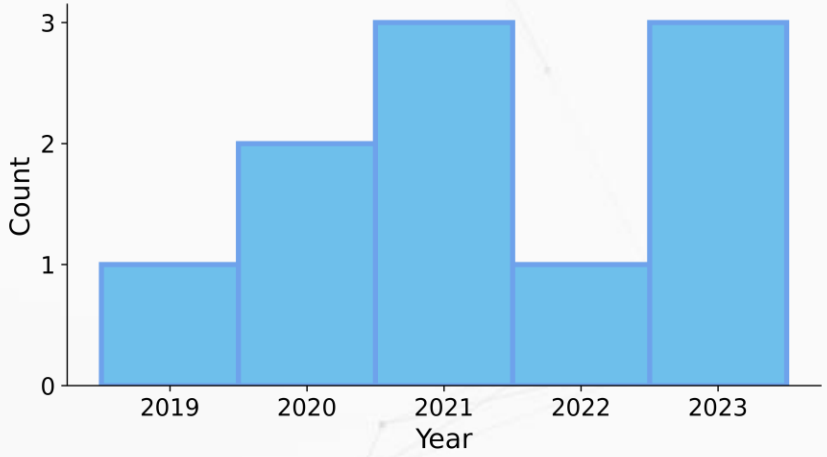
Validation: Examples from Literature

3. Calculate minimum linker length

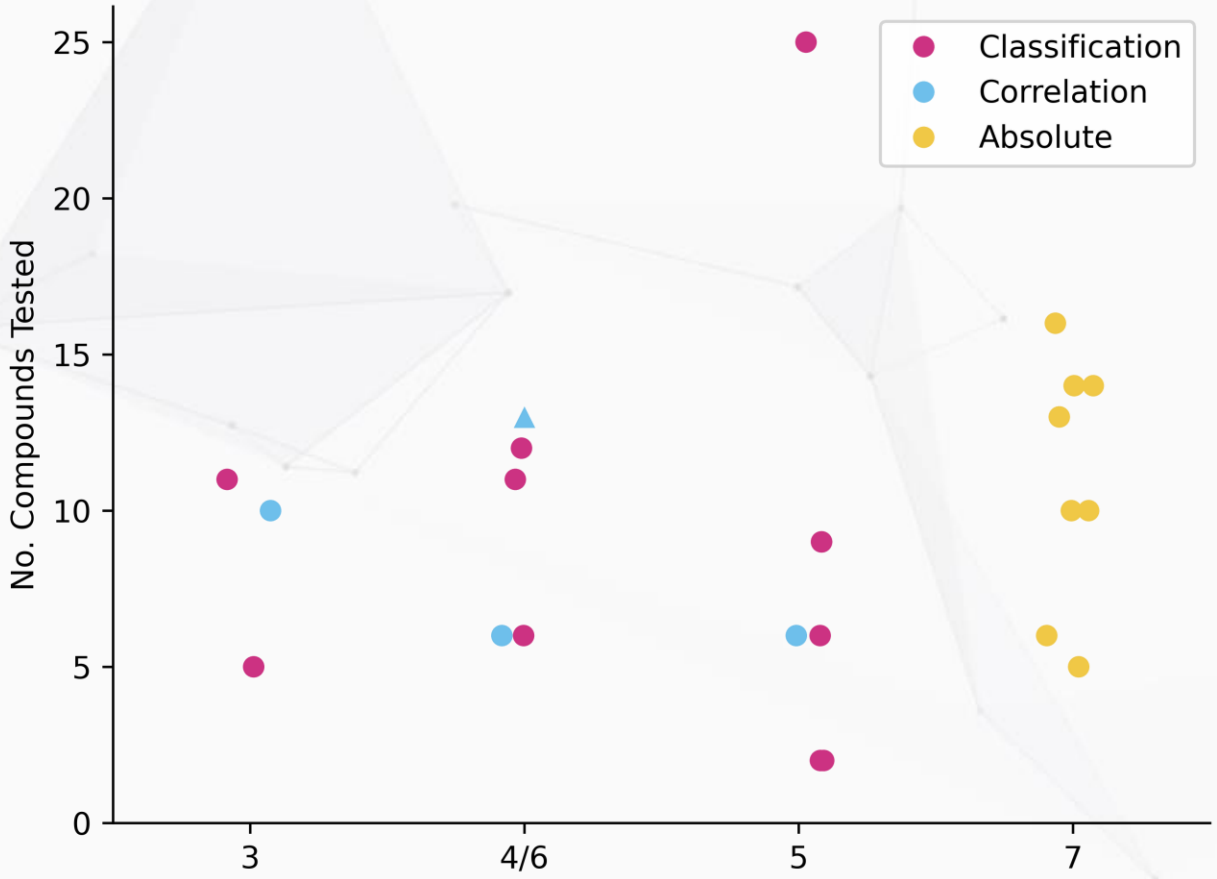
4. / 6. Decide which compounds to synthesize based on ternary complex formation ability

5. Decide which compounds to synthesize based on selectivity

7. Accurate structural prediction

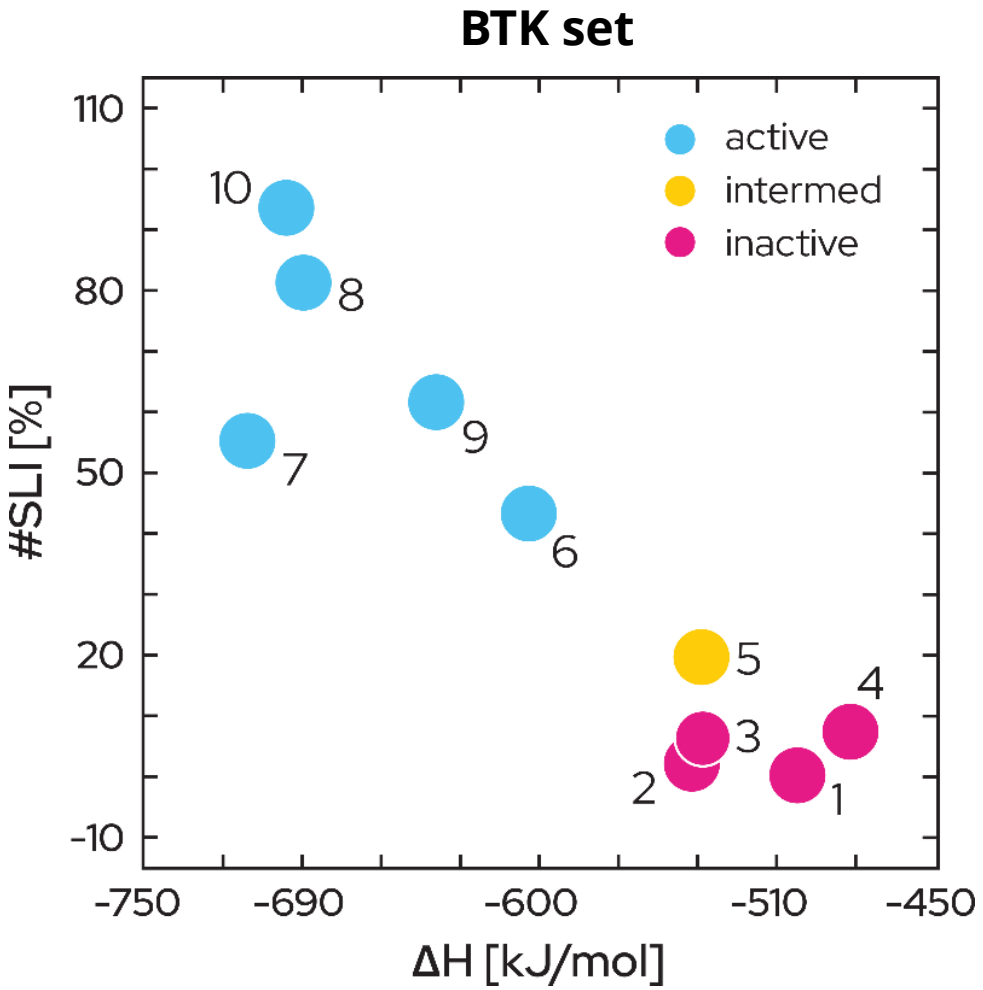


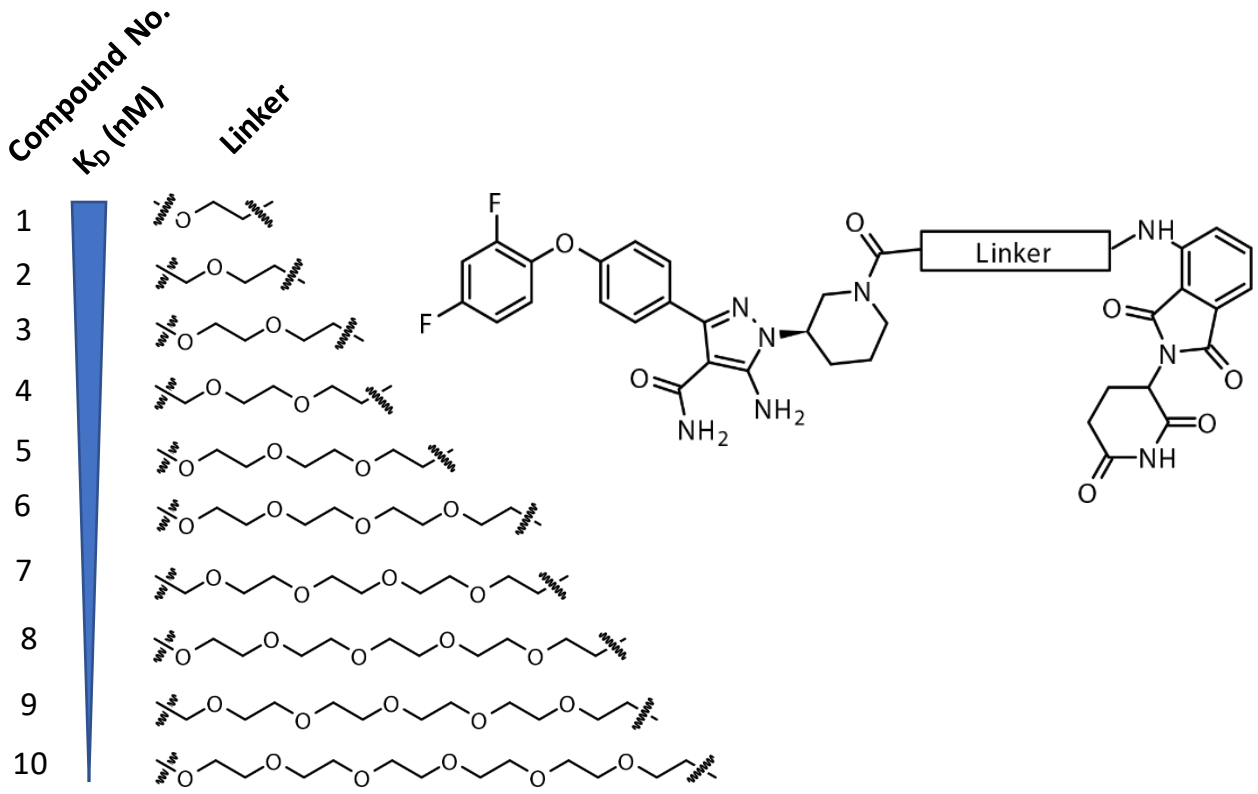
Histogram showing number of ternary complex pipeline method papers vs time



Plot showing number of compounds validated against in different type of ternary complex pipeline validation settings.

Validation: Tyrosine-protein kinase BTK





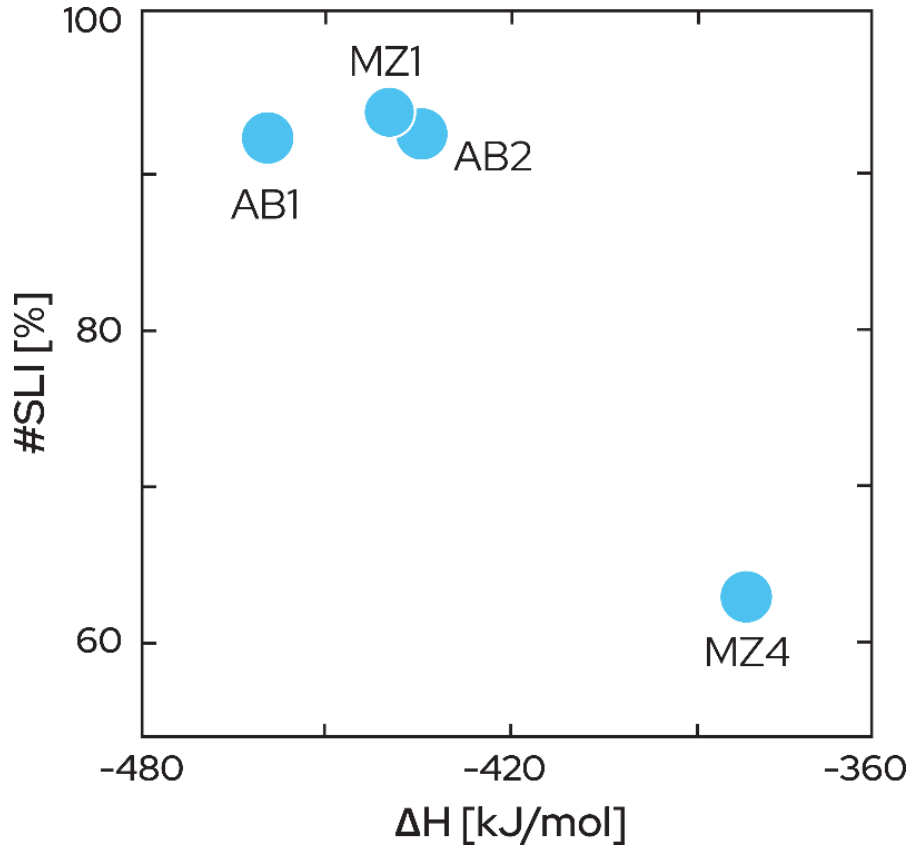
Validates "What is the shortest linker length?" + HitID virtual screen

Experimentally-determined data

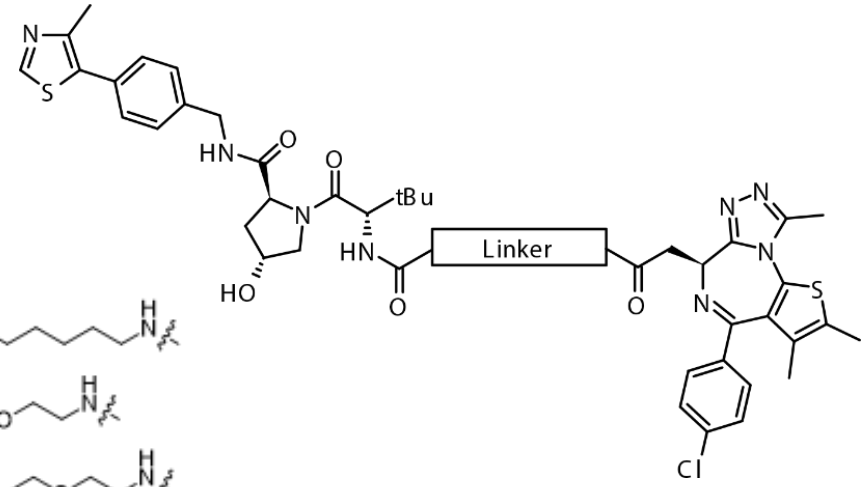
Zorba, A. et al. (2018). Delineating the role of cooperativity in the design of potent PROTACs for BTK. PNAS, 115(31), E7285–E7292.

Validation: BRD4

BRD4 set



Name	K_D (nM)	Linker
AB1	4	
AB2	3	
MZ1	4	
MZ4	26	



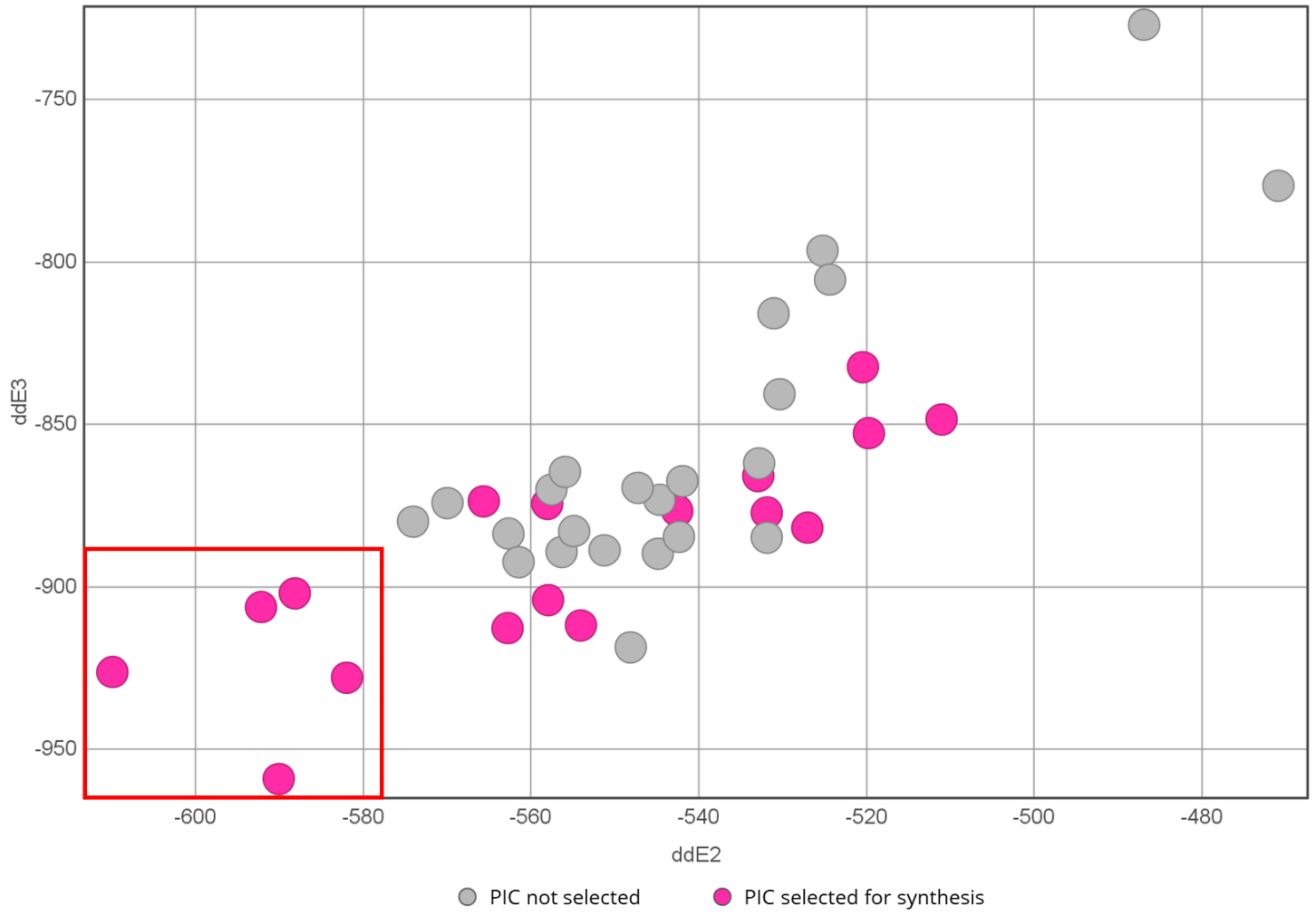
Validates "What is the shortest linker length?" + HitID virtual screen

Using TCP Pipeline For Hit Identification

Compound	#SLI (%)	# rotatable bonds	ΔH (kcal/mol)
Domain-X control (+ve)	33.4	21	-832
Domain-X control 1 (-ve)	8	7	-710
Domain-X control 2 (-ve)	2.4	9	-691
Domain-Y control (+ve)	27.2	11	-815
Domain-Y control (-ve)	12.6	7	-739

Validates HitID virtual screen

Using TCP Pipeline for Hit Expansion

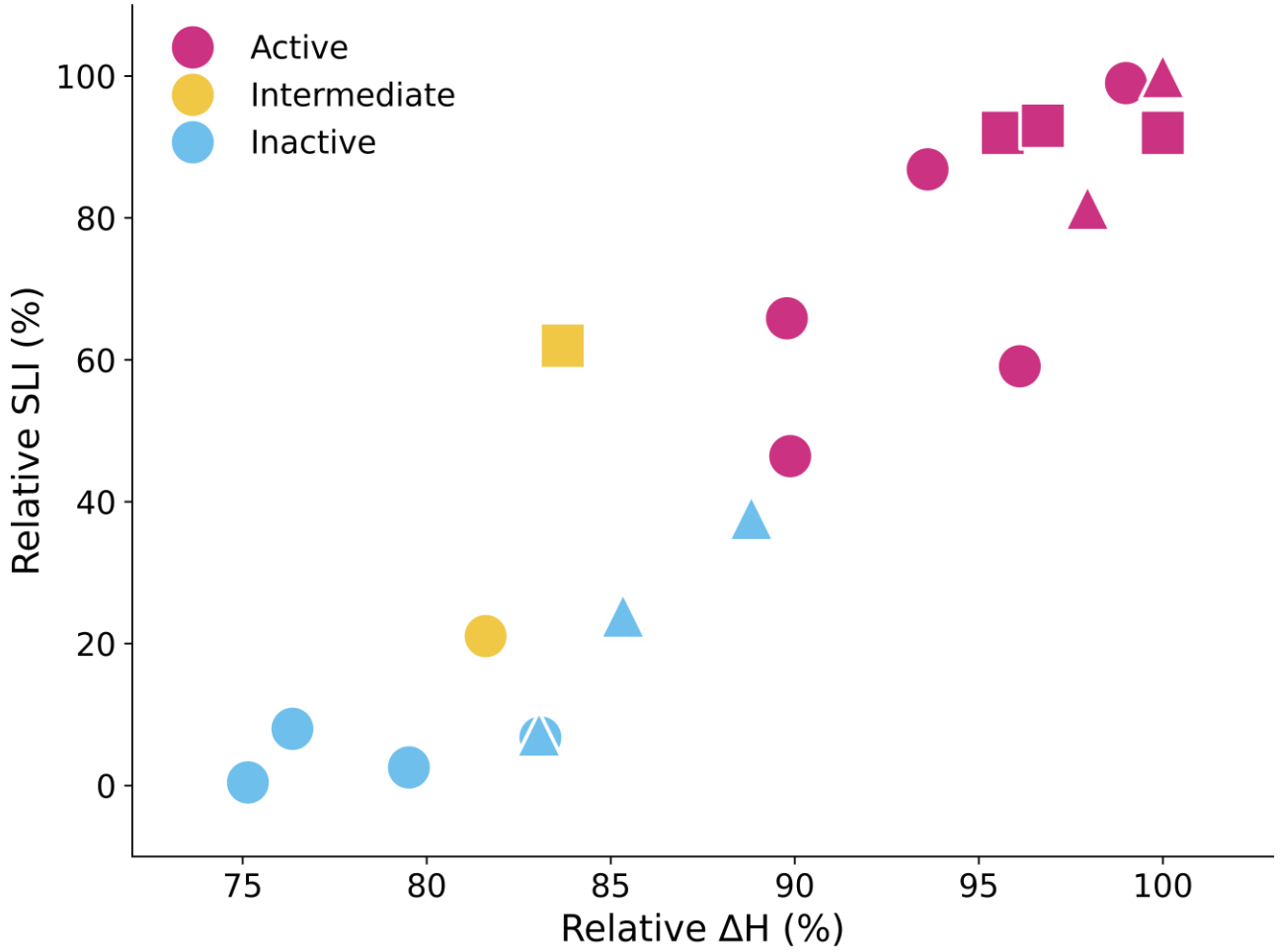


80 design ideas
from MedChem

40 compounds
for PBTCP runs

17 compounds
selected for synthesis

In-house TCP pipeline: Overall Validation Results



Validation of our PB-TCP method for different PROTACs.
Different marker shapes represent different targets.

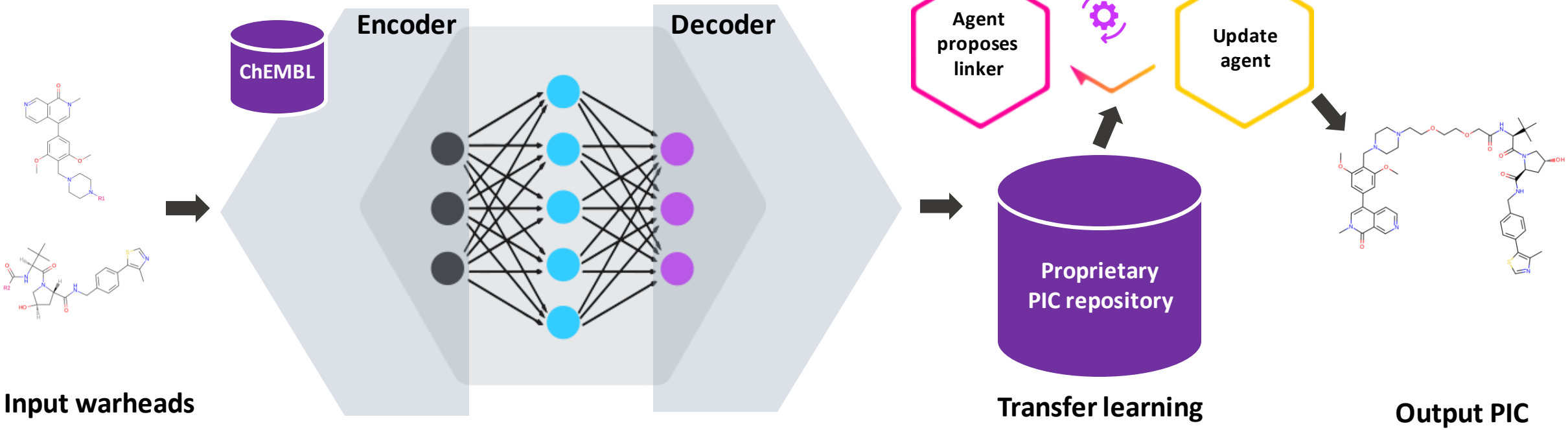
AGENDA

Generative AI for Molecular Linker Design

Input linker criteria

- Number of hydrogen bond donors
- Desired linker length
- Number of rotatable bonds

.....



Compilation of Training Data Set: Strategy

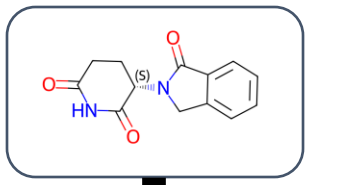
Substructure (E3 recruiter) screens

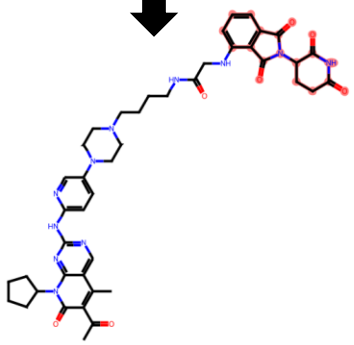
Structure standardization & integration

PhysChem filters application

Tanimoto similarity with Protac-DB

Structural query





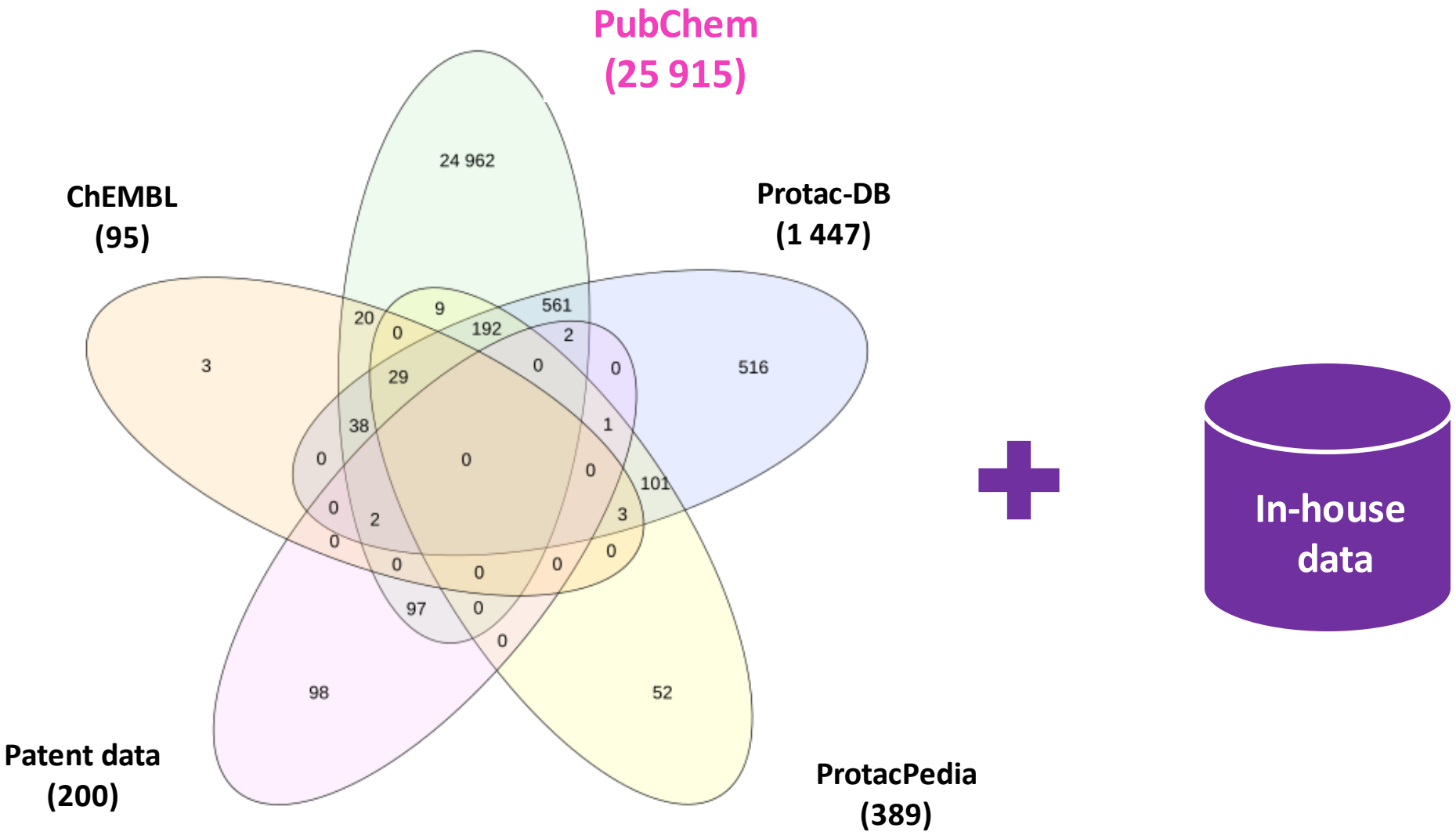
- Stripping of salts and fragments
- Neutralization of charges
- Filtering our entries with unusual elements
- Generating InChi, InChiKey, and SMILES (canonical)
- Data merging

- Molecular weight
- Hydrogen-bond donors
- Topological polar surface area / Molecular weight
- Topological polar surface area
- cLogP

Proprietary PIC repository

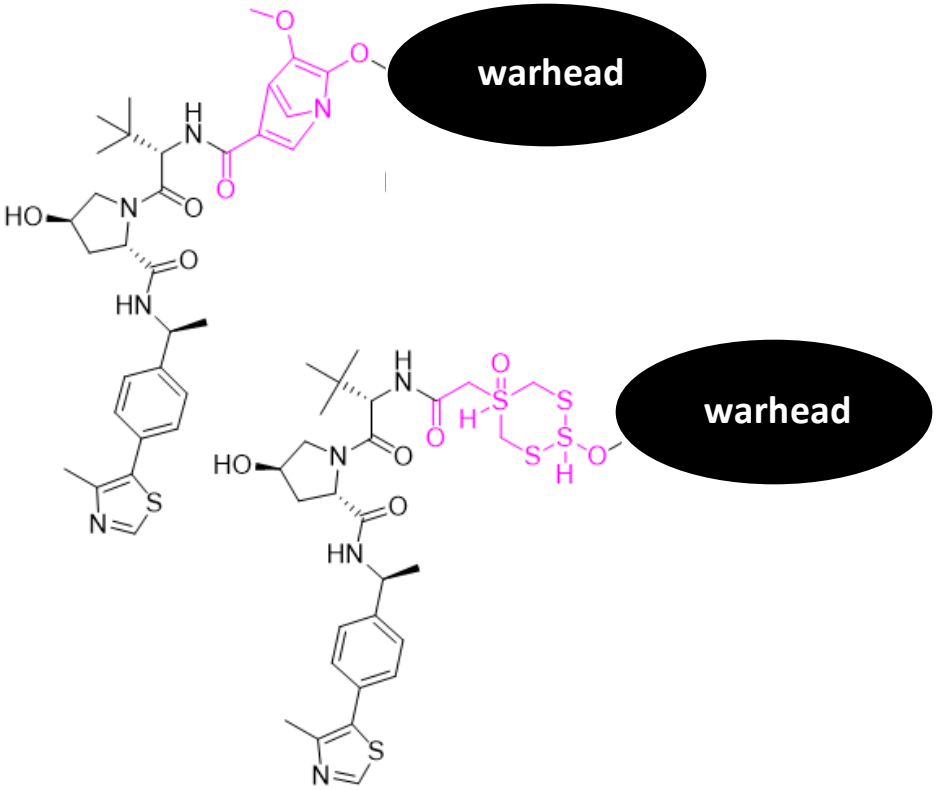
(ChEMBL, PubChem, PROTAC-DB, Patents, Protacpedia, in-house data)

Compilation of Training Data Set: Data Overlaps



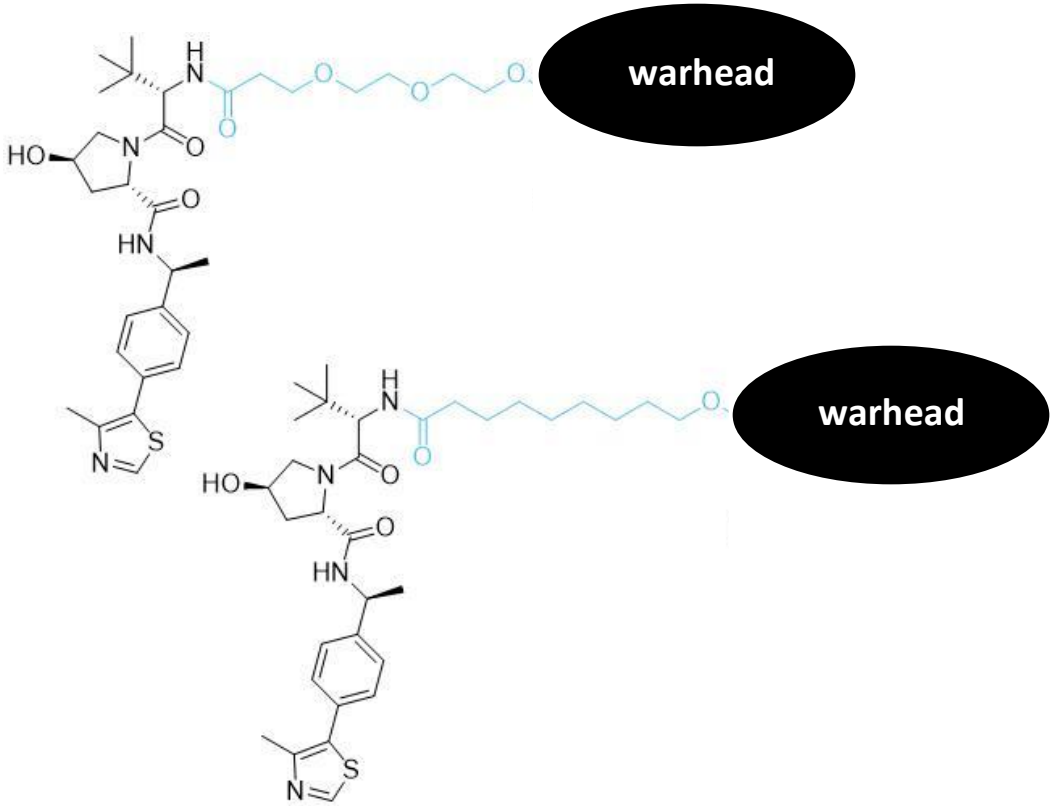
Linker Generation: Examples

creative & novel

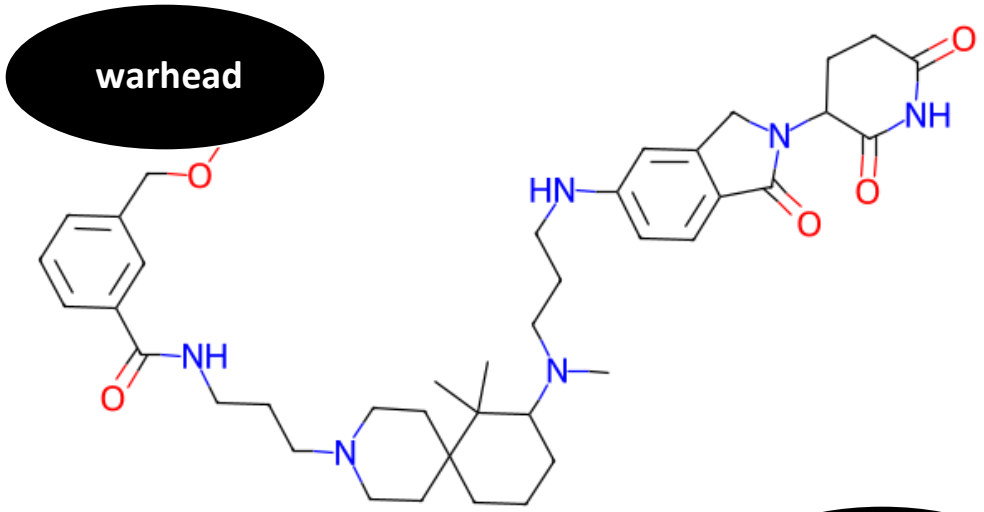


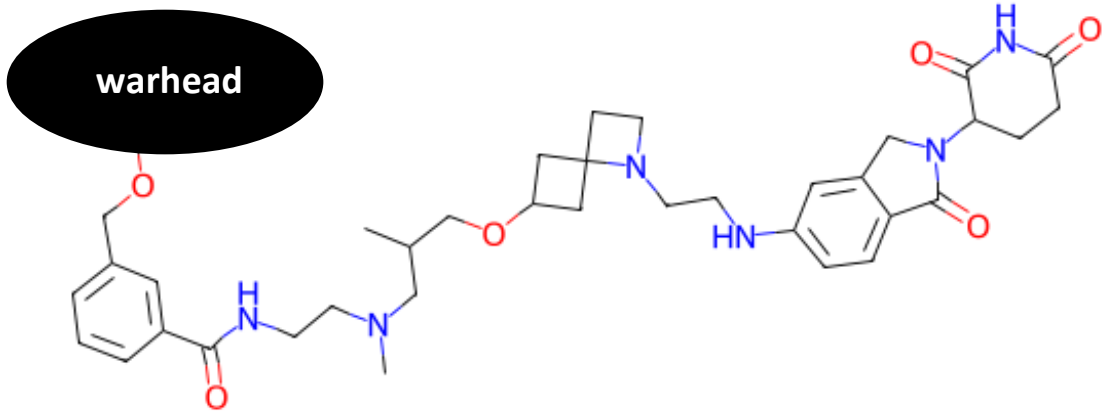
VS.

stable & synthesizable



Linker Generation: Examples





Linker generation model augmented by a subset of commercially available PIC building blocks

Conclusions

- The development of PICs can greatly benefit from the application of in silico approaches:
  - **Predicting ternary complexes** can be used for both **structure-based design**, and for **ranking compounds** during selection.
  - **Validation** of TCP approaches is crucial to ensure its applicability (by comparison of calculated TC stabilities with **ternary dissociation constants** or by obtaining additional experimental structures, preferably using **Cryo-EM or NMR** methods, as these methods capture protein conformational dynamics).

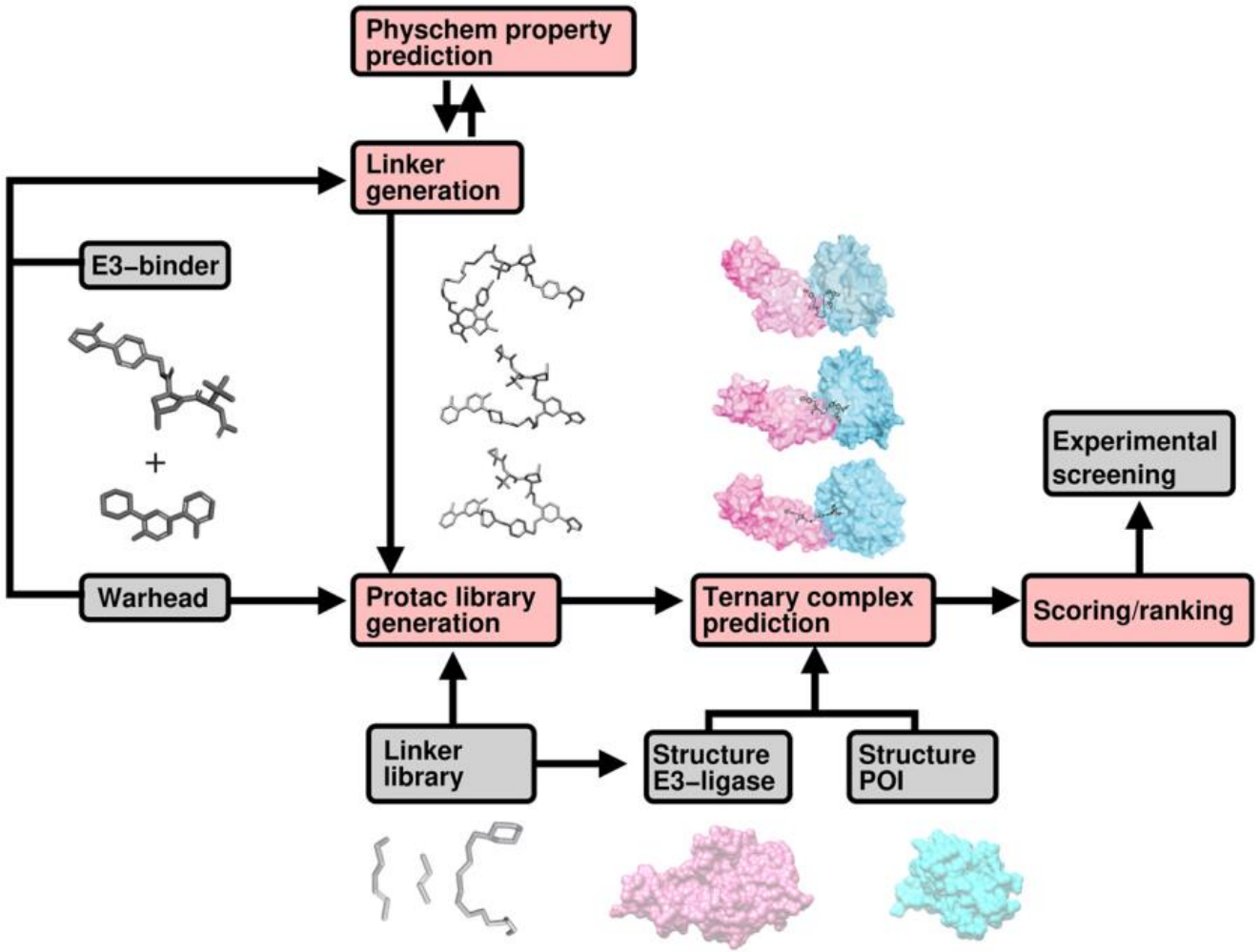
Conclusions

- **Linker generation tools** can provide valuable assistance to medicinal chemists in designing novel linkers by systematically optimizing different compound properties.
- Applying **data mining** approaches to augment current PIC datasets with **novel** and **standardized** data points is essential to enhance efforts in linker generation approaches.

Conclusions

- Coupling **ternary complex prediction** with **linker generation** has the potential to strengthen the current degrader discovery pipeline

Example of a PROTAC Screening Pipeline



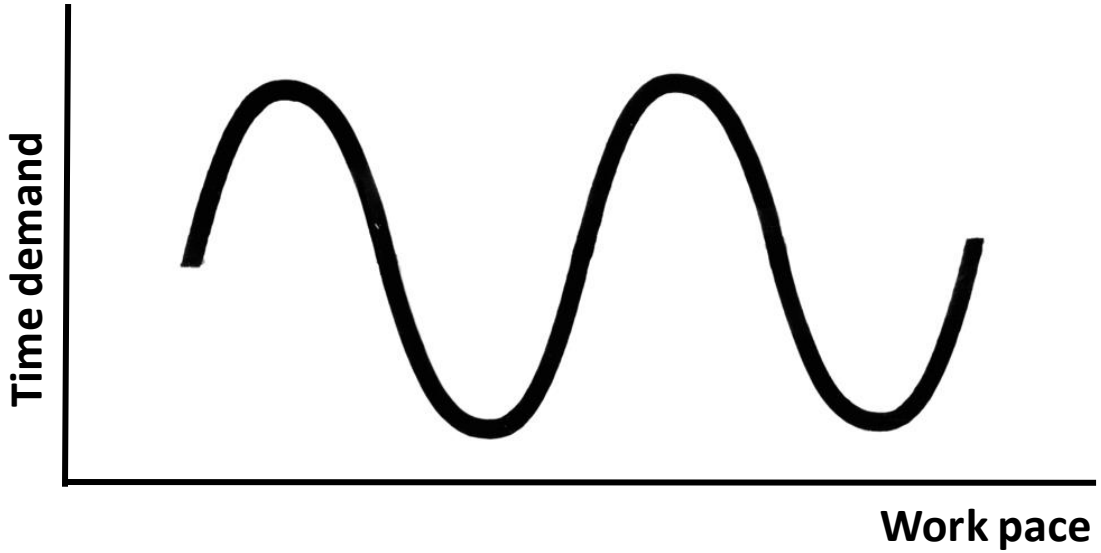
Gray boxes:
- experimental steps
- input data

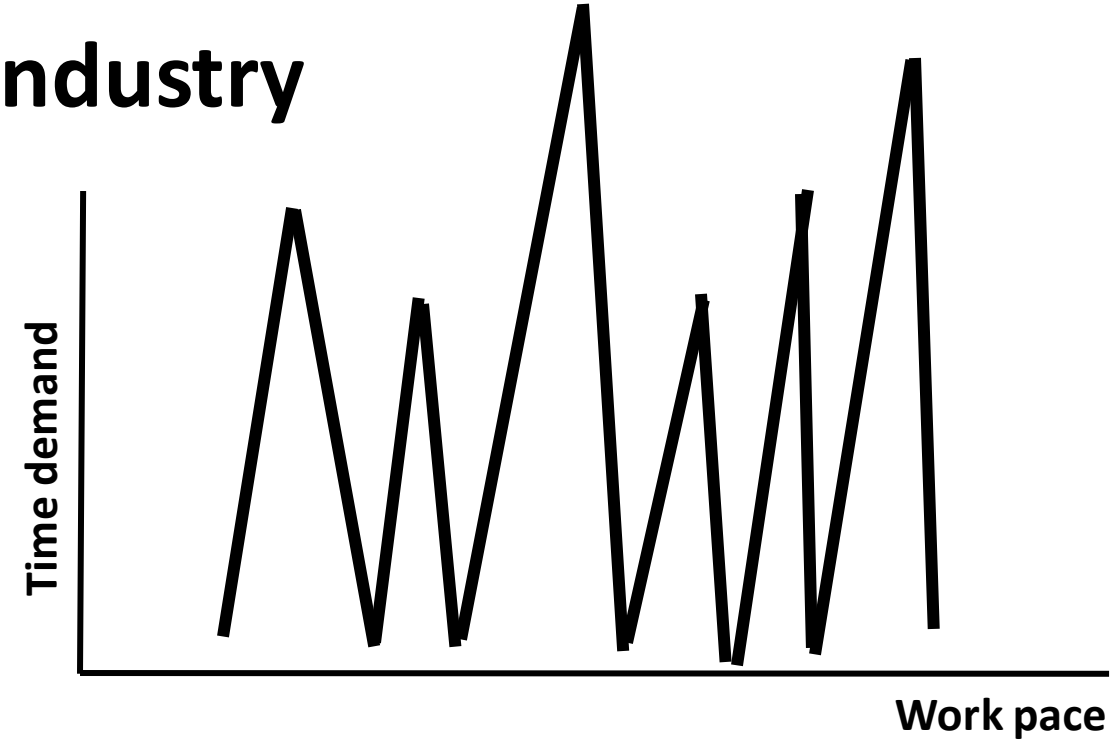
Pink boxes:
- in-silico steps

TUNJIC, Tin M.; WEBER, Noah; BRUNSTEINER, Michael. Computer aided drug design in the development of proteolysis targeting chimeras. *Computational and Structural Biotechnology Journal*, 2023.

AGENDA

Work model in Academia vs. Industry





	Academia	Industry
Work pace	Cyclic	Gas-Brake-Gas-Brake
Deliverable	Publication (patent)	Marketed product/service
Time frame	Long-term projects	Tight deadlines and project timelines
Team culture	Individualism	Cross-functional teams, frequent **coordination**
Decision-making structure	High degree of autonomy	Hierarchical

From Academia to Industry: Transferable Skills

Publishing research papers	**Data Visualization and Communication**
Public outreach (Researcher's night)	**Communication with non-experts** (investors, collaboration partners, etc.)
Customized code	**Documented code/ data pipelining tools** (KNIME, LiveDesign, etc.)
Teaching activities	**Cross-functional project team** communication (matrix organization)
Student (co-)supervision	**Managing a team, supervising junior team members/research interns**
Critical thinking, problem solving	**Task decomposition to manage workload**

Job interview process

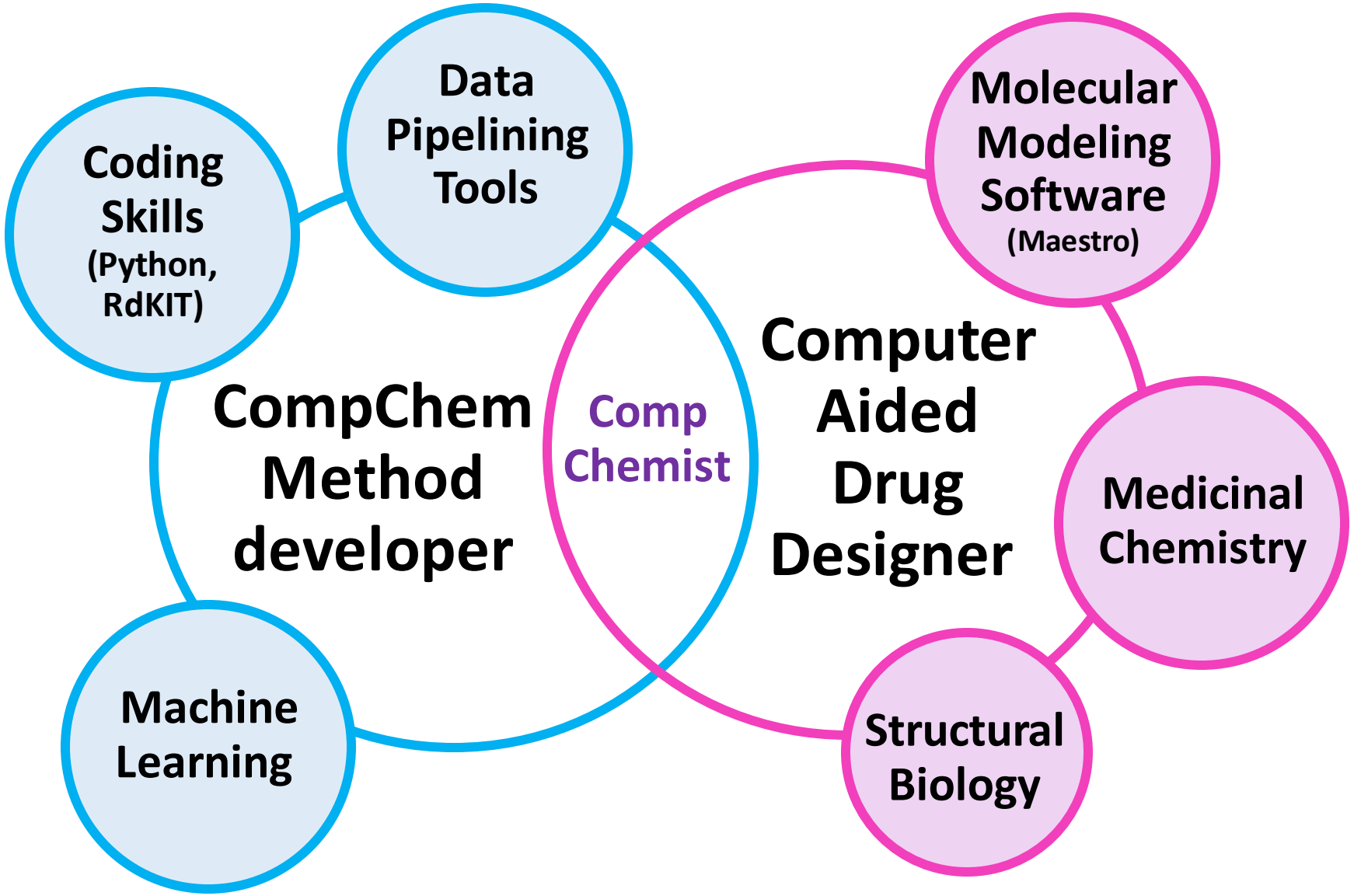
Role: **Senior/Principal Computational Chemist** (London, Graz)

Brief introduction
Background
Motivation
Expected salary
Expected start date

Meeting the team
MM questions
Conceptual tasks
Cultural fit

Medicinal chemistry
Biology
Synthetic chemistry

Different roles as a Computational Chemist



Thank you for attention!

CelerisTx

The PIC™ Engineering Company

linkedin.com/company/celeristx

celeristx.com

hi@celeristx.com